

# **Visualization Onboarding**

### Unterstützung von Benutzer\*innen beim Verständnis unbekannter Datenvisualisierungen

### DISSERTATION

zur Erlangung des akademischen Grades

### Doktorin der Technischen Wissenschaften

eingereicht von

#### **DI Christina Stoiber, BSc**

Matrikelnummer 11771178

an der Fakultät für Informatik

der Technischen Universität Wien

Betreuung: Priv.-Doz. Dipl.-Ing. Dr. Wolfgang Aigner, MSc Zweitbetreuung: Ao.Univ.Prof.in Mag.a Dr.in Margit Pohl

Diese Dissertation haben begutachtet:

Katy Börner

Samuel Huron

Wien, 7. August 2023

Christina Stoiber





# **Visualization Onboarding**

### Supporting Users in Understanding Unfamiliar Visual Representations

DISSERTATION

submitted in partial fulfillment of the requirements for the degree of

### Doktorin der Technischen Wissenschaften

by

DI Christina Stoiber, BSc

Registration Number 11771178

to the Faculty of Informatics

at the TU Wien

Advisor: Priv.-Doz. Dipl.-Ing. Dr. Wolfgang Aigner, MSc Second advisor: Ao.Univ.Prof.in Mag.a Dr.in Margit Pohl

The dissertation has been reviewed by:

Katy Börner

Samuel Huron

Vienna, 7<sup>th</sup> August, 2023

Christina Stoiber



# Erklärung zur Verfassung der Arbeit

DI Christina Stoiber, BSc

Hiermit erkläre ich, dass ich diese Arbeit selbständig verfasst habe, dass ich die verwendeten Quellen und Hilfsmittel vollständig angegeben habe und dass ich die Stellen der Arbeit – einschließlich Tabellen, Karten und Abbildungen –, die anderen Werken oder dem Internet im Wortlaut oder dem Sinn nach entnommen sind, auf jeden Fall unter Angabe der Quelle als Entlehnung kenntlich gemacht habe.

Wien, 7. August 2023

Christina Stoiber



## Acknowledgements

I want to express my deepest gratitude to my sister, whose unwavering support and encouragement have been instrumental in my academic journey. Her belief in me has been a constant source of motivation, and I could not have reached this point without her.

I am also immensely grateful to my husband, whose love and encouragement have been the foundation of my success. His support and patience have been a constant source of strength, and I cannot imagine completing this journey without him by my side. Furthermore, I want to thank my parents for supporting me in pursuing my dreams. Their support and guidance have been invaluable to me throughout my academic journey. I would also like to thank my neighbors Katrin, Florian & Kalea, who supported me during the intense writing phase. Their kindness has made my journey much easier, and I am grateful for their friendship.

Lastly, I would like to thank my colleagues, whose insights and expertise have enriched my work. Their support, encouragement, and feedback have helped me grow personally and professionally, and I am honored to have had the opportunity to work alongside such talented individuals.

I would like to express my deepest appreciation and gratitude to my primary supervisor, Wolfgang Aigner, for his guidance, support, and encouragement throughout my doctoral studies. His extensive knowledge, insightful feedback, and dedication have been instrumental in shaping my research and helping me to develop the skills necessary for a successful academic career. I would also like to thank my second supervisor, Margit Pohl, for her valuable insights and contributions to my work. Her feedback and suggestions have been immensely helpful in refining my research methodology and enhancing the quality of my thesis. I am grateful to both Wolfgang Aigner and Margit Pohl for their guidance and mentorship, which have been instrumental in my academic and personal growth. Their willingness to share their expertise and passion for research has inspired me to pursue excellence in my academic pursuits.

This work was supported by The Austrian Research Promotion Agency (FFG) via the SEVA project (no. 874018) and by the Austrian Science Fund as part of the Vis4Schools project (I 5622-N).



## Kurzfassung

Die effektive Kommunikation von Daten und Erkenntnissen ist in der heutigen datengesteuerten Welt von entscheidender Bedeutung, aber die Interpretation von Visualisierungen erfordert ein gewisses Maß an "Visualisierungskompetenz". Damit ist die Fähigkeit gemeint, Informationen durch visuelle Darstellungen zu lesen, zu interpretieren und zu kommunizieren. Mangelnde Fähigkeiten im Bereich der Visualisierung können Einzelne daran hindern, wertvolle Erkenntnisse zu erlangen. Das Visualisierungs-Onboarding macht die User\*innen mit den grundlegenden Konzepten, Werkzeugen und Techniken vertraut, die zur effektiven Interpretation von Datenvisualisierungen erforderlich sind, die zunehmend an Bedeutung gewinnen.

In dieser Dissertation wird untersucht, wie ein Visualisierungs-Onboarding-Konzept konzipiert und entwickelt werden kann, das in Visual-Analytics-Tools (VA) integriert wird und die den Visualisierungen zugrunde liegenden Daten nutzt. Diese Haupt-Forschungsfrage wird in dieser Dissertation untersucht: Wie können Onboarding-Methoden Nutzer\*innen mit geringer Visualisierungskompetenz bei der korrekten Interpretation von Datenvisualisierungen unterstützen? Um diese Frage zu beantworten, wurde ein deskriptiver Designraum entwickelt, um Schlüsselfragen im Zusammenhang mit Visualisierungs-Onboarding zu beantworten. Die Literaturstudie validierte den deskriptiven Designraum. Designüberlegungen für Visualisierungs-Onboarding-Konzepte wurden auch aus den empirischen Studien der von uns untersuchten Arbeiten abgeleitet. Als Nächstes untersuchten wir verschiedene Onboarding-Methoden, indem wir vier Onboarding-Konzepte entwickelten und sie in Usability-Studien mit MTurk-Mitarbeiter\*innen und Studierenden testeten, um ihre Auswirkungen auf die Leistung und Erfahrung der Nutzer\*innen zu bewerten. Außerdem haben wir in einer separaten Usability-Studie abstrakte und konkrete Onboarding-Anweisungen mit Studierenden verglichen. Darüber hinaus entwickelten wir ein Visualisierungskonzept für das Onboarding und validierten es iterativ durch ein kognitives Walkthrough mit Expert\*innen. Darauf folgte ein Rapid-Prototyping-Prozess zur Erstellung von VisAhoi, einer JavaScript-Bibliothek zur halbautomatischen Erstellung von Onboarding-Anweisungen für VA-Tools. Zur Anwendung unseres Visualisierungs-Onboarding-Konzepts haben wir zwei Designstudien durchgeführt. Die erste Designstudie konzentrierte sich auf die Integration von VisAhoi in ein visuelles Analysewerkzeug für die Analyse von Hochdurchsatz-Screeningdaten in der biomedizinischen R&D. Die zweite Designstudie befasste sich mit einem selbsterklärenden Visualisierungs-Onboarding-Konzept für datenjournalistische Anwendungsfälle. Schließlich stellen wir auf der Grundlage der

Erkenntnisse aus der Doktorarbeit Gestaltungsmaßnahmen zur Verfügung, die das Design und die Entwicklung von Visualisierungs-Onboarding leiten.

Visualisierungs-Onboarding ist nur ein Aspekt der Benutzerunterstützung. Eine weitere Komponente der Unterstützung ist Guidance, die darin besteht, der Benutzer\*in konkrete Vorschläge und Unterstützung für die nächsten Schritte bei der Exploration der Visualisierung zu geben. Wir veranschaulichen, wie das Visualisierungs-Onboarding und die Anleitung Benutzer\*innen in verschiedenen Analysephasen unterstützen können. Daher zeigen wir ein deskriptives Modell für die Integration von Visualisierungs-Onboarding und Guidance in den VA-Prozess. Um die praktische Anwendung unseres Modells zu demonstrieren, präsentieren wir ein Nutzungsszenario anhand eines Aktienhandels-Tools.

Zusammenfassen, die Hauptbeiträge dieser Arbeit sind ein Design Space für Visualisierungs-Onboarding und die Konzeption und Implementierung einer JavaScript-Bibliothek namens VisAhoi für die halbautomatische Generierung von Onboarding-Anweisungen auf Basis der der Visualisierung zugrunde liegenden Daten. Außerdem tragen wir mit den Ergebnissen zweier Designstudien in biomedizinischem R&D und Datenjournalismus bei, die die Anwendbarkeit von VisAhoi und dem Visualisierungs-Onboarding-Konzept zeigen. Darüber hinaus werden in dieser Dissertation eine Reihe von Gestaltungsmaßnahmen und ein Beschreibungsmodell vorgestellt.

### Abstract

Effective communication of data and insights are crucial in today's data-driven world. Albeit humans are visual beings and visual representations are easier to understand than other forms of data representations, we have to learn how to read and comprehend them. Interpreting visualizations correctly and taking full advantage of the visualization at hand requires a certain level of "visualization literacy". This refers to the ability to read, interpret, and communicate information through visual representations.

Visualization onboarding introduces individuals to the basic concepts, tools, and techniques required to interpret data visualization effectively, which is becoming increasingly important. This doctoral thesis explores how to design and develop a visualization onboarding concept that can be integrated into visual analytics (VA) tools and use the underlying data of the visualizations. The main research question of this doctoral thesis is: How can onboarding methods support users with low visualization literacy in correctly interpreting data visualizations?

To answer this question, we developed a descriptive design space to answer key questions related to visualization onboarding. The literature survey validated the descriptive design space by applying it to state-of-the-art onboarding methods. Design considerations for visualization onboarding concepts were also derived from the empirical studies of the papers we surveyed. Next, we explored different onboarding methods by developing four onboarding concepts. We tested them through usability studies with MTurk workers and students to evaluate their impact on user performance and experience. In a separate usability study, we also compared abstract and concrete onboarding instructions with students.

Furthermore, we developed a visualization onboarding concept and validated it iteratively through a cognitive walkthrough with experts. A rapid prototyping process followed this to create VisAhoi, a JavaScript library for semi-automatically generating onboarding instructions for VA tools. To apply our visualization onboarding concept, we conducted two design studies. The first study focused on integrating VisAhoi into a visual analytics tool for analyzing high-throughput screening data in biomedical R&D. The second study involved a self-explanatory visualization onboarding concept for data-journalistic use cases. Finally, we provide design actions based on the insights from the doctoral research to guide the design and development of visualization onboarding. Visualization onboarding is just one aspect of user assistance. Another component of user assistance is guidance, which involves providing users with concrete suggestions and support for the next steps in the exploration. We illustrated how visualization onboarding and guidance could assist users in various analysis phases. Hence, we introduce a descriptive model for integrating visualization onboarding and guidance into the VA process. To demonstrate the practical application of our model, we present a usage scenario using a stock trading tool.

The main contributions of this thesis are a design space for visualization onboarding and the conceptual design and implementation of a JavaScript library called VisAhoi for the semi-automatic generation of onboarding instructions based on the underlying data of the visualization. Besides, we contribute with the results of two design studies in biomedical R&D and data journalism, showing the applicability of VisAhoi and the visualization onboarding concept. Furthermore, a set of design actions and a descriptive model are presented.

# Contents

Kurzfassung				
A	bstra	let	xi	
Contents				
1	Intr	ntroduction		
	1.1	Motivation	2	
	1.2	Related Work & Background	4	
	1.3	Research Questions & Research Methodology	8	
	1.4	Contributions	11	
	1.5	Dissemination	12	
	1.6	Structural Overview	14	
	1.7	Conventions	14	
Ι	Exp	oloring Visualization Onboarding Approaches	15	
<b>2</b>	Descriptive Design Space			
	2.1	Introduction & Motivation	18	
	2.2	Related Work	19	
	2.3	Descriptive Design Space	24	
	2.4	Survey on Visualization Onboarding	28	
	2.5	Discussion & Conclusion	39	
3	Des	ign and Evaluation of Visualization Onboarding Concepts	41	
	3.1	Introduction & Motivation	42	
	3.2	Related Work	43	
	3.3	Research Design	44	
	3.4	Study 1: Visualization Onboarding and Effect on User Performance .	52	
	3.5	Study 2: Types of Visualization Onboarding Methods and the Effect on		
		User Performance	61	
	3.6	Study 3: In-situ Visualization Onboarding in Netflower	65	
	3.7	Discussion & General Lessons Learned	71	

xiii

	3.8	Conclusion	73
4	Abs teri 4.1 4.2 4.3 4.4 4.5 4.6	stract and Concrete Visualization Onboarding Instructional Ma- al for a Treemap Visualization         Introduction & Motivation         Related Work         Evaluation         Results         Discussion         Conclusion	<b>75</b> 76 76 80 85 92 94
5	Ref	lection & Lessons Learned	97
II	Des 10	signing and Implementing Visualization Onboarding Concepts	
6	Tow	vards a Design for Visualization Onboarding	103
	6.1	Introduction	104
	6.2	Visualization Unboarding Concept	104
	0.3		108
	6.5	Discussion & Lessons Learned	1109
	6.6	Conclusion	111
7	Vis	Ahoi:Design and Implementation of a Software Library	113
	7.1	Introduction & Motivation	114
	7.2	Method	115
	7.3	Objectives	115
	7.4	Related Work	116
	7.5	Visualization Onboarding Concept	120
	7.6	Design of the VisAhoi Library	124
	1.1	Conclusion & Future Work	130
	1.0		102
Π	I Ap	plying Visualization Onboarding Concepts in Design Studies	133
8	Des	ign Study 1: Authoring Tool for (Data)Journalists to Integrate	
	Vis	ualization Onboarding in a Treemap Visualization	135
	8.1	Motivation & Background	136
	8.2	Method	139
	8.3	Related Work	140
	8.4 0 5	Problem Unaracterization and Abstraction	141 145
	0.0		140

	8.6 8.7 8.8	Implementation	148 151 157				
9	9 Design Study 2: Visualization Onboarding in Visual Analytics tools						
	for 1	Biomedical Scientists	161				
	9.1	Motivation, Background & Related Work	162				
	9.2	Method	164				
	9.3	Problem Characterization and Abstraction	164				
	$9.4 \\ 9.5$	Conceptual Design & Implementation	168				
	9.6	Reflection Limitations & Lessons Learned	172				
	9.0 9.7	Conclusion	170				
IV	'Ger	neralization of Results & Outlook	181				
10	Des	ign Actions	183				
	10.1	Method	184				
	10.2	Design Actions	184				
	10.5	Summary	194				
11	Gui	dance, Visualization Onboarding, and Knowledge-assisted Visual					
11	Gui Ana	dance, Visualization Onboarding, and Knowledge-assisted Visual lytics	195				
11	<b>Gui</b> <b>Ana</b> 11.1	dance, Visualization Onboarding, and Knowledge-assisted Visual lytics Introduction & Motivation	<b>195</b> 196				
11	<b>Gui</b> <b>Ana</b> 11.1 11.2	dance, Visualization Onboarding, and Knowledge-assisted Visual         lytics         Introduction & Motivation	<b>195</b> 196 198				
11	Gui Ana 11.1 11.2 11.3	dance, Visualization Onboarding, and Knowledge-assisted Visual lytics Introduction & Motivation	<b>195</b> 196 198 202				
11	Guie Ana 11.1 11.2 11.3 11.4 11.5	dance, Visualization Onboarding, and Knowledge-assisted Visual lytics Introduction & Motivation	<b>195</b> 196 198 202 204				
11	<b>Gui</b> <b>Ana</b> 11.1 11.2 11.3 11.4 11.5	dance, Visualization Onboarding, and Knowledge-assisted Visual lytics Introduction & Motivation	<b>195</b> 196 198 202 204 210				
11	Gui Ana 11.1 11.2 11.3 11.4 11.5 11.6	dance, Visualization Onboarding, and Knowledge-assisted Visual lytics Introduction & Motivation	<b>195</b> 196 198 202 204 210 218				
11	Guia Ana 11.1 11.2 11.3 11.4 11.5 11.6 11.7	dance, Visualization Onboarding, and Knowledge-assisted Visual lytics Introduction & Motivation	<ul> <li>195</li> <li>196</li> <li>198</li> <li>202</li> <li>204</li> <li>210</li> <li>218</li> <li>222</li> </ul>				
11	Guia Ana 11.1 11.2 11.3 11.4 11.5 11.6 11.7 Out	dance, Visualization Onboarding, and Knowledge-assisted Visual         lytics         Introduction & Motivation         Related Work         Descriptive Model of Integrating Guidance and Visualization Onboarding         Aspects of Visualization Onboarding and Guidance         Usage Scenario on Visualization Onboarding and Guidance in the Financial         Domain          Discussion & Lessons Learned          look	<ul> <li>195</li> <li>196</li> <li>198</li> <li>202</li> <li>204</li> <li>210</li> <li>218</li> <li>222</li> <li>225</li> </ul>				
11 12 13	Guid Ana 11.1 11.2 11.3 11.4 11.5 11.6 11.7 Out Con	dance, Visualization Onboarding, and Knowledge-assisted Visual lytics Introduction & Motivation	<ul> <li>195</li> <li>196</li> <li>198</li> <li>202</li> <li>204</li> <li>210</li> <li>218</li> <li>222</li> <li>225</li> <li>229</li> </ul>				
111 122 133 V	Guid Ana 11.1 11.2 11.3 11.4 11.5 11.6 11.7 Out Con Bib	dance, Visualization Onboarding, and Knowledge-assisted Visual lytics Introduction & Motivation	<ul> <li>195</li> <li>196</li> <li>198</li> <li>202</li> <li>204</li> <li>210</li> <li>218</li> <li>222</li> <li>225</li> <li>229</li> <li>237</li> </ul>				
11 12 13 V	Guia Ana 11.1 11.2 11.3 11.4 11.5 11.6 11.7 Out Con Bib	dance, Visualization Onboarding, and Knowledge-assisted Visual         lytics         Introduction & Motivation         Related Work         Descriptive Model of Integrating Guidance and Visualization Onboarding         Aspects of Visualization Onboarding and Guidance         Usage Scenario on Visualization Onboarding and Guidance in the Financial         Domain         Discussion & Lessons Learned         Conclusion and Future Perspectives         look         clusion         liography, List of Figures & Tables, Glossary	<ul> <li>195</li> <li>196</li> <li>198</li> <li>202</li> <li>204</li> <li>210</li> <li>218</li> <li>222</li> <li>225</li> <li>229</li> <li>237</li> </ul>				
11 12 13 V Bi	Guia Ana 11.1 11.2 11.3 11.4 11.5 11.6 11.7 Out Con Bib	dance, Visualization Onboarding, and Knowledge-assisted Visual lytics Introduction & Motivation	<ul> <li>195</li> <li>196</li> <li>198</li> <li>202</li> <li>204</li> <li>210</li> <li>218</li> <li>222</li> <li>225</li> <li>229</li> <li>237</li> <li>239</li> </ul>				

List of Tables		
Glossary		
VIAppendix		
A Material Comparative Evaluation on Onboarding ApproachesA.1 Task CategoriesA.2 Hotjar Heatmaps	<b>3</b> 3 4	
B Material Abstract vs. Concrete Onboarding Instructions	7	
C Material Study Design Cognitive Walkthrough Inspection	11	
<ul> <li>D Material VisAhoi Library</li> <li>D.1 Screenshots of the Onboarding Concept using eCharts and Vega-lite .</li> </ul>	17 17	
E Material User Experience Evaluation of Design Study DDJ	19	
<ul> <li>F Material of Design Study Biomedical R&amp;D</li> <li>F.1 Screenshots of VA tool for HTS data analysis</li></ul>	<b>51</b> 94	
Curriculum Vitae		

# CHAPTER 1

## Introduction



In this chapter, we present the general motivation for this thesis (Chapter 1), which is based on visualization literacy as a fundamental skill to interpret more complex data visualizations correctly and how to support users in this by using visualization onboarding methods. Furthermore, we describe current work in the visualization literacy and onboarding fields (Chapter 1.2). The research questions and the general methodology for this thesis are outlined in Chapter 1.3. It will then go on to the contributions (Chapter 1.4), the dissemination activities related to this thesis (Chapter 1.5), and the structural overview in chapter 1.6.

#### 1.1 Motivation

Over the last decades, the amount and complexity of available data have grown tremendously. This increasing amount of available data offers immense opportunities to promote technological, economic, and societal success in many domains, such as industry, medicine, or science. However, the possibility of collecting and storing data increases faster than our ability to analyze and use it for decision-making. Visual interfaces, especially visualizations, are high bandwidth gateways for perceiving structures, patterns, or connections hidden in the data (Card et al., 1999). Therefore, **visualization** can help to alleviate the problem of information overload by taking advantage of the powerful human perceptual system, which is highly efficient in processing visual input to make sense of data, explore complex information spaces, or spot patterns and relationships.

Visualization has become more critical and widespread in science, business, and everyday contexts, such as data stories in newspapers, books, on the internet, or for personal data (e.g., sleep tracking, nutrition, sports, and more). Right now, the complexity and social relevance of the COVID-19 pandemic have put data visualization at the center of worldwide attention (Shneiderman, 2020). Since the outbreak, data visualization researchers and experts have been providing various data visualizations for public education. The general public got in touch with diverse data visualizations presenting medical data such as reproduction numbers, COVID-19 cases, hospitalization, and more. The size and complexity of today's datasets overwhelm traditional business charts such as bar charts, line charts, or pie charts. Therefore, more advanced visual representations are necessary to capture more complex data structures and significant amounts of data.

Visualization can be seen as transforming data into a visual form (Card et al., 1999). Card et al. (1999) define this visual structure as a set of marks (point, line, area, surface, volume), their retinal encoding (color, size, shape, gray-level, orientation, texture, connection, enclosure), and their positions in space and time (X, Y, Z, T). As a user, this transformation needs to be transparent to construct and decode the visual representations and correctly reason about the data. Even though humans are visual beings and visual representations are easier to understand than other data representations, users still need to learn how to create, read and comprehend them. Compared to reading and writing text, we usually do not learn how to construct, read and interpret visualizations in the course of our education apart from simple business charts (Börner et al., 2019). Hence, most people have difficulties interpreting and working with visual representations or comprehending characteristics of the underlying data (Börner et al., 2016; Galesic and Garcia-Retamero, 2010). This bears the risk of drawing wrong conclusions and leads to frustration or rejection of otherwise powerful data visualizations.

Boy et al. (2014) describe **visualization literacy** as "the ability to use well-established data visualizations (e.g., line graphs) to handle information in an effective, efficient, and confident manner". A recent examination of youth's and adults' ability to interpret and construct data visualizations by Börner et al. (2016) indicates that the general public has a low level of *visualization literacy*. Having limited visualization literacy skills can be



Figure 1.1: Two sides of the visualization process by (Cairo, 2013). The visuotextual encoding and transformations are mainly decided by the designer of the visual analytics (VA) system (left side). On the user side, the image needs to be decoded and understood (right side). This requires knowledge about the topic and familiarity with graphic forms (visualization literacy). Without this (meta-)knowledge, a user might be unable to interpret a visualization and draw correct conclusions. Visualization onboarding concepts can be applied to assist the user in closing the knowledge gap.

a severe handicap. It hinders people from accessing valuable information, which could help them learn and solve problems or make informed decisions. Understanding the *visual mapping* is the key to correctly constructing and interpreting visual representation and the underlying data, respectively. Visual mapping is the process of assigning data variables to visual variables resulting in a graphical representation (Card et al., 1999; dos Santos and Brodlie, 2004) (see Fig. 1.1). This process is the central component of all virtually known conceptual models of visualization, such as (Card et al., 1999; Chi, 2000; Wijk, 2006). However, especially for novice users, this task is often complicated and leads to wrong conclusions and insights concerning the data.

This doctoral research aims to develop **visualization onboarding** concepts that might be utilized to support users in reading, interpreting, and extracting information from visual representations of data (Stoiber et al., 2022b, 2021, 2022d). Firstly, we aim to explore visualization onboarding approaches from different perspectives. Drawing an overall picture of the current state of visualization onboarding, we developed a descriptive design space. Based on the resulting design space and derived design considerations, four different visualization onboarding concepts have been designed and evaluated to investigate onboarding in more detail. Moreover, we explored abstract and concrete onboarding instructions to improve the design and phrasing of instructions more efficiently. Secondly, we present the resulting visualization onboarding design and propose a JavaScript library, named VisAhoi, to generate self-explanatory visualization onboarding instructions semiautomatically utilizing the underlying data presented in the visualization. Finally, the resulting visualization onboarding concept and implementation have been applied to validate the design and technical implementation using two design studies in Datadriven journalism (DDJ) and biomedical research and development (R&D). Both design studies were carried out in the research project  $SEVA^1$ . Finally, we derived design actions (De Bruijn and Spence, 2008) for designing visualization onboarding concepts based on the lessons learned from the doctoral research. Finally, we present a descriptive

<sup>&</sup>lt;sup>1</sup>https://seva.fhstp.ac.at/en, Accessed: 2023-01-12

model to show how visualization onboarding should be utilized in different analysis phases to be effective and accepted by the user. It combines visualization onboarding and guidance as two concepts of *user assistance*, as well as knowledge-assisted VA, as both of them typically rely on the presence of a knowledge base.

#### 1.2 Related Work & Background

On the one hand, this section gives an overview of visualization literacy in general as a related concept to visualization onboarding. Moreover, on the other hand, it discusses visualization onboarding concepts in scientific literature, as well as the usage of it of commercial visualization tools.

#### 1.2.1 Background

Although the topics of onboarding and visualization literacy have been identified as future challenges by the VA community (e.g., in (Keim et al., 2010)), they have not received much attention so far (Börner et al., 2019, 2016; Firat et al., 2022a,b; Kirk, 2016).

In general, the term onboarding was originally coined in human resources management to support new employees in learning about the tasks that are part of their job within a particular company (Klein et al., 2015). This ongoing process aims to communicate formal knowledge about their tasks and informal knowledge about organizational culture and its unwritten rules to the new employees. This concept has been transferred to other domains such as human-computer interaction (HCI) (Banovic et al., 2012; Bergman et al., 2005; Carroll, 1987; Fernquist et al., 2011a; Grossman and Fitzmaurice, 2010a; Kelleher and Pausch, 2005; Palmiter and Elkerton, 1991). Different approaches have been investigated in depth since 1991 by Palmiter and Elkerton (1991), e.g., step-by-step wizards (Bergman et al., 2005; Fernquist et al., 2011b; Kelleher and Pausch, 2005), overlays (Fernquist et al., 2011b), Q & A forums (Matejka et al., 2011), or video-based tutorials (Banovic et al., 2012; Grossman and Fitzmaurice, 2010b; Pongnumkul et al., 2011). More recently, the focus of onboarding has shifted toward mobile applications. Hulik<sup>2</sup> introduced the concept of supporting users in learning smartphone applications and software tools. Kumar defined user onboarding as "the process of increasing the likelihood that new users become successful when adopting your product." (Kumar, 2017). Moreover, design practitioners such as Dominikus Baur & Moritz Stephaner explain that onboarding design is a critical step in making a successful visualization on the web (Baur and Stephaner, 2018). They proposed different concepts: tours, catchy headlines + one paragraph and gradual reveal (Reveal, 2019), storytelling, and narrative techniques. Tours can be described as tutorials with a predefined sequence of exploratory materials using, for example, overlays explaining the features of a visualization tool step-by-step. Catchy headlines and paragraphs use explanatory textual descriptions to onboard novice users. They also listed the "gradual reveal" concept to support users. "Gradual reveal works by

4

<sup>&</sup>lt;sup>2</sup>https://useronboard.com, Accessed: 2023-01-11

designing the order of the information so that one grouping of information can be shown, or focused on, at a time." (Reveal, 2019)

#### 1.2.2 Towards visualization onboarding

In general, there are only a few discussions about onboarding concepts for visualization techniques and visual analytics (VA) tools (Firat et al., 2022a,b). Dhanoa et al. (2022) proposed a process model for dashboard onboarding that formalizes and unifies onboarding strategies. Besides, Tanahashi et al. (2016) investigated top-down and bottom-up teaching methods and active or passive learning types in their introductory materials. The bottomup teaching method ("textbook approach") (Zeid et al., 2011) focuses on small, detailed pieces of information that students then combine to get a better understanding. Besides, a top-down teaching method is given when a broad overview first helps to understand the abstract, high-level parts of an idea/topic, which then provide context for understanding its components in detail (Tanahashi et al., 2016). Furthermore, a distinction can be made between active and passive learning types. Passive learning means that students only receive the information without participatory dialog. In contrast, active learning describes an active participation (Tanahashi et al., 2016). Their analysis shows exercise questions improve the performance of the participants. Active participation (using active learning) through questions had the best effect. Their results also showed that top-down exercises were more effective than bottom-up. And their materials utilizing active learning with top-down exercise questions were the most effective.

In their comparative study, Kwon and Lee (2016) explored the effectiveness of active



In parallel coordinates, **one line** represents **one row** of a table and **each axis** represents **each column**. In this page, you'll draw parallel coordinates by clicking points on each axis per car. Once you finish drawing parallel coordinates of three cars, you will proceed to the next page. Now, please click on the points in parallel coordinates corresponding to the highlighted cell in the table.

Figure 1.2: Onboarding for parallel coordinates plot(Kwon and Lee, 2016)



Figure 1.3: Learning-by-analogy: In-betweens of linear chart and spiral chart (Ruchikachorn and Mueller, 2015)

learning strategies. Three tutorial types—static, video-based, and interactive (shown in Figure 1.2)—were used to support the learning of parallel coordinates plot visualizations. They observed that participants who used interactive and video tutorials outperformed those who used static or no tutorials.

Ruchikachorn and Mueller (2015) found out that the learning-by-analogy concept is helpful as participants in their study could entirely or at least significantly understand the unfamiliar visualization methods better after they observed or interacted with the transitions from the familiar counterpart. They assessed four combinations and compared their difference in visual literacy: scatterplot matrix against hyperbox <sup>3</sup>, linear chart against the spiral chart (illustrated in Figure 1.3), hierarchical pie chart against treemap, and data table against parallel coordinates plots. As the demonstrations are visual only, it bridges any language barriers.

Additionally, Firat et al. (2020) developed an interactive pedagogical method for training and cognition of a treemap design and a treemap literacy test. The user study showed that students who interacted with the teaching tool outperformed those who learned through slides. Recently, Peng et al. (2022) present results of a study to evaluate six parallel coordinate literacy modules based on Bloom's taxonomy (Bloom et al., 1956) using videos, tests, and tasks. The intervention is based on a corrective feedback mechanism inspired by Mastery Learning (Block and Burns, 1976; McCane et al., 2017). It is an instructional strategy that requires the students to master a concept first before moving on to a more complex one. Their results show that students using the developed modules were more accurate and better understood the components and working of a parallel coordinates plot.

Moreover, Wang (2022) presents a toolkit for teaching flow visualization concepts (VisVi-

<sup>&</sup>lt;sup>3</sup>Hyperbox is a scatterplot matrix which allows non-orthogonal axes (Alpern and Carter, 1991).

sual), including web-based educational tools for instructors teaching and students learning graph visualizations (GraphVisual) (Martin et al., 2020), tree visualizations (TreeVisual) (O'Handley et al., 2022), and volume visualizations (Bao et al., 2021). Each tool covers two modes: *study mode* and *quiz mode* (Wang, 2022). These educational tools use questions to test students' general knowledge (quiz mode) about the given topic and support learning.

In addition to scientific literature, onboarding concepts are integrated into commercial visualization tools. Nowadays, most commercial visualization tools integrate onboarding concepts focusing on explaining features. Yalçın (2016) presented "HelpIn", a contextual in-situ help system design to explain parts of the visualization tool Keshif (2023). Chundury et al. (2023) propose data-driven, contextual, in-situ help features that can be integrated into visualization interfaces. They introduced five different modes: (1) contextual help on selected interface elements, (2) topic listing, (3) overview, (4) guided tour, and (5) notifications. The (4) guided tour mode is illustrated in Figure 1.4. Furthermore, IBM Cognos Analytics (2019) uses step-by-step tours with tooltips and overlays for onboarding new users. The commercial visualization tool Advizor (Advizor Solutions, 2023) uses a more traditional approach, which uses textual descriptions to explain the visual mapping for visualization techniques.



Figure 1.4: The Guided Tour mode (Chundury et al., 2023). The tour can be guided forward, backward, or to a specific step shown in a progress bar (1). The instructions are designed as tooltips and accessible by mouse-over on a visual glyph (2). Tooltips, as in (3), describe the main effect on the interface. The main tooltip (4) summarizes the tooltip and provides additional information. Related Topics (5) and (6) the context in which this topic can be accessed is also shown.

While existing learning environments and web platforms such as "The graphic continuum" (Swabish and Ribecca, 2014) or the "Data Viz Catalogue" (Ribecca, 2020) focus on the explanation of visualization techniques, the presented textual descriptions and illustrations are mainly manually developed by the visualization designer of the respective platform. This requires much effort from the visualization designer and developer. These onboarding concepts are generally tailored to specific visualization techniques and used data sets or are restricted as they are not reusable or integrated into existing visualization tools.

#### 1.2.3 Summary

In summary, a few onboarding methods exist in the literature using different strategies and educational theories, such as learning by doing (Kwon and Lee, 2016), learning by analogy (Ruchikachorn and Mueller, 2015), mastery learning (Peng et al., 2022), scaffolding (Bishop et al., 2020), or top-down and bottom-up teaching methods as well as active and passive learning types (Tanahashi et al., 2016). These concepts are stand-alone learning environments that are not integrable into VA tools. Only Yalçın (2016) and Chundury et al. (2023) developed an in-situ help for the visualization tool *Keshif* (Yalçın and Xharra, 2022). Nevertheless, this is restricted to the visualization tool. Further research is needed to identify effective onboarding concept designs and understand user behavior while using onboarding.

Next, we describe the thesis's research questions and methodology.

#### 1.3 Research Questions & Research Methodology

This doctoral thesis aims to study visualization onboarding concepts integrated into visual analytics (VA) tools to support users in effectively interpreting data visualizations. This leads to the following main research question:

• How can onboarding methods support users with low visualization literacy in correctly interpreting data visualizations?

The following research questions will be investigated to achieve this goal:

- What are the needs of domain experts when learning novel visual representations?
- How do onboarding concepts need to be designed and implemented to fulfill users' needs?
- How can existing educational theories be applied for visualization onboarding?
- How can a declarative visualization specification be used to (semi)-automatically generate onboarding idioms?

8

The thesis research generally uses well-established qualitative and quantitative methods following a user-centered design cycle to handle and consider research methods' different strengths and weaknesses. The research approach can be divided into four areas and is illustrated in Figure 1.5: (I) exploring visualization onboarding approaches, (II) designing and implementing onboarding concepts, and (III) the application of the developed visualization onboarding concept in design studies in the fields of biomedical R&D and data-driven journalism (DDJ), and (IV) the generalization of the results and outlook.

**Part I: Exploring Visualization Onboarding Approaches:** For developing the descriptive design space, we started with literature research for visualization onboarding concepts. We systematically surveyed the literature published in the leading venues in information visualization, visual analytics (VA), and human-computer interaction (Kitchenham, 2004). We also reviewed commercial visual analytic tools based on Behrisch et al. (2019). Finally, we used the identified literature to develop and validate our descriptive design space by applying it to describe and characterize state-of-the-art onboarding methods. We also derived design considerations for visualization onboarding concepts presented in Stoiber et al. (2022c).

We designed four onboarding concepts and developed proof-of-concept prototypes based on the design considerations. We conducted two quantitative comparative user studies with MTurk workers and a qualitative comparative usability study with students (Lazar et al., 2017, p. 25-44). The main aim of these studies was to investigate the effect of onboarding on user performance and evaluate the subjective user experience. Furthermore, we explored abstract and concrete onboarding instructions and assessed them in a quantitative comparative user study with students (Lazar et al., 2017, p. 25-44).

**Part II: Designing and Implementing Visualization Onboarding Concepts:** For the next part of the doctoral thesis, we iteratively designed a visualization onboarding concept and validated it along with a cognitive walkthrough with experts (Lazar et al., 2017, p. 269). Moreover, in a rapid prototyping process, we developed a JavaScript library, VisAhoi, to integrate and generate onboarding instructions for VA tools semi-automatically.

**Part III: Applying Visualization Onboarding Concepts in Design Studies** Visualization must be developed and tested in the context of specific tasks, users, and application areas (Munzner, 2009). To apply the visualization onboarding concept, we performed two design studies (SedImair et al., 2012).

We used a mixed-method approach for the first design study (Sedlmair et al., 2012) of the self-explanatory visualization onboarding concept for data-journalistic use cases. As is typical for design studies, we followed an iterative design and evaluation process to address a particular domain problem by involving collaborators and users from the domain (Sedlmair et al., 2012). The online interviews and a survey were chosen for the problem characterization and abstraction phase. Following the requirements analysis,

Exploring Visualization Onboarding Approaches									
Design Space	Design & Evaluation of different Onboarding Types	Abstract vs. Concrete Onboarding Instructions							
Systematic literature survey	Conceputal design & proof–of concept prototype Quantitative comparative user study with MTurk workers	Comparative quantitative user study with students							
	Qualitative comparative user study with students								

#### Designing and Implementing Visualization Onboarding Concepts

Visualization Onboarding Concept

Conceputal design & high-fidelity prototype

Cognitive walkthrough

VisAhoi JavaScript Library

Rapid prototyping

#### III Applying Visualization Onboarding Concepts in Design Studies

Design Study Datadriven journalism

Problem characterization and abstraction

Conceptual design & Implementation

User experience evaluation with domain experts

Design Study Biomedical R&D

Problem characterization and abstraction

Conceptual design & Implementation

Usage scenario development

#### IV Generalizing results & Outlook

#### Design Actions

onboarding

Creation of design actions describing design decisions for integrating visualization Descriptive Model

Model Building: Descriptive model of Onboarding, Guidance, and Knowledge-assisted visual analytics

Model validation by application scenario

Figure 1.5: Research methodology approach for the doctoral research

10

we produced high-fidelity prototypes based on the requirements and the results of our previous studies (Stoiber et al., 2022b, 2021, 2022d). We conducted a pilot study for the qualitative user evaluation with two experts in the field of HCI and Visualization and iteratively improved the prototype. Based on the feedback, we revised and improved the prototype for the following user experience evaluation. Finally, we analyzed the data to refine the proof-of-concept prototype and started with the reflection phase.

The second design study in the field of biomedical R&D presents the integration of VisAhoi in a VA tool for analyzing high-throughput screening (HTS) data interactively; we followed the well-known nested model for visualization design and validation proposed by Munzner (2009). In the first phase, problem characterization and abstraction (SedImair et al., 2012) were performed with domain experts to understand current practices and challenges better and further define requirements. Therefore, we conducted semi-structured interviews (Lazar et al., 2017) with domain experts in bioinformatics, biology, and chemistry. As a next step, we started with the visualization and interaction design followed by the algorithm design and implementation based on a user-centered design process (Preece et al., 2015). We interviewed experts using Thinking Aloud (Nielsen, 2012) to validate the developed visualization onboarding concept. We also show the applicability in a usage scenario (SedImair et al., 2012) by integrating the VisAhoi JavaScript library in the VA for analyzing HTS data. Finally, we analyzed the data to improve the library and started with the reflection.

We formulated take-home messages and lessons learned based on the insights gained from the design studies.

**Part IV: Generalization of Results & Outlook:** As a final step of this thesis, we present design actions (De Bruijn and Spence, 2008) for designing and developing visualization onboarding based on the lessons learned from the doctoral research. Besides, we present a descriptive model for how visualization onboarding and guidance as two concepts of user assistance—can be effectively integrated into the VA process (Stoiber et al., 2022a). Additionally, we describe a usage scenario of our model using the stock trading VA tool *Profit* to show how visualization onboarding and guidance can support users in different analysis phases.

#### 1.4 Contributions

This work explores visualization onboarding to support users in interpreting more complex data visualizations. Understanding how to design such concepts and integrate visualization onboarding technically and conceptually into a VA tool will benefit human-computer interaction, the visualization community, and designers and developers in biomedical R&D and DDJ. The contributions of this thesis are as follows:

• A design space for visualization onboarding (Chapter 2).

- A systematic literature review of existing visualization onboarding concepts and their classification into the design space (Chapter 2.4).
- Empirical results of three large-scale comparative user studies to explore different onboarding approaches and the effect on the user performance and subjective user experience (Chapter 3).
- Conceptual design and implementation of a JavaScript library (VisAhoi) for semiautomatic generation of onboarding instructions utilizing the underlying data (in Part II).
- A design study on semi-automatically generated onboarding instructions integrated into a treemap visualization for data-journalism use cases (Chapter 8).
- A design study on semi-automatically generated onboarding instructions for biomedical use cases and the integration in a VA tool (Chapter 9).
- Design actions for designing visualization onboarding methods that address the design, interaction concept, and phrasing of onboarding instructions (Chapter 10).
- A descriptive model of visualization onboarding and guidance as a concept of user assistance to show how these concepts can be utilized in different phases of the analysis to be effective and accepted by the user of a VA system (Chapter 11).

#### 1.5Dissemination

Parts of the presented results in this work have already been presented and published at scientific conferences and journals. All these results have been restructured, revised, and expanded, as well as set in the general context of this dissertation. In the following list, the publications are listed descending:

- Pohl, M., Potzmann, B., Stoiber, C., & Aigner, W. (2023). Visualization Onboarding Designed by Users: An Empirical Study. CELDA 2023. 20th International Conference on Cognition and Exploratory Learning in Digital Age, Funchal, Madeira, Portugal.
- Stoiber, C., Moitzi, D., Stitz, H., Girardi, D., Streit, M., & Aigner, W. (Under Review). Design of Visualization Onboarding Concepts for a 2D Scatterplot of Biomedical Visual Analytics Tools. VisComm Workshop, IEEE VIS, Melbourne, Australia.
- Stoiber, C., Pohl, M., & Aigner, W. (conditionally accepted). Design Actions for the Design of Visualization Onboarding Methods. EduVis Workshop, IEEE VIS, Melbourne, Australia.

- Stoiber, C., Pohl, M., & Aigner, W. (conditionally accepted). Design Actions for the Design of Visualization Onboarding Methods. EduVis Workshop, IEEE VIS, Melbourne, Australia.
- Stoiber, S., Grassinger, F., Moitzi, D., Girardi, D., Stitz, H., & Aigner, W. (Under Review). VisAhoi: Towards a Semi-Automated Generation of Visualization Onboarding Using High-level Visualization Grammars. Elsevier Visual Informatics.
- Stoiber, C., Radkohl, S., Grassinger, F., Moitzi, D., Girardi, D., Stitz, H., & Aigner, W. (2023). Authoring tool for Data Journalists integrating Self-Explanatory Visualization Onboarding Concept for a Treemap Visualization. CHIALY'23, Turin, Piedmont, Italy.
- Stoiber, C., Grassinger, F., & Aigner, W. (2022). Abstract and Concrete Materials: What to use for Visualization Onboarding. The 15th International Symposium on Visual Information Communication and Interaction (Vinci 2022). VINCI 2022, Chur, Switzerland. https://doi.org/10.1145/3554944.3554949
- Stoiber, C., Walchshofer, C., Pohl, M., Potzmann, B., Grassinger, F., Stitz, H., & Aigner, W. (2022). Comparative Evaluations of Visualization Onboarding Methods. Elsevier Journal of Visual Informatics. https://doi.org/10.1016/j.visinf.2022.07.001
- Stoiber, C., Ceneda, D., Wagner, M., Schetinger, V., Gschwandtner, T., Streit, M., Miksch, S., & Aigner, W. (2022). Perspectives of visualization onboarding and guidance in VA. Visual Informatics. https://doi.org/10.1016/j.visinf.2022.02.005
- Stoiber, C., Wagner, M., Grassinger, F., Pohl, M., Stitz, H., Streit, M., Potzmann, B., & Aigner, W. (2022). Visualization Onboarding Grounded in Educational Theories. In Visualization Psychology. Springer Nature. http://arxiv.org/abs/2203.11134
- Stoiber, C., Walchshofer, C., Grassinger, F., Sitz, H., Streit, M., & Aigner, W. (2021). Design and Comparative Evaluation of Visualization Onboarding Methods. VINCI'21 Short Papers, 1–5. https://doi.org/10.1145/3481549.3481558
- Gutounig, R., Radkohl, S., Goldgruber, E., & Stoiber, C. (2020). Datenjournalismus: Die Transformation journalistischer Arbeitsabläufe und Produkte durch Visualisierung und Analyse von Daten. In Springer Gabler.
- Stoiber, C., Grassinger, F., Pohl, M., Stitz, H., Streit, M., & Aigner, W. (2019). Visualization Onboarding: Learning How to Read and Use Visualizations. IEEE Workshop on Visualization for Communication. IEEE Symposium on Information Visualization, Vancouver, BC, Canada. https://doi.org/10/gh38zd
- Stoiber, C., Wagner, M., Ceneda, D., Pohl, M., Gschwandtner, T., Miksch, S., Streit, M., Girardi, D., & Aigner, W. (2019). Knowledge-assisted Visual Analytics meets Guidance and Onboarding. IEEE Application Spotlight. IEEE Symposium on Information Visualization, Vancouver, BC, Canada.

#### 1.6 Structural Overview

This thesis is divided into four parts (see Figure 1.5): Exploring visualization onboarding approaches, designing and implementing visualization onboarding concepts, applying visualization onboarding concepts in design studies, and generalization of results & outlook. The first part models a design space for visualization onboarding and applies it to the existing literature in the sense of a state-of-the-art report. We also report on the design and evaluation of different visualization onboarding concepts. In the second part of the doctoral thesis, we document the resulting visualization onboarding concept and implementation of a JavaScript Library. In the third Part III, we report on the results of two design studies applying the onboarding design. In the fourth Part IV, we generalize the results by presenting design actions (De Bruijn and Spence, 2008) and a descriptive model.

#### 1.7 Conventions

The work presented in this doctoral thesis was conducted in the context of the Austrian Research Promotion Agency (FFG)-funded research project "Self-Explanatory Visual Analytics for Data-Driven Insight Discovery" (SEVA) (the Austrian Ministry for Transport, Innovation and Technology (BMVIT) under the ICT of the Future program; no. 874018)<sup>4</sup>. The designs and evaluations were elaborated on and carried out by myself. The research and results have been shaped based on valuable discussions with my supervisors and collaboration partners of the SEVA project. For this reason, I decided to use the pronoun "we" instead of "I" in this thesis.

<sup>&</sup>lt;sup>4</sup>https://seva.fhstp.ac.at/en, Accessed: 2023-01-11

# Part I

# Exploring Visualization Onboarding Approaches



# CHAPTER 2

# **Descriptive Design Space**



This chapter describes the descriptive design space on visualization onboarding focusing on educational theories. We model the design space using a systematic literature review (Section 2.4). Finally, we derive design considerations from the collected publications and provide guidelines for applying educational theories for visualization onboarding (Section 2.4.4). This chapter is based on our published work (Stoiber et al., 2022c).

#### 2.1 Introduction & Motivation

Visualization onboarding aims to support end users in comprehending data visualizations and taking full advantage of the tools at hand (Stoiber et al., 2019a, 2022d). With effectively designed visualization onboarding methods, the knowledge gap of users could be filled. Thus, it makes sense to tap into the field of educational theories as well as identify how onboarding can benefit from explicit knowledge (see Figure 2.1).

In the literature, several different possibilities for how to realize onboarding have been suggested (see Table 7.1). Some of them are primarily based on cognitivist approaches (e.g., tutorials) (Kwon and Lee, 2016; Ola and Sedig, 2017) or Gestalt psychology (using analogy as a learning principle (Ruchikachorn and Mueller, 2015)). The educational theory on which these solutions are based is sometimes reflected explicitly and sometimes not. Informal evidence indicates that tutorials are often not read, and users proceed and start working and exploring system features themselves. Nevertheless, commercial systems often rely on tutorials and helpful websites as onboarding systems, e.g., (Advizor Solutions, 2023; IBM Cognos Analytics, 2019; Microsoft Power BI, 2019; Tableau Software, 2023). The discussion of this topic could help to clarify which approaches in the design of onboarding systems are more helpful than others.



Figure 2.1: Motivation for a descriptive design space for visualization onboarding

Therefore, we develop a descriptive design space, presented in Figure 2.4, covering aspects of visualization onboarding, especially focusing on educational theories. We conduct a systematic literature review to identify state-of-the-art visualization onboarding and categorize the work by summarizing existing onboarding concepts in scientific publications and commercial visualization tools using the *Five W's and How* (Hart, 1996, 2002).

WHY is visualization onboarding needed?
WHAT is visualization onboarding?
WHO is the user? Which knowledge gap does the user have?
HOW is visualization onboarding provided?
WHERE is visualization onboarding provided?
WHEN is visualization onboarding used?

We also derive design considerations from the collected publications and provide existing guidelines for applying educational theories for visualization onboarding in Section 2.4.4.

Overall, we can report that whether other approaches are better for onboarding or not is still an open question. Empirical research based on educational theories could help to gain more systematic information about this area.

#### 2.2 Related Work

As visualization onboarding aims to fill users' knowledge gaps by supporting the learning of new concepts, it makes sense to build upon knowledge from the fields of learning theories and cognitive science (see Figure 2.1). Therefore, we present the related work for visualization onboarding, educational theories in visualization and cognitive science, and how explicit knowledge relates to onboarding in the following subsections.

#### 2.2.1 Visualization Onboarding

In the literature, very little is known about designing and integrating visualization onboarding concepts in VA tools (Firat et al., 2022a; Stoiber et al., 2022c). Tanahashi et al. (2016) investigated *top-down* and *bottom-up teaching methods* and *active* or *passive learning types*. The bottom-up teaching method ("textbook approach") (Zeid et al., 2011) focuses on small, detailed pieces of information that students then combine to get a better understanding. Besides, a top-down teaching method is given when a broad overview first helps to understand the abstract, high-level parts of an idea/topic, which then provide context for understanding its components in detail (Tanahashi et al., 2016). Furthermore, a distinction can be made between active and passive learning types.



Figure 2.2: Examples of **bottom-up and top-down exercises** for a scatterplot (Tanahashi et al., 2016)

Passive learning means that students only receive the information without participatory dialog. In contrast, active learning describes an active participation (Tanahashi et al., 2016). In Figure 2.2, the left image shows the bottom-up material with an identical question, and the right one shows an example of a compare question for a top-down teaching method. The text and buttons below the visualizations illustrate the passive and active exercise types. For the active exercise, the users can select one of the four available answers and get feedback on whether they are correct. The passive exercise immediately answered the question.

In their comparative study, Kwon and Lee (2016) explored the effectiveness of *active learning* strategies. Three tutorial types—static, video-based, and interactive (see Figure 2.6 (3))—were used to support the learning of parallel coordinates plot visualizations. They observed that participants who used interactive and video tutorials outperformed those who used static or no tutorials. Their analysis indicates that top-down exercises were more effective than bottom-up and active learning types, with top-down tasks being the most effective.

Ruchikachorn and Mueller (2015) found out that the *learning-by-analogy* concept is helpful as participants in their study could entirely or at least significantly understand the unfamiliar visualization methods better after they observed the transitions from the familiar counterpart. They assessed four combinations and compared their difference in visual literacy: scatterplot matrix against hyperbox <sup>1</sup>, linear chart against spiral chart (see Figure 2.3), hierarchical pie chart against treemap, and data table against parallel coordinates plots. The authors also describe another advantage of learning-by-analogy over other forms of demonstrations, such as textual or oral descriptions, is the power of visuals, as they bridge any language barriers.

The educational community has also studied how students interpret and generate data visualizations (Baker et al., 2001). Furthermore, Alper et al. (2017) investigated how to teach bar charts in early grades using a tablet app called "C'est la vis" supporting elementary school pupils to learn how to interpret bar charts based on the *concreteness fading* approach. Concreteness fading is a pedagogical method to present new concepts with concrete examples at first before progressively abstracting them (see Figure 2.5 (2)). Recently, Bishop et al. (2020) developed a tablet-based tool called Construct-A-Vis, which supports elementary school children in creating a visualization based on free-form activities. They used scaffolding as a pedagogical method. Scaffolding is "the process that enables a child or novice to solve a problem, carry out a task, or achieve a goal which would be beyond his unassisted efforts" (Wood et al., 1976), issued in education to help children acquire new knowledge and skills. In detail, they integrated feedback mechanisms by showing if the visual mapping was correct.

Recently, Peng et al. (2022) present results of a study to evaluate six parallel coordinate literacy modules based on Bloom's taxonomy (Bloom et al., 1956) using videos, tests, and tasks. The intervention is based on a corrective feedback mechanism inspired by

<sup>&</sup>lt;sup>1</sup>Hyperbox is a scatterplot matrix which allows non-orthogonal axes (Alpern and Carter, 1991).


Figure 2.3: Learning-by-analogy: In-betweens of linear chart and spiral chart (Ruchikachorn and Mueller, 2015)

Mastery Learning (Block and Burns, 1976; McCane et al., 2017). It is an instructional strategy that requires the students to master a concept first before moving on to a more complex one. Their results show that students using the developed modules were more accurate and better understood the components and working of a parallel coordinates plot. Moreover, Wang (2022) presents a toolkit for teaching flow visualization concepts (VisVisual), including web-based educational tools for instructors teaching and students learning graph visualizations (GraphVisual) (Martin et al., 2020), tree visualizations (TreeVisual) (O'Handley et al., 2022), and volume visualizations (Bao et al., 2021). Each tool covers two modes: *study mode* and *quiz mode* Wang (2022). These educational tools use questions to test students' general knowledge (quiz mode) about the given topic and support learning.

In addition to scientific literature, onboarding concepts are also integrated into commercial visualization tools. Nowadays, most commercial visualization tools already integrate onboarding concepts focusing on explaining features; see Table 7.1. Yalçın (2016) presented HelpIn, a design of a contextual in-situ help system to explain features of Keshif (Keshif, 2023). More recently, Chundury et al. (2023) propose data-driven, contextual, in-situ help features that can be integrated into visualization interfaces based on the work by (Yalçın, 2016). Furthermore, IBM Cognos Analytics (2019) (see Figure 2.5 (1)), for example, uses step-by-step tours with tooltips and overlays for onboarding new users. A more traditional approach is used by the commercial visualization tool Advizor Solutions (2023), which makes use of textual descriptions to explain the visual mapping (seen in Figure 2.6 (1)) for visualization techniques.

Besides, there are platforms and websites available which can be categorized as external onboarding methods (Stoiber et al., 2019a) supporting users in understanding the visual mapping of various visualization techniques. For instance, *The graphic contin*- *uum* (Swabish and Ribecca, 2014) provides an overview of visualization types and supports design and method decisions. Similarly, the *Data Visualisation Catalogue* (Ribecca, 2020) seeks to support users in understanding different visualization types' encoding and building blocks. Furthermore, *From Data to Viz* (Holtz and Healy, 2018) aims to find an appropriate visualization type based on the input data using a decision tree. The catalog offers definitions, variations, the use of each visualization type, and potential issues that may arise during use and interpretation. These systems are not related to a particular visualization tool nor integrate any educational theories. In recent literature, Wang et al. (2020) presented a set of cheat sheets to support visualization literacy around visualization techniques inspired by infographics, data comics, and cheat sheets that are established onboarding methods in domains such as machine learning.

#### 2.2.2 Educational Theories in Visualization and Cognitive Science

Visualization onboarding supports users in learning new concepts (Stoiber et al., 2019a); therefore, it makes sense to build upon the knowledge from learning theories and cognitive science. We distinguish between three main educational theories: *behaviorism*, *cognitivism*, and *constructivism* (Ertmer and Newby, 2013). Behaviorism is an educational theory that only focuses on objectively observable behaviors and discounts any independent activities of the mind (Watson, 1997). It is based on positive and negative reinforcement techniques.

Besides, Cognitivism is a learning philosophy founded on the premise that learning can be modeled as a kind of information processing (Ertmer and Newby, 2013). Each of us generates our own "rules" and "mental models," which we use to make sense of our experiences. Therefore, learning is adjusting our mental models to accommodate new experiences. E-learning systems often integrate elements from different educational theories. This also applies to most onboarding systems.

Constructivist theories are the most appropriate for explaining learning processes with onboarding systems because they reflect on the application of learning in a practical context. Cognitive apprenticeship plays an essential role in constructivism (Duffy et al., 1993; Sawyer and Greeno, 2009). *Cognitive apprenticeship* is a kind of guided participation by learners in natural processes of knowledge generation. This is related to the concept of *scaffolding* (Holton and Clarke, 2006) where teachers gradually reduce the student's support level until the student can work autonomously. Cognitive apprenticeship and scaffolding can explain onboarding processes because the learners' goal is to solve a real task, while the guidance is gradually reduced.

Another theoretical framework relevant for onboarding is graph comprehension, a theory that aims to explain how users make sense of graphs. Most of the investigations in this context deal with simple, small graphs (Pohl et al., 2012). Nevertheless, the findings from graph comprehension yield exciting results that can inform the design of visualizations. This is especially valuable for onboarding systems because investigations in this area often address the issue of how to design graphs that are appropriate for use in educational

contexts. One of the most influential models in the context of the theory of graph comprehension describes this activity as consisting of three stages (Friel et al., 2001). These three stages are (1) reading the data (i.e., finding individual data values), (2) reading between the data (i.e., finding relationships between the data), (3) going beyond the data (i.e., interpreting the data, developing hypotheses about the data). Educational graphs are supposed to support all three stages, but the ultimate goal is to induce learners to "go beyond the data", that is, to reflect on the data and draw conclusions. Shah et al. (2005) argue that inexperienced users typically concentrate on single data points or single lines in line graphs, whereas experts can interpret patterns in the data. Peeck (1994) investigated whether it is possible to motivate learners to process graphs more comprehensively. In this context, the author successfully tested whether specific instructions for the processing of graphs support learning. The author also postulates that other measures, such as cues to draw the learner's attention or motivate the learner to solve simple tasks using graphs, are beneficial. Based on this approach, it can be recommended that onboarding should primarily support "going beyond the data" and that instructions and visual cues can help users better understand visualizations.

A further learning theory relevant for onboarding is *Microlearning*. Microlearning as an approach is a reaction to several technological developments. First, mobile technologies enable learners to learn flexibly, e.g., on the way to work, while traveling on public transport, or while waiting for a physician. In addition, microlearning is also relevant for workplace learning and continuing education (Schmidt, 2007). Employees in companies or other organizations do not need lengthy explanations but focused information necessary to continue their work. Microlearning has been defined as "special moments or episodes of learning while dealing with specific tasks or content and engaging in small but conscious steps." (Hug and Friesen, 2009). Microlearning usually encompasses small units of learning that never take longer than 15 minutes. The situation described for Microlearning in workplace learning is similar to that of users of complex information visualization systems.

Finally, Gestalt psychology is a theory that might be relevant for the design of onboarding systems. It is well-known that Gestalt psychology has made significant contributions to perceptual psychology. It is less well-known that Gestalt psychologists also conducted relevant research in the area of reasoning and problem-solving (see, e.g., (Mayer, 1995), (Holyoak and Thagard, 1995)). This is especially interesting for the design of visualizations as Gestalt psychologists conceptualized problem-solving as the (sudden) perception of structure in a problem domain. The so-called "Aha"-moment is when pieces fall into place, and a coherent structure is identified. In this context, analogies play an essential role because transferring structural knowledge from a well-known domain to an unknown domain is one of the learning methods that Gestalt psychologists suggested. Analogies can also support onboarding in improving the understanding of complex visualizations, as shown in the concept by Ruchikachorn and Mueller (2015).

#### 2.2.3 Knowledge Integration for Onboarding

In this section, we describe how user onboarding can benefit from *explicit knowledge* sources and contribute to generating new *(tacit) knowledge* and insights (Federico et al., 2017; Wang et al., 2009).

In this work, we mainly consider *explicit knowledge* (Wang et al., 2009) as the source for providing onboarding. Explicit knowledge describes the users' computerized representations of interests and domain knowledge. Usually, two types of prior knowledge are needed by a user to analyze data: operational knowledge (how to interact with the information visualization system), and *domain knowledge* (how to interpret the content) (Chen, 2005). While focusing on usability and perception- and cognition-aware design can alleviate the need for operational knowledge, domain knowledge cannot be easily replaced (Chen, 2005). Stoiber et al. (2019a) further enhanced the levels of the users' prior knowledge for visualization onboarding based on the nested model (Munzner, 2009) as 1) domain knowledge (e.g., vocabulary and concepts); 2) data knowledge (understanding the particular datatype); 3) visual encoding knowledge (understanding the visual mapping); 4) interaction knowledge (for performing tasks and understand relations in the data); and 5) analytical knowledge (knowledge of different automated data analysis methods) – see Section 2.3.2 for more details. However, Chen (2005) as well as Stoiber et al. (2019a) described the term *prior knowledge* at different granularities, whereby operational knowledge (Chen, 2005) can be seen as similar to the combination of visual encoding, interaction, and analytical knowledge (Stoiber et al., 2019a).

Based on the previously introduced terminology, we further characterize knowledge in Section 2.3.2, listing all possible knowledge types needed to support meaningful onboarding.

# 2.3 Descriptive Design Space

Covering different aspects of visualization onboarding in a descriptive design space can clarify which approaches for designing and integrating onboarding concepts in VA can be applied. In this section, we introduce the descriptive design space for visualization onboarding and discuss the role of educational theories in onboarding.

#### 2.3.1 Construction of Design Space

We structured the design space based on *Five W's* and the appended *How* (Hart, 1996, 2002). These questions frequently describe a matter from its most relevant angles in technical documentation and communication. Furthermore, the same questions have already been employed for structuring visualization for healthcare informatics (Zhang et al., 2013) and in a survey on the role of visual analytics in deep learning research (Hohman et al., 2019). We describe the space of visualization onboarding along the following questions:

24

**Tool-specific** 

Educational Theory e.g., concretness

fading, learning-bydoing, learning-by-

analogy, scaffolding,.

Yes | No

WHO is the user? Which knowledge gaps does the user have?HOW is visualization onboarding provided?WHERE is visualization onboarding provided?WHEN is visualization onboarding used?

Inside each dimension (question), we defined several categories described in detail in the section below. We followed an open coding approach for the survey of onboarding concepts, where we unified top-down approaches and bottom-up categorizations. Where available, we used existing taxonomies or frameworks, which we adapted to the specifics of visualization onboarding.

# 2.3.2 Design Space Dimensions

Visualization onboarding aims to support humans in dealing with large and complex information structures, make them more comprehensible, facilitate exploration, and enable knowledge discovery. Nevertheless, users often have problems reading and interpreting data from visualizations, particularly when they experience them for the first time. In this section, we present the design space dimensions of visualization onboarding and show its various aspects.

**4** HOW

Туре

n'm

context-sensitive

context-free

embedded

**5** WHERE

external

How is visualization onboarding provided?

vizards, help cen

mousover popups,

Where is visualization onboarding provided?

e.g., textual instructions, video, illustrations/figures, animations,

Context-sensitivity Interaction

e.g., guided tour, step-by-step

documentations, tooltips, overlays,

passive

active

reactive

#### **1** WHY

#### Why is visualization onboarding needed?

The aim of visualization is to support humans in dealing with large and complex information structures, to make them more comprehensible, facilitate exploration, and enable knowledge discovery. But, users often have problems in reading and interpreting data from visualizations, in particular when they experience them for the first time.

#### **2** WHAT

#### What is visualization onboarding?

Visualization onboarding is the process of supporting users on how to read, interpret, and extract information of visual representations of data.

#### **3** WHO

Who is the user? Which knowledge gap does the user have?

	Internal			
Domain knowledge				
Data knowledge	<b>O</b> WHEN			
Visual encoding & Analytical knowledge	When is visualization onboarding used			
	before while			

Figure 2.4: A visual **overview of the onboarding design space** and of how all six questions, "Why, What, Who, How, Where, and When" relate to one another.

#### WHO is the user?

Users need to understand the process and reasoning that lead to visual appearance, interactive behavior, and findings. Making this process transparent to the users is a central aspect of the design of visual analytics solutions. To conceptualize this aspect, we adopt the nested model by Munzner (2009) as the guiding framework for presenting different levels of knowledge. The nested model is a unified approach that splits the design into four levels and combines these with appropriate evaluation methods to mitigate threats to validity at each level. Analytical knowledge—such as different automated data analysis approaches, machine learning methods, or statistical methods applied to the data—is necessary to understand complex visualization interfaces and data. To be able to cover visual analytic approaches and include automated data analysis components, we expand the original model by adding analytical methods alongside visual encoding/interaction idioms. The model components represent the different levels of knowledge that (a) visualization users need to interpret (interactive) visualization artifacts correctly and (b) visualization designers have to consider when developing onboarding concepts. Figure 2.4 (3) keeps the nesting but shows an altered representation of the different levels.

Based on the definitions of our understanding of knowledge in the context of visual analytics (see Section 2.2.3), we define the following levels: domain knowledge, data knowledge, knowledge of visual encoding & interaction concepts, and analytical knowledge.

**Domain knowledge**: A specific domain is a particular field of interest by target users of a visualization tool (e.g., medicine, data journalism, bioinformatics). Each domain has a vocabulary for describing the data and problems, workflows, and how data can be used to solve a problem. Domain knowledge is also an ensemble of concepts, intellectual tools, and informational resources that a user can draw upon to contextualize the visualized data.

**Data knowledge**: Many visualization tools are specific to a particular *type of data*, such as multivariate, hierarchical, network, or time-oriented data.

Data knowledge refers to the necessary knowledge for understanding the data types and structures or statistical properties of the data. Users need to know how to get their data into a specific visualization tool as a first step. This relates to a more technical level of knowledge about a particular file format (e.g., CSV, JSON) or structure of the data – data format – (e.g., order and data types of individual variables).

Visual encoding knowledge: This type of knowledge is the most obvious in the context of visualization, as it concerns the visual appearance of the data. Data elements are mapped to visual marks and channels to form visualizations. Understanding this mapping is the basis for correctly interpreting the visualization.

**Interaction knowledge**: Interactivity is crucial for visualization tools. An interactive visualization tool can support the investigation at multiple levels of detail, such as a high-level overview or fully detailed views showing only a small data subset only (Munzner, 2009). Understanding the interaction concepts used in a visualization tool is important for users for an active discourse with the data, e.g., to perform tasks and understand connections and relationships in the data.

**Analytical knowledge:** is defined as the knowledge of different automated data analysis methods, for example, clustering (e.g., k-means) or data aggregation (e.g., dimensionality reduction). In certain cases, users need at least a basic understanding of their characteristics to choose or parameterize them correctly.

#### HOW is visualization onboarding provided?

Onboarding type, medium, context sensitivity, interaction, tool-specific, and educational theory are relevant aspects of the question of how visualization onboarding is provided. The onboarding type captures the used medium. The form of contextual aid is extremely important for applications (Grossman et al., 2009). The help system should be designed to guide users by demonstration in the context of their interface. Chilana et al. (2012) developed an approach to provide a new framework for integrating crowd-sourced contextual help into web applications. They also discussed the importance of contextual help and adaptive help systems in their work. Based on these results, we also integrate the aspect of context sensitivity into our framework for visualization onboarding. Fernquist et al. (2011a) introduced a set of the most relevant aspects for interactive tutorials for sketching software. Based on their design space for sketching software, we adopted the aspect of interactivity. Additionally, we integrated the category tool-specific category, indicating if the onboarding concept is connected to a visualization tool. Visualization onboarding supports users in learning new concepts. Hence, we integrated the category of educational theory.

**Onboarding type & medium:** Onboarding can be provided in different **types**, such as guided tours, step-by-step wizards, video-based tutorials, and help centers. We derived this terminology from our literature review and Pronovic's blog article about context-sensitive and embedded help formats (Contextual Help, 2019). A particular type of onboarding consists of a **medium** which can be, e.g., textual instructions, video, illustrations/figures, animations, and more.

**Context Sensitivity:** Context-sensitive help assists at a specific point in the current state of the tool. It is the smallest possible chunk of information the user needs to understand. Examples include application help centers, guided tours, mouseover popups, and instructional material. A type of context-sensitive help is **embedded help** which goes beyond basic information and explanations by detecting the user's need for help or offering a guided tour right on the interface. Examples are tooltips, instructions on the interface, or walkthroughs. **Context-free help** can be called at any state of usage and does not relate to the current state of help-seeking. Examples are online documentation and help videos.

**Interaction:** Interaction is applied within the onboarding process itself. We refer to Fernquist et al. (2011a) for defining the degree of interactivity in onboarding concepts. If users can try out the concepts, the onboarding concept is defined as **active**. Help systems can be **passive** if the user only consumes the learning material, such as reading an article or viewing a video. Active tutorials that are aware of the users' interactions and can respond to these are referred to as **reactive**.

Educational theory: Learning and educational theories are crucial when it comes to

onboarding approaches. Systematic categorization of educational theories was impossible as no taxonomy was available. Therefore, we collect educational theories, which authors described in their scientific publications (e.g., concreteness fading (Alper et al., 2017), learning by analog (Ruchikachorn and Mueller, 2015), and more.)

**Tool-specific:** The category describes if the onboarding concept is designed for a specific visualization tool (tool-specific) or if it is decoupled from it (non-tool-specific).

# 2.3.3 WHERE is visualization onboarding provided?

Based on Fernquist et al. (2011a) who introduced a set of the most relevant aspects, we also adopted the aspect of the integration of onboarding concepts by asking *Where is visualization onboarding provided?* – externally, internally. An onboarding system that is integrated **internally** into the visualization can be more helpful for users because they do not have to jump back and forth between two different systems. **External** sources for onboarding concepts can be defined as sources that can be reached independently of the current state of the tool. At the tightest level of integration, help systems can be provided **internally**. It should be pointed out, however, that integrating onboarding systems into the visualization or visual analytics tools is challenging and requires considerable effort.

#### 2.3.4 WHEN is visualization onboarding used?

The aspect of *WHEN* describes the temporal aspect of intended onboarding (see Figure 2.4 (6)). Onboarding concepts can be integrated **before** using the basic visualization tool (one time or repeated) or called up **while** the use of a specific tool, e.g., when support regarding a particular feature is needed.

In the next section, we describe the systematic literature review to identify state-of-the-art visualization onboarding and categorize the work by summarizing existing onboarding concepts in scientific publications and commercial visualization tools.

# 2.4 Survey on Visualization Onboarding

This section describes the method used for our systematic literature review in detail. Furthermore, we present the survey results based on our descriptive design space.

# 2.4.1 Method

To get a comprehensive overview of existing onboarding concepts, we systematically surveyed the literature published in the main venues in information visualization, visual analytics, and HCI. In addition to scientific publications, we reviewed commercial visual analytic tools based on a recent study about commercial systems by (Behrisch et al., 2019) (see Table 7.1). We focused on the following major conferences and journals: *IEEE InfoVis, IEEE VAST, EuroVis, Eurographics, EuroVA, IEEE TVCG, Information Visualisation* (IV), *ACM CHI*, and *ACM UIST*. Because the term *onboarding* is rarely used in the visualization community, we used the following keywords: *data visualization literacy, visualization literacy, instructional material, learning.* We scanned the title and abstract for the specific keywords.

We additionally examined papers published as part of various relevant workshops on the topic of *visualization literacy*, especially the *IEEE VIS DECISIVE Workshop*. We took into account both full and short papers. Moreover, we used snowballing (Wohlin, 2014) to identify the authors of the most relevant papers and included further publications by these researchers. We scanned through the related work sections of the relevant papers to find more literature related to our topic. We were able to identify a total of ten papers that focus on onboarding concepts and learning environments for visualization or visualization tools (Alper et al., 2017; Bishop et al., 2020; Firat et al., 2020; Kang et al., 2003; Kwon and Lee, 2016; Marai, 2015; Ola and Sedig, 2017; Peng et al., 2022; Ruchikachorn and Mueller, 2015; Tanahashi et al., 2016; Yalçın, 2016) and ten commercial tools that use a variety of onboarding methods and concepts (Advizor Solutions, 2023; IBM Cognos Analytics, 2019; Microsoft Power BI, 2019; QlikTech QlikView Visualization Tool, 2019; SAP Lumira, 2019; SAS JMP, 2019; SAS Visual Analytics, 2019; Tableau Software, 2023; TIBCO Jaspersoft, 2019; TIBCO Spotfire, 2019).

Two coders categorized every selected publication. After coding the ten papers, we discussed the coding criteria and matched our coding strategy. In the case of conflicting codes, coders discussed the reasons for decisions to resolve inconsistencies.

# 2.4.2 Results

We reviewed ten scientific publications and ten commercial tools, focusing on onboarding concepts summarized in Table 7.1. In the following sections, we discuss and highlight the most relevant factors of onboarding methods we discovered.

#### WHO – Who is the user? Which knowledge gap does the user have?

For both the scientific publications and the commercial tools, we recognized a strong emphasis on visual encoding and interaction knowledge, as well as data knowledge (Advizor Solutions, 2023; Alper et al., 2017; Bishop et al., 2020; Kang et al., 2003; Kwon and Lee, 2016; Marai, 2015; Microsoft Power BI, 2019; Ola and Sedig, 2017; Peng et al., 2022; QlikTech QlikView Visualization Tool, 2019; Ruchikachorn and Mueller, 2015; SAP Lumira, 2019; SAS JMP, 2019; SAS Visual Analytics, 2019; Tanahashi et al., 2016; TIBCO Jaspersoft, 2019; Yalçın, 2016). Interestingly, Kwon and Lee (2016), Ruchikachorn and Mueller (2015), Bishop et al. (2020) and the two visualization tools IBM Cognos Analytics (2019) and TIBCO Spotfire (2019) do not target data knowledge explicitly, which appears to be surprising as basic data knowledge is crucial to understand the visual encoding of a visualization. Only two publications (Alper et al., 2017; Kwon and Lee, 2016) cover analytical knowledge, while six of ten commercial tools provide support in this respect, e.g., classification and regression models (Advizor Solutions, 2023). We identified a lack of domain knowledge in all tools and the majority of scientific publications. Only

Name			Who?				How?				Where?	When?
	Domain Knowledge	Data Knowledge	Visual Encoding & Interaction Knowledge	Analytical Knowledge	Туре	Medium	Tool- specific	Context-sensitivity	Interactivity	Educational Theory	external / internal	before / while
Alper et al. (2017)	•	•	•	•	teaching tool	text, pictograms	-	context-free	active	concreteness fading	external	before, while
Ola and Sedig (2017)	n.a.	n.a.	n.a.	n.a.	video-tutorial	video	√	context-free	passive	n.a.	external	before
Kwon and Lee (2016)	0	0	•	٠	interactive tutorial walkthrough, video, static	screenshots, video, text, visual elements	-	context-free	active, passive, reactive	experimental learning model	external	before, while
Yalçın (2016)	•	•	•	0	topic listing, point & learn, guided tour, notifications, topic answers	text	~	context-sensitive, embedded	reactive	n.a.	internal	while
Tanahashi et al. (2016)	0	•	٠	0	InfoVis guide	text-plus-questions	_	context-free	active	top-down & bottom-up	external	before, while
Ruchikachorn and Mueller (2015)	0	0	٠	0	video tutorial	animated visualization sequences	-	context-free	active, passive	learning-by-analogy	external	before, while
Kang et al. (2003)	0	•	•	0	step-by-step overlays	text	~	context-sensitive, embedded	reactive	n.a.	internal	while
Bishop et al. (2020)	0	0	٠	0	free-form constructive visualization tool	visual elements	_	context-free	reactive	scaffolding via visual feedback, learning from shared experiences	external	before, while
Firat et al. (2020)	0		•	0	instructional software tool	visual elements, text	_	context-free	active	active learning	external	before, while
Peng et al. (2022)	0	•	۰	0	intervention with six modules	video, tests, and tasks	_	context-free	reactive	corrective immediate feedback-based learning (mastery learning)	external	before, while
SAS JMP (2019)	0	•	•	٠	website, video, step-by-step overlays	text, videos, images, visual element	ts 🗸	context-sensitive, -free	active, passive, reactive	n.a.	internal, external	before, while
Advizor Solutions (2023)	0	•	•	•	website, overlay	text, videos, images	√	context-sensitive, -free	passive	n.a.	external	before, while
SAP Lumira (2019)	0	•	٠	•	website, video	text, videos, images	~	context-free	passive	n.a.	external	before, while
IBM Cognos Analytics (2019)	0	0	•	0	website, video, interactive guided tour	text, videos, images, visual element	ts √	context-sensitive, -free	active, passive, reactive	n.a.	internal, external	before, while
TIBCO Jaspersoft (2019)	0	•	•	0	website, overlay	text, videos, images	~	context-sensitive, -free	passive	n.a.	internal, external	before, while
Microsoft Power BI (2019)	0	•	•	o	website, video, examples, ask questions in app	text, videos, images	~	context-sensitive, -free	passive	n.a.	external	before, while
SAS Visual Analytics (2019)	0	•	•	٠	website, video, interactive guided tour, courses, books	text, videos, images, visual element	ts √	context-sensitive, -free	active, passive	n.a.	internal, external	before, while
Tableau Software (2023)	0		•		website, video, courses, books	text, videos, images	√	context-free	passive	n.a.	external	before, while
TIBCO Spotfire (2019)	0	0	•	•	website, video, courses, books	text, videos, images	√	context-free	passive	n.a.	external	before, while
QlikTech QlikView Visualization Tool (2019)	0	•	•	0	website, videos, courses, overlays	text, videos, images	1	context-sensitive, -free	passive	n.a.	external	before, while

 $|\Sigma$ 

DESCRIPTIVE DESIGN SPACE

Table 2.1: Overview of available visualization onboarding approaches (rows), systematically characterized along the aspects of our conceptual framework (columns). The table is divided into academic research and concepts (upper half) and commercial tools (lower half), which use various onboarding concepts. With a primary focus on the questions, we took four of them and mapped them to categorize available approaches. The colors refer to the equivalent questions explained in the subsections of 3. • applicable  $\circ$  not applicable and *n.a.*: not available/unknown

two publications focus on *domain knowledge* in their onboarding concepts (Alper et al., 2017; Yalçın, 2016). The publication of Ola and Sedig (2017) was an exception, as we could not identify any knowledge gaps.







(3) Power BI Onboarding Page: Screenshots and text



(4) Instructional treemap tool interface with traditional tree structure (left) and linked treemap visualization (right).

Figure 2.5: **Onboarding approaches** (1) IBM Cognos Analytics (2019), (2) Alper et al. (2017) onboarding method based on the concreteness fading educational theory, (3) PowerBI external webpage with instructional material (screenshots and text) (Microsoft Power BI, 2019), and, (4) Educational instructional material for treemap visualization (Firat et al., 2020)

32



(1) Commercial Visualization Tool Advizor showing textual instruction to the use of various visualization techniques



(2) Learning-by-Analogy: In-betweens of linear chart and spiral chart

In parallel coordinates, one line represents one row of a table and each axis represents each column. In this page, you'll draw parallel coordinates by clicking points on each axis per car. Once you finish drawing parallel coordinates of three cars, you will proceed to the next page. Now, please click on the points in parallel coordinates corresponding to the highlighted cell in the table.



Figure 2.6: **Onboarding approaches** (1) Advizor Solutions (2023), (2) Learning by Analogy (Ruchikachorn and Mueller, 2015), (3) Interactive tutorial based on Experiential Learning Model (Kwon and Lee, 2016)

#### HOW – How is visualization onboarding provided?

In this dimension, we distinguish between five different aspects: onboarding type and medium, context-sensitivity, interactivity, tool-specific, educational theory (see Section 2.3.2). Regarding the onboarding type and medium, we found some similarities within the collection of publications. However, these have been the most difficult to gather, as the publications vary the most in their onboarding approaches. In the educational setting (Alper et al., 2017; Bishop et al., 2020; Firat et al., 2020; Peng et al., 2022), the teaching tools use *text*, visual elements, as well as pictograms as medium to educate students. In terms of documented onboarding type, Alper et al. (2017) introduced a "tool for teaching bar charts". More recently, Firat et al. (2020) developed an instructional software tool for treemap visualizations, and Bishop et al. (2020) introduced a "freeform constructive visualization tool". (Peng et al., 2022) used a combination of training videos, text, and tasks based on the cognitive levels outlined in the Bloom's taxonomy. Besides, Kang et al. (2003) and Yalçın (2016) only integrated *text* in their onboarding approaches on overlays. Kang et al. (2003) focused their concept on step-by-step overlays. In contrast, Yalçın (2016) used for his approach overlays including a combination of topic listing, point & learn, guided tour, notification, and topic answers. A further similarity is video and/or animation usage to onboard users. For example, Ola and Sedig (2017), (Peng et al., 2022), and Ruchikachorn and Mueller (2015) developed video-tutorials using animated visualization sequences (Ruchikachorn and Mueller, 2015) (see Figure 2.5 (3)) and a video (Ola and Sedig, 2017; Peng et al., 2022) to support users in learning. In addition, we identified other types, such as interactive walkthrough tutorials (Kwon and Lee, 2016) and InfoVis Guides using text-plus-questions (Tanahashi et al., 2016). In general, most of the collected onboarding approaches use a combination of different mediums and onboarding approaches.

All commercial tools could be systematically categorized only in terms of type and media using documentation/explanation websites with screenshots and textual descriptions (medium). Most tools also use videos as a medium to onboard users. SAS JMP (2019), IBM Cognos Analytics (2019) and SAS Visual Analytics (2019) integrate step-bystep tutorials or interactive guided tours and therefore also rely on visual elements (chart parts to interact with, applicable filters, etc.). TIBCO Jaspersoft (2019) and Advizor Solutions (2023) make use of an in-application help overlay using text and videos. Additionally, Microsoft Power BI (2019), SAS Visual Analytics (2019), Tableau Software (2023), TIBCO Spotfire (2019), and QlikTech QlikView Visualization Tool (2019) provide a combination of books and courses. One special method to highlight is the in-application ask questions of Microsoft Power BI (2019), which allows the users to ask a question related to the data set they are currently working on.

Tool-specific: For the scientific publications, we identified three onboarding approaches that can be categorized as tool-specific (Kandlhofer et al., 2019; Ola and Sedig, 2017; Yalçın, 2016). The remaining six are non-tool-specific (Alper et al., 2017; Bishop et al., 2020; Firat et al., 2020; Kwon and Lee, 2016; Ruchikachorn and Mueller, 2015; Tanahashi et al., 2016). We call these onboarding concepts *learning environments*, which are

independent of a specific visualization tool and can be used in general.

*Context sensitivity* refers to the three categories context-free, context-sensitive, and embedded concepts. Eight out of ten papers designed context-free onboarding concepts, while only Yalçın (2016) and Kang et al. (2003) use context-sensitive and embedded onboarding methods. On the other hand, three out of ten commercial tools integrate context-free onboarding concepts. The other commercial tools integrate context-free and context-sensitive methods using documentation websites and in-application overlays or guided tours. One example is Advizor Solutions (2023), which makes use of context-free and context-sensitive onboarding (see Figure 2.5 (2) for the design of the context-sensitive approach).

A more detailed investigation of the *interactivity* of the onboarding concepts described in publications revealed an excellent balance between the three types of interaction. The category *interactivity* is also connected with the used educational theory. Five of the ten onboarding concepts provide *reactive* onboarding (Bishop et al., 2020; Kang et al., 2003; Kwon and Lee, 2016; Peng et al., 2022; Yalçın, 2016). For the commercial tools, we observed a strong trend toward passive interactivity. Only two tools— IBM Cognos Analytics (2019); SAS JMP (2019)—cover all three interactivity types. *IBM Cognos Analytics*, for example, provides videos and a website (passive) as well as an interactive guided tour (reactive) to onboard users.

Regarding the integrated *educational theories*, we could not find any unique use of educational theories among the onboarding approaches presented in publications. Thus, we identified the following aspects: (1) *onboarding approach designed without the integration of educational theories* (Kang et al., 2003; Ola and Sedig, 2017; Yalçın, 2016); and (2) *onboarding approaches grounded in educational theories*: (2a) concreteness fading (Alper et al., 2017), (2b) Experiential learning model (Kwon and Lee, 2016), (2c) top-down & bottom-up (Tanahashi et al., 2016), (2d) learning-by-analogy (Ruchikachorn and Mueller, 2015), (2e) scaffolding via visual feedback, learning from shared experience (Bishop et al., 2020), (2f) active learning (Firat et al., 2020), and (2g) corrective immediate feedback-based learning (mastery learning) (Peng et al., 2022). In the following, we describe two examples of how visualization onboarding has been applied.

Example on Experiential Learning Model (2b): One example of a reactive onboarding is by (Kwon and Lee, 2016), who developed an online learning approach for parallel coordinates following the experiential learning model (see Figure 2.5 (6)). The model defines learning as the process in which knowledge is constructed via concrete experience and reflection on the experience (Kolb, 1984). Therefore, the interactive tutorial page integrates the experiential learning model's four stages (Concrete Experience, Reflective Observation, Abstract Conceptualization, and Active Experimentation)). The authors implemented the model as follows. The people are asked to complete a mission for the first stage, the Concrete Experience. For the Reflection Observation stage, the onboarding approach provides hints to the user interactions. Additionally, "the system shows the conceptual goal of the activity at a successful completion" (Kwon and Lee, 2016) (Abstract Conceptualization). For the fourth level – Active Experimentation – the learning approach suggests repeating the activity to strengthen the learning. They conducted a comparative evaluation with three tutorial types (static, video-based, and interactive tutorial walkthrough). They observed that participants using the interactive and video tutorials outperformed those with static or no tutorials.

Example on Learning-by-Analogy (2d): In addition to onboarding using the experiential learning model (Kwon and Lee, 2016), Ruchikachorn and Mueller (2015) proposed a concept for the teaching of unfamiliar visualizations by using the educational theory of *learning-by-analogy*. This is an example of a combination of *passive* and *active* onboarding systems. Based on animated visualization sequences (passive), the users were taught a more advanced visualization technique based on an easier one with transitions as presented in Figure 2.5 (3). The user could watch the sequences, which can be categorized as an passive interaction. The user could also start and stop animating the morphing (active).

#### WHERE – Where is visualization onboarding provided?

Our survey of existing work and commercial tools showed that the majority of onboarding solutions could either be classified as external or internal or a combination of both sources. Yalçın (2016) and Kang et al. (2003) designed an internal onboarding concept. All other solutions can be categorized as external onboarding approaches.

There is a relatively equal distribution for commercial tools between only *external* ones and those who are *external and internal*. Most commercial tools provide external material, such as documentation sites with text, images, and videos.

#### WHEN – When is visualization onboarding used?

Onboarding concepts can be integrated at different states of use – before or during. Ola and Sedig (2017) relied on a *before* approach, in contrast, Yalçın (2016), and Kang et al. (2003) provide their onboarding *while* the usage. Other onboarding approaches (Alper et al., 2017; Bishop et al., 2020; Firat et al., 2020; Kwon and Lee, 2016; Peng et al., 2022; Ruchikachorn and Mueller, 2015; Tanahashi et al., 2016) can be either used before or *while*. We detected a clear tendency for commercial tools, as all of the onboarding concepts can be used before and during the usage of the particular visualization tool.

#### 2.4.3Summary

Considering the **WHO** question we observed a strong tendency towards visual encoding  $\mathcal{E}$ interaction knowledge (Alper et al., 2017; Bishop et al., 2020; Firat et al., 2020; Kang et al., 2003; Kwon and Lee, 2016; Peng et al., 2022; Ruchikachorn and Mueller, 2015; Tanahashi et al., 2016; Yalçın, 2016). Data knowledge is also prominent in the literature (Alper et al., 2017; Firat et al., 2020; Kang et al., 2003; Peng et al., 2022; Tanahashi et al., 2016; Yalçın, 2016). However, domain knowledge (Alper et al., 2017; Yalçın, 2016) and analytical knowledge (Alper et al., 2017; Kwon and Lee, 2016) are covered only by two out of ten investigated papers. Only Alper et al. (2017) are targeting all knowledge

gaps. Regarding the question of **HOW** is onboarding provided? — we found a variety of different onboarding types.

This ranges from simple texts instructions (Yalçın, 2016) or videos (Ola and Sedig, 2017; Peng et al., 2022; Ruchikachorn and Mueller, 2015) to interactive visual elements (Alper et al., 2017; Bishop et al., 2020; Firat et al., 2020; Kwon and Lee, 2016) or step-by-step guides (Kang et al., 2003). Regarding context sensitivity, most of them use a context-free approach (Alper et al., 2017; Bishop et al., 2020; Firat et al., 2020; Kwon and Lee, 2016; Ola and Sedig, 2017; Peng et al., 2022; Ruchikachorn and Mueller, 2015; Tanahashi et al., 2016), with two exceptions that are context-sensitive and embedded in the visualization tool (Kang et al., 2003; Yalçın, 2016). Those two exceptions are also *internal* looking at the **WHERE** aspect. All others are designed as non-tool-specific onboarding approaches, i.e., not directly integrated into a visualization tool which are then *external*.

In the case of *educational theory*, however, no general statement can be made based on the categorization of the papers since each paper follows a different educational theory. However, we observed similarities regarding the educational theories presented in Section 2.4.4. In general, most of the collected onboarding approaches of the commercial tools are designed to be used *before* and *while* interacting with a particular visualization tool (WHEN).

None of the commercial tools address or attempt to explain the *domain knowledge* of the users. The tools mainly cover only the data knowledge (Advizor Solutions, 2023; Microsoft Power BI, 2019; QlikTech QlikView Visualization Tool, 2019; SAP Lumira, 2019; SAS JMP, 2019; SAS Visual Analytics, 2019; Tableau Software, 2023; TIBCO Jaspersoft. 2019) as well as the visual encoding  $\mathcal{C}$  interaction knowledge (Advizor Solutions, 2023; IBM Cognos Analytics, 2019; Microsoft Power BI, 2019; QlikTech QlikView Visualization Tool, 2019; SAP Lumira, 2019; SAS JMP, 2019; SAS Visual Analytics, 2019; Tableau Software, 2023; TIBCO Jaspersoft, 2019; TIBCO Spotfire, 2019). Generally, the tendency to convey analytical knowledge is much higher with commercial tools (Advizor Solutions, 2023; SAP Lumira, 2019; SAS JMP, 2019; SAS Visual Analytics, 2019; Tableau Software, 2023; TIBCO Spotfire, 2019) than with the scientific papers. Concerning tools, the type of **onboarding** mainly relies on help websites, video tutorials, or courses. There are a few exceptions (IBM Cognos Analytics, 2019; SAS JMP, 2019; SAS Visual Analytics, 2019) that also use visual elements offering more interaction. For *context-sensitivity*, it is about evenly distributed among the tools, but there is no single embedded one. Also, the *interactivity* in the tools is mostly *passive* since the help often is only provided on demand. Exceptions to this are the three approaches (IBM Cognos Analytics, 2019; SAS JMP, 2019; SAS Visual Analytics, 2019) that offer guides or tutorials directly or react to user interaction. Unfortunately, it was impossible to identify an *educational theory* for any of the commercial tools. The commercial tools have a balanced ratio in the question of WHERE. In terms of the WHEN-question, all the onboarding approaches can be used "while" or "before" using the actual visualization tool.

#### 2.4.4 Existing Design Considerations for Visualization Onboarding

In this section, we present existing guidelines derived from the collected papers. We focused on the given medium, the type of onboarding, and the education theory used to onboard users. Kwon and Lee (2016) developed an interactive guide for parallel coordinates plots based on a learning-by-doing approach. They followed the "Experiential learning model", which can be defined as the process in which knowledge is constructed via concrete experience and reflection on the experience (Kolb, 1984). The presented interactive tutorial walkthrough integrates textual descriptions as well as interactive visual elements (see Figure 2.6-(3)), where, for example, the user can click on points in integrated parallel coordinates, whereupon lines are drawn that than connect them. In their paper, Ruchikachorn and Mueller (2015) developed a teaching concept to learn and teach unfamiliar visualizations by linking them to a more familiar one. They followed the learning-by-analogy approach. Their approach overcomes language barriers as it uses visuals. The authors commented that their system could be helpful when the visualization method to be learned is inherently more powerful than its counterpart. The results of the study by Tanahashi et al. (2016) showed that tutorials, where users can directly interact with the visualization, will influence comprehension positively. They suggest using **active learning** type (participating actively in a corresponding dialog) with top-down exercises. In detail, this means asking participants to draw more advanced, less direct inferences from the data. Their study revealed that their approach of text-plus-question introductory tutorials is a valuable and practical way to onboard users to information visualizations. A recent study shows that there was a successful knowledge transfer to another concrete domain when concrete examples were given as opposed to abstract ones (De Bock et al., 2011). Alper et al. (2017) developed a tablet app teaching elementary school pupils how to create and interpret bar charts using the pedagogical method of **concreteness fading**. The tool provides a space with a reference line (x- and y-axis) and a free-form pictograph representing data in scattered illustrations. Children can stack the icons on top of each other and then watch animated transitions morphing the icons into more abstract representations of a bar chart (see Figure 2.5 (2)). Bishop et al. (2020) developed a free-form construction tool for tablets to engage pupils with the creation of visualization and to make the visual mapping of data more explicit. **Scaffolding** was integrated as educational theory. The results of their study highlight the advantage of scaffolding within the creation process of visualizations through visual feedback, configurability, and shared interaction.

When we sum up and generalize the results of the empirical studies of the papers, as well as the results of the analysis of the design space, we propose the following guidelines when it comes to design onboarding methods:

- explain the visual encoding & interaction concepts (Alper et al., 2017; Bishop et al., 2020; Kwon and Lee, 2016);
- 2. use interactive onboarding approaches, where users can interact with the visualization and with the instructional material (Kwon and Lee, 2016; Tanahashi et al.,

**TU Bibliothek**, Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar. WIEN Your knowledge hub The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

38

2016);

- 3. concrete experience and reflection can lead to higher understanding (Kwon and Lee, 2016; Tanahashi et al., 2016);
- 4. use animations or videos (Alper et al., 2017; Kwon and Lee, 2016; Ruchikachorn and Mueller, 2015) to show the data-to-visual mapping.

# 2.5 Discussion & Conclusion

We presented a descriptive design space for visualization onboarding and design considerations based on the existing empirical studies. The design space contains six aspects: WHY is visualization onboarding needed? WHAT is visualization onboarding? WHO is the user? Which knowledge gap does the user have? HOW is visualization onboarding provided? WHERE is visualization onboarding provided? WHEN is visualization onboarding used? We conducted a systematic literature review to develop the presented design space. We mainly focused on educational theories as the aspect of learning is essential when it comes to the design of visualization onboarding (see Table 7.1 and Section 2.4.4). We also reviewed commercial visualization tools listed in Table 7.1.

Ways to effectively support the learning process of users with different knowledge gaps can be considered by using educational theories. However, the literature needs educational theories focusing on onboarding concepts. Nevertheless, existing theories and results of educational research can be used to inform the design of onboarding systems. We identified guidelines based on the existing literature, which we presented in Section 2.4.4.

Onboarding systems can be designed like help systems, which implies a cognitivist approach, or they might use a scaffolding approach (Bishop et al., 2020; Peng et al., 2022), applying features such as prompts, tools to structure information or higherorder questions. The constructivist theory supports the assumption that cognitive apprenticeship or scaffolding supports higher-order reasoning processes and the ability to make inferences and draw conclusions from the data. Higher-order reasoning is the last stage in the model suggested by graph comprehension but also the ultimate goal of most visualization systems. Based on the papers, educational theories which support active learning and concrete experience are appropriate for onboarding.



# CHAPTER 3

# Design and Evaluation of Visualization Onboarding Concepts



This chapter presents the results of three studies with MTurk workers and students to understand if and how visualization onboarding affects user performance. We aim to investigate different visualization types (bar chart, horizon graph, change matrix, parallel coordinates plot, sankey diagram) and onboarding methods (a step-by-step guide, scrollytelling tutorial, video tutorial, and in-situ onboarding). We describe the research design in Section 3.3 and the results of the three studies (see Section 3.4, Section 3.5, and Section 3.6. Finally, we discuss the results and provide lessons learned. In Section 3.7.1, we propose design guidelines to facilitate the creation of visualization onboarding methods. The chapter is based on our publication (Stoiber et al., 2022d). All supplemental material can be found in Appendix A and retrieved from https://phaidra.fhstp.ac.at/o: 5188.

# 3.1 Introduction & Motivation

Understanding the visual mapping process is crucial for correctly decoding the visual representation and the underlying data. Limited visualization literacy skills can be a severe handicap as it hinders people from valuable information retrieval, which could be used to learn and solve problems or make informed decisions (Börner et al., 2016; Boy et al., 2014; Galesic and Garcia-Retamero, 2010). Visualization onboarding may alleviate this, empower users to understand the data better, and take full advantage of the visualization. We define visualization onboarding (Stoiber et al., 2019a, 2021) as follows: Visualization onboarding is the process of supporting users in reading, interpreting, and extracting information from visual representations of data. A few onboarding methods exist in the literature using different strategies and educational theories (Bishop et al., 2020; Kwon and Lee, 2016; Peng et al., 2022; Ruchikachorn and Mueller, 2015; Tanahashi et al., 2016). Nonetheless, further research is needed to identify effective designs for onboarding methods and to understand users' behavior while using onboarding methods. Therefore, the main objective is to understand how visualization onboarding affects user performance (see Figure 3.1). Therefore, we conducted two large-scaled studies with 388 and 145 MTurk workers (study 1 described in section 3.4 and study 2 in section 3.5) and one study with 63 students (study 3 presented in section 3.6). The qualitative results of study 2 have already been published in the short paper (Stoiber et al., 2021). We investigated various aspects of visualization onboarding (see Section 3.3), which we describe here in detail:

- 1. The effect of visualization onboarding using an interactive step-by-step guide by assessing the users' performance for four different interactive visualization types with varying complexity—a bar chart  $\mathbf{I}$ , a horizon graph  $\mathbf{I}$ , a change matrix  $\mathbf{I}$ , and a parallel coordinates plot  $\mathbf{I}$ , described in Section 3.4;
- 2. The effect on user performance for four different types of visualization onboarding methods. (i) a step-by-step guide, (ii) scrollytelling, (iii) a video tutorial, and (iv) an in-situ scrollytelling, see Section 3.5;
- 3. Differences in user performance based on three task types according to (Friel et al., 2001): reading the data, reading between the data, and reading beyond the data;
- 4. Differences in the subjective user experience and answer correctness between an in-situ visualization onboarding concept using scrollytelling and a tutorial with videos in the VA tool *Netflower* (Stoiber et al., 2019b) for a Sankey diagram and a bar chart
- 5. We present design actions (Spence, 2014) for the design of visualization onboarding methods, described in Section 3.7.1.

The studies aim to provide a comprehensive overview of onboarding approach characteristics using different types of visualizations and onboarding methods (see Figure 3.1). Therefore, we analyze quantitative data on the users' performance and qualitative data such as textual feedback or subjectively rated attitudes and preferences from participants. The analysis of different visualization onboarding methods showed that tutorials and videos are common ways to support users (Stoiber et al., 2019a). Hence, we started by exploring a traditional method — an interactive step-by-step guide (tutorial) — using textual descriptions and visual markers inspired by the metaphor of legends for four different interactive visualization types with varying complexity. Results of study 1 revealed that onboarding is needed for more complex visualization techniques such as parallel coordinate plots. We, therefore, further examined different onboarding methods. Besides the step-by-step guide, we developed a video tutorial as a traditional approach. Besides, we wanted to investigate scrollytelling in the context of visualization onboarding. As a promising onboarding approach resulting from study 2, we further examined the applicability of a scrollytelling onboarding by embedding it into a VA tool (see study 3).

# 3.2 Related Work

In Section 1 of this doctoral thesis and in the Chapter 2 "Descriptive Design Space", we discuss visualization onboarding-related work in detail.

In summary, there are studies in the educational community providing the first attempts at teaching students to interpret (Alper et al., 2017; Baker et al., 2001; Echeverria et al., 2017; Peng et al., 2022) and construct (Alper et al., 2017; Bishop et al., 2020; Firat et al., 2020; Peng et al., 2022) data visualizations. Also, research on the integration of storytelling as an educational theory reveals important contributions toward visualization onboarding design (Echeverria et al., 2017; Knaflic, 2015; Segel and Heer, 2010).

Overall, some of the given onboarding methods which exist in the literature use various *educational theories* (Bishop et al., 2020; Kwon and Lee, 2016; Peng et al., 2022; Ruchikachorn and Mueller, 2015; Tanahashi et al., 2016) and provide support for *different types of visualizations*, e.g., parallel coordinates plots (Kwon and Lee, 2016; Peng et al., 2022; Ruchikachorn and Mueller, 2015), hyperbox (Ruchikachorn and Mueller, 2015), spiral charts (Ruchikachorn and Mueller, 2015), tree maps (Ruchikachorn and Mueller, 2015; Tanahashi et al., 2016), network graphs (Tanahashi et al., 2016), and storylines (Tanahashi et al., 2016). In summary, several different possibilities for designing and realizing onboarding have been suggested in the literature. The studies report on the benefits for the users while using learning environments. Several aspects remain of how onboarding affects users' performance and how to design onboarding efficiently, of which relatively little is known.

# 3.3 Research Design

We conducted three usability studies to understand how visualization onboarding affects user performance at different levels. As a first step, we conducted a small-scale preliminary study to understand how people step-wise explain the *WHAT*, *WHY*, and *HOW* of visualization techniques. This study aimed to use the findings as a design basis for the first interactive step-by-step onboarding method.

To elaborate on the effect of onboarding methods and designs as outlined in Figure 3.1, we present five research questions  $(\mathbf{RQ})$ :

- 1. How does visualization onboarding affect the performance of users?
- 2. Is there a significant difference in user performance with onboarding concepts between different visualization types?
- 3. How does the type of visualization onboarding (step-by-step, scrollytelling, or video tutorial) concept affect the user performance?

	Study 1: Vis Onboarding & the Effect on User Performance	Study 2: Types of Vis Onboarding Methods & the Effect on User Performance	Study 3: In-situ Vis Onboarding in the VA Tool Netflower & the Effect on the User Experience		
Research questions	RQ1: How does visualization onboarding affect the performance of users?	RQ3: How does the type of visualization onboarding (step-by-step, scrollytelling, or video tutorial) concept affect the user performance?	RQ4: Is there any difference in the participant's performance between in-situ scrollytelling and in-situ video tutorial onboarding?		
	RQ2: Is there are a significant difference in user performance with onboarding concepts between different visualization types?		RQ5: Is there any difference in qualitative terms between the effects of in-situ scrollytelling and tutorial in the subjective attitude of the participants?		
	RQ-A: Is there any difference in user performance between different onboarding methods on the level of task difficulty?	RQ-A: Is there any difference in user performance between different onboarding methods on the level of task difficulty?	RQ-A: Is there any difference in user performance between different onboarding methods on the level of task difficulty?		
Participants	388 MTurk workers	145 MTurk workers	63 students		
Performance metrics	answer correctness & response time, behaviour analysis of interaction with the onboarding concept	answer correctness & response time, qualitative feedback	answer corretness, qualitative feedback		
Vis Type	bar chart horizon graph change matrix parallel coordinates plot	parallel coordinates plot	sankey diagram bar chart		
Onboarding Method					
	interactive step-by-step guide	interactive scrollytelling video tutorial	iexi + screensnots		

Figure 3.1: **Overview of the three studies** that were performed by highlighting the respective research questions (RQ), the origin of the participants, as well as the performance metrics, visualization types used, and the onboarding technique applied. We started with exploring a traditional method — an interactive step-by-step guide (tutorial) — using textual descriptions and visual markers inspired by the metaphor of legends for four different interactive visualization types with varying complexity. Results of study 1 revealed that onboarding is needed for more complex visualization techniques such as the parallel coordinates plots. We, therefore, further examined different onboarding methods. Besides the step-by-step guide, we developed a video tutorial as a traditional approach. Besides, we wanted to investigate scrollytelling in the context of visualization onboarding. As a promising onboarding approach resulting from study 2, we further examined the applicability of a scrollytelling onboarding by embedding it into a VA tool (see study 3).

- 4. Is there any difference in the participant's performance between in-situ scrollytelling and in-situ video tutorial onboarding?
- 5. Is there any difference in qualitative terms between the effects of in-situ scrollytelling and tutorial in the subjective attitude of the participants?

Moreover, we investigated the following research question for all three studies:

1. Is there any difference in user performance between different onboarding methods on the level of task difficulty?

In the following, we describe our preliminary study and elaborate on the different onboarding designs and the general study setup as illustrated in Figure 3.1.

**Preliminary study:** Previous studies (Huron et al., 2014a,b) examined how people create, update, and explain their visualizations using only tangible building blocks. Their main goal was to investigate how people construct their visualizations using physical tokens — wooden tiles taken from a learning toys kit designed by Manning (2005) for Kindergarten education. There are no studies available exploring how persons explain a particular visualization type. To provide the first step towards a deeper understanding, we conducted pilot interviews with 13 participants (m = 8, f = 5; Age: M = 30.23, SD = 3.32). The participants indicated they had moderate to high experience with visualizations and a background in computer science, accounting, and HCI. For the interviews, we developed three static visualizations— bar chart, horizon graph, and a change matrix. The bar chart shows sun hours in Innsbruck, Austria, from 1965 to 2019. For the horizon graph, we used random quantitative values from 0 to 30. Additionally, the change matrix visualized quantitative values ranging from -10 to 10. We asked the participants to explain the three visualizations. We took hand-written notes during the interviews. The systematic analysis showed that participants explained the data set and attributes before continuing with the visual encoding. For instance, for a bar chart: "The first thing I would say is that the diagram explains how many hours of sunshine the city of Innsbruck had in previous years." (translated from German to English). We have grouped the comments and explanations, and therefore we gained the following splitting of onboarding instructions: namely *reading*, using, and interacting with the chart, as described in Section 3.3.1.

User studies 1–3: Study 1 was conducted with MTurk workers using LimeSurvey between 10/2019-11/2019 and 02/2020 and aimed to understand (1) if and how visualization onboarding affects user performance with an interactive step-by-step guide; (2) if and how the user performance between four different visualization types (bar chart, horizon graph, change matrix, and parallel coordinates plot) varies; and (3) if and what differences in the user performance can be observed for different task types (Friel et al., 2001).

We conducted the second study (study 2) from 08/2020 to 09/2020. The objective was to investigate how different visualization onboarding methods (interactive step-by-step guide, scrollytelling, or video tutorial) affect user performance. Therefore, we designed two other onboarding methods: a scrollytelling (in-situ onboarding) and a video tutorial. In line with study 1, we examined if a difference in user performance based on different task types exists.

For both studies, we used a between-subject design, where each participant evaluated one of four visualization types/onboarding methods, either with or without onboarding (independent variables). The experience level for each visualization type was captured at the beginning of the survey. We posed questions directly before and after the interaction with the onboarding part to identify the effect of onboarding. The questions include tasks with different difficulties (Friel et al., 2001). We used different data sets for the questions before and after the onboarding part. We refer to the question set before the onboarding part as *pre-tasks* and to the questions after the onboarding as *post-tasks*. To compare the performance measures for study 1, we coded the answer correctness with "0" when the question was incorrectly answered and with "1" when it was correctly given.

Similarly, concerning the improvement measures, we coded answers that were correctly answered in the pre-question set and incorrectly answered in the post-question set with "0", answers that were initially incorrectly answered and then in the post-question set correctly answered with "1", and questions that were equally correct or equally incorrect answered with "-99" as no improvement or deterioration was achieved. Similar to response time improvements, we evaluated the time difference between pre-and post-tasks. Thus, the more negative the value is, the faster the post-task was answered compared to the pre-task counterpart.

Lastly, from 06/2020 to 07/2020 and 10/2020 to 11/2020, we conducted the third study (study 3) to understand how in-situ visualization onboarding in the VA tool *Netflower* (Sankey diagrams and bar charts (Stoiber et al., 2019b)) affects both the user performance and the user experience. For this study, we aim to enable comparative evaluation with students between the provided onboarding in the Netflower tool and the in-situ method.

#### 3.3.1 Onboarding Methods

This section describes the four visualization onboarding methods and their design decisions: a step-by-step guide, scrollytelling tutorial, video tutorial, and in-situ onboarding using Netflower.

**Step-by-step Guide:** The design of our onboarding concept for the *step-by-step guide* is inspired by interactive legends of maps (e.g., thematic maps and cartographic communication) and statistical charts. Printed maps traditionally use static legends to encode the meaning of symbols and colors used (Göbel et al., 2018). In general, legends are an essential design element for cases where data can no longer be labeled directly in



Figure 3.2: • Step-by-step guide for a Parallel Coordinates Plot. The step-bystep guide is based on textual descriptions and in-place annotations anchors and consists of four parts: a brief textual introduction providing contextual information about the visualization  $\square$ , navigational elements  $\square$  to go through the step-by-step instructions  $\square$ , and the visualization itself  $\square$ . We divide the textual descriptions into *Reading the Chart*, for explaining the visual encoding, *Interacting with the Chart*, for explaining the interaction concept, and *Using the Chart*, for providing insights. • Scrollytelling tutorial. Users can incrementally scroll through the instructions  $\square$  while the visualization  $\square$  changes to the text appropriately.

visualizations (Few, 2012). We used the metaphor of legends and developed an interactive visualization onboarding concept consisting of step-by-step textual descriptions for visual encoding and interaction techniques (e.g., hovering, selecting, etc.). We combined them with in-place annotations anchors in the main visualization. As described in Section 3.3, we divided the onboarding instructions into three parts: *reading, interacting, and using* the chart, illustrated in Figure 3.2 (left). In the following paragraphs, we elaborate on the different parts.

**Introduction A**: Due to the importance of title elements and legends (Borkin et al., 2016), we integrated an introductory sentence at the top of our onboarding method to provide general information about the data and describe the area of application for the respective visualization type. For the parallel coordinates plot, for instance, the purpose of the visualization is to identify relationships between attributes.

**Navigation E**: The selected step within **E** the step-by-step instructions is highlighted in the respective sections (*reading, interacting, using*) as well as in the visualization itself. In contrast, the other steps are greyed out (see Figure 3.2 (left)). Users can navigate through the descriptions in a step-wise manner either by using the *Next* and *Previous* buttons on the top right (Figure 3.2 **E**) or by direct manipulation through clicking on the numbered textual description within the step-by-step guide. As an alternative for those who do not need a step-by-step explanation, we provide a *Show All* toggle element to hide and show all hints at a glance (see Figure 3.2 (right)).

**Step-by-Step Messages G**: Based on the explanatory sequences of our interview participants, the step-by-step guide incorporates textual descriptions on how to *read*,

interact with, and use the chart. The first part contains information, e.g., about the visualizations' shape, axes, or color coding. The second part *interacting* with the chart emphasizes the applied interaction techniques, e.g., indicating how to re-order axes or filter attribute values. For part Using the Chart, we provide three examples using the low-level typology (identification, comparison, and summarization task) by Munzner (2009).

Visualization with in-place annotation anchors **D**: The design is organized in numbered textual descriptions with highlighted attributes in combination with in-place annotation anchors (Zhi et al., 2019) (numbers and symbols) to indicate the connection between the highlighted text elements and the visual encoding. Inspired by different annotation designs by Lu (2022), we use circular anchors with numbers related to the selected step. To understand the visual encoding, we highlight *data*, *quides*, and *marks*, as used in the Vega grammar (Satyanarayan et al., 2016). Hence, in each step, we accentuate words related to visual properties of the visualization, encoding, data transformation, pattern, or finding. Similarly, to explain patterns (e.g., positive or negative correlations between attributes), we use rectangles overlayed in the visualizations as shown in step 3 in Figure 3.2. In addition, to demonstrate how to interact with the visualization, we introduce small icons, e.g., for brushing  $\square$  and reordering axes \$

Scrollytelling Tutorial: As outlined in Figure 3.2 (right), we adapted the visual representation of the step-by-step guide. We applied it to a scrollytelling tutorial, where users had to scroll from top to bottom to see descriptions and hints within and beside the chart. Scrollytelling (Amabili, 2019) is a powerful narrative format to package and transmit complex information (Riche et al., 2018; Segel and Heer, 2010). Based on the principles of construction by Nolan and Perrett (2016), text elements can be displayed incrementally by scrolling up and down the screen (Riche et al., 2018, pp. 95). Numbers next to the text elements indicate the current and total number of instructions to give users an idea about the length of the scrollytelling.

The content and organization (description number and assignment to one of the three sections) align with the step-by-step guide to ensure comparability. Our scrollytelling prototype (learning environment cf. (Stoiber et al., 2019a)) for study 2 shows **D** the interactive visualization with visual cues (annotations) on the left and **G** ten instruction steps on the right side of the user interface, see Figure 3.2.

**Video Tutorial:** As a third method, we introduce video tutorials and YouTube videos commonly used in onboarding and help systems (Banovic et al., 2012; Grossman and Fitzmaurice, 2010a; Kwon and Lee, 2016; Lafreniere et al., 2013; Ola and Sedig, 2017; Peng et al., 2022; Pongnumkul et al., 2011; Ruchikachorn and Mueller, 2015). Video tutorials rely on passive instructions, as interaction with the visualization is not supported. Similar to the explanation and structure of the step-by-step guide, we show each mark and interaction element and describe each step of the explanation with a voice-over. Textual descriptions with subtitles are provided to enable auditory-impaired people to understand these explanations.

The step-by-step guide, the scrollytelling tutorial, and the video tutorial can be categorized as "learning environments" (Stoiber et al., 2019a), independently used by a VA tool. The integration of visualization onboarding concepts into a VA tool is not explored in depth besides (Chundury et al., 2023; Yalçın, 2016). Therefore, we integrated an in-situ (internal onboarding cf. (Stoiber et al., 2019a)) onboarding into the VA tool *Netflower*, using a scrollytelling approach. Results of study 2 in Section 3.5 indicated scrollytelling as a promising approach. In the next paragraph, we describe the design of the in-situ onboarding.



Figure 3.3: In-situ onboarding for *Netflower*. The main visualization  $\square$  and the onboarding panel  $\square$  on the side. The user can incrementally scroll through the onboarding messages. All additional components are hidden and only revealed in tandem with their explanation. To combat change blindness, the current onboarding message and the explained elements are highlighted in red  $\square$ . Additional help diagrams  $\square$  are overlaid when appropriate, using the visualized data points.

**In-situ Onboarding using Netflower:** Netflower<sup>1</sup> is an interactive web application for visually exploring dynamic networks (Stoiber et al., 2019b). Its landing page already contains an integrated onboarding consisting of annotated screenshots, textual descriptions, and videos, presented in Figure 3.4.

<sup>&</sup>lt;sup>1</sup>https://netflower.fhstp.ac.at, Accessed: 2023-01-23



Figure 3.4: Netflower onboarding consisting of multiple sections including text along annotated screenshots and a video with a verbal explanation of the functionality: loading data; reading the visualization; filtering, sorting, ordering; using tags; and notebook. https://netflower.fhstp.ac.at

For our comparative evaluation, we compared the already existing onboarding (external onboarding cf. (Stoiber et al., 2019a)) (shown in Figure 3.4) with an in-situ developed scrollytelling (internal onboarding cf. (Stoiber et al., 2019a)) (shown in Figure 3.3). When scrolling from top to bottom, text elements are highlighted incrementally at the right side of the interface  $\Box$ , next to the main visualization  $\Box$ , see Figure 3.3. As the side pane is linked with the interactive visualization, each paragraph is highlighted in red respectively. The scrollytelling behavior in the in-situ onboarding is similar to the scrollytelling tutorial in study 2, as the scrolling guides the user through the instructions. The difference is that the in-situ scrollytelling shows all the instructions in the side pane and triggers the user to highlight the respective paragraph. However, the scrollytelling tutorial provides step-by-step instructions without showing all the instructions at a glance.

The text for this in-situ onboarding is the same as in the already existing onboarding

integrated into the landing page of the Netflower tool, as shown in Figure 3.4. The Sankey Diagram is interactive; e.g., clicking on a connection line in the Sankey diagram displays a detailed view showing the flow size between the two nodes. To onboard the user to those features, we simulate the interaction and show the detail view, for example, while scrolling through the onboarding messages.

#### 3.3.2 Implementation

The prototypes used for this contribution are built with web technologies such as JavaScript, HTML, and CSS. We rely on React (2023) as the basic front-end framework to provide a responsive user interface. For the visual design, layout, and explanation steps, we benefit from the styling framework Ant Design (2023), as it offers many user interface elements and guarantees a responsive grid. For rendering the charts, we rely on Vega and the Vega-Lite wrapper for React Vega (2022) and D3.js (D3 Language, 2023) to integrate the in-place annotations anchors. As we faced limitations concerning the supported interaction techniques (e.g., brushing, reordering, and tooltips) with Vega (Satyanarayan et al., 2016) and Vega-Lite (Satyanarayan et al., 2017) while developing the parallel coordinates plot, we modified the library to enable the interactive exploration (Kai, 2021). For the prototype used in study 3, we relied heavily on the code base of Netflower, which is available under the MIT license at https://github.com/VALIDproject/netflower. We forked it and included our in-situ scrollytelling without additional libraries. All prototypes can be accessed here: https://onboarding-methods.netlify.app/.

#### 3.3.3 Task Categories

To determine differences regarding task difficulty, we use the model by Friel et al. (2001) that distinguishes between three stages of graph comprehension: 1) reading the data, 2) reading between the data and 3) reading beyond the data. As outlined by Lee et al. (2017), reading beyond the data can only be assessed using open-ended questions with qualitative answers. We designed tasks for all three studies following the Visualization Literacy Assessment Test (VLAT) (Lee et al., 2017), which defines a set of questions for tasks such as retrieving the value, finding extremum, determining range, or making comparisons for 12 different visualization types. However, not all visualization types used within this contribution are covered by the VLAT. Therefore, we adapted questions for not yet described visualization types (e.g., parallel coordinates plot) using the VLAT questions as a template. All task questions and analysis tables can be retrieved from from the url https://phaidra.fhstp.ac.at/o:5188 and the Appendix A.1.

Examples of the parallel coordinates plot are the following:

Retrieve Value: What is the percentage of cacao in Palette de Bine?

Find Extremum: Which company had the highest sales?

**Determine Range**: What was the average price per kg for cacao beans with a 70% cacao percentage?

*Make Comparison*: Comparing Caribbeans, Kah Kow, and Sibu, which of these companies shows the lowest price per kg and the highest sales volume (in Euro)?

*Correlation*: In general, there is a (negative | positive) correlation between the review date and the sales volume.

**Similarity**: What are the most similar companies in 2015 regarding the rating and price per kg?

Further and inline with (Freedman and Shah, 2002; Kosslyn, 1989; Lee et al., 2016), we use easy-to-understand, concrete, and time-oriented data sets, see Sections 3.4.3 and 3.6.4 for a detailed description of the used data sets in each study.

# 3.4 Study 1: Visualization Onboarding and Effect on User Performance

As outlined in Figure 3.1, we designed study 1 to understand how visualization onboarding affects the users' performance (1) using different visualization types (2) and task difficulty levels (1). Therefore, we analyzed four visualization types (bar chart, horizon graph, change matrix, and parallel coordinates plot), with two conditions (with and without a step-by-step guide as onboarding method) and an interchanged data set for pre- and post-tasks, resulting in 16 conditions. Consequently, we report on differences in the performance measures answer correctness and improvement (in percent), response time and response time improvement (in seconds), the interaction time with the onboarding part (in seconds), and interaction patterns, as well as qualitative feedback on the onboarding approach. Thus, we followed a mixed-method approach to increase the comprehension of visualization onboarding usage.

#### 3.4.1 Participants and Apparatus

We conducted 16 experiments on the crowdsourcing platform Amazon Mechanical Turk. Each worker was compensated with \$3 for surveys with onboarding, as it was more effort to participate in, and \$1.5 for surveys without onboarding. We randomly assigned MTurk workers to the surveys. Furthermore, we required the workers to have at least a Bachelor's Degree and fulfill the MTurk definition of a Master <sup>2</sup>, meaning that they consistently showed good survey results. Because we released the surveys sequentially, we excluded workers if they had already participated earlier. We recruited 400 MTurk workers (25 workers per condition), but we had to exclude 12 due to an incomplete survey. Thus, 388 MTurk workers (Gender: m = 241, f = 149, prefer not to say=2); Age: M = 38.23, SD = 9.60) participated in our survey.

We conducted the experiment using LimeSurvey<sup>3</sup> where we made use of the time recording and question randomization features. We designed and implemented the visualization for

<sup>&</sup>lt;sup>2</sup>https://www.mturk.com/worker/help, Accessed: 2023-01-23

<sup>&</sup>lt;sup>3</sup>http://www.limesurvey.org/ - Accessed: 2023-01-23



Figure 3.5: **Procedure of study 1 and study 2 with MTurk workers**. The surveys followed a uniform structure with two conditions—"with" and "without" onboarding—where the "with" onboarding condition included an interactive onboarding part and a questionnaire to gather feedback about the onboarding concept.

the pre-and post-tasks and onboarding parts as described in Section 3.3.2. We linked the onboarding part to LimeSurvey and integrated the visualization parts (pre- and post-tasks) into the questionnaire using iFrames.

To assess interaction time with the onboarding concept based on videos and click events, we used the behavior analytics tool Hotjar (2023). Participants used their own devices. Since our onboarding concept was designed for desktop and mouse-based devices only, we excluded the usage of mobile devices not only in the textual description and programmatically. Two co-authors inspected the videos independently to identify different aspects to answer the following questions: "How long do users, on average, interact with the onboarding concept? Was there any interaction with the visualizations? Which interaction behaviors can be observed?"

#### 3.4.2 Procedure

Figure 3.5 shows both survey conditions, one with and the other without visualization onboarding. Both surveys were structured equally, except the one with onboarding incorporates the interactive onboarding part and seven questions about the ease of use, confidence, comprehensibility, style, interpretation, interaction, and frequency of the onboarding part itself.

Each survey started with a question about the level of experience for the given visualization type using a 5-point Likert scale.

After assessing the experience level, we continued with the pre-task set: four questions for a bar chart, horizon graph, and change matrix, and six questions for parallel coordinate



Figure 3.6: Level of experience per visualization type for study 1 on a 5-point Likert scale, where 0 means no experience and four highly experienced. The percentage value indicates how many participants subjectively rated their level of experience. The 'AVG' bar also outlines the average experience level per visualization type from 0-100%.

plots. We extended the pre-task set with two questions regarding a positive and negative correlation. The questions were posed in randomized order to avoid a selection bias. As outlined in Section 3.3.3, the questions were inspired by the VLAT (Visualization questions: e.g., retrieve the value, find extremum, determine the range, find anomalies, etc.). Next, participants were obliged to interact with the onboarding part—if assigned to the onboarding group. To ensure that they interacted with the onboarding, we added an obligatory confirmation question before continuing the survey. The set of questions of the post-tasks followed this. Finally, we completed the survey with questions about demographics (e.g., age, gender, profession).

#### 3.4.3 Data Sets

To follow the concept of easy-to-understand and concrete data sets, we used weather, car data, and Olympic medals distributions for study 1. For the weather data, we have chosen the publicly available data set about the daily 20th-century surface air from the European Climate Assessment & Data set project (ECA&D, 2019). We calculated the average temperature of European cities for selected years (1990, 1991, 2018) using this data set for the bar chart and horizon graph onboarding. For the change matrix, we used the Olympic data set (IOC, 2023) showing the medal distributions between 1990 and 1991. The data sets for the parallel coordinates plot required more dimensions. Thus, we decided to take the car data set from PACO (Heinrich and Broeksema, 2023) for our visualization onboarding which we extended by the production year of the cars to add the time aspect. For the pre-and post-tasks, we used easy-to-understand, time-oriented data sets (e.g., Spotify data based on song titles and various characteristics (Kaggle.com, 2023b), chocolate bar rankings (Kaggle.com, 2023a), amount of steps over time, weather data (ECA&D, 2019), and Olympic medals (IOC, 2023) distribution over time).

Performance Measure Onboarding	Answer Cor With	rectness (in %) Without	<b>Response</b> With	<b>time (in s)</b> Without			
Bar Chart	<b>0.88</b> ±.33		30.25 s	±26.44			
	0.87±.33	$0.88 \pm .32$	$33.92 \text{ s} \pm 28.02$	$26.58 \text{ s} \pm 24.31$			
Horizon	<b>0.89</b> ±.32		$38.26 \text{ s} \pm 28.21$				
	<b>0.90</b> ±.302	$0.87 \pm .34$	$38.86 \text{ s} \pm 27.80$	$37.65 \text{ s} \pm 28.69$			
Change Matrix	<b>0.76</b> ±.43		rix 0.76±.43		<b>43.73 s</b> ±72.61		
	0.81±.39	<b>0.70</b> ±.46	$52.78 \text{ s} \pm 95.56$	<b>34.50</b> s±34.49			
Parallel Coordinates	<b>0.55</b> ±.50		57.21 s	±78.00			
Plot	$0.52 \pm .50$	$0.58 \pm .49$	$65.91 \text{ s} \pm 41.70$	$50.13 \text{ s} \pm 99.60$			

Table 3.1: Averages (and SD) of performance measures for all four visualization types of study 1, n=388.

#### 3.4.4 Results

**Visualization Type:** We started by assessing the differences between all four visualization types. We identified significant differences in the answer correctness between the four visualization types. The highest answer correctness for post-questions was given for the horizon graph (M = .89; SD = .32), followed by the bar chart (M = .88; SD = .33), the change matrix (M = .76; SD = .43), and the parallel coordinates plot (M = .55; SD = .50), F(3, 1742) = 70.557, p = .000. Concerning the response time, the fastest answers were given for the bar chart ( $M = .30.25 \ s$ ; SD = .26.44), then the horizon graph ( $M = .38.26 \ s$ ; SD = .28.21), the change matrix ( $M = .43.73 \ s$ ; SD = .72.61), and then the parallel coordinates plot ( $M = .57.21 \ s$ ; SD = .70.00) survey, F(3, 1743) = 18.523, p = .000.

The results based on answer correctness or response time improvement reveal that statistically significant differences between the visualization types exist for the answer correctness improvement,  $\tilde{\chi}^2 = 6.643$ , p = .084, df = 3. In contrast, for the response time improvement, participants answered post-questions faster with the horizon graph  $(M = -20.58 \ s; SD = 234.42)$ , then with the parallel coordinates plot  $(M = -19.32 \ s; SD = 95.01)$ , followed by the bar chart  $(M = -2.315 \ s; SD = 42.15)$  and the change matrix  $(M = 1.63 \ s; SD = 100.09)$  — irrespective of the onboarding concept, F(3, 1742) = 3.190, p = .23,  $\eta^2 = .005$ . The more negative the value is, the faster participants answered the post-question set. Thus, questions about horizon graphs and bar charts could be answered more correctly than the other visualization types but the average response time, for instance, for the horizon graph, was longer for the post- than for the pre-question set.

**Overall Onboarding:** We continued by analyzing the differences in the measures between a step-by-step onboarding guide and our control group without any onboarding at all. We hypothesized that the answer correctness is higher for surveys with onboarding than without onboarding. Using all surveys from all visualization types, we could not determine statistically significant differences between the answer correctness between surveys with (M = 879.25) or without onboarding (M = 867.85), U = 376064.00, z = -.624, p = .533. In contrast, the participants conducting surveys in the control condition showed statistically significant faster answers  $(M = 38.74 \ s; SD = 63.43)$  than participants in the onboarding condition  $(M = 49.73 \ s; SD = 56.41)$ , t(1742) = 3.821, p = .000,  $eta^2 = 0.008$ . The improvement measures did not show significant differences between both conditions. Thus, we could not identify differences or improvements across all visualization types.

**Onboarding per Visualization Type:** We, therefore, decided to investigate if the utilization of onboarding influences the measures for each visualization type independently. We used a one-way multivariate analysis of variance (MANOVA) with answer correctness and response time as dependent variables and the onboarding concept as well as the visualization type as fixed factors. There was a statistically significant difference in user performance based on the onboarding concept for each visualization type,  $F(6, 3470) = 2.607, p = .016; Wilk's\Lambda = .991, \eta_p^2 = .004$ . However, only for the answer correctness ( $F(3, 1736) = 3.310; p = .019; eta_p^2 = .006$ ) and not the response time ( $F(3, 1736) = 1.798; p = .145; eta_p^2 = .003$ ). Regarding the performance improvement, no significant differences could be observed, neither in the answer correctness improvement nor the response time improvement. To investigate the effect further on the visualization type level, a significant difference in the performance could be observed for the bar chart, F(2, 389) = 3.835, p = .022;  $Wilk's\Lambda = .981$ . However, only the response time (F(1, 392) = .7.662; p = .006) between the two conditions but not for the answer correctness (F(1, 392) = .094; p = .759). This means that surveys without any onboarding were answered faster  $26.58 \ s \pm 24.31$  than surveys with an onboarding concept (33.92  $s \pm 28.02$ ). For the horizon graph, neither in the answer correctness nor in the response time, any performance differences could be determined between surveys with and without onboarding, F(2,373) = .485, p = .616;  $Wilk's\Lambda = .997$ . In contrast, for the change matrix, both performance measures show statistically significant differences, F(2,393) = 6.698, p = .001; Wilk'sA = .967. Interestingly, the answer correctness was on average higher for surveys with an onboarding  $(0.81 \pm .39)$ , whereas the condition without any onboarding shows lower answer correctness  $(0.70 \pm .46)$ , F(1, 396) = 6.675; p = .010. Additionally, concerning the response time, surveys with onboarding (52.78  $s \pm 95.56$ ) were answered on average slower than surveys without onboarding  $(34.50 \ s \pm 34.49)$ , F(1, 396) = 6.357; p = .012. In summary, participants using the change matrix invested more time with the visualization and achieved higher answer correctness. Lastly, also for the parallel coordinates plot, a statistical difference in the performance between surveys for the two conditions could be assessed,  $F(2, 577) = .3.506, p = .031; Wilk's\Lambda = .988.$ However, only the response time shows differences, F(1, 580) = 6.048; p = .014, where surveys without onboarding (50.13  $s \pm 99.57$ ) showed on average faster response times than surveys with onboarding (65.91  $s \pm 41.70$ ). The answer correctness without onboarding showed surprisingly more correct answers  $(0.58 \pm .49)$  than for surveys with the onboarding  $(0.52 \pm .50)$ , but with a large variance.

56
**Onboarding and Task Difficulty:** Subsequently, we investigated the influence of the question type difficulty on the users' performance outcome. Therefore, we assessed differences between *reading the data*, *reading between the data*, and *reading beyond the data*. Irrespective of the onboarding method and the visualization type, *reading the data* questions showed the highest mean rank (*meanrank* = 955.13), followed by *reading between the data* (*meanrank* = 813.33) and *reading beyond the data* question types (*meanrank* = 803.58),  $\tilde{\chi}^2 = 64.044$ , p = .000, df = 2. Similarly, the response time is in accordance with the answer correctness, where the fastest answers were given for the *reading the data* ( $M = 28.0305 \ s; SD = 34.70184$ ), then for the *reading between the data* questions ( $M = 55.25 \ s; SD = 48.46$ ), and lastly for the *reading beyond the data* questions ( $M = 58.82 \ s; SD = 88.22$ ), F(2, 1741) = 53.651, p = .000,  $\eta^2 = .058$ . An additional Tukey HSD post-hoc test reveals a statistically significant difference between *reading the data* and both *reading between the data* and *reading beyond the data* (p = .610) over all visualization types exists.

Analyzing the performance measures using a three-way ANOVA between onboarding concept, visualization type, and task difficulty level shows statistically significant differences in the response time but not for the answer correctness, F(6, 1722) = 1.485, p = .180,  $\eta_p^2 =$ .005. Interestingly, concerning the response time improvement, differences could be determined, F(6, 1722) = 6.016, p = .000,  $\eta_p^2 = .021$ . Based on the response time improvements, we would argue for a learning effect for all visualization types for the participants that did not receive an onboarding, as all question types (Friel et al., 2001) were answered faster  $(M = -15.07 \ s; SD = 188.30)$  for the post-questions rather than the pre-question set. In contrast, participants who were in the onboarding condition showed diverging results. In particular, the reading between the data question  $(M = 57.12 \ s; SD = 118.31)$ for change matrices with onboarding took statistically significantly longer to answer than all the other question types (M = 22.33 s; SD = 488.90). Moreover, the reading beyond the data questions  $(M = -55.06 \ s; SD = 137.53)$  for the parallel coordinate plots were answered faster by participants that received an onboarding  $(M = -23.53 \ s; SD = 105.65)$ . This goes along with the feedback as "Once I read the step-by-step guide, using the graph felt way easier" (St55). Similarly, the reading between the data task for horizon graphs was answered faster by participants that had an onboarding to the visualization type.

#### 3.4.5 User Behavior

We used the recordings of mouse and click events of the participants' sessions to investigate the interaction with our onboarding concepts. Due to blocking software in browsers or other available options which protect users from tracking or recording their actions, it was only possible to collect interaction data from 91 out of 145 participants across all studies who were supposed to do the onboarding. Find all heatmaps for the visualization types in Appendix A. **Interaction Time:** Based on the recorded videos, we analyzed the *interaction time* by checking when the user stopped interacting with the onboarding prototype. Interaction time is defined as *no mouse movement any longer*. We found that the MTurk workers interacted with the onboarding concept independently of the visualization type ( $M = 102.4 \ s; SD = 213.71$ ). Interaction time with bar chart ( $M = 65.96 \ s; SD = 75.40$ ) and horizon graph ( $M = 60.19 \ s; SD = 65.59$ ) was moderately low. In contrast, interaction time for change matrix ( $M = 204.05 \ s; SD = 446.41$ ) and parallel coordinates plot ( $M = 99.86 \ s; SD = 96.33$ ) was higher.

**Interaction Behaviors:** Extracting information from the mouse tracking data helped us to determine hot spots in the onboarding concept. We evaluated these hot spots as attentive areas with our four areas of interest: A introduction, A navigation, Step-by-step guide, and visualization (Figure 3.2).

As the area for *introduction* does not contain interactive elements, hardly any mouse movement and no click events could be observed. When analyzing patterns in the visualization itself, we could observe that the hover feature was frequently used for all visualization types. In particular when the step-by-step guide referred to attributes or exact measures from the data set. Interestingly, we found that besides the mouse movements, more click events were ascertained with the parallel coordinates plot. This can be traced back to the introduced interaction possibilities, e.g., brushing and re-ordering axes (see Figure 3.8). Sessions from the bar chart, horizon graph, and change matrix show that most participants interacted with the navigation pane and the step-by-step guide in a balanced manner (see Figure 3.7). In contrast, we could observe that users predominantly clicked on the text within the step-by-step guide for the parallel coordinates plot. Hardly any superimposed mouse movements were detected with our navigation pane. Moreover, when assessing the usage of the Show All button, we recognized that it was only used by 10 out of 91 participants (5.08%) for all visualization types. Caused of the active usage of the step-by-step instructional descriptions, we analyzed the explanation sections in further detail. We noticed less interaction or interest in the Using the Chart section than with *Reading the Chart* and *Interacting with the Chart* indicated by fewer mouse movements or click events and by a lower retention time. This applies to all visualization types. Particularly noteworthy was this behavior for the parallel coordinates plot, which had an additional section for interaction and, thus, more steps. We noticed that 42participants out of 56 total records for parallel coordinates plot went much faster over the Using the Chart section or completely ignored it. It could be noted across all studies that this section seemed to be less interesting to the participants.

Overall, it can be said that familiarity with a visualization type has an impact on the answer's correctness. Thus, for the rather unfamiliar visualization types, the change matrix, and the parallel coordinates plot, it can be seen that onboarding is needed to improve performance. In contrast, for highly familiar visualization types, onboarding is not required and does not affect the performance measures. Therefore, we continued to investigate different onboarding concepts in study 2 using the parallel coordinates plot.



Figure 3.7: Mouse movement heatmap on the step-by-step guide for a bar chart, horizon graph, and matrix received from Hotjar (2023). In Appendix A we attached the Hotjar heatmaps separately.



Figure 3.8: Mouse movement heatmap on the step-by-step guide for parallel coordinates plot received from Hotjar (2023).

60

# 3.5 Study 2: Types of Visualization Onboarding Methods and the Effect on User Performance

As we identified a need for onboarding for the more unknown visualization type—the parallel coordinates plot—we aimed to observe differences between the means of how onboarding can be provided to a user. We therefore implemented and assessed three methods: (i) an interactive step-by-step guide, (ii) a scrollytelling tutorial, and (iii) a video tutorial with voice-over. Tutorials and videos are common ways to support users. Hence, we wanted to explore scrollytelling as a new approach to the more traditional concepts of a tutorial and videos. We build upon the previous study and compare the measures for the parallel coordinates plot. Thus, we aim to understand how the different onboarding methods affect the users' performance (3) and if a difference in user performance can be assessed for different task types (1). Further, we applied a sentiment analysis on the qualitative feedback to understand differences in the perception of the three onboarding methods. We, therefore, made use of the sentiment analyzer by MonkeyLearn (MacCaw et al., 2022) — a machine learning platform that aims to retrieve and classify text information — while also evaluating each qualitative feedback ourselves.

We used the same study design as for study 1. The results of the sentiment analysis are already published by Stoiber et al. (2021).

#### 3.5.1 Participants and Apparatus

We conducted a between-subject design and employed additionally 100 MTurk workers, where 25 were assigned to each condition for the yet unknown onboarding methods (scrollytelling and video). In total, we had to exclude 5 participants, hence answers from 145 participants (m = 94, f = 50, prefer not to say=1; Age: M = 35.46, SD = 9.29) were additionally taken into account. As Hotjar builds heatmaps on top of screenshots, we could not assess the scrolling behavior and thus did not assess mouse movements for the method comparison.

#### 3.5.2 Results

**Overall Onboarding:** First, we started to determine differences in the performance measures (answer correctness and response time) for all four conditions (the baseline without onboarding, step-by-step guide, scrollytelling tutorial, and video guide). No statistically significant difference in the answer correctness for the post-question set could be determined. However, the response time showed that faster answers were given after a scrollytelling tutorial ( $M = 41.17 \ s; SD = 46.55$ ), followed by the step-by-step guide ( $M = 49.97 \ s; SD = 59.77$ ), the condition without any onboarding ( $M = 58.07 \ s; SD = 104.90$ ), and the video tutorial ( $M = 58.39 \ s; SD = 71.47$ ), F(3, 1134) = 3.3785; p = .010; Eta = .096.

**Onboarding and Task Difficulty:** For the parallel coordinates plots, differences in the task complexity level could be observed for both performance measures (answer correctness: F(2, 1134) = 37.081; p = .000; response time: F(2, 1134) = 36.66; p = .010). A Kruskal-Wallis test showed that the highest mean rank for the dichotomous variable answer correctness was achieved by the easiest task (669.00), followed by the intermediate (550.50) and advanced task (501.00), H(2) = 69.33, p = .000. As expected, an ANOVA with post-hoc SNK test showed that *reading the data* tasks were answered the fastest ( $M = 30.918 \ s; SD = 64.53$ ), followed by *reading between the data*  $M = 49.76 \ s; SD = 56.01$ ) and *reading beyond the data*  $M = 74.79 \ s; SD = 104.99$ ). However, similar results for each onboarding technique could be observed, see Table 3.2.

	Answer	Correctn	less $(\%)$	Response Time (s)			
Task Difficulty*	1	2	3	1	2	3	
Baseline		$.52 \pm .50$			58.07 s±104.90		
No Onboarding	$.71 \pm .46$	$.46 \pm .50$	<b>.39</b> ±.49	<b>33.56 s</b> ±30.53	$54.19 \ s{\pm}58.87$	86.45 $s \pm 165.59$	
Step-by-Step		$.52 \pm .50$			$49.97 \text{ s}{\pm}59.78$		
Guide	<b>.74</b> ±.44	$.47 \pm .50$	$.35 \pm .48$	$27.11 \text{ s} \pm 34.26$	$48.20 \ s \pm 52.78$	<b>74.61</b> s±75.45	
Scrollytelling		<b>.42</b> ±.50			$41.17 \text{ s} \pm 46.55$		
Tutorial	$.65 \pm .48$	$.47 \pm .50$	$.42 \pm .50$	$28.57 s \pm 28.76$	$48.20 \ s \pm 52.78$	<b>53.85</b> s±54.99	
Video		<b>.49</b> ±.50			58.39 s±71.47		
Tutorial	.60±.49	$.49 \pm .50$	$.37 \pm .49$	$34.44 \text{ s} \pm 42.62$	55.75 s $\pm 62.45$	<b>84.98</b> s±91.73	

\* 1= Reading the data; 2= Reading between the data; 3 = Reading beyond the data

Table 3.2: **Overview of performance measures for all four onboarding conditions** (baseline with no onboarding, step-by-step guide, scrollytelling, and video tutorial) and three task difficulty levels. *Note*: Not all participants answered these questions as they were not mandatory.

	Positive		N	Negative		Neutral		No answer	
Baseline			-	-	-	-	-	-	
Step-by-Step Guide	13	27.66%	3	6.38%	4	8.51%	27	57.45%	
Scrollytelling Tutorial	16	32.65%	6	12.24%	9	18.37%	18	36.73%	
Video Tutorial	21	42.86%	1	2.04%	6	12.24%	18	42.86%	

Table 3.3: The results of the **text classification** using the sentiment analysis for study 2 are presented. Total indicates the number of submitted positive/negative/neutral feedback in relation to the total number of participants assigned to a method. *Note*: Not all participants answered these questions as they were not mandatory.

**Qualitative Feedback:** Overall, the results show that 34.48% of the responses can be classified as positive, 6.90% as negative, 13.10% as neutral statements, and 45.52% of

the participants did not submit any feedback on the onboarding as it was not mandatory (see Table 3.3). Participants decisively appreciated the condensed, structured, and grouped explanation steps of each of the approaches. On closer examination, the highest positive feedback was given for the video tutorial (42.86%), followed by the scrollytelling tutorial (32.65%), and the step-by-step guide (27.66%). Noteworthy is that participants highlighted learning new features during the video tutorial (V26, V42, V50), e.g., "The video was helpful and showed me some features that I wasn't familiar with [..]" (V50). More specifically, "[..] I liked knowing that I could move columns next to one another, etc." (V60), which relates to the re-arrangement of axes. However, the automatically generated voice-over in the video guide was described as unattractive as it sounded robotic (V3). In contrast, the step-by-step guide did not support the usage and understanding of interactive elements (e.g., filtering and moving axes) or the interpretation of correlations as St34 described: "I have trouble with correlations, but I don't think that is the fault of the quide—although examples would be good" (St34). This may also explain why the results of study 1 were not significantly improved for the onboarding condition as anticipated. Regarding the scrollytelling tutorial, users "[..] enjoy that the images on the left do not appear until I need to see them, which prevents confusion" (Sc38). Hence, the display of information on-demand through scrolling and enriching the visualization with information increased the participants' perceived satisfaction.

Although statistically, no significant differences can be found in the different onboarding methods, the qualitative feedback from the participants supports the statement of needing onboarding predominantly for unfamiliar visualizations. In addition, the video tutorial shows a high resonance with 42.86% positive feedback since the storytelling approach and guidance can be adopted here. As the results in Table 3.2 show that the average answer correctness for the scrollytelling tutorial was relatively low (but with high variance), we decided to investigate the effect of onboarding even more with study 3, where onboarding was embedded in the software. Furthermore, we wanted to investigate new ways of supporting users using the scrollytelling approach.

In the next paragraphs, we present the lessons learned from studies 1& 2.

#### 3.5.3 Lessons Learned

Lessons Learned Study 1: Based on the qualitative feedback of the onboarding concept (study 1), onboarding was described to be easy to use and easy to understand. Moreover, the easy to understand example (monthly changes in weather) supported the visualization literacy for new kind of graphs. Also, the results of the quantitative measures showed that the participants ranked the onboarding concepts as easy to use, easy to comprehend, and easy to follow about the design. Furthermore, participants would be willing to use the onboarding concept when working with new visualization types.

Onboarding elevates the improvement of response accuracy. However, assessing the effects on visualization type level, we could not observe differences in improvements between surveys with and without onboarding. On the contrary, the results indicate that

the highly familiar visualization type — bar chart — does not require onboarding and results in faster answers after multiple question iterations (pre- and post-tasks). Here, we can refer to learning effects based on repetition. Items practiced a lot acquire a high activation in our memory and are retrieved faster (Budiu, 2016).

Feedback on the widely known visualization type—the bar chart—indicates to be "selfexplanatory" or "straightforward". This finding also coincides with a low interaction time and hardly any mouse movements with the onboarding. The horizon graph, for instance, showed, despite an intermediate level of familiarity, both a high improvement in response accuracy and response time. Further, participants of the horizon graph and the change matrix emphasized "the organization of the onboarding" and the "choice of colors" particularly positively. Similarly, their interaction time with the onboarding concept increased. Similar to the improvements, the interaction time with the onboarding concepts increased gradually with the level of familiarity. Thus, the higher the experience with visualization, the lower the interaction time with the onboarding. Interestingly, for the most unfamiliar visualization type—the parallel coordinates plot—the feedback on onboarding diverged. On the one hand, participants appreciated that the *essential* words have been marked so that you could more easily find out what the step was about". which goes along with our consideration of basing visualization literacy descriptions on the Vega-Lite specification and highlighting respective visual properties. On the other hand, some participants were overwhelmed by the information displayed due to a lack of literacy. To reduce this information overload, features should be *"introduced slowly and* not at the same time". This leads to a gradual approximation of new functionality with highly unfamiliar and complex visualization types.

One participant mentioned the following: "Unfortunately, I still don't think I have a handle on how to understand and read these charts. I still have questions. I need to be able to redo the questions I was first asked to see if I can get them right. I need to not only read the step-by-step guide but put what I have learned into practice to make sure I truly understand what I learned." This approach appears to be highly interesting as integrating challenges and small tasks in the onboarding concept can provide a format to check if the text was understood correctly.

Lessons Learned Study 2: Reflecting on the feedback from study 2, we summarize the most important insights. We realized that independently of the visualization type and method applied, an easy-to-understand data set and concrete examples on how to read the chart, support, and increase comprehension are vital. Furthermore, qualitative feedback indicated that the to-the-point descriptions make it easier to absorb information. Based on the comments, the competence to learn new interaction techniques can be increased through video tutorials. Likewise, the understanding of introductions is enabled by interactive and linked descriptions of the visualizations  $\mathbf{E}$  and the described steps  $\mathbf{D}$  (as used for the step-by-step guide and the scrollytelling tutorial). Finally, it is important to consider which visualization types (predominantly unknown and new ones) require an introduction to support the user when needed.

## 3.6 Study 3: In-situ Visualization Onboarding in Netflower

Our last study aimed to compare an external and an in-situ onboarding method (see Figure 3.1). We added in-situ onboarding to an existing VA tool to see if differences in performance (4) and attitude (5) of users exist compared to an external onboarding. We use the term *attitude* as it is used in psychology to describe a user's way of thinking about something and their corresponding behaviors, beliefs, and emotions towards it (Gerrig et al., 2015). Inspired by the results of the previous studies (study 1 & 2), we additionally investigated if we can detect any difference with regards to tasks of different complexity as described by Friel et al. (2001) (Research Question 1). In the following, we present the study design, participants, apparatus, and material, as well as describe the procedure and the results of our comparison.



Figure 3.9: **Procedure for study 3 with students**. The flowchart shows the structure of the experiment. The table shows how we split the participants into groups A-D, which groups were assigned which onboarding system, and how many people participated in parentheses.

#### 3.6.1 Study Design

The study was performed in two rounds, the first in spring 2020 and the second in fall 2020, illustrated in Figure 3.9. In each round, the students were split into two groups. Each group completed the experiment in two sessions, with three weeks in between each session, and using a different data set each session. In the first round, one group was assigned the external onboarding, and one group was assigned the in-situ onboarding. They used the same onboarding for both sessions to establish the existence or absence of a learning effect. In the second round, one group used in-situ onboarding for the first session and external onboarding for the session. In contrast, the second group did the opposite to enable within-subject comparison. The experiment was the same for each

session: The participants were given a survey, which included an introduction, different tasks, and a questionnaire to be answered after completing those tasks.

#### 3.6.2 Participants and Apparatus

We recruited 65 students (Gender: m = 42, f = 23, prefer not to say=0) from the first year of the international Master's programs "Data Science" and "Media Informatics" at the TU Wien, where the study was conducted in the context of a lecture. The participants were required to have at least basic knowledge of data visualizations. For round one, 24 participants completed both sessions, with an additional two only completing the first one. In round two, 39 participants completed both sessions, but 2 participants collaborated in the second session, so the results of only one were taken for further analysis. The exact distribution of participants in each session and round is shown in Figure 3.9.

In line with the other two studies, we used LimeSurvey to present the tasks and the questionnaire. The participants had to open the prototype (described in Section 3.3) parallel to the survey.

#### 3.6.3 Procedure

The survey started with an introduction to the study and the Netflower tool (Stoiber et al., 2019b). The participants were given a scenario in which they should imagine themselves as data journalists and asked to complete 14 tasks. The first 11 tasks were presented as multiple-choice aimed for a distinct answer, e.g., "Compare Malaysia, Sweden, and Austria in 2016. Which country received the most asylum seekers?" In contrast, the last three questions were posted as free-text answers to understand the analysis behavior, e.g., "Compare the development of asylum seekers between 2000 and 2016 between Austria, Germany, and Sweden. Is the development in Austria more similar to Sweden or to Germany? Explain why you think that this is the case?" The tasks were chosen based on VLAT (Lee et al., 2017), as discussed in section 3.3. After the tasks, they were given a questionnaire of six to nine questions, part multiple-choice and part open-ended, that aimed to make the participants reflect and give feedback. Some questions referred to the usage of the onboarding system, some referred to the quality of the onboarding system and others were more open-ended. The final questions asked about the participants' user experience concerning the onboarding system. The whole survey can be found here: https://phaidra.fhstp.ac.at/o:4872 and in Appendix A.

#### 3.6.4 Data Sets

Netflower is a visual exploration tool for investigating flows between networks over time (Stoiber et al., 2019b), see Figure 3.3. For our study, we used data about asylum seekers published by the UNHCR (the UN refugee agency) that shows asylum applications from the years 2000 to 2016 in 44 different countries (UNHCR, 2020) as well as media transparency data (RTR, 2022) about money flows between government entities and media institutions. The media transparency data set is compiled by the Rundfunk und Telekom Regulierungs-GmbH and records money given to media institutions by Austrian government entities (RTR, 2022). In each first session, the asylum seeker data set was used, and in each second session, the media transparency data set was used.

#### 3.6.5 Results from Tasks

We assessed answer correctness (in percent) on multiple-choice questions, as well as open-ended questions. As study 1 found significant differences in answer correctness based on task difficulty as categorized by Friel et al. (2001), we also checked if this holds true for study 3. For the analysis of the free-text questions, we defined two main categories:

**Category 1**: Answers, where the correct usage of the system is apparent, for example, by having the correct answers or by having written descriptions of how the system was used.

**Category 2**: Any other answers, for example, wrong answers, no answers, answers not rooted in the visualization. An example of an answer for task 12 that was judged to be this category would be "Austria is more similar to Germany due to the geographic location". This could be a valid argument, but by itself does not show that they used the system correctly to come to this conclusion.

We assigned these categories to the free-text answers and evaluated them like the multiplechoice answers (in percent).

**Overall Comparison:** Comparing all results of participants using the scrollytelling approach for the first time (n = 51, M = .72; SD = .159) with those using the tutorial for the first time (n = 52, M = .75; SD = .159), we can see no difference in answer correctness (p = .317). Here, we use data from session one in round one, as well as both sessions in round two.

**Learning Effect:** The goal in the first round (in spring 2020) was to establish the properties of our test setup. We wanted to know how the data set influences responses, as well as how well the participants retain their skills of using Netflower after a month. Comparing the answers of session 1 (n = 26, M = .77; SD = .103) and session 2 (n = 24, M = .78; SD = .116) in spring 2020, we can see that there is no appreciable difference in answer correctness (p = .878).

Within-Subject Comparison: We can compare within subjects by looking at only the answers in round two. However the answers of participants using the in-situ onboarding (n = 38, M = .71; SD = .171) compared to those using the external one (n = 39, M = .74; SD = .176) show no appreciable difference in answer correctness either (p = .517).

**Task Difficulty:** We could not find a significant difference in answer correctness between the external onboarding and the in-situ scrollytelling when grouping by task difficulty. A t-test reveals p = .343 for elementary tasks, p = .483 for intermediate tasks and p = .598 for comprehensive tasks. When analyzing the task difficulty using the taxonomy by Lee et al. (2017), we noticed that the taxonomy only comprises fairly simple interaction processes. The added additional tasks based on the model by Friel et al. (2001) include more complex interaction processes as, for example, relating values from different visualizations in Netflower and then deriving conclusions from this process. This task posed some problems to inexperienced users although information about these interaction processes could be found in the onboarding systems. In accordance with the model developed by Friel et al. (2001), the simplest cognitive processes (identifying simple values) were easiest. Comparison processes were more difficult, and the most difficult cognitive processes were "compute derived value" which forced participants to "go beyond the data". In general, we noticed that participants made many mistakes as, in some cases, more than 60% of the answers were incorrect. This contradicts the subjective impression of some of the participants that they could understand Netflower without studying the onboarding system and their confidence in the correctness of the results.

#### 3.6.6 Results from the Questionnaire



Figure 3.10: **Participants' attitude towards onboarding concepts**. *Note*: Not all participants answered these questions as they were not mandatory.

The questionnaire consisted of two parts, one part containing multiple-choice questions and another part consisting of open-ended questions (see Figure 3.10).

**Multiple-choice questions:** When asked whether they started to study the onboarding system, or on the other hand, they started to explore the system on their own, almost 50% of the participants reported that they read the text in the scrollytelling system first while only about one-fifth of the participants in the tutorial group did. We could not detect any difference with regards to the intensity of usage of the onboarding system: In both conditions, about one-fifth of the participants reported that they used the onboarding system "often" and about 50% of them reported having "never" used it at all. The rest reported having used the system "sometimes". We were also interested in the acceptability of the amount of information displayed. Therefore, we asked if there was too much text in the onboarding system. Again, no differences between the two conditions could be determined. About one-third claimed that there was too much text for both methods, and the rest did not agree with this statement. For the scrollytelling condition, we additionally asked about the helpfulness of the gradual appearance of the elements, which was perceived by about 73% as helpful. Additionally, the simple introduction to the system was mentioned to be supportive.

**Open-ended questions:** The open-ended questions were grouped and treated as one question that can be assigned to multiple categories. The questions that had more varied responses were further evaluated using qualitative content analysis (Mayring, 2010; Schreier, 2012) and discussed here. The qualitative content analysis aimed to analyze textual or other material systematically. It is applied when the analysis process does not yield entirely obvious categories and an extended interpretation of the material is necessary. The analysis consisted of repeatedly coding the material and results in a few categories describing the most important insights that can be derived from the material.

In this part of the questionnaire, whether participants appreciated the onboarding system (either in-situ scrollytelling or tutorial) yielded interesting results. The participants were specifically asked to discuss the advantages and disadvantages. Therefore, the answers tend to include both positive and negative aspects.

In general, participants liked scrollytelling more than the tutorial. Six participants appreciated scrollytelling very much, 14 liked it, nine were undecided, and seven did not like it. In contrast, only one participant liked the tutorial very much, six appreciated it, 18 were undecided, and 15 did not like it. Note that the open-ended questions were not answered by all participants as they were not mandatory.

The following statements describe the positive attitudes concerning scrollytelling. "I liked it because it visually highlighted the important parts, and I was less overloaded with all the possibilities in the beginning because it showed me one feature at a time" (P3). "I like that it is quick and that it doesn't require too much time to be learned and understood. I liked that the text was not too invasive, being on the side and having the possibility to hide it" (P16).

The most important positive aspects of the system that were mentioned in the answers were: (1) It visually highlights the most important parts and shows one feature at a time (11 mentions); in this way, it also provides an inherent structure of the system. (2) It is directly inside the system, therefore, using it does not disrupt the workflow (7 mentions). (3) It is a simple introduction to the system (13 mentions).

Participants also outlined drawbacks of the system: (1) The most important negative aspect was technical and usability problems (10 mentions). The scrollytelling system was a prototype, and sometimes usability issues occurred. A few of the participants mentioned that they did not need scrollytelling and discovered the system's features themselves (5 mentions). One example of a negative attitude concerning scrollytelling is the following:

"I did not really like it. The system to use is quite intuitive, and the scrollytelling is, in comparison, too long and convoluted. At the same time, it did not cover enhanced topics like notes and tags. I prefer an approach where one starts immediately using it, and later gets the possibility to learn more details" (P12).

The tutorial was appreciated less than the scrollytelling method. Nevertheless, it was stated that "I liked the tutorial because it allowed me to start working quickly. Whenever I struggled, I could go back and read a little more" (P22). The following positive features of the tutorial were mentioned: (1) The tutorial is easy to comprehend, it is nice as a reference, and it has a good overview (15 mentions). (2) Some of the participants also mentioned that they liked the videos (8 mentions).

Several participants had negative attitudes concerning the tutorial: "The tutorial was not really needed because the visualization was quite intuitive, and one could guess what functions and filters were available by trying them out. Still, the tutorial gives a good overview about what tools are available" (P11). In the following, we present the most important negative remarks: (1) Many of the participants mentioned that the tutorial was not needed and that they just started working (8 mentions). This is especially interesting because this was not the case with scrollytelling. The explanatory text was the same, but participants invested much more time interacting with scrollytelling than with the tutorial. Apparently, scrollytelling is more motivating than the tutorial. (2) Several of the participants also mentioned that the tutorial only explained simple things and did not address the more complicated features of the system. They would also have wanted sophisticated examples (9 mentions). (3) Very few participants also mentioned that they did not like videos (3 mentions). Apparently, the participants' attitude toward videos is ambiguous. Some like them very much, and others do not like them at all.

#### 3.6.7 Lessons Learned

In this investigation, we distinguished between the performance measures and attitudes of the participants. In general, we found that there were no significant differences in the performance of participants using the tutorial and scrollytelling. This may be because the visualization used was not that demanding for computer science students. Many participants mentioned that they figured out how the visualization features worked. Nevertheless, there were single tasks that were difficult for the participants. Even in these cases, participants sometimes did not use the onboarding systems to solve these tasks. Integrating solved tasks (worked examples) and small challenges might help to overcome this problem.

On the other hand, the attitude of the participants was decidedly in favor of scrollytelling. Participants reported using scrollytelling more often than the tutorial before starting to work. Many of them stated that they found scrollytelling more appealing than the tutorial. Features they found especially appealing were the fact that single components of the visualization appeared gradually and that the onboarding system was integrated into the system and did not interrupt the workflow. The positive features of the tutorial were that it provided a nice overview and contained videos. The attitude concerning videos is mixed; some participants explicitly stated that they did not like videos.

### 3.7 Discussion & General Lessons Learned

Introducing users to a new visual layout and guiding them to a higher visualization literacy is the fundamental goal of visualization onboarding. However, as data sets are increasing in complexity, well-known conventional business charts (e.g., line charts, bar charts, and pie charts) are no longer sufficient, and new forms of visualization—especially designed to deal with this increasing complexity—need to be applied. The lack of onboarding and the reduced focus on how to read a chart lead us to this contribution. Thus, we wanted to explore different aspects of visualization onboarding: the general effect of visualization onboarding on user performance (study 1), the effect on user performance of different types of visualization onboarding (study 2), and the use of in-situ visualization onboarding in the VA tool Netflower (study 3).

**Experience Level and Answer Correctness.** Besides the limitation of the study design, we can report on lessons learned regarding the answer correctness (in %) in general and the answer correctness concerning the subjectively-rated level of experience of MTurk workers. In line with the previous work by (Börner et al., 2019, 2016; Galesic and Garcia-Retamero, 2010), we also see poor answer correctness over different onboarding concepts (cf. Study 2 3.3.1 and Table 3.2). Our results show answer correctness around 50%. For the Baseline (no onboarding), we can report the answer correctness of the tasks of 52%. This is an alerting result and indicates the need for visualization onboarding.

The analysis revealed discrepancies when comparing the subjective assessment of the experience with different visualization techniques to the answer correctness of prequestions. When assessing the influence of the subjectively rated experience levels with the visualization types (pre-questions), we observed that highly experienced participants did not show high answer correctness. In contrast, rather inexperienced users achieved higher performance measures.

A similar phenomenon occurred with the computer science students in study 3. They were confident about their ability to solve tasks with the Netflower tool. Nevertheless, in certain task categories requiring less obvious modes of interaction, the results were fairly mixed. This indicates that users should consult the onboarding system more often than they do. In future work, we plan to address this issue, that is, how to motivate users to work with the onboarding system more intensively and check their solutions more often. Possible solutions for this might be to add tasks with solutions (worked examples) or small challenges to the onboarding system so that users can check whether their method of problem-solving was appropriate.

**Summary of Lessons Learned from all Studies.** Taking together all the results of the three studies, we can formulate the following lessons learned regarding visualization onboarding:

- Subjective experiences levels per visualization type do not correlate with high answer correctness.
- Participants prefer the video tutorial to learn how to interact with the parallel coordinates plot (over scrollytelling and step-by-step guide).
- In-situ scrollytelling is highly preferred over a tutorial due to its integration into the Netflower tool and gradual guidance through the instructions.

#### 3.7.1 Towards Design Actions for Visualization Onboarding Methods

To facilitate the creation of visualization onboarding methods, we summarize some core advice we have learned during our research. In the following, we list the design actions (Spence, 2014).

- In general, onboarding systems that are **integrated** into the visualization system are appreciated. Users prefer systems that do not disrupt the workflow. In such systems, important **features** should be **highlighted**.
- Use an **easy-to-understand data set** and **concrete examples** on how to read the charts, support, and increase comprehension is vital.
- To-the-point descriptions make it easier to absorb information (step-by-step).
- Use **videos** or small animations as an onboarding method to explain interaction techniques. In contrast, videos are appreciated by a minority. Other users have a negative attitude. It depends on the target group whether they can be successfully deployed.
- Some users tend to ignore onboarding systems, even if they struggle. These users have to be **motivated to use onboarding**.

#### 3.7.2 Limitations

**Study Design.** A limitation of our conducted studies 1 and 2 is the study design. We designed the study, as illustrated in Figure 3.5, integrating pre- and post-questions to assess the improvement of participants. Participants assigned to the with or without onboarding condition had to answer those questions. We couldn't find any statistical significance for answer correctness between the groups. We assume that the participants without onboarding conditions could increase their answer correctness because of the repetitive nature of the questions. So there was already a learning effect visible for participants who did not interact with the visualization onboarding.

**Target Group.** Additionally, the analysis of the interaction time with the visualization onboarding (see Study 1) showed that the MTurk workers may not take enough time with the onboarding concept itself to carefully read through the text, for example.

This is a limitation of our study as we should have applied a scaling incentive system where participants get paid more for correct answers to encourage them to engage more with the onboarding concept. Participants who used MTurk worked fairly briefly and got a visualization with reduced interaction possibilities. The computer science students in study 3 worked for a longer time and used a fairly complex tool with many different possibilities to solve tasks. This difference might be responsible for differences in attitude to onboarding systems in studies 1 and 2 and in study 3. The scrollytelling method is probably more useful for complex visualization tools when users need more detailed information about the tool and flexible access. All three studies indicate that an integrated, step-by-step approach was appreciated more by users than a stand-alone onboarding system. The integrated system helps users to relate the information of the onboarding system to the features of the visualization, and the step-by-step procedure avoids information overload for inexperienced users. Especially in study 3, we found that participants sometimes had difficulties knowing when they needed help. They were overconfident about their ability to solve problems and primarily relied on the possibility to find out about the features of the visualization by trial and error. Therefore, they had specific difficulties to solve tasks that required them to use features of the system in a way that was not obvious. Future work should clarify how this problem could be solved. There is some indication that tasks with solutions and small challenges can help to overcome this problem.

**Hotjar tracking.** We assessed the interaction behaviors with the Hotjar (Hotjar, 2023) tool. The blocking software in the browser and other available options protect users from tracking or recording their actions in the browsers, restricted our tracking and we could just collect data from 91 out of 145 participants of all studies. Furthermore, as Hotjar builds heatmaps on top of screenshots, we could not assess the scrolling behavior in our scrollytelling onboarding method and thus did not assess mouse movements for the method comparison.

## 3.8 Conclusion

We presented the results of three studies with MTurk workers as well as students (see Figure 3.1) to investigate how visualization onboarding affects users' performance at different levels. Firstly, we assessed the effect of visualization onboarding in performance measures between different visualization types — for a bar chart, horizon graph, change matrix, and parallel coordinates plot (study 1). We identified significant differences in the answer correctness between the four visualization types. Horizon graphs and bar charts show more correct answers than the other visualization types. Relating the response time, the fastest answers were given to the bar chart, horizon graph, change matrix, and parallel coordinates plot. We could not determine statistically significant differences between the answer correctness in surveys with or without onboarding over all four visualization types. The participants in the onboarding condition. Thus, we could not identify differences or improvements based on all visualization types. Furthermore,

we also assessed differences between task types of *reading the data, reading between the data,* and *reading beyond the data.* Irrespective of the onboarding method and the visualization type, *reading the data* questions showed the highest mean rank, followed by *reading between the data,* and *reading beyond the data* question types. Overall, the onboarding influences answer correctness for the change matrix. The results for the most unfamiliar visualization type — the parallel coordinates plot — show that onboarding is needed to improve performance while interacting with the chart.

We conducted a further study with MTurk workers to investigate the effect on user performance for four different onboarding types (step-by-step, scrollytelling, video tutorial, and in-situ scrollytelling in Netflower) in studies 2 and 3. The results of study 2 showed that no statistically significant difference in the answer correctness for the post-questions could be determined between the different onboarding types. However, the response time showed that faster answers were given after a scrollytelling tutorial, followed by the step-by-step guide, the condition without any onboarding, and the video tutorial. The analysis of onboarding and task difficulty by Friel et al. (2001) showed that participants *reading the data* tasks were answered the fastest, followed by *reading between the data*, and *reading beyond data*. The analysis of the qualitative feedback revealed that independently of the visualization type and method applied, an easy-to-understand data set and concrete examples on how to read the chart, support, and increase comprehension are vital.

For study 3, we gathered data on user experience by using the in-situ scrollytelling for the VA tool Netflower (Stoiber et al., 2019b). The results of the evaluation with students showed that while there was no significant difference in answer correctness (see 3.6.5), we did find a difference in attitude: Participants reported preferring in-situ onboarding to an external one.

We learned that concrete examples of how to read the chart could support and increase comprehension. Therefore, we further investigate onboarding instructions and how to phrase them. In the literature, there is a discussion about concrete vs. abstract examples to teach new concepts (De Bock et al., 2011; Kaminski et al., 2008). Thus, we conducted a study with 40 students to investigate abstract and concrete onboarding instructions and assess which is more appropriate for users unfamiliar with a treemap visualization. In the next chapter 4, we describe the study design, the instructional design, and the results.

74



# Abstract and Concrete Visualization Onboarding Instructional Material for a Treemap Visualization



This chapter presents the results of a comparative study to investigate abstract and concrete onboarding instructions. We introduce abstract and concrete onboarding instructions for a treemap visualization (see Section 4.3.1). Furthermore, we describe in Section 4.4 the results of the comparative study with 40 students in detail. All materials used, such as the collected data from Google Forms and the participants' comments, are included in Appendix B. This chapter is based on our published work (Stoiber et al., 2022b).

# 4.1 Introduction & Motivation

In the literature (De Bock et al., 2011; Kaminski et al., 2008), there are discussions about concrete vs. abstract examples to teach new concepts, including examples for both approaches. According to Kaminski et al. (2008), concrete examples would hinder learners from transferring and generalizing new concepts. However, more recent studies provide counter-evidence. For instance, De Bock et al. (2011) shows a more successful knowledge transfer using concrete examples. Similarly, when introducing new visualizations, it is recommended to use an easy and understandable data set that can be assumed to be well-known by the general public (Freedman and Shah, 2002; Kosslyn, 1989; Lee et al., 2016).

To fill this research gap, we aim to investigate abstract and concrete onboarding instructions to assess which is more appropriate for users unfamiliar with a particular visualization technique. Therefore, we conducted a within-subject study with 40 students. We collected different data, which we describe in the following in more detail. The chapter presents the following contributions:

- We present abstract and concrete onboarding instructions for a treemap visualization for two different data sets: Biden's tax overhaul (Committee for Responsible Federal Budget, 2021; Probasco, 2021) and The Austrian federal budget of 2022 (Federal Ministry Republic of Austria Finance, 2022) (see Section 4.3.1). In our previous study (Stoiber et al., 2021), the need for onboarding concepts was visible, especially for more complex visualization techniques. Treemaps (Shneiderman, 1992) are well suited to represent hierarchical data with a quantitative attribute. However, treemaps are not particularly well-known among audiences and might be hard to interpret initially. Therefore, we decided to generate onboarding messages for a treemap visualization.
- Furthermore, we present the results of the comparative study with 40 students. First, we show the results of a 7-point Likert scale towards the quality of the onboarding instruction per condition abstract and concrete.
- Additionally, we present the categorization schema (Section 4.4.1) we used to assess the quality and the value of the descriptions and insights the participants had to provide. We show the results of this analysis in Section 4.4. Besides, we present the sentiment analysis results of the subjective feedback in Section 4.4.5.
- We derived design implications for the phrasing of onboarding instructions, which we list in Section 4.5.1.

# 4.2 Related Work

Recommendations regarding on the used data set tend to be mixed when explaining visualization systems and the underlying data structures. On the one hand, literature

calls for abstract explanations to be more generalizable and transferable to a novel context (Kaminski et al., 2008). On the other hand, more recent scientific work provides counter-evidence. For instance De Bock et al. (2011) shows a more successful knowledge transfer using concrete examples. Similarly, when introducing new visualizations, it is recommended to use an easy and understandable data set that can be assumed to be well-known by the general public (Freedman and Shah, 2002; Kosslyn, 1989; Lee et al., 2016). In the following sections, we discuss the related literature using concrete vs. abstract material for teaching in general, as well as for teaching visualizations. Furthermore, we present studies in the field of visualization onboarding.

#### 4.2.1 Concrete vs. abstract materials

A longstanding debate concerns using concrete versus abstract instructional materials. particularly in domains such as mathematics and science (Fyfe et al., 2015). Concrete materials are widely used, including physical, virtual, and pictorial objects (Bryan et al., 2007). Concrete materials can provide a practical context and activate real-world knowledge (Schliemann and Carraher, 2002). A second advantage is that concrete material can enhance memory and understanding by inducing physical or imagined action (Glenberg et al., 2004). Brown et al. (2009) also found that concrete materials enable learners to construct their knowledge about abstract concepts. Concrete exemplification enhances transfer in a new concrete domain more than abstract exemplification (De Bock et al., 2011). However, studies report disadvantages such as distracting learners from the relevant information (Kaminski et al., 2008) or constrained transfer of knowledge (Sloutsky et al., 2005). In contrast, abstract material eliminates extraneous perceptual details and can increase generalizability to various contexts (Son et al., 2008). Learners can focus on the structure and representative aspects rather than the surface features (Uttal et al., 2009). One weakness of abstract learning material is that learners manipulate meaningless symbols without conceptual understanding (Nathan, 2012). To summarize. the mentioned research outlines no clear direction for using concrete or abstract learning materials.

#### 4.2.2 Visualization Onboarding

Educational Community: Little is known about the effectiveness of concrete vs. abstract materials for visualization onboarding instructions. The educational community has studied how students interpret and generate data visualizations (Baker et al., 2001) and how to teach bar charts in early grades (Alper et al., 2017) using a tablet application called "C'est la vis", supporting elementary school pupils to learn how to interpret bar charts based on the concreteness fading approach (Fyfe et al., 2015). Concreteness fading is a pedagogical method in which new concepts are presented with concrete examples before progressively abstracting them. In recent literature, Wang et al. (2020) presents a set of cheat sheets to support visualization literacy around visualization techniques inspired by infographics and data comics, which are well-established onboarding methods in domains such as machine learning. The cheat sheets use a combination of abstract and

#### Description

Also known as a Scatter Graph, Point Graph, X-Y Plot, Scatter Chart or Scattergram.

Scatterplots use a collection of points placed using Cartesian Coordinates to display values from two variables. By displaying a variable in each axis, you can detect if a relationship or correlation between the two variables exists.

Various types of correlation can be interpreted through the patterns displayed on Scatterplots. These are: **positive** (values increase together), **negative** (one value decreases as the other increases), **null** (no correlation), **linear**, **exponential** and **U-shaped**. The strength of the correlation can be determined by how closely packed the points are to each other on the graph. Points that end up far outside the general cluster of points are known as **outliers**.

Lines or curves are fitted within the graph to aid in analysis and are drawn as close to all the points as possible and to show how all the points were condensed into a single line would look. This is typically known as the **Line of Best Fit** or a **Trend Line** and can be used to make estimates via interpolation.

Scatterplots are ideal when you have paired numerical data and you want to see if one variable impacts the other. However, do remember that correlation is not causation and another unnoticed variable may be influencing results.



Figure 4.1: Scatterplot - abstract explanation (Ribecca, 2020)

concrete instructions. A concrete data set (nutrition values like calories, calcium, water, etc.) explains the anatomy, construction, and introduction to the visualization technique.

Scientific community: Furthermore, Kwon and Lee (2016) followed the "Experiential Learning Model" (Kolb, 1984). The knowledge is constructed via concrete experience and reflection on the experience. They used the well-known car data set (Heinrich and Broeksema, 2023). Approaches by (Ruchikachorn and Mueller, 2015; Tanahashi et al., 2016) use concrete onboarding by using concrete data sets to explain a certain visualization technique and facilitate learning. For example, Ruchikachorn and Mueller (2015) explored the learning-by-analogy concept by demonstrating an unfamiliar visualization method by linking it to another more familiar one. The authors found that the learning by analogy concept is useful as participants in their study could fully or at least significantly better understand the unfamiliar visualization methods after they observed or interacted with the transitions from the familiar counterpart.

Additionally, Tanahashi et al. (2016) investigated top-down and bottom-up teaching methods and active or passive learning types. The bottom-up teaching method ("textbook

78

approach") (Zeid et al., 2011) focuses on small, detailed pieces of information that students then combine to get a better understanding. Besides, a top-down teaching method is given when a broad overview first helps to understand the abstract, high-level parts of an idea/topic, which then provide context for understanding its components in detail (Tanahashi et al., 2016). Furthermore, a distinction can be made between active and passive learning types. Passive learning means that students only receive the information without participatory dialog. In contrast, active learning describes active participation. Their analysis indicates that top-down exercises were more effective than bottom-up and active learning types, with top-down exercises being the most effective.

**Visualization Tools & Visualization Platforms:** Besides scientific literature, onboarding concepts are also integrated into commercial visualization tools, which focus on explaining features using abstract instructions.

For example, Advizor Solutions (2023) uses highly abstract textual descriptions to explain the visual mapping for visualization techniques, e.g., "A Scatterplot shows the interaction of two fields. They may be continuous (numbers) or categorical (names), although two continuous fields work best. Size and color of data points can be used to show additional dimensions". In contrast, IBM Cognos Analytics (2019) uses step-by-step tours with tooltips and overlays to onboard new users. The tour includes the creation of visualizations using a concrete data set — donations per year in Mio \$.

Furthermore, the *Data Visualisation Catalogue* (Ribecca, 2020) seeks to support users to understand the encoding and building blocks of different visualization types using abstract instructions and data, see Figure 4.1 for a Scatterplot. Furthermore, *From Data to Viz* (Holtz and Healy, 2018) aims to find an appropriate visualization type based on the input data using a decision tree. The catalog offers definitions, variations, and the use of each visualization type, as well as potential issues that may arise during use and interpretation. All the definitions and explanations are abstract, though.

Therefore, we developed two sets of onboarding instructions and conducted a comparative evaluation to assess which ones are more appropriate. The concrete instructions refer to the data, and the abstract instructions are phrased without referring to the data set. Section 4.3.1 presents the onboarding instructions in detail.

### 4.2.3 Hypotheses

We developed our hypotheses based on the existing literature and open questions of other conducted studies (Stoiber et al., 2019a, 2021).

- **H-Quality:** We expect participants to rate the concrete onboarding instruction as more *understandable*, *helpful*, and *complete* compared to the abstract ones. The participants also assess the *length of instructions as short enough*.
- **H-ExperienceLevel**: We expect that the participant's experience level affects the statements.

- **H-Value:** We expect participants to generate more valuable insights and descriptions after reading the concrete onboarding instructions than reading the abstract instructions. We believe that referring to the data shown in the visualization and giving concrete examples improves insight generation.
- **H-Preference:** We expect *concrete onboarding is subjectively preferred over abstract onboarding.* This means participants are more likely to read the instructions and feel more confident interpreting and reading the visualization.

# 4.3 Evaluation

This study uses Google Forms to compare abstract and concrete onboarding instructions in a within-subject study design setting. We were generally interested in assessing the usefulness and understandability of the abstract and concrete visualization onboarding instructions for a treemap visualization. As data visualizations aim to support users in gaining valuable insights into the data (Card et al., 1999), participants had to identify and describe insights. As visualization onboarding can support users in interpreting data visualizations, especially when they are not experienced in visual data analysis, we invited students in the first and third semesters of the study program Media Technology at St. Pölten University of Applied Sciences. We also carried out a pilot study with two experts in usability and visualization to ensure the suitability and correctness of the onboarding instructions.

### 4.3.1 Visualization Onboarding Instructions

In the following section, we describe the design, the general layout of the prototypes, the onboarding messages, and the implementation.

**Design and Architecture:** We include four figures (Figure 4.3 and 4.2) of all four prototypes, collected here: https://treemaps.netlify.app/. All the prototypes are shown on the left side of the treemap visualization and the right side of the onboarding messages (see Figure 4.2, 4.3, 4.5, and 4.4).

The data story by Reuters <sup>1</sup> visualizing the Biden's tax overhaul inspired us. The data set in the visualization area is either Biden's tax overhaul (Committee for Responsible Federal Budget, 2021; Probasco, 2021) or the Austrian federal budget of 2022 (Federal Ministry Republic of Austria Finance, 2022), which is shown as a hierarchical treemap. We used similar data sets to provide comparability between conditions. In comparison, we selected a similar data set — the Austrian federal budget of 2022 (Federal Ministry Republic of Austria Finance, 2022). Both data sets are publicly available.

Each cluster of tiles has a specific color that is unique to each category. Hovering over a rectangle reveals detailed information about the current entity, which includes the

<sup>&</sup>lt;sup>1</sup>https://www.reuters.com/graphics/USA-BIDEN/INVESTMENTS/xlbvgkbxlvq/, Accessed: 2023-01-30



Figure 4.2: Abstract onboarding instructions for the treemap showing Biden's tax overhaul (Committee for Responsible Federal Budget, 2021; Probasco, 2021). The visualization area (1) with the treemap visualization — Biden's tax overhaul (Committee for Responsible Federal Budget, 2021; Probasco, 2021). The onboarding area (2) contains abstract onboarding messages split into three thematic sections: *Reading the chart* (2.a) provides instructions on how to read the treemap. In *Interacting with the chart* (2.b) the reader learns about the interaction possibilities, and in *Analyzing the chart* (2.c) examples and insights are provided. Prototypes: https://treemaps.netlify.app/

name, a short description (if available), and the exact value. Above each category, the total values and the category name are displayed. The messages are tailored to the respective data set. We distinguish between **concrete** and **abstract** onboarding messages, while the former tries to include concrete facts from the visualizations, e.g., values or insights — "The size of each rectangle represents the spendings in US-Dollars (e.g., the Housing rectangle, representing \$213 billion, is approx. twice as large as the Clean drinking rectangle, representing \$111 billion.)" The latter tries to provide a more generic description of the treemap which doesn't include any concrete values from the visualization, e.g., "The size of each rectangle represents a quantitative value associated with each element in the hierarchy." In general, the onboarding messages are grouped into Reading the chart — explaining the general encoding (e.g., size, color), Interacting with the chart — explaining the possible interactions with the visualization, and Analyzing the chart — trying to guide the reader towards further insights (e.g., making comparisons). The structure is based on our previous work (Stoiber et al., 2021, 2022d). In Figure 4.2, exemplary, the abstract onboarding prototype for Biden's tax overhaul plans can be seen.

**Implementation:** We used HTML, CSS, and JavaScript as our main languages to create the four prototype websites. For bootstrapping the applications and creating the

# 4. Abstract and Concrete Visualization Onboarding Instructional Material for a Treemap Visualization



Figure 4.3: Concrete onboarding instruction for the treemap showing Biden's tax overhaul (Committee for Responsible Federal Budget, 2021; Probasco, 2021). The layout and structure are the same as in Figure 4.2.

basic setup, we used Vite JS (You, 2019) — a frontend build tool. For the visualization, we used the standard and well-known library D3.js (D3 Language, 2023). The treemap was created with d3, and only one additional library was used to enhance the user experience. The library is called d3-v6-tip (https://www.npmjs.com/package/d3-v6-tip, Accessed: 2023-01-30) and is used to generate the tooltips for the treemap.

#### 4.3.2 Study Design & Data Collection

This study used a within-subject design with students. We collected the following data during the study: (1) Task 1: *demographic information* such as gender, age, and color blindness, and assessed their experience with data visualizations and the treemap visualization. (2) Task 2: descriptions of the treemap in participants' own words; (3) Task 3: at least three insights; (4) rating based on a 7-point Likert scale of four statements to assess the quality of the onboarding instructions; and (5) answers to open questions to gather more comments, suggestions, and remarks regarding the onboarding instructions.

There were two independent variables: onboarding instruction type (concrete or abstract) and data set (Biden's tax overhaul plan (B) or Austrian federal budget payments (A)).



Figure 4.4: Concrete onboarding instruction for the treemap showing Austrian federal budget of 2022 (Federal Ministry Republic of Austria Finance, 2022). The layout and structure are the same as in Figure 4.2.

The dependent variables were question scores (4) to four statements regarding the textual instructions. We used Google Forms to set up the online questionnaire. Students had to enroll in the test with their student ID. Then, we randomly assigned students to the concrete and abstract onboarding (counterbalancing) and randomly assigned them to the two data sets. We used the university's internal Moodle platform to create groups per condition and publish the link to the study.

#### 4.3.3 Participants

There were 40 students (Abstract A: 18, B: 22; Concrete A: 22, B:18) (Gender: m = 18, f = 20, prefer not to say = 2) conducting the study between 2022/01/25 to 2022/02/01. The students were from the first and third-year bachelor study program Media Technologies at St.Pölten University of Applied Sciences. The participants were between 18 and 27

# 4. Abstract and Concrete Visualization Onboarding Instructional Material for a Treemap Visualization



Figure 4.5: Abstract onboarding instruction showing Austrian federal budget of 2022 (Federal Ministry Republic of Austria Finance, 2022). The layout and structure are the same as in Figure 4.2.



Figure 4.6: (A) Level of **experience for data visualizations** and for the (B) **treemap visualization** on a 7-point Likert scale, where 1 means no experience at all and 7 is highly experienced. The blue dashed line shows 4 (experienced) as the median (Mdn) for data visualization in general and 3 (little experienced) as the median (Mdn) for the treemap visualization.

years old. Two of the participants stated that they have red-green/blue-purple color blindness. The survey started with a question about the level of experience with data

84

visualization in general (Figure 4.6) as well as the level of experience with the treemap visualization (Figure 4.6) by using a 7-point Likert scale. In general, the participants rated their level of experience "little" (3 and 4 (Mdn) on the Likert scale).

#### 4.3.4 Procedure

At first, the participants had to fill in demographic information. As the next step, they continued with Task 2 (description) and Task 3 (insights).

Additionally, they had to fill out a questionnaire to assess the quality of the onboarding instructions based on a 7-point Likert scale. We presented four statements they had to assess: **(S1)** The textual instructions were easy to understand. **(S2)** The textual instructions were short enough. **(S3)** The textual instructions were helpful in understanding the treemap visualization. **(S4)** All relevant aspects to understanding the treemap were described.

We also integrated two open questions to gather more comments, suggestions, and remarks regarding the onboarding instructions. As it was a within-subject design, the participants started by assessing the second condition with another onboarding type and data set with the same tasks and questions.

# 4.4 Results

All 40 participants completed the study. We had to exclude the answers of one participant on task 2 and task 3 as the participant was not working on the tasks. We downloaded the data from Google Forms and included all the data in one analysis spreadsheet. The analysis table can be found in the supplemental material. We used R to analyze the data as well as produce the plots.

### 4.4.1 Categorization of descriptions and insights

We defined categories for the data analysis of task 2 and task 3. For task 2 (descriptions), we assessed the quality of the answers of all participants on a scale of 1 - 3. Whereas 1 means trivial description, 2 intermediate descriptions, and 3 highly understandable explanations of the treemap visualization. Furthermore, we also coded if the descriptions contain the following information with Yes and No: description of *color* and *size* of rectangles of the treemap, *interaction*, and an example based on a distinct value. Insights characteristics (task 3) were inspired by the paper by Saraiya et al. (2005). We define insight as an individual observation of the data by the participant. The following quantifiable characteristics of each insight were then to be encoded:

• **Observation:** The actual finding about the data. We counted distinct data observations by each participant. As the participant had to find at least three insights, the number of observations did not vary much.

- Value of Insight: The value, importance, or significance of the insight. Simple observations such as "Most money has been spent on public pension schemes." are assessed as trivial, whereas observations comparing two different subcategories "The president's biggest spending section would be the child tax credit expansions in which he aims to extend the child tax credit expansion and to make the child tax credit expansion fully refundable." earned 2 points. And insights that included domain knowledge/experience or created new hypotheses earn 3 points.
- Correctness of Insight: Some insights were incorrect observations that resulted from a misinterpretation of the visualization. The coders code this with Yes/No per insight.
- Task Difficulty Taxonomy: We also categorized the insights along the Task Difficulty taxonomy by Friel et al. (2001) including (1) reading the data (identify one value, or biggest or smallest rectangle in the sub-category of the treemap), (2) reading between the data (make comparisons between different sub-categories in the treemap), and (3) reading beyond the data (own interpretation).

Three coders separately categorized the answers to task 2 and task 3 alone. In an online meeting, we discussed the samples we disagreed on and concluded.

#### 4.4.2 Quality

The quality of the onboarding instructions (**H-Quality**) was assessed by participants along a 7-point Likert Scale. Figure 4.7 shows the boxplots of the four statements per condition and data set. Table 4.1 gives an overview of the mean and standard deviation for a 7-point Likert scale [1-7 strongly disagree to agree strongly].

Statement		Ranking							
	Abstract A	Concrete A	Abstract B	Concrete B					
S1 (understandable)	$6.17 (SD \ 0.62)$	$5.68 (SD \ 1.48)$	$5.55 (SD \ 1.22)$	$6.11 (SD \ 0.68)$					
S2 (helpful)	$5.83 (SD \ 1.47)$	$5.05 (SD \ 1.33)$	$6.18 (SD \ 1.05)$	$5.00 (SD \ 1.53)$					
S3 (length)	$6.11 (SD \ 0.9)$	$6.23 (SD \ 1.02)$	$5.82 (SD \ 1.14)$	$6.33 (SD \ 1.08)$					
S4 (complete)	$6.33 (SD \ 0.77)$	6.09 (SD 1.06)	$5.95 (SD \ 1.36)$	$6.33 (SD \ 0.77)$					

Table 4.1:	Mean	and	SD	$\mathbf{per}$	condition	and	data	$\mathbf{set}$
------------	------	-----	----	----------------	-----------	-----	------	----------------

A Friedman test was conducted to determine whether statement scores (Likert data) differ between onboarding instruction types and one of the four statements and the data set. The results show significant differences in the type of onboarding in statement **S2** (length of onboarding messages) (p-value = 0.0002607) and in statement **S3** (helpfulness) (p-value = 0.04123). Participants rated the length of the concrete onboarding instructions



Figure 4.7: Boxplots showing the results of the four statements per condition. A Friedman test show significant differences S2 (length of onboarding instruction) p-value = 0.0002607 and S3 (helpfulness) p-value = 0.04123

with 5 (Mdn) (see Figure 4.7). For the abstract onboarding conditions, participants rated the statement S2 as 6 (Mdn), which is higher than for the concrete one.

Additionally, the statement S3 regarding helpfulness to better understanding was ranked better in the concrete onboarding condition as in the abstract one. Therefore, the concrete onboarding instructions are more *helpful* than the abstract onboarding instructions for both data sets.

#### 4.4.3 Experience Level

The aim was to see if there is a difference between the level of experience

(H-ExperienceLevel) and the score of each of the four statements. Based on the analysis of the visual exploration of the experience levels (see Figures 4.6), we decided to use three categories not experienced, moderately experienced, and very experienced. We assigned the scores 1 to 3 in the category not experienced, 4 and 5 in moderately experienced, and 6 and 7 in very experienced. A Kruskal Wallis test revealed no significant effect on the four statements, neither on the general experience level nor on the experience level regarding treemap visualizations.

#### 4.4.4 Value

**Results Analysis of Description (Task 2):** As concrete onboarding provides distinct examples and refers to the user data set, we expected participants to generate more valuable descriptions and insights in contrast to participants reading the abstract onboarding instructions (**H-Value**). Based on the categorization of the descriptions along the categories we described in Section 4.4.1 we can report on the following results.

The bar chart in Figure 4.8 shows that participants reading the abstract onboarding instructions submitted more high-qualitative descriptions (11) than in the concrete condition (5). Slightly more intermediate descriptions could be found in the concrete condition (14) than in the abstract (10). Most descriptions were categorized as trivial and high-level in both conditions.

Then, we analyzed the descriptions in more detail towards the content. Based on our onboarding instructions, we derived categories (size, color, interaction) and analyzed if they were also part of the participant's descriptions. In general (both Abstract and Concrete), 53.75% of the participants explained the color of the categories in the treemap, 73.75% the size of the rectangles, 41.25% the interaction concept (hovering), and 21.25% used specific examples and values.

Split up in the abstract and concrete condition, we can report the following: In the *abstract* condition, 53.49% explained the color, 51.72% the size, 57.58% the interaction in the abstract condition, and 47.06% used distinct values in their explanation. In the *concrete* condition, 46.51% of the participants stressed the color of the rectangles, 48.28% the size, 57.58% the interaction possibility of the treemap visualization, and 52% distinct values. Together these results indicate no difference between the particular explanation in the concrete or abstract onboarding instruction.

Interestingly, 21.52% of the participants used distinct values for their explanations overall. Split up in the conditions, the usage of distinct values to explain the treemap is nearly balanced (Abstract 47.06% and Concrete 52.94%).

Analyzing the data on a participant level, we found some interesting insights. Participant 23 does not relate to the data in the abstract condition: "You can see rectangles divided into categories, which are then divided into sub-categories within themselves.". In the concrete condition, the participant indicates to the data — the first sentence is "The treemap is about Biden's tax overhaul, which is divided into two categories, jobs plan and families plan." Furthermore, participants 13, 20, 31, 32, and 37 used a concrete example in the concrete condition to explain the treemap, but not in the abstract condition: "The jobs plan is divided into Community Infrastructure, Workforce Development, Transportation and Eldercare with a budget of 2,3 trillion dollars." (P31). Participant 20 added the USD in the instructions but not the exact value. In contrast, participants 27, 30, and 32 used examples in the abstract condition but not in the concrete one.

**Results Analysis of Insights (Task 3):** In general, participants in both conditions submitted primarily trivial observations, which give a high-level view of the data, see



Figure 4.8: (A) Quality of Descriptions (Task 2), where 1 means trivial descriptions, high-level instructions, 2 intermediate, and 3 highly understandable descriptions. (B) Value of Insights (Task 3), where 1 means trivial observation; 2 intermediate value of insight, and 3 insight that confirms or creates a hypothesis, own interpretation, or integration of domain knowledge. (C) Categorization along Task Taxonomy by Friel et al. (2001) where 1 is reading the data, 2 is reading between the data, and 3 is reading beyond the data.

Figure 4.8. The used data set and resulting treemap are not complex. Therefore, the results are not surprising. However, based on the categorization of the value of the insight, we see no difference in abstract and concrete conditions (numbers are nearly the same), illustrated in Figure 4.8 (B). In the concrete condition (28), eight more intermediate observations could be found than in the abstract (20). Participant 27 wrote: Over half of the money from the Austrian Federal Ministry of Finance goes into the Work, Social Affairs, Health, and Family sector, whereas the most money in that category went into Public Pension Schemes. Slightly more valuable insights could be found in the concrete condition (12 compared to 9). Participants integrated their prior knowledge and generated insights that we ranked as highly valuable insights. For example, "As you can see in Austria planning to invest most of the budget into Work, Social Affairs, Health, and Family. Which shows how important the government thinks it is to spend money for their people." (P38).

Furthermore, we categorized the insights along the task difficulty level by Friel et al. (2001): reading the data, reading between the data, and reading beyond the data. We wanted to quantify the insights along the difficulty level to get a better overview of the general insights the participants generated. We couldn't find any difference between the two conditions. In general, more insights are ranked as "reading the data". In the concrete condition, more insights are categorized as reading between the data (26 to 17), illustrated in Figure 4.8 (C).

**Results Analysis of Insights Correctness:** All the insights of the participants were correct.

#### 4.4.5 Subjective feedback

As part of the study, we also evaluated comments on visualization and onboarding to assess user preferences (H-Preference) of users. We extracted 32 considered usable

Sentiment Analysis									
	$\begin{tabular}{lllllllllllllllllllllllllllllllllll$								
Concrete	16	8	50.00	2	12.50	6	37.50		
Abstract	17	7	41.18	4	35.29	6	23.53		
Sum	33	15	45.45	6	38.71	12	18.18		

Table 4.2: Results of sentiment analysis

comments related to the onboarding instructions. The exact breakdown can be seen in Table 4.2. While reviewing each comment manually, we performed a sentiment analysis using OpenAI API <sup>2</sup>. The comments are categorized as positive, neutral, and negative. More details about the parameters used and the analysis setup can be found in the Appendix B.

**Positive comments:** The results of the analysis (see Table 4.2) did not show a clear difference between abstract and concrete onboarding **(H-Preference)**. 50.00% of the comments concerning the concrete onboarding are rated positive, and 41.18% are related to the abstract onboarding.

Participants highlighted that the abstract onboarding concept in general "[...] was good, and I think that a person who is unfamiliar with this type would easily understand it with those instructions."–(P32 | Abstract A); and another subject commented: "I like the division of the instructions with headings like "Reading the chart" and think it's a very good idea to include instructions when it comes to visualizations in general. I think they can be very helpful, especially if you don't know how to read graphs. It also makes the chart look less intimidating or overwhelming if that makes any sense." – (P3 | Abstract B);

However, if we look at the comments in detail, participants stated: "...I preferred this concrete one to the abstract one. Because there were examples included in the instructions... My understanding was definitely supported more in this prototype..." – (P3 | Concrete A); "The analyzing the chart section was more helpful than in the other example because it provided two actual examples." – (P6 | Concrete A); The only noteworthy aspect here is the change of mind of a user who preferred the abstract onboarding: "In retrospect, I find this treemap much clearer and I also like the description better. I first thought the examples in "Analyzing the chart" would be helpful in the other treemap, but I honestly didn't miss them here." – (P39 | Abstract A).

<sup>&</sup>lt;sup>2</sup>https://openai.com/api/, Accessed: 2023-04-20

**Negative comments:** 37.50% of the feedback in the concrete onboarding condition and 35.29% of the abstract condition was categorized as negative in the sentiment analysis.

Participants stated the following in the concrete onboarding condition: "In my opinion, the descriptions are a bit too long. Especially the second paragraph includes information I would not have needed to understand the visualization. Nevertheless, the information concerning the items is good, because it is easier to understand." – (P2 | Concrete A); Participants commented on the length (P24 | Concrete B) and level of detail of the onboarding message (P3 | Concrete A), as well as perceived the onboarding messages as "bloated and frustrating to read" (P11 | Concrete A).

Furthermore, in the abstract condition, P31 (Abstract A) mentioned that "The messages are in my opinion too concrete and could use a more casual language." while P23 (Abstract A) stated "I had to concentrate a lot to read them, maybe shorter and more general sentences would be easier." which should be the case in abstract onboarding anyway. Contrary some are seeking more information in the abstract condition: "It is not entirely obvious how these rectangles are divided up in terms of size. A more detailed description would be appropriate here." – (P8 | Abstract B); "More Information about reading the chart correctly." – (P22 | Abstract B);

Subjects commented negatively to the section Analyzing the chart of the onboarding instruction. "I don't fully understand the purpose of the last set of messages (Analyzing the chart)." (P6 | Abstract B). While P5 (Abstract B) phrased the following: "...I would also argue that analyzing and reading the chart are more or less the same thing in this context." Besides, one participant didn't like the phrasing of the sentence: Seek out the largest rectangular values (or the largest collected group of rectangles). The subject indicated that it 'sounds more like a command rather than a how to analyze." (P5 | Abstract B).

**Neutral comments:** Some of this suggestive neutral feedback includes suggestions toward the structure and design of the onboarding messages: "Maybe adding subheadings to the paragraphs in the "reading the chart" section could help the reader find relevant information faster." – (P11 | Concrete A); "I would highlight the most important parts of the text before, for example by making some words bold." – (P33 | Concrete B, Abstract A).

Additionally, P3 (Abstract B) suggested: "I'd maybe simplify the instructions a bit more or use examples if it's intended for users without much experience when it comes to understanding visualizations.". Additionally, P15 (Abstract B) suggested that "In the instructions, it should also say that rectangles can be stretched in different dimensions and still represent the same value". The neutral feedback can be seen as suggestions for improvement or additional features.

**Summary:** Based on the sentiment analysis, there is no clear trend toward concrete or abstract onboarding instructions. However, concrete and abstract onboarding instructions

# 4. Abstract and Concrete Visualization Onboarding Instructional Material for a Treemap Visualization

are often described as simple, easy to understand, and helpful. Participants highlighted the helpfulness of the examples in the section "Analyzing the chart", especially in the concrete condition. On the contrary, abstract onboarding is seen as more straightforward and shorter, although the specific examples are often missing, and the messages could be phrased more easily.

## 4.5 Discussion

This section summarizes our findings, design implications for onboarding instructions, limitations, and future work (see Table 4.3).

H-Quality: Concrete onboarding messages are more *helpful* than the abstract onboarding instructions, whereas the length of the abstract messages is preferred over the concrete one. We have to reject the hypothesis **H**-Quality. The statistical analysis results of the statements show that the concrete onboarding messages were ranked more helpful than the abstract ones. However, the length of the onboarding messages was ranked better in the abstract condition. We tested the onboarding messages' readability (ATOS level (Renaissance, 2022)). The ATOs Level is a measure of readability designed to guide students to appropriate-level books. ATOS considers the most important predictors of text complexity—average sentence length, average word length, and word difficulty level. The results are provided in a grade-level, easy-to-use and easy-to-understand scale. The concrete onboarding messages were ranked 11.7 (dataset A) (11th grade), 11.2 (dataset B), and the abstract messages 10.8 (end 10th grade). The complexity of readability is nearly balanced between all the conditions. Therefore, this result may be explained by the fact that we integrated concrete examples in the section "Reading the chart" and this resulted in longer onboarding messages, e.g., "The size of each rectangle represents the outflow of cash and cash equivalents in  $\notin$  Euro (e.g., the public pension scheme rectangle, representing € 12,468.8 million, is approx. twice as large as the Family and youth rectangle, representing  $\notin$  7,687.1 million.)". Therefore, participants might have perceived the abstract onboarding messages as less overwhelming and long.

**H-ExperienceLevel:** The experience level does not influence the score of statements. Our analysis showed that there is no significant effect on the four statements neither on the level of experience for visualization in general nor on the level of experience regarding treemap visualizations. Therefore, we have to reject the hypothesis H-ExperienceLevel.

H-Value: Abstract onboarding messages lead to more valuable descriptions. Based on the categorization of the descriptions, participants wrote more valuable descriptions while reading the abstract onboarding instructions than with the concrete ones. We expected to have more examples/distinct values in the descriptions. However, the usage of distinct values is nearly balanced between the conditions.

H-Value: Both concrete and abstract onboarding messages can lead to highly
valuable insights As in Figure 4.8 illustrated, there is no trend visible towards concrete onboarding instructions leading to higher valuable insights. Therefore, we have to reject this hypothesis (H-Value). Previous research has shown that concrete and abstract learning materials have both strengths and weaknesses (De Bock et al., 2011; Kaminski et al., 2008). Manually generating concrete onboarding instructions is costly and effortful. For example, they must be developed by the visualization designer because concrete onboarding instructions must be tailored to a specific visualization technique and dataset.

In contrast, abstract onboarding instructions can be used in different treemap visualizations with different data sets. Based on our results and the fact of generalization of onboarding instructions to other data sets, we suggest using abstract onboarding instructions along the following structure: Reading the chart, Interacting with the chart, and Analyzing the chart, with highlighting of essential words. As the subjective feedback emphasized the importance of examples/insights, we recommend integrating those in the "Analyzing" section refers to the used data set.

H-Preference: Concrete and abstract onboarding instructions are both rated helpful and simple. There is no clear trend visible. We have to reject this hypothesis as there was no clear trend in the participants' preferences. We expect concrete onboarding to be subjectively preferred over abstract onboarding. The sentiment analysis revealed that both instructions are simple, easy to understand, and helpful. Abstract onboarding is seen as more straightforward and shorter. But the specific examples are often missing. Participants emphasized phrasing the abstract onboarding instructions more easily.

Further research might explore the semi-automatic generation of onboarding instructions for several visualization techniques with the possibility of adjusting and customizing the automatically generated abstract onboarding messages.

#	Hypothesis	Accept/Reject	
H-Quality	We expect participants to rate the concrete onboarding instructions as more	x	
	understandable, helpful, and complete compared to the abstract one.		
H-ExperienceLevel	We expect the participant's experience level to affect the statements.	х	
H-Value	We expect participants to generate more valuable insights and descriptions after reading		
	the concrete onboarding instructions than reading the abstract one.	х	
H-Preference	We expect concrete onboarding is subjectively preferred over abstract onboarding.	x	

Table 4.3: Overview of our hypotheses with indicating four rejected ones.

#### 4.5.1 Phrasing of Onboarding messages

Our results can be used to discuss and inform the phrasing of onboarding instructions. The results of our study reflect the discussion in the field of abstract or concrete materials for teaching (De Bock et al., 2011; Kaminski et al., 2008). The presented findings do not answer whether abstract or concrete instructions are more appropriate for onboarding instructions. According to the data, we can infer that both concrete and abstract

#### 4. Abstract and Concrete Visualization Onboarding Instructional Material for A TREEMAP VISUALIZATION

instructions help us understand visualizations better. The instructions should meet the following points:

- structure the onboarding messages with headlines (e.g., Reading, Interacting, Analyzing the chart) to highlight important parts and lead the reader through the instructions. Headlines can increase the contrast (Williams, 2014) and guide the reader.
- onboarding instructions should be placed **integrated into** the visualization to link the reading instructions to the visualization itself directly.
- highlight important words, e.g., by making them bold or highlighting them with color. In a previous study, we introduced ways to highlight important words and lead the user through the onboarding (Stoiber et al., 2021).
- understandable instructions: easy to understand and short sentences should be considered (Shneiderman, 2022). For guidelines for writing understandable sentences see Janice Ginny (2014).

#### 4.5.2Limitations

Participants and context: The audience and context of the study naturally limit our results. The participants in our study were students in the first and third semester of the bachelor study program Media Technology, in nearly the same age group, with nearly the same pre-knowledge about visualizations and the used data sets. Some of the participants seemed less engaged with the study. Although students could earn extra points for their lectures, some subjects did not invest that much effort in their answers.

Study Material: In previous studies (Stoiber et al., 2021), we presented onboarding concepts, including an interactive interface design to interact with the onboarding instructions. In this study, we focused on the phrasing of onboarding instructions and the question of abstract vs. concrete for a treemap visualization. Further research might explore an adapted version of the onboarding instructions incorporated in an interactive onboarding concept (Stoiber et al., 2021) in different domains such as journalism, medicine, or education using different visualization techniques.

#### Conclusion 4.6

There are discussions about using concrete or abstract material to foster learning. Especially for visualization onboarding, it is an open question of whether to use abstract or concrete onboarding messages. This study aimed to evaluate two types of onboarding messages — concrete and abstract — and investigate which are more appropriate to onboard users to a treemap visualization. The study with 40 participants revealed the following results: (1) Concrete onboarding messages are more helpful than abstract messages, whereas the shorter length of the abstract messages is perceived to be more

94

appropriate. (2) Abstract onboarding messages lead to more valuable descriptions. (3) Both concrete and abstract onboarding messages can lead to highly valuable insights.



## CHAPTER 5

### **Reflection & Lessons Learned**



## Reflection & Lessons Learned

In this chapter, we reflect on the results of the analysis of empirical studies, development of the descriptive design space (see Chapter 2), and the conducted user studies (see Chapter 3, and Chapter 4). Part II will present the resulting visualization onboarding design and the implementation of the VisAhoi JavaScript library.

#### Design Consideration for Visualization Onboarding

We structured our lessons learned along three questions:

- Why is visualization onboarding needed?
- Which parts of a visualization need to be explained?
- How to phrase onboarding instructions?

**Method** We collected all lessons learned from the literature (see Chapter 2) and empirical studies (see Chapter 3, and Chapter 4). We summarized the main objectives along with the three questions described above. In the following paragraphs, we report on the design considerations for visualization onboarding.

Why is visualization onboarding needed? The results of our studies reveal a need for visualization onboarding. In line with the previous work by (Börner et al., 2019, 2016; Galesic and Garcia-Retamero, 2010), we see **poor answer correctness** over different onboarding concepts (cf. Study 2 presented in Section 3.3.1 and Table 3.2). Moreover, we realized that the subjective experience levels of our participants per visualization type do not correlate with high answer correctness. When comparing the subjective assessment of the experience with different visualization techniques to the answer correctness, the analysis revealed discrepancies. We observed that highly experienced participants did not show high answer correctness. In contrast, relatively inexperienced users achieved higher performance measures. Therefore, users should consult the onboarding system more often than they do.

Which parts of a visualization need to be explained? In our comparative user studies (see Chapter 3), we realized that independently of the visualization type and method applied, an easy-to-understand data set and concrete examples on how to read the chart support and increase comprehension are vital. The generalization of the results of the related empirical studies and the design space analysis shows that the visualization onboarding should explain the visual encoding & interaction concepts (Alper et al., 2017; Bishop et al., 2020; Kwon and Lee, 2016).

Furthermore, we investigated in our study (Stoiber et al., 2022b), described in Chapter 4, abstract and concrete onboarding instructions revealing the following results. **Concrete onboarding messages** are **more helpful** than abstract ones. In contrast, the **shorter length** of the abstract messages is perceived to be more appropriate. Furthermore, abstract onboarding messages lead to more valuable descriptions. However, both concrete and abstract onboarding messages can lead to precious insights. In summary, it is essential to consider which visualization types (predominantly unknown and new ones) require instructions to support the user when needed.

How is visualization onboarding provided? In our study, participants prefer the video tutorial to learn how to interact with the parallel coordinates plot (over scrollytelling and step-by-step guide) (see results of study 3). In-situ scrollytelling is highly preferred over a tutorial due to its integration into the tool and gradual guidance through the onboarding instructions. Furthermore, qualitative feedback (see Chapter 3, study 3.5) indicated that the to-the-point descriptions make it easier to absorb information (see study results in study 3.5). Likewise, the understanding of introductions is enabled by interactive and linked descriptions of the visualizations and the described steps (as used for the step-by-step guide and the scrollytelling tutorial).

How to phrase onboarding instructions? Additionally, the results of this concreteabstract onboarding instruction study (Chapter 4) can be used to discuss and inform the phrasing of onboarding instructions. The results reflect the discussion in the field of abstract or concrete materials for teaching (De Bock et al., 2011; Kaminski et al., 2008). According to the data, we can infer that concrete and abstract instructions help to understand visualizations better. The instructions should meet the following points:

- structure the onboarding messages with headlines (e.g., Reading, Interacting, Analyzing the chart) to highlight important parts and lead the reader through the instructions. Headlines can increase the contrast (Williams, 2014) and guide the reader.
- onboarding instructions should be placed **integrated into** the visualization to link the reading instructions to the visualization itself directly.
- highlight important words, e.g., by making them bold or highlighting them with color. In a previous study, we introduced ways to highlight important words and lead the user through the onboarding (Stoiber et al., 2021).
- understandable instructions: easy to understand and short sentences should be considered (Shneiderman, 2022). For guidelines for writing understandable sentences see Janice Ginny (2014).

We aim for visualization onboarding that can be integrated into VA tools and provide a blueprint for developers and visualization design on phrasing easy-to-understand onboarding instructions. Therefore, we improved our visualization onboarding concept iteratively based on the lessons learned. Moreover, we also introduced a JavaScript library to generate onboarding instructions semi-automatically for interactive visualization, which shows generated onboarding messages as tooltips based on the data shown in the visualization (Chapter 7). In the next part II, we provide a detailed description of the resulting onboarding design and underlying design decisions.



### Part II

## Designing and Implementing Visualization Onboarding Concepts



# CHAPTER 6

### Towards a Design for Visualization Onboarding



This chapter describes the design of a visualization onboarding concept, exemplarily integrated into visual analytics (VA) tool for analyzing high-throughput screening (HTS) data (see Section 6.2). Furthermore, we report the results of the cognitive walkthrough inspections with three experts in Section 6.3. The clickable high-fidelity prototype can be accessed here: https://bit.ly/3iTNpRm In Appendix C includes the study design of the cognitive walkthrough inspection.

#### 6.1 Introduction

In part I, we reported on the results of exploring different onboarding approaches, resulting in lessons learned, described in Chapter 5. The analysis and reflection show that there is a need for visualization onboarding concepts that are integrated into VA tools, as recommended by the user and empirical studies (see Chapter 3 and Chapter 4).

In chapter 5, we described how to design visualization onboarding concepts in terms of onboarding type, parts that need to be explained, and the phrasing of onboarding instructions. Along with our derived lessons learned, we further improved our onboarding design by creating a high-fidelity clickable prototype using Sketch <sup>1</sup>. Thus, we present the resulting **visualization onboarding design concept**. We integrated our onboarding concept exemplarily in a VA tool to analyze high-throughput screening (HTS) data. The HTS analysis application holds data from HTS experiments where the binding kinetics of roughly 4 Mio. chemical compounds against potential drug targets are measured. The measurement for each compound is a time series curve that describes the binding behavior of the compound against the target over time. The 4 Mio time series measurements are projected into a two-dimensional latent space with a multidimensional scaling approach. The 2D projection is visualized with a scatterplot. Each dot represents a single compound, and the topology in the 2D space reflects the topology of the time series curves in the high-dimension space. Domain experts search for clusters in these scatterplots with compounds showing promising binding behavior.

These two scatterplots visualize the data side-by-side as illustrated in Figure 6.1. Therefore, we target our onboarding concept to the scatterplot visualizations and users of this VA tool. Furthermore, we also report on the results of a task-based cognitive walkthrough (Salazar, 2022) with experts to improve the design and gain valuable feedback iteratively. Finally, we discuss our findings and present lessons learned.

#### 6.2 Visualization Onboarding Concept

This section presents the major design decisions that led to the final design of the visualization onboarding concept for a scatterplot visualization. The clickable high-fidelity prototype can be accessed here: https://bit.ly/3iTNpRm

The presented onboarding concept is based on the results of the study by Stoiber et al. (2022b,d) and has been actively refined. For onboarding integration in VA tools, we further developed the approach of the step-by-step guide (Stoiber et al., 2022d) based on lessons learned presented in Chapter 5. In the following, we list the design considerations (DC) we derived and explain the design decisions when integrating visualization onboarding in a VA tool:

• **DC1:** Use concrete examples of how to read the chart (cf. Chapter 3 and our published work (Stoiber et al., 2022b))

<sup>1</sup>https://www.sketch.com, Accessed: 2023-01-31

- DC2: Explain the visual encoding and interaction concept (cf. Chapter 3)
- **DC3:** Use in-situ to-the-point onboarding instructions to make it easier to absorb information (cf. Chapter 3 and Chapter 4)
- DC4: Provide interactively and linked onboarding instructions (cf. Chapter 3)

Our previous onboarding concept, the step-by-step guide, and the scrollytelling approach (cf. Chapter 3) are not easily integrable into a VA tool, as it is space-filling, with the onboarding instructions side-by-side with the visualization. Therefore, we surveyed different design alternatives on how to provide the onboarding message more efficiently. On the one hand, integrating it into the main visualization interface does not distract users from exploring the data. On the other hand, it provides context-sensitive and to-the-point instructions. Therefore, in the current design, a floating action button (Material Design, 2022) (Material Design Pattern) is used to access the onboarding instructions (see Figure 6.1). In the following paragraphs, we elaborate on the different parts in more detail.

**Navigation:** In general, we designed the onboarding concept so that users can be either (1) guided through the different onboarding messages or (2) self-explore it and seek help at any stage of their interpretation and interaction with the visualization. We consider the user's learning preference by providing both ways of exploring the onboarding. At any point of exploration, the user can interact with the visualizations. Users are not restricted to step-wise navigate through the onboarding instructions as known from other user onboarding libraries such as aSimpleTour (2020); GuideChip (2022); Hopscotch (2022); intro.js (2020); Orient (2019); Tourguide.js (2023); webTour.js (2022). The onboarding concept provides context-sensitive help, as we found is one main design consideration (DC1) when integrating onboarding to VA tools (cf. Chapter 2 and Chapter 5).

A floating action button (Material Design, 2022) is the main navigation element, as illustrated in Figure 6.1 (1). As default, we positioned it on the right side of the visualization. Thus, we can independently add onboarding instructions to different visualizations (see Figure 6.1). Clicking on the button reveals three circular buttons, representing the three default stages of the onboarding process: *Reading, Interacting, and Analyzing* (Figure 6.1 (2)). The onboarding stage *Reading the Chart* explains the visual encoding (DC2) (e.g., size, color); *Interacting with the Chart* onboarding the user to the interaction concept. Moreover, *Analyzing the Chart* guides the user towards further insights (e.g., making comparisons) (DC1). The step-by-step navigation widget represents the number of in-place annotation anchor/onboarding messages to implement the guided approach (Figure 6.1 (3))(DC4).

Above the onboarding button, illustrated in Figure 6.1 (3), arrow icons allow users to browse the onboarding messages to support the users in receiving guided help. Users who seek help for a specific part of the visualization, e.g., interaction, can self-explore the onboarding messages by opening a stage (Reading, Interacting, or Analyzing) in the



Figure 6.1: Visualization onboarding concept integrated into a scatterplot visualization. The navigation comprises different elements to provide a guided tour and a self-exploratory experience. By clicking on the floating action button (1), the menu expands outward from a central point and reveals three onboarding stages: Reading, Interacting, and Analyzing (2). We decided to integrate a step-by-step navigation widget (3) that represents onboarding messages (4) and arrow icons to navigate through it (3). The onboarding messages are designed as tooltip in combination with in-place annotation anchors (5) to indicate the connection between the textual instructions and the visual encoding.

106

main navigation menu and then clicking on the specific in-place annotation anchor in the visualization.

**Onboarding messages & in-place annotation anchor for anchoring onboarding messages:** Based on our comparative study (Stoiber et al., 2022b) between abstract and concrete onboarding messages for a treemap visualization, we found that both concrete and abstract onboarding messages can lead to precious insights. Concrete onboarding messages are tailored to the respective data set by including concrete facts from the visualizations, e.g., values or insights (DC1). As the data used in this scatterplot is highly complex, we included concrete onboarding messages to provide context-sensitive onboarding instructions. The structure of the onboarding messages is based on our previous work (Stoiber et al., 2021, 2022d).

The design is organized in textual instructions (onboarding messages) in the form of tooltips in combination with in-place annotation anchors (Zhi et al., 2019) (numbers and symbols) to indicate the connection between the text elements and the visual encoding (DC3, DC4). The tooltip (onboarding message) illustrated in Figure 6.1 (4) consists of a headline and textual instruction. Important words are highlighted in color in the instructional text, which we found as an essential (Stoiber et al., 2022b) (cf. Section 5). Inspired by different annotation designs by Lu (2022), we use circular anchors with numbers related to the onboarding message.

Together with a domain expert— the product owner of the VA tool— we developed the onboarding messages. We first provided a draft for the onboarding messages, which the experts further revised and improved. The following paragraphs present the resulting onboarding messages integrated into the clickable prototype. We developed *concrete* onboarding messages referring to the visualized data and examples. The first onboarding instruction in the *Reading* stage explains the data visualized in the scatterplot due to the importance of title elements and legends (Borkin et al., 2016).

• Latent space: The VA tool uses variational autoencoders (VAE), a kind of neural network model, to derive a lower dimensional representation (points in 2D) of high dimensional input data (kinetic curves). This lower dimensional representation called latent space is shown here. During training, the VAE model has been constrained to push data of neutral control samples as close as possible toward the origin of the latent space, thus shaping the latent representation of the data set.

Further, we explain the visual encoding (DC2) (see Chapter 5). Therefore, onboarding messages explaining the x and y-axis (in-place annotations Figure 6.1) and the dots meaning and position (anchors (Munzner, 2014)) are explained in more detail. The onboarding messages for the meaning and position of the dot are phrased as the following:

• **Dot meaning:** Experiments are performed in microtiter plates (MTP). Thus the input data set consists of one **experimental observation** (kinetic curves) for

each well of the MTP. Each dot corresponds to a 2D representation of these higher dimensional kinetic curves. Notice: Each substance well contains two distinct chemical substances at once.

• Dot position: Measurements of neutral control samples (e.g., DMSO) are constant towards the origin of the latent space. Therefore, the **further** a point is **from the origin** the more **dissimilar** its corresponding experimental observation is to a neutral control sample.

In the onboarding stage Interacting with the chart, we explain how to interact with the scatterplot (DC2), e.g., panning, zooming, selecting, and how to download a screenshot. In this example, we explained the pre-defined interaction possibility by Plotly.js (Plotly.js, 2023) for a scatterplot visualization. For example, for the selection interaction, we phrased instructions as the following Drag the mouse over the dots in the scatterplot to select a specific subset of data. Additionally, we included small icons in the tooltips.

Supporting users in analyzing the HTS data in the scatterplot, we provide two concrete examples (DC1) of how to analyze the data and gain insights by explaining the dot's position in the scatterplot. We included these two onboarding messages:

- Closest to origin: This is the dot closest to the origin (0/0). The compound kinetic curve is the most similar to the neutral control.
- Farest to origin: This dot exhibits the greatest distance to the origin. The corresponding experimental observation for this well on the MTP differs strongly from the observation for a neutral control sample. Dots exhibiting a very large distance to the origin could potentially be outliers due to experimental artifacts.

In the next section, we describe the results of the cognitive walkthrough inspection with domain experts.

#### 6.3 Evaluation

The main aim of the evaluation was to gain feedback regarding the usability of the onboarding concept and feedback on the onboarding messages. We also observed how subjects interacted with the in-place annotation anchors and step-wise navigation widgets.

#### 6.3.1 Method & Participants

We conducted a task-based cognitive walkthrough (six tasks) and semi-structured interviews (Lazar et al., 2017) with experts. We recruited three experts between the ages of 32 to 33 (gender: m=3, f=0). One expert from the field of Human-computer Interaction (HCI) and two domain experts — data scientist and laboratory head — for the evaluation. The participants subjectively rated their experience with visualization

108

tools and scatterplot visualizations along a 5-point Likert scale. Table 6.1 present the demographics. Participants generally rated their experience with visualization tools as "average experienced" (3 median (Mdn) on the Likert scale) and with scatterplot visualizations as "experienced" (4 Mdn on the Likert scale). We performed four pilot studies before the cognitive walkthrough to review the study design. We used Zoom to conduct the usability inspection. The test took approximately 25 min.

The inspection was structured as follows: (1) Introduction to the study, (2) collection of demographic data and level of experience with visualization tools and scatterplot visualizations, (3) five minutes test phase to familiarize the participants with the onboarding concept, (4) working through several tasks using the interface by opening the clickable prototype on their own device and sharing their screen, (5) a semi-structured interview (to gain more feedback regarding the onboarding concept itself). We recorded the interviews and demonstrations using Zoom and took notes during the observations. The participants worked independently through the tasks on their laptops, activating screen sharing of Zoom. We asked the participants to think aloud while working through the tasks during the evaluation.

The tasks for the inspection were the following:

Tasks 1: Open the onboarding instructions
Tasks 2: Open the onboarding stage "Reading"
Tasks 3: Read the onboarding messages in the "Reading" stage
Tasks 4: Read the onboarding messages in the "Interacting" stage
Tasks 5: Read the onboarding messages in the "Analyzing" stage
Tasks 6: Close the onboarding instructions

In Appendix C, find the guide to the inspection, including the tasks and the question of the semi-structured interview part.

#### 6.4 Results

Subjects performed the six tasks while using the clickable prototype. Regarding the interaction with the onboarding concept, expert 1 used the in-place annotation anchors and the step-by-step navigation widget. Interestingly, the domain experts (experts 2 and 3) only used in-place annotation anchors to navigate the onboarding messages.

The semi-structured interview results highlighted the need for onboarding to be "highly relevant for users of the VA tool" (expert 2). Subjects also liked the design of the onboarding concept, as it is clear and fits the VA tool. They assessed the onboarding concept as "intuitive and easy" (expert 3) and "logical" (expert 2). Furthermore, the experts categorized the length of the onboarding messages as appropriate. They highlighted the onboarding stages (Reading, interacting, and Analyzing) as suitable and

Table 6.1: **Demographic information of experts**. The reported experience level with visualization tools and scatterplot visualization along a 5-point Likert scale where 1 means "no experience" and 5 "very experienced".

Expert	Gender	Age	Profession	Experience w. visualiza- tion tools	Experience w. scatterplot
1	male	32	HCI expert	not experienced at all (1)	average experience (3)
2	male	33	data scientist	experienced $(4)$	experienced $(4)$
3	male	33	laboratory head	average experi- ence (3)	experienced $(4)$

highly emphasized the "Analyzing the chart" stage. Expert 2 mentioned providing a more detailed description of the underlying data to support users better. One expert commented that the font size in the onboarding message is too small and needs to be adjusted. Expert 1 mentioned increasing the size of circles in the step-by-step navigation widget as it is hard to click them in the clickable prototype. Furthermore, the experts call for the improvement of the step-by-step navigation widget (see Figure 6.1 (3)) to make it more scaleable to a higher number of onboarding messages.

Besides visualization onboarding, they also mentioned integrating onboarding into the interface of the VA tool to support users in this respect.

#### 6.5 Discussion & Lessons Learned

The cognitive walkthrough inspections with three experts showed that the onboarding concept is generally well-designed. We identified some potentials for the improvements of the concept, which we list here: (1) increase the size of the circles in the step-by-step navigation widget, (2) improve the onboarding messages to give a better introduction to the underlying data. (3) make the step-by-step navigation widget more scalable to a higher number of onboarding messages.

The generalizability of the results is subject to certain limitations. For instance, we integrated onboarding instructions in a VA tool analyzing a specific type of data — high-throughput screening data HTS, which is used in biomedical R&D. Therefore, we cannot quickly transform the results one-to-one to another domain, as we invited two domain experts for the cognitive walkthrough and just one HCI expert. However, we found general usability issues, e.g., increasing the size of the circles in the step-by-step navigation widget through the inspection, which needs to be considered.

110

#### 6.6 Conclusion

This chapter presents a visualization onboarding design, exemplarily integrated into the VA tool for analyzing HTS data. We conducted cognitive walkthrough inspections to evaluate the onboarding concepts regarding usability and also gain feedback on the onboarding messages.

We aim to integrate the visualization onboarding concept into VA tools using the underlying data of the visualizations. Available products for user onboarding mainly focus on the overall user interface rather than visual representations, as demonstrated by sources such as AppCues (2020) and Intercom (2020). However, most user onboarding JavaScript libraries such as a SimpleTour (2020); GuideChip (2022); Hopscotch (2022); intro.js (2020); Orient (2019); Tourguide.js (2023); tutoBox (2020); webTour.js (2022) are not fully suitable for visualization onboarding. This is because they are limited to guided tours that explain elements in the user interface, which can block the whole interface and prevent users from interacting with the visualization while seeking help. This is especially problematic when exploring visualizations because users should be able to interact with both the visualization and the onboarding simultaneously. Additionally, the existing onboarding libraries and solutions do not utilize the underlying data of the visualization, which is a critical factor in introducing unfamiliar visualizations to users.

Hence, there is a need for a JavaScript library to integrate visualization onboarding into VA tools that utilize the visualization's underlying data to explain the visualization and not block the whole interface while using the onboarding instructions. Therefore, we implement a JavaScript library called VisAhoi presented in the next chapter 7. The library covers the visualization onboarding design presented in this chapter. The in-place annotations allow users to see onboarding instructions based on the underlying data of the visualization. The onboarding message introduces specific visualization parts, e.g., axis, visual mapping, and legends, which are directly integrated into the interface, reducing the need to switch back and forth between the visualization and a separate tutorial or help section. The floating action button gives users a persistent and accessible way to access instructions throughout their interaction with the visualization.

Developing onboarding messages is a vital step in the design of visualization onboarding. Using a semi-automatic approach, the VisAhoi library leverage existing data and metadata about the visualization to generate onboarding messages and support developers and visualization designers with a blueprint that can be further customized.

In summary, the library provides a solution for visualization designers and developers to semi-automatically generate onboarding instructions based on the underlying data with textual descriptions (onboarding messages) and in-place annotation anchors.



## CHAPTER 7

## VisAhoi: Towards a Semi-Automated Generation of Visualization Onboarding Using High-level Visualization Grammars



This chapter presents the implementation of the VisAhoi JavaScript library, which can be accessed here: https://datavisyn.github.io/visAhoi/. In detail, we present the results of a systematic analysis of onboarding libraries (see Table 7.1), objectives towards a visualization onboarding library in Section 7.3, and the design and functionality of the VisAhoi library in Section 7.6. Furthermore, we discuss the library's limitations (in Section 7.7.1).

#### 7.1 Introduction & Motivation

Concerning available products for user onboarding, solutions mostly focus on the overall user interface rather than the visual representations in particular (e.g., AppCues (2020) or Intercom (2020)). Most user onboarding JavaScript libraries (aSimpleTour, 2020; GuideChip, 2022; Hopscotch, 2022; intro.js, 2020; Orient, 2019; Tourguide.js, 2023; tutoBox, 2020; webTour.js, 2022) are not feasible for visualization onboarding as they are limited to guided tours explaining elements in the user interface, which block the whole interface. Especially when exploring visualizations, it is crucial to allow users to interact with the visualization and onboarding simultaneously, not to block the entire interface when seeking help. Moreover, the current onboarding libraries and solutions do not leverage the underlying data of the visualization, which is an essential aspect of introducing unfamiliar visualizations to users. While existing learning environments and web platforms such as "The graphic continuum" (Swabish and Ribecca, 2014) or the "Data Viz Catalogue" (Ribecca, 2020) focus on the explanation of visualization techniques, the presented textual descriptions and illustrations are mainly manually developed by the visualization designer of the respective platform. This requires much effort from the visualization designer and developer. These onboarding concepts are generally tailored to specific visualization techniques and used datasets or are restricted as they are not reusable or integrated into existing visualization tools.

Therefore, we aim to create a versatile and reusable solution for visualization designers and developers to generate visualization onboarding instructions semi-automatically with textual descriptions (onboarding messages) and in-place annotation anchors (see Figure 7.3 and 7.4) that allow the users to interact with the visualization while interacting with the onboarding instructions. We aim to allow the user to self-explore the onboarding instructions and receive context-sensitive help at any point of their exploration with the VA tool. Furthermore, our onboarding solution is designed to be reusable and can be integrated into custom visualizations. The goal is not to create static visualization tutorials but to enable the developers to integrate the onboarding seamlessly into the productively used visualizations. The end users can switch on the onboarding as they need it directly in their currently used visualization with the actual loaded data. To achieve this, we leverage high-level declarative visualization grammars, which aim to simplify programming tasks by specifying "what" visualization results should be rather than "how" they should be computed (Satyanarayan et al., 2016). They can simplify visualization specifications while supporting a high degree of expressiveness and customization. We use the power of high-level declarative visualization grammars to extract and describe visualization onboarding instructions.

We propose a JavaScript library called *VisAhoi* and build upon three of the mostused (NPM Trends, 2020) visualization libraries with high-level declarative grammars— Plotly.js (Plotly.js, 2023), Vega-Lite (Satyanarayan et al., 2017), and Apache ECharts (Li et al., 2018)—to explore how we can semi-automatically generate onboarding instructions based on the currently shown dataset. We show our approach for a bar chart, a change matrix (Niederer et al., 2017), a horizon graph (Heer et al., 2009), a scatterplot (Friendly

114

and Denis, 2005), and treemap (Shneiderman, 1992) (see Figure 7.3). We decided to demonstrate our approach using five visualization types with diverging levels of familiarity (Borkin et al., 2013; Börner et al., 2016; Boy et al., 2014; Galesic and Garcia-Retamero, 2010; Grammel et al., 2010; Perkhofer et al., 2019) to cover visualization literacy at different levels. We also demonstrate our approach for visualizations with different visual attributes (Mackinlay, 1986) (e.g., xPos, yPos, area).

#### 7.2 Method

We present the results of a **systematic analysis** of onboarding libraries (see Table 7.1) and derived **objectives** resulting from the analysis in Section 7.3. The main part is the **design and architecture** of the VisAhoi JavaScript library following a rapid prototyping approach (SedImair et al., 2012). We discuss extracting information regarding the visual encoding of the high-level declarative description, the JavaScript runtime object, and DOM node from Vega-Lite, Plotly.js, and ECharts to exemplarily generate onboarding instructions based on textual descriptions combined with in-place annotation anchors based on the currently shown dataset. We present our approach for different visualization techniques in Figure 7.3 and Figure 7.4: a bar chart, a change matrix (Niederer et al., 2017), a horizon graph (Heer et al., 2009), a scatterplot (Friendly and Denis, 2005), and treemap (Shneiderman, 1992). Finally, we discuss the limitations of our library and the results.

In Part III of this thesis, we present two design studies showing the applicability of the VisAhoi library in two different domains. We integrated the VisAhoi library into a VA tool for biomedical experts, presented in Chapter 9, and into the visualization tools Flourish (2022) to show the applicability in a data-journalistic use case (Chapter 8).

#### 7.3 Objectives

We defined five objectives a visualization onboarding library should meet from a developer's perspective. The derived objectives resulted from a systematic analysis of the literature on visualization onboarding in general and existing JavaScript libraries and tools for onboarding (see Table 7.1). Additionally, our experience from a recent research project and published work (Stoiber et al., 2019a) lead to the following objectives:

- 1. **Integrable.** Adding the onboarding library to existing charts and integrating it into an existing application plays an important role in gaining fast adoption and making visualization onboarding available for potential users of a VA tool.
- 2. Extensible. The onboarding library should be extensible to support multiple declarative visualization frameworks to provide broad coverage.
- 3. **Reusable.** The onboarding library should provide reusable onboarding messages for different datasets for the same visualization type (e.g., horizon graph). Visualization

designers/developers should be able to reuse onboarding messages across various visualization projects.

- 4. Semi-Automated. The template for the onboarding message should be provided with the onboarding library to semi-automatically generate onboarding instructions based on the given visualization type and dataset.
- 5. **Customizable.** The developer should be able to customize the appearance of the onboarding messages as well as the annotation style.

This version of VisAhoi supports all five objectives. We discuss the limitations regarding these objectives in Section 7.7.

Our work relates to (1) information extraction form visualization, (2) visualization grammar, and (3) user onboarding tools and libraries, which we present in the next section.

#### 7.4 Related Work

#### 7.4.1 Information Extraction from Visualizations

Visualization onboarding requires information about the visualization to create, for instance, tailored messages or set the position of the annotations. A body of previous work describes approaches to extract information from visualizations as bitmap images to reverse engineer them (Poco and Heer, 2017), create re-usable visualization templates that can be used with different datasets (Harper and Agrawala, 2018), or create textual summaries for visually impaired people (Obeid and Hoque, 2020). Most techniques employ machine learning algorithms to extract structural information (e.g., axes, marks), visual encoding, and labels from the visualization.

In contrast, Harper and Agrawala (2014) presents an approach to deconstruct D3 charts to extract their underlying structure, i.e., data, marks, and the mapping between them. Similar to our approach, they focus on SVG-based visualizations but extract the information directly from the DOM structure and data attached to the DOM nodes. This method leads to limitations when, for instance, deconstructing non-linear functional mappings, e.g., log scale in a scatterplot. In *VisAhoi*, we employ the declarative description, the JavaScript runtime object, and the DOM structure. As a result, we have access to internal information and functions of the visualization library, which can be used for tailored onboarding messages.

#### 7.4.2 Visualization Grammar

One of the first examples of a grammar-based approach to visualization was Wilkinson's *The Grammar of Graphics* (Wilkinson, 1999), an abstraction that makes thinking, reasoning, and communicating about graphics easier. Wickham's popular *ggplot2* (Wickham,

2010) and qqvis (Wickham, 2018) packages implement variants of Wilkinson's model in the R statistical language. Furthermore, Shih et al. (2019) developed a declarative grammar to specify advanced volume visualizations effectively. ggplot2 and ggvis are for specific types of visualizations. In contrast, languages such as *Protovis* (Bostock and Heer, 2009) or Vega (Satyanarayan et al., 2016) provide basic building blocks for a wide variety of visualization designs: data loading and transformation, scales, map projections, axes, legends, and graphical marks such as rectangles, lines, plotting symbols, etc. Interaction techniques can be specified using reactive signals that dynamically modify a visualization in response to input event streams. Veqa-Lite (Satyanarayan et al., 2017) expands the basic interaction features of Vega and provides a high-level specification of interactive data visualizations. It enables specifying visualizations in terms of marks, layout, and data. Encodeable (Wongsuphasawat, 2020) is a new configurable grammar independent from rendering where a component author can declare a grammar for the encoding channel of his/her component, which looks like a subset of the Vega-Lite grammar. Li et al. (2018) introduced *ECharts* as a web-based framework for the rapid construction of cross-platform visualizations. The ECharts library employs an all-in-one JSON format option to declare the components, styles, data, and interactions, resulting in a logicless and stateless mode. *Plotly js* is a high-level, declarative charting library that is highly used and covers 40 visualization types, including 3D charts, statistical graphs, and SVG maps (Plotly.js, 2023).

To validate the applicability of our approach, we selected the most-used (NPM Trends, 2020) visualization libraries Vega-Lite, ECharts, and Plotly.js to prototypically demonstrate the extraction of information from declarative visualization specifications and generate onboarding instructions in an automated manner. We decided to use those three visualization grammars due to the following characteristics: (1) all libraries are **high-level declarative visualization grammars**, (2) we explicitly target visualization libraries with **SVG** rendering, as it is important for employing our onboarding annotations in the DOM structure, and (3) all libraries are **open-source**.

#### 7.4.3 User Onboarding Tools and Libraries

Several commercial software tools are available for integrating onboarding elements into user interfaces. For instance, AppCues (2020), a tutorial-based onboarding service, can provide contextual help. Besides web applications, there are some frameworks and JavaScript libraries available that can be used to onboard the user. As seen from Table 7.1, most libraries provide guided tours to introduce the user to an application. Hopscotch (2022) is a powerful framework for building a simple welcome tour (illustrated in Figure 7.1 (C)) using an exchangeable JSON format system. Other libraries for guided tours are aSimpleTour (2020), intro.js (2020), Tourguide.js (2023), webTour.js (2022), Sequence (2022), Shepherd.js (2022), Orient (2019), GuideChip (2022), tutoBox (2020), and driver.js (2020), which allow the integration of animated guided tours. All of those libraries integrate buttons and/or step-by-step widget interaction elements to provide a guided experience for the users. Therefore, supporting the self-exploratory approach we propose

iv Perfect Product Experience reat ve guided product tours in minutes with GuideChimp the most non-technical friendly, lightweight and extendable library.	B Chardin.js
	The busine tagging the vertices of the transmission of the transmi
<pre>characterization (http://http:/h</pre>	PRACTICE An annotations are made of just three parts, a note, a connector, and a subject.

Figure 7.1: Examples of user and visualization onboarding JavaScript libraries (A) GuideChip (2022) providing guided tours with step-wise instructions. (B) chardin.js (2022) uses static labels to onboard different user interface elements. (C) Hopscotch (2022) uses guided tours with buttons to navigate the steps. And (D) the d3-annotation library by (Lu, 2022) integrates labels that can be positioned in the visualization itself.

with our VisAhoi library is impossible. Other available libraries, like chardin.js (2022), provide visually appealing labels but no interactive elements as shown in Figure 7.1 (B). The presented libraries are feasible for onboarding different web applications but are limited to guided tours and static labels. Therefore, unusable for our proposed visualization onboarding concepts with in-place annotations anchors and onboarding instructions as tooltips.

Additionally, most onboarding libraries do not allow users to interact with the underlying interface while reading the onboarding instructions. Especially when it comes to exploring visualizations, it is crucial to allow users to interact with the visualization and onboarding simultaneously, not to block the whole interface. VisAhoi allows users to open the onboarding instructions at any point (self-exploratory). The presented JavaScript libraries restrict the user from navigating step-wise (guide) through the onboarding instructions. With GuideChip (2022), developers can extend the basic library with a plugin to also provide animated beacons (see Figure 7.2) in the interface to open the onboarding instructions in a self-exploratory way. Moreover, after closing the onboarding, users have to start from scratch. So, the onboarding libraries don't provide context-sensitive help (cf. Chapter 2), as the user has to navigate through all the instructions to find the most related one to their problem.



Figure 7.2: Self-exploratory animated beacons provided by the library GuideChip (2022).

However, there are also more visualization-related libraries and approaches to create annotations (see Table 7.1). d3-annotation (Lu, 2022) is a library closer to visualization, illustrated in Figure 7.1 (D). It can onboard users to visualizations developed in D3.js (D3 Language, 2023). Additionally, the libraries swoopyDrag.js (2020) and labella.js (2022) can create data-driven annotation and labeling for visualizations. These libraries are limited as they can create static labels in the visualization but cannot generate interactive onboarding instructions which provide richer textual explanations, as we propose in this thesis. Additionally, swoopyDrag is (2020) and labella is (2022) are no longer updated or maintained. These annotations libraries could be integrated into our proposed VisAhoi Library as an extension (see Figure 7.5). Besides annotations via programming, there are codeless tools such as ChartAccent (Ren et al., 2017) and VisAnnotator (Lai et al., 2020). ChartAccent (Ren et al., 2017) is a tool that allows users to quickly and easily augment charts via a palette of annotation interactions that generate manual- and data-driven annotations. More recently, Lai et al. (2020) developed Vis-Annotator, an automatic annotating technique for promoting efficient data presentation with visualizations based on Machine Learning and a Natural Language Interface. The presenter can upload a visualization image with the corresponding textual description and get a series of annotated animations.

In summary, several products and services support developers in integrating onboarding elements. Approaches to automatically create annotations for visualization in the field of data stories are present in the literature (Hullman et al., 2013; Lai et al., 2020). However, to our knowledge, none is mainly aimed at onboarding for data visualization, and the semi-automated creation of onboarding messages and in-place annotation anchors cannot be achieved.

Table 7.1: This table summarizes the **related work in terms of visualization onboarding libraries** (from 2019-2023). It highlights the field of application (general purpose/visualization-specific) and the onboarding type the respective library supports. We present objectives in Section 7.3, which our proposed VisAhoi library should meet. In column *Objectives*, we present the results of the categorization of the related work to our objectives. X = meets the objective, (X) = partially meets the objective, and blank cell = does not meet an objective. We discuss the limitations of the current version of VisAhoi in Section 7.7.1.

		On	boarding t	ypes			Objectives	5	
Library	Context	Guide Tour	Tooltip	Label	Integratable	Extensible	Reusable	Semi-Autor	nated Customitab
aSimpleTour (2020)	general purpose	x			х				x
intro.js (2020)	general purpose	х			x				x
driver.js (2020)	general purpose	x			x				x
Tourguide.js (2023)	general purpose	x			x				x
webTour.js (2022)	general purpose	х			x				х
Sequence (2022)	general purpose	х			x				х
Shepherd.js (2022)	general purpose	x			x				x
Orient (2019)	general purpose	х			x		x		x
GuideChip (2022)	general purpose	х			x	х			x
tutoBox (2020)	general purpose	x			x				x
Hopscotch (2022)	general purpose	х			x				x
chardin.js (2022)	general purpose			x	x				x
smartour (2019)	general purpose		x		x				x
d3-annotation (Lu, 2022)	visualization specific			x	x				x
VisAhoi	visualization specific		x		x	x	x	x	x

#### 7.5 Visualization Onboarding Concept

This section describes the central design aspect of the onboarding concept exemplarily with five different visualization techniques: a bar chart, horizon graph, change matrix, scatterplot, and treemap. We iteratively improved the onboarding design resulting in this VisAhoi.js library. We identified and described design considerations in Chapter 5, which we integrated into the visualization onboarding concept framework proposed and presented in Chapter 6. This section shows screenshots of all five visualization techniques using the high-level declarative grammars—Plotly.js (Plotly.js, 2023) (see Figure 7.4 and Figure 7.3. The screenshots for the declarative grammar Vega-Lite (Satyanarayan et al., 2017) and ECharts (Li et al., 2018) can be accessed in Appendix D. The demo page includes more examples and can be accessed at https://datavisyn.github.io/visAhoi/demos/.

As explained in detail in Chapter 6, the main design decisions for an onboarding design that allow developers and designers to integrate into a VA tool or visualization tool are as the following (see Figure 7.3 and Figure 7.4): (1) a **floating action button** as the primary navigation element to access onboarding messages; (2) **tooltips and in-place annotation anchors** to indicate the connection between the text elements and the visual encoding (guided vs. self-exploratory); (3) **concrete onboarding messages** based on the underlying data of the visualization, e.g., for the horizon graph in the reading stage, default messages look like *The horizon graph shows the Average temperature in Oslo*, *Norway in 2018.* and *Light green areas indicate a moderate positive average temperature in* °C and dark green areas a high positive average temperature in °C.; (4) **interact simultaneously** with the underlying **visualization** and the **onboarding**.

For the visualizations illustrated in Figure 7.4 and Figure 7.3, we used different data sets: the average temperature (ECA&D, 2019) in Oslo, Norway, in 2018 for the horizon graph and bar chart. For the treemaps visualization, we used the Biden Jobs plan data (Probasco, 2021), and for the scatterplot visualization, the car data set (Heinrich and Broeksema, 2023).

In the next section, we describe the design of the VisAhoi.js library.

#### 7. VISAHOI: DESIGN AND IMPLEMENTATION OF A SOFTWARE LIBRARY









Figure 7.3: VisAhoi.js integrated into a bar chart, change matrix, scatterplot, and treemap visualization using Plotly.js

122



Figure 7.4: VisAhoi.js integrated into a horizon graph using Plotly.js: (A) onboarding activated, showing the floating action button and four different onboarding stages. (B) in-place annotation anchors are positioned in the visualization. The concrete onboarding message explaining the color encoding is open. The italic text elements are semi-automatically extracted from the Plotly.js declarative grammar. (C) Onboarding button with the step-by-step widget to guide users through the onboarding messages.



Figure 7.5: Workflow for visualization developers: Based on a visualization specification (e.g., Plotly.js (Plotly.js, 2023), Vega-Lite (Satyanarayan et al., 2017), etc.) and the definition of a visualization type (e.g., bar chart, change matrix, etc.), the VisAhoi Library semi-automatically generates in-place annotation anchors and onboarding messages in the form of tooltips.

#### 7.6 Design of the VisAhoi Library

In this section, we describe the design decisions of the development of VisAhoi, motivated by our objectives (see Section 7.3).

VisAhoi, in general, supports declarative visualization libraries, which create the visualization based on a specification (e.g., JSON) that describes the visual encoding, the interactive behavior, and the data of the visualization. Those libraries transform the input into a library-specific runtime object and render the visualization in the provided DOM node (see Figure 7.5).

VisAhoi is developed using the JavaScript framework Svelte (2023). We chose Svelte, as it compiles plain JavaScript and allows developers to integrate VisAhoi in applications with different frontend frameworks. VisAhoi currently supports three popular visualization libraries: Plotly.js (Plotly.js, 2023), Vega-Lite (Satyanarayan et al., 2017), and Apache ECharts (Li et al., 2018). The VisAhoi repository (https://datavisyn.github .io/visAhoi/) contains the adaptors for each supported visualization library and the VisAhoi core, which contains functionality that can be shared independently of the visualization library. An additional demo package is included to demonstrate the VisAhoi integration for different visualization types for each supported visualization library (https://datavisyn.github.io/visAhoi/demos/).

The current implementation of our VisAhoi library supports the following visualization types: a bar chart, a change matrix (Niederer et al., 2017), a horizon graph (Heer et al., 2009), a scatterplot (Friendly and Denis, 2005), and treemap (Shneiderman, 1992). A generic type is available so that developers are not limited to the supported visualization techniques of the library. The generic type initializes the onboarding with an empty set of onboarding messages, which the developers can then define.

To initialize VisAhoi, developers only need to fetch the runtime object of the created visualization and pass it on to VisAhoi when calling the ahoi () function, together with the



Figure 7.6: The VisAhoi workflow: (1) a visualization generated by a visualization library; the library renders a visualization from a given dataset, visualization specification, and DOM node and provides this information as the runtime object, which is passed to the VisAhoi adapter in (2). (2) The VisAhoi adapter specific to the visualization library extracts information, e.g., values and anchors, from the visualization specification and runtime object, and provides initial onboarding messages that can optionally be customized, (3) the VisAhoi core generates onboarding messages; The values are applied to the onboarding messages. Due to missing values or anchors, incomplete messages are filtered out, (4) the onboarding can be further customized in terms of visual appearance, and (5) VisAhoi renders the onboarding interface with in-place annotation anchors at the given position in the visualization.

visualization type. The ahoi() function provides a callback function (customizeOnboardingMessages()) as a parameter, which provides the default onboarding messages created by VisAhoi for the specific visualization type. Developers can either rely on those, customize them, or add new onboarding messages (see Section 7.6.4). Every onboarding message needs to be assigned to an onboarding stage, which is basically the group where all related onboarding messages will be rendered.

All onboarding messages will then be added to the state of VisAhoi and, finally, rendered in the specified onboarding stage in the onboarding user interface (UI) created by VisAhoi.

In the following, we explain the VisAhoi architecture in more detail.

#### 7.6.1 VisAhoi Architecture

The library consists of two main components: the visualization-library-specific adapter and the core, illustrated in Figure 7.6 (2) and (3).

**VisAhoi adapter:** The VisAhoi adapter is visualization-library specific, which means we provide one for Plotly.js, Apache ECharts, and Vega-Lite, each. It considers the DOM architecture of the visualization and the specification object, which differ for each visualization-specific library. Depending on which library developers use, they need to import the respective package. When developers initialize the onboarding by calling ahoi (), showing in Listing 1, the function in the respective package is called, to which the visualization type and the runtime object need to be passed. Additionally, a configuration object and a function to customize the onboarding messages can be passed if the developer needs further configuration or customization (see Section 7.6.4).

```
import { ahoi, EVisualizationType } from '@visahoi/plotly';
1
    // create plotly visualization
2
    const runtimeObject = await new Plotly.newPlot(element, traces, layout);
3
4
    // initialize visAhoi
    const visAhoi = ahoi({
5
     visType: EVisualizationType.SCATTERPLOT,
6
      chart: runtimeObject,
7
     ahoiConfig: {}, // additional configuration
8
      customizeOnboardingMessages: (defaultOnboardingMessages: IOnboardingMessage[],
9
         contextKey: string) =>
        customizeHexbinplotMessages(defaultOnboardingMessages, contextKey,
10
           runtimeObject, numberOfDatapoints),
11
    });
```

Listing 1: Each **visualization library adapter** exports an ahoi() function that takes care of extracting the information from the runtime object for the supported visualization types and generates the onboarding. This listing shows an example for initializing visAhoi for a Plotly.js visualization.

The adapter provides functionality for each supported visualization type. It creates the default onboarding messages by considering the visualization type specified by the developer and extracting the necessary information from the runtime object. The runtime object contains the DOM element, the chart specification, and the data contained in the visualization, which is used to get the necessary information for individual in-place annotations.

Each visualization type-specific function across all visualization library-specific extractors returns an object, the onboarding specification, with the same structure, which will later be passed on to our core package. The onboarding specification contains information about how to extract specific information from various visualization elements and where to place the respective anchor. Listing 2 shows an example of extracting information for the visualization title and the minimum value along the y-axis.

In-place annotation anchors can be positioned in two ways. The information is stored in the anchor attribute, which you can see in Listing 2 in lines 3, 11, and 21. On the one hand, one can pass coordinates, meaning the anchor will be absolutely positioned in the visualization (see Listing 2, yMin). On the other hand, the anchor can be attached to a DOM node inside the visualization. To attach the anchor to a DOM node, we again have implemented two different approaches. Either VisAhoi looks for the DOM node by the given content, e.g., the text in the visualization title (see Listing 2, chartTitle). To do that, the flag findDomNodeByValue in the anchor object needs to be set. Or, the selector is defined, which VisAhoi will then use to find the corresponding DOM node in the visualization (see Listing 2, yAxis, line 21). We generally recommend the

```
1
    chart title: {
\mathbf{2}
         value: runtimeObject?.layout?.title?.text,
3
         anchor: {
4
           findDomNodeByValue: true,
5
           offset: { left: -20, top: 10 },
6
         },
\overline{7}
    },
8
9
    yMin: {
10
       value: min.toFixed(2),
11
       anchor: {
12
         coords: {
           x: xGrids[minIndex]?.getBoundingClientRect().x,
13
           y: traceNodes[2].childNodes[0].getBoundingClientRect().y,
14
15
         },
16
       },
17
    },
18
    yAxis: {
19
       value: chart.layout.yaxis.title.text,
       anchor: {
20
21
         sel: ".infolayer .ytitle",
22
       },
23
    },
```

Listing 2: This listing shows the **positioning of the in-place annotation anchors** exemplary for the chart title and y-axis.

latter approach, as we made the experience that DOM elements inside the visualization can have the same text content in some cases, which would lead to wrongly positioned anchors. The created onboarding specification for the current visualization is then passed on to the core package, where the onboarding messages will be created.

**Core:** Our core package includes functionality that is shared across different visualization libraries. When the onboarding specification is created in the corresponding adapter, it will be passed to the core, where the onboarding messages will be created (see Figure 7.6). Onboarding messages will be the same across different visualization libraries. In case the adapter cannot correctly extract all the necessary information from the runtime object, the onboarding messages cannot be displayed. The core method generateMessages () receives the runtime object, the onboarding specification, and the visualization type. Depending on the visualization type, onboarding messages will be created appropriately. To display a specific message, specific information about the visualization is needed. Depending on whether VisAhoi has this information or not, the onboarding message will be rendered in the end. For example, finding the visualization title inside the visualization is necessary to explain the visualization title. If the adapter is not able to extract this information from the runtime object, the anchor and onboarding message for the visualization title will not be rendered. The core package defines which information is needed for each attribute of the visualization. Listing 3 shows an example of this object for the visualization title. The attribute requires in line 3 defines which properties are needed from the onboarding specification, which is passed from the

```
1
    {
2
    anchor: getAnchor(spec.chartTitle, visElement),
3
         requires: ['chartTitle'],
         text: `The horizon graph shows the <i>${spec.chartTitle?.value}</i>.`,
4
         title: 'Reading the chart',
5
6
         onboardingStage: reading,
7
         marker: {
           id: `visahoi-marker-${contextKey}-8`
8
9
         },
10
         id: `visahoi-message-${contextKey}-8`,
         order: 1
11
12
       }
```

Listing 3: The **core package** of VisAhoi defines requirements for the default onboarding messages. In this case, to show the onboarding message for the visualization title, the attribute chartTitle needs to be given in the onboarding specification that the core gets passed from the adapter (see Listing 2).

adapter. In this case, the attribute charttitle , defined in the adapter (see listings 2 and 3), is needed to display the onboarding message in the UI.

Eventually, the generateMessages() function returns an object containing all possible onboarding messages for the specific visualization type. Those messages will be passed back to the developer through the callback function customizeOnboardingMessages() mentioned above, where they can be filtered, updated, extended, and customized.

#### 7.6.2 Internal State of VisAhoi

As VisAhoi is developed using the frontend framework Svelte, it also uses a Svelte state internally to store all the information for onboarding. It is necessary to have a state per visualization because everything related to the onboarding is stored in this state, starting from the chart-specific onboarding messages or stages and any custom configuration. We solved this by having a global state that holds all states of all visualizations as a key-value pair. The key (called contextKey ) is the unique identifier that is used to find the associated onboarding for a visualization. Developers can either pass their own contextKey when initializing the onboarding by passing it to the ahoi() function. In case no contextKey is passed, VisAhoi will create one and return it back to the developer. Whenever customizing or retrieving the onboarding of a visualization, this key needs to be passed to the respective VisAhoi function. Having this key, VisAhoi will then store the update to the corresponding internal state of the visualization.

#### 7.6.3 Rendering of VisAhoi UI elements

Onboarding stages, onboarding messages, the in-place annotation anchor, the tooltip, and the onboarding navigation are part of the core package of VisAhoi and can be customized by the developer. When the adapter extracts all necessary information,
it passes the onboarding specification to the core package, generating the onboarding messages. Those messages are the basis and will be used to render the onboarding in the DOM finally. Therefore, VisAhoi creates an SVG element with the same width and height as the visualization. This information can be extracted from the passed runtime object. The created SVG element will be positioned exactly on top of the visualization so that the onboarding, which will be rendered inside, exactly overlaps the visualization. This way, the following described UI elements of the VisAhoi interface are rendered correctly. To keep the possibility to interact with the original visualization as the SVG of the onboarding is rendered right above, we added pointer-events: none. This way, users can interact with the visualization even when the onboarding is opened.

#### 7.6.4 Customization

VisAhoi allows a variety of customization options for the developer. It allows adapting the general UI of the onboarding, including icons and colors, to make it easy to integrate onboarding into a given visual analytics tool.

**Onboarding stage:** By default, VisAhoi provides three onboarding stages: Reading, Analyzing, and Interacting with the chart. The developers can adjust onboarding stages by changing the color, the icon, and the stage title or even deleting or adding new onboarding stages.

**Onboarding message:** The onboarding message consists of a title, text, an anchor (=position where it should be placed in the visualization), an in-place annotation (=annotation in the chart), an order (order within an onboarding stage), a tooltip position (should the tooltip be rendered on top/bottom/left/right of the anchor), and required information to display the message. Additionally, an onboarding message needs to have a corresponding onboarding stage. The developer can customize onboarding messages: Changing the onboarding instructions (text), order, and corresponding onboarding stage.

**Onboarding navigation:** Developers can decide if users should be able to show the stepper navigation in the onboarding navigation. The position of the onboarding navigation can also be customized: it can be unfolded upwards or to the left.

**In-place annotation anchor:** In-place annotations are attached to visualization elements to highlight that onboarding is available. In-place annotation anchors can have text, e.g., numbers, guiding the user through onboarding. The developer can decide if the numbers are needed.

**Tooltip:** Tooltips are attached to in-place annotation anchors and reveal the onboarding information when clicking on the anchor. The tooltip includes a title and the onboarding message, which the developer can adjust. Furthermore, the developer can highlight important words in the onboarding message by formatting it using HTML.

All these configuration options can be passed to the ahoiConfig object in the ahoi() function when initializing the onboarding for a visualization.

#### 7.6.5 Inline-edit mode

VisAhoi provides an edit mode for users of visualizations where onboarding is available. This feature was specifically implemented to be used for the data-journalistic use case with Flourish (see Chapter 8). VisAhoi provides a function setEditMode() that enables developers to enable or disable the edit mode. When enabled, the user interface of the onboarding of VisAhoi changes slightly. Users, e.g., of an authoring tool, can adjust the onboarding, for example, edit or delete the onboarding message, edit the tooltip title, or delete an onboarding stage inline. It is not recommended to generally enable this feature for all use cases as users could edit the onboarding of visualizations they are not familiar with.

**Summary:** VisAhoi generates onboarding messages semi-automatically by extracting information, e.g., values, anchors from the visualization specification and runtime object, and provides default onboarding messages that the developer can customize. The onboarding can be further customized by changing onboarding messages, onboarding stages, titles, icons, and colors of the onboarding concept. The library allows developers to integrate onboarding instructions into VA tools, extend it as it uses declarative visualization frameworks, reuse onboarding messages across visualization projects, semi-automate the generation of onboarding messages, and highly customize it.

## 7.7 Discussion & Future Work

In this section, we discuss how VisAhoi can support the objectives described in Section 7.3 and present limitations of the VisAhoi library.

VisAhoi is a Typescript library designed with a developer's experience in mind. Adding semi-automatically generated onboarding to visualizations can be done with a few lines of code (see Listing 1) (Objective Integratable). The basic onboarding messages, semi-automatically (Objective Semi-automated) generated by VisAhoi, can be used as a blueprint to extend and further adjust towards users' needs (see Section 7.5).

The customization possibility (Objective Customizable) of VisAhoi allows developers and visualization designers to flexibly adjust the onboarding concept to meet the user's needs of the VA tool. Developers can adjust onboarding messages and stages, for example.

The inline-edit mode, which the VisAhoi library provides, allowed us to develop a visual interface for an authoring tool. We used this functionality in our design study to develop an authoring tool for data journalists (see Chapter 8). The provided function setEditMode() enables users to edit or delete the onboarding message, the tooltip title, an onboarding stage, etc., via the UI.

VisAhoi is reusable as developers can reuse onboarding messages. Developers can access the state of the onboarding for all visualizations using the context key (Objective Reusable). They can also retrieve the onboarding messages from VisAhoi and store them in their application state to reuse them for other visualizations in the same application.

Furthermore, VisAhoi is extensible (Objective Extensible). First, it currently supports three different visualization libraries Vega-Lite, Plotly.js, and Apache ECharts. Second, different visualization techniques are supported such as bar charts, change matrices, horizon graphs, scatterplots, and treemaps. We showed the extensibility of the VisAhoi library in the usage scenarios in Chapter 9. We also implemented a generic type to provide onboarding for other visualization techniques. The only restriction is that there are no semi-automatically generated onboarding messages available as we do not know the exact DOM tree. The developer has to define them manually.

#### 7.7.1 Limitation

A general limitation of VisAhoi is the fragility of the visualization adapters, as they depend on the runtime objects of the visualization libraries. The adapter might break or return different onboarding messages with different versions of the visualization library when the internal data structure changes. As mentioned in Section 7.6, we require the runtime object to access the actual values used to render the visualization. A solution for this limitation would be a standardized application programming interface (API) for runtime objects that the visualization libraries provide.

#### 7.7.2 Future Work

**Canvas Support.** VisAhoi explicitly targets visualization libraries with SVG rendering, which facilitates the selection of anchors for the in-place annotations by employing the DOM structure. In contrast, canvas rendering has a better performance but does not provide additional structure. The anchor coordinates must be retrieved exclusively from the runtime objects that often lack this information. Thus, we positioned the onboarding message manually in the hexbin plot to provide the instruction for the position and meaning of the hexagon, as we could not retrieve the DOM element due to the canvas restriction. In a future iteration, we want to analyze the canvas renderer of the three selected visualization libraries. The approach could then be applied to other libraries, such as the popular Chart.js (2020) library, which uses canvas rendering only.

Semi-automatic Onboarding for Interactions. Explaining available user interactions with a visualization, such as, for instance, brushing or item selection, is a major aspect besides visual encoding. In contrast to visual encoding, the declaration of interactions differs between visualization libraries and can be more complex in multiple coordinated views (MCV). ECharts and Vega-Lite describe interactions as part of the visualization specification (Satyanarayan et al., 2017). The visualization library adapter could identify and extract the information necessary to generate matching onboarding messages. For Plotly.js, VisAhoi provides onboarding messages for all interaction possibilities in the modebar. When creating a single visualization, developers can decide which interactions to allow and show in the modebar and VisAhoi adapts the corresponding onboarding message. In a MCV setup, Plotly.js requires custom JavaScript code using Plotly's selection event listener for MCV with linked visualizations. Hence, it is more challenging to implement onboarding messages for custom event listeners than to build a functionality that can be accessed using the runtime objects.

**Data Insights.** With the visualization onboarding in place, the infrastructure can provide insights into the dataset using quality metrics (Behrisch et al., 2018; Bertini et al., 2011). Quality metrics can reveal one or multiple interpretable visual patterns for a specific visualization type and dataset. Onboarding can explain the findings to the user in a comprehensive manner. Our current prototype provides a glimpse of this feature by displaying extracted statistical values from the visualization library, e.g., minimum and maximum values. Showing onboarding messages for data insights could be especially useful in combination with the user interaction creating a feedback loop that could guide users in their visual analysis process (Ceneda et al., 2017). Once the user interacts with the visualization, the quality metric is re-computed. It reveals new, potentially exciting patterns that can be presented to the user as onboarding messages and/or highlights. This aligns with our vision, where users get an onboarding tailored to a visualization and a dataset rather than a generic tutorial.

## 7.8 Conclusion

We presented VisAhoi, a JavaScript library that semi-automatically extracts and describes onboarding annotations in the respective visualization using the high-level declarative visualization grammars Vega-Lite, Plotly.js, and ECharts. VisAhoi is an open source system available at https://github.com/datavisyn/visAhoi.

The VisAhoi library is easy to integrate and extensible, as the developers can write their own adapters to use other visualization libraries. Besides, the library is reusable and highly customizable. The current prototype demonstrates that it can generate and integrate onboarding information semi-automatically, which provides a solid foundation for future developments.

In Part III, we present two design studies to show the applicability of the VisAhoi library. Chapter 8 presents the integration of VisAhoi a Flourish template to provide an authoring tool for data-journalistic use cases for a treemap visualization. Furthermore, we show the results of a design study in the field of biomedical R&D in Chapter 9. We integrated the VisAhoi library into a VA tool to analyze high-throughput screening (HTS) data.

# Part III

# Applying Visualization Onboarding Concepts in Design Studies



# CHAPTER 8

# Design Study 1: Authoring Tool for (Data)Journalists to Integrate Visualization Onboarding in a Treemap Visualization



This chapter starts with a general motivation for the design study about the selfexplanatory visualization onboarding approach for data-journalistic use cases (Section 8.1). We describe our approach (Section 8.2) and continue with related work (Section 8.3)), discuss current visualization onboarding concepts in the scientific community, particularly in the data-journalistic context. We also present the visualization onboarding design and implementation of the proof-of-concept prototype for a treemap visualization. The results of iterative user experience evaluation with (data)journalists and their reflections on them are also part of this chapter. We provide a detailed description of lessons learned. Find the study design in Appendix E. Download supplemental material, including Excel sheets with the categorization of the transcript and a video showing the usage of the authoring tool at https://phaidra.fhstp.ac.at/o:5219. This chapter is based on our published work (Stoiber et al., 2023).

# 8.1 Motivation & Background

Data journalists face the challenge of using simple data visualizations to avoid overwhelming readers (Feigenbaum et al., 2016; Riche et al., 2018). In any case, a certain level of visualization literacy (Boy et al., 2014; Lee et al., 2017) is required to analyze and interpret data successfully. Visualization literacy refers to the ability to extract information from data visualizations and gain insights in an effective, efficient, and safe manner (Lee et al., 2017).

Readers of data stories need to read and understand statistical information and data visualizations to gain insights from them (Börner et al., 2016; Engel et al., 2019). For example, in Figure 8.1, a data story is shown, which is published by Reuters (Harman and Lange, 2022), including a data visualization representing the split delegations in the US Senate. However, many users need help interpreting and working with novel visual representations they are not familiar with (Grammel et al., 2010; Perkhofer et al., 2019). This bears the risk of drawing incorrect conclusions and leads to frustration or the rejection of otherwise powerful data visualizations (Börner et al., 2016). Right now, the complexity and social relevance of the COVID-19 pandemic have put data visualization at the center of worldwide attention (Shneiderman, 2020). In this context, data visualization researchers, journalists, and experts have provided various data visualizations for public education. The general public got in touch with diverse data visualizations presenting medical data such as reproduction numbers, COVID-19 cases, hospitalizations, etc. The size and complexity of today's data sets overwhelm traditional business charts such as bar charts, line charts, or pie charts. Therefore, more advanced visual representations are necessary to capture more complex data structures and significant amounts of data. To address this challenge, visualization onboarding methods might be utilized to support users in reading, interpreting, and extracting information from visual representations of data (Stoiber et al., 2022b, 2021, 2022d).

Annotations in the data visualization can be used to explain how to read the graph, introduce the context, and highlight interesting insights (Hullman et al., 2013). Our user research with data journalists has shown that journalists integrate tooltips and text, reading examples, storytelling, or step-by-step guides to onboard users to more unfamiliar data visualizations in their data stories (see Section 8.4). As many journalists lack programming skills (Reilly, 2017, p.8), they use tools such as Flourish (2022) and Datawrapper (2023) to create their data visualizations (Abellán, 2022). Flourish (2022) is a platform that allows us to create data visualizations easily and is widely adopted in the data journalism (DDJ) community. Flourish (Abellán, 2022) only partially integrates some of the onboarding mentioned above approaches, such as annotations (Flourish Annotations, 2023) or storytelling (Flourish – create story, 2022). However, these approaches exhibit weaknesses as they have to be designed fully manually, are not related to the data, cannot be used as a guided tour, and affect more experienced users in their exploration as they are always visible. We developed a JavaScript library, VisAhoi, to generate onboarding instructions semi-automatically and easily integrate them into VA tools. The technical description of the library is presented in the previous chapter 7.



By Travis Harman and Jason Lange PUBLISHED DEC. 1, 2022

The U.S. Senate increasingly features state delegations where both members come from the same party, a sign of hardening regional divisions between Democrats and Republicans.

Only six states currently have split delegations, the lowest number since a 1913 constitutional amendment allowed for direct election of U.S. senators

Following Democrat John Fetterman's victory last month in a Pennsylvania contest for retiring Republican Senator Pat Toomey's seat, the number could fall to five if Democratic Senator Raphael Warnock successfully defends his Georgia seat in a Dec. 6 run-off contest against Republican Herschel Walker.





Figure 8.1: **Exemplary data story** by Reuters (Harman and Lange, 2022) showing every state's senate partial breakdown by year from 1914 to 2022.

# 8. Design Study 1: Authoring Tool for (Data)Journalists to Integrate Visualization Onboarding in a Treemap Visualization

**Design study contribution:** In this design study, we show the applicability of the VisAhoi library by reporting on the design, development, and evaluation of an authoring tool for (data)journalists to integrate self-explanatory visualization onboarding into a treemap visualization. We used a treemap visualization to show the authoring tool in practice exemplarily. Treemaps (Shneiderman, 1992) are well suited to represent hierarchical data with a quantitative attribute. However, treemaps are not particularly well-known among audiences (Börner et al., 2019) and might be hard to interpret initially. Therefore, we decided to demonstrate the validity of our approach and the functionality of our authoring tool based on a treemap representation.



Flourish Authoring tool for integrating selfexplanatory visualization onboarding semiautomatically generated using VisAhoi is library

Figure 8.2: Flourish authoring tool components

The authoring tool combines a Flourish template and the VisAhoi library, as shown in Figure 8.2. It provides a user interface to add onboarding and customize it: (1) editing, reordering, and/or removing the semi-automatically generated onboarding messages and onboarding stages, and (2) adding own messages and stages. In our proof-of-concept implementation, we used a treemap visualization with Biden's tax overhaul (Probasco, 2021) data as an example data set.

Over the last two and a half years, we collaborated closely with (data)journalists, thereby iteratively developing a better understanding of the problem and designing, implementing, and evaluating a prototype to address it. Methodologically, we relied on interviews and user experience evaluation with domain experts to fit our time-pressured target audience.

In summary, the main contributions of this design study are:

• a problem characterization and abstraction (see Chapter 8.4);

- the design and implementation of an authoring tool providing self-explanatory visualization onboarding, which includes semi-automatically generated onboarding messages for a treemap visualization that can be adjusted and expanded by the journalist, as described in Chapter. 8.5;
- the results of four iterative user experience evaluations with seven Austrian data journalists (see Chapter 8.7) to gain qualitative feedback from the domain experts on our visualization onboarding authoring tool regarding usability and user experience;
- a set of general lessons learned which we derived from our project, such as the fact that (data) journalists like the authoring tool to provide self-explanatory visualization onboarding. In their opinion is, the onboarding concept appropriate for readers, whereas the onboarding messages need to be simplified and less complex to avoid overwhelming the recipients (see Chapter 8.8).

## 8.2 Method



Figure 8.3: Timeline of the design, development, validation, and reflection phases from fall 2020 to November 2022

We decided to validate our Flourish-based prototype with the integrated VisAhoi library using a series of user studies with (data) journalists. As is common for design studies, we followed a process of iterative design and evaluation to address a particular domain problem by involving collaborators and users from the domain (SedImair et al., 2012). The online interviews for the problem characterization and abstraction phase were chosen to reach out to our busy domain experts easily. A timeline illustrating the *Flourish template* development and the iterative validation phases is shown in Figure 8.3. We followed a mixed-methods approach for the **user research** to understand the challenges in data journalism in terms of data, user, and tasks, as well as current practices, problems, and used tools. We derived implications for the design of onboarding concepts and the development of the proof-of-concept prototype (Flourish template). The user research, including ten qualitative interviews and a survey with data journalists (trainers) (15 answers), was conducted from fall 2020 to spring 2022. Following the requirements analysis, we produced high-fidelity prototypes based on the design requirements and the results of our previous studies (Stoiber et al., 2022b, 2021, 2022d). Following the **conceptual design** phase, we started with the **implementation** of the **VisAhoi library** (November 2021 - October 2022) and later with the **implementation** of the **Flourish template** with the included VisAhoi library. We conducted a pilot study for the **qualitative user evaluation** with two experts in the field of HCI and Visualization and iteratively improved the prototype.

Furthermore, we performed **four iterative user experience evaluations** either in a one-on-one or group setting. Based on the feedback, we revised and improved the Flourish template for the following evaluations. Finally, we analyzed the evaluation data to improve the proof-of-concept prototype and started with the reflection.

Next, we outline related work for the addressed areas of interest.

# 8.3 Related Work

This section covers related work about visualization onboarding concepts, particularly in the data-journalistic context.

### 8.3.1 Visualization Onboarding in Data Journalism

In recent years, integrating annotations in data visualizations has become increasingly important (Riche et al., 2018, p.214-215) to provide the context of the data story to the audience, explain how to read an unfamiliar data visualization (visualization onboarding), and highlight insights to support users in insight generation. So far, data journalists integrate tooltips and text (Baur and Stephaner, 2018), reading examples, and use storytelling or step-by-step guides to onboard readers of their data story (see Section 8.4). The most used visualization tools in data journalism (DDJ) (Abellán, 2022) such as Flourish (2022) and Datawrapper (2023) provide ways to annotate the developed data visualization (Flourish – create story, 2022; Flourish Annotations, 2023; Muth, 2022). Flourish (2022) allow users to add annotations (see Figure 8.5) or add animated transitions to create richer stories (Flourish – create story, 2022). Flourish Annotations (2023) can be added as labels, blocks of text, or connector lines to restricted data visualizations, e.g., line charts, bar charts, or pie charts only.

Also, Datawrapper (2023) provides text annotations to highlight data insights, illustrated in Figure 8.4. These approaches have to be fully manually integrated by the data journalists. Additionally, the annotations could affect more experienced users in their exploration as they are always visible and are restricted to a small number of visualization techniques.

Approaches to automatically create annotations for visualization in data stories are present in the literature (Hullman et al., 2013; Lai et al., 2020). However, these approaches cannot be integrated into journalists' most used visualization tools.

Kim et al. (2021) explored how readers of line charts gather takeaways when looking at the data visualization and caption together. They proposed guidelines on how to improve



Figure 8.4: DataWrapper annotation example (Muth, 2022)

the way conveying their message to the reader. By highlighting or adding annotations in the caption, the authors can emphasize the key messages of the data visualization, for example.

Furthermore, data visualization and motion graphics are used in *data videos* to tell a data story (Segel and Heer, 2010). Animation can encode data attributes such as revealing relationships, showing uncertainties, or conveying emotions (Chevalier et al., 2016; Huron et al., 2013), which could be used to onboard users as well. Shi et al. (2021) presented a design space for animated visual narrative in data videos. Furthermore, Yang et al. (2022) introduced a design space as practical guidance on achieving narrative intents, organizing data facts, and selecting visual design techniques through story creation.

Most data journalists have no coding skills and need to integrate visualization onboarding into their data visualizations. We support journalists in easily adding and editing onboarding instructions using the authoring interface in Flourish's visualization tool for a treemap visualization. Our onboarding concepts take the knowledge gap of the reader into account by providing interactive, context-sensitive, and self-explanatory onboarding instructions (see Chapter 6.2).

# 8.4 Problem Characterization and Abstraction

In the first phase of the design study, we conducted interviews and a survey with data journalists and data journalism trainers to collect data regarding the domain experts' needs and challenges and current practices regarding visualization onboarding. In Chapter 6.2, we present a detailed description of the visualization onboarding concept. Journalists must also understand the chosen onboarding method from the readers' point of view and, above all, find it suitable and useful so that they would also use/incorporate it. Thus, in the design study's problem characterization and abstraction phase, we gathered information regarding the visualization onboarding design to identify if our design is also appropriate in data journalism.



Figure 8.5: Flourish annotation example (Bider, 2020)

#### 8.4.1 Method & participants

Most of the participants were not only doing training in the field of data journalism but also worked as (data) journalists daily. Therefore, we decided to conduct our research in data journalism training as both fields – onboarding and training – mainly focus on learning processes. We conducted ten interviews with data journalists (trainers) in the German-speaking area of Europe (Germany, Austria, Switzerland) between fall 2020 and spring 2021 (gender: m = 6, f = 4). The interview material was analyzed using an approach of content structuring/content analysis according to Kuckartz (Kuckartz, 2014) by two coders. We further developed a survey for the same target group based on our findings. We received 15 complete answers from fall 2021 till spring 2022 (gender: m =14, no answer = 1).

#### 8.4.2 Results

As our interview partners indicated, data visualizations are used to facilitate data work in the journalistic workflow. They can help to get an overview or find a story when data sets are large or highly complex. Furthermore, data visualization is used to gain first insights, display data over time, recognize data failures, compare data, or show the distribution of data. Nevertheless, data visualizations are most often seen as the final product of a data story. They are often excluded from data journalism as many media companies have their own graphics and design departments that are responsible for data visualizations. Data visualizations are therefore seen as visual assistance when working with a data set as well as an end product.

Building data visualizations is described as less challenging than data work. Participants assess their familiarity with data visualizations as high, and the knowledge of visualization types is considered important as well (the same tendencies emerged in (Abellán, 2022)).

The most frequently mentioned challenge is finding the right data visualization for a data set and the story (data)journalists want to tell. Oftentimes, tools are therefore misleading as they offer too many options. The decision of whether to use a tool or build a custom data visualization has been mentioned as another challenge. Besides spreadsheet programs, participants mention using various tools such as Datawrapper (2023), Q (q nzz, 2023), Flourish (2022), Chartbuilder by Quartz (quartz, 2022), or Raw Graphs (2023). Working with a tool might be faster but has limitations, and data sets often must be further cleaned and processed. In contrast, custom data visualizations need programming skills and take more effort and time. Thus, both approaches have advantages and disadvantages, and it depends highly on the story's context or the data visualizations is to display a data set quickly and to tell a story with little effort.

Regarding current practices in visualization onboarding, data journalists mentioned the following: Data journalists use a variety of possibilities to explain data visualizations. The easiest way is to use text close to the data visualization, titles, and subtitles. Data journalists also mentioned including legends, note fields, colors, and annotations within a data visualization to make it understandable. Concerning the context in which onboarding should be provided, journalists found it very important that onboarding methods are context-sensitive rather than context-free. Looking at where the onboarding should take place, the answers show that internal (inside the data story) onboarding is the preferred method. Moreover, onboarding should be provided during the interaction with the data visualization. This makes it very important for data journalists to be able to change the content and design of messages to meet the needs of readers but also corporate design requirements.

Participants are still determining whether data and visualization literacy have increased in the population. They generally use rather simplistic and presumably understandable visualizations not to overwhelm the audience. However, participants would also be interested in working with more unfamiliar data visualizations to the audience, such as scatterplots, network diagrams, logarithmic representations, radar charts, or heat maps. The COVID-19 pandemic, in particular, has created opportunities for more elaborate data visualizations as readers were increasingly exposed to data stories (cf. (Shneiderman, 2020)). Participants concluded that, on the one hand, it is now possible to expect readers to understand more unfamiliar data and data visualizations. However, on the other hand, they still have concerns about using more unfamiliar data visualizations. All in all, they mention a new awareness of their power and responsibility as well as of the role of media within society. Some similarities allowed us to derive the key requirements listed in the following section.

#### 8.4.3 Design Requirements

Based on the results of previous studies (Stoiber et al., 2022b, 2021, 2022d) (see Chapter 5 and Chapter 6.2) and the problem characterization (interviews & survey), we identified five key design requirements (DR). We present design requirements related to the **authoring** 

**phase**, which the journalist performs, and a **reading phase** targeted to the reader of the data story. The authoring phase is when the journalist adds and edits the onboarding instructions using our authoring tool in Flourish. When the data story is published, the reader is consuming and interacting with the visualization onboarding. We present the design requirements covering the *reading phase* to show the applicability and suitability of the VisAhoi library/onboarding design in the authoring tool from the journalist's perspective, as they are the ones incorporating it into their data stories.

#### Authoring phase (data journalist)

- **DR-A1** Integrate visualization onboarding into a well-known visualization tool: The interviewee named data visualization software, such as Datawrapper and Flourish, as accessible and simple online tools to visualize data. The Flourish SDK (2023) allows developers to create their templates and upload them to Flourish, so we decided to use it to implement our self-explanatory visualization onboarding prototype.
- **DR-A2** Semi-automated generation of onboarding messages: Data projects are often lengthy and difficult to reconcile with the fast-paced daily editorial routine (Appelgren and Nygren, 2014). Journalists also have to deal with the challenge of using self-explanatory and simple data visualizations to avoid overwhelming readers (Feigenbaum et al., 2016; Riche et al., 2018). However, they would be interested (and the COVID-19 pandemic has opened opportunities to do so) in using more complex graphics. A semi-automated generation of onboarding messages can support journalists in setting up onboarding more efficiently as they function as a blueprint and quickly explain unfamiliar visualizations to their readers.
- **DR-A3** Customize onboarding: As the background and knowledge of readers are diverse and the onboarding messages must fulfill corporate design requirements, an onboarding tool needs to provide possibilities to customize the onboarding, e.g., change color, icons, and onboarding instructions.

#### Reading phase (audience/reader)

- **DR-R1** Support readers in interpreting unfamiliar visualizations: In the interview, journalists mentioned reading examples, short textual explanations, and tooltips to onboard users to data visualizations. The visualization onboarding concept should support the readers with textual instructions by using tooltips that can be integrated into the data visualization.
- **DR-R2** Provide context-sensitive onboarding which can be called up during exploration: Context-sensitive means to provide onboarding instructions at a specific point in the current state of the tool (Stoiber et al., 2022c). An onboarding system integrated internally into the data visualization helps users not jump back

and forth between two different systems. Additionally, onboarding should be called up while using the data visualization for insight generation.

Two of the five essential requirements (DR-R1, DR-R2) are related directly to the design of visualization onboarding concepts and features. The presented VisAhoi.js library covers DR-R1, DR-R2, and DR-A2, DR-A3 (see Chapter 7). DR-A1 emphasizes the importance of using a tool widely used by data journalists. Therefore, we opted to use Flourish as our platform, as it is also open to developing templates where other components can be integrated.

## 8.5 Conceptual Design & Implementation

In this section, we present the main design of the authoring tool based on the design decisions for onboarding (explained in detail in Chapter 6 and Chapter 7) and the design requirements (DR) in particular for the domain of data journalists presented in Section 8.4.3.

#### 8.5.1 Visualization onboarding concept

The floating action button (onboarding button) is the main navigation element to access onboarding instructions. The floating action button is positioned next to the data visualization. We position it to the bottom right next (see Figure 8.7 (6)).

The onboarding design uses textual instructions (onboarding messages) in the form of tooltips in combination with in-place annotation anchors (DR-R1) to indicate the connection between the text elements and the visual encoding, illustrated in Figure 8.9. We structured the onboarding messages in three default onboarding stages: *Reading*, *Interacting*, and *Analyzing*. The onboarding stages are the three circular buttons when clicking the floating action button, illustrated in Figure 8.6. Some onboarding messages and stages are generated semi-automatically by default based on the visualized data (DR-A2). We provide a more generic description of the treemap, which does not include any concrete values from the data visualization, e.g., "*The size of each rectangle represents a quantitative value associated with each element in the hierarchy*." Based on our comparative study (Stoiber et al., 2022b) between abstract and concrete onboarding messages for a treemap visualization, we found that both concrete and abstract onboarding messages to provide a blueprint for the journalists, which can be reused more easily for other treemap visualizations.

We designed the onboarding to allow users to be (1) **guided** through the different onboarding messages or (2) **self-explore** it and seek help at any stage of their interpretation and interaction with the data visualization (DR-R2). Therefore, we included a **step-by-step navigation widget** representing the number of visual markers/onboarding messages and **arrow icons** that allow the user to browse through the onboarding message and

# 8. Design Study 1: Authoring Tool for (Data)Journalists to Integrate Visualization Onboarding in a Treemap Visualization



Figure 8.6: **Onboarding button** – Floating action button (1). The navigation is composed of different elements to provide a guided tour as well as a self-exploratory experience. By clicking on the floating action button (1), the menu expands outward from a central point and reveals three onboarding stages: Reading, Interacting, and Analyzing (2). We decided to integrate a step-by-step navigation widget that represents onboarding messages as well as arrow icons to navigate through it (3).

receive guided help (see Figure 8.6). Users get help for a specific element or part of the data visualization, e.g., interaction, or visual encoding, when clicking on a specific visual marker. That way, we can provide a self-exploratory experience. A toggle button below the float action button deactivates the step-by-step navigation widget, illustrated in Figure 8.7 (6).

The onboarding concept allows users to **interact simultaneously** with the underlying **data visualization** and the **onboarding** (DR-R2).

#### 8.5.2 Conceptual design of Flourish template

To illustrate our approach, we use a treemap visualization presenting Biden's tax overhaul (Probasco, 2021) data as an example data set (see Figure 8.7). In the following paragraphs, we describe the different parts and features of the Flourish template in more detail (DR-A1). In Section 8.6, we present the technical parts of the implementation.

Layout: As illustrated in Figure 8.7, the basic layout of the template is predefined by



Figure 8.7: Authoring tool. The proof-of-concept implementation of a self-explanatory visualization onboarding approach for a treemap visualization. The interface is divided into two areas: (A) preview with the data visualization (5), the onboarding button (6), and (B) the settings panel of Flourish (1) with the settings regarding onboarding customizations (2, 3, 4). Also, journalists can change the data of the treemap (7) provided by the Flourish template.

the Flourish SDK (2023) which uses a preview window (1) (visualization area - left part of the interface) and a settings panel on the right-hand side (2, 3, 4). The most important part of Flourish is the settings panel, where users can modify the data visualization and appearance (1). We added another set of onboarding-specific settings to *add* (2), *customize* (3), and *save* (4) the current onboarding instructions. We included a GIF overlay when opening the Flourish template for the first time to onboard the Flourish user to the saving process. Journalists can change the data of the treemap (7) which is provided by the Flourish template.

**Customizing onboarding messages and stages:** Journalists can add the onboarding concept using the settings panel (DR-A3). They can either use the predefined semi-automatically generated onboarding messages and stages and adjust them by activating the editing mode in the settings panel of Flourish or add new ones through the edit mode. When activating edit mode, a button appears in the preview area in the right corner to *add* and *customize* the onboarding stages and messages in a separate overlay (see Figure 8.7 (3)). In the same overlay, users can edit the name, color, icons, and order

# 8. Design Study 1: Authoring Tool for (Data)Journalists to Integrate Visualization Onboarding in a Treemap Visualization

Ahoi - Customization				×		
Add Onboarding Stage						
Title	+ Add (	🖁 Edit 🗧	2 Reorder,	adjust color	and char	nge icon for onboarding s
Enter a title	Edit O	nboardin	g Stages	?		
The title of the stage. Color The color of the stage and it's message	Explanation Drag Click Click Warr	on: and Drop the a on the 🕑 icon on the 面 icon f <mark>sing!</mark> This will a	stage (list item) t to edit a stage. to delete a stage. Iso delete all me:	o reorder it.	associated	with the stage.
Enter a icon name	Title:	Reading	Color:	Icon:	69	Actions: 🗹 💼
Only FontAwesome 5 icons are availab	Title:	Interacting	Color:	Icon:	<b>5</b>	Actions: 🗹 🛅
https://fontawesome.com/v5/search						

Figure 8.8: Flourish graphical user interface. Interface to add onboarding stages and onboarding messages (1) as well as to edit onboarding stages such as color, icon, title of the stage, and order (2).

of the onboarding stages, as illustrated in Figure 8.8 (2). Onboarding stages can also be deleted using the dustbin icon. While in edit mode, onboarding messages can be directly adjusted in the data visualization. Small icons appear next to the title of the onboarding message to either *delete* the whole message or *edit* the onboarding text directly in the tooltip, as illustrated in Figure 8.9.

# 8.6 Implementation

The VisAhoi library and the Flourish template are developed using JavaScript, including other JavaScript libraries. A detailed explanation can be found in the respective subsections.

#### VisAhoi Library

VisAhoi provides interactive onboarding for data visualizations generated from highlevel declarative visualization grammars such as Vega-Lite (Satyanarayan et al., 2017), Plotly.js (Plotly.js, 2023), and ECharts (Li et al., 2018) (see Chapter 7). VisAhoi is



Figure 8.9: Edit mode of an onboarding message. Edit and delete onboarding messages and the respective in-place annotation in the data visualization using the small icons next to the title of the onboarding message



Figure 8.10: **Overview of VisAhoi.** The VisAhoi adapter extracts the information from Plotly.js' runtime object and DOM nodes and passes them to the VisAhoi core. The core generates the basic onboarding stages and messages that can be customized and renders the onboarding user interface.

implemented using Svelte (Svelte, 2023), which compiles Svelte-specific code to plain JavaScript. The compiled source code is framework-agnostic and can be used in combination with other front-end frameworks such as React (React, 2023), Vue (vue, 2023), or Angular (Angular, 2023). In chapter 7, we describe the design of the JavaScript library in more detail.

Figure 8.10 shows the main components of the onboarding generation process for a treemap visualization in Flourish:

1. The treemap visualization is created from a high-level declarative visualization grammar by a visualization library such as Plotly.js in Flourish. To render the visu-

alization, Plotly.js creates a runtime object from the grammar and adds computed values about the chart (e.g., the position of data items).

- 2. The VisAhoi adapter for Plotly.js extracts information such as the chart title or positions and values of data items for the onboarding from the runtime object and the DOM nodes of the data visualization.
- 3. The VisAhoi core generates basic onboarding stages and messages by filling the extracted information from the adapter into predefined templates. Stages that do not have messages are filtered out.
- 4. Users can customize the onboarding by adding, editing, and deleting the provided onboarding stages and messages.
- 5. VisAhoi generates the onboarding user interface based on the customized onboarding stages and messages. It integrates the visual markers at the given position in the data visualization and renders the onboarding.

Currently, VisAhoi is extensible to multiple visualization libraries and supports onboarding for scatter charts, horizon charts, bar charts, heat maps, and treemaps for each supported visualization libraries. We also implemented a generic type where developers can add onboarding to any visualization technique. This is restricted, as the developer doesn't get any semi-automatically generated onboarding messages as a blueprint.

#### Flourish template

The onboarding logic and basic functionality have been taken from the VisAhoi library (see Sec. 8.6), which we integrated into the Flourish template. We used the Flourish SDK (2023) for creating custom templates. In general, the templates are divided into a fully customizable visualization area (see Figure 8.7 (A)) and the settings panel (see Figure 8.7 (B)), which has limited customization possibilities. We also use Plotly.js (Plotly.js, 2023) to render the data visualization and the VisAhoi library to create the semi-automatically generated onboarding instructions. The Flourish template is open-source and can be accessed with a Flourish account at http://bit.ly/3gBqrxv. We provide the code for the authoring tool in the GitHub repository at https://github.com/doomsay er2/Onboarding\_DDJ\_UseCase.

We placed our custom overlay menu for authoring onboarding instructions in the visualization area (see Figure 8.7 (3)). The overlay mainly uses Bootstrap (bootstrap, 2023) for the styling.

We used the settings menu behavior to automatically save the customizations of the onboarding once the user interacts with one of the controls in the settings panel of Flourish. We provided two text fields — for the onboarding stages and messages separately — where the user has to copy a JSON string automatically generated when pressing the "save" button. The separation was necessary for VisAhoi to restore the previous state of

the Flourish template when reloaded. This workaround was needed as only settings in the Flourish settings panel are saved. A GIF animation support users when opening the Flourish template for the first time.

## 8.7 Validation & Results

From August to October 2022, we conducted four user experience evaluations (with seven participants), including semi-structured interviews (Lazar et al., 2017) in different settings with domain experts to (1) evaluate the usability of the Flourish template iteratively (see Figure 8.7) and (2) gain feedback on the onboarding concept itself. Before the user experience evaluations, we conducted two pilot studies with experts in HCI and visualization. The following subsections describe the method as well as the results.

Table 8.1: **Demographic information of journalists.** Reported experience level with Flourish along a 5-point Likert scale

Participant	Gender	Age	Profession	Experience w. Flourish	Experience w. visualization tools
P1 (study 1)	male	35	journalist	experienced $(4)$	Datawrapper, Flourish
P2 (study 2)	male	40	journalist	not experienced $(1)$	Datawrapper, Tableau, Flourish
P3 (study $3$ )	male	42	data journalist	very experi- enced (5)	Datawrapper, Rid- dle
P4 (study $3$ )	male	29	data journalist	little experience (2)	ggplot, qgis, d3
P5 (study $3$ )	male	32	data journalist	experienced $(4)$	Datawrapper, D3/Javascript
P6 (study $3$ )	female	30	data journalist	not experienced $(1)$	Datawrapper, Tableau, Excel
P7 (study 4)	female	31	data journalist	experienced $(4)$	Datawrapper, Flourish

#### 8.7.1 Method & participants

We recruited seven Austrian (data) journalists (in four studies) aged 29 to 42 (gender: m = 5, f = 2) for the study. Demographics are presented in Table 8.1. The subjects reported having experience with the following visualization tools in the context of journalism: (Datawrapper, 2023), Flourish (2022), Tableau (Tableau Software, 2023), Riddle, ggplot (Wickham et al., 2022), qgis(QGIS, 2022), d3.js (D3 Language, 2023), and Excel. The participants subjectively rated their experience with Flourish along a 5-point

Likert scale. The participants generally rated their experience "experienced" (4 Mdn on the Likert scale), whereas two participants with "not experienced"

Due to the COVID-19 pandemic, the usability studies were conducted online using MS Teams. The tests took approximately 40 min. Sessions 1, 2, and 4 were in a one-on-one setting, whereas study 3 was in a group setting with four journalists (see Figure 8.11). In the one-on-one settings, the usability study was structured as follows: (1) introduction to the study, (2) collection of demographic data and level of experience with the tool Flourish as well with other visualization tools, (3) 5 minutes test phase to familiarise the participants with the Flourish template interface, (4) working through several tasks using the interface, (5) open questions (to gain more feedback regarding the Flourish template as well as the onboarding concept itself). In the group setting, we adjusted the test designs, as illustrated in Figure 8.11, by using Google Forms to include demographic data and level of experience (2) as well as the tasks (4).

The tasks were as follows:

- Tasks 1: Read saving overlay
- Tasks 2: Add onboarding to the treemap visualization
- Tasks 3: Position onboarding button
- Tasks 4: Read through the onboarding instructions
- Tasks 5: Add the onboarding stage
- Tasks 6: Add onboarding message to the added onboarding stage
- Tasks 7: Adjust the text in the onboarding message
- Tasks 8: Delete the onboarding message
- Tasks 9: Delete the onboarding stage "Interacting"
- Tasks 10: Save

Find the study design in Appendix E.

The participants worked independently through the tasks on their laptops using Google Forms. Finally, we asked open questions in a focus group discussion setting. We recorded the interviews and demonstrations using MS Teams and took notes during the observations. Subsequently, the interviews were transcribed. We asked the participants to think aloud while working through the tasks during the evaluation. Based on the collected material, we performed a content analysis.

#### 8.7.2 Data analysis strategy

We categorized the usability problems along their **severity** (Nielsen, 2022). Besides, the **effort** can be described as the estimated time needed to implement the user feedback into



Figure 8.11: **Procedure of the usability study with journalists in online oneon-one or group settings.** In the group setting, we used google forms to collect data regarding demographics and experience and presented the tasks to the participants. In the final part of the evaluation, we conducted a focus group discussion instead of a semi-structured interview, having the same set of questions.

the system (see Table 8.2 for details about the used scales). We also analyzed the semistructured interviews and the thinking-aloud part of the usability study using qualitative content analysis by Kuckartz (2014). We used the mixed approach, including a deductive and inductive category application. One coder started with the definition of categories. Furthermore, we went through the content of the categories and started to split them into sub-codes. Two other coders categorized the transcripts again and revised the set of categories. Please refer to the URL https://phaidra.fhstp.ac.at/o:5190 for supplemental material, including details about the used categories.

Table 8.2: Categorization	scheme for	usability issues
---------------------------	------------	------------------

Ranking	Severity	Effort
1	minor usability problem: does not affect the user experience	$\mathbf{low}$ effort to implement
2	<b>major usability problem</b> : important to fix; average impact on the usage of the system	<b>average</b> effort to implement
3	<b>usability disaster:</b> necessary to fix because the problem affects the user experience and the usage of the system	$\mathbf{high}$ effort to implement

#### 8.7.3 Results

We report on the qualitative feedback in the following paragraphs. All comments and quotes of the participants have been translated from German into English by the authors.

We asked data journalists to report on their current practice when integrating onboarding

Task	Usability problem	Severity	Effort
1 – read saving overlay	GIF is too small	1	1
5 – add a stage	text next to the settings button is small	3	2
5 – add a stage	could not find the setting button, or it was not clear how to activate the toggle button	2	1
5 - add a stage	difference between onboarding stage and message	2	1
9 – delete onboarding stage	not intuitive to delete it using the icon instead of the settings panel	2	3
10 – save	saving is complicated	2	3

Table 8.3: Categorization of identified problems

concepts into data stories. Some of the participants mentioned that they do not integrate any onboarding at all (P2, P3). They mentioned using reading examples accompanying the data visualization when it becomes more complex for the recipient. Additionally, they integrate explanatory textual descriptions in the text of the data story. Furthermore, interactive tooltips in data visualization are integrated, and storytelling elements guide readers.

Our qualitative analysis indicated that the usage of the Flourish template was "clear", "straightforward", as well as "fast". A participant in P7 commented that "It was a little different from what I was used to. It's not the typical Flourish editing process. But most of it is quite intuitive." In study 3 (P3, 4, 5), data journalists mentioned that "[...] it just took me a few minutes to get used to it." Furthermore, the participants highlighted the well-designed onboarding concept. They liked the integrated step-by-step and ordered instructions that helped recipients to interpret the visualization (P1, P7) successfully. Data journalists also mentioned using the onboarding concept to explain insights and report on missing data values narratively.

*Usability problems:* Table 8.3 shows the usability problems by tasks along the defined schemes (severity and effort). We report on usability issues, remarks, and feedback regarding the onboarding design, onboarding stages, messages, feature requests, and general remarks.

The GIF in the overlay explaining the saving process was found to be too small. Participants in study 3, in particular, highlighted this issue. The severity was defined as a "minor problem" as we explained how to save customization in the Flourish template. The effort to fix this problem is defined as "low" as it can be eliminated by increasing the size of the overlay and the presented GIF.

Furthermore, some participants did not activate the toggle button in Flourish's settings panel; therefore, the button used to add an onboarding stage did not appear in the preview editor (task 5 - studies 2 & 3). We ranked this usability problem as "major" as it affects the usage of the whole system. The effort is "average" to set up documentation on how to use the Flourish template and increase the affordance of the button by, i.e.,

increasing the size or adding some animation. Moreover, the text box next to this button was also considered too small (study 3). We defined the severity of this problem as "minor" as it affects the user experience only. The effort was defined as "low" as it is easy to increase the font size of the text.

Two participants needed help understanding the difference between onboarding messages and stages (P1, P2). We resolved this issue before conducting the other two test sessions. A screenshot explaining the onboarding stage and message was integrated into the interface to help users understand the difference. We defined the severity level as "major" and the effort as "low".

Task 9, deleting an onboarding stage, needed to be more intuitive. The subjects mentioned that they had yet to see the dustbin icon next to the onboarding stage in editing mode as they searched for the functionality in the settings overlay where they had added a stage before. In the first test sessions, the subjects identified an issue with the order of icons in the onboarding stage button, which we solved before study 4 (P7). This usability problem was ranked as "major" and "high effort". There are more possibilities to solve this by creating documentation and changing the interface and structure.

Another usability issue appeared in the saving task (Task 10). Due to the technical restrictions of Flourish, we had to implement this workaround to copy the declaration of the overlay in the settings panel of Flourish to save the adaptations made in the template. Therefore, it was not surprising that the participants had some issues. By providing the GIF and including a detailed description, we help the user perform this task successfully. The subject (P2) mentioned "[...] saving is a bit complicated, but it works once you have done it." Therefore, we ranked it as a "major" usability issue and "high" effort to solve it, as it would require getting in touch with the developer of Flourish to integrate this feature in the main Flourish SDK (Flourish SDK, 2023). In the following paragraphs, we present the feedback and comments on the thinking-aloud method.

Onboarding design: The participants mentioned that the functionality of the small toggle button below the onboarding button (see Figure 8.6) is unclear. They also suggested allowing the user to use the arrow keys to navigate through the step-by-step navigation (P1, P4, P7). Remarks regarding the onboarding design included the affordance of the onboarding button. Some data journalists highlighted that it could be a problem for readers to recognize the onboarding button (studies 1 & 3). They recommended opening the Reading section of the onboarding to point the recipient towards the onboarding when reading the data visualization for the first time. Another suggestion is adding pulsating animation to the button to draw attention.

Onboarding stage: The onboarding messages are divided into the onboarding stages Reading, Interacting, and Analyzing. Participants mentioned that they would reduce the complexity by just providing one onboarding stage for more common, less complex visualizations: "[...] I think on a day-to-day basis, I would handle it on one level. [...] I don't want to confuse them with chapters." (P7) The data journalist in study 1 also stated: "[...] but I think that's enough to leave that at one level."

# 8. Design Study 1: Authoring Tool for (Data)Journalists to Integrate Visualization Onboarding in a Treemap Visualization

Onboarding message: The integrated VisAhoi library semi-automatically generates abstract onboarding messages as a blueprint for journalists. We used these messages in our proof-of-concept prototype. Participants highlighted that onboarding messages should be available in German as well. Furthermore, there were some suggestions that the onboarding messages were too "technical and abstract" and "[...] put that in simpler terms to make it reader-friendly. So that it is aimed at a general audience. I would make it more practical, in that case, with examples." (P7) The participant (P2) also said that the onboarding messages were too complex, especially onboarding message four.

Feature request: During the interaction with the Flourish template, subjects requested some features to improve the user experience. In the first study, the journalists wanted to adjust the onboarding concept's color to fit the company's corporate design. Moreover, although changing onboarding messages was easy for the participants, they recommended adding a limitation of characters to avoid overwhelming the readers. Participants emphasized the need to change onboarding stages. Changing the order of onboarding stages was impossible in the proof-of-concept version for the evaluations. In our final version, changing the order of the stages, the icons, and the color is possible. Furthermore, the subjects recommended adding the feature to reorder and adjust the visual markers of the onboarding messages. It is only possible to delete the onboarding message and add it again.

*General remarks:* Participants recommended providing some best practices on onboarding, as well as explaining in more detail what is visible for the readers and which parts are just visible in the Flourish template. We also asked the participants whether they thought readers would understand the onboarding as an open question at the end of our study. They commented that it could be "helpful" (study 3), and they would use it (P1, P3-P7). In contrast, they also mentioned reducing the number of onboarding stages to avoid overwhelming readers when integrating our visualization onboarding into their data stories.

#### Take-home messages:

- Due to the onboarding design with integrated step-by-step and ordered instructions, the data journalists would use the template in their daily routines.
- Data journalists like the possibility of adjusting the onboarding concept to their needs.
- Data journalists highlighted the need to reduce the number of onboarding stages to avoid overwhelming readers.
- Data journalists suggest to simplify onboarding messages and referring to concrete examples and insights of the data in the data visualization narratively.

### 8.8 Reflection, limitation, and future work

The user studies showed that journalists liked the step-by-step, ordered instructions and would use them for their data stories by providing just one level of onboarding stage so as not to overwhelm readers and simplifying the semi-automatically generated onboarding messages. In our user experience evaluation, we focused on finding usability issues and validating the user experience of data journalists using the Flourish template with VisAhoi. The data journalists ranked the template's usage as "straightforward and clear". Using the template in an actual data visualization project and validating the visualization onboarding concept in more detail with recipients would be possible further steps. Further research might also explore the usage behaviors of the recipients of a data story. In particular: How often did the readers seek help? Did the reader gain a deeper understanding of the visualized data? Can the onboarding concept increase the time readers spend exploring the data visualization? Next, we present our lessons learned and the limitations of the design study.

#### 8.8.1 Lessons Learned

The visualization onboarding approach can support the narrative structure journalists are used to. The Flourish template (DR1-A1) is the first step toward an authoring tool to integrate self-explanatory onboarding instructions into a treemap visualization. Data journalists are experienced in using Flourish and can integrate visualization onboarding in their data visualization without any coding skills.

Furthermore, the data journalists highlighted that the visualization onboarding concept could support their narrative structure to present insights to readers. The semiautomatically generated onboarding messages are a blueprint to provide a basic version for treemap visualizations (DR-A2). As the authoring tool allows journalists to adjust the onboarding stages and onboarding messages, it is feasible for their purposes. In the "Analyzing" onboarding stage, we refer to the data as we, e.g., identify the maximum value. Nonetheless, the domain experts emphasized the need for a more simplified version of the messages not to overwhelm the readers. Besides, they highlighted the need to change the messages in such a way that they refer to the data and highlight insights.

Visualization onboarding design similar to Flourish annotations but different in a way to provide a self-explanatory and guided version. Flourish provides annotations integrated into the visualization, which are static and always visible. These can be used to point out insights. They are unsuitable for onboarding readers and do not meet the requirements of using tooltips (DR-R1) or providing context-sensitive, internal, and interactive onboarding instructions (DR-R2). Besides, the Flourish annotations are less customizable than our provided visualization onboarding template (DR-A3). We have learned from our user studies that the domain experts liked the design of the visualization onboarding approach, see Chapter 8.7, as it is based on tooltips as well as step-by-step and ordered onboarding instructions (DR-R1). They also mentioned that we meet the readers' mental model as they are used to guided tours from smartphone applications.

# 8. Design Study 1: Authoring Tool for (Data)Journalists to Integrate Visualization Onboarding in a Treemap Visualization

The onboarding approach is designed to be integrated into the data visualization using the data visualized in their onboarding messages to cover context sensitivity (DR-R2). By using the floating action button, readers can call the onboarding instructions at any exploration stage to search for instructions and explanations.

**Increase affordance of onboarding button.** Our domain experts highlighted the need to work on the affordance of the onboarding button to point the reader to it more strongly. The onboarding button is positioned next to the data visualization not to distract users. The domain experts noted that the readers might overlook the instructions. Therefore, they suggested opening the "Reading" onboarding stage and showing the onboarding messages to make the readers aware of the onboarding concept when approaching the data visualization.

The semi-automated generated onboarding messages are too abstract and need to be simplified. The semi-automated generation of onboarding messages can support journalists in setting up onboarding more efficiently as they function as a blueprint. Data journalists highlighted the need for more straightforward onboarding messages in the user experience evaluations not to overwhelm readers. We met this requirement by providing customization options for the data journalists by editing and deleting onboarding stages and messages. We provide a blueprint of what and how to explain a treemap effectively (compare results of a study by Stoiber et al. (2022b)).

Use cases for the usage of the Flourish template. The data journalists (P7) gave the example of using it to visualize the budget data of Austria and updating the data yearly to see changes in the data. The data journalists mentioned using the authoring tool to reuse data visualization and onboarding and updated the data. As the Flourish template can be saved and adjusted, it is easy to adjust the onboarding messages and visual markers to new insights and data.

#### 8.8.2 Limitations

The provided authoring tool (Flourish template + VisAhoi.js) is currently limited to a treemap visualization. However, experienced journalists in programming can exchange the treemap for any visualization technique. VisAhoi supports a generic type to support any visualization technique besides the bar chart, horizon graph, change matrix, scatterplot, or treemap visualization (see Section 8.6). Journalists can upload their data in the data upload (see Figure 8.7 (7)) of Flourish and customize the onboarding stages and messages for the selected data visualization.

Our target user group of journalists, especially those interested in data-driven journalism, is a small group of experts. Moreover, they have very tight schedules. Thus, seven data journalists participated in our qualitative evaluations.

### 8.8.3 Summary

We presented the results of a design study of the design, development, and evaluation of an authoring tool providing visualization onboarding for data journalists for a treemap visualization. The proof-of-concept implementation uses Flourish and the visualization onboarding JavaScript library VisAhoi. We validated the prototype based on data journalists' qualitative user experience evaluation. We have learned that:

- Data journalists can use the design of the semi-automated onboarding instructions integrated into the Flourish template in their daily routines.
- They liked the various customization possibilities in the Flourish authoring tool, e.g., adding, reordering, deleting, and editing onboarding stages and messages.
- The onboarding messages are too complex and need to be simplified to meet the audience's needs.
- The design of the onboarding buttons needs to be improved.
- In general, the Flourish template can support the narrative structure of data journalists to present insights to readers.

In the next Chapter 9, we present our second design study that integrates the developed VisAhoi library into a visual analytics (VA) tool to analyze high-throughput screening data for biomedical researchers.



# CHAPTER 9

# Design Study 2: Visualization Onboarding in Visual Analytics tools for Biomedical Scientists



This chapter starts with a general motivation for the design study about the selfexplanatory visualization onboarding approach for biomedical R&D. We describe our methodological approach (Section 9.2), the problem characterization & abstraction with the resulting requirements for the design and implementation in Section 9.3. Furthermore, we present the visualization onboarding concept and implementation (see Section 9.4). For the evaluation, we present in Section 9.5 a usage scenario showing the integration of the VisAhoi library into the visual analytics (VA) tool for analyzing high-throughput screening (HTS) data for biomedical researchers. Finally, we discuss the results and reflect on them (Section 9.6). The interview guide and screenshots are included in Appendix F.

# 9.1 Motivation, Background & Related Work

Biomedical research is an intensely data-driven discipline. Even before the big data and data analytics hype reached production companies (Industry 4.0) and other disciplines (trade, banking, retail, etc.), biological experiments resulted in large amounts of data that needed to be explored. This was mainly driven by introducing high-throughput techniques such as micro-arrays, whole genome sequencing, and mass spectrometry (Gehlenborg et al., 2010). The yielded data are highly heterogeneous regarding their origin, structure, and associated visualization techniques. Furthermore, it comprises information about various entities and concepts, including diseases, proteins, genes, chemicals, pathways, small molecules, ontologies, sequences, structures, and expression data (Pavlopoulos et al., 2015). All this data is analyzed and explored mainly using Visual Analytics (VA). The variety of applied visualization methods is enormous. It ranges from network and pathway visualizations over gene expression mappings (Kerpedjiev et al., 2018) and phylogenetic trees to conventional methods such as bar charts, line charts, boxplots, and more. The amount and importance of VA applications are enormous and constantly increasing. In most cases, these visualizations are applied to generate static images. Still, they can be linked and interactive, allowing active data exploration. Pavlopoulos et al. (2015) and Gehlenborg et al. (2010) provides a good overview of data visualization methods, applications, and frameworks in the field of biomedical research. Several competitive conferences and journals also reflect this in this field, such as BioVis or VIZBI. Nevertheless, the visualization of data and the subsequent interpretation of these visualizations are not always part of the domain experts' education and training. It is known that data literacy — including data visualization and interpretation — needs special training, even among highly trained specialists (Federer et al., 2016). Albeit humans are visual beings and visual representations are easier to understand than other forms of data representation, users still need to learn how to read and comprehend them. Biomedical scientists are highly trained experts in their domains. Still, most users are not experts in VA and have difficulties interpreting and working with novel visual representations or comprehending the characteristics of data analytics methods. This bears the risk of drawing wrong conclusions about the data or failing to tap into the power of data analytics, leading to frustration or rejection of essentially powerful data tools.

Visualization onboarding methods might support users in reading, interpreting, and extracting information from visual representations of data (Stoiber et al., 2022b, 2021, 2022d). Therefore, we aim to design visualization onboarding concepts for this domain to support them better in interpreting complex data.

We developed a JavaScript library, VisAhoi, to generate semi-automatically onboarding instructions and easily integrate them into VA tools. In the previous chapter 7, we presented the technical description of the JavaScript library.

In this design study, we show the applicability of the library by reporting on the design, development, and evaluation of the self-explanatory visualization onboarding approach

for the biomedical R&D use case. The VisAhoi library provides interactive onboarding for visualizations composed of in-place annotation anchors, a floating action button to call the onboarding, and semi-automatically generated onboarding messages as tooltips based on the data shown in the visualization. The onboarding concepts will be demonstrated by the example of a data visualization platform for high-throughput screening (HTS) data. The HTS analysis application holds data from HTS experiments where the binding kinetics of roughly 4 Mio. chemical compounds against potential drug targets are measured. The measurement for each compound is a time series curve that describes the binding behavior of the compound against the target over time. The 4 Mio time-series measurements are projected into a two-dimensional latent space with a multidimensional scaling approach. The 2D projection is visualized with a hybrid hexbin plot, which bins single points close to each other to girded hexagons to prevent overplotting, illustrated in Figure 9.2 (left). The density of single points behind a hexagon is encoded in color. Individual hexagons can be dissolved to show the actual compounds as scatter points behind them. The topology in the 2D space reflects the topology of the time series curves in the high-dimensional space.

Conversely, the tool offers a scatterplot, where each dot represents a single compound, as shown in Figure 9.2 (right). The attributes shown on the x and y axis and the color encoding are fully customizable by the tool user. Both visualizations are synchronized, which means that when selecting, e.g., a hexagon on the left, the respective points are highlighted in the scatter plot on the right and vice versa. Domain experts use those visualizations to search for clusters with compounds showing promising binding behavior.

Domain experts have been used to work with the initial scatterplot shown in Figure 9.2 on the right side for many years. The latent space embedding, shown on the left, is a new technique that needs to be discovered by most users of the VA tool. Therefore, onboarding is optional to explain what a scatterplot, by definition, is, but rather what the data means, how it was pre-processed, and how to correctly interpret it. Furthermore, domain experts are used to reading from static plots, where this VA tool offers many interaction possibilities to dive deep into the visualizations. Therefore, additional onboarding is needed to explain how to interact with and analyze the given visualization, at least for the first few times the user works with this tool.

Over the last year, we collaborated closely with domain experts in the field of biomedical R&D, thereby iteratively developing a better understanding of the problem and designing and evaluating the library and onboarding design to address it. Methodologically, we specifically relied on interviews with domain experts.

In summary, the main contributions of this design study are:

- a problem characterization and abstraction (see Section 9.3);
- the design of a self-explanatory visualization onboarding approach in Section 9.4,
- a usage scenario (SedImair et al., 2012) with the integration of the VisAhoi library in the VA tool to analyze HTS data in a hexbin plot and scatterplot visualization as described in Section 9.5;

• a set of general lessons learned that we derived from our project see Section 9.6).

In Chapter 6.2, we present the design of visualization onboarding exemplarily integrated into a VA tool to analyze high-throughput (HTS) screening data. We also report on the results of the cognitive walkthrough inspection with domain experts. For this design study, we used the same VA tool to show the applicability of our VisAhoi library in biomedical R&D.

## 9.2 Method



Figure 9.1: Timeline of the development, validation, and reflection phases

With a usage scenario, we show the library integration into the VA tool to analyze HTS data interactively. As typical for design studies, we followed a process of iterative design and evaluation process to address a particular domain problem by involving collaborators and users from the domain (Sedlmair et al., 2012). The six online interviews for the problem characterization and abstraction phase were chosen to reach out to our busy domain experts easily. In these interviews, we gathered data to understand the challenges in biomedical R&D regarding data, users, tasks, current practices, problems, and how onboarding was done so far. Furthermore, we validated the visualization onboarding concept through a clickable high-fidelity prototype and collected valuable feedback on our onboarding design. A timeline illustrating the development and the iterative validation phases is shown in Figure 9.1. The user research and evaluation phases were conducted between November and December 2021. Following the adjustment of the conceptual design and revision of the VisAhoi Javascript library, we started with integrating our self-explanatory onboarding approach in the VA tool to analyze HTS data interactively. We present the integration results as a usage scenario (Sedlmair et al., 2012). Finally, we reflected on integrating the library into the VA tool and discussed lessons learned.

## 9.3 Problem Characterization and Abstraction

In the first phase of the design study, we conducted interviews with biomedical researchers to collect data regarding the domain experts' needs and challenges and current practices regarding visualization onboarding.
#### 9.3.1 Method & participants

The six interviews (see Table 9.1 for demographic information about our domain experts') took place online via Zoom calls from 2021/10/21 to 2021/11/25. The interviews had an average duration of 35 minutes. Notes were taken during the interviews and the audio using the Zoom recording. We shared our screen showing the interview guide to provide a structured way to review the questions. The interviews were held in German.

Table 9.1: Demographic information of interviewees. Gender: f = 3, m = 3)

Participant Gender		Gender	Educational background	Job description/tasks	
	P1	female	molecular biology	lab head at Boehringer Ingelheim (early phase of drug development)	
	P2	male	biology	head of research group at Boehringer Ingel heim (early oncology research)	
	P3	female	molecular biotechnology	lab scientist at Boehringer Ingelheim (planning and conducting experiments)	
	P4	female	biology	senior scientist at Bayer (substance testing and characterization from HTS to clinical sub- stance)	
	Р5	male	biochemistry	scientist at Bayer (analysis of high-throughput screening and characterization)	
	P6	male	biology focused on informatics	senior scientist at Bayer (early research (screening) with a focus on chemoinformatic and compound logistics)	

The questionnaire was structured as follows (see Appendix F). (1) After welcoming the interviewees, we asked fundamental questions regarding using interactive visualizations and background and current job descriptions. (2) The second part was designed to evaluate the current visualization literacy. Furthermore, we collected data about familiarity and experience in constructing visualization techniques. We also asked the interviewees which tools they use by providing a list of visualizations and visual analytics tools. (3) The central part of this interview was to get an idea of how onboarding is currently happening and where it is needed. As onboarding concepts should support the users and not distract them, we wanted to know where and when onboarding is needed. Therefore, we designed the next part of this interview to gather information regarding visualization tools, which materials were used, and which problems occurred. (4) Finally, we gathered feedback on the visualization onboarding concept. We observed the subjects using the clickable prototype on their own device and collected their comments using thinkingaloud. The experts used the clickable high-fidelity prototype, which can be accessed here: http://bit.ly/3iTNpRm. This clickable prototype was also used in the cognitive walkthrough inspection. For the usage scenario presented in this chapter, we utilized the improved visualization concepts using the VisAhoi library (Chapter 7).

For the analysis of the interview data, we used an open coding strategy for the analysis

of the transcripts. We translated all comments from German to English.

#### 9.3.2 Results

The data analysis regarding familiarity with different visualization techniques indicates that business charts like bar charts, line charts, and pie charts are the most familiar. More than 80% of the participants' rated bar charts and line charts as "extremely familiar". Pie charts and scatter charts/bubble charts show a level of familiarity of 60%, followed by heatmaps and dendrograms (50% extremely familiar). Boxplots, violin plots, parallel coordinates plots, treemaps, and node-link network plots are rated under 50%. Radar/sunburst charts and the Sankey diagram are the least familiar ones. All interviewees said they had already created bar charts, line charts, pie charts, and heatmaps (100%). Any of the participants did not create Sankey diagrams. 83% of the participants said they also created scatter charts, boxplots, and dendrogram/tree visualizations. Node-link network plots, parallel coordinates plots, violin plots, radar/sunburst charts, and treemaps are the visualization techniques created the least.

The results indicate that the domain experts have a high level of visualization literacy (familiarity and the construction of visualizations). They said they need visualization in their daily work for different tasks such as project planning, explorative data analysis, and presentation purposes. Visualization techniques are used less to visualize network data (e.g., node-link, Sankey diagrams).

As domain experts exhibit a high level of visualization literacy, onboarding to explain, the visual encoding of a specific visualization is not that important. They emphasized the need for onboarding concepts explaining the data and interaction concepts (see knowledge gaps (Stoiber et al., 2019a)). We asked the interviewees where and when onboarding was needed. They answered to have in-app onboarding (where) and on-demand (when). One participant said: "So what I found very exciting is your comment from above, with in-app help. I mean, it is often the case that you have the help files stored in a structured way, and you can navigate through them, but I would say in-app interactive. It is exciting, although I cannot think of how I would imagine that right now if I were to run a program like that and say, oh, try this, do this, try this, I would find that exciting." Another interviewee emphasized onboarding on-demand: "[...] I want to tell myself; I need help, and then just contextually." Another comment to on-demand onboarding: "I would prefer on demand. In the tool itself, or possibly linked to an external source."

We also asked the interviewees how they usually learned new visualization tools and which materials were used. All the interviewees said that they usually use screenshots and text. 83% indicated using In-app help like in Excel. Tutorial videos were used by half of the participants. Interviewees also emphasized the "help of colleagues" as one important source to get help with a particular task or problem. One participant commented: "Yes, trial and error doesn't always lead to the desired success. [...] You have to use other sources like tutorials and videos. The formulation of the question is often a problem [...]. And it is time-consuming." Participants also talked about how a video is not that much

liked: "If I have to? Videos too, but I don't like them." They rated tutorials with textual instructions and screenshots as more valuable: "I think tutorials are better." Importantly, users also mentioned that they first start with Trial and Error and then continue to search for tutorials and videos or ask colleagues for help. Also, google search, for example, was mentioned in the interviews as an essential source. General courses, for example, using Excel, are less valuable as they cover only a general view of the tool. They do not use data or work on problems domain experts have as their domain is quite specific. For learning in-house tools, interviewees reported that they participate in courses. There is also a service department to contact for any questions about in-house tools.

The thinking-aloud session at the end of the interviews showed that the participants liked the onboarding messages separated into the three onboarding stages: *reading*, *interacting*, and analyzing the chart. The domain experts highlighted the intuitive design and interaction concept and design of the visualization onboarding. They also stressed that the integration of onboarding in the VA tool and interface is appropriate. Subjects mentioned providing the possibility to move the onboarding messages around not to overlap essential data points in the scatterplot. They emphasized that onboarding should not distract during data exploration.

Additionally, they highlighted the need for onboarding for tool-specific features besides visualization. They also mentioned working on the understandability of the onboarding messages by using wording established in the domain. While observing the subjects interacting with the clickable high-fidelity prototype, we realized they first needed help finding the onboarding button. They mentioned searching for it in the main menu bar, as they are used to the help in Microsoft Excel. Participants also mentioned that they would not use numbers in the in-place annotation anchors not to limit the users. In summary, they also commented on the usage of onboarding. They highlighted that they would use the onboarding instructions for first use only.

In conclusion, the interviews show that the domain experts using VA tools in their daily routine have a high visualization literacy (interpretation and creation). It also shows that the integrated and context-sensitive manner of the visualization onboarding is appropriate for this domain, as they want to interact with the scatterplot while having the onboarding messages open. Visualization onboarding, integrated into the HTS screen analysis tool, should describe the underlying data as well as the features of the tool itself. Additionally, more design possibilities must be considered to meet the users' mental model when opening onboarding. Also, short and easy-to-understand onboarding messages are vital.

#### 9.3.3 Design Requirements

We identified three design requirements (DR) based on the problem characterization and abstraction (interviews):

#### DR1 Explain the underlying data and provide data-related onboarding instructions: The interviewees emphasized the importance of explaining the under-

lying data as it is new to them. The VA tool offers many interaction possibilities, which need to be explained to the domain expert.

- DR2 Support domain experts with tooltips and textual explanations: The domain experts mentioned providing on-demand onboarding instructions and tutorials. The VisAhoi library is based on in-place annotation anchors and tooltips with textual instructions integrated into the visualization (cf. Chapter 7, and Chapter 6). From other studies (Stoiber et al., 2022b, 2021, 2022d), we know that the onboarding concept should provide different ways of catering to different knowledge gaps (Stoiber et al., 2022c). Therefore, we allow users to self-explore and be guided through the onboarding instructions to cover different knowledge gaps and not distract the user while exploring the data in the visualization (cf. Chapter 6).
- **DR3** Provide context-sensitive onboarding which can be called up during exploration: Context-sensitive means to provide onboarding instructions at a specific point in the current state of the tool (Stoiber et al., 2022c). An onboarding system integrated internally into the visualization helps users not jump back and forth between two different systems. Additionally, onboarding should be called up while using the visualization for insight generation. This design requirement is already included in the VisAhoi library.
- DR4 Position the onboarding button in the toolbar to easily approach it when seeking help: The domain experts mentioned positioning the onboarding button in the toolbar provided by the Plotly library to meet the user's mental model.

## 9.4 Conceptual Design & Implementation

This section presents the major onboarding design based on the design decisions (explained in detail in Chapter 6 and Chapter 7) and the design requirements (DR), presented in Section 9.3.3, for the integration into the VA tool for HTS data analysis. The interface design of the VA tool for HTS data analysis is described in the usage scenario Section 9.5.

- The floating action button (onboarding button) is the main navigation element to access onboarding instructions. Users can independently access onboarding to different visualizations simultaneously by activating the onboarding mode in the plotly toolbar at the top right of the visualization (DR4), illustrated in Figure 9.2 (1). The floating action button then appears next to the data visualization. We position it to the bottom right next to the hexbin plot (left) and scatterplot visualization (right) (see Figure 9.2).
- During the interviews with domain experts (results presented in Section 9.3.2), we realized that the users had difficulty finding the onboarding button in the interface,

as we positioned it next to the visualization. Users of this VA tool for HTS data analysis are used to the toolbar provided by the Plotly library. Therefore, we positioned it there and used a shiny pink color to attract attention. The appearance of a new element when activating the onboarding guide the users and provide affordance.

- The onboarding design uses textual instructions (onboarding messages) as tooltips in combination with in-place annotation anchors (DR2) to indicate the connection between the text elements and the visual encoding, illustrated in Figure 9.2.
- We designed the onboarding to allow users to be (1) guided through the different onboarding messages or (2) self-explore it and seek help at any stage of their interpretation and interaction with the visualization. Therefore, we included a step-by-step navigation widget representing the number of in-place annotation anchor/onboarding messages and arrow icons that allow the user to browse through the onboarding message and receive guided help (see Figure 9.2). Users get help for a specific element or part of the visualization, e.g., interaction, or visual encoding, when clicking on a specific anchor. That way, we can provide a self-exploratory experience. A toggle button below the float action button deactivates the step-by-step navigation widget.
- We provide **concrete onboarding messages**, which we extract from the declara-• tive grammar of Plotly is and adjust to the domain experts' needs. The onboarding messages for the scatterplot visualization were developed in close collaboration with a domain expert and product owner of the VA tool. We structured the onboarding messages in three default onboarding stages: Reading, Interacting, and Analyzing. The onboarding stages are the three circular buttons when clicking the floating action button, illustrated in Figure 7.4 (A). The first onboarding message in the *Reading* onboarding stage explains the data visualized in the hexbin plot and scatterplot, which the users emphasized in the interviews (DR1). Here is the example for the scatterplot visualization: Latent space: The VA tool uses variational autoencoders (VAE), a kind of neural network model, to derive a lower dimensional representation (points in 2D) of high dimensional input data (kinetic curves). This lower dimensional representation called **latent space**, is shown here. During training, the VAE model has been constrained to push data of neutral control samples as close as possible toward the origin of the latent space, thus shaping the latent representation of the data set. The onboarding messages are developed in close collaboration with an expert.
- The onboarding concept allows users to **interact simultaneously** with the underlying **visualization** and the **onboarding** (DR2), illustrated in Figure 9.2 (B). Onboarding messages can be relocated by drag&drop not to cover essential visualization elements.

# 9. Design Study 2: Visualization Onboarding in Visual Analytics tools for Biomedical Scientists



Figure 9.2: Visualization onboarding concept integrated into the VA tool for analyzing HTS data — side-by-side a hexbin plot (left) and scatterplot (right). (A) The navigation comprises different elements to provide a guided tour and a self-exploratory experience. First, in the toolbar of both Plotly visualizations, the onboarding button is positioned to activate the onboarding and shows the floating action button (3). By clicking on the floating action button (1), the menu expands outward from a central point and reveals three onboarding stages: Reading, Interacting, and Analyzing. We decided to integrate a step-by-step navigation widget (3) that represents onboarding messages (4) as well as arrow icons to navigate through it (3). The onboarding messages are designed as tooltips in combination with in-place annotations anchors (4) to indicate the connection between the textual instructions and the visual encoding. The user can open both onboarding instructions for both visualizations at the same time, as illustrated in (B).



Figure 9.3: **Overview of VisAhoi components.** The VisAhoi adapter extracts the information from Plotly.js' runtime object and DOM nodes and passes them to the VisAhoi core. The core generates the basic onboarding stages and messages.

#### 9.4.1 Implementation

In general, VisAhoi provides interactive onboarding for visualizations generated from high-level declarative visualization grammars such as Vega-Lite (Satyanarayan et al., 2017), Plotly.js (Plotly.js, 2023), and ECharts (Li et al., 2018) (cf. Chapter 7). VisAhoi is implemented using Svelte (Svelte, 2023), which compiles Svelte-specific code to plain JavaScript. The compiled source code is framework-agnostic and can be used in combination with other front-end frameworks such as React (React, 2023), Vue (vue, 2023), or Angular (Angular, 2023). In Chapter 7, we presented the technical design and functionality of the VisAhoi library in more detail.

Figure 9.3 shows the main components of the onboarding generation process for the hexbin plot and scatterplot visualization in the VA tool.

- 1. The scatterplot and hexbin visualization is created from a high-level declarative visualization grammar by the visualization library Plotly.js (Plotly.js, 2023). To render the visualization, Plotly.js creates a runtime object from the grammar and adds computed values about the chart (e.g., the position of data items).
- 2. The VisAhoi adapter for Plotly.js extracts information such as the chart title or positions and values of data items for the onboarding from the runtime object and the DOM nodes of the visualization.
- 3. The VisAhoi core generates basic onboarding stages and messages by filling the extracted information from the adapter into predefined templates. Onboarding stages that do not have messages are filtered out.
- 4. Developer/visualization designer can customize the onboarding by adding, editing, and deleting the provided onboarding stages and messages.

5. VisAhoi generates the onboarding user interface based on the customized onboarding stages and messages. It integrates the in-place annotation anchor at the given position in the visualization and renders the onboarding.

VisAhoi is currently extensible to multiple visualization libraries and supports onboarding for scatter charts, horizon charts, bar charts, heat maps, and treemaps for each supported visualization library.

In the next section, we explain the applicability of VisAhoi by discussing the integration of the library into the VA tool for biomedical scientists analyzing HTS data. We describe the general design of the VA tool and explain how we integrated the library into the VA tool to onboard domain experts. We also list the concrete onboarding messages. Furthermore, we discuss the limitations and reflect on them in Section 9.6.

## 9.5 Usage Scenario on Integrating into a VA tool for Biomedical Scientists analyzing HTS data



Figure 9.4: VA tool for analyzing HTS data — side-by-side a hexbin plot (left) and scatterplot (right) with integrated VisAhoi library to provide onboarding instructions for the domain experts.

As explained in Section 9.1, biomedical scientists use this VA tool to analyze highthroughput screening (HTS) data (see Figure 9.2). The domain experts use this tool mainly to search for clusters with compounds showing promising binding behavior. We present this usage scenario (SedImair et al., 2012) showing the applicability of the VisAhoi library in this domain.

The VA tool contains data from HTS experiments that measure the binding kinetics of approximately four million chemical compounds against potential drug targets. Each compound is represented by a time series curve that describes its binding behavior against the target over time. A multidimensional scaling approach is used to project the four million time series measurements into a two-dimensional latent space, which is visualized using a hybrid hexbin plot (see Figure 9.4 (A)). This visualization bins single points close to each other into hexagons to prevent overplotting and encodes the density of single points in color. The topology of the 2D space reflects the topology of the time series curves in the high-dimensional space.

Additionally, the tool offers a scatterplot, illustrated in Figure 9.4 (B), where each dot represents a single compound, and the attributes shown on the x and y-axis, as well as the color encoding are fully customized by the user. The plots are implemented as multiple-coordinated views, which means a selection in one of the plots is highlighted in the other. Domain experts use these plots to search for clusters of compounds showing promising binding behavior.

Most domain experts are accustomed to working with static scatterplots on the right side (cf. results of problem characterization and abstraction in Section 9.3). At the same time, the latent space embedding on the left is a relatively new and unfamiliar technique for the users (DR1). Therefore, onboarding is necessary to explain the data, preprocessing methods, and how to interpret the visualization correctly. Additionally, this VA tool offers various interaction possibilities that require onboarding to effectively explain how to analyze and interact with the visualization.

Although having semi-automatically generated onboarding messages for new visualization types is convenient, in this particular use case, most of the basic onboarding messages provided by the library were too general because domain experts already knew how to read scatterplots in general (cf. results of problem characterization and abstraction in Section 9.3). Therefore, the primary focus of the onboarding was to explain how the data was pre-processed and displayed, especially in the hexbin plot, and how the plot's interaction work. Hence, it was crucial that the onboarding messages were customized and developers added additional ones to meet the domain experts' needs.

The developer of the VA tool did the technical integration after the user interface was done. The VisAhoi Library provides interactive onboarding for data visualizations generated from high-level declarative visualization grammars such as Vega-Lite (Satyanarayan et al., 2017), Plotly.js (2023), and ECharts (Li et al., 2018) (see Chapter 7). As both data visualization are developed with Plotly.js (2023), it was easy to integrate them into the VA tool.

In the following, we explain how we integrated the VisAhoi.js library technically and present the onboarding messages in detail.

**Technical Integration & Onboarding Messages:** The developer integrating VisAhoi in the VA tool was an experienced full-stack developer who gained experience in various languages over the past years. She is very experienced in for this use case relevant programming language, Typescript.



Figure 9.5: Interacting with the chart onboarding stage with its in-place annotation anchors and tooltips. The onboarding instructions explain how to interact with the hexbin plot (left) and scatterplot (right). We integrated small icons that relate to the interaction bar provided by Plotly.js. Both visualizations are highly interactive and users need instructions on how to interact with the charts.

In this particular usage scenario, the pre-built (default) onboarding messages provided by the VisAhoi library for a basic understanding of scatterplots were deemed unnecessary, as domain experts already have this knowledge. Instead, the focus was on explaining how the data was processed and presented, particularly in the hexbin plot and the visualization's interactive features. Therefore, it was crucial that the developers could customize the onboarding messages and add new ones as needed. To accomplish this, the customizeOnboardingMessages() function of the ahoi() function was examined to filter out irrelevant messages and focus only on onboarding messages related to visualization interaction. These messages were then customized to suit the use case and additional messages were created for the *Reading the chart* and *Analyzing the chart* onboarding stages. All messages were added to an array and returned via the customizeOnboardingMessages() callback function in the ahoi(), which will be passed to VisAhoi later on.

We packed all this functionality into a function in our code called customizeHexbinplotMessages() and customizeScatterplotMessages(), which can be re-used anytime.

To initialize the onboarding for our two visualizations, we call the ahoi() function, which is exposed by the VisAhoi library, and pass the visualization type (scatterplot in both cases), the runtime object, and the function that returns the customized onboarding messages. ahoi() returns the instance of the onboarding, enabling us to show and hide the onboarding as the user interacts with the application. We added a button to our user

interface that hides or shows the VisAhoi onboarding on click (see pink "?" icon in the top right corner of the visualizations in Figure 9.2). This way, new users of the tool can easily access the onboarding on demand, and power users do not feel disturbed by it.

The onboarding instructions for both visualizations for the onboarding stage *Interacting* with the chart are illustrated in Figure 9.5. We used the same small icons as the Plotly.js interaction bar to make it easier for domain experts to relate to it when interacting with the visualization. For "Zooming" and "Dissolving hexagons" in the hexbin plot, we phrased the following onboarding message:

- Zoom: With those buttons or with the mouse wheel, you can zoom in and out in the visualization to get a more detailed view of the data.
- **Dissolving hexagons**: Select any hexagons in the visualization and press this button to dissolve them and see the real data point behind them.

We integrated the following onboarding message to the hexbin plot to explain the underlying data and explain how to read the corresponding visualization (Reading the chart). The other screenshots and onboarding messages are attached to Appendix F.1.

- **Hexbinplot**: This hexbin plot shows a projection of 877,048 compounds. As this high amount of points would cause overplotting in a normal scatterplot, the data points are binned into gridded hexagons. Each hexagon represents several points, where the distribution of points is encoded in color.
- **Reading the chart:** A dark green color indicates a high density of points, whereas the medium green color reflects a medium amount of points. This light green color represents very sparse areas.
- **Reading the chart:** It is possible to dissolve individual hexagons (explained in Interacting) to see the real data point behind them. when doing so, the blue color represents the substance samples, gray represents the neutral samples, and red represents the positive samples. Black points have been excluded.
- Latent space: The VA tool uses variational autoencoders (VAE), a kind of neural network model, to derive a lower dimensional representation (point in 2D) of high dimensional input data (kinetic curves). This lower dimensional representation, called latent space, is shown here. During training, the VAE model has been a constraint to push data of neural control samples as close as possible towards the origin of the latent space, thus shaping the latent representations of the data set.
- Dots of resolved hexagons: Experiments are performed in microtiter pates (MTP). Thus the input data set consists of one experimental observation (kinetic curve) for each well of the MTP. Each dot corresponds to a 2D representation of these higher dimensional kinetic curves. Notice: Each substance well contains two distinct chemical substances at once.

• **Position along the axis:** Measurements of neutral control samples (e.g., DMSO) are constrained towards the origin of the latent space. Therefore, the further a point is from the origin, the more dissimilar it's corresponding experimental observation is to a neutral control sample.

Lessons Learned: In summary, integrating VisAhoi into the VA tool was quite simple and straightforward. Adding the basic onboarding to visualizations is easy and can be done with a few lines of code. Having basic onboarding messages was very helpful as it allows the developer to see what the onboarding could look like. This is a good starting point and makes it easier to decide what to customize in the next step. As the default onboarding messages were not enough for the use case, additional ones have been added, which was not an issue as they can simply be passed to the ahoi () function as an additional parameter. To find the right position for the anchor of new messages, the developer has to dive deeper into the runtime object of the visualization. As the visualizations of this VA tool hold huge amounts of data, canvas support (discussed in section 7.7.2) would be helpful to annotate individual data points better (e.g., min/max values along an axis), which is currently only supported for SVG. Nevertheless, the developer placed all needed onboarding messages properly in the visualization.

### 9.6 Reflection, Limitations & Lessons Learned

The interviews with domain experts in the problem characterization and abstraction phase (see Section 9.3) reveal a strong understanding of interpreting and creating visualizations. Therefore, the onboarding process for the HTS screen analysis VA tool focuses on explaining the underlying data and the features of the VA tool, especially the interaction possibilities, as it is new to the domain experts. We provide context-sensitive onboarding, presented in Section 9.4, by integrating our developed VisAhoi library into the VA tool. Our approach is suitable for this domain, as users want to interact with the visualizations while having the onboarding messages available at any exploration stage (cf.Section 9.3). The usage scenario (SedImair et al., 2012) shows the technical integration of the VisAhoi library in a VA tool for analyzing HTS data. The VA tool was developed, and the VisAhoi library was integrated later into the tool to onboard the biomedical scientist to the new visualization technique, the hexbin plot, and the interaction possibilities of both visualization techniques.

The visualization onboarding design meets the requirements derived from the problem characterization and abstraction phase. In the following paragraphs, we go into detail and reflect on it.

The floating action button (onboarding button) is the primary way to access onboarding instructions. The onboarding button appears next to the data visualization, positioned to the bottom right of the hexbin plot and scatterplot visualization. Users can simultaneously access onboarding instructions for different visualizations by activating onboarding mode in the Plotly toolbar (DR4). Users had difficulty finding the onboarding button next to

the visualization in our interviews, so we positioned it in the toolbar provided by the Plotly.js library and used a shiny pink color to attract attention. The appearance of a new element when activating the onboarding also guides users.

The onboarding design combines textual instructions as tooltips with in-place annotation anchors to indicate the connection between the text elements and the visual encoding (DR2). The onboarding concept allows users to interact simultaneously with the visualization and the onboarding. Onboarding messages can be relocated by drag&drop not to cover essential visualization elements, which is important when integrated into a dense VA tool.

Users can be guided through the onboarding messages or self-explore them and seek help at any stage of their interpretation and interaction with the visualizations. A step-by-step navigation widget and arrow icons allow users to browse the onboarding messages. Onboarding messages are extracted from the declarative grammar of Plotly.js and adjusted to the domain experts' needs (DR3) by referring to the underlying data (concrete onboarding messages). We realized during the interviews (cf. results of problem characterization and abstraction in Section 9.3) with the domain experts that they are experienced in reading visualization, in particular static scatterplot visualizations (DR1). At the same time, the latent space embedded in the hexbin plot is relatively unfamiliar to the domain experts. Therefore, explaining the data, preprocessing methods, and how to interpret the visualization correctly is necessary. Additionally, this VA tool offers various interaction possibilities that require onboarding to effectively explain how to analyze and interact with the visualization.

#### 9.6.1 Lessons Learned

**Biomedical scientists have a high visualization literacy.** The interviews show that the domain experts use VA tools in their daily routine, and therefore have a high visualization literacy (interpretation and creation). Nevertheless, integrated into the HTS screen analysis tool, visualization onboarding is needed as there is a new visualization technique- the hexbin plot- integrated, and the data pre-processing is new to the domain experts. Additionally, the visualizations are highly interactive and provide linking and brushing, which needs to be explained to allow users to explore the data effectively.

VisAhoi provides different customization possibilities for the developer to make it easy to integrate it into a given VA tool. VisAhoi is a Typescript library designed with a developer's experience in mind. Adding semi-automatically generated onboarding to visualizations can be done with a few lines of code (see Listing 1) (Objective 7.3: Integratable).

Besides, The VisAhoi library allows the developer to customize the onboarding instructions and therefore make it easy to integrate it into a given VA tool (see Chapter 7, Section 7.3). The developer can rephrase the semi-automatically generated onboarding messages and stages, customize the position of the onboarding button next to the visualization, change the color and icons of onboarding stages, sort the onboarding stages, and place the in-place annotation anchors to elements in the visualization.

The customization possibility of VisAhoi allows developers and visualization to flexibly adjust the onboarding concept to meet the user's needs of the VA tool. The usage scenario showed that customization possibilities, e.g., adjusting onboarding messages and stages, are essential for domain experts' needs to adjust the onboarding concept. Biomedical scientists are used to interpreting scatterplot visualizations illustrated in Figure 9.4 (right). Therefore the onboarding focuses on explaining the interaction and how to read it by including. The developer provided onboarding stages for *Reading the chart* and *Interacting with the chart* only. Unlike the hexbin plot onboarding, the developer provides onboarding instructions to all three onboarding stages: Reading, Interacting, and Analyzing.

VisAhoi allows the developer to provide concrete and targeted onboarding messages for the particular target group. VisAhoi provides onboarding messages semi-automatically, which can be used as a blueprint. The developer can easily adjust the onboarding message to meet the domain experts' needs. As described in the usage scenario, we provided concrete. We targeted onboarding messages for biomedical scientists, explaining the data, pre-processing methodology, and how to read and interact with the unfamiliar hexbin plot and the scatterplot visualization. Further research should be carried out to evaluate the phrased onboarding messages with domain experts using this VA tool for HTS data analysis in the wild. Additionally, user experience evaluations with domain experts need to be conducted to observe the behavior with the onboarding instructions.

#### 9.6.2 Limitations

The usage scenario shows the applicability of the integration of the VisAhoi library technically from the point of view of a developer. We are still determining if the onboarding messages for the hexbin plot are phrased correctly for the domain experts using the VA tool. During our interviews in the problem characterization and abstraction phase, we already gained some feedback on the onboarding messages and general onboarding concept for a scatterplot visualization (cf. results presented in Section 9.3). As the onboarding button should be easily accessible and attention-grabbing, we positioned it into the toolbar of the Plotly.js visualization to meet the mental model of the domain experts. More research is needed to validate if users of the VA tool can seek help and understand what the shiny pink onboarding button means. Future studies that explore the different levels of affordance might be beneficial.

As mentioned in Chapter 7, the VisAhoi is designed to work with visualization libraries that use SVG rendering. It allows for a more manageable selection of anchors using the DOM structure. However, canvas rendering, which has better performance, does not provide this structure, and the anchor coordinates must be obtained from runtime

objects, which often lack this information. Thus, we positioned the onboarding message manually to the hexbin plot to provide the instruction for the position and meaning of the hexagon, as we could retrieve the DOM element due to the canvas restriction. In the future, we plan to investigate the runtime objects of the three selected visualization libraries that use canvas rendering. This approach could be extended to other libraries that use canvas rendering, such as the widely used Chart.js (2020) library.

### 9.7 Conclusion

Biomedical research generates large amounts of highly heterogeneous data, and their analysis and exploration mostly rely on Visual Analytics (VA). However, the interpretation of VA is only sometimes part of domain experts' education and training, which increases the risk of drawing incorrect conclusions and missing out on the benefits of data analytics. To address this, visualization onboarding methods can support users in interpreting complex data. The VisAhoi JavaScript library is developed to generate semi-automated onboarding instructions and integrate them into VA tools, demonstrated through a design study on high-throughput screening (HTS) data analysis. We performed six interviews for the problem characterization and abstraction phase with the domain experts. A usage scenario describes how VisAhoi integrates into a VA tool for HTS data analysis for biomedical scientists. We learned that integrating the VisAhoi library was easy for the developer as the visualization in the VA tool was developed in Plotly, is, and our library supports this visualization library. The design of the visualization onboarding concept is appropriate to be integrated into a finished VA tool, as it relies on a floating action button, which can be positioned easily next to the visualization. The developer can position in-place annotation anchors to provide highly customizable onboarding. Onboarding messages can be called up at any stage of the domain experts' exploration with the hexbin plot and scatterplot. During the design study, we realized that it is crucial to start thinking about integrating onboarding into a VA tool early while designing the VA tool. Domain experts also highlighted providing onboarding for general user interface elements, especially for the first use of the VA tool. Design Actions can help developers and visualization designers to design and integrate visualization onboarding into VA tools.

In the next section, we provide design actions from our research (Chapter 5, Chapter 8, and Chapter 9).



## Part IV

# Generalization of Results & Outlook



# CHAPTER 10

# **Design** Actions



**Design Actions** 

This chapter presents design actions (De Bruijn and Spence, 2008) for designing visualization onboarding concepts that can be integrated into visual analytics (VA) tools and data visualizations. This chapter is based on our published work (Stoiber et al., pted). We critically reflect on the results and the process within the two design studies performed for this thesis (Chapter 8 and Chapter 9). We derived implications, which we formulated as design actions based on the lessons learned. These design actions should help visualization designers and developers better understand how to design and integrate onboarding into VA tools.

#### 10.1 Method

We conducted an in-depth reflection on the results and the process of the two design studies (Chapter 8 and Chapter 9).

The design actions comprise structured guidelines on integrating and designing visualization onboarding concepts along the framework by Spence (De Bruijn and Spence, 2008), including a (1) **description** clarifying the **title**, the (2) **effect** of the design action towards onboarding, advantages, and trade-offs ((3) **upside** and (4) **downside**), (5) **issues** describing the application of the design action, and (6) theory a reference to the cognitive **theory** is provided.

The presented design actions are structured along our guiding questions from our design space (see Chapter 2): WHO is the user, and which knowledge gap does the user have? Which parts of a visualization need to be explained? How to phrase onboarding instructions? HOW, WHERE, and WHEN is visualization onboarding provided?

#### 10.2 Design Actions

WHO is the user, and which knowledge gap does the user have? Which parts of a visualization need to be explained? How to phrase onboarding instructions?

ID DA1

Title

e Be aware of different knowledge gaps of users

Description When designing and integrating onboarding instructions into a visual analytics (VA) tool, the different knowledge gaps among your users and their needs and prior knowledge need to be considered. Novice users may need more or a different kind of support as expert users in interpreting data visualizations (Friel et al., 2001; Grammel et al., 2010). In our design space presented in Chapter 2 and previous work (Stoiber et al., 2022a,c), we present different knowledge gaps which should be considered when designing visualization onboarding: *domain knowledge* (e.g., vocabulary and concepts of a particular target group), *data knowledge* (understanding the particular datatype), *visual encoding* (understanding the visual mapping) and *interaction knowledge* (for

performing tasks and understand relations in the data), and *analytical* knowledge (knowledge of different automated data analysis methods).

- Effect This affects the type of onboarding messages, the number and order of onboarding stages, and for which parts of the VA tool onboarding are needed. See related design actions DA3.
- Upside Targeted onboarding messages can help users of the VA tool to have a better user experience. By collecting data during the domain problem analysis, visualization designers can benefit later when designing onboarding instructions.
- Downside To phrase and design onboarding based on the user's needs is demanding. Gathering data during the domain problem analysis and interaction and visual encoding phase (Munzner, 2009) is rather timeconsuming.
- Issues Consider the user's prior knowledge by gathering information in your user research phase.
- Theory/Evidence Prior knowledge about graphs influences the level of graph comprehension (Curcio, 1987; Freedman and Shah, 2002). Kosslyn (1989) pointed out that a visualization reader's previous knowledge of the data visualization and content is one of the major factors when interpreting data visualizations. Therefore, when designing visualization onboarding, it is important to consider the user's prior knowledge and knowledge gap. In our descriptive model of user assistance (Stoiber et al., 2022a), we propose to design and adjust visualization onboarding and guidance (Ceneda et al., 2017) based on the state of the knowledge of the user. Our interviews with domain experts (cf. Chapter 9) showed that they prefer onboarding instructions targeted to their prior knowledge and experience.

#### ID DA2

Title

#### Explain how to read, interact, and analyze the data visualization

Description Visualization onboarding instructions should explain how the data visualization is read, how to interact with it, and what insights you can gain (analyzing the chart). *Reading the chart* explains the general visual encoding (e.g., size, color), *Interacting with the chart* the possible interactions with the data visualization (e.g., zooming, selecting, or dissolve), and *Analyzing the chart* - guiding the reader towards further insights (e.g., making comparisons).

Effect	The structure of onboarding messages allows users to seek help for different aspects of the onboarding based on their prior knowledge and current knowledge gap (see Design Space dimension in Chapter 2).		
Upside	<ol> <li>Onboarding instructions along the structure of reading, interacting, and analyzing the visualization cover essential aspects of a visualization, which the user has to understand to interpret and interact with it successfully.</li> <li>Users can select which onboarding instructions are needed based on their prior knowledge or current knowledge gap.</li> <li>Besides, smaller chunks (microlearning) of information are less overwhelming for users (Hug and Friesen, 2009).</li> <li>The separation provides a self-exploratory experience for the user (see the related design action DA5).</li> </ol>		
Downside	(1) If the structure of onboarding instructions is too detailed, it could annoy and overwhelm the user. (2) The structure could feel too restricted for the user. (3) Instructional split-attention (Ayres and Sweller, 2005) could occur when users have to split their attention between onboarding messages assigned to the structure of reading, interacting, and analyzing.		
Issues	For some data visualization tools, providing more onboarding instruc- tions that expand the structure of reading, interacting, and analyzing can be relevant. For example, providing onboarding instructions re- garding the data pre-processing might make sense as it is new to the user. The structure can also be dissolved and the wording adjusted if the use case requires it.		
Theory/Evidence	The structure (reading, interacting, and analyzing) also reflects the stages of comprehending (Friel et al., 2001) (1) reading the data (i.e., finding individual data values), (2) reading between the data (i.e., finding relationships between the data), (3) going beyond the data (i.e., interpreting the data, developing hypotheses about the data). Whereas the stages of comprehending reading between the data and going beyond the data are covered by the category "Analyzing the chart". Small information units can support learning (microlearning) (Hug and Friesen, 2009). Besides, our studies show that users liked the structured way of the onboarding instructions and the self-explanatory experience of it (cf. Chapter 4).		

#### ID DA3 Title Use concrete onboarding instructions where possible & needed Description When phrasing onboarding instructions, refer to the underlying data visualized and provide examples if possible and needed by the users (see design action DA1) (concrete onboarding instructions). However, the onboarding instructions should be phrased as easy to understand, and short sentences should be considered (Shneiderman, 2022). For guidelines for writing understandable sentences, see (Janice Ginny, 2014). Furthermore, the wording of the specific domain in the onboarding is designed needs to be considered when phrasing onboarding instructions. Effect Using data/concrete examples for onboarding instructions the user is working with can enhance the understanding and support to-the-point assistance targeted to the visualized data set and the real problem the user wants to solve. Upside Short, distinct, and understandable onboarding messages improve understanding and increase the engagement of users. Concrete examples can provide a practical context and activate real-world knowledge (Schliemann and Carraher, 2002). Besides, it can enhance memory and understanding (Glenberg et al., 2004) and enhance transfer in a new concrete domain more than abstract exemplification (De Bock et al., 2011). Downside When phrasing concrete onboarding messages, the text could be too long and feel overwhelming for users (cf. study presented in chapter 4). Another downside could be the distraction from the relevant information (Kaminski et al., 2008) or constrained transfer of knowledge (Sloutsky et al., 2005). Issues Concrete onboarding messages should be short, distinct, and simple for the users. Therefore, iterative feedback from users should be gathered before publishing onboarding in the primary visualization tool. Theory/Evidence A longstanding debate concerns using concrete versus abstract instructional materials (Fyfe et al., 2015). Both concepts have their advantages and disadvantage, which have to be considered. Using concrete examples and referring to the underlying data enable learners to construct their knowledge about abstract concepts (Brown et al., 2009). Concrete exemplification enhances transfer in a new concrete domain more than abstract exemplification (De Bock et al., 2011).

Constructivist theories reflect on the application of learning in a practical context (concrete examples), which is relevant for the design of visualization onboarding instructions (Duffy et al., 1993; Sawyer and Greeno, 2009). Domain experts (cf. Section 9.3) emphasized the need for context-sensitive and concrete onboarding instructions (Schliemann and Carraher, 2002).

#### HOW is onboarding provided?

#### ID DA4

Title

Use floating text boxes with onboarding messages and inplace annotation anchors

- Support users with *floating text boxes*, including onboarding messages Description (textual explanations) with *in-place annotation anchors* in the visualization, which the user can interact with. The in-place annotation anchors, e.g., numbers and symbols, indicate the connection between the onboarding instructions and visual encoding. The floating text boxes should be positioned close to the in-place annotation anchor to follow the principles of Gestalt psychology of Proximity (Wagemans et al., 2012). In contrast, it is vital to allow the users to relocate them by drag&drop not to cover essential visualization elements. The onboarding concept can provide a guided (step-by-step) or selfexploratory experience for the users by using numbers and symbols or just symbols to show the connection between the onboarding instruction and the visual encoding (see related design action DA6). The text boxes and in-place annotation anchors can be assigned to the structure of reading the chart, interacting with the chart, and analyzing the chart presented in design action DA2.
- Effect The in-place annotation anchors provides context-sensitive and tothe-point instructions to the specific element in the visualization, e.g., legend, x-axis, or a particular artifact in the data. It makes it easier to absorb information.
- Upside The onboarding concept is easy to integrate into a VA tool as it is not space-filling. The in-place annotation anchors provide to-the-point descriptions and can be easily accessed by the user. The floating text boxes stand out from the rest of the user interface.
- Downside Due to the number of different elements, the visualization interface could be cluttered and overwhelm the user. The separation of onboarding instructions in text boxes, presenting small chunks of information, can mentally increase users' cognitive load by the need to integrate multiple sources of information (Ayres and Sweller, 2005). Instructional split-attention could occur when users have to split their attention between the different floating text boxes to gather all information essential for understanding.

Issues

The positioning of the onboarding concept elements, such as the floating text boxes and anchors, must be done correctly. The mental model of the users and the prior knowledge should be taken into account. The text boxes can consist of a headline, the instructional text, and essential words in the explanation can be highlighted visually.

Theory/Evidence Visualization onboarding that uses floating text boxes and in-place annotation anchors provided to-the-point descriptions, which make it easier to absorb information (Wagemans et al., 2012) (see results in study 3.5). Likewise, the understanding of introductions is enabled by interactive and linked descriptions of the visualizations. The text boxes present small information units that support learning (microlearning) (Hug and Friesen, 2009). Scaffolding (Holton and Clarke, 2006) is a key feature and part of constructivist learning, which can be applied in visualization onboarding by using text boxes and in-place annotation anchors. Scaffolding in visualization onboarding continually adjusts the level of help to the learner's performance by giving hints and cues and providing textual instructions to different elements of the visualization along the structure: reading, interacting, and analyzing described in design action DA2.

#### DA5

ID

#### TitleProvide self-exploratory and guided onboarding

Description The onboarding concept should be designed so that users can be both (1) guided (step-by-step) through the different onboarding messages and (2) self-explore it and seek help at any stage of their interpretation and interaction with the visualization. A guided (step-wise) instruction can be implemented using numbers and symbols instead of a "tour" with predefined steps. The user can gradually follow the numbers to be guided through the different onboarding messages. Users preferring to self-explore the onboarding instructions can open any onboarding messages individually at any exploration stage.

Effect Users' preferences in terms of learning can be taken into account by providing both concepts.

Upside First-time users, for example, can use the guided approach by getting an idea of the visualizations and the provided onboarding instructions. At the same time, a more experienced user can access onboarding instructions individually to their knowledge gap. As a martini glass structure (Segel and Heer, 2010), onboarding instructions can be designed to explore freely and then provide a guided way through the onboarding instructions to support first-time users, for example.

Downside	Covering both ways (self-exploratory and guided) could be challenging
	for the visualization designer as it requires more interaction possibili-
	ties with the onboarding concept, which has to be considered when
	designing visualization interfaces. Besides, there is the risk of not
	building a coherent mental model (Braasch et al., 2012) with the self-
	exploratory approach, as users can browse the onboarding messages
	according to their knowledge gap and interest.

Issues This design action relates to DA2 and DA4.

Theory/Evidence Both concepts refer to constructivism and cognitivism learning theories. Constructivist theories are the most appropriate for explaining learning processes with onboarding systems because they reflect on the application of learning in a practical context (Duffy et al., 1993). The self-exploratory approach can be assigned to constructivism, as users can freely explore onboarding messages and build their understanding.

> Cognitivism is a learning philosophy founded on the premise that learning can be modeled as a kind of information processing (Ertmer and Newby, 2013). Teachers are experts who provide knowledge and guide them in developing their cognitive abilities. This can be applied to the *guided* visualization onboarding, where the in-place annotation anchors with numbers guide the user through the onboarding messages.

#### ID DA6

Title

# Allow users to hide and show onboarding instructions to call them on demand

Description The onboarding concept should allow users to hide and show onboarding instructions to support on-demand help. This can be realized with an activation element to access the onboarding instructions on demand. Users should be able to interact with the whole visualization interface when onboarding is activated.

Effect Users can explore the data and open the onboarding when they need help ("pull" model (Horvitz, 1999)). This way, they are flexible and not restricted to using it when there is no need.

Upside As there are different knowledge gaps among users, this concept can support self-explanatory experiences and give more freedom to the user.

Downside	First-time users of an interface may need help finding the onboarding instructions or need to be made aware that any are available. A pull-based system requires the user to request assistance if needed actively.
Issues	The positioning of the onboarding activation element is essential. It needs to meet the users' mental model, and the design needs enough affordance if it is not activated. One possible solution is to open some onboarding instructions for first-time users to clarify that instructions are available. A floating action button (Material Design, 2022) (Material Design Pattern) can be integrated as an activation element.
Theory/Evidence	This follows the principles of the "pull" model, where a user explicitly asks for help (Horvitz, 1999) by activating the onboarding messages if

needed.

#### WHERE, and WHEN is visualization onboarding provided?

ID	$\mathbf{DA7}$

Title

## $Provide \ in-situ \ onboarding$

- Description Onboarding instructions should be integrated (in-situ) into the visualization to allow users to access onboarding while exploring the data (cf. Chapter 2). As described in the design action DA4, floating text boxes with onboarding messages and in-place annotation anchors can be used to integrate onboarding in a VA tool. At any point in the exploration (while using the VA tool), the user can activate onboarding instructions and get explanations to the point.
- Effect By providing in-situ onboarding, users do not have to switch back and forth between the VA tool and instructional material.
- Upside In-situ onboarding instructions integrated into the VA tool provides tothe-point explanations, which are easier for users to absorb information than external instructions on a website, for example. The onboarding messages are directly integrated into the interface, reducing the need to switch back and forth between the visualization and a separate tutorial or help section. In-situ onboarding instructions can reduce cognitive load for users. The onboarding instructions can be positioned near the element that needs to be explained. It supports the connection between textual explanation and visual encoding.
- Downside The additional visual elements of the integrated onboarding could overwhelm users.
- Issues When providing onboarding, the positioning of elements needs to be considered. Visualization onboarding should be designed parallel to the visualization interface to make positioning onboarding elements easier. Also, onboarding should reflect the design language of the underlying visualization tool, such as the same color and icons and color design.
- Theory/Evidence Integrating onboarding and using the underlying design language can reduce the extraneous cognitive load (Chandler and Sweller, 1991) in translating instructions to the current interface state. Splitattention (Ayres and Sweller, 2005) risk can be reduced as the onboarding instructions are integrated into the user interface of the VA tool and the switching between the VA tool and instructions is omitted.

Table 10.1: Overview of our design actions that are categorized along our guiding questions of the Design Space.

ID	Design Action	Design Space
DA1	Be aware of different knowledge gaps of users.	Who?
DA2	Explain how to read, interact, and analyze the data visualization.	Who?
DA3	Use concrete onboarding instructions where possible and needed.	Who?
DA4	Use floating text boxes with onboarding messages and in-place annotation anchors.	How?
DA5	Allow users to hide and show onboarding instructions to call them on-demand.	How?
DA6	Provide self-exploratory and guided onboarding.	How?
DA7	Provide in-situ onboarding.	Where? When?

#### 10.3 Summary

This chapter presents our design actions for designing visualization onboarding integrated into a VA tool or data visualization. Table 10.1 shows an overview of the design guidelines and categorizes them along our guiding questions from our design space (see Chapter 2): WHO is the user, and which knowledge gap does the user have? Which parts of a visualization need to be explained? How to phrase onboarding instructions? HOW, WHERE, and WHEN is visualization onboarding provided? We derived the design actions from results and the process within the two design studies performed for this thesis (Chapter 9 and Chapter 8). With these design actions, visualization designers and developers can better understand how to design and integrate onboarding into their data visualization applications.

Visualization onboarding is one aspect of user assistance. Another component is guidance (Ceneda et al., 2017). Next, we present our descriptive model on how visualization onboarding and guidance can be effectively integrated into the VA process to support users.

# CHAPTER **11**

# Guidance, Visualization Onboarding, and Knowledge-assisted Visual Analytics



This chapter describes the descriptive model of how visualization onboarding and guidance can be effectively integrated into the VA process. To do so, we confront the concepts of visualization onboarding and guidance and analyze their suitability in different analysis settings using Why/Who/When/Where/How questions (Section 11.4). We also present a usage scenario to show how visualization onboarding and guidance can support users in different analysis phases. We discuss the model's power (Beaudouin-Lafon, 2004) and provide an outlook. This chapter is based on our published work (Stoiber et al., 2022a).

## 11.1 Introduction & Motivation

Effective visualization onboarding supports users in learning how to read, interpret and extract information from visualizations (Stoiber et al., 2022b,c,d). However, visualization onboarding is just one aspect of user assistance. Another component of user assistance is guidance (Ceneda et al., 2020, 2017, 2018), which involves providing users with concrete suggestions and support for the next steps in the exploration. Both concepts of user assistance differ in the way they support users. Visualization onboarding targets teaching the user how to use the VA tool, whereas guidance provides interactive assistance while using the tool. This chapter explores the integration of visualization onboarding and guidance, examining their complementary aspects and developing a descriptive model to combine these two concepts.

The practical uptake of Visual Analytics (VA) approaches has increased substantially in recent years. Similarly to what happens in any data analysis process—and VA is no exception in this respect—analysts using VA typically alternate between a data *foraging loop*, in which bits of information are collected from a larger data pool, and a *sensemaking loop*, in which hypotheses are generated and explored by making sense of the data collected in the aforementioned process (Pirolli and Card, 2005).

However, Pirolli and Card (2005), who first described the sensemaking loop, mention that this process is not free from challenges: a set of *pain points*—or leverage points—typically affect the analysis and add to its total cost if no solution is provided to alleviate them. For instance, collecting information comes at the expense of exploring sufficient portions of the data and follow-up searches. Making sense of the data is limited by human working memory and the user's attention span.

The literature published in the last decade can be seen as a generalized attempt to mitigate such problems. Improvements in visual representations and modern graphical interfaces are all meant to reduce the costs associated with data exploration and augment human working memory. The advancements in machine learning have led to a plethora of "smart-agents" that aim to expand the options available to the users, allowing them to "discover the unexpected" while cutting down the effects of possible biases. Knowledge can also be reused and exploited to reduce the efforts of analyzing domain-specific problems. The threshold to gain new insights is usually set higher due to the initial knowledge required to analyze the data in the first place.

Transforming data into insights and knowledge is a challenging and time-consuming process. What we can observe, though, is that the success of any data analysis process relies on a small set of ingredients: (1) a well-designed user interface supporting effortless data exploration, (2) appropriate visual data encoding to account for the cognitive capabilities of the human visual system and promote, in the best way possible, the completion of tasks, and (3) a sufficient amount of knowledge, either in the user's mind or externalized in a knowledge base, to make sense of the data.

However, these three ingredients are often insufficient for successfully concluding the

analysis. When the user's knowledge does not match the expertise required to perform the analysis tasks, a knowledge gap prevents the completion of the task. In recent times, smart approaches, namely visualization onboarding and guidance techniques, have been studied to account for such situations and to address different *knowledge gaps*. These can be applied to extend the power of VA, enhance the user's experience, and broaden the audience for VA. Although different approaches to visualization onboarding and guidance in VA exist, there is a lack of research on designing and integrating them effectively and efficiently.

Both visualization onboarding and guidance are concepts easily understood, as they are also part of our daily experience. Imagine a cooking recipe with a list of ingredients but no instructions. Depending on the dish's complexity, you might assume that flour, water, and eggs should be mixed to prepare the dough. However, this might lead to mistakes. Cooking instructions are a good analogy for guidance, clarifying what to do with the ingredients to reach the expected result (the cake). Anyone with some cooking experience knows recipes are not always easily reproducible, even with detailed instructions. How can the cook know, for instance, that the "light brown" color of the bread in the oven matches the one expected in the recipe? Thus, onboarding is similar to detailed instructions and visual aids for recipes. The onboarding is successful when a person can follow these instructions and reproduce the dish.

Starting from this example, we argue in this chapter that visualization onboarding and guidance are complementary and decisive dimensions of any successful VA approach. While visualization onboarding and guidance have received a fair share of attention in their own right, they have not been considered together, nor has it been described how they concur to help the user complete any given task. In this chapter, we go in this precise direction, showing their different uses and scopes.

We build on well-known models to pinpoint the subtle (but important) differences that characterize visualization onboarding and guidance across domains, describe the scenarios in which it makes more sense to use them and show how they should be utilized for being effective and accepted by the user (Section 11.4).

In Chapter 11, we discuss a usage scenario in which we apply our framework to the financial domain, using Why/Who/When/Where/How questions and showing how onboarding and guidance can facilitate two everyday tasks in financial analysis: building a portfolio, and swing trading. The object of our analysis is the software *Profit* (Nelogica, 2021). While we do not provide information regarding design rationales for visualization onboarding and guidance (as discussed in our previous work (Ceneda et al., 2017, 2018; Stoiber et al., 2019a, 2022d)), our overall goal is to support the understanding and the intertwining of these two concepts to foster a well-considered usage of visualization onboarding and guidance techniques in practical scenarios, thereby reducing the cost of making sense of data, as initially promoted by Pirolli and Card (2005). In summary:

• We present a descriptive model of how visualization onboarding and guidance can be effectively integrated into the VA process (Section 11.3).

- We confront the two concepts of visualization onboarding and guidance and analyze their suitability in different analysis settings using Why/Who/When/Where/How questions (Section 11.4).
- We describe a usage scenario (SedImair et al., 2012) of our model using the stock trading VA tool *Profit* to show how visualization onboarding and guidance can support users in different phases of the analysis (Section 11.5).

## 11.2 Related Work

In this section, we revisit the literature addressing guidance. Visualization onboarding and guidance typically also rely on the presence of a knowledge base. Therefore, we complement our discussion by presenting knowledge-assisted VA as well.

#### 11.2.1 Guidance

Scientific work related to guidance goes back to developing the first human-computer interfaces (Dix et al., 2004; Engels, 1996; Smith and Mosier, 1986). In this initial phase, guidance was mentioned only as abstract guidelines for an effective design of user interfaces. The term guidance has also been used in other domains. For instance, in decision theory, Mark Silver described how to assist decision support systems (Silver, 1991).

Although using many different names, a considerable number of guidance approaches have been developed even in the visualization community. A preliminary work by Schulz et al. aimed at bringing all such approaches under the common name of guidance and described an initial set of characteristics they all had in common (Schulz et al., 2013). Among many others, one renowned example of guidance is the Scented Widgets (Willett et al., 2007). The approach uses visual cues integrated into interface widgets to guide the user through the exploration of a large information space. Gotz and Wen (2009) describe how to assist users in creating their visualizations and solving specific visual analysis tasks. The approach automatically understands the user's task by matching the interaction patterns with a database of previously created visualizations and accordingly proposes improvements to the current design. More recently, Ceneda et al. (2018) describe guidance to support the analysis and exploration of cyclical patterns in time-oriented data. In particular, the VA tool they designed can provide the user with indications about the most prominent cycles to explore, the length of the cycles, and possible repetitions in time. More recently, a literature review by the same authors categorized these and many other guidance approaches focusing on the mixed-initiative nature of guidance methods (Ceneda et al., 2019). Guidance can be seen as a dual process in which not only the system supports the user's analysis but also in which the user supports and provide valuable feedback to the analytical system and steer the analysis.

Although there has been much progress in the field, there is still a lot of confusion about distinguishing approaches that qualify as actual guidance from those that do not. Regarding guidance in VA, the confusion might also originate from a vague definition of the types of knowledge gaps that can be addressed using guidance (Ceneda et al., 2017). For this reason, we propose a unified model, clarifying the extent of visualization onboarding and guidance in VA.

#### 11.2.2 Knowledge in Visual Analytics

VA's main goals are discovering, acquiring, and generating new knowledge. According to Thomas and Cook (2005, p. 42), the final task of the analytical reasoning process is to create some kind of knowledge product or direct action based on gained insights. However, this process is not always straightforward. When approaching data analysis, we can imagine that the user has a certain amount of *prior* knowledge (Chen, 2005) that is based on previous experiences. We refer to this as *tacit* knowledge (Federico et al., 2017; Wang et al., 2009), as it is usually gained through incidental activities and is mainly contained in the user's mind. However, as described in Section 11.5, in increasingly complex data analysis scenarios, this knowledge is not sufficient to fully exploit the features of the analysis tools and, as a consequence, to solve all tasks and complete the analysis. This problem stems from a mismatch between the knowledge required to perform the task and the knowledge possessed by the analyst. We refer to this issue as the *knowledge gap* (Ceneda et al., 2017; Federico et al., 2017; Stoiber et al., 2019a, 2022c). Therefore, user assistance is intrinsically related to generating knowledge in that it actively contributes to closing knowledge gaps.

The process of closing a knowledge gap through visualization onboarding and guidance greatly benefits from the use of external sources of knowledge. As previously mentioned, a certain amount of knowledge is typically required to operate tools, namely *operational* knowledge (how to interact with the visualization system) and *domain* knowledge (how to interpret the content) (Chen, 2005). This knowledge can be externalized, stored in a database, and actively used to support the user when a knowledge gap is detected (Federico et al., 2017; Rind et al., 2019; Wang et al., 2009). We call this process *externalization*, and its output is *explicit* knowledge (Federico et al., 2017; Wang et al., 2009).

Stoiber et al. (2019a, 2022c) further enhanced these two levels of knowledge: (1) domain (e.g., vocabulary and concepts); (2) data (understanding the particular data type); (3) visual encoding (understanding the visual mapping); (4) interaction concepts (for performing tasks and understanding relationships between the data); and 5) analytical knowledge (knowledge of different automated data analysis methods). Operational knowledge (Chen, 2005) can be seen as a combination of visual encoding, interaction, and analytical knowledge (Stoiber et al., 2019a, 2022c).

We will use the aforementioned different types of knowledge to show how visualization onboarding and guidance deal with them.

**Domain Knowledge:** This kind of knowledge relates to a specific field of interest (e.g., healthcare, journalism, cyber security). Each domain has its vocabulary (Munzner, 2009)

# 11. Guidance, Visualization Onboarding, and Knowledge-Assisted Visual Analytics

for describing the data, workflows, conventions, and how data can be used to solve a problem (Miksch and Aigner, 2014). Domain knowledge is also an ensemble of concepts, intellectual tools, and information resources a user can draw upon to contextualize the visualized data.

An appropriate solution to a lack of domain knowledge is a knowledge base containing an externalized and computer-readable version of such knowledge (i.e., explicit knowledge  $\overline{K^{\epsilon}}$ ) (Federico et al., 2017; Wagner, 2017; Wang et al., 2009). One example would be a collection of computer-executable medical guidelines. On the one hand, onboarding techniques can use this knowledge to support the user's contextual understanding. For instance, by describing the vocabulary used, the domain's conventions, and workflows and explaining why certain choices should be/have been made. Onboarding might provide information as to why a particular treatment plan was selected based on the blood test results. On the other hand, the knowledge base can be exploited for providing guidance, for instance, to support the discovery of insights or to decide on future steps, such as choosing the next medical treatment (Federico et al., 2015). Guidance might provide information on which treatment method promises the greatest success for a certain evaluation result of a blood test in combination with the patient's other key data.

**Data Knowledge:** The first step in any VA process is data selection and manipulation (Keim et al., 2010). Therefore, when approaching a problem, the user must first know about the data at hand. Onboarding techniques tackle a lack of low-level information about the data, such as understanding principles of the specific data format, data types, or data structure. In contrast, guidance leads the user through the VA process for manipulating a specific data set as well as showing or highlighting interesting aspects of the data, such as outliers, missing data points, uncertainties, and inconsistencies.

Visualization and Interaction Knowledge: This type of knowledge deals with understanding the visual mapping (Card et al., 1999) and the interaction with the VA environment. By avoiding confusion about how to use the provided interaction features or which effects these interactions have, onboarding and guidance can fill a knowledge gap regarding the visualization of the interactions.

When dealing with interactive visual representations, onboarding mainly addresses problems related to understanding visual encoding and interaction concepts, performing tasks, and understanding relationships within the data. Typically, onboarding methods aim to fill this knowledge gap by using contextual menus, labels, or tutorials so that the user can understand their use and function (Stoiber et al., 2019a).

In contrast, guidance deals mainly with formalizing and structuring execution plans using these controls and interactions as building blocks. Thus, guidance might be used to help the user gain new insights from the visualized data, e.g., by highlighting deviations from normal medical conditions in the blood test data indicating a specific disease.
Analytical Knowledge: This knowledge is necessary to understand the analytical methods that are used by the given VA approach and how to set the parameters (e.g., machine learning techniques for prediction or classification). The user needs to understand the analytical methods and their characteristics to choose, parameterize, or utilize them effectively. This is where onboarding techniques come into play (Stoiber et al., 2019a). They establish a basic understanding and introduce the user to the different possibilities of the analytical methods. Furthermore, assisting the user in choosing the most appropriate of these analytical methods and parameters concerning the task at hand (Bouali et al., 2016) is a typical usage scenario for guidance. Sometimes, this might also include the comparison of performance metrics or an estimation of user preferences.

In this work, we describe how user assistance can benefit from explicit knowledge sources and contribute to generating new knowledge and insights. Based on the previously introduced terminology, we further characterize knowledge in Section 11.4 by listing all possible gaps that might arise during analysis and showing how user assistance can contribute to their resolution.

### 11.2.3 Model Building

Up to now, several papers have presented different methods for integrating knowledge into the VA process. Thereby, we can distinguish between descriptive models and mathematical models. In the view of descriptive models, Lammarsch et al. (2011) described that combining automated analysis methods with interactive visualizations is "a necessary step". The framework by Sacha et al. (2014) describes the process of knowledge generation based on three dedicated loops (evaluation loop, verification loop, and knowledge generation loop) when a human analyst is using VA tools. Furthermore, Sacha et al. (2016) have illustrated how uncertainties arise, propagate, and impact human knowledge generation processes by relating the concepts of uncertainty, awareness, and trust. A recent paper by Andrienko et al. (2018) presents a framework in which the VA process is considered as a goal-oriented workflow producing a model as a result. Besides, Chen et al. (2020) proposed a general framework bridging the gap between VA and storytelling. They introduced a story synthesis phase that extends the VA workflow with storytelling.

van Wijk (2005) presents a generic visualization model discussing the cost and gains of VA and the integration of knowledge on the user's side of the model based on mathematical equations. As an extension, Wang et al. (2009) described in their paper that integrating knowledge into the visualization process for solving analytical tasks is a fast-growing area. By integrating the expert's knowledge into the visualization process, "the experts are more capable of performing complex analytical processes" (Wang et al., 2009, p. 616).

The presented descriptive model (see Section 11.3) is also conceptually grounded in the visualization model introduced by van Wijk (2005) and on the "knowledge-assisted visual analytics Model" by Federico et al. (2017); Wagner (2017).

Summarizing, we discussed the work related to the three main topics: visualization onboarding (Stoiber et al., 2019a, 2022d), guidance (Ceneda et al., 2017), and knowledge-

# 11. Guidance, Visualization Onboarding, and Knowledge-Assisted Visual Analytics



Figure 11.1: The overview model of user assistance (drawn in black) shows the incorporation of user assistance (U) into the "Knowledge-Assisted Visual Analytics Model" (Federico et al., 2017) with connections to data D, explicit knowledge  $\overline{K^{\epsilon}}$ , and specification S. The detailed model representation (drawn in orange) shows that visualization onboarding O and guidance G are contained in the user assistance (U) process. Additionally, O and G are utilizing options Opt, cues C or external resources Ext contained in the image I and the exploration (E) and the perception P shown by dotted orange lines.

assisted VA (Federico et al., 2017). Onboarding can support users in learning and correctly interpreting applied VA methods (Stoiber et al., 2022b, 2019a, 2022d), whereas the intuitive or pervasive nature (Harman, 2010) of guidance (Ceneda et al., 2017) supports users throughout the visual exploration. Additionally, with knowledge-assisted VA (Federico et al., 2017), implicit knowledge of the user can be extracted and provided as machine-readable knowledge used for various aspects of onboarding and guidance. In the next section, we present the descriptive model in detail.

## 11.3 Descriptive Model of Integrating Guidance and Visualization Onboarding

To promote a discussion of visualization onboarding and guidance analogies and differences and to foster a better design and integration of such concepts into the VA process, we present a general descriptive model (Beaudouin-Lafon, 2004) combining and describing the aspects of **visualization onboarding** and **guidance** under the umbrella term of

what we might generically call user assistance (see Figure 11.1).

To illustrate the role of visualization onboarding and guidance in VA, we added the user assistance (U) element—containing onboarding (O) and guidance (G)—to the machine space of the model, as it is aimed at semi-automated generation rather than provided by a human expert (see Figure 11.1).

For providing user assistance (U), connections between extracted and stored explicit knowledge  $\overline{K^{\epsilon}}$ , data D, and the system's specification S are needed. In particular:

- 1. Data (data tables) is used as input for user assistance  $(D \to U)$  (e.g., important meta-information, such as the dimensions of the data, the included data types, and the data structure).
- 2. Explicit knowledge is another potential input to user assistance  $(\overline{K^{\epsilon}} \to (U))$ . For instance, the knowledge gap of the current user might be identified and closed based on the stored explicit knowledge of other users. Thus, user assistance can be provided concerning the state of knowledge.
- 3. User assistance and specification are connected in both directions  $(\bigcup \leftrightarrow \underline{S} \leftarrow \underline{E})$ , and the exploration state is an input to the specification. For example, the current exploration state defines the VA system's specification. If the user needs assistance, the current specification settings are used to determine appropriate user assistance. Based on the assistance steps, the user interacts with the system, thereby changing the specification.

Figure 11.1 shows the two user assistance (U) concepts, integrating elements of guidance (G) and onboarding (O), that are highlighted in orange. Here, user assistance (U) is directed from the machine side towards human perception (P). The onboarding (O) and guidance (G) processes both exploit similar input information. Based on their goals, different types of assistance can be provided (see Section 11.4).

In the following subsections, we discuss how to integrate visualization onboarding and guidance into the VA process.

#### 11.3.1 The Integration of Visualization Onboarding

The detailed description of the onboarding process in combination with all other elements is illustrated in Figure 11.1:

$$\{ \ \overline{K^{\epsilon}} \ , \ \overline{D} \ , \ \overline{S} \ \} \ \rightarrow \ \widehat{O} \ \rightarrow \ \{ \ \overline{C} \ , \ \overline{Ext} \ \} \ \rightarrow \ \widehat{P}$$

Onboarding can be either directly integrated into the interface of the VA tool in the form of visual cues  $\boxed{C}$  or independently as an external source  $\boxed{Ext}$ —for example, a website of the VA tool with instructional material. Visual cues  $\boxed{C}$  for onboarding may be

represented by tooltips, annotations, overlays, labels, highlighting in contextual menus, etc. (Stoiber et al., 2019a, 2022c) or, as presented in Figure 11.7, sticky notes.

Additionally, external resources  $\boxed{Ext}$ —commonly websites or documents describing the VA tool—provide explanations based on screenshots, videos, textual descriptions, etc. (see Figure 11.4).

In general, onboarding methods support the user in understanding different aspects of a VA system, such as the visual encoding, interaction concepts provided for the exploration (E), the underlying data D, or analytic methods (A) (Stoiber et al., 2019a). Explicit knowledge  $K^{\epsilon}$  may act as the basis for adaptive onboarding methods, supporting the closing of the user's knowledge gap, for instance, understanding the visual encoding of the presented visualization technique.

#### 11.3.2 The Integration of Guidance

 $\{ \ \overline{K^{\epsilon}} \ , \ \overline{D} \ , \ \overline{S} \ \} \ \rightarrow \ \overline{(G)} \ \rightarrow \ \{ \ \overline{Opt} \ , \ \overline{C} \ \} \ \rightarrow \ \{ \ \overline{P} \ , \ \overline{E} \ \}$ 

Different types of guidance might be advised for different users. For instance, visual cues  $\boxed{C}$  could be provided that is either integrated directly into the visualization itself or as part of the analysis environment. Visual cues usually present a set of unordered suggestions on how to proceed with the analysis. Other visual cues might be useful for orientation purposes, i.e., showing users where they currently are. Through the perception process P (i.e., perceiving these visual cues), they help the analyst to make sense of the data and thus to gain insights. In addition, guidance could be communicated by providing the analyst with various options  $\boxed{Opt}$  of how to proceed with the analysts. Usually, these options are ordered, for instance, by importance, and help the analyst to navigate through the exploration process. Finally, guidance could also provide a step-by-step description of how to proceed. This kind of guidance autonomously influences the specification ( $\textcircled{G} \rightarrow \boxed{S}$ ) of the visualization and the analytical processes. In contrast to guidance, onboarding does not provide any options giving concrete suggestions on how to proceed in a given situation.

## 11.4 Aspects of Visualization Onboarding and Guidance

A common goal of visualization onboarding and guidance is to support the user in effectively utilizing a specific VA tool.

Besides this common goal, however, they differ in how they support it. While onboarding is targeted at teaching the user how to use the tool, the guidance provides interactive assistance while using the tool. In the following, we present aspects of visualization onboarding and guidance based on five questions (Hart, 1996, 2002). Figure 11.2 gives an overview of the differences and similarities of visualization onboarding and guidance and illustrates how they complement each other.



Figure 11.2: Aspects of visualization onboarding and guidance along the questions of Why do we provide visualization onboarding and guidance?, Who can benefit from visualization onboarding and guidance?, When to provide visualization onboarding and guidance?, Where to provide visualization onboarding and guidance?, and How to provide visualization onboarding and guidance?

Why do we provide visualization onboarding and guidance? Although the general goal of both concepts is clear, different types of support should be provided concerning the context in which it is required. Visualization onboarding and guidance serve different purposes. Users have difficulties interpreting and working with novel visual representations or comprehending characteristics of the underlying data. Therefore, onboarding mainly deals with providing hints to understand and get acquainted with the interactive visualization environment—often before the actual use of the tool—or to tackle a lack of general understanding during the first use. Moreover, guidance aims at assisting the user in overcoming knowledge gaps regarding how to find or reach a specific analysis goal during an interactive analysis session. Finally, onboarding mainly involves perception and interaction, whereas guidance encompasses cognition and reasoning for knowledge discovery. Both concepts can help to reduce the costs of working with a VA tool by providing assistance at different stages of the usage of the tool.

Who can benefit from visualization onboarding and guidance? Any user of the VA environment can benefit from visualization onboarding and guidance. The proper implementation of both concepts can be critical for a solution to support a wide variety of users. Individuals have different amounts of prior knowledge, and therefore different *knowledge gaps* (Ceneda et al., 2017; Federico et al., 2017; Stoiber et al., 2019a, 2022d) when using a VA system to complete a certain task. Visualization onboarding and guidance do not target the designer of the VA systems. However, designers can benefit from the descriptive model, as they can use it as a blueprint (described in Section 11.6.2).

When to provide guidance and/or onboarding? Another important difference between visualization onboarding and guidance is the moment in which they are required (see Figure 11.2). Usually, onboarding means are offered immediately before or during the analysis, although many take place during the first use of a specific element, for instance, an introductory tutorial (Stoiber et al., 2019a, 2022c). Guidance, on the other hand, is provided mainly during the analysis. However, these questions can be tricky to answer for guidance techniques since all analysis states are typically not known in advance. Guidance is especially beneficial at the moment when the user has to make a decision (Silver, 1991).

Where to provide visualization onboarding and guidance? This question aims at defining the locations where assistance should be offered—either onboarding and/or guidance. Typical locations for providing both visualization onboarding and guidance are found in the visualization environment itself (Kang et al., 2003; Streit et al., 2011; Yalçın, 2016) (i.e., integrated into the visualization or as a separate component). An example of internal onboarding can be seen in Figures 11.7 and 11.3. Kang et al. (2003) presented approaches to help users get started with visualization interfaces using the sticky notes (see Figure 11.7) metaphor. Ymap is a map-based visualization tool where users can click on the map to see facts about the selected area in a table.

Additionally, they can select multiple areas and zoom in on the map. A dynamic query to filter the map according to a list of criteria is integrated. A scatterplot coupled with the map and table shows relationships between these criteria. However, onboarding methods can also be offered externally (Microsoft Power BI, 2019; Ola and Sedig, 2017; QlikTech QlikView Visualization Tool, 2019; Tableau Software, 2023; TIBCO Spotfire, 2019). For instance, common external onboarding resources are websites of the VA tool itself or offline documentation. A special case is represented by learning environments (Alper et al., 2017; Firat et al., 2022b; Ruchikachorn and Mueller, 2015; Tanahashi et al., 2016). These can be small games that users may play before using an application to become familiar with the basic concepts or websites such as Data Visualisation Catalogue (Ribecca, 2020) and FromDataToVis (Holtz and Healy, 2018) which are libraries of different information visualization types explaining the visual encoding. These learning environments cannot be considered internal or external, as they are not directly related to the VA tool. For guidance, an external source would be represented by users providing assistance in real time during the analysis. However, it is usually preferred to avoid additional users as sources of guidance since they may introduce biases. When this happens, the necessary input for guidance should be adequately formalized and included in the knowledge base.

How to provide visualization onboarding and guidance? The last and most important question concerns the modalities for providing visualization onboarding and guidance. Answering this question correctly directly influences the effectiveness of the



Figure 11.3: The guided tour mode (Yalçın, 2016): (1) The tour progress is visible, and users can direct it forwards, backward, or to a specific step. (2) The tooltip of the main action, which is a mouse-over on a visual glyph, is highlighted by color. (3) Additional tooltips describe the effect of this action on other interface components. (4) A detailed topic description presents an easy-to-read summary of tooltips and additional information. Related topics and the context in which this topic applies can also be viewed on demand.

provided assistance. However, the 'how' is context-sensitive and should therefore be dynamically generated while also considering possible user preferences.

**Onboarding** procedures should be tailored to the users' preferences. Some might prefer videos or hands-on tutorials, while others might find textual descriptions or offline documents more useful (Stoiber et al., 2019a, 2022c). An example of how to integrate onboarding using visual cues can be seen in the approach of Kang et al. (2003) using the sticky notes metaphor. These cues are context-sensitive and embed an onboarding method that supports the users by highlighting the main functions of the interface. The sticky notes provide a series of steps to accomplish a task, such as highlighting unconventional features or showing where to click in the interface). Furthermore, Yalçın (2016) used overlays including a combination of topic listing, point & learn, guided tour, notification, and topic answers for his approach. Once the help material is selected, its answer is presented in-situ; the material is fully integrated into the interface. Relevant components are highlighted, tooltips to describe the actions are shown, and textual instructional descriptions are provided to the user (see Figure 11.3).

Commercial visualization tools, in particular, use documentation/explanation websites with screenshots, videos, and textual descriptions (QlikTech QlikView Visualization Tool,

# 11. Guidance, Visualization Onboarding, and Knowledge-assisted Visual Analytics



Figure 11.4: **External help website** including videos, screenshots, and instructional text to explain the features of Microsoft Power BI (2019).

2019; SAP Lumira, 2019; SAS JMP, 2019; Tableau Software, 2023; TIBCO Spotfire, 2019). In Figure 11.4, the external help website of Microsoft Power BI (Microsoft Power BI, 2019) is presented, including videos, textual descriptions, and instructional texts.

The degree of **guidance** can also be varied (i.e., low to a high level of guidance) and should be chosen concerning the users' needs. Different degrees of guidance could range from visual cues for orientation to specific step-by-step instructions (Ceneda et al., 2017). In general, using visual cues serves the goal of providing subtle suggestions to the user, who, however, must discern and reason about them before blindly following suggestions. For instance, Luboschik et al. (2012) use visual highlighting cues to indicate data cases that are worth analyzing based on a heterogeneity metric and hence, foster profitable data exploration. Streit et al. (2011) present possible future analysis paths to

🚊 IBM Cognos Analy	ytics Wilkommen 🗸	Buy now 🕺 🗘 🚊 🕜
Startseile Q. Suchen Eigener inhalt Teaminhalt	Willkommen bei IBM Cognos Analytics Offnen Sie zuerst ein Dashboard, einen Bericht oder eine Story.	Kurzübersicht
Califizit versendete. Propiste	Create new content From uploading files to creating a new tashbaard, report, or exploration, you an launch them all from here. Der Inhalt wird hier angezeigt. Tom Ehsting können Sie Daten Hinzuffigen ober des Bespuese Neten.	Getting Started         In 5 minutes you're going to be a Cognos         Superstart         %         Office         Take a tour of Cognos Analytics         Oreate a Dashboard         Discover the new Explore
Q Verwalten	Next	No thanks

Figure 11.5: **Step-by-step tour** in IBM Cognos Analytics (IBM Cognos Analytics, 2019) to onboard first-time users on demand using tooltips and overlays to explain the general features of the tool.



Figure 11.6: Visualization onboarding by analogy. To introduce a spiral chart, a linear chart is gradually morphed into a spiral layout (Ruchikachorn and Mueller, 2015).

the analyst. Ceneda et al. (2018) provide suggestions of potentially interesting cyclical patterns in time series data to users. Interesting cycle lengths are highlighted directly in the interface widgets that are used for exploring the data.

Guidance can also provide a set of alternative options for the next analysis steps. The user may or may not choose to follow these options. A typical example of this kind of assistance is recommender systems. The recommendations produced by such systems are directly related to the task the user is currently solving. In comparison to lower

# 11. Guidance, Visualization Onboarding, and Knowledge-assisted Visual Analytics

degrees of assistance, like visual cues, such assistance is considered stronger and directly tailored towards steering the course of the analysis. An example of this kind of assistance is presented by Kandel et al. (2011, 2012). In their work, they designed Profiler and Wrangler tools for supporting data profiling and transformation. The tool works by recommending a set of useful functions to perform data cleansing and transformation operations based on the type of data that is analyzed. Another approach that aims to redirect the analysis is presented by May et al. (2012). Their approach aims at assisting the user in exploring large graphs by incorporating visual glyphs to direct the exploration towards interesting areas of the graph or towards data that has not been analyzed yet.

One of the least intrusive forms of guidance is exemplified by approaches improving the user's orientation. In data analysis, this typically translates into supporting the construction of a mental map, which refers to the ability of a user to build an internal representation of the data under analysis. Guidance works by directly targeting the user's perception by exploiting low-level information extracted from the dataset, e.g., statistical values, and mapping such values to basic pre-attentive features, e.g., color hue, position, etc., drawing the user's attention towards certain data characteristics (Ceneda et al., 2019). Such orientation cues could be added on top of the visualization itself or in the interface widgets. In both cases, the goal is to guide and address the interaction, foster understanding, and take the first steps toward gaining insights. The fact that orientation guidance works at such a low level means that it has to target basic human abilities. Typically, this involves the user's pre-attentive skills. For instance, highlighting data items or parts of the interface is typical of orientation guidance. The contrast between the color hue or the color intensity of the highlighted elements and the surrounding ones should immediately strike the user's attention, avoiding tedious search operations. May et al. (2011) show an instance of such a technique: they provide guidance to feature selection for model building. This is achieved by changing the color hue, i.e., highlighting of data columns that the system deems interesting.

While context-free visualization onboarding and guidance are theoretically possible, they should be avoided, as unspecific assistance might be even worse than doing nothing. In addition, the combination of multiple types of assistance might be indicated. For instance, a visualization tool could offer visualization onboarding to facilitate the use of the visual interface while at the same time providing guidance options to complete a given task.

## 11.5 Usage Scenario on Visualization Onboarding and Guidance in the Financial Domain

In this section, we describe how visualization onboarding and guidance can be effectively utilized to support domain tasks actively. To the best of our knowledge, not many approaches in the literature describe how to exploit and combine the full extent of visualization onboarding and guidance support. However, we were able to identify multiple aspects of visualization onboarding and guidance in the professional trading and stock market analysis software *Profit* (see Figure 11.8). Therefore, we provide



Figure 11.7: Sticky notes (Kang et al., 2003) used to onboard users to the Ymap tool. This is an example of using visual cues to support users in understanding the main functions of the interface.

conclusions regarding the financial domain from our descriptive model perspective along the Why/Who/When/Where/How questions in Section 11.5.1, as we did it in Section 11.4 to show the general aspects of visualization onboarding and guidance.

The financial domain is complex and hard to grasp fully, both in theory and practice. It is riddled with jargon and tangles legal, economic, and political aspects that can be both national and/or international. State of the art in investing is also never truly reflected in the published literature for obvious reasons. Due to its competitive nature, it is not in anyone's best interest to reveal optimal practices because they might provide economic advantages. Furthermore, it is commonly accepted that there is a division between the types of stakeholders that make up the economy (also known as "Wall Street vs. Main Street"(Lamin and Zaheer, 2012)), which can cause wildly varying pictures of what good investments are at any given time. In this section, we show how visualization onboarding and guidance are essential for countering this complexity and providing efficient VA for the financial domain.

In general, financial analysis is, broadly speaking, divided into Fundamental Analysis (FA) and Technical Analysis (TA). FA places emphasis on *fundamentals*, qualities or quantities that can be measured about a stock, such as a company's equity, its brand value, or the skill of the executive team. To figure out a good investment strategy, investors can look at all data available to better gauge values. TA, on the other hand, assumes that the current price approximates the real value of a stock, all things considered. Its modus operandi is the visual analysis of graphs. It does not try to avoid overpricing, speculation, random events, and emotion as noise over the price signal. Rather, finding good investments involves using them tactically. Therefore, it is a useful analogy to think that FA is more of a strategic tool, while TA is more tactical: FA believes that fundamentals are the true latent variables that control the price over time, therefore it plans to win a war. TA

# 11. Guidance, Visualization Onboarding, and Knowledge-Assisted Visual Analytics



Figure 11.8: Starting screen of the *Profit* (Nelogica, 2021) software in the default PT-BR language. The central window  $\square$  shows the valuation of the Brazilian stock exchange (BOVESPA) through the IBOV index in a candlestick chart. Around it, a myriad of green and white numbers display other stock variations  $\square$ . On the right  $\square$ , controls for placing others and adding markers to the graph can be seen. This is the first dashboard presented to new users but it is highly customizable, with hundreds of choices between different interfaces, indicators, buying interfaces, and views. With all the different parameters and possibly financial instruments to be displayed, the possible dashboard configurations are virtually infinite.

believes that the stock signals have local behavior that can be predicted over shorter periods of time, allowing the user to corner the price and generate profit over many trade battles. These two approaches are not exclusive, and most investors will have a good understanding of both while favoring one of them.

All trading tasks involve analyzing price behavior: inspecting the past, gauging the present, and estimating the future. Figure 11.13 shows three optimal trading patterns over the same candlestick series, all of which involve buying and selling stocks and profiting from the price difference, but with different strategies or restrictions. From images a to c, the number of actions performed increases, and the estimated profit increases, but so does the risk involved. These optimal patterns would not be achievable in a normal scenario, as it is impossible to determine the exact behavior of the price series. The object of our analysis is the software *Profit* (Nelogica, 2021). It is the largest platform for stock trading in Brazil, with over six million trades daily. This is a considerable share of all BOVESPA (Brazil's stock exchange) transactions on any given day.

*Profit's target audience is personal traders* instead of large institutions and hedge funds, which represent the bulk of the financial volume. This makes it interesting for our exposition, as it is tailored to ease novice users in the financial domain into trading, which is reflected in the visualization onboarding procedures. Figure 11.8 shows *Profit's* main screen when opened for the first time. This is the first dashboard presented to new

users, but it is highly customizable, with hundreds of choices between different interfaces, indicators, buying interfaces, and views. With all the different parameters and possibly financial instruments to be displayed, the possible dashboard configurations are virtually infinite. The main screen, illustrated in Figure 11.8, shows the valuation of the Brazilian stock exchange (BOVESPA) through the IBOV index in a candlestick chart. The reader might notice that it has the appearance of an analysis tool for temporal data. Still, the complexity of the domain can already be grasped by the amount of information displayed and encodings that might be unfamiliar, such as the candlestick chart used for displaying price variation (Figure 11.12). Therefore, there is a need for visualization onboarding and guidance to support the users in exploration tasks with the *Profit* tool. To illustrate our discussion, we propose two trading tasks that can be predominantly associated with FA and TA, respectively: building a portfolio and swing trading. Figures 11.9 and 11.10 show two desktops configured to perform the respective task within *Profit*. In the first example, a stock is bought and held for a long period to be sold for a small but safe margin of profit (Figure 11.13 (a)). The average value of a stock's price tends to increase slowly over time, so holding will probably result in gains even against local lows and fluctuations (i.e., the two drops in price between buying and selling). Swing trading, on the other hand, tries to use these fluctuations to benefit on many "buy low, sell high" trades over time, having a higher profit margin potential (Figure 11.13 (b)). Both personalized desktops are shown in Figure 11.9 and 11.10 subtly reflect these particularities of the task. In the first one, most of the screen space is dedicated to identifying good candidate stocks, and there is no button for buying stocks visible, as this action is done minimally and can be performed through a pop-up window. On the right, however, an integrated component for configuring and placing orders is strategically placed next to the graph.

#### 11.5.1 Descriptive Model illustrated in the Financial Domain

In this section, we tie the two example tasks from FA and TA to our Why?, Who?, When?, Where?, and How? questions and provide conclusions regarding the financial domain from our framework's perspective. In Figure 11.11, we present the aspect of visualization onboarding and guidance compared to the two tasks of FA and TA, according to the aspects described in Section 11.4.

Why do we provide visualization onboarding and guidance? When evaluating the financial domain from a VA perspective, such as understanding the data, users, and tasks (Miksch and Aigner, 2014), two core aspects can be highlighted. First, it essentially uses time-series data, and second, most tasks have a standard key performance indicator: profit. All the diversity of investing and trading strategies will fall under the scope of this umbrella. Still, they do not fully characterize what can be considered data analysis in the financial domain. A data journalist who writes for a financial newspaper can perform the same tasks for analysis as an investor, but only to gather information for publication, without the commitment of trading and, therefore, without risk. This is an important distinction because it is evident that both visualization onboarding and

# 11. Guidance, Visualization Onboarding, and Knowledge-assisted Visual Analytics



Figure 11.9: The **adapted desktop for performing the task of building a portfolio** which we present as more in line with FA. Most of the screen space is used to visualize stocks to identify good candidates.



Figure 11.10: **Swing trading desktops configuration**, which is rather associated with TA tasks. In this setup, the screen shows different information about the stocks to find good times to buy and sell stocks to which price.



Figure 11.11: Figure showing the comparison of **two high-level tasks** in the context of building a portfolio (Fundamental Analysis (FA)) and performing swing trading (Technical Analysis task (TA)) in the *Profit* VA tool, respectively, according to the criteria described in Figure 11.2. Exemplarily, for the question of **When** to provide guidance and/or onboarding? guidance and onboarding can be provided differently. For example, guidance is integrated into the tool providing user assistance **while** using the tool, contrary, onboarding also supports users **before** starting the data exploration with the *Profit*. In the light of the using the *Profit* tool for building a **portfolio/performing swing trading**, the integrated **onboarding** concepts can support the user when opening the *Profit* tool for the first time, or when some real-time events occur, and/or when new features are activated. Besides, **guidance** supports users **while** selecting stocks, monitoring portfolio growth (FA), and also selecting financial instruments to operate.

guidance for these two users (the data journalist and the investor) in an *ideal scenario* would be similar but cannot be the same.

When a task involves taking action with risk, this should be reflected both in visualization onboarding (e.g., guaranteeing that the risk and its factors are communicated) and guidance (e.g., providing options and plans to reduce the risk). Risk, together with the two core aspects discussed previously—time and profit—characterize the financial domain well as dimensions of analysis. Time and money are exchangeable resources, and risk modulates the efficiency or control of these exchanges. Virtually all tasks in the financial domain involve estimating time-profit-risk relations between entities (stocks, commodities, contracts, currencies, players, one's portfolio), which amounts to understanding the environment, and then maybe performing an action that alters

## 11. Guidance, Visualization Onboarding, and Knowledge-Assisted Visual Analytics

that state (placing a buy or sell order). Since the market is a chaotic system, it is impossible to fully understand the environment and the consequences of one's actions, so all measurements and actions contain some uncertainty.

In our example, the users of the *Profit* tool are investors performing two tasks: building a portfolio and swing trading. In Figure 11.9 and 11.10, we provide the respective desktop to do the specific tasks. The goal of building a portfolio is to select promising, stable stocks that the investor will hold for a long time before selling. A user building a portfolio performs an exploratory analysis to find good candidate stocks, and the presented desktop is optimized to help him, or her find them. On the right-hand view, one can see stocks with larger volumes, which can be re-sold more easily in the future, and the ones marked in the green present a lower risk since they have a positive valuation. In the bottom left corner, the most popular stocks are displayed, while in the top center field, the user can monitor their portfolio evolution over time.

Swing trading, on the other hand, requires an active effort to identify good moments to buy and sell stocks. This desktop in Figure 11.10 then shows different information about a stock that helps the trader identify not only good times to buy and sell but also how much and which stock. Besides the candlestick chart at the bottom enriched with moving average indicators, most of the data displayed are statistics about the order book, where buyers and sellers make their offers, and which is the actual engine behind stock trading. The price graph is an abstraction of what is happening in the book, indicating the average price of the last trades that happened, which is generally a good summary of stock, but it is far from the complete picture. A user operating this desktop will navigate this interface to gauge potentially good trades, then place orders using the interface in the bottom right corner and monitor them both on the graph and in the central view.

To successfully perform those two tasks, the investor has to have different types of knowledge  $\overline{K^{\epsilon}}$  (see Section 11.2.2 for the different knowledge types): domain knowledge about the financial domain, data knowledge, visual encoding, and interaction knowledge to understand candlestick charts, treemaps, and line charts, and the different interaction possibilities of the *Profit* tool. If there is a knowledge gap in one of those areas, visualization onboarding (gain understanding) and guidance (gain insight) can help the user to overcome it ( $\overline{K^{\epsilon}} \rightarrow (\underline{U})$ ). With the integration of visualization onboarding  $(\underline{O})$  concepts into the *Profit* tool, we can support the users in understanding the domain, data, visualization, and interaction. Guidance  $(\underline{G})$  can help the user to choose strategies and support the exploration for building a portfolio.

 $\{ \ \overline{K^{\epsilon}} \ , \ \overline{D} \ , \ \overline{S} \ \} \ \rightarrow \ O, \ \widehat{G} \ \rightarrow \ \{ \ \overline{Opt}, \ \overline{C} \ , \ \overline{Ext} \ \} \ \rightarrow \ P$ 

Who can benefit from visualization onboarding and guidance? When we think about the two tasks of building a portfolio and performing swing trading with the *Profit* tool, visualization onboarding can support new users of the tool (investors), but also experienced investors who have never built a portfolio before or have no experience with

**TU Bibliothek**, Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar. WIEN Your knowledge hub The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.

swing trading. Guidance can support long-term investors and mid- to short-term traders in the respective tasks (data knowledge gap; see Section 11.2.2). An investigator, such as the financial journalist mentioned above, or anyone just wanting to perform some data analysis on stock data could also be supported by guidance for the first task, as it is largely an exploration task (data & visualization and interaction, in greater detail described in Section 11.2.2). To sum up, the financial journalist or the investors using the *Profit* tool most probably don't have a knowledge gap in the *domain*, as they are experts in the field. A gap in *data, visualization, and interaction* can occur.

How to provide visualization onboarding and guidance? A new user is offered the option of using the tutorial, which is comprised of nine steps. In Figure 11.14, we present four screenshots highlighting specific parts of the tutorial. *Profit* integrates *visual cues*  $\boxed{C}$  (tooltips) with textual descriptions into three parts:

$$\{ [K^{\epsilon}], [D], [S] \} \rightarrow (O) \rightarrow \{ [C] \} \rightarrow (P)$$

These three parts are: learning how to take a position by configuring and placing orders (steps 1 to 4, Figure 11.14a), learning how to monitor and close ones position (steps 5 and 6, Figure 11.14b), and learning how to explore more options and get support (steps 7 to 9, Figures 11.14c, 11.14d). In the last step (9, Figure 11.14d), the user is directed towards more in-software onboarding tutorials (Visualization Onboarding /Onb) or to additional study material available in multiple media.

$$\{ \ \overline{K^{\epsilon}} \ , \ \overline{D} \ , \ \overline{S} \ \} \ \rightarrow \ O \ \rightarrow \ \{ \ \overline{Ext} \ \} \ \rightarrow \ P$$

In the YouTube channel  $\boxed{Ext}$ , for instance, content is published daily with guidance on how to use the software for different purposes. These onboarding concepts can be used to support the investor in building a portfolio as well as in performing swing trading.

Guidance (G) is not explicitly provided in the *Profit* tool. However, various supporting mechanisms support the user in building a portfolio or performing swing trading. In light of our framework, this can be considered as guidance. First, *Profit* provides different desktop configurations the user can save, load, and share Opt. As exemplified in Figure 11.9 and Figure 11.10, this can be used to guide users towards suitable setups for performing certain tasks. Furthermore, most components and indicators are hyperparametrized, and this parametrization is task-sensitive. Providing preset parameters for certain strategies and tasks is a weak form of guidance, which dynamic, user-centered recommendations could improve. A user who has never used a certain indicator, but is performing swing trading, for instance, might identify this preset from the list and have a better chance of gaining insights for this task.

 $\{ \ \overline{K^{\epsilon}} \ , \ \overline{D} \ , \ \overline{S} \ \} \ \rightarrow \ \widehat{G} \ \rightarrow \ \{ \ \overline{Opt} \ \} \ \rightarrow \ \widehat{P}$ 

Finally, *Profit* provides a simulator that allows the user to replay the behavior of a stock at a particular time and test strategies. While this is not explicitly visualization onboarding

or guidance, it provides support for both, as one can gain both understanding and insights from the simulation. Testing a strategy on past data is strongly recommended, and unexpected results and patterns can be uncovered.

When to provide guidance and/or onboarding? Visualization onboarding and guidance are required at different points in the data exploration with *Profit*. For the task of building a portfolio, visualization onboarding can support first-time users or during use when new features are activated. Guidance can be used to choose the right parameters for indexes and encodings so that the user can easily search for patterns. For the task of *performing swing trading*, visualization onboarding of first-time users is even more important, as there is more risk involved. Additionally, when some real-time event occurs, visualization onboarding can be called up. Guidance, on the other hand, is provided mainly at two times: before the operation, for example, while selecting an interesting stock or financial instrument to operate, and during operation, actively assisting the trading process by showing potential entry and exit points for trades.

Where to provide visualization onboarding and guidance? For visualization onboarding, the *Profit* tool integrates a nine-step onboarding process within the main interface. Then, in the last step of these in-place onboarding instructions, the system points to additional study material, which is available in multiple media, such as additional in-software tutorials, social media, and an integrated support hotline. In the YouTube channel, for instance, content is published daily with support on how to use the software for different strategies. This can be helpful while performing both tasks of building a portfolio and swing trading. Most features also have a help component, which at first sight can be considered pure onboarding, but can provide guidance when linked to additional support content. For instance, if the user has never used Renko charts and clicking the help button on a Renko chart of the XYZ index suggests an introductory YouTube video on this topic, this qualifies as guidance.

In summary, there are many ways in which our framework can be applied to make use of visualization onboarding and guidance whenever a user has to interact with data to perform a task.

## 11.6 Discussion & Lessons Learned

In the following, we present lessons learned and limitations of our descriptive model.

### 11.6.1 Integration of Guidance & Visualization Onboarding

Our descriptive model shows theoretically how visualization onboarding and guidance can be integrated into the VA process and how their different components interact with each other (see Section 11.3).



Figure 11.12: (a) Basic anatomy of a **candlestick chart** with all its essential components: opening price (O), closing price (C), highest (H), and lowest (L). The white or green color is associated with a price rise (C > O), and a red or black color is used for a fall in price (C < O). The painted area between O and C is the "real" body of the candlestick, and the lines that stretch to H and L are its shadow or tail. (b) Comparison between two visually similar patterns with the same real body but highly distinct shadows represents different stories. Both candles show the same positive price variation (C > O). Still, on the right, the price oscillation was much smaller than on the left, indicating that market sentiment is more stable on how to price this stock during the period.

In terms of current approaches to visualization onboarding, we can list some solutions of commercial VA tools, such as (Kang et al., 2003; QlikTech QlikView Visualization Tool, 2019; SAP Lumira, 2019; SAS JMP, 2019; Tableau Software, 2023; TIBCO Spotfire, 2019). When it comes to guidance, some approaches can also be found in the literature (Ceneda et al., 2018; Luboschik et al., 2012; May et al., 2011, 2012).

However, hardly any VA system integrates both concepts. Therefore, we present how visualization onboarding and guidance can be integrated into a VA system to support users in the financial domain based on a usage scenario with the VA tool *Profit*. In Section 11.5, we describe the use of visualization onboarding and guidance along the five questions of our framework. We decided on the financial domain as it is complex and hard to understand fully in detail. For novices in the financial domain (e.g., in financial analysis or a journalist for a newspaper), basic concepts, terms, and visual encoding (e.g., sticks) have to be explained to perform financial analysis successfully.

Lessons Learned from the Usage Scenario: When applying the Why?, Who?, When?, Where?, and How? analysis to both tasks (Figure 11.11) within *Profit*, we can outline the differences and similarities between them, but also between the visualization onboarding and guidance dimensions. On the "Why" criterion, for instance, visualization onboarding and guidance are diametrically opposed on the understanding/insight axis, but both tasks share the same motivation for visualization onboarding. The "Who", however, has both visualization onboarding and guidance are diametrically and guidance matching on the user/designer.

## 11. Guidance, Visualization Onboarding, and Knowledge-Assisted Visual Analytics



Figure 11.13: **Distinct trading scenarios** with the same candlestick patterns but different strategies. (a) A stock is bought and held for a long time, then sold. (b) Multiple trades are performed in the same period, taking advantage of low and high price variations to optimize profit. (c) If shorting is allowed, all movements can be turned into profit, but at a higher risk. The rectangles encapsulate the space where each position is open and can be cued to understand the risk in each situation. Profit can be determined by the difference in price between entering and exiting operations, the orange arrows, which are essentially their vertical component, are indicated on the right by the dotted arrows.

220

**TU Bibliothek**, Die approbierte gedruckte Originalversion dieser Dissertation ist an der TU Wien Bibliothek verfügbar. Wien wurknowledge hub The approved original version of this doctoral thesis is available in print at TU Wien Bibliothek.



Figure 11.14: Screenshots of *Profit's* first-time user onboarding, which is comprised of nine steps. Here we show four out of nine steps:  $\Box$  Overview of the interface, the main graph, and how to place orders.  $\Box$  Monitoring trading results.  $\Box$  In the financial domain, extreme importance is attached to live support, which is why it is presented as an essential tool during onboarding.  $\Box$  Further options, both internal and external to the software, are provided to proceed with the learning curve. The images are all in screen resolution to reproduce the users' experience. They should be zoomed in on and read left to right, top to bottom.

axis, but the user profiles are different. Finally, the best way to distinguish both is in the "How", which can be characterized well in the financial domain for the active operational aspect (also linked to "When"), even if it has the idiosyncrasy of providing human-based

visualization onboarding and guidance through an external call center. In summary, there are many ways that our framework can be applied to make use of **guidance** and **visualization onboarding** whenever a user has to engage with data to perform a task.

#### 11.6.2 Descriptive Model of Visualization Onboarding and Guidance

According to Beaudouin-Lafon (Beaudouin-Lafon, 2004), models can exhibit three types of characteristics: (1) *descriptive*: the ability to describe a wide range of existing methods; (2) *evaluative*: enable the assessment of multiple design alternatives, and (3) *generative*: help in designing new methods. Our proposed model is of the descriptive kind (see Section 11.4), as it can systematically describe and present relevant aspects of visualization onboarding and guidance.

Lessons Learned from building the Descriptive Model: Thus, our model helps visualization designers who can use it as a blueprint for future VA tools integrating visualization onboarding and guidance components. First, they can decide how and when to integrate visualization onboarding or/and guidance into their VA system (see Figure 11.2).

On the machine side of the descriptive model (Figure 11.1), the designer can see that the data, for example, can serve as input for the planned onboarding. Furthermore, s/he can decide on how to provide the onboarding perceived by the user through integrating visual cues, for example.

#### 11.6.3 Limitations

One limitation of the model is that it does not provide a step-by-step architecture (e.g., design patterns, algorithms, data structures) or concrete design guidelines for visualization onboarding and guidance concepts (e.g., instructional material in the form of videos, texts, etc., highlighting concepts for leading the focus to interesting elements in the VA system). In Chapter 10, we present design actions for the design and integration of visualization onboarding in VA tools, which is not part of the current descriptive model presented in this chapter. Besides, Ceneda et al. (2020) present a framework that points design considerations are also not incorporated in the presented descriptive design model. Since this model is a high-level blueprint, a limitation is a possible depth of describing the user's cognitive processes, perception, and tacit knowledge generation. Moreover, additional research needs to be carried out to extend the model beyond its descriptive nature to a more generative and evaluative one for visualization onboarding and guidance.

## 11.7 Conclusion and Future Perspectives

In this paper, we contributed a descriptive model of visualization onboarding and guidance in VA. When it comes to such models, we assess their differences and similarities along the questions of Why/Who/When/Where/How thereby leading to their effective and thoughtful use in VA. We also show an application of the model in the financial domain, illustrating the importance of visualization onboarding and guidance for an effective analysis. The in-depth discussions and comparisons opened up several future perspectives:

- 1. Effective assistance in the form of guidance and/or onboarding during the visual exploration process needs to be evaluated in more detail with a particular focus on real-world and real-time scenarios. This calls for appropriate evaluation metrics and processes, which are open challenges in VA (Isenberg et al., 2013);
- 2. Visualization onboarding and guidance concepts require flexible and adaptive approaches according to the prior knowledge of the users, which the knowledge gap could determine the kinds of exploration, as well as the interaction behavior of the users;
- 3. Tacit and explicit knowledge are crucial for visualization onboarding and guidance. Consequently, the reliability of such knowledge needs to be assessed.

Next, we discuss the outlook based on our research (Chapter 12).



# CHAPTER 12

# Outlook



This chapter describes future directions based on our research results of Part I, Part II, and Part III and the design actions we present in Chapter 10.

Based on our research results, we propose future directions that guide researchers, designers, and developers to onboard users efficiently to their VA tools. We extract these future directions from the two design studies, presented in Part III.

**Extension of Design Space with Evaluation Space** In general, evaluation is essential to develop meaningful, easy-to-use, and understandable visualization onboarding concepts that enhance learning and support users in using the VA tool. To fully understand how users interact with the visualization onboarding concept in detail, more research is needed with actual users. Firat et al. (2022a) provide a literature review of visualization literacy and an overview of different evaluation methods used in visualization literacy studies. They identified potential research areas where they propose to develop new experimental settings, including different parameters or materials to understand barriers to visualization literacy.

Our current Design Space does not cover this aspect, as it is focused on who the user is (knowledge gap) and how to provide onboarding. In the future, we can extend our design space with an Evaluation Space to give a structured overview of evaluation methods for measuring learning in visualization onboarding.

Impact of Visualization Onboarding on the Visualization Literacy Level of Users More recent attention has focused on assessing the visualization literacy level. Lee et al. (2017) developed a visualization literacy assessment test including twelve data visualization types and eight visualization tasks. Furthermore, Börner et al. (2016) tried to understand the general public's literacy level by showing museum participants 20 different printed data visualization. In line with the previous work by (Börner et al., 2019, 2016; Galesic and Garcia-Retamero, 2010), we see poor answer correctness over different onboarding concepts (cf. Study 2 presented in Section 3.3.1 and Table 3.2). Moreover, we realized that the subjective experience levels of our participants per visualization type do not correlate with high answer correctness.

In future work, we will extend our research to evaluate if and how visualization onboarding can increase a broader audience's visualization literacy level.

**Data Analytics Onboarding** Our developed visualization onboarding concept covers different knowledge gaps (Stoiber et al., 2019c, 2022c) such as data, visual encoding, interaction knowledge by providing onboarding messages targeted to these knowledge gaps of the user. Due to the extreme variety of data and requirements (clearing, automated processing, pre-processing, etc.), meaningful data analytics includes some requirements and skills: "[...] graph and network analysis, machine learning, mathematics, statistics, signal, and text processing, among others" (Markl, 2014). Besides, data scientists should also be well prepared to deal with big data and capable of scalable systems programming. Trying to onboard a newcomer to all the above disciplines simultaneously is impossible.

With the current visualization onboarding concept, we cover data analytics onboarding only partially by explaining the pre-processing of data (cf. Chapter 9). Hence, further work

is needed to understand how to design visualization onboarding to support inexperienced users in understanding the involved methods and the limitations of the respective data analytics methods and interpreting the results accordingly.

**Discoverability of Visualization Onboarding** In this doctoral thesis's proposed visualization onboarding concept, we follow the principles of the "pull" model, where a user explicitly asks for help (Horvitz, 1999) by activating the onboarding messages if needed. Further research could explore various design alternatives besides a floating action button to provide a seamless and unobtrusive experience for the user. In the future, our goal is to ensure that users know the onboarding instructions and how to activate them without causing any disruption or annoyance to power users. The emphasis is on making the onboarding process easily accessible and intuitive while catering to the needs of experienced users who may choose to dismiss it.

**Extension of the Model of Visualization Onboarding and Guidance** In Chapter 11, we present our descriptive model of how visualization onboarding and guidance can be effectively integrated into the VA process (Section 11.3). We also confront the concepts of visualization onboarding and guidance and analyze their suitability in different analysis settings. A usage scenario of our model using a stock trading VA tool shows how visualization onboarding and guidance can support users in different analysis phases.

According to Beaudouin-Lafon (Beaudouin-Lafon, 2004), models can exhibit three types of characteristics: (1) *descriptive*: the ability to describe a wide range of existing methods; (2) *evaluative*: enable the assessment of multiple design alternatives, and (3) *generative*: help in designing new methods. Our proposed model is of the descriptive kind (see Section 11.4), as it can systematically describe and present relevant aspects of visualization onboarding and guidance. Therefore, additional research must be carried out to extend the model beyond its descriptive nature to a more generative and evaluative one for visualization onboarding and guidance.

**Improving Synergy between Interface Design and Help Design** From the perspective of interface designers and developers, our integrated visualization onboarding approach enables preparing and maintaining the onboarding material along with the design and implementation of the interface (cf. (Chundury et al., 2023)). Visualization onboarding should be designed and developed in parallel with the VA tool. The field of visualization literacy has yet to be a very visible sub-field, nor is visualization onboarding. More actions to promote the importance of it can be done by conducting more research.

**Expand visualization onboarding concepts to networks and maps** In this thesis, we explored onboarding instructions for bar charts (see Chapter 6), hexbin plots (see Section 9), scatterplots (see Chapter 6 and Chapter 9), parallel coordinates plots (see Chapter 3), Sankey diagrams(see Chapter 3), treemaps (see Chapter 4 and Chapter 8), horizon graphs, and change matrices (see Chapter 6). We did not cover networks or maps yet. The study by Börner et al. (2016) shows that for example, casual users are more familiar with basic charts and maps, but few are familiar with network layouts, and most cannot read network visualizations. Therefore, further research might explore integrating visualization onboarding instructions in network visualizations, e.g., node-link diagrams or geospatial maps.

# CHAPTER 13

# Conclusion



In this chapter, we conclude this thesis with a summary of our work and answer the research question raised in Chapter 1.

This thesis aims to design, implement, and evaluate self-explanatory visualization onboarding concepts that can be integrated into VA tools to support users in interpreting data visualizations and effectively using the tool. We wanted to gain a deeper understanding of how visualization onboarding concepts need to be designed to support users. Therefore, we raised research questions in Section 1.3, which we can answer in detail and results in the following contributions.

The main contributions of this thesis are a *design space* for visualization onboarding (in Part I) and the *conceptual design and implementation* of a JavaScript library called VisAhoi (in Part II) for the semi-automatic generation of onboarding instructions based on the underlying data of the visualization. Besides, we contribute with the results of two *design studies* in biomedical R&D and data journalism, showing the applicability of VisAhoi and the visualization onboarding concept (in Part III). Furthermore, a set of *design actions* and a *descriptive model* are presented in this doctoral thesis (in Part IV).

## Main Question: How can onboarding methods support users with low visualization literacy in correctly interpreting data visualizations?

In our descriptive design space, we discussed the fundamental design dimensions and aspects of visualization onboarding using the *Five W's and How* (Hart, 1996, 2002). WHY is visualization onboarding needed? WHAT is visualization onboarding? WHO is the user? Which knowledge gap does the user have? HOW is visualization onboarding provided? WHERE is visualization onboarding provided? WHEN is visualization onboarding used?

We used the identified literature of our survey to develop our descriptive design space and validate it by applying it to describe and characterize state-of-theart onboarding methods. We also derived design considerations for visualization onboarding concepts in Stoiber et al. (2022c). The results of the empirical studies of the papers and the results of the analysis of the design space are the following:

- explain the visual encoding & interaction concepts (Alper et al., 2017; Bishop et al., 2020; Kwon and Lee, 2016);
- use interactive onboarding approaches, where users can interact with the visualization and with the instructional material (Kwon and Lee, 2016; Tanahashi et al., 2016);
- **concrete experience** and reflection can lead to higher understanding (Kwon and Lee, 2016; Tanahashi et al., 2016);
- use **animations** or **videos** (Alper et al., 2017; Kwon and Lee, 2016; Ruchikachorn and Mueller, 2015) to show the data-to-visual mapping.

To answer the main question further, we first give the answers to our sub-questions.

## **Sub Question 1:** What are the needs of domain experts when learning novel visual representations?

We applied our developed VisAhoi Javascript library to semi-automatically generate onboarding in two design studies (see Part III). In the first design study, we developed an authoring tool for data journalists as a Flourish template for treemap visualizations to integrate onboarding instructions for data-journalistic use cases. We worked closely with Austrian data journalists and iteratively designed and developed our authoring tool. We chose interviews and surveys with domain experts for the problem characterization and abstraction phase. We collected data regarding the domain experts' needs and challenges and current practices regarding visualization onboarding. Based on the results of previous studies (Stoiber et al., 2022b, 2021, 2022d) (see Chapter 5 and Chapter 6.2) and the problem characterization (interviews & survey), we identified five key design requirements (DR) divided into authoring-phase (data journalist) and the reading phase (audience/reader) (see Section 8.4.3). The data journalists' needs regarding the authoring tool are the following:

- Integrate visualization onboarding into a well-known visualization tool;
- Semi-automated generation of onboarding messages;
- Customize onboarding;
- Support readers with tooltips and textual explanations in interpreting complex visualizations;
- Provide context-sensitive onboarding which can be called up during exploration.

In our second design study, we integrated VisAhoi into a VA tool for analyzing highthroughput screening (HTS) data interactively (see Chapter 9). We conducted six interviews with biomedical researchers to collect data regarding the domain experts' needs and challenges and current practices regarding visualization onboarding. Based on the results, we also derived design requirements:

- Explain the underlying data and provide data-related onboarding instructions;
- Support domain experts with tooltips and textual explanations;
- Provide context-sensitive onboarding which can be called up during exploration;
- Position the onboarding button in the toolbar to easily approach it when seeking help.

In summary, the design requirements in both domains are similar as both groups suggest context-sensitive onboarding called up during the exploration with the VA tool. Also, both domains requested tooltips with textual explanations. In contrast, the level of visualization literacy varies between the group of biomedical researchers and the general audience (readers). In our interview with biomedical researchers, we realized they have a high visualization literacy level and therefore need instructions for unfamiliar visualizations, interaction possibilities, and data pre-processing. We know from our study (cf. Chapter 3) and empirical studies (Börner et al., 2019, 2016; Galesic and Garcia-Retamero, 2010) that the visualization literacy level of the general audience is relatively low. Data journalists in our first user study (Chapter 8) would even reduce the complexity of onboarding messages by providing only one onboarding stage not to overwhelm the reader.

## **Sub Question 2:** How do onboarding concepts need to be designed and implemented to fulfill users' needs?

In the first part I of our research, we explored different approaches to answer this subquestion. We designed four onboarding concepts (a step-by-step guide, scrollytelling tutorial, video tutorial, and in-situ onboarding). We also developed proof-of-concept prototypes based on the design considerations (see Chapter 3). We conducted two quantitative comparative studies with MTurk workers and a qualitative comparative user study with students (Lazar et al., 2017, p. 25-44). The main aim of these studies was to investigate the effect of onboarding on user performance and evaluate the subjective user experience. We learned during our research that onboarding systems that are **integrated** into the system are appreciated. Users prefer systems that do not disrupt the workflow. In such systems, important features should be highlighted. Besides, using an easy-to-understand data set and concrete examples on how to read the charts, support, and increase comprehension is vital. **To-the-point descriptions** make it easier to absorb information (step-by-step). Furthermore, videos are appreciated by a minority but can be used to explain interaction techniques along with small animations. Other users have a negative attitude. It depends on the target group whether they can be successfully deployed. Finally, some users ignore onboarding systems, even if they struggle. These users have to be **motivated to use onboarding**.

To explore visualization onboarding in more detail, we explored abstract and concrete onboarding instructions, described in Chapter 4. We assessed them in a quantitative comparative user study with students (Lazar et al., 2017, p. 25-44). The results of our study reflect the discussion in the field of abstract or concrete materials for teaching (De Bock et al., 2011; Kaminski et al., 2008). The presented findings do not answer whether abstract or concrete instructions are more appropriate for onboarding instructions. According to the data, we can infer that both concrete and abstract instructions help us understand visualizations better.

In Chapter 5, we reflect on the results of the analysis of empirical studies, development of the descriptive design space (see Chapter 2), and the conducted user studies (see Chapter 3, and Chapter 4).

The final onboarding design, which is presented in Chapter 6.1, consists of a floating action button (Material Design, 2022) (Material Design Pattern) used to access the

onboarding instructions (see Figure 6.1), tooltips with textual instructions, and in-place annotation anchors. Clicking on the button reveals three circular buttons, representing the three default stages of the onboarding process: *Reading, Interacting,* and *Analyzing* (Figure 6.1 (2)). The onboarding stage *Reading the Chart* explains the visual encoding (DC2) (e.g., size, color); *Interacting with the Chart* onboarding the user to the interaction concept. Moreover, *Analyzing the Chart* guides the user towards further insights (e.g., making comparisons). The step-by-step navigation widget represents the number of in-place annotation anchors/onboarding messages to implement the guided approach (Figure 6.1 (3)). Above the onboarding button, illustrated in Figure 6.1 (3), arrow icons allow users to browse the onboarding messages to support the users in receiving guided help. Users who seek help for a specific part of the visualization, e.g., interaction, can self-explore the onboarding messages by opening a stage (Reading, Interacting, or Analyzing) in the main navigation menu and then clicking on the specific in-place annotation anchor in the visualization.

Based on our two design studies, we derived implications for designing and integrating visualization onboarding by formulating design actions. These design actions help visualization designers and developers better understand how to design and integrate onboarding into VA tools. Chapter 10 presented these design actions.

- DA1: Be aware of different knowledge gaps of users.
- DA2: Explain how to read, interact, and analyze the data visualization.
- DA3: Use concrete onboarding instructions where possible and needed.
- DA4: Use floating text boxes with onboarding messages and in-place annotation anchors.
- DA5: Allow users to hide and show onboarding instructions to call them on-demand.
- DA6: Provide self-exploratory and guided onboarding.
- DA7: Provide in-situ onboarding.

Finally, we present our descriptive model on how visualization onboarding and guidance, as two concepts of user assistance, can be effectively integrated into the VA process in Chapter 11. We assess their differences and similarities of visualization onboarding and guidance along the questions of Why/Who/When/Where/How thereby leading to their practical and thoughtful use in VA. We also apply the descriptive model in the financial domain, illustrating the importance of visualization onboarding and guidance for practical analysis.

## **Sub Question 3:** How can existing educational theories be applied for visualization onboarding?

While developing our descriptive design space, presented in Chapter 2, we particularly surveyed education theories. We integrated this aspect in our design space under the question, "How is onboarding provided?". Learning and educational theories are crucial when it comes to onboarding approaches. Systematic categorization of educational theories was impossible as no taxonomy was available. Therefore, we collect educational theories, which authors described in their scientific publications (e.g., concreteness fading (Alper et al., 2017), learning by analog (Ruchikachorn and Mueller, 2015), and more.) In our analysis, we could not find any unique use of educational theories among the onboarding approaches presented in publications. For our developed visualization onboarding approach, we based it on the concept of microlearning (Hug and Friesen, 2009). Users gain small units of onboarding instruction to a different aspect of the visualization (see Chapter 6.2).

#### **Sub Question 4:** How can a declarative visualization specification be used to (semi)automatically generate onboarding idioms?

We aim for visualization onboarding that can be integrated into VA tools and provide a blueprint for developers and visualization design on phrasing easy-tounderstand onboarding instructions. Therefore, we improved our visualization onboarding concept iteratively based on the lessons learned (see Chapter 6.2). Moreover, we propose a JavaScript library called VisAhoi and build upon three visualization libraries with high-level declarative grammar — plotly.js (Plotly.js, 2023), Vega-lite (Satyanarayan et al., 2017), and eCharts (Li et al., 2018) — to explore how we can semi-automatically generate onboarding instructions based on the currently shown dataset. The resulting VisAhoi library semi-automatically generated onboarding messages as tooltips and in-place annotation anchors. In part II, we describe the resulting onboarding design and underlying design decisions.

We applied the VisAhoi library in two design studies by integrating it into a VA tool for analyzing high-throughput screening (HTS) data interactively (Chapter 9); and by integrating it into a Flourish template to provide a visual authoring tool for visualization onboarding for data journalists (Chapter 8).

This doctoral research investigates the use of visualization onboarding as a means of helping users better understand unfamiliar data visualizations. The thesis explores how to design and integrate such concepts into VA tools and how this can benefit HCI, the visualization community, and designers and developers in biomedical R&D and data journalism (DDJ). The contributions of this thesis include the following:

- A design space for visualization onboarding (Chapter 2).
- A systematic literature review of existing visualization onboarding concepts and their classification into the design space (Chapter 2.4).
- Empirical results of three large-scale comparative user studies to explore different onboarding approaches and the effect on the user performance and subjective user experience (Chapter 3).

- Conceptual design and implementation of a JavaScript library (VisAhoi) for semiautomatic generation of onboarding instructions utilizing the underlying data (in Part II).
- A design study on semi-automatically generated onboarding instructions integrated into a treemap visualization for data-journalism use cases (Chapter 8).
- A design study on semi-automatically generated onboarding instructions for biomedical use cases and the integration in a VA tool (Chapter 9).
- Design actions for designing visualization onboarding methods that address the design, interaction concept, and phrasing of onboarding instructions (Chapter 10).
- A descriptive model of visualization onboarding and guidance as a concept of user assistance to show how these concepts can be utilized in different phases of the analysis to be effective and accepted by the user of a VA system (Chapter 11).

Overall, this doctoral thesis makes important contributions to the field of visualization and human-computer interaction (HCI) by emphasizing the significance of visualization literacy and providing first solutions to support users in interpreting, reading, and extracting information from unfamiliar visual representations. It presents a visualization onboarding concept that can be integrated into VA by using the JavaScript library VisAhoi which semi-automatically generates onboarding instructions based on the underlying data of the visualization. Based on the insights gained from two design studies, design actions, and a descriptive model are proposed to guide the design and development of visualization onboarding methods in the future. Besides, further research is required to explore it in more detail. There are still open research challenges as described in Chapter 12, e.g., the impact of visualization onboarding on the visualization literacy level of users, integrating data analytics onboarding, and discoverability of visualization onboarding.


### Part V

## Bibliography, List of Figures & Tables, Glossary



## Bibliography

- Abellán, A. (2022). State of Data Journalism Survey 2021: 11 surprising findings. https: //datajournalism.com/read/blog/data-journalism-survey-2021. Accessed: 2022-11-14.
- Advizor Solutions (2023). https://www.advizorsolutions.com/. Accessed: 2013-01-19.
- Alper, B., Riche, N. H., Chevalier, F., Boy, J., and Sezgin, M. (2017). Visualization Literacy at Elementary School. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, CHI '17, page 5485–5497, New York, NY, USA. Association for Computing Machinery.
- Alpern, B. and Carter, L. (1991). The hyperbox. In Proceedings of the 2nd Conference on Visualization '91, VIS '91, page 133–139, Washington, DC, USA. IEEE Computer Society Press.
- Amabili, L. (2019). From Storytelling To Scrollytelling: A Short Introduction and Beyond. https://medium.com/nightingale/from-storytelling-to-scrollyt elling-a-short-introduction-and-beyond-fbda32066964. Accessed: 2022-01-31.
- Andrienko, N. V., Lammarsch, T., Andrienko, G. L., Fuchs, G., Keim, D. A., Miksch, S., and Rind, A. (2018). Viewing Visual Analytics as Model Building. *Computer Graphics Forum*, 37.

Angular (2023). https://angular.io/. Accessed: 2022-11-16.

Ant Design (2023). https://ant.design/. Accessed: 2023-01-31.

AppCues (2020). https://www.appcues.com/. Accessed: 2020-12-03.

Appelgren, E. and Nygren, G. (2014). Data Journalism in Sweden: Introducing new methods and genres of journalism into "old" organizations. *Digital Journalism*, 2(3):394– 405.

aSimpleTour (2020). http://alvaroveliz.github.io/aSimpleTour/. Accessed: 2023-01-03.

- Ayres, P. and Sweller, J. (2005). The Split-Attention Principle in Multimedia Learning, page 135–146. Cambridge Handbooks in Psychology. Cambridge University Press.
- Baker, R. S., Corebett, A. T., and Koedinger, K. R. (2001). Toward a Model of Learning Data Representations. In *Proceedings of the Annual Meeting of the Cognitive Science Society*, pages 45–50.
- Banovic, N., Grossman, T., Matejka, J., and Fitzmaurice, G. (2012). Waken: Reverse engineering usage information and interface structure from software videos. In *Proceed*ings of the 25th Annual ACM Symposium on User Interface Software and Technology, UIST '12, pages 83–92. ACM.
- Bao, X., Han, J., and Wang, C. (2021). VolumeVisual: Design and Evaluation of an Educational Software Tool for Teaching and Learning Volume Visualization. In 2021 ASEE Virtual Annual Conference Content Access, Virtual Conference. ASEE Conferences.
- Baur, D. and Stephaner, M. (2018). Everything except the chart. http://webvis-d o.minik.us/. Accessed: 2023-01-11.
- Beaudouin-Lafon, M. (2004). Designing Interaction, Not Interfaces. In Proceedings of the Working Conference on Advanced Visual Interfaces, AVI '04, pages 15–22, New York, NY, USA. ACM.
- Behrisch, M., Blumenschein, M., Kim, N. W., Shao, L., El-Assady, M., Fuchs, J., Seebacher, D., Diehl, A., Brandes, U., Pfister, H., Schreck, T., Weiskopf, D., and Keim, D. A. (2018). Quality metrics for information visualization. *Computer Graphics Forum*, 37:625–662.
- Behrisch, M., Streeb, D., Stoffel, F., Seebacher, D., Matejek, B., Weber, S. H., Mittelstaedt, S., Pfister, H., and Keim, D. (2019). Commercial Visual Analytics Systems-Advances in the Big Data Analytics Field. *IEEE Transactions on Visualization and Computer Graphics*, 25(10):3011–3031.
- Bergman, L., Castelli, V., Lau, T., and Oblinger, D. (2005). DocWizards: A system for authoring follow-me documentation wizards. In *Proceedings of the 18th annual ACM* symposium on User interface software and technology, pages 191–200. ACM.
- Bertini, E., Tatu, A., and Keim, D. (2011). Quality Metrics in High-Dimensional Data Visualization: An Overview and Systematization. *IEEE Transactions on Visualization* and Computer Graphics, 17(12):2203–2212.
- Bider, L. (2020). Stop hesitating, start annotating. https://flourish.studio/bl og/introducing-annotations/. Accessed: 2023-02-03.
- Bishop, F., Zagermann, J., Pfeil, U., Sanderson, G., Reiterer, H., and Hinrichs, U. (2020). Construct-A-Vis: Exploring the Free-Form Visualization Processes of Children. *IEEE Transactions on Visualization and Computer Graphics*, 26(1):451–460.

- Block, J. H. and Burns, R. B. (1976). Mastery learning. Review of Research in Education, 4(1):3–49.
- Bloom, B., Engelhart, M., Furst, E., W.Hill, and Krathwohl, D. (1956). Taxonomy of Educational Objectives, Handbook 1: Cognitive Domain. Longman, New York; New York; London, 2nd edition edition.

bootstrap (2023). https://getbootstrap.com/. Accessed: 2023-01-19.

- Borkin, M. A., Bylinskii, Z., Kim, N. W., Bainbridge, C. M., Yeh, C. S., Borkin, D., Pfister, H., and Oliva, A. (2016). Beyond Memorability: Visualization Recognition and Recall. *IEEE Transactions on Visualization and Computer Graphics*, 22(1):519–528.
- Borkin, M. A., Vo, A. A., Bylinskii, Z., Isola, P., Sunkavalli, S., Oliva, A., and Pfister, H. (2013). What makes a visualization memorable? *IEEE Transactions on Visualization* and Computer Graphics, 19(12):2306–2315.
- Börner, K., Bueckle, A., and Ginda, M. (2019). Data visualization literacy: Definitions, conceptual frameworks, exercises, and assessments. *Proceedings of the National Academy of Sciences*, 116(6):1857–1864.
- Börner, K., Maltese, A., Balliet, R. N., and Heimlich, J. (2016). Investigating Aspects of Data Visualization Literacy Using 20 Information Visualizations and 273 Science Museum Visitors. *Information Visualization*, 15(3):198–213.
- Bostock, M. and Heer, J. (2009). Protovis: A graphical toolkit for visualization. *IEEE Transactions on Visualization and Computer Graphics*, 15(6):1121–1128.
- Bouali, F., Guettala, A., and Venturini, G. (2016). VizAssist: An Interactive User Assistant for Visual Data Mining. The Visual Computer: International Journal of Computer Graphics, 32(11):1447–1463.
- Boy, J., Rensink, R. A., Bertini, E., and Fekete, J. D. (2014). A Principled Way of Assessing Visualization Literacy. *IEEE Transactions on Visualization and Computer Graphics*, 20(12):1963–1972.
- Braasch, J. L. G., Rouet, J.-F., Vibert, N., and Britt, M. A. (2012). Readers' use of source information in text comprehension. *Memory & Cognition*, 40(3):450–465.
- Brown, M. C., McNeil, N. M., and Glenberg, A. M. (2009). Using Concreteness in Education: Real Problems, Potential Solutions. *Child Development Perspectives*, 3(3):160–164.
- Bryan, C., T., W., and B., P. (2007). Comparison and contrast: similarities and differences of teachers' views of effective mathematics teaching and learning from four regions. *ZDM Mathematics Education*, 39:329–340.

- Budiu, R. (2016). The Power Law of Learning: Consistency vs. Innovation in User Interfaces. https://www.nngroup.com/articles/power-law-learning/. Accessed: 2022-01-31.
- Cairo, A. (2013). The Functional Art: An introduction to information graphics and visualization. New Riders.
- Card, S. K., Mackinlay, J. D., and Shneiderman, B. (1999). Readings in Information Visualisation. Using Vision to Think.: Using Vision to Think. Morgan Kaufmann.
- Carroll, J. M. (1987). Interfacing thought: Cognitive aspects of human-computer interaction. Interfacing thought: Cognitive aspects of human-computer interaction. The MIT Press.
- Ceneda, D., Andrienko, N., Andrienko, G., Gschwandtner, T., Miksch, S., Piccolotto, N., Schreck, T., Streit, M., Suschnigg, J., and Tominski, C. (2020). Guide Me in Analysis: A Framework for Guidance Designers. *Computer Graphics Forum*, 39(6):269–288. \_\_eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/cgf.14017.
- Ceneda, D., Gschwandtner, T., May, T., Miksch, S., Schulz, H.-J., Streit, M., and Tominski, C. (2017). Characterizing Guidance in Visual Analytics. *IEEE Transactions* on Visualization and Computer Graphics (VAST'16), 23(1):111–120.
- Ceneda, D., Gschwandtner, T., and Miksch, S. (2019). A review of guidance approaches in visual data analysis: A multifocal perspective. *Computer Graphics Forum*, 38(3):861– 879.
- Ceneda, D., Gschwandtner, T., Miksch, S., and Tominski, C. (2018). Guided visual exploration of cyclical patterns in time-series. In Proceedings of the IEEE Symposium on Visualization in Data Science (VDS). IEEE Computer Society.
- Chandler, P. and Sweller, J. (1991). Cognitive load theory and the format of instruction. Cognition and Instruction, 8(4):293–332.
- chardin.js (2022). http://heelhook.github.io/chardin.js/. Accessed: 2023-02-01.
- Chart.js (2020). Chart.js | open source html5 charts for your website. https://www.chartjs.org/. Accessed: 2023-02-03.
- Chen, C. (2005). Top 10 unsolved information visualization problems. *IEEE Computer Graphics and Applications*, 25(4):12–16.
- Chen, S., Li, J., Andrienko, G., Andrienko, N., Wang, Y., Nguyen, P. H., and Turkay, C. (2020). Supporting story synthesis: Bridging the gap between visual analytics and storytelling. *IEEE Transactions on Visualization and Computer Graphics*, 26(7):2499– 2516.

- Chevalier, F., Riche, N. H., Plaisant, C., Chalbi, A., and Hurter, C. (2016). Animations 25 years later: New roles and opportunities. In *Proceedings of the International Working Conference on Advanced Visual Interfaces*, AVI '16, page 280–287, New York, NY, USA. Association for Computing Machinery.
- Chi, E. H. (2000). A Taxonomy of Visualization Techniques Using the Data State Reference Model. In Proceedings of the IEEE Symposium on Information Visualization (Info Vis), pages 69–76, Washington, DC, USA. IEEE Computer Society.
- Chilana, P. K., Ko, A. J., and Wobbrock, J. O. (2012). LemonAid: selection-based crowdsourced contextual help for web applications. In *Proceedings of the 2012 ACM* annual conference on Human Factors in Computing Systems - CHI '12, page 1549. ACM Press.
- Chundury, P., Yalçin, M. A., Crabtree, J., Mahurkar, A., Shulman, L. M., and Elmqvist, N. (2023). Contextual in situ help for visual data interfaces. *Information Visualization*, 22(1):69–84.
- Committee for Responsible Federal Budget (2021). What's in president biden's american families plan? https://www.crfb.org/blogs/whats-president-bidens-american-families-plan. Accessed: 2023-01-19.
- Contextual Help (2019). An overview of context-sensitive and embedded help formats. https://pronovix.com/blog/overview-context-sensitive-and-emb edded-help-formats. Accessed: 2019-03-14.
- Curcio, F. R. (1987). Comprehension of mathematical relationships expressed in graphs. Journal for Research in Mathematics Education, 18(5):382–393.
- D3 Language (2023). D3 Language. https://d3js.org/. Accessed: 2023-01-18.

Datawrapper (2023). https://www.datawrapper.de/. Accessed: 2023-01-19.

- De Bock, D., Deprez, J., Van Dooren, W., Roelens, M., and Verschaffel, L. (2011). Abstract or Concrete Examples in Learning Mathematics? A Replication and Elaboration of Kaminski, Sloutsky, and Heckler's Study. *Journal for Research in Mathematics Education*, 42(2):109–126.
- De Bruijn, O. and Spence, R. (2008). A New Framework for Theory-Based Interaction Design Applied to Serendipitous Information Retrieval. ACM Transactions on Computer-Human Interaction, 15(1).
- Dhanoa, V., Walchshofer, C., Hinterreiter, A., Stitz, H., Groeller, E., and Streit, M. (2022). A Process Model for Dashboard Onboarding. *Computer Graphics Forum* (*Euro Vis '22*), 41(3):501–513.
- Dix, A., Finlay, J., Abowd, G., and Beale, R. (2004). *Human-Computer Interaction*. Pearson Education, 3rd edition.

- dos Santos, S. and Brodlie, K. (2004). Gaining understanding of multivariate and multidimensional data through visualization. *Computers & Graphics*, 28(3):311–325.
- driver.js (2020). https://github.com/kamranahmedse/driver.js. Accessed: 2023-01-03.
- Duffy, T. M., Lowyck, J., and Jonassen, D. H., editors (1993). Designing Environments for Constructive Learning. Nato ASI Subseries F:. Springer-Verlag.
- ECA&D (2019). European Climate Assessment & Dataset project. http://www.ecad .eu. Accessed: 2023-01-19.
- Echeverria, V., Martinez-Maldonado, R., and Shum, S. B. (2017). Towards data storytelling to support teaching and learning. In *Proceedings of the 29th Australian Conference on Computer-Human Interaction*, OZCHI '17, page 347–351, New York, NY, USA. Association for Computing Machinery.
- Engel, J., Biehler, R., Frischemeier, D., Podworny, S., Schiller, A., and Martignon, L. (2019). Zivilstatistik: Konzept einer neuen Perspektive auf Data Literacy und Statistical Literacyd. AStA Wirtschafts- und Sozialstatistisches Archiv, 13(3):213–244.
- Engels, R. (1996). Planning tasks for knowledge discovery in databases performing task-oriented user-guidance. In Proc. of Conference on Knowledge Discovery and Data Mining (KDD), pages 170–175. AAAI Press.
- Ertmer, P. A. and Newby, T. J. (2013). Behaviorism, cognitivism, constructivism: Comparing critical features from an instructional design perspective. *Performance Improvement Quarterly*, 26(2):43–71.
- Federal Ministry Republic of Austria Finance (2022). Federal ministry of finance the budget of austria, in figures and charts. https://bit.ly/3s6gJFC. Accessed: 2023-01-19.
- Federer, L. M., Lu, Y.-L., and Joubert, D. J. (2016). Data literacy training needs of biomedical researchers. Journal of the Medical Library Association, 104(1):52–57.
- Federico, P., Unger, J., Amor-Amorós, A., Sacchi, L., Klimov, D., and Miksch, S. (2015). Gnaeus: Utilizing clinical guidelines for knowledge-assisted visualisation of EHR cohorts. In Bertini, E. and Roberts, J. C., editors, *EuroVis Workshop on Visual Analytics (EuroVA)*. The Eurographics Association.
- Federico, P., Wagner, M., Rind, A., Amor-Amorós, A., Miksch, S., and Aigner, W. (2017). The role of explicit knowledge: A conceptual model of knowledge-assisted visual analytics. In *Proceedings IEEE Conference on Visual Analytics Science and Technology (VAST)*. IEEE.

- Feigenbaum, A., Einar, T., Weissmann, D., and Demirkol, O. (2016). Visualising data stories together: Reflections on data journalism education from the Bournemouth University Datalabs Project. *Journalism Education*, 2(5):59–74.
- Fernquist, J., Grossman, T., and Fitzmaurice, G. (2011a). Sketch-sketch revolution: An engaging tutorial system for guided sketching and application learning. In *Proceedings* of the 24th annual ACM symposium on User interface software and technology, pages 373–382. ACM.
- Fernquist, J., Grossman, T., and Fitzmaurice, G. (2011b). Sketch-sketch revolution: An engaging tutorial system for guided sketching and application learning. In *Proceedings* of the 24th Annual ACM Symposium on User Interface Software and Technology, UIST '11, pages 373–382, New York, NY, USA. ACM.
- Few, S. (2012). Show Me the Numbers: Designing Tables and Graphs to Enlighten. Analytics Press, 2nd edition.
- Firat, E. E., Denisova, A., and Laramee, R. S. (2020). Treemap Literacy: A Classroom-Based Investigation. In Romero, M. and Sousa Santos, B., editors, *Eurographics 2020 - Education Papers*. The Eurographics Association.
- Firat, E. E., Joshi, A., and Laramee, R. S. (2022a). Interactive visualization literacy: The state-of-the-art. SAGE Information Visualization, 21(3):0285–310.
- Firat, E. E., Joshi, A., and Laramee, R. S. (2022b). Visite: Visualization literacy and evaluation. *IEEE Computer Graphics and Applications*, 42(3):99–107.

Flourish (2022). https://flourish.studio/. Accessed: 2022-11-14.

- Flourish create story (2022). https://help.flourish.studio/article/13creating-a-story. Accessed: 2022-10-25.
- Flourish Annotations (2023). https://help.flourish.studio/article/180-h ow-to-add-annotations-to-your-visualization. Accessed: 2023-03-25.
- Flourish SDK (2023). https://developers.flourish.studio/sdk/introdu ction/. Accessed: 2023-01-19.
- Freedman, E. G. and Shah, P. (2002). Toward a model of knowledge-based graph comprehension. In Hegarty, M., Meyer, B., and Narayanan, N. H., editors, *Diagrammatic Representation and Inference*, Lecture Notes in Computer Science, pages 18–30. Springer.
- Friel, S. N., Curcio, F. R., and Bright, G. W. (2001). Making sense of graphs: Critical factors influencing comprehension and instructional implications. *Journal for Research* in Mathematics Education JRME, 32(2):124 – 158.

- Friendly, M. and Denis, D. (2005). The early origins and development of the scatterplot. *Journal of the History of the Behavioral Sciences*, 41(2):103–130. \_\_eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/jhbs.20078.
- Fyfe, E. R., McNeil, N. M., and Borjas, S. (2015). Benefits of "concreteness fading" for children's mathematics understanding. *Learning and Instruction*, 35:104–120.
- Galesic, M. and Garcia-Retamero, R. (2010). Graph Literacy: A Cross-Cultural Comparison. *Medical Decision Making*, 31(3):444–457. Publisher: SAGE PublicationsSage CA: Los Angeles, CA.
- Gehlenborg, N., O'Donoghue, S. I., Baliga, N. S., Goesmann, A., Hibbs, M. A., Kitano, H., Kohlbacher, O., Neuweger, H., Schneider, R., Tenenbaum, D., and Gavin, A.-C. (2010). Visualization of omics data for systems biology. *Nature Methods*, 7(3):S56–S68.
- Gerrig, R. J., Zimbardo, P. G., Campbell, A. J., Cumming, S. R., and Wilkes, F. J. (2015). Psychology and life. Pearson Higher Education AU.
- Glenberg, A., Gutierrez, T., Levin, J. R., Japuntich, S., and Kaschak, M. P. (2004). Comparison and contrast: similarities and differences of teachers' views of effective mathematics teaching and learning from four regions. *Journal of Educational Psychol*ogy, 96:424–436.
- Göbel, F., Kiefer, P., Giannopoulos, I., Duchowski, A. T., and Raubal, M. (2018). Improving map reading with gaze-adaptive legends. In *Proceedings of the 2018 ACM Symposium on Eye Tracking Research & Applications*, ETRA '18, pages 29:1–29:9, New York, NY, USA. ACM.
- Gotz, D. and Wen, Z. (2009). Behavior-driven visualization recommendation. In Proceedings of the 14th International Conference on Intelligent User Interfaces, IUI '09, page 315–324, New York, NY, USA. Association for Computing Machinery.
- Grammel, L., Tory, M., and Storey, M. (2010). How information visualization novices construct visualizations. *IEEE Transactions on Visualization and Computer Graphics*, 16(6):943–952.
- Grossman, T. and Fitzmaurice, G. (2010a). ToolClips: An investigation of contextual video assistance for functionality understanding. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 1515–1524. ACM.
- Grossman, T. and Fitzmaurice, G. (2010b). Toolclips: An investigation of contextual video assistance for functionality understanding. In *Proceedings of the SIGCHI Conference* on Human Factors in Computing Systems, CHI '10, pages 1515–1524, New York, NY, USA. ACM.
- Grossman, T., Fitzmaurice, G., and Attar, R. (2009). A survey of software learnability: Metrics, methodologies and guidelines. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '09, pages 649–658. ACM.

- GuideChip (2022). https://github.com/samhynds/Orient-JS. Accessed: 2023-02-01.
- Harman, G. (2010). Technology, objects and things in Heidegger. Cambridge Journal of Economics, 34(1):17–25.
- Harman, T. and Lange, J. (2022). States with split delegations in u.s. senate at historic low. https://www.reuters.com/graphics/USA-ELECTION/SPLIT-DELEGA TION/mopaknrdopa/. Accessed: 2023-04-15.
- Harper, J. and Agrawala, M. (2014). Deconstructing and restyling d3 visualizations. In Proceedings of the 27th Annual ACM Symposium on User Interface Software and Technology, UIST '14, page 253–262, New York, NY, USA. Association for Computing Machinery.
- Harper, J. and Agrawala, M. (2018). Converting Basic D3 Charts into Reusable Style Templates. *IEEE Transactions on Visualization and Computer Graphics*, 24(3):1274– 1286. Conference Name: IEEE Transactions on Visualization and Computer Graphics.
- Hart, G. (1996). The five W's: An old tool for the new task of task analysis. *Technical communication*, 43(2):139–145.
- Hart, G. (2002). The five w's of online help systems. http://www.geoff-hart.com/articles/2002/fivew.htm. Accessed: 2023-01-12.
- Heer, J., Kong, N., and Agrawala, M. (2009). Sizing the horizon: The effects of chart size and layering on the graphical perception of time series visualizations. In *Proceedings* of the SIGCHI Conference on Human Factors in Computing Systems, CHI '09, page 1303–1312, New York, NY, USA. Association for Computing Machinery.
- Heinrich, J. and Broeksema, B. (2023). Car data. http://parallelcoordinates. de/paco/. Accessed: 2023-01-19.
- Hohman, F., Kahng, M., Pienta, R., and Chau, D. H. (2019). Visual analytics in deep learning: An interrogative survey for the next frontiers. *IEEE Transactions on Visualization and Computer Graphics*, 25(8):2674–2693.
- Holton, D. and Clarke, D. (2006). Scaffolding and metacognition. International Journal of Mathematical Education in Science and Technology, 37(2):127–143.
- Holtz, Y. and Healy, C. (2018). From data to Viz | Find the graphic you need. https: //www.data-to-viz.com/. Accessed: 2023-01-19.
- Holyoak, K. J. and Thagard, P. (1995). Mental leaps: Analogy in creative thought. A bradford book. MIT press, Cambridge, MA, US; London, England.
- Hopscotch (2022). https://github.com/LinkedInAttic/hopscotch. Accessed: 2023-02-01.

Horvitz, E. (1999). Principles of mixed-initiative user interfaces. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI '99, page 159–166, New York, NY, USA. Association for Computing Machinery.

Hotjar (2023). https://www.hotjar.com/. Accessed: 2023-01-19.

- Hug, T. and Friesen, N. (2009). Outline of a microlearning agenda. *eLearning Papers*, 16(16).
- Hullman, J., Diakopoulos, N., and Adar, E. (2013). Contextifier: Automatic generation of annotated stock visualizations. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '13, page 2707–2716, New York, NY, USA. Association for Computing Machinery.
- Huron, S., Carpendale, S., Thudt, A., Tang, A., and Mauerer, M. (2014a). Constructive visualization. In *Proceedings of the 2014 Conference on Designing Interactive Systems*, DIS '14, page 433–442, New York, NY, USA. Association for Computing Machinery.
- Huron, S., Jansen, Y., and Carpendale, S. (2014b). Constructing visual representations: Investigating the use of tangible tokens. *IEEE Transactions on Visualization and Computer Graphics*, 20(12):2102–2111.
- Huron, S., Vuillemot, R., and Fekete, J.-D. (2013). Visual sedimentation. IEEE Transactions on Visualization and Computer Graphics, 19(12):2446–2455.
- IBM Cognos Analytics (2019). https://www.ibm.com/at-en/marketplace/bu siness-intelligence. Accessed: 2019-03-14.
- Intercom (2020). https://www.intercom.com/. Accessed:2020-12-03.
- intro.js (2020). Intro.js lightweight, user-friendly onboarding tour library. https://introjs.com/. Accessed: 2023-01-03.
- IOC (2023). Olympic World Library. https://library.olympic.org/. Accessed: 2023-01-19.
- Isenberg, T., Isenberg, P., Chen, J., Sedlmair, M., and Möller, T. (2013). A systematic review on the practice of evaluating visualization. *IEEE Trans. Visualization and Computer Graphics*, 19(12):2818–2827.
- Janice Ginny, R. (2014). Guidelines for Document Designers. Technical Report A Product of Document Design Project Funded by National Institute of Education, American Institutes for Research.
- Kaggle.com (2023a). Chocolate Bar Rating. https://www.kaggle.com/rtatman /chocolate-bar-ratings. Accessed: 2023-01-20.
- Kaggle.com (2023b). Spotify Classification. https://www.kaggle.com/geomack/s potifyclassification. Accessed: 2023-01-19.

- Kai (2021). Parallel Coordinates. https://github.com/syntagmatic/paralle l-coordinates. Accessed: 2023-01-019.
- Kaminski, J. A., Sloutsky, V. M., and Heckler, A. F. (2008). The advantage of abstract examples in learning math. *Science*, 320(5875):454–455.
- Kandel, S., Paepcke, A., Hellerstein, J., and Heer, J. (2011). Wrangler: Interactive visual specification of data transformation scripts. In *Proceedings of the SIGCHI Conference* on Human Factors in Computing Systems, CHI '11, pages 3363–3372, New York, NY, USA. Association for Computing Machinery.
- Kandel, S., Parikh, R., Paepcke, A., Hellerstein, J. M., and Heer, J. (2012). Profiler: Integrated statistical analysis and visualization for data quality assessment. In *Proceedings of the International Working Conference on Advanced Visual Interfaces*, AVI '12, page 547–554, New York, NY, USA. Association for Computing Machinery.
- Kandlhofer, M., Steinbauer, G., Lassnig, J., Baumann, W., Plomer, S., Ballagi, A., and Alfoldi, I. (2019). Enabling the Creation of Intelligent Things: Bringing Artificial Intelligence and Robotics to Schools. In *IEEE Frontiers in Education Conference* (*FIE*).
- Kang, H., Plaisant, C., and Shneiderman, B. (2003). New Approaches to Help Users Get Started with Visual Interfaces: Multi-layered Interfaces and Integrated Initial Guidance. In *Proceedings of the 2003 Annual National Conference on Digital Government Research*, dg.o '03, pages 1–6. Digital Government Society of North America.
- Keim, D. A., Kohlhammer, J., Ellis, G., and Mansmann, F. (2010). Mastering the information age solving problems with visual analytics. *Eurographics*.
- Kelleher, C. and Pausch, R. (2005). Stencils-based tutorials: Design and evaluation. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI '05, pages 541–550. ACM.
- Kerpedjiev, P., Abdennur, N., Lekschas, F., McCallum, C., Dinkla, K., Strobelt, H., Luber, J. M., Ouellette, S. B., Azhir, A., Kumar, N., Hwang, J., Lee, S., Alver, B. H., Pfister, H., Mirny, L. A., Park, P. J., and Gehlenborg, N. (2018). Higlass: web-based visual exploration and analysis of genome interaction maps. *Genome Biology*, 19(1):125.
- Keshif (2023). Keshif | Data Made Explorable. https://www.keshif.me/. Accessed: 2023-01-18.
- Kim, D. H., Setlur, V., and Agrawala, M. (2021). Towards understanding how readers integrate charts and captions: A case study with line charts. In *Proceedings of the* 2021 CHI Conference on Human Factors in Computing Systems, CHI '21, New York, NY, USA. Association for Computing Machinery.
- Kirk, A. (2016). Data Visualisation: A Handbook for Data Driven Design. Sage Publications Ltd.

- Kitchenham, B. (2004). Procedures for performing systematic reviews. Keele, UK, Keele University, 33(2004):1–26.
- Klein, H. J., Polin, B., and Sutton, K. L. (2015). Specific onboarding practices for the socialization of new employees. *International Journal of Selection and Assessment*, 23(3):263–283.
- Knaflic, C. N. (2015). Storytelling with Data: A Data Visualization Guide for Business Professionals. Wiley, Hoboken, New Jersey, 1. edition edition.
- Kolb, D. (1984). Experiential Learning: Experience As The Source Of Learning And Development, volume 1. Prentice-Hall. Journal Abbreviation: Journal of Business Ethics Publication Title: Journal of Business Ethics.
- Kosslyn, S. M. (1989). Understanding charts and graphs. Applied Cognitive Psychology, 3(3):185–225.
- Kuckartz, U. (2014). Qualitative Text Analysis: A Guide to Methods, Practice & Using Software. SAGE Publications Ltd, 1 Oliver's Yard, 55 City Road, London EC1Y 1SP United Kingdom.
- Kumar, A. (2017). Why user onboarding is essential for your application? https: //yourstory.com/mystory/41d0f708f7-why-user-onboarding-is-es sential-for-your-application-. Accessed: 2023-01-21.
- Kwon, B. C. and Lee, B. (2016). A comparative evaluation on online learning approaches using parallel coordinate visualization. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '16, pages 993–997. ACM Press.
- labella.js (2022). https://twitter.github.io/labella.js/. Accessed: 2023-01-03.
- Lafreniere, B., Grossman, T., and Fitzmaurice, G. (2013). Community enhanced tutorials: Improving tutorials with multiple demonstrations. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '13, page 1779–1788, New York, NY, USA. Association for Computing Machinery.
- Lai, C., Lin, Z., Jiang, R., Han, Y., Liu, C., and Yuan, X. (2020). Automatic annotation synchronizing with textual description for visualization. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, CHI '20, page 1–13, New York, NY, USA. Association for Computing Machinery.
- Lamin, A. and Zaheer, S. (2012). Wall street vs. main street: Firm strategies for defending legitimacy and their impact on different stakeholders. Organization Science, 23(1):47–66.

- Lammarsch, T., Aigner, W., Bertone, A., Miksch, S., and Rind, A. (2011). Towards a concept how the structure of time can support the visual analytics process. In Miksch, S. and Santucci, G., editors, *EUROVA*, pages 9–12. Eurographics, Eurographics.
- Lazar, J., Feng, J. H., and Hochheiser, H. (2017). Research Methods in Human-Computer Interaction. Morgan Kaufmann, Boston, second edition.
- Lee, S., Kim, S.-H., Hung, Y.-H., Lam, H., Kang, Y.-A., and Yi, J. S. (2016). How do people make sense of unfamiliar visualizations?: A grounded model of novice's information visualization sensemaking. *IEEE Transactions on Visualization and Computer Graphics*, 22(1):499–508.
- Lee, S., Kim, S.-H., and Kwon, B. C. (2017). VLAT: Development of a visualization literacy assessment test. *IEEE Transactions on Visualization and Computer Graphics*, 23(1):551–560.
- Li, D., Mei, H., Shen, Y., Su, S., Zhang, W., Wang, J., Zu, M., and Chen, W. (2018). ECharts: A declarative framework for rapid construction of web-based visualization. *Visual Informatics*, 2(2):136–146.
- Lu, S. (2022). D3 Annotation. https://d3-annotation.susielu.com/. Accessed: 2022-01-31.
- Luboschik, M., Maus, C., Schulz, H.-J., Schumann, H., and Uhrmacher, A. (2012). Heterogeneity-based guidance for exploring multiscale data in systems biology. In *BIOVIS*, pages 33–40. IEEE.
- MacCaw, A., Blum, S., Fishkin, R., and Cabane, G. (2022). Text Analysis. https://monkeylearn.com/. Accessed: 2023-01-31.
- Mackinlay, J. (1986). Automating the design of graphical presentations of relational information. ACM Trans. Graphics, 5(2):110–141.
- Manning, J. P. (2005). Rediscovering Froebel: A Call to Re-examine his Life & Gifts. Early Childhood Education Journal, 32(6):371–376.
- Marai, G. E. (2015). Visual scaffolding in integrated spatial and nonspatial analysis. In *EuroVis Workshop on Visual Analytics (EuroVA)*. The Eurographics Association.
- Markl, V. (2014). Breaking the chains: On declarative data analysis and data independence in the big data era. Proc. VLDB Endow., 7(13):1730–1733.
- Martin, I., Wenqing, C., Shuzhan, W., Christine P., T., and Chaoli, W. (2020). GraphVisual: Design and Evaluation of a Web-Based Visualization Tool for Teaching and Learning Graph Visualization. In 2020 ASEE Virtual Annual Conference Content Access, Virtual Online. ASEE Conferences. https://peer.asee.org/34715.

- Matejka, J., Grossman, T., and Fitzmaurice, G. (2011). Ambient help. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI '11, pages 2751–2760, New York, NY, USA. ACM.
- Material Design (2022). Buttons: floating action button. https://material.io/co mponents/buttons-floating-action-button. Accessed: 2022-10-18.
- May, T., Bannach, A., Davey, J., Ruppert, T., and Kohlhammer, J. (2011). Guiding feature subset selection with an interactive visualization. In *Proc. IEEE Symp. Visual Analytics Science and Technology (VAST)*, pages 111–120. IEEE.
- May, T., Steiger, M., Davey, J., and Kohlhammer, J. (2012). Using signposts for navigation in large graphs. CGF, 31:985–994.
- Mayer, R. E. (1995). The search for insight: Grappling with Gestalt psychology's unanswered questions. In Sternberg, R. J. and Davidson, J. E., editors, *The nature of insight*, pages 3–12. MIT Press, Cambridge, MA, US.
- Mayring, P. (2010). Qualitative Inhaltsanalyse. In Mey, G. and Mruck, K., editors, Handbuch Qualitative Forschung in der Psychologie, pages 601–613. VS Verlag für Sozialwissenschaften, Wiesbaden.
- McCane, B., Ott, C., Meek, N., and Robins, A. (2017). Mastery learning in introductory programming. In *Proceedings of the Nineteenth Australasian Computing Education Conference*, ACE '17, page 1–10, New York, NY, USA. Association for Computing Machinery.
- Microsoft Power BI (2019). https://powerbi.microsoft.com/de-de/. Accessed: 2019-03-14.
- Miksch, S. and Aigner, W. (2014). A matter of time: Applying a data-users-tasks design triangle to visual analytics of time-oriented data. *Computers & Graphics*, 38:286–290.
- Munzner, T. (2009). A nested process model for visualization design and validation. IEEE Transactions on Visualization and Computer Graphics (InfoVis '09), 15(6):921–928.
- Munzner, T. (2014). Visualization Analysis and Design. AK Peters, Boca Raton.
- Muth, L. C. (2022). DataWrapper text annotations. https://blog.datawrapper .de/better-more-responsive-annotations-in-datawrapper-data-v isualizations/. Accessed: 2023-02-03.
- Nathan, M. J. (2012). Rethinking formalisms in formal education. *Educational Psychologist*, 47(2):125–148.
- Nelogica (2021). Nelogica Tecnologia e Informação para o Mercado Financeiro. https: //www.nelogica.com.br/, Accessed: 06/2021.

- Niederer, C., Stitz, H., Hourieh, R., Grassinger, F., Aigner, W., and Streit, M. (2017). TACO: Visualizing Changes in Tables Over Time. *IEEE Transactions on Visualization* and Computer Graphics (InfoVis '17), 24(1):677–686. tex.ids: NiedererTACOVisualizingChanges2018.
- Nielsen, J. (2012). Thinking aloud: The #1 usability tool. https://www.nngrou p.com/articles/thinking-aloud-the-1-usability-tool/. Accessed: 2022-12-7.
- Nielsen, J. (2022). Severity ratings for usability problems. https://www.nngroup. com/articles/how-to-rate-the-severity-of-usability-problems/. Accessed: 2022-10-12.
- Nolan, D. and Perrett, J. (2016). Teaching and Learning Data Visualization: Ideas and Assignments. The American Statistician, 70(3):260–269.
- NPM Trends (2020). Npm trends. https://www.npmtrends.com/chart.js-vsecharts-vs-plotly.js-vs-vega-lite. Accessed: 2020-12-03.
- Obeid, J. and Hoque, E. (2020). Chart-to-Text: Generating Natural Language Descriptions for Charts by Adapting the Transformer Model. arXiv:2010.09142 [cs].
- O'Handley, B. J., Wu, Y., Duan, H., and Wang, C. (2022). TreeVisual: Design and Evaluation of a Web-Based Visualization Tool for Teaching and Learning Tree Visualization. *Proceedings of American Society for Engineering Education AnnualConference*.
- Ola, O. and Sedig, K. (2017). Health literacy for the general public: Making a case for non-trivial visualizations. *Informatics*, 4(4):33.
- Orient (2019). https://github.com/samhynds/Orient-JS. Accessed: 2023-02-01.
- Palmiter, S. and Elkerton, J. (1991). An evaluation of animated demonstrations of learning computer-based tasks. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '91, pages 257–263, New York, NY, USA. ACM.
- Pavlopoulos, G. A., Malliarakis, D., Papanikolaou, N., Theodosiou, T., Enright, A. J., and Iliopoulos, I. (2015). Visualizing genome and systems biology: technologies, tools, implementation techniques and trends, past, present and future. *GigaScience*, 4(1).
- Peeck, J. (1994). Enhancing graphic-effects in instructional texts: Influencing learning activities. In Schnotz, W. and Kulhavy, R. W., editors, Advances in Psychology, volume 108 of Comprehension of Graphics, pages 291–301. North-Holland.
- Peng, I., Firat, E. E., Laramee, R. S., and Joshi, A. (2022). Evaluating Bloom's Taxonomy-based Learning Modules for Parallel Coordinates Literacy. In Bourdin, J.-J. and Paquette, E., editors, *Eurographics 2022 - Education Papers*. The Eurographics Association.

- Perkhofer, L. M., Hofer, P., Walchshofer, C., Plank, T., and Jetter, H.-C. (2019). Interactive visualization of big data in the field of accounting: A survey of current practice and potential barriers for adoption. *Journal of Applied Accounting Research*, 20(4):497–525.
- Pirolli, P. and Card, S. (2005). The sensemaking process and leverage points for analyst technology as identified through cognitive task analysis. In Proc. of the Int. Conf. on Intelligence Analysis, volume 5, pages 2–4.
- Plotly.js (2023). https://plotly.com/javascript/. Accessed: 2023-01-10.
- Poco, J. and Heer, J. (2017). Reverse-Engineering Visualizations: Recovering Visual Encodings from Chart Images. *Computer Graphics Forum*, 36(3):353–363.
- Pohl, M., Smuc, M., and Mayr, E. (2012). The user puzzle explaining the interaction with visual analytics systems. *IEEE Transactions on Visualization and Computer Graphics*, 18(12):2908–2916.
- Pongnumkul, S., Dontcheva, M., Li, W., Wang, J., Bourdev, L., Avidan, S., and Cohen, M. F. (2011). Pause-and-play: Automatically linking screencast video tutorials with applications. In *Proceedings of the 24th Annual ACM Symposium on User Interface* Software and Technology, UIST '11, pages 135–144, New York, NY, USA. ACM.
- Preece, J., Rogers, Y., and Sharp, H. (2015). Interaction Design: Beyond human-computer interaction (4th ed). John Wiley & Sons, Chichester.
- Probasco, J. (2021). What's in joe biden's \$2.3 trillion american jobs plan? https://www.investopedia.com/what-s-in-joe-biden-s-usd2-trillion-americanjobs-plan-5120273. Accessed: 2022-05-02.
- q nzz (2023). https://www.nzz.ch/storytelling/tech. Accessed: 2023-01-16.
- QGIS (2022). https://www.qgis.org/de/site/. Accessed: 2023-01-09.
- QlikTech QlikView Visualization Tool (2019). https://www.qlik.com/us/produc ts/qlikview. Accessed: 2023-01-18.
- quartz (2022). Chartbuider by quartz. https://quartz.github.io/Chartbuil der/. Accessed: 2023-01-19.
- Raw Graphs (2023). https://www.rawgraphs.io/. Accessed: 2023-01-19.
- React (2023). https://reactjs.org/. Accessed: 2022-11-16.
- React Vega (2022). https://github.com/vega/react-vega. Accessed: 2022-01-31.
- Reilly, S. (2017). The need to help journalists with data and information visualization. *IEEE Computer Graphics and Applications*, 37(2):8–10.

- Ren, D., Brehmer, M., Bongshin Lee, Höllerer, T., and Choe, E. K. (2017). Chartaccent: Annotation for data-driven storytelling. In 2017 IEEE Pacific Visualization Symposium (Pacific Vis), pages 230–239.
- Renaissance (2022). Atos readability formula. https://www.renaissance.com/products/acceleratedreader/atos-and-text-complexity/. Accessed: 2022-05-04.
- Reveal, G. (2019). https://www.juiceanalytics.com/design-principles /gradual-reveal. Accessed: 2023-01-14.
- Ribecca, S. (2020). The data visualisation catalogue. https://datavizcatalogue .com/. Accessed: 2020-03-12.
- Riche, N. H., Hurter, C., Diakopoulos, N., and Carpendale, S. (2018). Data-Driven Storytelling. A K Peters/CRC Press, Boca Raton.
- Rind, A., Wagner, M., and Aigner, W. (2019). Towards a structural framework for explicit domain knowledge in visual analytics. In *IEEE Workshop on Visual Analytics* in Healthcare (VAHC), pages 33–40.
- RTR (2022). Medientransparenz Datenmeldung. https://www.rtr.at/medie n/was\_wir\_tun/medientransparenz/startseite.de.html. Accessed: 2023-01-31.
- Ruchikachorn, P. and Mueller, K. (2015). Learning visualizations by analogy: Promoting visual literacy through visualization morphing. *IEEE Transactions on Visualization* and Computer Graphics, 21(9):1028–1044.
- Sacha, D., Senaratne, H., Kwon, B. C., Ellis, G., and Keim, D. A. (2016). The role of uncertainty, awareness, and trust in visual analytics. *IEEE Transactions on Visualization* and Computer Graphics, 22(1):240–249.
- Sacha, D., Stoffel, A., Stoffel, F., Kwon, B. C., Ellis, G., and Keim, D. (2014). Knowledge generation model for visual analytics. *IEEE Trans. Visualization and Computer Graphics*, 20(12):1604–1613.
- Salazar, K. (2022). Evaluate interface learnability with cognitive walkthroughs. https: //www.nngroup.com/articles/cognitive-walkthroughs/. Accessed: 2022-12-7.
- SAP Lumira (2019). https://saplumira.com/. Accessed: 2019-03-14.
- Saraiya, P., North, C., and Duca, K. (2005). An insight-based methodology for evaluating bioinformatics visualizations. *IEEE Transactions on Visualization and Computer Graphics*, 11(4):443–456.
- SAS JMP (2019). https://www.jmp.com/en\_us/software/data-analysis-s oftware.html. Accessed: 2019-03-14.

- SAS Visual Analytics (2019). https://www.sas.com/en\_us/software/visualanalytics.html. Accessed: 2019-03-14.
- Satyanarayan, A., Moritz, D., Wongsuphasawat, K., and Heer, J. (2017). Vega-lite: A grammar of interactive graphics. *IEEE Transactions on Visualization and Computer Graphics*, 23(1):341–350.
- Satyanarayan, A., Russell, R., Hoffswell, J., and Heer, J. (2016). Reactive vega: A streaming dataflow architecture for declarative interactive visualization. *IEEE Transactions* on Visualization and Computer Graphics, 22(1):659–668.
- Sawyer, R. K. and Greeno, J. G. (2009). Situativity and learning. In Aydede, M. and Robbins, P., editors, *The Cambridge Handbook of Situated Cognition*, pages 347–367. Cambridge: Cambridge University Press.
- Schliemann, A. D. and Carraher, D. W. (2002). The Evolution of Mathematical Reasoning: Everyday versus Idealized Understandings. *Developmental Review*, 22(2):242–266.
- Schmidt, A. (2007). Microlearning and the Knowledge Maturing Process: Towards Conceptual Foundations for Work-Integrated Microlearning Support. In Hug, T., Lindner, M., and Bruck, P. A., editors, *Proceedings of the 3rd International Microlearning 2007 Conference*, Innsbruck, Austria. Innsbruck University Press.
- Schreier, M. (2012). Qualitative Content Analysis in Practice. Sage Publications.
- Schulz, H.-J., Streit, M., May, T., and Tominski, C. (2013). Towards a Characterization of Guidance in Visualization. Poster at IEEE Conference on Information Visualization (InfoVis).
- Sedlmair, M., Meyer, M., and Munzner, T. (2012). Design Study Methodology: Reflections from the Trenches and the Stacks. *IEEE Transactions on Visualization and Computer Graphics*, 18(12):2431–2440.
- Segel, E. and Heer, J. (2010). Narrative visualization: Telling stories with data. IEEE Transactions on Visualization and Computer Graphics, 16(6):1139–1148.
- Sequence, T. (2022). https://tooltip-sequence.netlify.app/. Accessed: 2023-02-01.
- Shah, P., Freedman, E. G., and Vekiri, I. (2005). The comprehension of quantitative information in graphical displays. In *The Cambridge Handbook of Visuospatial Thinking*, Cambridge Handbooks in Psychology, pages 426–476. Cambridge University Press.

Shepherd.js (2022). https://shepherdjs.dev/. Accessed: 2023-02-01.

Shi, Y., Lan, X., Li, J., Li, Z., and Cao, N. (2021). Communicating with motion: A design space for animated visual narratives in data videos. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, CHI '21, New York, NY, USA. Association for Computing Machinery.

- Shih, M., Rozhon, C., and Ma, K. (2019). A declarative grammar of flexible volume visualization pipelines. *IEEE Transactions on Visualization and Computer Graphics*, 25(1):1050–1059.
- Shneiderman, B. (1992). Tree visualization with tree-maps: 2-d space-filling approach. ACM Trans. Graph., 11(1):92–99.
- Shneiderman, B. (2020). Data visualization's breakthrough moment in the COVID-19 crisis. https://medium.com/nightingale/data-visualizations-bre akthrough-moment-in-the-covid-19-crisis-ce46627c7db5. Accessed: 2020-09-14.
- Shneiderman, B. (2022). Content design london readability guidelines. https://
  readabilityguidelines.co.uk/clear-language/simple-sentences/.
  Accessed: 2022-05-04.
- Silver, M. S. (1991). Decisional guidance for computer-based decision support. MIS quarterly, pages 105–122.
- Sloutsky, V., Kaminski, J., and Heckler, A. (2005). The advantage of simple symbols for learning and transfer. *Journal of Educational Psychology*, 12:508–513.
- smartour (2019). https://github.com/jrainlau/smartour. Accessed: 2023-01-31.
- Smith, S. and Mosier, J. (1986). Guidelines for designing user interface software. Technical Report ESD-TR-86-278, Mitre Corporation, Bedford MA.
- Son, J. Y., Smith, L. B., and Goldstone, R. L. (2008). Simplicity and generalization: Short-cutting abstraction in children's object categorizations. *Cognition*, 108:626–638.
- Spence, R. (2014). Information Visualization. Springer, Cham, 3 edition.
- Stoiber, C., Ceneda, D., Wagner, M., Schetinger, V., Gschwandtner, T., Streit, M., Miksch, S., and Aigner, W. (2022a). Perspectives of visualization onboarding and guidance in VA. Visual Informatics, 6(1):68–83.
- Stoiber, C., Grassinger, F., and Aigner, W. (2022b). Abstract and concrete materials: What to use for visualization onboarding for a treemap visualization? In *Proceedings* of the 15th International Symposium on Visual Information Communication and Interaction, VINCI '22, New York, NY, USA. Association for Computing Machinery.
- Stoiber, C., Grassinger, F., Pohl, M., Stitz, H., Streit, M., and Aigner, W. (2019a). Visualization onboarding: Learning how to read and use visualizations. In *IEEE Workshop on Visualization for Communication*. IEEE.
- Stoiber, C., Pohl, M., and Aigner, W. (conditionally-accepted). Design Actions for the Design of Visualization Onboarding Methods. In *IEEE VIS Workshop on Visualization Education, Literacy, and Activities*, Melbourne, Australia. IEEE.

- Stoiber, C., Radkohl, S., Grassinger, F., Moitzi, D., Stitz, H., Goldgruber, E., Girardi, D., and Aigner, W. (2023). Authoring tool for Data Journalists integrating Self-Explanatory Visualization Onboarding Concept for a Treemap Visualization. In *Proceedings CHItaly* 2023, Turin, Piedmont, Italy. ACM.
- Stoiber, C., Rind, A., Grassinger, F., Gutounig, R., Goldgruber, E., Sedlmair, M., Emrich, S., and Aigner, W. (2019b). netflower: Dynamic network visualization for data journalists. *Comput. Graph. Forum*, 38(3):699–711.
- Stoiber, C., Wagner, M., Ceneda, D., Pohl, M., Gschwandtner, T., Miksch, S., Streit, M., Girardi, D., and Aigner, W. (2019c). Knowledge-assisted visual analytics meets guidance and onboarding. In *IEEE Workshop on Application Spotlight*.
- Stoiber, C., Wagner, M., Grassinger, F., Pohl, M., Stitz, H., Streit, M., Potzmann, B., and Aigner, W. (2022c). Visualization Onboarding Grounded in Educational Theories. In Visualization Psychology. Springer Nature.
- Stoiber, C., Walchshofer, C., Grassinger, F., Stitz, H., Streit, M., and Aigner, W. (2021). Design and comparative evaluation of visualization onboarding methods. In *Proceedings of the 14th International Symposium on Visual Information Communication* and Interaction, VINCI '21, page 9, New York, NY, USA. Association for Computing Machinery.
- Stoiber, C., Walchshofer, C., Pohl, M., Potzmann, B., Grassinger, F., Stitz, H., Streit, M., and Aigner, W. (2022d). Comparative evaluations of visualization onboarding methods. *Visual Informatics*.
- Streit, M., Schulz, H.-J., Lex, A., Schmalstieg, D., and Schumann, H. (2011). Modeldriven design for the visual analysis of heterogeneous data. *IEEE Trans. Visualization* and Computer Graphics, 18(6):998–1010.
- Svelte (2023). https://svelte.dev/. Accessed: 2022-11-16.
- Swabish, J. and Ribecca, S. (2014). The graphic continuum: A poster project for your office. https://policyviz.com/2014/09/09/graphic-continuum/. Accessed: 2020-03-25.
- swoopyDrag.js (2020). http://lwheel.github.io/swoopy-drag/. Accessed: 2023-01-03.
- Tableau Software (2023). https://www.tableau.com/products/desktop. Accessed: 2023-01-18.
- Tanahashi, Y., Leaf, N., and Ma, K. (2016). A study on designing effective introductory materials for information visualization. *Computer Graphics Forum*, 35(7):117–126.
- Thomas, J. J. and Cook, K. A., editors (2005). Illuminating the Path: The Research and Development Agenda for Visual Analytics. IEEE.

TIBCO Jaspersoft (2019). https://www.jaspersoft.com/. Accessed: 2013-01-14.

- TIBCO Spotfire (2019). https://www.tibco.com/products/tibco-spotfire. Accessed: 2019-03-14.
- Tourguide.js (2023). https://github.com/LikaloLLC/tourguide.js. Accessed: 2023-02-01.
- tutoBox (2020). https://bocovo.github.io/tutobox-site/. Accessed: 2023-02-01.
- UNHCR (2020). UNHCR Refugee Statistics. https://www.unhcr.org/refugeestatistics/. Accessed: 2022-01-31.
- Uttal, D. H., O'Doherty, K., Newland, R., Hand, L. L., and DeLoache, J. (2009). Dual representation and the linking of concrete and symbolic representations. *Child Development Perspectives*, 3:156–159.
- van Wijk, J. J. (2005). The value of visualization. In Proc. IEEE Visualization (VIS 05), pages 79–86.
- vue (2023). https://vuejs.org/. Accessed: 2022-11-16.
- Wagemans, J., Feldman, J., Gepshtein, S., Kimchi, R., Pomerantz, J. R., van der Helm, P. A., and van Leeuwen, C. (2012). A century of Gestalt psychology in visual perception: II. Conceptual and theoretical foundations. *Psychological Bulletin*, 138:1218–1252. Place: US Publisher: American Psychological Association.
- Wagner, M. (2017). Integrating Explicit Knowledge in the Visual Analytics Process: Model and Case Studies on Time-oriented Data. PhD Thesis, Vienna University of Technology, Vienna.
- Wang, C. (2022). Visvisual: A toolkit for teaching and learning data visualization. IEEE Computer Graphics and Applications, 42:20–26.
- Wang, X., Jeong, D. H., Dou, W., Lee, S.-W., Ribarsky, W., and Chang, R. (2009). Defining and applying knowledge conversion processes to a visual analytics system. *Computers & Graphics*, 33(5):616–623.
- Wang, Z., Sundin, L., Murray-Rust, D., and Bach, B. (2020). Cheat sheets for data visualization techniques. In *Proceedings of the 2020 CHI Conference on Human Factors* in Computing Systems, CHI '20, page 1–13, New York, NY, USA. Association for Computing Machinery.

Watson, J. B. (1997). Behaviorism. Taylor & Francis, 1st edition.

webTour.js (2022). https://votch18.github.io/webtour.js/. Accessed: 2023-02-01.

- Wickham, H. (2010). A layered grammar of graphics. Journal of Computational and Graphical Statistics, 19(1):3–28.
- Wickham, H. (2018). ggvis 0.4 overview. http://ggvis.rstudio.com. Online; Last accessed 2019-01-25.
- Wickham, H., Chang, W., Henry, L., Lin Pederson, T., Takahashi, K., Wilke, C., Woo, K., Yutani, H., and Dunnington, D. (2022). ggplot2. Accessed: 2022-11-09.
- Wijk, J. J. v. (2006). Views on visualization. IEEE Transactions on Visualization and Computer Graphics, 12(4):1000–433.
- Wilkinson, L. (1999). The Grammar of Graphics. Springer.
- Willett, W., Heer, J., and Agrawala, M. (2007). Scented widgets: Improving navigation cues with embedded visualizations. *IEEE Transactions on Visualization and Computer Graphics*, 13(6):1129–1136.
- Williams, R. (2014). Non-Designer's Design Book. Peach Pit, San Francisco, California, 4 edition.
- Wohlin, C. (2014). Guidelines for snowballing in systematic literature studies and a replication in software engineering. In Proceedings of the 18th international conference on evaluation and assessment in software engineering, pages 1–10.
- Wongsuphasawat, K. (2020). Encodable: Configurable Grammar for Visualization Components. arXiv:2009.00722 [cs].
- Wood, D., Bruner, J. S., and Ross, G. (1976). The Role of Tutoring in Problem Solving\*. Journal of Child Psychology and Psychiatry, 17(2).
- Yalçın, M. A. (2016). A Systematic and Minimalist Approach to Lower Barriers in Visual Data Exploration. Dissertation, University of Maryland, College Park.
- Yalçın, M. A. and Xharra, B. (2022). Keshif Data Made Explorable. https://www. keshif.me/. Accessed: 2022-01-31.
- Yang, L., Xu, X., Lan, X., Liu, Z., Guo, S., Shi, Y., Qu, H., and Cao, N. (2022). A design space for applying the freytag's pyramid structure to data stories. *IEEE Transactions* on Visualization and Computer Graphics, 28(1):922–932.
- You, E. (2019). Vite. https://vitejs.dev/. Accessed: 2023-01-18.
- Zeid, A., Kamarthi, S., Duggan, C., and Chin, J. (2011). CAPSULE: An Innovative Capstone-Based Pedagogical Approach to Engage High School Students in STEM Learning. Proceedings of the ASME 2011 International Mechanical Engineering Congress and Exposition, 5:305–314.

- Zhang, Z., Wang, B., Ahmed, F., Ramakrishnan, I. V., Zhao, R., Viccellio, A., and Mueller, K. (2013). The five ws for information visualization with application to healthcare informatics. *IEEE Transactions on Visualization and Computer Graphics*, 19(11):1895–1910.
- Zhi, Q., Ottley, A., and Metoyer, R. A. (2019). Linking and layout: Exploring the integration of text and visualization in storytelling. *Comput. Graph. Forum*, 38:675–685.

## List of Figures

1.1	Two sides of the visualization process by (Cairo, 2013)	3
1.2	<b>Onboarding for parallel coordinates plot</b> (Kwon and Lee, 2016)	5
1.3	Learning-by-analogy: In-betweens of linear chart and spiral chart (Ruchika-	
	chorn and Mueller, $2015$ )	6
1.4	Guided Tour mode by (Chundury et al., 2023)	7
1.5	Research methodology approach for the doctoral research $\ldots$ .	10
2.1	$\mathbf{Motivation}$ for a descriptive design space for visualization onboarding $\ .$	18
2.2	Examples of <b>bottom-up and top-down exercises</b> for a scatterplot (Tana-	
	hashi et al., 2016) $\ldots$	19
2.3	Learning-by-analogy: In-betweens of linear chart and spiral chart (Ruchika-	
	chorn and Mueller, 2015) $\ldots$	21
2.4	A visual overview of the onboarding design space and of how all six	
	questions, "Why, What, Who, How, Where, and When" relate to one	
	another	25
2.5	Onboarding approaches: IBM Cognos, PowerBI, C'est la vis, and learning	
0.0	environment for treemaps	32
2.6	Onboarding approaches: Advizor, Learning-by-analogy, interactive tutorial	
	for parallel coordinates plot	33
3.1	Overview of three studies that were performed by highlighting the respective	
	research questions $(RQ)$	44
3.2	Step-by-step guide for a Parallel Coordinates Plot & Scrollytelling tutorial	47
3.3	In-situ onboarding for Netflower	49
3.4	Netflower onboarding consisting of multiple sections including text along an-	
	notated screenshots and a video with a verbal explanation of the functionality	50
3.5	Procedure of study 1 and study 2 with MTurk workers	53

$\begin{array}{c} 3.6\\ 3.7\end{array}$	Level of experience per visualization type for study 1 on a 5-point Likert scale Mouse movement heatmap on the step-by-step guide for a bar chart, horizon graph and matrix received from Hotiar (2023). In Appendix A we	54
38	attached the Hotjar heatmaps separately	59
3.0	coordinates plot received from Hotjar (2023)	60
3.9	Procedure for study 3 with students	65
3.10	Participants' attitude towards onboarding concepts	68
4.1	Scatterplot - <b>abstract explanation</b> (Ribecca, 2020)	78
4.2 4.3	Abstract onboarding instructions for the treemap showing Biden's tax overnaul	81
1.0	tax overhaul (Committee for Responsible Federal Budget, 2021; Probasco,	
	2021). The layout and structure are the same as in Figure 4.2.	82
4.4	Concrete onboarding instruction for the treemap showing Austrian	
	nance, 2022). The layout and structure are the same as in Figure 4.2.	83
4.5	Abstract onboarding instruction showing Austrian federal budget	
	of 2022 (Federal Ministry Republic of Austria Finance, 2022). The	~ (
16	layout and structure are the same as in Figure 4.2.	84
4.0	Boxplots showing the results of the four statements per condition	84 87
4.8	Bar charts showing results the analysis of insights	89
6.1	Visualization onboarding concept integrated into a scatterplot visualization.	106
7.1	Examples of user and visualization onboarding JavaScript libraries	118
(.2	(2022)	119
7.3	VisAhoi.js integrated into a bar chart, change matrix, scatterplot,	110
	and treemap visualization using Plotly.js	122
7.4	VisAhoi.js integrated into a horizon graph using Plotly.js	123
7.5	Workflow for visualization developers using VisAhoi	124
7.6	The VisAhoi workflow	125
8.1	Exemplary data story by Reuters	137
8.2	Flourish authoring tool components	138
8.3	Timeline of the design, development, validation, and reflection phases from full 2020 to November 2022	120
8.4	DataWrapper annotation example (Muth 2022)	139 141
8.5	Flourish annotation example (Bider, 2020)	142
8.6	Onboarding button – Floating action button	146
8.7	Authoring tool in Flourish	147

8.8	Flourish graphical user interface. Interface to add onboarding stages and onboarding messages (1) as well as to edit onboarding stages such as	1.40
	color, icon, title of the stage, and order $(2)$	148
8.9	Edit mode of an onboarding message	149
8.10	Overview of VisAhoi workflow	149
8.11	Procedure of the usability study with journalists in online one-on-one or group settings	153
9.1	Timeline of the development, validation, and reflection phases .	164
9.2	Visualization onboarding concept integrated into the VA tool for analyzing	
	HTS data	170
9.3	Overview of VisAhoi components	171
9.4	VA tool for analyzing HTS data — side-by-side a hexbin plot (left) and scatterplot (right) with integrated VisAhoi library to provide onboard-	170
0.5	Ing instructions for the domain experts	1(2
9.5	interacting with the chart onboarding stage with its in-place annotation	174
		174
11.1	The overview model of user assistance	202
11.2	Aspects of visualization onboarding and guidance	205
11.3	The guided tour mode	207
11.4	<b>External help website</b> including videos, screenshots, and instructional text to explain the features of Microsoft Power BI (2019)	208
11.5	<b>Step-by-step tour</b> in IBM Cognos Analytics (IBM Cognos Analytics, 2019) to onboard first-time users on demand using tooltips and overlays to explain	200
11.0	the general features of the tool	209
11.0	Visualization onboarding by analogy. To introduce a spiral chart, a	
	Mueller 2015)	200
117	Sticky notes (Kang et al. 2003) used to onboard users to the Vman	203
11.1	tool. This is an example of using visual cues to support users in understanding the main functions of the interface	911
11.0	Starting screen of the Drofit	211 919
11.0	Draft: Adapted dealter for performing the task of building a portfolio	212 914
11.9	Droft: Swing trading desktops configuration	214 914
11.10	Figure showing the comparison of two high level tools in the context.	214
11.1.	of building a portfolio (Fundamental Analysis (FA)) and performing swing trading (Technical Analysis task (TA)) in the <i>Profit</i> VA tool	215
11.12	2Basic anatomy of a candlestick chart	219
11.13	BDistinct trading scenarios with the same candlestick pattern but different strategies	220
$11.1_{-1}$	4Screenshots of Profit's first-time user onboarding	221

A.1	Mouse movement heatmap on the step-by-step-guide for bar chart received	
	from (Hotjar, 2023)	4
A.2	Mouse movement heatmap on the step-by-step-guide for horizon graph received	
	from (Hotjar, 2023)	5
A.3	Mouse movement heatmap on the step-by-step-guide for matrix received	
	from (Hotjar, 2023)	6
D 1	VisAhoi: Onboarding Message using eCharts (Li et al. 2018)	17
D 2	VisAhoi: Onboarding Message using Vega-lite (Satvanaravan et al. 2017)	18
D.2	visition. Onboarding message using vega nee (sauyanarayan ee al., 2017)	10
F.1	Onboarding Message Reading the chart $(1)$	94
F.2	Onboarding Message Reading the chart (2)	95
F.3	Onboarding Message Reading the chart $(3)$	96
F.4	Onboarding Message Reading the chart (4)	97
F.5	Onboarding Message Reading the chart $(5)$	98
F.6	Onboarding Message Reading the chart (6)	99
F.7	Onboarding Message Interacting with the chart $(1)$	100
F.8	Onboarding Message Interacting with the chart $(2)$	101
F.9	Onboarding Message Interacting with the chart $(3)$	102
F.10	Onboarding Message Reading the chart $(1)$	103
F.11	Onboarding Message Reading the chart $(2)$	104
F.12	Onboarding Message Reading the chart $(3)$	105
F.13	Onboarding Message Interacting with the chart $(1)$	106
F.14	Onboarding Message Interacting with the chart $(2)$	107
F.15	Onboarding Message Interacting with the chart $(3)$	108

## List of Tables

2.1	Overview table of available visualization on boarding approaches	30
3.1 3.2 3.3	Averages (and SD) of performance measures for all four visualization types of study 1, n=388	55 62 62
$4.1 \\ 4.2 \\ 4.3$	Mean and SD per condition and data set	86 90 93
6.1	Demographic information of experts	110
7.1	Overview table showing related work in terms of visualization onboarding libraries	120
8.1 8.2 8.3	Demographic information of journalists. Reported experience level with Flourish along a 5-point Likert scale	151 153 154
9.1	<b>Demographic information of interviewees</b> . Gender: $f = 3, m = 3$ ).	165
10.1	Overview of our design actions that are categorized along our guiding questions of the Design Space.	194



## Glossary

- guidance Guidance is a computer-assisted process that aims to actively resolve a knowledge gap encountered by users during an interactive VA session (Ceneda et al., 2017).. 194
- in-place annotation anchor Inspired by different annotation designs by Lu (2022), we use circular anchors with numbers related to the onboarding message. The onboarding messages are designed as tooltip in combination with the in-place annotations anchors to indicate the connection between the textual instructions and the visual encoding in the visualization.. 48, 106, 189
- knowledge-assisted visual analytics Knowledge-assisted visual analytics is the science of extracting the tacit knowledge of users to make it machine-readable as extracted knowledge to enrich and enhance the VA Process by user-shared knowledge integration (Federico et al., 2017; Wagner, 2017)... 201
- tooltip As a common UI pattern, tooltips (callouts) are simple text labels that offer brief information next to a UI component on demand (such as on mouse hover). However, they generally present static (non-contextual) descriptive information, and they are not indexed for navigation (Chundury et al., 2023). It consists of a headline and textual instruction. Important words are highlighted in color in the instructional text, which we found as an essential (Stoiber et al., 2022b) (cf. Section 5).. 106, 107, 267
- visualization onboarding Visualization onboarding is the process of supporting users in reading, interpreting, and extracting information from visual representations of data (Stoiber et al., 2022c, 2021, 2022d).. 194



Part VI Appendix



# Appendix A

## Material Comparative Evaluation on Onboarding Approaches

### A.1 Task Categories

A.1.1 Bar Chart

Retrieve Value: How many extreme temperature events have happened in 2014?

Find Extremum: In which year were the most droughts registered?

**Determine Range**: What is the worldwide minimum and maximum of recorded earthquakes from 1990 - 2018?

*Make Comparison*: How much did the number of recorded extreme temperature events increase from 2003 to 2012?

### A.1.2 Horizon Graph

Retrieve Value: How many steps were measured on 01/06/2019 from Madison?

Find Extremum: On which day did Anna take the highest number of steps?

**Determine Range**: What is the minimum and maximum number of recorded steps of Justin between 01/17/2019 and 01/23/2019?

*Make Comparison*: How many steps did Justin take less on 01/04/2019 than on 01/10/2019?

### A.1.3 Change Matrix

**Retrieve Value**: How big is the value change in the discipline "200m Breaststroke Men" in the USA?

*Find Extremum*: Which discipline had the highest change in the number of medals across all countries?

**Determine Range**: What is the minimum and maximum of value changes across all countries and disciplines?

*Make Comparison*: Which countries showed an increase in the number of won medals in the discipline "100m Freestyle Women"?

### A.2 Hotjar Heatmaps



Figure A.1: Mouse movement heatmap on the step-by-step-guide for bar chart received from (Hotjar, 2023)


Figure A.2: Mouse movement heatmap on the step-by-step-guide for horizon graph received from (Hotjar, 2023)

#### A. MATERIAL COMPARATIVE EVALUATION ON ONBOARDING APPROACHES



Figure A.3: Mouse movement heatmap on the step-by-step-guide for matrix received from (Hotjar, 2023)

# APPENDIX **B**

# Material Abstract vs. Concrete Onboarding Instructions

# Concrete A - Abstract B [SEVA] Onboarding Messages Treemap Survey

Dear students, I appreciate that you participate in this survey.

This is a survey to assess the usefulness and understandability of visualization onboarding messages for a treemap visualization. Visualization onboarding is defined as the process of supporting users in reading, interpreting, and extracting information from visual representations of data.

ž

:

First, fill out the questionnaire (below) starting with demographic data such as age, gender, experience with visualization etc. After, continue with the tasks 1-3 in the questionnaire.

Task 1: Open the prototype and perform the following task: Try to understand the treemap visualization using the textual instructions next to it! Please, take as much time as necessary until you feel you understand the tree map visualization.

Task 2: Write down what you see and how you would describe it to someone who has no idea how to read it (up to 5-8 sentences).

Task 3: Find out at least three facts about the shown data and describe them here

In total, the survey will take around 30min to 1 hour. We kindly ask you to conduct the survey on your desktop computer due to incompatibilities on mobile devices.

Matr	ikelnr. *
Kurza	antwort-Text
Plea	se indicate your gender in the box below. *
0	female
0	nale
0	Dther
0	Prefer not to say
Plea	se indicate your age in years in the box below. $$ *
Kurza	antwort-Text
Do y	ou ever suffer from color blindness (e.g. cannot distinguish red and green)? *
0	/es
$\overline{\mathbf{O}}$	Νο
If vo	u anoward the questions above with "VEC" places indicate the type of color deficiency
II yo	a answered the questions above with TES, please indicate the type of color deliciency.
KUIZ	
Plea	se indicate the experience with interpreting data visualizations. $^{\star}$
	1 2 3 4 5 6 7
no	experienced in interpreting

Please indicate the ex								
Birds	;		Man	Pet Shop nmals		Fish		Reptiles
Zebra Finches	Canaries	D	ogs		Gold	fish Coy	Carp	Geckos
	Cockatiels				Ange	lfish Catfi	sh	
Budgerigars								Snakes Chameleon
	Lovebirds	Ci	its	Rabbits Mi	Cichli	d Molly	Tetra	
				Rats <sup>Guinea</sup> <sup>IN</sup>	Gappy	Losch Betta	Glofish	Bearded Dragon Manifer
	1	2	3	4	5	6 7	,	
not experienced	0	$\bigcirc$	0	0	0	0 0	)	highly experience
h Abschnitt 1 Weiter z	um nächs	ten Absc	hnitt			*		
oschnitt 2 von 3								
Feedback to Prototyp	e 1							×
Task 1: Open the link to	the proto	type and	perform t	he followi	ng task: 1	Try to unde	erstand	the treemap
visualization using the t understand the tree ma	extual ins p visualiza	tructions ation.	s next to it	Please, t	ake as m	uch time a	s nece	ssary until you feel
Prototype 1: <u>https://chu</u>	isi.netlify.a	pp/ (CA	)					
Task 2 <sup>.</sup> Write down wha	t vou see	and how	you woul	d describe	it to son	neone who	has n	o idea how to read
rask 2. Write down write	it you see		you woul		11 10 301	leone who	1103 11	
Task 3: Find out at least Task 2: Please write ( no idea how to read it .angantwort-Text	t three fac down wha (up to 5-	ts about at you so 8 senter	the shown ee and ho nces). (Pr	n data and	d describe	e them her	e. some	one who has *
Task 3: Find out at least Task 2: Please write on o idea how to read it angantwort-Text Task 3: Please find ou (Prototype 1)	t three fac down what (up to 5- it at least	ts about at you s 8 senter three fa	the shown ee and ho nces). (Pr	n data and ww you we rototype 1 t the sho	d describe ould des 1) wn data	e them her	e. some	one who has *
Task 3: Find out at lease Task 2: Please write on o idea how to read it angantwort-Text Task 3: Please find ou (Prototype 1) angantwort-Text	t three fac down whi (up to 5-	ts about at you s 8 senter three fa	ee and ho nces). (Pr	n data and ow you we tototype 1 it the sho	d describe ould des I) wn data	e them her	e. some	one who has *
Task 3: Find out at lease Task 2: Please write on idea how to read it angantwort-Text Task 3: Please find ou (Prototype 1) angantwort-Text	t three fac down what (up to 5-	ts about at you s 8 senter	ee and honces). (Pr	n data and ow you we ototype 1	ould des 1) wn data	e them her	e. some	one who has *
Task 3: Find out at lease Task 2: Please write on the idea how to read it angantwort-Text Task 3: Please find ou (Prototype 1) angantwort-Text The textual instruction	t three fac down wha (up to 5- it at least	at you so 8 senter three fa	ee and ho nces). (Pr acts abou	n data and ow you we ototype 1 t the sho d. *	ould des 1) wn data	e them her	e. some	one who has *
Task 3: Find out at lease Task 2: Please write on the idea how to read it angantwort-Text Task 3: Please find ou (Prototype 1) angantwort-Text The textual instruction	t three fac down whi (up to 5- ut at least ns were e	at you so 8 senter three fa easy to u	ee and ho nces). (Pr acts abou	w you we ototype 1 t the sho d. * 4	ould des 1) wn data	e them her	e. some iribe th	one who has *
Task 3: Find out at lease Task 2: Please write on the idea how to read it angantwort-Text Task 3: Please find out (Prototype 1) angantwort-Text The textual instruction Strongly Disagree	t three fac down what (up to 5- it at least ns were e	at you so 8 senter three fa	ee and honces). (Pr acts about understant 3	w you we ototype 1 t the sho d. * 4	ould des 1) wn data	e them her cribe it to and desc 6	e. some ribe th	one who has *
Task 3: Find out at lease Task 2: Please write ( no idea how to read it .angantwort-Text Task 3: Please find ou (Prototype 1) .angantwort-Text The textual instruction Strongly Disagree The textual instruction	t three fac down wh: (up to 5- it at least ns were e 1 0	at you si 8 senter three fa easy to u 2	ee and honces). (Pr acts about understan 3 	n data and ow you we ototype 1 t the sho d. * 4	ould des 1) wn data	e them her	e. some	one who has *
Task 3: Find out at lease Task 2: Please write ( no idea how to read it angantwort-Text Task 3: Please find ou (Prototype 1) angantwort-Text The textual instruction Strongly Disagree The textual instruction	t three fac down wh: (up to 5- it at least ns were e 1 0	ts about at you si 8 senter three fa easy to L 2 0	ee and honces). (Pr acts about understan 3 	t the sho	ould des 1) wn data	e them her	e. some	one who has *
Task 3: Find out at lease Task 2: Please write on the idea how to read it angantwort-Text Task 3: Please find ou (Prototype 1) angantwort-Text The textual instruction Strongly Disagree The textual instruction	t three fac down what (up to 5- it at least ns were e 1 0 ns were s 1	at you so 8 senter three fa easy to u 2 0 thort end 2	ee and honces). (Pr acts about understant 3 	t the sho	ould des 1) wn data 5 5	e them her	e. some	one who has *
Task 3: Find out at lease Task 2: Please write on the idea how to read it angantwort-Text Task 3: Please find ou (Prototype 1) angantwort-Text The textual instruction Strongly Disagree The textual instruction Strongly Disagree	t three fac down what (up to 5- it at least it at least 1 0 ns were s 1 0	at you so 8 senter three fa easy to u c hort end 2 0	ee and honces). (Pr acts about understan 3 Ough. * 3 O	an data and ow you we ototype 1 t the sho d. * 4 0	ould des 1) wn data	e them her	some	one who has *
Task 3: Find out at lease Task 2: Please write on the idea how to read it angantwort-Text Task 3: Please find ou (Prototype 1) angantwort-Text The textual instruction Strongly Disagree The textual instruction Strongly Disagree The textual instruction	t three fac down what (up to 5- it at least ins were e 1 0 ns were s 1 0	ts about at you si 8 senter three fa easy to u 2 0 whort end 2 0 0	ee and honces). (Pr acts about understan 3 ough. * 3 ounderst	and the t	ould des ould des ) wn data 5 5 5 0 reemap	e them her	e. some	one who has *
Task 3: Find out at lease Task 2: Please write on the idea how to read it angantwort-Text Task 3: Please find ou (Prototype 1) angantwort-Text The textual instruction Strongly Disagree The textual instruction Strongly Disagree The textual instruction	t three fac down what (up to 5- it at least ins were e 1 0 ns were s 1 0 1	ts about at you so 8 senter three fa easy to L 2 0 thort end 2 0 elepful to 2	ee and honces). (Pr acts abou understan 3 Ough. * 3 Ough. * 3 Ough. *	w you we ototype 1 t the sho d. * 4 0 and the to 4	ould des ould des ) wn data 5 5 0 reemap ' 5	e them her	e. some ribe th 7 0 7 0 7 0	one who has *
Task 3: Find out at lease Task 2: Please write on the idea how to read it angantwort-Text Task 3: Please find out (Prototype 1) angantwort-Text The textual instruction Strongly Disagree The textual instruction Strongly Disagree The textual instruction	t three fac down whi (up to 5- it at least ns were e 1 0 ns were s 1 0 1	ts about at you si 8 senter three fa easy to u 2 0 thort end 2 0 1 thort end 2 0 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	ee and honces). (Pr acts about inderstan 3 ough. * 3 o underst 3	and the tr	s ould des 1) wn data 5 0 5 0 7 reemap v 5	e them her	e. some rribe th 7 0 7 0	one who has *
Task 3: Find out at lease Task 2: Please write on the idea how to read it angantwort-Text Task 3: Please find out (Prototype 1) angantwort-Text The textual instruction Strongly Disagree The textual instruction Strongly Disagree The textual instruction Strongly Disagree	t three fac down what (up to 5- it at least ns were e 1 0 ins were s 1 0 1 0 1	ts about at you si 8 senter three fa easy to u 2 0 thort end 2 0 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	ee and honces). (Pr acts about acts acts about acts acts about acts acts about acts acts about acts	and data and ow you we ototype 1 t the sho d. * 4 0 4 0 4 0 4 0 4 0 4 0	ould des ould des ) wn data 5 5 0 reemap 7 5 0 1 1 1 1 1 1 1 1 1 1 1 1 1	e them her	e. some ribe th 7 0 7 0	one who has *
Task 3: Find out at lease Task 2: Please write on the idea how to read it angantwort-Text Task 3: Please find ou (Prototype 1) angantwort-Text The textual instruction Strongly Disagree The textual instruction Strongly Disagree The textual instruction Strongly Disagree All relevant aspects to	three fac down whi (up to 5- it at least ns were e 1 0 ns were s 1 0 1 0 1 0 1 0 1 0 1 0 1 0 1 0 1 0 1	at you so 8 senter three fa easy to u c c c c c c c c c c c c c c c c c c c	ee and honces). (Print acts about inderstanding acts about inderstandin	and the sho d. * 4 0 and the tr 4 0 awere des	scribed. *	e them her	e. some ribe th 7 0 7 0	one who has * em here. * strongly agre
Task 3: Find out at lease Task 2: Please write on the idea how to read it angantwort-Text Task 3: Please find out (Prototype 1) angantwort-Text The textual instruction Strongly Disagree The textual instruction Strongly Disagree The textual instruction Strongly Disagree All relevant aspects to	t three fac down whi (up to 5- it at least ns were e 1 0 ns were s 1 0 1 0 1 0 1 0 1 0 1 0	ts about at you si 8 senter three fa easy to u 2 0 thort end 2 0 elpful to 2 0 and the 2	ee and honces). (Pr acts about inderstan 3 ough. * 3 o underst 3 o underst 3 o underst 3 o underst 3 o underst 3 o underst	and data and ow you we ototype 1 t the sho d. * 4 0 4 0 4 0 4 0 4 0 4 0 4 0 4 0 4 0 4	ould des ) wn data 5 5 0 5 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7	e them her	e. some ribe th 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7	one who has * uem here. * strongly agre
Task 2: Please write on idea how to read it angantwort-Text Task 3: Please find ou (Prototype 1) angantwort-Text The textual instruction Strongly Disagree The textual instruction Strongly Disagree The textual instruction Strongly Disagree The textual instruction	three fac down what (up to 5- it at least ins were e 1 0 ns were s 1 0 1 0 1 0 1 0 1 0 1 0 1 0 1 0 1 0 1	ts about at you si 8 senter three fa easy to L 2 0 thort end 2 0 elpful to 2 0 and the 2	ee and honces). (Pr acts about acts acts about acts	and the tr d. * 4 0 and the tr 4 0 and the tr 4 0 and the tr 4 0 and the tr	ould des ) wn data 5 5 5 7 7 8 7 8 7 8 7 8 7 7 7 7 7 7 7 7	e them her	e. some ribe th 7 0 7 0 1 0 7 0 1 0 7 0 7	one who has *

		1 2	3	4	5	6	7		
Not knowledgeable abo	out (	50	0	0	0	0	0	Very kno	wledgeable ab
What should we improve	e? What	is miss	ing? *						
Langantwort-Text									
If you have any further c	commen	nts. suad	restions	or rem	arks re	egardi	na the	onboardin	a *
messages, please write	them in	the box	below.			- <b>j</b>			5
Langantwort-Text									
h Abschnitt 2 Weiter zum	n nächste	en Absch	nitt				-		
bschnitt 3 von 3									
Feedback to Prototype 2	2								ž
Task 1: Open the link to the	e prototy	/pe and p	perform t	he follo	wing ta	sk: Try	to und	erstand the	treemap
visualization using the text understand the tree map v	tual instr /isualizat	ructions tion.	next to it	! Please	, take a	is muc	h time	as necessa	ry until you fee
Prototype 2: <u>https://vasuki</u>	<u>i.netlify.</u> a	app/ (Ab)	)						
Task 2: Write down what y	ou see a	ind how	you woul	d descri	be it to	some	one wh	o has no id	ea how to read
Task 3: Find out at least th	nree facts	s about t	he show	n data a	nd des	cribe t	hem he	ere.	
Langantwort-Text	at least 1	three for	cts abo	It the of		lata a	nd dee	cribe them	here <sup>1</sup>
(Prototype 2)	at redăt l	ance Id	us auul	a ane Sf	.owii C	aud d	na ues	Since mem	HEIC.
Langantwort-Text									
Langantwort-Text	(next to	o the tree	emap vis	sualizat	ion) w	ere ea	asy to u	understand	.*
Langantwort-Text	(next to	o the tree 2	emap vis 3	sualizat 4	ion) w	ere ea	asy to u 6	understand 7	*
Langantwort-Text The textual instructions Strongly Disagree	(next to 1	o the tree 2	emap vis 3	sualizat 4	ion) w	ere ea	asy to t 6	understand 7	. * strongly ag
Langantwort-Text The textual instructions Strongly Disagree The textual instructions	(next to 1 (next to	the tree	emap vis 3 O	sualizat 4 O sualizat	ion) w	ere ea	asy to t 6 O	understand 7 Ough. *	. * strongly ag
Langantwort-Text The textual instructions Strongly Disagree The textual instructions	(next to 1 (next to 1	o the tree 2 O the tree 2	emap vis 3 O emap vis 3	sualizat 4 O sualizat	ion) w	ere ea 5 ) ere sh	asy to a 6 O nort en 6	understand 7 Ough. * 7	.* strongly ag
Langantwort-Text The textual instructions Strongly Disagree The textual instructions Strongly Disagree	(next to 1 (next to 1 0	o the tree 2 o the tree 2 o the tree 2	emap vis 3 O emap vis 3 O	sualizat 4 O sualizat 4	ion) w ! ( ion) w !	ere ea 5 ) ere sh	6 Onort en 6	understand 7 Ough. * 7 O	. * strongly ag
Langantwort-Text The textual instructions Strongly Disagree The textual instructions Strongly Disagree The textual instructions treemap visualization	(next to 1 (next to 1 (next to (next to	o the tree 2 0 the tree 2 0 0 the tree	emap vis 3 emap vis 3 O	4 	ion) w ! ( ion) w ! ( ion) w	ere ea 5 ) ere sh 5 ) vere h	asy to t 6 0 nort en 6 0	understand 7 ough. * 7 0	strongly ag
Langantwort-Text The textual instructions Strongly Disagree The textual instructions Strongly Disagree The textual instructions treemap visualization	(next to 1 (next to 1 (next to 1 1 1	o the tree 2 0 the tree 2 0 the tree 2 0 the tree 2	emap vis 3 0 emap vis 0 emap vis 3 0	4 4 5 5 5 5 5 5 5 5 5 5 5 5 5	ion) w ( ion) w ( ion) w	ere ea 5 ) ere sh 5 ) vere h	asy to t 6 0 nort en 6 0 elpful 1 6	understand 7 Ough. * 7 Ough. *	strongly ag
Langantwort-Text The textual instructions Strongly Disagree The textual instructions Strongly Disagree The textual instructions treemap visualization Strongly Disagree	(next to 1 (next to 1 (next to 1 (next to 1 (next to	o the tree 2 0 the tree 2 0 0 the tree 2 0	emap vis 3 • • • • • • • • • •	sualizat 4 0 sualizat 4 0 sualizat 4	ion) w ( ( ( ( ( ( (	ere ea 5 	asy to the formation of	understand 7 Ough. * 7 Oundersta 7 O	strongly ag
Langantwort-Text The textual instructions Strongly Disagree The textual instructions Strongly Disagree The textual instructions treemap visualization Strongly Disagree All relevant aspects to u	(next to 1 (next to 1 (next to 1 (next to 1 (next to 1 (next to	o the tree 2 0 the tree 2 0 0 the tree 2 0 0 0 the tree	emap vis 3 () emap vis 3 () emap vis 3 () reemap	sualizat 4 0 sualizat 4 0 sualizat	ion) w ( ion) w ( ion) w ( c escribe	ere ea 5 ) ere sh 5 ) vere h 5 5 )	asy to t 6 0 nort en 6 0 elpful 1 6	understand 7 ough. * 7 0 xo understa 7 0	strongly age
Langantwort-Text The textual instructions Strongly Disagree The textual instructions Strongly Disagree The textual instructions treemap visualization Strongly Disagree All relevant aspects to u	(next to 1 (next to 1 (next to 1 1 understa 1	o the tree 2 0 the tree 2 0 the tree 2 0 1 the tree 2 1 the tree 2 1 the tree 2 1 the tree 1 the tree 1 the tree 2 1 the tree 1 the tree 2 1 the tree 1 the tree 2 1 the tree 1	emap vis 3 	sualizat 4 3 sualizat 4 0 sualizat 4 0 vwere du 4	ion) w ( ion) w ( ion) w ( c escribe	ere ea 5 	asy to t 6 0 nort en 6 0 elepful 1 6 0	understand 7 ough. * 7 0 xo understa 7 0	. * strongly ag and the *



# Material Study Design Cognitive Walkthrough Inspection

[Aufzeichnung starten] - wenn Sie anonym sein wollen, dann bitte Name in Zoom direkt ändern und die Kamera ausschalten

# Welcome (5 min)

Mein Name ist Christina Stoiber und ich werde Sie durch diese Evaluierung begleiten. Ziel dieser Studie ist es, unsere entwickelten Onboarding Konzepte für HIDDEN auf die Usability zu evaluieren. Ihr Experten-Feedback soll dazu beitragen, die Onboarding-konzepte zu verbessern. Visualization Onboarding definieren wir in unserem Projekt als den Prozess Benutzer\*innen beim Lesen, Interpretieren und Extrahieren von Informationen aus visuellen Darstellungen von Daten zu unterstützen.

Die Studie wird folgendermaßen ablaufen: Zuerst werde ich Sie nach ihren demographischen Daten fragen und ihnen dann einen Link zur Verfügung stellen. Ich werde Sie bitten den Klick-dummy vorerst selbst ein bisschen durchzuschauen um sie vertraut zu machen. Dann stelle ich ihnen 6 Aufgaben. Ich bitte Sie während der Evaluierung ihre Gedanken laut auszusprechen. Nach dem Durcharbeiten der Aufgaben werde ich ihnen noch ein paar Fragen stellen.

Wir werden die Evaluierung aufzeichnen sowie handschriftliche Notizen anfertigen. Die Studie wird ca 35 min dauern.

# Demografische Daten (3 min)

Alter: derzeitige Position: höchste abgeschlossene Ausbildung:

### Erfahrung mit Visualization Tools:

sehr wenig Erfahrung	wenig Erfahrung	Durchschnittliche Erfahrung	Erfahren	Sehr erfahren

Erfahrung mit Scatterplot Visualisierungen:

sehr wenig Erfahrung	wenig Erfahrung	Durchschnittliche Erfahrung	Erfahren	Sehr erfahren

# Walkthrough (3 min)

Bitte geben Sie jetzt den Bildschirm frei. Der Button befindet sich unten in der Leiste. Ich schicke Ihnen nun den Link zum Klick-Dummy im Chat. https://sketch.cloud/s/5a130ac2-bf59-4fe8-a467-228c0f614091

https://www.sketch.com/s/5a130ac2-bf59-4fe8-a467-228c0f614091/a/dl1rg33/play

Öffnen Sie nun den diesen Link und klicken Sie auf die linke obere Schaltfläche um den Prototypen zu öffnen.

Machen Sie sich bitte jetzt mit dem Interface vertraut. Klicken Sie gerne herum und probieren sie die Schaltflächen aus. Ich möchte Sie darauf hinweisen, dass es sich um einen Klick-Dummy handelt und noch nicht um das finale Umsetzung.

Was ist Ihr erster Eindruck:

# Aufgaben (10 min)

Bitte stellen Sie sich folgendes Szenario vor:

Sie sind Data Scientist\*in und arbeiten mit dem HIDDEN-Tool. Eine neue Visualisierung wurde integriert und ist nun in der neuesten Version verfügbar. Sie benötigen Hilfe bei der Visualisierung.

Ich werde die einzelnen Fragen in den Chat von Zoom kopieren, damit Sie sie besser verfolgen können.

## Aufgabe 1 - Hilfe (Onboarding) öffnen

Bitte öffnen Sie das Onboarding für die Scatterplot-Visualisierung (linke Bildschirmseite).

- War Hilfestellung notwendig um die Aufgabe zu lösen?
- Traten falsche Annahmen auf?
- Unmutsäußerungen der Testperson

Notizen:

## Aufgabe 2 - Öffnen des Reading-Parts

Öffnen Sie nun den "Reading" Part.

- War Hilfestellung notwendig um die Aufgabe zu lösen?
- Traten falsche Annahmen auf?
- Unmutsäußerungen der Testperson

Notizen:

#### Aufgabe 3 - Reading Part lesen

Lesen Sie sich nun alle Punkte im "Reading" Bereich durch.

- War Hilfestellung notwendig um die Aufgabe zu lösen?
- Traten falsche Annahmen auf?
- Unmutsäußerungen der Testperson
- Wie wurden die Tooltips aufgerufen?
  - Klick auf Visual Markers
  - Über das Navigationspanel (Pagination)

Notizen:

### Aufgabe 4 - Interaction Part lesen

Lesen Sie sich nun alle Punkte im "Interacting" Bereich durch.

- War Hilfestellung notwendig um die Aufgabe zu lösen?
- Traten falsche Annahmen auf?
- Unmutsäußerungen der Testperson
- Wie wurde zur nächsten Section gewechselt?
- Wie wurden die Tooltips aufgerufen?
  - Klick auf Visual Markers
  - Über das Navigationspanel (Pagination)

Notizen:

## Aufgabe 5 - Analyzing Part lesen

Lesen Sie sich nun alle Punkte im "Analyzing" Bereich durch.

- War Hilfestellung notwendig um die Aufgabe zu lösen?
- Traten falsche Annahmen auf?
- Wie wurde zur nächsten Section gewechselt?
- Unmutsäußerungen der Testperson
- Wie wurden die Tooltips aufgerufen?
  - Klick auf Visual Markers
  - Über das Navigationspanel (Pagination)

### Notizen:

## Aufgabe 6 - Schließen vom Onboarding Mode

Bitte schließen Sie das Onboarding nun.

- War Hilfestellung notwendig um die Aufgabe zu lösen?
- Traten falsche Annahmen auf?
- Unmutsäußerungen der Testperson

### Notizen:

Vielen Dank für das Durcharbeiten der Aufgaben. Ich habe jetzt noch ein paar Fragen zu Nutzung.

# Interview (10 min)

- Wie haben Sie die Navigation der Hilfestellung empfunden? (innerhalb der Sections bzw. Öffnen & Schließen des Hilfe-Modus)
- Was sagen Sie zu den Onboarding-Texten (Texte in den Tooltips)?
  - Sind sie verständlich?
  - kurz genug?
  - Würden Sie etwas ergänzen / fehlt etwas?
- Was hat gut funktioniert?
- Was könnten wir noch verbessern?
- Was fehlt?
- Gibt es sonst noch Feedback oder Anmerkungen?

Vielen Dank. Wir sind jetzt fertig.

[STOP RECORDING]



# APPENDIX D

# Material VisAhoi Library

## D.1 Screenshots of the Onboarding Concept using eCharts and Vega-lite



Figure D.1: VisAhoi: Onboarding Message using eCharts (Li et al., 2018)

#### D. MATERIAL VISAHOI LIBRARY





Figure D.2: VisAhoi: Onboarding Message using Vega-lite (Satyanarayan et al., 2017)

18



# Material User Experience Evaluation of Design Study DDJ

# Evaluation Datenjournalist\*innen Herbst 2022 - der Standard

# Interview Guide

Vielen Dank für Ihre Zeit. Ich möchte mich kurz vorstellen und den Ablauf erklären.

Wir haben eine Javascript Library entwickelt die semi-automatisiert Onboarding Texte und visuelle Marker für Plotly.js, echarts und vega-lite generiert. In dem Flourish Template für eine Treemap Visualisierung haben wir diese Library integriert und bieten verschiedene Features an um die Onboarding Texte bzw. visuelle Marker zu adaptieren.

Sie können die Navigation adaptieren, sowie die Texte und visuelle Marker hinzufügen und wieder löschen. Die Änderungen können in dem Flourish Template auch gepeichert werden in der aktuellen Session.

Ziel der Studie ist es das Flourish Template für eine Treemap Visualisierung auf die Benutzbarkeit zu überprüfen sowie Feedback zum Onboarding Konzept zu erhalten aus der Domäne.

Die Studie wird folgendermaßen ablaufen: Ich bitte Sie den Fragebogen durchzuarbieten und zum Schluss werden ich Ihnen noch ein paar Fragen stellen.

Ich werde die Evaluierung/Interview aufzeichnen sowie handschriftliche Notizen anfertigen.

Backup-Links: <u>https://treemaps.netlify.app/</u> -> all Onboarding concepts <u>https://flock-1571.students.fhstp.ac.at/Reuters/</u>

Link zum Flourish Template: https://app.flourish.studio/@cubicmagnet45/visahoi\_treemap\_fhstp\_beta/0.0.8 Fragebogen (10 min)

https://forms.gle/PNCoqJ8ksZVfaUDv9

# Abschließende Fragen & Feedback:

[entweder one-to-one oder als Fokusgruppen-Diskussion]

- 1. Wie haben Sie die Benutzung empfunden?
- 2. Wie bewerten Sie die Einbettung von der Onboarding Library in ein Flourish Template?
- 3. Welche Funktionen fehlen?
- 4. Wie haben Sie das generelle Onboarding Konzept empfunden? (die Navigation, Texte, Sections der Hilfestellung)
- 5. Was ist ihre Einschätzung zu den Onboarding Konzept (Zielgruppe LeserInnen)?
- 6. Was hat Ihnen gut gefallen?
- 7. Was hat weniger gut funktioniert? Was fehlt bei der Library/Design?
- 8. Gibt es sonst noch Feedback oder Anmerkungen?

# Evaluation Datenjournalist\*innen Herbst 2022 (One-to-one) Katrin Nussmayr

# Interview Guide

Vielen Dank für Ihre Zeit. Ich möchte mich kurz vorstellen und den Ablauf erklären.

Ziel der Studie ist es das Flourish Template für eine Treemap Visualisierung auf die Benutzbarkeit zu überprüfen sowie Feedback zum Onboarding Konzept zu erhalten aus der Domäne.

Die Studie wird folgendermaßen ablaufen: Zuerst werde ich Sie nach ihren demographischen Daten fragen und Ihnen kurz unser Forschungsprojekt sowie den derzeitigen Stand präsentieren. Als nächstes schicke ich Ihnen dann einen Link zum Flourish Template mit dem Onboarding. Ich werde Sie bitten sich vorerst selbst ein bisschen durchzuschauen um sich vertraut zu machen. Dann stelle ich ihnen 10 kurze Aufgaben. Ich bitte Sie während der Benutzung ihre Gedanken laut auszusprechen. Nach dem Durcharbeiten der Aufgaben werde ich ihnen noch ein paar Fragen stellen.

Ich werde die Evaluierung aufzeichnen sowie handschriftliche Notizen anfertigen. Die Studie wird ca 30min dauern.

### Aufnahme starten

# Demographische Daten

- Alter:
- Geschlecht:
- Welchen Ausbildung haben Sie?
- Was sind Ihre derzeitigen Aufgabe?

Erfahrung mit Flourish

sehr wenig wenig Erfahrung Erfahrung	Durchschnittliche Erfahrung	Erfahren	Sehr erfahren
---	--------------------------------	----------	---------------

Anderen Visualisierungstools (e.g., Datawrapper, Tableau)

sehr wenig Erfahrung	wenig Erfahrung	Durchschnittliche Erfahrung	Erfahren	Sehr erfahren
-------------------------	--------------------	--------------------------------	----------	---------------

- Welche Visualisierungstools verwenden Sie sonst?
- Wie wird Visualization Onboarding (Lesehilfe für Leser\*innen) bis jetzt umgesetzt?

# **TRY-OUT Flourish Template**

## Short intro to SEVA & Onboarding (3min)

Wir haben eine Javascript Library entwickelt die semi-automatisiert Onboarding Texte und visuelle Marker für Plotly.js, echarts und vega-lite generiert. In dem Flourish Template für eine Treemap Visualisierung haben wir diese Library integriert und bieten verschiedene Features an um die Onboarding Texte bzw. visuelle Marker zu adaptieren.

Sie können die Navigation adaptieren, sowie die Texte und visuelle Marker hinzufügen und wieder löschen. Die Änderungen können in dem Flourish Template auch gepeichert werden in der aktuellen Session.

Backup-Links: <u>https://treemaps.netlify.app/</u> -> all Onboarding concepts <u>https://flock-1571.students.fhstp.ac.at/Reuters/</u>

Walkthrough (5 min)

Link zum Flourish Template: https://app.flourish.studio/@cubicmagnet45/visahoi\_treemap\_fhstp\_beta/0.0.8

Bitte geben Sie jetzt den Bildschirm frei.

Öffnen Sie nun den diesen Link. Machen Sie sich bitte jetzt mit dem Interface vertraut.

Was ist Ihr erster Eindruck zum Flourish Template?

# Aufgaben (10 min)

## Aufgabe 0 - Öffnen vom Flourish Template

Öffnen Sie folgenden Link und "Erstellen sie eine Visualisierung". Link zum Template: https://app.flourish.studio/@cubicmagnet45/visahoi treemap fhstp beta/0.0.1

- War Hilfestellung notwendig um die Aufgabe zu lösen?
- Traten falsche Annahmen auf?
- Unmutsäußerungen der Testperson

## Aufgabe 1 - Overlay zum Speicherprozess durchlesen

Lesen Sie nun die Information zum Speichern durch. Wenn Sie es verstanden haben, dann machen Sie weiter mit Aufgabe 2.

- War Hilfestellung notwendig um die Aufgabe zu lösen?
- Traten falsche Annahmen auf?
- Unmutsäußerungen der Testperson

## Aufgabe 2 - Onboarding hinzufügen zur Treemap Visualisierung

Fügen Sie die Lesehilfe (Onboarding) zur Treemap Visualisierung hinzu.

- War Hilfestellung notwendig um die Aufgabe zu lösen?
- Traten falsche Annahmen auf?
- Unmutsäußerungen der Testperson

## Aufgabe 3 - Onboarding Button verschieben

Adaptieren Sie die Position vom Onboarding Button auf X-Axis Offset: 50 px.

- War Hilfestellung notwendig um die Aufgabe zu lösen?
- Traten falsche Annahmen auf?
- Unmutsäußerungen der Testperson

Aufgabe 4 - Onboarding Interface Elemente durchklicken Öffnen Sie nun das Onboarding.

Welche Onboarding Stages können Sie sehen?

Lesen Sie jetzt die Onboarding Messages durch in der "Reading" Section.

- War Hilfestellung notwendig um die Aufgabe zu lösen?
- Traten falsche Annahmen auf?
- Unmutsäußerungen der Testperson

## Aufgabe 5 - Hinzufügen von Stage

Sie wollen nun das die Onboarding adaptieren. Fügen Sie einen Stage hinzu mit dem Namen "Remember". Verwenden Sie dazu eine beliebige Farbe und folgendes Icon: "fa fa-book".

- War Hilfestellung notwendig um die Aufgabe zu lösen?
- Traten falsche Annahmen auf?
- Unmutsäußerungen der Testperson

## Aufgabe 6 - Hinzufügen von Onboarding Text zur gerade erstellten Stage

Fügen Sie nun folgende Onboarding Message zum neu erzeugten Stage hinzu. Verwenden Sie dazu die unten angeführten Informationen.

Titel: Remember key facts

Message: This is a dummy onboarding message explaining how to remember key facts.

Postion: Verwenden Sie den Picker um beliebige Koordinaten auszuwählen.

- War Hilfestellung notwendig um die Aufgabe zu lösen?
- Traten falsche Annahmen auf?
- Unmutsäußerungen der Testperson

## Aufgabe 7 - Adaptieren von Onboarding Text

Ändern sie nun den Onboarding Text in der Section "Reading" Nummer 3. Fügen Sie das Wort "Adaptions" am Ende hinzu.

- War Hilfestellung notwendig um die Aufgabe zu lösen?
- Traten falsche Annahmen auf?
- Unmutsäußerungen der Testperson

Aufgabe 8 - Löschen von Onboarding Text

Löschen Sie nun die Onboarding Message des visuellen Marker 4 in der Section "Reading".

- War Hilfestellung notwendig um die Aufgabe zu lösen?
- Traten falsche Annahmen auf?
- Unmutsäußerungen der Testperson

Aufgabe 9 - Löschen von Onboarding Stage Interacting

Löschen Sie nun den Onboarding Stage "Interacting".

- War Hilfestellung notwendig um die Aufgabe zu lösen?
- Traten falsche Annahmen auf?
- Unmutsäußerungen der Testperson

### Aufgabe 10 - Speichern

Speichern Sie nun die Änderungen im Template.

- War Hilfestellung notwendig um die Aufgabe zu lösen?
- Traten falsche Annahmen auf?
- Unmutsäußerungen der Testperson

# Abschließende Fragen & Feedback:

[entweder one-to-one oder als Fokusgruppen-Diskussion]

- 1. Wie haben Sie die Benutzung empfunden?
- 2. Wie bewerten Sie die Einbettung von der Onboarding Library in ein Flourish Template?
- 3. Wie haben Sie das generelle Onboarding Konzept empfunden? (die Navigation, Texte, Sections der Hilfestellung)
- 4. Was ist ihre Einschätzung zu den Onboarding Konzept (Zielgruppe LeserInnen)?
- 5. Was hat Ihnen gut gefallen?
- 6. Was hat weniger gut funktioniert? Was fehlt bei der Library/Design?

Elouriah Int	onio	,	0/00			~
Flourish int	erview	/ 2022	2/08			^
Wir haben eine Javascrip Plotly.js, echarts und veg diese Library integriert ur adaptieren.	t Library entv a-lite generie 1d bieten vers	wickelt die se rt. In dem Fl schiedene Fe	emi-automat ourish Temp eatures an ur	isiert Onboa late für eine m die Onboa	rding Texte u Treemap Vis rding Texte b	ind visuelle Marker f ualisierung haben v izw. visuelle Marker
Ziel der Studie ist es das überprüfen sowie Feedba	Flourish Tem ack zum Onbo	nplate für ein oarding Kon:	e Treemap V zept zu erhal	/isualisierun Iten aus der	g auf die Ben Domäne.	utzbarkeit zu
Bitte geben Sie in dem	unten stehe	enden Feld I	hr Alter in J	lahren an.	*	
Kurzantwort-Text						
Bitte geben Sie in dem	unten stehe	enden Feld I	hr Geschlee	cht an. *		
O Weiblich						
Männlich						
Ich möchte lieber nic	cht sagen.					
<u> </u>	-					
Bitte geben Sie in dem Kurzantwort-Text	nachstehen	iden Feld Ih	ren aktuelle	en Tätigkeiti	en an. *	
Bitte geben Sie Ihre Erf	ahrung mit	der Erstellu	ng von Date	envisualisie	rungen an. *	•
	1	2	3	4	5	
wenig erfahren	$\bigcirc$	0	$\bigcirc$	0	$\bigcirc$	sehr erfahren
Bitte geben Sie Ihre Erf	ahrung mit I	Flourish an	.*			
	1	2	3	4	5	
wenig erfahren	0	0	0	0	0	sehr erfahren
				*		
	tools verwe	enden Sie so	onst noch?			
Welche Visualisierungs						
Welche Visualisierungs Kurzantwort-Text						

Langantwort-Text



#### Aufgabe 4 - Onboarding Interface Elemente durchklicken

Öffnen Sie nun das Onboarding.

Welche Onboarding Stages können Sie sehen? Bitte fügen Sie die Antwort unten im Textfeld hinzu.

Lesen Sie jetzt die Onboarding Messages durch in der "Reading" Section.

Kurzantwort-Text

#### Aufgabe 5 - Hinzufügen von Stage

Sie wollen nun das die Onboarding adaptieren. Fügen Sie einen Stage hinzu mit dem Namen "Remember". Verwenden Sie dazu eine beliebige Farbe und folgendes Icon: "fa fa-book".

Beschreibung (optional)

#### Aufgabe 6 - Hinzufügen von Onboarding Text zur gerade erstellten Stage

Fügen Sie nun folgende Onboarding Message zum neu erzeugten Stage hinzu. Verwenden Sie dazu die unten angeführten Informationen.

Titel: Remember key facts

Message: This is a dummy onboarding message explaining how to remember key facts.

Postion: Verwenden Sie den Picker um beliebige Koordinaten auszuwählen.

Beschreibung (optional)

#### Aufgabe 7 - Adaptieren von Onboarding Text

Ändern sie nun den Onboarding Text in der Section "Reading" Nummer 3. Fügen Sie das Wort "Adaptions" am Ende hinzu.

Beschreibung (optional)

#### Aufgabe 8 - Löschen von Onboarding Text

Löschen Sie nun die Onboarding Message des visuellen Marker 4 in der Section "Reading".

Beschreibung (optional)

#### Aufgabe 9 - Löschen von Onboarding Stage Interacting

Löschen Sie nun den Onboarding Stage "Interacting".

Beschreibung (optional)

#### Aufgabe 10 - Speichern

Speichern Sie nun die Änderungen im Template.

#### User Experience Evaluation Transcript 1

A: Genau, also ganz kurz nur zum Ablauf, was quasi jetzt da die nächsten paar Minuten passieren wird. Also zuerst quasi werde ich so einen kleinen Fragebogen zu demographischen Daten, dann eben wollen wir kurz präsentieren, wie das Flourish Template grob ausschaut. Dann würden wir gerne diese – den Link zu diesem Flourish Template sharen mit dir, dass du da quasi auf deinem eigenen Device das anschauen kannst, dann habe ich so kleine Aufgaben, also einfach so kleine Tasks, die du dann durchführen kannst, um dann bisschen hineinzukommen in das, wie wir uns das gedacht haben. ob das quasi Sinn macht und dann habe ich noch ein paar abschließende Fragen. Wo es eben darum geht, wie die Benutzung ist, auf was wir noch aufpassen könnten, welche Features noch sinnvoll wären. Genau. Eben aufgezeichnet wird das Ganze, ich werde dann auch schriftliche Notizen anfertigen und das Interview wird wahrscheinlich so zwischen einer halben Stunde und Stunde dauern, je nachdem, wie schnell wir sind.

#### B: Passt.

A: Passt, dann komme ich kurz zu den demographischen Daten. Du hast eh schon gesagt, quasi, wo du arbeitest, was deine Aufgaben sind, aber darf ich vielleicht noch dein Alter wissen?

B: 35. Ich weiß es tatsächlich manchmal nicht, ich habe das verdrängt.

A: Ja, ich habe auch das Problem hin und wieder. Geschlecht ist männlich.

#### B: Männlich. Ja.

A: Passt. Dann ist meine Frage, wie viel Erfahrung hast du schon mit dem Flourish Tool gesammelt? Also sehr wenig Erfahrung, wenig, durchschnittlich, erfahren oder sehr erfahren?

B: Also erfahren würde ich sagen. Und jetzt habe ich gerade so den anderen Code kopiert, weil ich einen Workshop für die Teams entwickle. Also erfahren bis sehr erfahren würde ich sagen.

A: Perfekt. Hast du Erfahrungen mit anderen Visualisierungstools auch, wie Data Rapper, Tableau oder so?

B: Data Rapper, ja, das haben wir im Zuge unserer Dossier Academies, benutzen wir das gerne.

A: Tableau, wird das auch verwendet?

- 32 B: Nein.
- 33 A: zu groß?
- 34 B: zu Business, zu wenig Journalismus.
- 35 A: Okay. Passt, dann habe ich die nächste Frage eigentlich auch schon. Also Data
- 36 Rapper und Flourish quasi, in eurer Routine. Okay. Gibts bei euren Visualisierungen,
- 37 die ihr da für Leserinnen quasi erstellt, gibt's da schon irgendwelche Lesehilfen für
- 38 Leserinnen und wenn ja, wie wird das quasi umgesetzt, also konzeptionell, technisch,
- 39 so auf der Basis?
- 40 B: Es ist Lesehilfe eigentlich quasi der Text, der die Grafik begleitet. Also das ist keine
- 41 große Sache, weil es bei uns meistens doch relativ easy Sachen sind. Und auch ja,
- 42 ansonsten, klassischerweise geben wir halt das Lesebeispiel. Wenn es komplexer
- 43 wird. Aber ja, also da setzen wir schon eher auf Fließtexte, der mehrere Zeilen lang
- 44 sein kann und einführt in die Materie und somit die Grafik eher selbsterklärend ist.
- 45 A: Okay. Passt. Also das war sogar mal der 1. Teil.
- 46 B: Kurz zu den Tools, die wir auch hin und wieder verwenden oder was für mich dann
- 47 eine Grundlage ist für aufwendigere Visualisierungen Progress.
- 48 A: Okay, ja.
- 49 B: Ja, weil es, aber nicht so ausgereift ist wie die anderen Beispiele, die ihr angeführt
- 50 habt, aber Visualisierungstypen unterstützen, die jetzt andere nicht unterstützen,
- 51 andere Tools.
- A: Mhm. Würdet ihr andere wie (unv.) auch verwenden, oder wird das rein mit visuellenTools gemacht.
- 54 B: Ja, schon, schon, wobei da haben wir eben Flourish installiert. Wir haben relativ
- 55 viel selber gebaut früher. Dann hat halt auch quasi der Datenjournalist das (unv.)
- 56 ausgetauscht hinter dem Chart. Das war nicht groß anpassbar, das war unflexibel und
- 57 ja, das war ein Mitgrund, warum wir dann auf Flourish umgestiegen sind. Oder durch
- 58 diese Google News Initiative. Ja.
- 59 A: Okay. Passt. Gut. dann würde ich jetzt meinen Screen teilen und zeige einfach mal
- 60 her, was wir uns da gedacht haben. Mit dem Flourish Template. Und zwar, also wir
- 61 haben jetzt quasi, also der Florian, ist hauptsächlich quasi bei der technischen
- 62 Entwicklung dieses Flourish Templates also das ist die (unv.) Visualisierung, die der

31

Florian vorher schon angesprochen hat. Also wir haben uns gedacht, wir starten einfach mit einer eher komplexeren Visualisierung für Leserinnen. Weil wir sind quasi durch unsere Studien, die wir gemacht haben, auch draufgekommen, dass guasi diese Leserhilfen oder das Onboarding oder die Hilfestellung eigentlich bei sehr einfachen Visualisierungstechniken eigentlich gar nicht so wirklich verwendet wird, bzw. wird gar nicht so wirklich gebraucht, sage ich jetzt mal. Und wir haben da jetzt eben ein Template gebaut. Also wir - ich öffne das mal, du kennst eh, wie das funktioniert. Haben wir gleich mal die Daten für diese G Map geladen. Wenn man das Onboarding, sagt, okay, man hat diese Map Visualisierung, man möchte vielleicht seine eigenen Daten auch noch hochladen. Aber wir sagen mal, wir haben das jetzt guasi schon. Wir sagen, okay, diese Visualisierung braucht jetzt eigentlich eine Lesehilfe, weil das ist doch eher komplex. Dann kann man da quasi auf dieses show Onboarding klicken in der Sidebar, eh bekannt von Flourish und wir haben uns überlegt, rein konzeptionell zum Onboarding, es gibt da unten guasi diesen Fragezeichen Button. Ich kann drauf klicken, habe dann verschiedene Tool Tipps, zeige ich dir gleich so und er geht das dann ein. Also das guasi sind abstrakt geschriebene Onboarding Messages mit diesen visuellen Markern. Wir haben diese ganzen Marker in 3 Sections quasi aufgeteilt, quasi es gibt Text oder Onboarding Messages, wie wir sie nennen, how to read the charts, quasi, aber auch, wie man damit interagiert und auch, wie man diese Charts analysieren kann. Und wir haben da guasi unsere Sections, die visuellen Marker, da klickt man drauf und quasi kann man sich da durch navigieren als Leserin und Leser und kann guasi ein bisschen so, nicht geführt, sondern eher rein explorativ diese Dinge anschauen. Ja, also das ist quasi mal so die Grundidee von diesem Onboarding Konzept und dann gibt's da oben auch noch natürlich diese Onboarding Messages, sind von uns quasi geschrieben und wir vermuten, dass es sozusagen die Need gibt, dass man diese Messages, die wir da anbieten, auch adaptieren kann. Das soll dann quasi so passieren, dass es einen edit mode gibt, das ist dann halt implementiert, quasi, dass es einen edit mode gibt mit einem Toggle Button da oben, da kann man die Texte öffnen und adaptieren und die werden dann wieder gespeichert.

C: Ich wollte nur sagen, du könntest das Video herzeigen, das eine, zweite Video.

A: Ja, was man noch machen kann, ist quasi dieses Onboarding auch zu solchen Stages hinzuzufügen, das sind quasi Reading, Interacting und Analyzing. Kann man GIFS hinzufügen, darfst du quasi alles nachher noch probieren, die Onboarding

- 96 messages, das wird dann im Template gespeichert. Das ist quasi mal die Grundidee
- 97 und ich würde dich einfach hast du einen Flourish Account?
- 98 B: Ja.
- 99 A: Passt, dann schicke ich dir den Link zu dem Template, ich würde dich bitten, dass
- 100 du einfach deinen Screen sharest und dann würde ich dich bitten, ein paar Sachen
- 101 durchzuführen, auszuprobieren, weil dann können wir am besten nachvollziehen, wie
- 102 das funktioniert. Und während du das machst, vielleicht könnte ich dich auch bitten,
- 103 dass du einfach laut deine Gedanken aussprichst. Sachen, die dir komisch
- 104 vorkommen, also gerne einfach ansprechen. Der Link sollte schon da sein.

105 C: Darf ich noch was sagen zum Speichern oder ist das eher nachher dann?

- 106 A: Nachher dann.
- 107 B: Wo finde ich bei Teams -
- 108 A: Ach so, ganz oben ist neben dem Verlassen Button sind so kleine Icons und das ist
- 109 im Zweiten also bei mir ist es das Zweite. Ganz oben, also wenn du das Fenster
- 110 hast, quasi, dann ist rechts oben so ein Verlassen-Button. Hast du den?
- 111 B: Nein.
- 112 A: Okay, dann schicke ich dir das per E-Mail. Okay, ich schicke einfach per Mail.
- 113 Einfach so ganz, den Link nur ganz formlos, bevor wir da lange suchen.
- 114 B: Gut. Passt. Gut, seht ihr meinen Bildschirm?
- 115 A: Noch ist grau. Aber langsam system preferences sehe ich.
- 116 B: Okay, seht ihr das?
- 117 A: Nein. Es ist der Chrome da, aber sehen tue ich noch nichts.
- 118 B: Okay... Komisch...
- 119 A: Vielleicht kannst du nochmal quasi das sharen entsharen.
- 120 B: ... .Jetzt?
- 121 A: Ich sehe immer nur grau. Florian, siehst du das auch, grau?
- 122 C: Ja, ich komme auf Teams grau.
- 123 B: Okay, ich mache das über die Share Tray. Dann sage ich: meinen Bildschirm...
- 124 A: Komisch.

B: Ja, ich verwende so gut wie nie Teams, da muss ich dem irgendwelche Rechte geben... Komisch... A: Normalerweise müsste das eigentlich funktionieren so, ich glaube nicht, dass man irgendwelche Rechte braucht. Glaube ich. C: Welches Tool wird sonst verwendet zum – sonst können wir – oh nein, das ist blöd, weil wir aufzeichnen. A: Ich kann sonst gerne einfach meinen Bildschirm nochmal teilen und du sagst mir einfach, wo ich hinklicken soll, vielleicht probieren wir das so, wenn es nicht funktioniert. B: Ich probiere noch kurz aus Teams aussteigen und wieder rein. A: Okay, passt. (...) A: Die Technik ist echt furchtbar, hin und wieder. C: Ah, zulassen. So. B: Okay, bin ich jetzt zweimal drinnen. Gut, probieren wir das mal. Ist schon wieder alles grau. Und dann ... mah ... A: Komisch. B: Jetzt probiere ich noch eines... A: Ah ja, jetzt da. B: So, jetzt bin ich über den Browser drinnen, okay.... C: Wir können es auch gerne so machen -B: Ich glaube, das ist jetzt geschafft. Seht ihr es jetzt? C: Nein. Wir können das auch so machen und jetzt sollte man meines sehen. A: Ich sehe gar nichts mehr. C: Sekunde. A: Ah jetzt, jetzt ist mal das gesharte weg. C: Okay, Teams, was ist los. A: Ah, jetzt sieht man deinen Screen. Ja, passt. Vielleicht machen wir das so -

- 153 C: Ich kann dir die Steuerung übergeben.
- 154 A: Passt. Mir? Geht's extern nicht.
- 155 C: Da sollte es so einen Button geben, Steuerung anfordern, gibt's das nicht -
- 156 B: Weiß ich nicht.
- 157 C: Okay. Dann müssen wir klicken.
- 158 A: Ja, vielleicht kannst du einfach sagen, wo du hinklicken möchtest und der Florian
- 159 klickt für dich und wir gehen das so durch. Genau, also die Idee wäre quasi, dass -
- 160 Florian, magst du es vielleicht mal öffnen. Das Template öffnen. Genau. Quasi, ich
- 161 würde jetzt ein paar Sachen sagen, die du dem Florian sagst. Es sind nur 4 ganz kurze
- 162 Punkte, ich bin es vorher kurz durchgegangen, halt blöd, dass du das jetzt nicht so für
- 163 dich ausprobieren kannst, aber der Link ist verfügbar, den kannst du gerne nachher
- 164 noch testen. Also die 1. Aufgabe quasi wäre, dass du für diese Visualisierung das
- 165 Onboarding aktivierst.
- 166 B: Da klicke ich rechts auf Show Onboarding.
- 167 A: Perfekt. Genau. Und du möchtest jetzt da die Position von diesem Onboarding
- 168 Button möchtest du verändern, z.B. weiter nach rechts gehen. Wie würdest du das
- 169 machen?
- 170 B: Was ich mich jetzt gefragt habe, was sieht der User, die Userin. Das wäre mir nicht
- 171 klar. Also rechts -
- A: Ja, passt, also wir haben Florian, kannst du das herzeigen, diese Reuters-Story,
  bitte.
- 174 C: Okay. Das, was der User bzw. die Userin sieht, also das kennst du ja, diese
- 175 Template Creation und dieses Setting Menü sieht ja der User und die Userin nicht.
- 176 A: Nein, nein, die Leserin, du meinst die Leserin, was die sieht von dem Onboarding,
- 177 oder?
- 178 B: ja.
- 179 C: Also, wenn man das z.B. einbaut von Reuters -
- 180 B: Wir haben diesen Artikel hergenommen, haben versucht, zu schauen, wie das
- 181 ausschaut, wenn wir das da einbetten.
- 182 C: Da hast du den Originalartikel -

C: Das ist der Original, das ist der Interaktive. Und das ist der ganze Artikel nochmal interaktiv und da ist jetzt einfach ein onboarding dabei und dann kannst du da herklicken und das ist halt nur das default onboarding, da sind keine customized dabei, das ist einfach gepublished und das Menü verschwindet. Also, dieses Menü.

B: Okay, verstehe, das heißt, das entspricht eh dem, weil im Flourish sehe ich ja das, was die Leserin sieht dann.

C: Das Menü brauchen wir nur zusätzlich, weil das sieht man dann ja nicht, das Menü verschwindet dann.

B: Okay.

A: Ja.

B: Also um die Position zu ändern muss ich da erst onboarding anklicken rechts.

#### A: Genau.

B: Weil es üblicherweise dort – genau. Okay und ich kann jetzt einen Abstand definieren. Egal, irgendwas. Minus 500, um zu schauen, was passiert. Okay. Dann 500 werde ich mal ausprobieren.

A: Genau. Also ja. Also da ist quasi der Abstand zur Visualisierung.

C: ich glaube, 40 ist gut. wenn man das 30 Mal probiert während dem Testen, dann weiß man das gut.

A: Okay, genau, das wäre mal die 2. Aufgabe, die dritte wäre, du möchtest jetzt quasi diese generierten Onboarding, diese Steps, die du in diesem Menü siehst, denen möchtest du quasi einen Step hinzufügen, wie würdest du das machen?

B: Da würde ich bei dem Punkt (unv.) Stages den Edit-Mode aktivieren.

A: der macht jetzt gerade noch nichts, aber guter Punkt. Also wir haben das oben in diesem Menü, rechts oben, wo die kleine Notiz steckt, da wird das versteckt sein. Wenn du vielleicht mal draufklicken kannst, Florian.

C: Was man dazu sagen muss, dieser Edit-Mode, der nachher dabei ist, da wird es möglich sein, eben, was die Christina vorher gesagt hat, dass man die Nachrichten bearbeiten kann, aber auch neue hinzufügen. Das versteckt sich dann alles unter dem

- 212 Menüpunkt, das war schon richtig, aber ich kann dann hinzufügen, Neue, und
- 213 bearbeiten. Also hinzufügen Neuer ich mache das mal ein bisschen größer.
- 214 A: Genau.
- 215 C: Ja, habe es.
- 216 A: Passt, genau, also quasi die Aufgabe ist da, dass du da eine Onboarding Stage
- 217 hinzufügst, also kannst welche und dann eine Stage dazu.
- 218 B: Also gleich vorweg, ich wüsste nicht, was eine Stage ist, ich wüsste nicht den
- 219 Unterschied zwischen Stage und Message
- 220 A: Okay. Stage ist die Sections quasi und die Message ist dann in dem Visuellen, also
- 221 wenn du auf den Einser so draufklickst, ist das die Message. Aber danke für den
- 222 Hinweis.
- 223 B: Okay. Ja, da würde ich dann, entweder, wenn ich welche hinzufügen möchte, da
- bin ich eh schon, dann hinzufügen, da gibt's eben den 4. Button, das habe ich
- 225 verstanden.
- 226 A: Warte mal, Florian, nicht so schnell, bleib mal da oben, da gibt's ja mehr Sachen.
- 227 C: Entschuldigung.
- 228 B: Und die Aufgabe war jetzt, eine Stage -
- 229 A: Genau, wir können das machen, wir können gerne einen Titel hinzufügen.
- B: Titel, okay... Icon List, da würde ich auf die Icons klicken.
- 231 C: Hoppala, das geht natürlich auf in einem neuen Tab.
- 232 B: Okay, ja, das ist sehr sinnvoll, nehmen wir halt irgendwas.
- 233 C: Genau, Slowflake.
- B: Wunderbar. Ja, das hätte ich kopiert, da rein, die Farbe kann man noch... ja.
- A: den sieht man gar nicht, den Overlay. Ja.
- 236 B: Ich glaube, das ist bei uns so -
- 237 C: Ja, das sieht man nicht, das wird nicht geshared in Teams.
- 238 B: Aber wunderbare Farbe.

C: Da sieht man gleich eine Message, man muss bitte das Format, das Fonds – also man muss noch einmal FA dazu schreiben, das ist noch wichtig. Und dann sollte da stehen die Stage ist (unv.)

#### B: Okay.

A: Und jetzt bei der Stage, die du erstellt hast, quasi den Onboarding Text hinzufügen.

B: Ja, wenn ich jetzt Onboarding – ich würde mal runter scrollen, kann ich da auswählen, welche Stage – genau, da schaue ich mal, ob die Stage schon drinnen ist, die neue.

C: Ja, Entschuldigung, Reading, Analyzing, Interacting und Message, also sind alle da.

B: Ja und Text, würde ich selbst eingeben. Da wäre vielleicht eine Zeichenbeschränkung sinnvoll, weil alle Menschen dazu neigen, viel zu viel Text zu schreiben immer.

#### A: Ja, guter Punkt.

B: Und dann ist praktisch, dann würde ich mit diesem Fadenkreuz die Position auswählen, irgendwo. Ja, genau da. Und add message. Würde ich mal zumachen und schauen, ob das funktioniert hat.

C: Das Icon habe ich falsch geschrieben, anscheinend, aber -

#### B: Okay.

A: Genau. Ja, also das wären mal quasi die ganzen kleinen Aufgaben gewesen. Danke für deine Hinweise und das Feedback. Vielleicht gehen wir gleich zu den Fragen und der Florian hat dann auch gesagt, er möchte noch auf die Features, die wir geplant haben, möchten wir noch gerne eingehen. Aber ich möchte dich mal fragen, wie hast du die Nutzung empfunden? Das war jetzt zwar nicht auf deinem Device, aber was sagst du zu dem Template, zu der Implementation? Zu dem, was wir uns dabei gedacht haben, quasi, vielleicht kannst du da ganz kurz darüber reden.

B: Also prinzipiell, ich glaube, dass Leute, die Flourish nutzen, das ist ja relativ komplex teilweise, dass die damit gut klarkommen. Also ich finde das gut gelöst. Was für mich neu war, war quasi – also wenn ich rechts oben über das Editier-Icon von euch, also dass ich Einstellungen nicht nur in der rechten Spalte vornehmen und da nicht nur

269 meine Interaktion habe, sondern das ist ein Overlay da drinnen. Aber das hat einen

270 technischen Hintergrund, oder?

271 A: Ja.

B: Das ist so Gewöhnungssache, da kommt man rein. Also prinzipiell habe ich mich
zurechtgefunden. Bis auf die Frage, aber das habe ich eh schon gesagt, was ist
überhaupt eine Stage.

275 A: Okay. Also wir haben jetzt quasi von den Funktionen her, ist eben angedacht, dann 276 kannst du das ganz kurz bewerten. Also man kann eben diesen Edit-Mode, wie du 277 vorher gesehen hast, aktivieren. Dann kann man diese Semi-automatischen 278 Onboarding Messages, die es von der Library gibt, editieren. Man kann die Sections 279 quasi hinzufügen, man kann eigene Texte quasi schreiben für die Erklärung und dann die Idee ist, dass man diese dann reordern kann in diesem Menü. Und dass man 280 281 Sachen weglöschen kann, also deleten. Florian, habe ich noch was vergessen von den Features? 282

C: Ja, das, was man nachher noch machen können sollte, generell im Edit-Mode –
was man generell machen können wird, ist im Edit-Mode, dass man reordern kann,
also die Appearance im Menü, das bedeutet, dass man das da ändern kann. Also dass
zuerst z.B. Interacting vor Reading kommt. Dass dann dieser Stepper kommt, der ist
noch nicht dabei und vielleicht die Farbe ändern und deleten ist auch dabei, dass man
Stages auch deleten kann.

A: Ja, was sagst du dazu, ist das sinnvoll für dich, fehlt dir noch was, sollten wirirgendwas nicht machen?

291 B: Mir ist das eher zu viel. Für mich wäre gut, wenn ich sage, kann ich diese Stages 292 löschen, dass ich nur eine habe. Ich glaube, dass das LeserInnen überfordert, wenn 293 ich an dieser Stelle drei unterschiedliche Leseebenen habe und der dann selber 294 entscheiden muss, welche ist, die mit Begriffen codiert sind, die nicht selbsterklärend 295 sind wie read, analyze. Also das sind meine 5 Cent, ich würde sagen, aber die Idee ist 296 super, dass ich Schritt für Schritt und auch geordnet Hinweise kriege, wie ich diese Grafik lesen kann. aber ich glaube, das reicht, das auf einer Ebene zu lassen. Und ich 297 298 habe das Fragezeichen. Da kann man sagen, default ist sogar schon die Erklärung -299 ich kann das wegklicken, wenn es mich nicht interessiert. Kannst du mal auf Reading 300 klicken? Ja, genau, so ist, wie ich Usability am besten fände und auch, weil ich in dem 301 Bereich arbeite - ich habe diese 4 Punkte angezeigt, das erste ist schon offen, das

gibt mir den Hinweis, okay, ich kann da was öffnen, ich kann mir da was anzeigen lassen. Das ist jetzt nur, dass mein Blick geführt wird von 1, 2, 3, 4 und kann das aber dann schließen rechts unten, wenn es mich nicht interessiert.

C: Ja.

A: Ja. Florian, mal nichts sagen. Passt, okay. Was fällt dir noch auf so?

B: Also was mir schon wichtig wäre, dass ich die Farbe anpassen kann von dem Button, damit das wieder – das ist jetzt bei uns speziell so, glaube ich, obwohl das war bei zwei Kunden jetzt so, dass die das an ihr Design anpassen wollten, da genügt es, wenn ich die Farbe anpassen kann.

A: Ja, das geht dann eh.

B: Ansonsten... ja, also wie gesagt, ich würde es eher vereinfachen jetzt.

A: Okay. Kannst du dir vorstellen, dass du das quasi verwendest, also dieses Onboarding –

B: Tatsächlich, also wenn ich das so anpassen könnte, wie ich gerade beschrieben habe, gibt's sicher – also wenn Datenpunkte fehlen und ich darauf hinweisen will oder erzählerischer, mit 1, 2, 3, 4, als was herausheben. Durchaus, auf jeden Fall. Aber dann wäre es gut, wenn cih das mit jeder Flourish Visualisierung aktivieren könnte. Das ist absolut ein Konzept, das ja, Flourish interessieren könnte. Habt ihr mit denen schon geredet mal?

A: Nein, also eigentlich nicht. Aber ich glaube, ich weiß nicht, wie einfach das ist, quasi, mit Flourish Kontakt aufzunehmen.

B: Ganz sicher, wenn ihr fertig seid mit dem Projekt, das steht und ihr zufrieden seid, dann würde ich das machen.

A: Ja, gute Idee. Ja, bei den Interviews ist halt herausgekommen, dass die Welt des Datenjournalismus so divers ist, dass wir auch zuerst ein bisschen gebraucht haben quasi, um zu evaluieren, welches Tool können wir jetzt wirklich verwenden, das Sinn macht. Und es ist dann eh Flourish herauskristallisiert, dass das quasi das Tool ist, dass viele verwenden. Aber ja, vielleicht ist es eh eine gute Idee, dass wir da mal Schritte setzen, denen das präsentieren.

B: Ja, tut nicht weh. Habt ihr vor -

333 B: ich stelle nur neugierige Fragen. Habt ihr vor, das zu open sourcen?

- 334 A: Ja, die (unv.) Library ist open source, also prinzipiell baut sie auf - also das ist 335 Plotage GS, weil wir haben quasi diese (unv.) Library auf Grammatikalisierungs (unv.) wie Plotage und WeCharts, weil wir eben aus dieser Visualisierung das so 336 337 herausnehmen, dass sie quasi auf die Daten quasi abzielt und nicht jetzt nur abstrakte Dinge guasi präsentiert. Und darum haben wir guasi, dieses Konzept haben wir 338 verfolgt. Auch deswegen, in dem anderen Usecase wird sehr viel mit Plotage S 339 340 gearbeitet, deswegen kann man diese Library dann für Plotage S, WeCharts oder (unv.) verwendet. Und diese 3 sind auch sehr verbreitet. Und ja, wir können dann, 341 342 wenn es dich interessiert, den Link sharen zu (unv.), also wir sind quasi noch in der Entwicklung. Das heißt, ziemlich fertig schon mit der Library und würden das dann, ich 343 kann dir dann den Link schicken, wo du das dann verwenden kannst. Genau, ja. 344 B: Super. 345 A: Gibt's sonst noch Fragen von dir? 346
- 347 B: Jetzt habe ich dich kurz nicht gehört.
- 348 A: Gibt's sonst noch eine Frage an uns?
- 349 B: Gerade nicht.
- 350 A: Okay, okay, ja. Dann wäre das quasi, ich wäre mit meinen Fragen durch, du hast
- 351 eh quasi beim Erzählen alle Fragen beantwortet, die mich quasi interessieren. Genau,
- 352 gibt's sonst irgendwie noch allgemeineres Feedback oder Anregungen, das ist meine
- 353 letzte Frage. Irgendwas, auf das wir noch aufpassen sollen oder keine Ahnung, was
- 354 vielleicht anderen Kolleginnen helfen würde?
- 355 B: Nicht wirklich. Also ich war jetzt eh nicht zurückhaltend im Feedback, ich glaube,
- 356 ich habe eh alles gesagt, was mir eingefallen ist.

A: Perfekt, ja, super, super, dann sage ich danke. Ich würde jetzt da die Aufnahmebeenden.

- 359
- 360
- 361
- 362

#### User Experience Evaluation Transcript 2

A: Okay, sollte jetzt schon funktionieren quasi. Okay, passt, gut. dann würde ich gerne zum nächsten Punkt gehen und zwar zu den demographischen Daten. Darf ich Ihr Alter und Geschlecht wissen bitte?

B: 40 Jahre alt und männlich.

A: Welche Ausbildung haben Sie, also welche höchste abgeschlossene Ausbildung?

B: ich habe Volkswirtschaft studiert, also einen Magister und ich habe einen Master in Journalismus und neue Medien.

A: Okay. Was ist ihre derzeitige Aufgabe?

B: Ich habe gerade gekündigt bei Dossier, aber ich war dort Datenjournalismus, aber ich mache gerade Neuorientierung und mache Workshops in Datenjournalismus, bei verschiedenen Medien und Institutionen.

A: Okay. Super. Okay. Darf ich Sie als Einschätzung bitten und zwar, wie viel Erfahrung haben Sie mit Flourish? Also sehr wenig Erfahrung, wenig, durchschnittlich erfahren oder sehr erfahren?

B: Flourish, ich habe es ein paar Mal gesehen, also sehr wenig.

A: Okay. Haben Sie mit anderen Visualisierungstools, wie z.B. Data Rapper, Tableau mehr Erfahrung, können Sie da kurz darüber sprechen quasi?

B: Ja, am meisten, in den Workshops zeige ich auch, also weil es am schnellsten ist und weil es gratis ist, verwende ich meistens Data Rapper, um das den Leuten zu zeigen. Ich habe auch Tableau mal gesehen, Flourish mal gesehen, ein bisschen was gemacht, aber hauptsächlich verwende ich eigentlich Data Rapper.

A: Okay. Haben Sie vielleicht auch Lesehilfen für Ihre Leserinnen in Ihre Datenvisualisierungen eingebaut –

B: Nein, da waren wir noch hinten nach, das war nicht.

A: Okay. Perfekt. Das war mal der 1. Teil, jetzt, wo wir wissen, wer wir sind. Und jetzt würde ich quasi unser Projekt und die Idee hinter unserem Flourish-Template erklären, weil Sie eh noch nicht so viel Erfahrung haben. Und zwar, wir haben eben in diesem Seva-Forschungsprojekt das Ziel, so Lesehilfen für einerseits Leserinnen, die journalistische Inhalte und Datenvisualisierung quasi lesen und eben Biomediziner und wir haben auch von der Fachhochschule Joanneum, die haben auch Interviews

durchgeführt und wir haben da eruiert, dass eben Flourish sehr oft in dieser Domäne

verwendet wird, deshalb haben wir uns gedacht, wir wollen die Journalisten dabei 32 33 unterstützen und entwickeln ein Flourish Template. das kann man öffnen und wir haben eine (unv.) Library entwickelt, die so Lesehilfen, also die soll so Visualization 34 Onboarding, also die sich automatisch zu einer Datenvisualisierung hinzufügt. Also wir 35 haben uns konzentriert auf z.B. Grammatik, also wir können unsere Library für Plotage 36 37 S, Wecharts und (Unv.) Light verwenden. Da gibt's schon ein Set an Datenvisualisierungen, das kann man hernehmen und da hat man das Onboarding 38 39 schon integriert. Und die Idee ist, dass wir das in das Flourish Template integrieren und man hat verschiedene Lesehilfen über die Features, die man hinzufügen kann. Ich 40 würde das jetzt herzeigen, ich würde meinen Screen sharen, dass man das mal sehen 41 42 kann, was ist das ist. Ich glaube, Sie müssten jetzt meinen Screen sehen. Also das ist 43 quasi die Startseite von dem Flourish Template, wenn man da auf great visualization 44 geht, dann kommt man zu dieser Anwendung von uns. Es gibt da die Information, wie 45 man diese Änderungen, die man da in diesem Template macht, auch speichern kann. das können Sie sich nachher eh nochmal durchschauen. Wir haben da als ersten 46 Ansatz eine (unv.) Visualisierung. Das sind die Biden's Tax Roll Daten. Einfach in einer 47 48 (unv.) Visualisierung, einer interaktiven integriert, man kann dann das onboarding 49 hinzufügen. Man hat dann da unten so einen Button, wo man verschiedene Sections hat. Es werden dann so visuelle Marker quasi in diese Visualisierung integriert. Man 50 51 kann dann drauf klicken, es gibt dann einen Tool mit Text. Und man kann sich dann 52 als Leser durchklicken. Das ist guasi die Hauptnavigation. und im Zuge von diesem 53 Template kann man dann auch noch verschiedene Sachen adaptieren von diesem 54 onboarding, aber da wären wir quasi dann eh in der Evaluierung noch mehr drauf 55 eingehen. Also das ist das Konzept, das sieht die Leserin in der Visualisierung. Man kann diesen Button natürlich auch dann noch verschieben und das Ganze auch 56 57 speichern. Aber das ist quasi das, was der Leser dann für Informationen kriegt, wenn er halt Hilfe braucht, mit dieser (unv.) Visualisierung. Passt. Das war mal guasi ganz 58 59 kurze Einführung. Man kann eben – man kann durch diese Navigation navigieren. Man kann diese Texte, die visuellen Marker adaptieren, das geht alles quasi mit diesem 60 61 Flourish Template, das wir da entwickelt haben. ich würde da gerne den Link teilen, ich würde da im Chat den Link hinzufügen. Ich hoffe, der Chat ist verfügbar. 62 B: Ja. 63 A: Perfekt. 64

#### B: Habe ihn.

A: Dann würde ich Sie bitten, dass Sie vielleicht den einmal öffnen, den Link und dann mal Ihren Screen sharen.

B: Ja.

- A: Meine Frage ist auch, haben Sie einen Flourish Account?
- B: Ja, aber nur vom Dossier, also den sollte ich eigentlich nicht mehr verwenden.

A: Ist kein Problem, ich habe einen Testaccount, ich würde einfach den da in den Chat, die Benutzerdaten quasi und das Passwort da in den Chat gerne posten, dann können Sie meinen verwenden.

B: Passt, ich schaue jetzt noch, wo ich meinen Screen sharen kann.

A: Ganz rechts oben sollte so ein Button sein mit teilen, neben dem großen roten Verlassen-Button.

B: Da habe ich auflegen, aber den Screen Share sehe ich noch nicht – teilen, ah ja, jetzt habe ich es. Funktioniert es?

- A: Ja, perfekt, ja, super. Perfekt, okay.
- B: Sign in.
- A: ja, genau. Ich glaube, dass das nämlich sonst nicht funktioniert.
- B: Passt, dann muss ich nochmal zum Chat....
- A: Können Sie vielleicht einfach den Link nochmal öffnen, den ich geschickt habe.
- B: Ja.

A: Perfekt, genau. Ja, perfekt. Also die Idee ist jetzt, dass Sie sich ganz kurz mit diesem Interface quasi bekannt machen. Dass Sie da einfach auf create visualization klicken, dass Sie sich das in Ruhe anschauen, wie was wann und dann starten wir quasi mit den Aufgaben.

#### B: Okay.

A: Ja? Passt, genau, also das auf der rechten Seite, was Sie da sehen, das rechte Panel, das wird quasi von Flourish bereitgestellt, also das haben wir quasi, da kann man noch Informationen hinzufügen und das, linke Seite, ist das, was wir dann quasi entwickelt haben, nur von der Übersichtlichkeit quasi. Genau. Man kann dann auch,

- 94 wie es vielleicht eh von Flourish bekannt ist, auch die Daten ändern, wir arbeiten heute
- 95 mit denen, die quasi angezeigt werden. Okay, passt, perfekt, gut. dann würde ich gerne
- 96 mit den Aufgaben starten. Und zwar, Sie haben jetzt quasi die 1. Aufgabe, Sie erstellen
- 97 eine Visualisierung, haben Sie schon durchgeführt. Ich würde Sie jetzt bitten, dass Sie
- 98 nochmal Sie haben den Saving Screen durchgelesen, der, was als erster war, quasi,
- 99 wo die Information war?
- 100 B: Ja.
- 101 A: Passt, dann haben wir die Aufgabe auch schon erledigt, dann geht's schon weiter.
- 102 Und zwar, die Aufgabe wäre jetzt, dass Sie diese Lesehilfe, zu dem Onboarding, zur
- 103 Visualisierung hinzufügen. Perfekt. Nächste Aufgabe wäre, dass Sie die Position für
- 104 diesen Onboarding Button mit der X ist Offset auf 50 Pixel stellen. Perfekt. Das sind
- 105 ganz kurze, kleine Aufgaben. Also das jetzt die nächste Aufgabe wäre, dass Sie das
- 106 Onboarding öffnen und können Sie mir sagen, wie viele Onboarding Stages Sie da 107 sehen?
- 108 B: Onboarding Stages, das sagt mir jetzt nichts.
- A: Ja, okay. Okay. Also die Stages nennen wir, das Reading, Interacting and
  Analyzing, das sind quasi Sections, die wir definiert haben, die wir auch durch Studien
- belegt haben, dass die quasi die Aufteilung von den Informationen, die man braucht,
- 112 um das zu verstehen, dass man da durchklicken kann, wenn z.B. der User sagt, okay,
- 113 ich weiß nicht mehr, wie man da interagiert, dann kann er auf Interacting klicken und
- 114 sieht quasi nur die Informationen, die er braucht zur Interaktion mit dieser (unv.)
- 115 Visualisierung. Also dass die LeserInnen nicht nur so durchgeführt werden wie mit
- 116 einem Guide, wo man Schritt für Schritt durchgehen, sondern dass man explainatory
- 117 diese Stages durchklicken kann. Und wir nennen diese 3 Sections Stages. Okay.
- 118 Könnten Sie jetzt die Reading Section, also die Reading Stage öffnen und die
- 119 Onboarding Messages durchlesen und ganz kurz Feedback geben, ob die verständlich
- 120 sind.
- 121 B: Okay, reading... da würde ich noch einen Abstand machen auf dem Punkt.
- 122 A: Ah ja, danke.
- 123 B: represents quantitative... ja. Das finde ich gut verständlich.
- 124 A: Okay. Passt. Die nächsten Punkte alle bitte.
- 125 B: Ja.

#### A: Okay. 3 und 4 auch noch.

B: Das finde ich ein bisschen kompliziert... Ja, das verstehe ich nicht ganz. Also wie sagt man, closely back together... ach so, na, doch. Einfach ein größerer Abstand, wenn es nicht die gleiche Kategorie ist, ja. Bisschen habe ich gebraucht, aber ja.

A: Aber ist gut, vielleicht müssen wir da ein bisschen bei der Formulierung arbeiten. Okay, passt. Das war quasi schon die 4. Aufgabe, jetzt geht's zur nächste. Und zwar, Sie wollen nun dieses Onboarding, wie das quasi wieder zur Verfügung stellen, aber adaptieren. Und zwar, Sie wollen eine Stage hinzufügen mit dem Namen remember. Verwenden Sie einfach eine beliebige Farbe und einfach ein Icon, das Sie dann einfach ganz kurz Bescheid geben, wie das heißt. Wie würden Sie da vorgehen?

B: Also... edit onboarding messages und stages... Okay. Wie soll es heißen?

A: Es soll heißen remember.

B: Ja.

A: Sie sind jetzt in der falschen Section. Kein Problem. Scrollen Sie einfach ein bisschen nach oben. Also dass links, dass sich die Visualisierung – also auf der Seite nach oben scrollen. Ja, genau.

#### B: ah.

A: Genau. Perfekt. Also, der onboarding Stage soll heißen remember und eine beliebige Farbe auswählen und die – genau. Und das Icon können Sie einfach von da unten, da ist so ein kleiner Text einfach, das kopieren, das Format, z.B. das Feuer da. Einfach als Test. Also unter dem Textfeld, Entschuldigung, steht, wie das Icon heißen soll und da einfach nur das, wie – was da z.B. steht, rauskopieren. Also das FAFA thumbs up z.B.

B: Das sehe ich jetzt gerade nicht.

A: Also bei den markierten, bei dem markierten Satz, den Sie da haben.

B: Ach so, das da.

A: Okay, einfach das eine Example. Genau, perfekt. Genau.

B: So. Add Stage oder ich schaue noch kurz, was es da gibt. Okay, ich würde es zu reading machen, oder? Nein –

- 155 A: Nein. Perfekt. Okay. Das war schon die Aufgabe. Die nächste Aufgabe und zwar,
- 156 Sie wollen jetzt zu dieser neuen Stage, die Sie da quasi erstellt haben, eine Info, also
- 157  $\,$  einen Text hinzufügen. Und zwar, ich speichere das da ganz kurz quasi. Ich kopiere
- das dann in den Chat rein, wie das heißt. Also der Titel soll sein, remember Keyfacts,
- 159 Message, also das ist eine dummy Onboarding Message, how to remember Keyfacts
- 160 und die Position können Sie einfach so wählen, dass die Koordinaten beliebige
- 161 Koordinaten sind.

162 B: Können Sie das nochmal wiederholen.

A: Ja, sicher. Also zu dem Onboarding Stage remember fügen Sie jetzt diese
Onboarding Message, also den Text, diese Tool Tipps hinzu. Wie würden Sie da
vorgehen?

166 B: ja. Also, hier haben ich remember.

- A: Genau, perfekt. Und ich habe jetzt da soll heißen, remember keyfacts, ich habe
  das in den Chat gepostet, wie der Titel wäre von der Message. Da können Sie das
  einfach rauskopieren. Genau, den Text. Genau, perfekt. Und die Position einfach
  beliebig, wo Sie den hinsetzen wollen, den Marker. Sie können den Picker da
  verwenden, der da zur Verfügung steht.
- 172 B: Okay.

A: Perfekt. Perfekt, genau. Okay. So, also die nächste Aufgabe wäre, Sie wollen jetzt 173 den Onboarding Text ändern in der Reading Section und zwar beim Marker 3. Sind 174 175 quasi da draufgekommen, irgendwas stimmt nicht - genau, beim Marker 3, und Sie wollen am Ende hinzufügen: Test. Perfekt. Gut, dann sind wir schon fast am Ende und 176 zwar, nächste Aufgabe wäre, dass Sie die onboarding Message löschen, beim 177 visuellen Marker 4 in der Reading Section. Also.... Den 4er wollen Sie löschen. Perfekt. 178 179 So, jetzt haben wir schon die vorletzte Aufgabe und zwar, löschen Sie den onboarding stage Interacting. 180 181 B: Okay... schauen wir mal da her... das komplette, ah ja, da ist es. Okay.

- 182 A: Perfekt. Okay. So, letzte Aufgabe ist, dass Sie jetzt diese Änderung, die Sie da
- 183 durchgeführt haben, speichern.
- 184 B: In den Settings... Settings... oh... ah. Hab es. Ich mach mal. Mach ich gerade sehr
- 185 schlechte Sachen?

#### A: Alles gut.

B: Okay.

A: Perfekt, genau.

B: Okay.

A: Genau, also das ist jetzt gespeichert, Sie können jetzt gerne diese Seite einfach nochmal neu laden im Browser, dann schauen Sie mal quasi, ob Ihre Settings noch da sind.

B: Okay, also das müsste ich nochmal gescheiter lesen nächstes Mal.

A: Können anschauen –

B: Aber ja, das eine ist weg, remember ist noch da.

A: Perfekt, okay, passt. Ja, gut, das waren jetzt quasi die Aufgaben, danke fürs Durchführen. Wir würden jetzt zu den abschließenden Fragen schon kommen. Und zwar, wie haben Sie die Benutzung quasi von diesem Template empfunden?

B: Ja, kann man was mit machen, aber insgesamt geht's schon gut, also es ist viel intuitiv, dass ich da auf viel hinklicken kann und was dazu schreiben kann. ich bräuchte ein... irgendeine Erklärung, was hier gut wäre, also wie viele ich machen sollte, irgendwie best practice oder so. Dass ich die Leute nicht überfordere oder auch nicht zu wenig erkläre. Ja, das Speichern ist ein bisschen kompliziert, aber geht auch, wenn man das ein, zweimal gemacht hat. Aber nein, sonst ist es – ja, ein bisschen bin ich jetzt verwirrt zwischen hier diese Einstellungen und Safe und dann habe ich hier die da nochmal, mir ist nicht ganz bewusst, was jetzt genau was ist. Also das habe ich einfach noch nicht durchschaut. Genau, nein, aber sonst ist das sicher hilfreich. Ja.

A: Okay. Wie bewerten Sie jetzt da das Konzept quasi, diese Lesehilfe? Was sagen Sie da dazu quasi? Also können Sie sich vorstellen, dass das Leserinnen dabei unterstützt?

B: Ja und das wäre jetzt z.B. bei online interaktiven Grafiken oder so, hätte ich dann diesen Button dabei und das können dann die Leser und Leserinnen sich noch anschauen. Nein, finde ich auf jeden Fall sinnvoll, weil ich kann es eben, ich finde gut, dass ich das trennen kann, weil eben, Informationen über die Grafik könnte ich ja, soweit ich weiß, in Flourish selbst dazu, also wenn es um den Inhalt geht und die Daten, also dass ich das quasi getrennt habe und ein eigenes Ding habe, dass mir so

- 217 Metainformationen gibt, was ist das eigentlich und was kann das und was nicht, das
- 218 finde ich grundsätzlich gut.
- 219 A: Wie bewerten Sie jetzt die Einbettung und das Flourish Template im Sinne von
- 220 Funktionen? Fehlen welche?
- 221 B: Uh, ja, also ich glaube, also das kann jetzt nur eine Spontanantwort sein, ich glaube,
- 222 da müsste ein paar Stunden ich mich herumspielen, bis ich draufkomme, ob noch was
- 223 fehlt. Aber bis jetzt geht's flott und schnell, das passt schon, ich bin wirklich dann,
- 224 wo war das nochmal mit der Farbe, genau... die kann ich da oben aussuchen.
- 225 Messages kann ich hier...das ist auch einfach. Und das also nein, mir fehlt jetzt
- 226 nichts grundsätzlich, auf die Schnelle.
- 227 A: Okay.
- 228 B: Wie kommt man auf die Icons?
- 229 A: Die da? Die wir da -
- 230 B: Ah, okay.
- 231 A: Genau. Da rechts, wenn man auf die -
- 232 B: Ah, die heißen.
- 233 A: genau, die kann man, wenn Sie den Link öffnen, dann können Sie sich die Icons
- 234 aussuchen und da die Formate dort so beschrieben sind, da wird das reinkopiert und
- 235 dann kann man für jede Stage das Icon verwenden.
- 236 B: Okay, das möchte ich noch anschauen. Ah, da sind sie und das kopiere ich dann,
- 237 ja, okay, macht auch Sinn.
- 238 A: Okay. Eine Funktion haben wir noch nicht ganz, vielleicht könnten Sie nochmal, da
- 239 unten neben dem Onboarding Button ist ein kleiner (unv.) Button. Würden Sie den nur
- 240 kopieren, also noch weiter runter.
- 241 B: Ich komme gerade nicht weiter runter.
- 242 A: Sehen Sie den Button unter dem pinken Button?
- 243 B: Ah, das habe ich nicht erkannt.
- 244 A: Kein Problem. Würden Sie jetzt mal die Reading Section öffnen. Also genau, da
- 245 kommt jetzt eine Funktion, die haben wir vorher nicht wirklich verwendet, weil die
- 246 wahrscheinlich ein bisschen versteckt ist, diese Funktion. Man kann da quasi Circles

hinzufügen, dann kann man quasi durch die einzelnen Onboarding Messages durchnavigieren anhand von diesen Circles, da hat man oben so Pfleile, da wäre die Idee, dass man sagt, okay, man startet beim ersten und lässt sich so durchführen und das wird für die User auch so sichtbar sein, der Toggle Button, bevor sie das Template publishen.

B: Ja, jetzt verstehe ich, das sind alle blauen und alle reading. Die Reihenfolge habe ich nur noch nicht ganz verstanden. Das ist ein Bug, oder, 2, 1.

A: Eigentlich sollte es sein, bei 1 beginnend, 1, 2, 3 und dann die nächste Stage 2 usw. Aber die Reihenfolge passt noch nicht ganz. Können Sie vielleicht kurz was dazu sagen zu dieser Navigation?

B: Ja, das passt – also der Button war sehr versteckt und sonst ist es super, weil ich einen Überblick habe mit allen Sachen, die ich geschrieben habe und da kann ich am Schluss durchklicken und alles kontrollieren.

A: Okay, perfekt. Super, dann wären wir quasi mit den Fragen, den Aufgaben durch und ich habe noch eine kurze Frage. Und zwar, wie würden quasi nach dieser Evaluierung dieses Flourish Template noch ein bisschen adaptieren und dann wäre die Idee, dass wir damit eine Direct Study durchführen, wo die Journalisten ein bisschen mehr damit arbeiten, vielleicht eigene Texte hinzufügen, vielleicht eigene Daten hochladen. Hätten Sie vielleicht da in nächster Zeit Ressourcen, dass Sie sich, weiß nicht, eine Stunde oder so beschäftigen, oder im Zuge von einem anderen Projekt oder einem Workshop oder so?

B: Ja, gerne, also am liebsten würde ich das wirklich in einen Workshop einbauen, da muss ich aber noch überlegen, was als nächstes kommt und was da sinnvoll wäre. Also nächstes Semester wäre wieder FH Joanneum, da macht das glaube ich auch am meisten Sinn. Weil die anderen Workshops sind doch viel Stoff in wenigen Tagen, aber wenn wir ein ganzes Semester Zeit haben, kann ich das gerne mal einbauen und den Leuten zeigen, hey, das gibt es auch noch und schauen wir uns das mal an.

A: Okay, das wäre super, weil vielleicht können das die Studierenden dann auch benutzen, vielleicht können wir, wenn es da konkrete Pläne gibt, uns nochmal zusammenschließen, weil ich würde auch Material vorbereiten und wir haben eh direkte Kontakte auch zum FH Joanneum und vielleicht können wir das im Zuge von dieser Lehrveranstaltung dann nochmal im Detail evaluieren.

279 B: Voll gerne, das wäre dann wieder das nächste Sommersemester. Also jetzt ist

280 nichts, das kommende Semester, aber dann, im Sommersemester, das wäre kein

281 Problem -

282 A: Ja, super, das wäre kein Problem, wenn das erst im Sommersemester ist und wenn

283 es dann Termine gibt, dann reden wir uns zusammen, wie wir das in dieser

- 284 Lehrveranstaltung integrieren könnten und es wäre voll nett, wenn Sie mir da kurz
- 285 Bescheid geben könnten, wenn Sie da Infos haben.
- 286 B: Ja, super. Ja, da bräuchten wir FLourish Accounts für alle. Oder?
- 287 A: Ja, ich kann da schauen oder mit dem Robert nochmal reden, wie das ist mit den
- 288 Studierenden, ob das vielleicht in einer anderen Lehrveranstaltung schon verwendet
- 289 wird, aber um das können wir uns kümmern, dass das dann funktioniert. Ja, perfekt,
- 290 dann schauen wir, dass wir in Verbindung bleiben, vielleicht einfach eine E-Mail
- 291 schreiben, dann können wir uns vielleicht nochmal zusammentelefonieren oder so. Bis
- 292 nächstes Sommersemester und sonst sage ich danke für die Zeit. War sehr wertvolles
- 293 Feedback. Danke.
- 294 B: Bitte, ja, danke auch
- 295 A: Genau, die Einverständniserklärung darf ich noch kurz erinnern, bitte die noch kurz
- an mich senden. Und perfekt, dann sage ich danke schön für die Zeit und schönen
- 297 Tag noch und wir hören uns.
- 298 B: Ja, Ciao, super, danke, Baba.
- 299

301

302

303

- 300

- 304
- 305
- 306
- 307
- 308
# User Experience Evaluation Transcript 3 - group setting

#### A: Bitte?

B: Es fühlt sich irgendwie nach Schultest an.

A: (lacht) Nein, es ist kein Test, also ich hoffe – also es geht uns alle, wir werden alle getestet, also eigentlich müsste ich voll nervös sein.

#### (...)

B: Frage ist vielleicht, wenn wir jetzt doch alle dieselben Vorlagen verwenden, ob wir uns da jetzt nicht in die Quere kommen?

A: Ich kann gerne da in den Chat nochmal posten, meinen Testaccount, dann kann den vielleicht noch jemand verwenden, aber generell, so ausprobieren – ich schicke es zur Sicherheit nochmal rein, weil, wenn es nicht geht, dann...

#### (...)

A: Gibt's Fragen oder steht ihr irgendwo an, oder geht's eh?

B: Ja, ist wahrscheinlich -

- C: Sollen wir denn das Formular schon absenden?
- A: Bitte, ja, wenn ihr fertig seid, einfach senden.

C: Okay.

B: Also einen Verbesserungspunkt, glaube ich, hätte ich schon.

#### A: Ja?

B: Ich habe den Marker für den remember stage auf – müsste ich jetzt nochmal nachschauen... - auf welcher Position weiß ich nicht mehr, auf jeden Fall beim orangen und letzten kleinen Segment gesetzt. Und wenn ich da jetzt in der mobilen Version das Thema Flourish vorschlägt, wenn man das so schmal macht, dann sieht man den gar nicht mehr.

#### A: Okay.

B: Sollte man vielleicht, weiß nicht, ob das möglich ist, aber dass man eine Warnung kriegt, dass dieser Marker im Mobilen sich außerhalb des sichtbaren Bereichs befindet.A: Okay. Habe ich notiert, danke. Wir warten nur ganz kurz, bis die Kollegin da fertig ist.

30 A: Okay, ich glaube, jetzt haben alle, ich sehe gerade, dass 4 Antworten zum

- 31 Fragebogen gekommen sind. Danke. Ich würde jetzt gerne auf die Fragen eingehen
- 32 und quasi euer Feedback dazu erfragen. Und zwar, wie ist es euch allgemein
- 33 gegangen mit der Benutzung, wie habt ihr das empfunden, das Flourish Tamplet?
- 34 B: Ja, das war eigentlich ziemlich klar und straight. Also, ich habe mich die ersten paar
- 35 Minuten, ja, schon mal zurechtfinden müssen und gefragt, was halt bzw. weil auch das
- 36 Formular ich habe mir gedacht, also ich habe immer vorschauen müssen, was sind
- 37 die nächsten Punkte, die kommentiert werden, dann war für mich auch klarer, was die
- 38 Aufgaben werden. Aber jetzt, wo ich es einmal gemacht habe, brauche ich glaube ich
- 39 keinen 2. Ablauf, dass ich weiß, wie das funktioniert und dass ich das nächste Mal
- 40 solche Marker setzen kann mit der Beschreibung.
- A: Okay, wie ist es den anderen gegangen dabei, vielleicht könnt ihr auch was dazusagen?
- 43 C: Ich habe ein bisschen gestruggelt zuerst, weil ich bin auf dem Laptop da, mit so
- 44 einem 13 Zoll Screen und dieses Settings und Safety, rechts oben und der unten,
- 45 dieser onboarding button, wie der Screen dann aufgeht, das ist bei mir nicht
- 46 gleichzeitig im Viewpoint und ich habe diesen Setting Button nicht mehr gefunden, also
- 47 ich habe jetzt nicht gewusst, wo ich die Stage dazu geben soll, da habe ich jetzt48 gespickt.
- 49 A: (lacht) Okay.

D: Ja, mir gings ähnlich, obwohl ich auf einem größeren Bildschirm bin, aber ich 50 brauchte ein paar Minuten, um rein zu kommen. Und ich fand auch der Text daneben 51 war ein bisschen klein vielleicht, neben diesem Drehrad. Und auch das GIF ganz am 52 53 Anfang zum Speichern war ein bisschen klein vielleicht, aber danach war eigentlich alles selbsterklärend. Ach so, und was ich ein bisschen so - ich weiß nicht, ob ich da 54 was falsch verstanden habe - so ein bisschen unintuitiv fand, dass man den Code da 55 reinkopieren musste zum Speichern. Und dann kein so - Also man hat zwar den 56 Speicherbutton da oben, aber der Text bleibt halt solange drinnen und dass das ein 57 bisschen das Gefühl erzeugt -58

59 A: Ja, da kann ich was dazu sagen und zwar ist das eine technische Einschränkung,

- die wir haben. wir können bei Flourish da keine Buttons und so, wir können auf die
- 61 Rechte Settings nicht zugreifen, deswegen haben wir da einen Work around
- 62 gebraucht, deswegen ist das Speichern jetzt so umständlich. Aber das ist die einzige

9 (...)

Möglichkeit, die wir gefunden haben, dass das funktioniert. Wir haben schon bei Flourish selbst angefragt, aber die antworten nicht. ich weiß nicht, ob die nicht supporten oder so viele Anfragen kriegen. Aber jedenfalls haben wir eine Lösung gebraucht und das war die einzige, die wir gefunden haben, dass man das speichern kann. Deswegen ist es so umständlich. Aber ja, es stimmt. Es ist leider kein Feedback da, ob das jetzt gespeichert worden ist oder nicht.

D: Das ist in dem Template gespeichert dann, wenn man das neu lädt -

A: Genau, wenn du die Seite neu lädst, sollten die Änderungen immer noch da sein. Wenn du es richtig gespeichert hast, quasi. Also nach dem Speichern sollte die eine Stage weg sein und das remember – das ist leider ein Work around, der ein bisschen nicht schön ist, oder nicht intuitiv ist, ja. Aber derzeit gibt's keine andere Lösung für das. Okay, vielleicht darf ich noch ein ganz kurzes Feedback –

B: Genau, was mir noch aufgefallen ist, der Button unter dem (unv.) der war mir nicht ganz klar, was der macht, ich bin dann draufgekommen, dass man da die ganzen Stages untereinander anzeigen kann, also genau. Ja.

A: Genau. Ich zeige es kurz her. Ja. Genau. Im Endeffekt, ich verschiebe den Button mal rüber, und wenn ich das jetzt aufmache, dann habe ich da eine Leiste, da ist die Reihenfolge noch falsch. Also da ist der 1., 2., 3., 4., im Reading, dann hüpft man weiter zu Acting und Analyzing. Die Idee war da, dass man das so fürs Onboarding verwenden kann und sagt, okay, ich weiß jetzt nicht, wie man interagiert und man kann quasi gleich interacting gehen und lest sich die Message durch. Oder man sagt, okay, ich möchte mich durch guiden lassen. Ich möchte mal von Anfang an schauen und lasse mich durch guiden, deswegen kann man das aus oder einblenden, wenn man das möchte. Weil manchmal man das für die Visualisierung nicht braucht. Weil man vielleicht eh nur eine Stage hat und das für die Visualisierung nicht braucht, deswegen kann man das da (unv.) aber er ist ein bisschen versteckt und es stimmt, wenn man das nicht weiß, dann weiß man das nicht.

B: Wird das dann bei den Usern auch angezeigt?

A: Ja.

B: Okay. Und was mir noch aufgefallen bei ist bei dem Löschen der Stage, passen die lcons dann nicht mehr.

- 94 A: Ja, danke, das stimmt voll. Da hast du den Bug gefunden. Voll. Ja, die rutschen
- 95 nämlich rauf und passen nicht mehr, das stimmt. Haben wir schon auf unserer Liste.
- 96 D: Was cool wäre, was mir aufgefallen ist bei den Knöpfen, die aufspringen, wenn man
- 97 da die hoch und runter Tasten benutzen könnte.
- 98 B: Lassen sich eigentlich die Reihenfolgen der Stages, lassen die sich verändern?
- 99 A: In der Version noch nicht, aber wir haben eine Version, also die Idee ist, dass man
- 100 die re-ordern kann und dass man dann im Nachhinein auch die Farben und
- 101 Reihenfolge ändern kann. Aber das ist in der Version noch nicht, weil es da noch einen
- 102 Bug gibt in den Levels. Aber gut, dass du das ansprichst, weil da wäre eh die Frage
- 103 quasi, ob das sinnvoll ist, dass man das auch anbietet. Sonst muss man es irgendwie
- 104 löschen und immer neu anlegen ist auch blöd. Darum quasi gibt's diese Reorder.
- 105 B: Okay, cool.
- 106 C: Ein Punkt, der mir also bei der Aufgabe, Stage zu löschen, war mir nicht ganz
- 107 klar, wie das möglich ist. Ich bin intuitiv auf die Settings gegangen, aber da kann ich
- 108 nur welche hinzufügen. Wenn ich dort eine Übersicht hätte, welche es gibt und ich
- 109 kann es dort löschen, dann wäre es mir logischer vorgekommen. Ich habe erst danach
- 110 gesehen, dass ich so eine Stage anklicken muss und dann daneben der Mülleimer
- 111 erscheint. Ich habe das so gemacht, dass ich den Marker, den letzten verbleibenden,
- 112 wenn du den rausnimmst, dann löscht du natürlich die gesamte Stage.
- 113 A: Okay. Danke, habe ich dokumentiert.
- 114 E: Ich hatte am MacBook das Problem, dass der Save Button nicht geladen hat. Also
- 115 ich habe einmal geglaubt, ich habe gesafed und dann habe ich neu geladen, dann war
- 116 er nicht mehr da. Also ich bin sicher, dass einige Dinge dann auch weg waren. Aber
- 117 ich habe einen sehr alten, deswegen vielleicht -
- 118 A: Welchen Browser verwendest du momentan?
- 119 E: Ich bin Firefox momentan.
- 120 A: Ja, danke, kann sein, dass das am Browser liegt, weil wir haben das bisher nur in
- 121 Chrome getestet mittlerweile. Aber danke, wir werden das dann für andere Browser
- 122 auch durchchecken und für ältere Geräte, ob dann der gleiche Fehler nochmal auftritt.

B: Ich habe mich noch gerade gefragt, wie schaut es denn aus, wenn man das published, kann man dann das Onboarding gar nicht mehr bearbeiten oder kann man das im Editor eh ganz normal bearbeiten.

A: Also man kann es ganz normal publishen, wie man das in Flourish kennt und man hat dann den Onboarding Button da, wie man das adaptiert hat (?). man muss halt so zum Template gehen, weil es gespeichert ist, kann man es von dort wieder weiter adaptieren. Also das Gepublishte kann man als (unv.) haben, das ist dann quasi gepublished, da kann man nichts verändern. Aber im Template kannst du wieder zurückgehen und dann die Sachen weiter adaptieren.

B: Und noch, gibt's einen Modus, wo man quasi das Onboarding direkt refillen kann, wenn man das anschaut. Weil so muss ich ja als User auf das Fragezeichen klicken und dann die Stages durchgehen, damit ich ins Onboarding komme.

A: Ja, genau, wenn ich z.B. Leserin bin von, dann muss ich die Stages durchgehen, damit ich zum Onboarding komme. Habe ich das richtig verstanden?

B: Genau. Genau, ich habe mich gefragt, ob es auch eine Möglichkeit, ob ihr vorgesehen habt, dass man quasi direkt in die 1. Stage irgendwie reinkommt. Also sobald das Onboarding schon da ist, genau.

A: Also dass man es quasi gleich aktiviert und gleich z.B. in Reading reinkommt und nicht die Übersicht von den 3 verschiedenen Stages hat?

B: Ja, dass überhaupt schon quasi die erste Message offen ist oder so.

A: Okay. Ja, also macht wahrscheinlich mehr Sinn, wenn man nur eine Stage hat, wenn man nur das Reading hat, dass sich das gleich öffnet, weil, wenn man 3 hat oder mehrere, dann bin ich nicht sicher, wo startet man dann. Das haben wir uns eh schon gefragt, startet man beim Reading oder Analyzing, aber ja, das wäre schon eine Idee, dass gleich die erste Stage oder die Marker sichtbar sind.

C: In dem Dialogfeld, wo das GIF dann drinnen ist oder das ganze Video, wo man den Button drücken muss, das ist jetzt nur für uns vorgesehen oder ist das auch bei Usern im Onboarding?

A: Nein, nur für die User in der Onboarding Stage, dadurch, dass das Speichern eben so überhaupt nicht intuitiv ist, haben wir gedacht, wir müssen irgendwas einbetten, dass man das richtig macht, dass man das speichern kann. Und das ist quasi nur für diese Person, die das Flourish Template verwendet. Das ist für die Leser überhaupt nicht sichtbar, das ist nur für uns gemacht, für den, der das halt gemacht hat, aber
sonst im Endeffekt ist nichts sichtbar. Genau. Genau, was sagt ihr zu dem Onboarding
Konzept an sich, also dass es quasi diese 3 Sections hat, dass man die Marker hat in
der Visualisierung und die Texte? Wie schätzt ihr das ein, ist das für die Leser hilfreich,
würdet ihr das auch einbauen in eure Visualisierungen oder ist das zu detailliert oder
so?

B: Ich kann mir schon vorstellen, dass man sowas verwendet. Die Gretchen oder 161 162 wichtige Frage, weiß ich nicht, ist, nutzen es die Leute dann auch? Wir haben jetzt keine Eyetracking Tests oder was auch immer gemacht, aber es heißt immer, wenn 163 164 irgendwelche Datenvisualisierungen sind, man kann z.B. umschalten von prozentuell auf absolut - die schauen halt hin und wenn sie das verstehen, dann schauen sie 165 vielleicht 3 Sekunden länger drauf und sonst woanders hin. Ob das wirklich in jeden 166 Tag Aktionen so breit verwendet wird. Ja, weiß ich nicht, ob man das jemals rausfinden 167 168 kann. Und davon wird aber abhängen, wie viel Zeit man drauf verwendet, das zu 169 machen und zu erklären, zu überlegen, wie führt man das am besten durch. 170 A: Ja, könnten wir wahrscheinlich vielleicht nochmal eine Schulung machen darüber. 171 Oder wenn ihr sagt, ihr habt noch was - oder noch ein Punkt, der vielleicht

- 172 ansprechend ist, wir haben vor, dass wir so eine direct Study machen, wo Leute halt
- 173 länger dieses Template verwenden und vielleicht mal probehalber eine Visualisierung
- 174 bauen, die halt vielleicht auch gerade in eine Story oder so passt. Aber ja, wenn da
- 175 Interesse besteht, können wir das sicher anschauen. Weil vielleicht, wenn man die
- 176 Hilfe dann doch verwendet, vielleicht würden die Leute dann länger bleiben. Wenn die
- 177 Leute sagen, okay, ich kenne mich nicht aus, wenn das hilft, dann bleibe ich dann
- 178 vielleicht doch länger als 3 Sekunden. Aber das ist nur eine Hypothese von mir. Weil
- 179 das doch immer mehr wird mit News Snacking und die Leute immer weiter klicken und
- 180 immer weiter scrollen und sich eben nicht wirklich was genau anschauen. Also das
- 181 Problem ist bekannt.
- B: Habt ihr Erfahrungswerte damit, wie viele Leute dann das Onboarding entdeckenund damit interagieren?
- 184 A: Im Zuge von so einer Datenstory noch nicht. Also wir haben eher also die letzten
- 185 Evaluierungen, die wir gemacht haben, waren so auf die Texte formuliert. Also wie ist
- 186 die Formulierung von den Texten, ist die kurz genug, verständlich, aber die
- 187 Anwendung in der Domäne, also für die Leserinnen, ist noch nicht getestet. Das wäre

noch so ein Punkt in unserem Forschungsprojekt, dass wir dann halt auch die Leserinnen gerne evaluieren lassen. Und sagen, okay, ist es jetzt hilfreich z.B. und vielleicht kann man dann diesen Aspekt sich auch noch anschauen. Also da ist auf alle Fälle Potenzial da. Und wir haben das Projekt jetzt auf nächstes Jahr verlängert, also wir haben noch ein Jahr länger und haben auch noch etliche Ressourcen, wenn es um Evaluierung geht, aber so Erfahrungen aus Datenstorys habe ich leider nicht.

C: Ich kann es mir schon hilfreich vorstellen. Also, wir haben halt hin und wieder solche (unv.) Prompts gemacht, wo man einfach eine Maus sieht oder sowas, dass man dort hinklicken kann. Also subjektiv, keine Ahnung, wie hilfreich das dann war. Aber genau, ich frage mich nur, ob es vielleicht ein bisschen zu versteckt ist, dass man das sofort präsentiert, könnte ich mir schon vorstellen, dass das recht hilfreich ist.

B: Dass der Button mit dem Fragezeichen vielleicht pulsiert oder so.

A: Okay, ja, ist ein guter Punkt. Ja, das habe ich noch gar nicht so gedacht, weil ich es noch nicht so im Setting beachtet habe von einer Story. Aber stimmt, vielleicht ist sogar der Button ein bisschen unsichtbar, muss man schöner darauf hinweisen. Okay, habe ich aufgeschrieben. Okay. Dann würde ich gerne die Frage stellen, gibt's im Flourish Template noch irgendwelche Funktionen, die fehlen, außer das Reordern von den Stages? Weil das habe ich schon aufgeschrieben. Gibt's noch was, wo ihr sagt, das wäre noch super, wenn man das noch einstellen kann oder so?

B: Bis auf das, was du glaube ich eh schon gesagt hast, dass man die Position z.B. von den einzelnen Punkten noch nachträglich verändern kann. Ich glaube, das wäre noch – aber das hast du eh gemeint, dass ihr das, also dass in der neueren Version schon –

A: Ja, genau, das soll es geben. Okay. Ja, gut, dann also von meiner Seite quasi alles gefragt, was ich mir vorgenommen habe. Ja, gibt's sonst noch irgendein Feedback oder Anmerkung von der Runde?

B: Wann ist die Beta-Phase abgeschlossen und das Template zu verwenden?

A: Also ich möchte das quasi dann bald mal abschließen, also wir sind eben, jetzt auch durch Evaluierungen mit ein paar anderen Journalistinnen draufgekommen, dass da noch ein paar Bugs sind, die wir weiterentwickeln. Also ich möchte das schon Anfang Oktober fertig kriegen, weil es wird daraus dann eine Publikation entstehen, die ich im Dezember einreichen möchte und da muss ich sowieso die Daten auswerten auch.

- 220 Also das muss quasi fertig sein. Sobald das fertig ist, kann ich nochmal eine E-Mail
- schicken, weil die Libraries sind auf (unv.) open source. Und das Flourish Template
- 222 wird dann natürlich auch aktualisiert, also ich möchte das quasi Ende Oktober finalisiert
- 223 haben, dass das dann quasi fertig ist.
- 224 B: Cool. Wie wird das dann integriert in Flourish?
- 225 A: Die Library?
- 226 B: Ja, dass, wenn ich halt irgendeine beliebige -

227 A: Ja, also in den Libraries, prinzipiell, setzt auf so Visualisierungspraktiken wie Plotage S, Wecharts und (unv.), also von den 3 Plattformen, also wenn man die 228 229 verwendet für die Visualisierung, dann kann man auch ein Onboarding dazu 230 generieren. Das passiert eigentlich eh automatisch, also wir nehmen die Daten aus der Visualisierung her und schreiben dazu Texte, die man als Entwickler adaptieren 231 232 kann: und im Template ist jetzt quasi die Onboarding Library drinnen, die verwenden wir da und wir haben quasi die Settings und Saving, haben wir draufgesetzt auf das 233 234 Flourish Template. Das greift quasi alles ineinander. Bisher gibt's nur die (unv.) Visualisierung in Flourish, aber wenn Sie sagen, Sie hätten jetzt gerne irgendeine 235 236 bestimmte Visualisierungstechnik, keine Ahnung, weil – dann können wir gerne den 237 Code zur Verfügung stellen oder wir können einfach die Visualisierung bauen, dann 238 können Sie das verwenden und einmal das Onboarding adaptieren. Also das ist so, 239 wie es zusammenhängt. Also wenn man da - mit den (unv.) Visualisierungen geht's leider nicht, also wenn es nur die gibt, dann kann man unser Onboarding leider nicht 240 verwenden, weil das nur bei den dreien geht. 241 B: Nein, ich habe jetzt einfach gedacht, weiß nicht, ich mache ein beliebiges Flourish 242 243 Template auf -aber da gibt's keine Add ons oder so, dass ich sage, ich mache das

- 244 Onboarding dazu. Das heißt, ich kann das nicht mit beliebigen verwenden, sondern
- 245 mit (unv.) oder mit euch in Kontakt treten.

A: Genau. Oder wenn ihr eh quasi Entwicklerinnen im Team habt, dann können wir unseren Code sharen, den wir da quasi haben und man kann dann quasi die Visualisierung austauschen und dann auch die Library. Weil die Idee ist das, also wir haben aus den Interviews auch mit anderen Journalistinnen herausgefunden, dass halt viele nicht jetzt da so programmieraffin sind, jetzt haben wir uns überlegt, wir brauchen irgendwas, das man auch mit klicken und quasi mit Visualisierung bauen kann. weil,

252 wenn man Entwickler ist, dann macht man einfach den Code auf, nimmt ein Plotage

S, eine Grafik und bettet es quasi ein. Ganz normales Programmierprojekt, sage ich mal. Aber wenn das jetzt nicht der Fall ist, dann haben wir uns eben überlegt, dass das Flourish Template guasi als Alternative zum Programmieren funktioniert. Sonst die ganzen anderen Funktionen kann man mit Flourish auch haben, man kann Onboarding ändern, Stages ändern, Farben ändern, das funktioniert alles über die Library, aber wenn man das halt nicht programmatisch quasi machen kann, dann ist quasi das Flourish Template da. Das ist quasi so die Grundidee dahinter, warum es jetzt quasi so zwei Dinge gibt.

B: Von unserer Seite gibt's sonst nichts mehr.

C: Dann bin ich gespannt, wie das ausschauen wird.

A: Ja, wenn ich darf, dann schicke ich die fertigen Sachen, wenn sie fertig sind, dann schicke ich vielleicht eine E-Mail und vielleicht können wir uns sonst nochmal was ausmachen, wenn ihr das vielleicht mal verwenden wollt, dass wir quasi für die Forschung auch was haben, dass wir quasi eine kleine Studie machen, dass ihr uns einfach Feedback gebt, wenn ihr das mal benutzt habt, weil dann vielleicht noch Sachen aufkommen, die noch nicht so gut funktionieren. Aber da können wir dann noch reden, wenn es soweit ist. Dann kann ich quasi (unv.) mal schicken, wenn er fertig ist. Super, dann sage ich danke für die Zeit, ein bisschen haben wir eh überzogen. Ich hoffe, das war jetzt kein Problem.

B: Nein, nein.

A: Okay. Passt, gut.

B: Schönen Tag noch.

C: Wiedersehen.

#### User Experience Evaluation Transcript 4

A: Ich würde gerne am Anfang noch ein paar demographische Daten erheben. Darf ich dein Alter wissen bitte?

B: 31.

- A: Was hast du für eine höchste abgeschlossene Ausbildung?
- B: Ein Bachelorstudium FH.
- A: Was hast du da genau gemacht?
- B: Das war an der FH Joanneum, Journalismus und PR.
- A: Danke. Was ist deine derzeitige Aufgabe und Rollenbeschreibung?
- B: Ich bin Kulturredakteurin für das Feuilleton der Presse.

A: Okay, Perfekt, Dann habe ich noch zwei kurze Einschätzungen von dir. Kannst du deine Erfahrung mit Flourish allgemein einschätzen? Sehr wenig Erfahrung, wenig Erfahrung, durchschnittliche Erfahrung, erfahren oder sehr erfahren.

B: Ich würde sagen erfahren. Da gibt es sicher noch eine Steigerungsstufe.

A: Verwendest du andere Visualisierungstools auch? Kannst du auch sagen, wie erfahren du bist mit diesen Tools?

B: Ja. Erfahrener. Sehr erfahren bin ich mit Data Wrapper. In den letzten Jahren habe ich eigentlich nur diese beiden. Data Wrapper und Flourish.

- A: Hast du mit Tableau schon gearbeitet?
- B: Früher einmal, aber ietzt schon lange nicht mehr.

A: Hast du bei deinen Visualisierungen, die du erstellst hast, so Lesehilfen für Leser;innen umgesetzt? Wenn ja, wie hast du das gemacht?

B: Ich habe mir einen kurzen Screenshot angeschaut von dir. Nicht in diesem Sinne. Ich habe oft in der Beschreibung oder im Subheader versucht, eine Erklärung zu geben, was man auf diesem Diagramm sieht. Mit Tooltips habe ich immer wieder gearbeitet, dass man drüberfährt und dass man nicht nur einen Zahlenwert bekommt, sondern auch einen Satz. der erklärt, was man da sieht. Bei Flourish kann man ja auch diese Stories machen. Das habe ich auch immer genützt, um eine Erzählung zu machen, in die dann Daten einfließen und Diagramme einfließen.

A: Okay, super. Dann war das der erste Erhebungsteil. Ich habe dir eh schon ganz kurz erklärt, worum es da geht. Ich würde dir gerne den Link zu unserem Flourish Template schicken, ich hoffe, du hast Zugriff zu diesem Chat in Teams.

#### B: Ja.

A: Perfekt, dann schicke ich dir den Link, Möchtest du den Link öffnen und deinen Screen bitte sharen?

B: Es lädt gerade noch. Ich habe ganz viele Tabs offen, die brauche ich aber noch, die möchte ich nicht zumachen. Siehst du es?

A: Ja, sehe ich. Dann möchte ich dich bitten, dass du es öffnest.

B: Auf Create?

- 39 A: Ja. Du bist eh erfahren mit Flourish, du weißt eh, mit der Sidebar und so, aber dass du
- 40 dich ganz kurz zurechtfindest und schaust, was da zu sehen ist. Ich möchte dich bitten, dass
- 41 du laut aussprichst, was du denkst.
- B: Es ist schon eine Weile her, dass ich mit Flourish gearbeitet habe. Ich muss schauen, ob 42
- 43 das nicht irgendwie eingefroren ist. Ich schaue mir kurz die dahinterliegenden Daten an,
- 44 dass ich weiß, was ich da habe. Okay, also da gibt es verschiedene Kategorien, die Werte 45 und das ist zugeordnet.
- 46 A: Ja, die Daten haben wir von beiden Tags, das war eine Datengeschichte von Reuters
- glaube ich, die haben die Steuern in den USA in einer hierarchischen Liste dargestellt. So 47
- 48 sind die Daten. Ganz ohne Story, nur die Visualisierung.
- 49 B: Okay, gut, Dann schaue ich mir die Lavoutoptionen durch. Die Schriftart ist mir wurscht.
- 50 Die Direction schaue ich mir an. Da scheint sich nichts zu ändern. Ah ja, genau. Da könnte ich jetzt einen Titel reinschreiben. 51
- 52 A: Genau
- B: Subtitle und Text war, was ich sonst auch verwendet habe, um dem Leser näher zu 53
- erklären, was es ist. Ich mache da irgendetwas. 54
- A: Ja. Aber du hast schon mit Flourish gearbeitet, du weißt guasi, wo man reingeht. 55
- B: Grundsätzlich schon. Ich würde nicht immer auf Anhieb alles finden, aber grundsätzlich ist 56
- 57 alles da. Bei dieser Diagrammart gibt es auch nicht wahnsinnig viele Optionen, um das
- anzupassen, weil die Größen eh automatisch vorgegeben sind. 58
- 59 A: Ja. Na passt. Dann würde ich sagen starten wir in die kurzen Aufgaben. Wir haben jetzt
- diese Trivial Visualisierung und du möchtest eine Leserhilfe mit unserem integrierten 60
- Onboarding zu dieser Trivial Visualisierung hinzufügen. Wie würdest du da vorgehen? 61
- B: Ich weiß gar nicht, was euer Tool kann. Ich denke mir, es wäre ganz klasse, wenn man 62 63 Schritt für Schritt durchaeführt wird.
- A: Genau. Im Endeffekt ist es so, du wirst es gleich sehen. Wir haben im rechten Bereich bei 64
- den General Settings, wo du gerade bist, da gibt es unten so einen Toggle-Button, der hat 65
- Add Onboarding. Wenn du den anklickst, dann erscheint so ein Fragezeichen. Schau dir das 66
- 67 einfach an. Wir haben drei Sections definiert, wo man Informationen kriegt zum Reading, wie
- liest man das, zum Interacting, wie interagiert man mit dieser Trivial Visualisierung und 68
- Analysing, was man daraus hinauslesen. Die visuellen Marker werden in diese Trivial 69
- Visualisierung eingebaut. Man kann auf diese Marker klicken und kriegt einen kurzen Text. 70
- So wird man Schritt für Schritt durchgeführt. Unser Konzept sind die visuellen Marker mit 71
- 72 diesen Tooltips. Das ist das Onboarding-Konzept an sich. Unter dem Reading-Button rechts
- gibt es noch einen kleinen Toggle-Button. Die visuellen Marker sind wie so kleine Circles 73
- 74 angeordnet, du kannst so auch durchgehen und mit den Pfeilen kannst du auch durchgehen.
- Das ist eine Art geguidedes Onboarding, wo man so durchgeht. Wenn man das nicht mag, 75
- dann geht man auf die Sections, die man braucht und schaut sich die ganzen anderen 76 77
  - Sachen an.
- B: Okay. Diesen Toggle-Button unten fand ich jetzt nicht so intuitiv, weil da auch nicht steht, 78 79 was der macht.
- 80 A: Ja, voll. Das stimmt. Darf ich dich noch bitten, magst du den Onboarding-Button von der
- Position verschieben? Und zwar auf die x-Achse auf 50 Pixel. Wie würdest du das machen? 81
- 82 B: Da habe ich es schon.

#### A: X auf 50.

#### B: Dann rückt das so hin.

A: Ja, die nächste Aufgabe hast du schon fast gemacht. Ich möchte dich bitten, dass du den Onboarding-Button öffnest. Wie viele Stages kannst du da sehen?

B: Hier jetzt drei.

A: Ja, okay.

B: Dann gibt es innerhalb dieser Stages diese Unterbuttons.

A: Genau, perfekt. Kannst du die Reading Section durchlesen und mir ganz kurz Feedback geben, was du von diesen Texten hältst. Sind die kompliziert zu lesen?

B: Es ist auf Englisch. Für meinen Zweck im Alltag wäre es auf Deutsch besser. Ich glaube, dass es für einen Leser, der nicht versiert ist, zu viele Fachbegriffe sind. Hierarchische Daten, quantitative Werte und so. Das müsste man einfacher ausdrücken, damit es leserfreundlich ist. Also wenn es sich an ein allgemeines Publikum richtet. Ich würde es praktischer machen, in dem Fall mit konkreten Beispielen. Ich würde nicht sagen, das repräsentiert das. Sondern: 200 von dem und dem sind so und so.

#### A: Okay.

B: So würde ich das machen. Ja, da wieder das Gleiche, das ist eine sehr technische und abstrakte Beschreibung. Ich glaube, dass es eine Beschreibung ist, mit der die Leute, die diese Diagramme herstellen, etwas anfangen können. Ich glaube nicht, dass diejenigen, die die Diagramme am Ende lesen sollen, sich auskennen, also wenn das ein allgemeines journalistisches Publikum. Ich gehe jetzt einfach weiter und jede Ebene tiefer hinein. Selbe Farbe. Ja. Es ist eine Beschreibung, die im Grunde erklärt, was eine Treemap ist und nicht wirklich, was diese Treemap mir zeigt. Wenn ich das jetzt nützen würde und da die Möglichkeit hätte, diese Punkte einzeln zu bearbeiten, dann würde ich versuchen, damit eine Geschichte zu erzählen, die mit diesem konkreten Datensatz zu tun hat.

A: Ja, super. Danke dir. Dann geht es zur nächsten Aufgabe, wir sind schon bei Aufgabe 5, es läuft dahin. Ich möchte dich bitten dieses Onboarding so wie du es gerade angedeutet hast, zu adaptieren. Und zwar möchte ich dich bitten, dass du zu den Stages, die du eh schon angesprochen hast, eine hinzufügst mit dem Namen "Remember". Wie würdest du das machen?

B: Ich schaue bei Adding Onboarding, ob sich da etwas tut. Ah, und jetzt geht es da oben offenbar weiter.

A: Ja, genau. Stage hinzufügen mit dem Namen "Remember" und du kannst eine billige Farbe auswählen und einfach ein Icon verwendet. Du kannst das gerne rauskopieren was unten dem Text steht.

B: Ein Icon darf ich mir suchen.

A: Genau, da öffnet sich ein Fenster. Du kannst einfach irgendeines aussuchen. Es ist ganz egal.

#### B: Dann nehmen wir das Wasser.

A: Vom Format muss es so stehen wie unterbei. Klicke das Wasser einmal an. Genau. Da kopiere das einfach raus, was in der Klasse steht.

- 125 A: Ja, dann sollte es funktionieren. Perfekt.
- 126 B: Schön, dass man da gleich Feedback bekommt.
- 127 A: Nächste Aufgabe,-
- 128 B: Ich schaue mir das noch an.
- A: Gerne. Jetzt möchtest du zu der Stage, die du gerade hinzugefügt hast, eine OnboardingMessage hinzufügen. Wie würdest du da vorgehen?
- 131 B: Add Onboarding Message. Dann würde ich schauen für welche Stage. Für die
- 132 Remember.
- 133 A: Auch vom Titel her, ich schicke dir da im Chat Informationen, die kannst du reinkopieren,
- 134 da musst du dir nichts überlegen. Ein paar Keyfacts. Genau, einfach reinkopieren und eine
- 135 beliebige Information auswählen. Genau.
- 136 B: The message was added. Okay.
- 137 A: Das kannst du gleich anschauen, ob es funktioniert hat.
- 138 B: Ja.
- 139 A: Perfekt. Super. Dann kommen wir schon zur letzten Aufgabe. In der Reading Section
- 140 möchtest du jetzt den Text ändern. In der Reading Section, Marker 3. Du hast zuvor gesagt,
- dass die Texte sehr komplex sind. Du möchtest jetzt einfach einen von den vorgegebenen
- 142 Onboarding Texten für deine Story adaptieren. Wie würdest du das machen?
- 143 B: Es ist lustig, ich habe nämlich zuerst gedacht, dass man das zuerst auch in den Settings
- 144 machen muss. Aber da kann ich nur Adding. Jetzt habe ich zufällig gesehen, indem ich
- 145 draufgeklickt habe, dass sich da ein Bearbeitungssymbol auftut.
- 146 A: Genau.
- 147 B: Ah, sollte ich da einmal speichern?
- 148 A: Kannst du später machen. Das Speichern ist ein bisschen komplizierter, weil es technisch
- 149 nicht anders geht. Zum Reading und die Nummer 3. Genau, zum Schluss schreibst du Test
- 150 oder so dazu. Du kannst irgendetwas dazuschreiben. Da musst du runterscrollen. Perfekt.
- 151 B: Okay.
- 152 A: Die nächste Aufgabe ist, dass du die Onboarding Message, die Nummer 4 da im Reading
- 153 löscht. Die brauchst du nicht, du willst nur drei. Die vorgegebenen Onboarding Stages
- 154 möchtest du bei den Onboarding Stages Interacting löschen. Wie machst du das?
- 155 B: Die Stage will ich löschen?
- 156 A: Ja, die Interacting Stage möchtest du löschen.
- 157 B: Hmm, das schaut mir radikal aus, aber gut.
- 158 A: Okay, passt.
- 159 B: Was ist, wenn ich das doch wieder brauche? Kann ich es zurückholen?
- 160 A: Nein, dann musst du es neu erstellen. Ich kann dir nachher noch etwas dazu sagen. Es
- 161 gibt ein neues Feature, das ist aber noch nicht integriert. Die nächste Aufgabe ist, dass du
- 162 die Flourish-Änderungen jetzt speicherst. Wie würdest du da vorgehen?

24 B: Nur die Klasse.

B: Da würde ich jetzt auf den Speicherbutton klicken und dann...Was? Ah, okay, jetzt verstehe ich wie euer Template funktioniert. Ihr habt da eine ausgelagerte Bearbeitung und die kopiert man dann rüber.

A: Ja, weil anders ist es technisch nicht möglich, dass man diese Änderungen speichern kann. Darum ist es ein Workaround, sagen wir so.

B: Okay.

A: Passt, also wenn du die Seite neu lädst, solltest du die Änderungen haben, die du gemacht hast.

- B: Die ganze Seite?
- A: Ja, probiere es einmal.
- B: Ah, hier steht das nochmal.

A: Wenn du das Onboarding nochmal aufmachst...Ja, genau. Passt. Das waren die Aufgaben. Danke für das Durcharbeiten.

B: Wenn ich das hier weggebe und hinzufüge, dann bleibt es auch erhalten, oder?

- A: Genau.
- B: Und hier hole ich mir die Maske.

A: Es ist nur der Onboarding Button da. Genau. Perfekt. Gut, dann kommen wir schon zu den abschließenden Fragen. Wie hast du die Benützung des Flourish Templates empfunden? Wie war das für dich?

B: Es war ein bisschen anders als ich es gewöhnt war. Es entspricht nicht dem typischen Flourish-Bearbeitungsprozess. Aber das meiste ist recht intuitiv. Was nicht ganz intuitiv ist, ist dass man das Bestellen und Bearbeiten auf zwei völligen unterschiedlichen Wegen macht. Wenn ich eine Stage hinzufügen will, dann ist sie auf einmal da. Um sie zu bearbeiten, muss ich einen anderen Weg nehmen. Aber es ist nicht schwierig und recht nahe, also wie drücke ich das aus? Man ist mit der Bearbeitung sehr nahe am Endprodukt. What you see is what you get.

A: Wie bewertest du die Einbettung vom Onboarding im Flourish-Template? Ist das nützlich für dich? Macht das Sinn? Kannst du dir vorstellen, dass du das verwendest?

B: Ja, das kann ich mir vorstellen.

A: Okay. Dann habe ich eine Frage zum Onboarding-Konzept selber. Quasi dieses Menü mit diesen Marken und so. Wie hast du die Benützung empfunden? Kannst du einschätzen, ob das deine Leser:innen verwenden können oder würden?

B: Grundsätzlich schon. Ich bin mir nicht sicher, ob es Sinn macht, das in zwei Ebenen aufzuteilen, dass man die Stages hat und dann in jeder Stage nochmal die Unterpunkte. Das ist für komplexe Grafiken. Ich glaube, dass ich für die allermeisten Grafiken versuchen würde, das in einer Ebene abzuhandeln. Wenn ich dem Leser versuche, etwas einfacher zu machen, dann will ich ihn nicht mit Kapiteln verwirren. Ich glaube, dass ich das im Alltag auf einer Ebene abhandeln würde. Ansonsten ist es schon recht cool. Vermutlich würde ich dann, wenn ich weiß, dass es diese Stages gibt und wenn ich weiß, dass die auf allen Geräten funktionieren, versuchen, noch reduzierter zu werden bei den Titeln und bei den eingebauten Texten. Da würde ich noch weniger Infos anbieten, um das in einer erzählerischen Variante in die Buttons einzubauen.

- 205 A: Okay. Super. Fehlt dir irgendetwas bei der Flourish Onboarding Geschichte? Hast du
- 206 Features, noch super wären, wenn du sie in einer Story einbaust?
- 207 B: Was ganz cool wäre, ist, wenn man mit Pfeiltasten arbeitet. Wenn man unten den Toggle-
- 208 Button aktiviert, dann hat man da die Pfeile, aber ich finde das nicht so intuitiv, dass es von
- 209 oben nach unten geordnet ist. Grundsätzlich ist die Idee, dass man sich so durchklickt, so
- 210 gelernt von allen möglichen Hilfeseiten, die es im Internet ja gibt. Wenn man eine neue App
- 211 installiert, dann hat man das ja auch so in fünf Schritten durch. Wenn man das optisch noch
- 212 mehr diesem gelernten Modus annähert, dann wäre es noch intuitiver für den Leser.
- 213 Ansonsten finde ich das mit dem Durchklicken eigentlich ganz cool. Wenn man da nur eine
- 214 Ebene hat und nicht die zwei verschiedenen, dann wäre es auch schon einfacher. Ich glaube
- 215 nicht, dass irgendein Leser checkt, dass man mit dem Button das aktivieren kann. Das ist
- 216 recht versteckt.
- 217 A: Ja, vermutlich.
- 218 B: Die Frage ist, ob man es verstecken muss, oder ob man mir als Erstellerin die Wahl
- geben kann, ob ich das haben will, oder nicht. Und dann fixiere ich, ob es das gibt oder nicht.
- A: Hast du noch ein Feedback, ob wir noch etwas verändern können oder was dir nochgefallen hat?
- 222 B: Nein, ich habe eine Frage: Wird die Schrift automatisch übernommen? Wir die Schrift
- 223 eingesetzt, die ich auch sonst in meinem Diagramm habe?
- 224 A: Ja. Würdest du die gerne adaptieren wollen?
- 225 B: Nein, ich würde mit einer Schriftart eh durchgehend kommen. Aber ich finde es cool, dass
- 226 es das automatisch anpasst. Nein, das ist eh besser. Wenn man da anfängt, noch eigene
- 227 Schriften zu definieren, das fände ich lästig.
- 228 A: Wir haben dann noch ein weiteres Feature geplant. Die Idee ist, dass man die Stages re-
- 229 ordern kann. Du hast gesagt, dass man das vielleicht nicht braucht, oder nur bei
- 230 komplexeren Visualisierungen. Also dass es in der Ansicht oben noch einen Reiter gibt mit
- 231 re-ordering, dass du die Stages re-ordern kannst und die Farben und die Icons noch im
- 232 Nachhinein ändern kannst. Das ist noch nicht in der Version, aber das ist auch geplant, dass
- 233 man es noch mehr adaptieren kann und dass man mehr die Freiheit hat.
- B: Was mir jetzt noch einfällt: Kann ich die Position verändern?
- 235 A: Derzeit musst du es löschen und neu machen. Aber du willst die Position der
- 236 bestehenden Message ändern?
- 237 B: Ja, das wäre cool. Wenn ich zum Beispiel draufkomme, dass ich da bei den Daten etwas
- 238 ändern und auf einmal liegt das nicht mehr gut, dann wäre es lästig, wenn ich alles
- 239 verändern müsste. Oder wenn ich zum Beispiel, das ist eine Budgetgrafik und im nächsten
- 240 Jahr kommt das neue Budget, dann möchte ich ja die Grafik behalten, aber die Positionen
- 241 ändern.
- 242 A: Okay, perfekt. Gibt es sonst noch Anmerkungen von dir?
- 243 B: Nein, ansonsten fällt mir überhaupt nichts ein.
- A: Ja, dann sage ich danke. Wir sind jetzt durch mit dem Fragebogen. Danke für deine Zeit.
- 245 Es ist jetzt so, du kannst es gerne verwenden, keine Ahnung, ob du es brauchen kannst. Es
- kommt dann noch eine Version raus. Ich kann dir gerne noch einen Link schicken zur
- 247 neuesten Version. Wir reichen dann auch ein Paper ein, wo die Evaluierung dabei ist und wo

es fertig verfügbar ist. Wenn du es wirklich einmal verwendest, dann wäre es spannend, wenn du einmal Feedback geben könntest.

B: Ja, gerne. Ich mache in letzter Zeit total wenig datenjournalistisch im Alltag. Aber wenn, dann gebe ich dir gerne Bescheid. Es ist eine coole Sache.

A: Super, perfekt. Danke für deine Zeit. Es war voll super, dass wir uns wieder gesehen haben. Zwar nur online, aber trotzdem.

B: Tut mir leid, dass es so lange gedauert hat, einen Termin zu finden.

A: Kein Stress. Ich war dann eh hartnäckig und habe nochmal nachgefragt. Dann sage ich danke nochmal.





# APPENDIX **F**

# Material of Design Study Biomedical R&D

#### Interview 1

A: OK. Ja, vielen, lieben Dank nochmal für die Bereitschaft für das Interview, ich hoffe, Du siehst den Interviewleitfaden.

#### B: Ja, sehe ich.

A: OK, sehr gut. Darf ich Dich mal mit einer ganz offenen Einleitungsfrage fragen: Welchen Bezug hast Du zu interaktiven Visualisierungen, wofür setzt Du diese ein?

B: Interaktive Visualisierungen benutze ich eigentlich in meinem Arbeitsleben für, sagen wir komplexe Datensätze, also mehrdimensionale Datensätze, die ich dann eben gegeneinander auftrage und daran versuche, bestimmte Entscheidungen zu treffen. Also eigentlich unterstützt mich das bei der Entscheidungsfindungen, welche Schlüsse ich dann konkret aus meinen Ergebnissen ziehe.

#### A: Hmhm.

B: Ja, also im Prinzip in unserem Business natürlich die Unterscheidung, Substanzen, die sich lohnen, weiterzubearbeiten, und Substanzen, die wir aufgrund verschiedener Counter-Datensätze eigentlich ausschließen möchten oder weniger priorisieren möchten. Also eigentlich ein Filter- und Priorisierungstool hauptsächlich.

A: Hmhm. OK. Holger, ich glaube, Du schreibst mit, oder?

C: Dann kannst Du Dich besser darauf konzentrieren, ist OK.

#### B: (lacht)

A: Genau. OK, super, vielen Dank, sehr spannende Antwort. Vielleicht ganz kurz: Welchen Educational Background?

B: Ganz kurz, ich möchte noch ganz kurz vielleicht ergänzen, eine zweiet Anwendung ist natürlich auch Präsentation von komplexen Daten.

#### A: Hmhm.

B: Also, wenn ich die Ergebnisse vorstelle, dann ist es natürlich nicht ganz interaktiv, aber ich kann ja, wenn ich zum Beispiel zwei Plots habe und eine Selektion im einen treffe, dann im anderen markieren und eben dann auch meine Auswertestrategie oder meine Gedankengänge nochmal präsentieren und anderen Teammitgliedern vorstellen.

A: Hmhm.

- B: Also Präsentationszwecke ist die zweite Antwort.
- A: Ja, alles klar.
- C: Ganz spannend.
- A: OK, danke schön, vielleicht kurz formal: Welchen Educational Background bzw. welche Aufgabe hast Du bei Dir im Unternehmen?
- B: Educational Background ist, ich habe Biochemie studiert und promoviert.
- A: Hmhm.

- 38 B: Und bin dann bei Bayer eingestiegen als (?)stock in der biochemischen Abteilung,
- 39 habe dort dann in Testentwicklung für biochemische (?)-Formate gemacht und bin
- 40 dann aber in die HTS-Truppe gewechselt.
- 41 A: Hmhm.
- 42 B: Und seit da ist, seit 11 Jahren, 12 Jahren beschäftige ich mich jetzt eben mit
- 43 Auswertung von HTS-Datensätzen und dann eben auch die Charakterisierung von den
- 44 Hits, die wir gefunden haben und da kommt dann die Mehrdimensionalität rein, weil
- 45 wir machen für die Charakterisierung natürlich zig verschiedene (?), holen uns noch
- 46 andere Informationen rein, die verschiedene Eigenschaften von Compounds dann
- 47 beschreiben, (?) Properties, Löslichkeit, molecular (?), whatever und dann hat man
- 48 eben einen multidimensionalen Datensatz.
- 49 A: Hmhm. Super, vielen Dank. Jetzt kommen ein paar sehr fragebogenähnliche
- 50 Fragen. Die erste bezieht sich auf eine Reihe von Chart-Typen, die Du hier auch kurz
- 51 siehst, ich hoffe, die Vergrößerung reicht. Bitte beurteile selbst, wie gut Du mit welchen
- 52 Visualisierungen vertraut bist, von 1 überhaupt nicht, bis 5 äußerst vertraut.
- B: Dann sage ich mal beim Bar Chart 4.
- 54 A: Ja. Line Charts?
- 55 B: Eine 5.
- 56 A: Pie?
- 57 B: Pie nehmen wir mal eine 2.
- 58 A: OK. Scatter Charts, Bubble Charts?
- 59 B: Scatter Charts ist eine 5.
- 60 A: Hmhm.
- B: Bubble Charts ist eigentlich eine 1, ehrlicherweise. (lacht)
- 62 A: OK, guter Punkt.
- B: Weil liegt eigentlich auch da dran, da wir derzeit keine guten Tools haben, um da in
- 64 so einem Scatter Plot die dritte Dimension reinzukriegen. Wir haben schon
- 65 Möglichkeiten, aber wirklich nicht gut ausgebaut und wenn wir das hätten, würde ich
- das sicher auch mehr nutzen. Anders ausgedrückt, ich kann damit umgehen, ich weiß
- 67 sie zu lesen, aber ich arbeite sehr wenig damit.
- 68 A: OK. Scher auch spannend für Triage, wäre leicht möglich in Triage.
- B: Ja, also dreidimensionale Scatter Plots finde ich super spannend, in welcher Form
   auch immer die dritte Dimension dann reinkommt.
- 71 A: Hmhm. Hmhm. OK, super. Box Plots?
- 72 B: Eine 3.
- 73 A: Violin Plots?
- 74 B: 1.

\_\_\_\_\_01

**T** 

**ب** 

**9** 

៣ភ្ចំ

- A: Radar-Sunburst Charts?
- B: Eigentlich auch eine 1.
- A: Heatmaps?
- B: 5.
- A: (?) Gramme, Baumvisualisierung?
- B: Auch eine 5.
- A: 5?
- B: Hmhm.
- A: Parallele Koordinatensysteme?
- B: 1.
- A: Treemaps?
- B: 1. (lacht) Spannend. (lacht)
- A: Sankey-Diagramme?
- B: Auch eine 1.
- A: Und am Schluss, auf der nächsten Seite haben wir noch Node Link-Netzwerk, also Netzwerkvisualisierungen als Graphen?
- B: 2.
- A: OK. Vielen Dank. Verwendest Du andere Visualisierungsarten als die, die wir hier genannt und aufgelistet haben?
- B: Nein, eigentlich nicht. Also die, meine Haupttools sind die, wo ich die 5 gesagt habe.
- A: Hmhm.
- B: Das sind die, die ich auch natürlich tagtäglich benutze.
- A: Hmhm. OK. Gut. Vielen Dank. Bitte gib an, welche Visualisierungen? (Geräusch) Entschuldigung?
- B: Alles gut. (lacht) Jetzt.

A: OK. Bitte gib an, welche Visualisierung Du bereits selbst erstellt hast, sei es jetzt für Präsentation oder Kollegenmanagement, Konferenzen und dergleichen? Also einfach Ja oder Nein, ob es schon mal selbst gemacht wurde?

- B: Ja. Ja. Ja. Ja. Auch ein Ja. Nein. Nein. Ja. Ja. Und der Rest wäre dann Nein.
- A: Hmhm.
- B: Ach, den haben wir noch. Auch ein Nein, genau, ja. Genau.
- A: Welche der bestehenden Tools verwendest Du, ich lese vor und gehe es kurz durch. Tableau?

- 108 B: Nein.
- 109 A: Spotfire?
- 110 B: Ja, aber selten.
- 111 A: Hmhm. KNIME?
- 112 B: Nein.
- 113 A: (?)?
- 114 B: Nein.
- 115 A: Excel?
- 116 B: Ja.
- 117 A: Clickview?
- 118 B: Nee.
- 119 A: SAS?
- 120 B: Nein.
- A: Inhouse Tools, also Eigenimplementierung, sei es jetzt durch Drittleister, wie Holmes oder von eurer eigenen IT?
- 123 B: Ja. Und unter other hätte ich noch Graph Pad Prism.
- 124 A: Wie schreibe ich das?
- 125 B: Graph, G-R-A-P-H.
- 126 A: Hmhm.
- 127 B: Pad. Leerzeichen. Prism, P-R-I-S-M.
- 128 A: So?
- 129 B: Ja. Prism. Ah, R-S-M. S-M.
- 130 A: Sorry.
- 131 B: Genau Graph Pad Prism. Ja, sehr gut.
- 132 A: OK.
- 133 B: Und IGOR.

134 A: OK. Gut, vielen Dank. Jetzt kommen wir zum eigentlichen Thema. Bei unserem

- 135 Forschungsprojekt geht es ja um Visualisation Onboarding, also wie können wir Leute,
- die Domain-Experten sind, aber nicht notwendigerweise immer die Visualisierung oder
- 137 die aktuelle Implementierung der Visualisierung, welcher Button macht jetzt genau
- 138 was, kennen, onboarden. Darum würde es uns jetzt interessieren, diese Tools, die Du
- verwendest, wie wurden die gelernt? Also wir haben hier Selbststudium Trial & Error,ja oder nein, bzw.
- 141 B: Ja.

#### A: Deinen Kommentar dazu?

B: Ja. Und das, würde ich sagen, umfasst den, wenn ich die so durchlese, wird das wahrscheinlich die Hauptmethode sein, um die Tools anzuwenden.

A: Hmhm. Nach Hilfe durch Kollegen oder Kolleginnen?

B: Ja.

- A: Das Lesen von Tutorials?
- B: Ja.
- A: Das Ansehen von Videos?
- B: Nee, eigentlich nicht, nee.

A: Hmhm. Und strukturierte Kurse, eine Schulung, eine LVA vielleicht noch, zurück an der Uni-Zeit?

B: Nein. Uni-Zeit kann ich mich nicht mehr erinnern, das ist zu lange her. (lacht)

A: Hmhm. OK. Das ist für mich sehr, sehr spannend, dass es on der App selbst passiert, weil das ist genau das, wo wir mit unserem Forschungsprojekt hinwollen. Wie können wir es genau?

B: Hmhm.

A: Dort an(?).

B: Also noch ein kurzer Kommentar, ich bin da eher der, der einfach erst mal wild rumklickt.

#### A: Hmhm.

B: Und einmal kuckt, was passiert. Und was ich da auch immer sehr hilfreich finde, sind Mouseovers.

#### A: Hmhm.

B: Wenn ich einfach in einem kurzen Satz schon, wenn ich irgendwo drüberfahre und mal eine Sekunde draufbleibe und ein ganz kurzer Satz kommt, was jetzt eigentlich passiert.

A: Hmhm. Sehr schön. Was ist bei dieser Methode schwierig, welche Probleme treten da auf und was funktionier gut dabei?

B: Die Probleme sind, wenn man eine sehr spezifische Fragestellung hat und diese dann lösen möchte.

### A: Hmhm.

B: Da ist man dann eigentlich doch auf Tutorials angewiesen, auf die Hilfestellung, die einem im Programm gegeben wird, und das wird dann, also die erste Methode führt dann für mich oft nicht zum Ziel, als da brauche ich dann echt Hilfe, meist wird sie ja vom Programm irgendwie gegeben und man kann sich da durchhangeln.

177 A: Hmhm.

- 178 B: Dann das wilde Herumklicken finde ich sehr spannend, auch wenn man was Neues
- 179 hat, eine neue Software hat und einfach mal probiert, was möglich ist, was die kann.
- 180 Auch wie intuitiv die Software ist.
- 181 A: Hmhm.
- 182 B: Heißt, fühle ich mich da drin wohl, habe ich den richtigen Bezug dazu und kann sie
- 183 durch, im Prinzip ist es ja Vorwissen von anderen Softwares oder wie ähnlich ist das
- 184 da dazu, wie kann ich das da wiederfinden, was ich in anderen schon gemacht habe
- 185 oder ist es was komplett Anderes?
- 186 A: Hmhm.
- 187 B: Funktioniert auch nur bedingt, ich sage mal, kommt darauf an, wie mächtig dieses
- 188 Tool ist, also wenn ich an Spotfire denke, da führt diese Methode eigentlich nicht zu
- 189 viel Erfolg.
- 190 A: Hmhm.
- B: Weil das dann doch schon sehr komplex ist und für mein Empfinden auch rechtkomplex anzuwenden.
- 193 A: Hmhm. OK. Sehr schön. Vielen Dank. Genau, welche Ressourcen, Webseiten oder
- 194 Plattformen verwendest Du, wenn Du Hilfe brauchst? Also da geht es jetzt mal hier um
- 195 das Format. Eher Videos, Bild und Text, also Wiki, Tutorials, PDFs,
- 196 Schulungsunterlagen oder interaktiv bzw. in der Applikation selbst?
- 197 B: Also fast ausschließlich Bild und Text, bei mir.
- 198 A: Hmhm.
- 199 B: Vielleicht mal ganz selten ein Video und In-App eigentlich gar nicht.
- 200 A: Sorry Holger.
- 201 C: Ich schreibe weiter. (lacht)
- 202 A: Kennst Du eine der folgenden Webseiten, ich kann sie auch kurz mal aufmachen.
- 203 Also sagen Dir diese URLS von den Namen etwas?
- B: Nein.
- A: Genau, ich zeige Dir mal kurz, worauf wir hinauswollen. Das wären Webseiten, wie 205 hier zum Beispiel, wo Diagrammtypen erklärt sind. Da kann man nochmal draufklicken, 206 mit Beispielen und eine Erklärung und dann gibt es diese anderen Webseiten, die 207 quasi von der Problem-Domain her kommen, das heißt, ich sage, ich habe numerische 208 Daten, ich will, habe eine, mehrere oder mehrere Variablen, was will ich da zeigen, 209 sind die geordnet, sind die nicht geordnet und so weiter und so fort, also das wären 210 Tools, die so über Entscheidungsbäume von der Problemstellung ausgehend, 211 Visualisierungsarten empfehlen und dieser Data Vis-Katalog ebenfalls, das heißt, ich 212 kann mich hier über die Visualisierung schlau machen, ich kann es aber auch über die 213 Problem-Domain wieder, ich will eine, zwei Dinge vergleichen, ich will die Proportionen 214 darstellen, ich will Part-to-Whole-Relations, ich will Verteilungen darstellen. Also das 215 wären Ressourcen, die man hinzufügen, hinzuziehen kann, wenn man nicht weiß, 216 217 welche Visualisierung soll ich überhaupt nehmen?

#### B: OK, habe ich noch nie, habe ich noch nie drauf gekuckt.

A: Wären die, wären die sinnvoll, diese, diese, also würdest Du so etwas verwenden oder ist für Dich die Fragestellung ohnehin, Du weißt, was Du, welche Daten Du wie visualisieren musst und das Problem liegt vielleicht, wie mache ich es im konkreten Tool?

B: Eigentlich in den meisten Fällen ja, aber es gibt sicher auch Fälle, wo ich erst mal ein bisschen rumprobieren muss, welche dieser Visualisierungen jetzt am besten passt. Und was mir da jetzt auf Anhieb sehr gut gefallen hat, ist dieser Entscheidungsbaum.

#### A: Ja.

B: Ja, der ausgeht von der Problemstellung und dann eben die Möglichkeiten aufzeigt, was man da ausprobieren könnte.

A: Hmhm. Ich kann Dir die Links gerne nachher schicken.

B: Das wäre super, ja.

A: Schicke ich aber nur Dir, weil sonst habe ich, beeinflusse ich die Umfrage für die (?).

#### B: (lacht)

A: Und Stefan. OK, welche zusätzlichen Ressourcen, und da sprechen wir jetzt auch, um die Tools oder die Visualisierungen besser bedienen zu können. Würdest Du Dir wünschen, was fehlt aus Deiner Sicht?

B: Also was ich sehr spannend fand, ist euer Kommentar von oben, mit In-App-Hilfe. Ich meine, oft ist es ja so, dass Du dann eben die Hilfedateien irgendwo dann strukturiert abgelegt hast und Dich dann da durchhangeln kannst, aber ich sage mal, In-App interaktiv. Ist eigentlich spannend, wobei mir jetzt gerade auch nicht einfällt, wie ich mir das vorstelle, aber wenn ich so ein Programm (lacht) dann irgendwie durchführt und sagt, ach mal das, mach mal das, probiere mal das, fände ich spannend.

A: Hmhm. Das passt uns jetzt sehr gut, weil genau in die Richtung geht nämlich unser Onboarding. Wann und wo, das heißt, da geht es jetzt, die Frage, wenn Du eben direkt im Tool selbst oder extern auf anderen Webseiten hast Du vorhin schon beantwortet, eher in. Und dann die Frage on demand oder willst Du das auch beim ersten Use oder immer zur Verfügung haben?

B: Aug jeden Fall on demand.

A: Hmhm.

B: Als ich bin immer sehr genervt von irgendwelchen Hilfe-Fenstern, die aufgehen und mir irgendwas sagen wollen, wo ich im Prinzip was ganz Anderes machen will. (lacht) Microsoft ist da ja ein Meister drinnen, dass sie Dir da immer irgendwelche Pop-Ups dann reingeben, di Du eigentlich gar nicht sehen willst oder gar nichts mit anfangen kannst.

A: OK. Sehr, sehr schön. Grundsätzlich vielleicht noch, bevor wir jetzt zu, Dir kurz unsere Sketches zeigen wollen und was wir unter Onboarding verstehen. Visual

- 258 Onboarding ist für uns der Prozess, die BenutzerInnen beim Lesen interpretieren und
- 259 extrahieren, von Informationen aus visuellen Darstellungen zu unterstützen. Ich
- schicke Dir jetzt mal kurz diesen Link, wir könne dann auch das Screensharing
- 261 umdrehen, wir würden Dich bitten, Dir das mal anzusehen und vielleicht einfach Think
- aloud-Session zu machen, ob das für Dich eine Art und Weise wäre, wie man so ein
- 263 Onboarding betreiben kann?
- 264 B: Hmhm.
- A: Ich habe Dir den Link jetzt hier in den Chat reingetan, das ist ein Clip-Prototyp, das
- ist keine echte App, es sind nur Screenshots von Apps, die aber mit so einem
- 267 Onboarding überlagert sind. Vielleicht, wenn Du bitte Deinen Bildschirm dann teilen
- 268 könntest, wenn Du das offen hast?
- 269 B: Ja. Der Host hat die Freigabe des Teilnehmerbildschirms deaktiviert.
- 270 A: Ach. Sorry, das muss ich kurz machen. Share screen. Jetzt müsste es gehen?
- B: Ja, funktioniert, warte ich hole es kurz rein. Könnt ihr es sehen?
- 272 A: Hmhm. Ja. Ja.
- 273 C: Jetzt sehen wir es.
- 274 A: Du wirst die Screenshots vielleicht erkennen, die sind aus dem Hidden Tool.
- 275 B: OK, ja.
- A: Und jetzt stelle Dir vor, Du arbeitest mit diesem Web-Tool, wo würdest Du Dir jetztHilfe suchen, bei den Visualisierungen?
- B: Noch mal ganz kurz draufschauen. Ich überlege gerade, also wenn ich die Plots
- 279 jetzt sehe, was ich damit (lacht) machen würde oder was ich machen will? Zum einen
- 280 natürlich interaktiv, aber das werden sie ja sowieso sein, zum Beispiel kucke ich mir
- mal die untere Ecke hier an und kucke mir mal die Ecke hier an und kucke, was die da
- 282 drüben machen.
- A: Genau, also im echten Tool wären die natürlich interaktiv, das sind jetzt nur Screenshots dahinter, aber wenn Du jetzt Hilfe brauchst, wo würdest Du diese suchen?
- B: Also intuitiv mache ich erst mal immer einen Rechtsklick, das ist wahrscheinlich einfach Microsoft biased, aber.
- 287 A: Ja.
- B: Also entweder mouseover an irgendeiner Stelle, wo ich mir jetzt, wo ich dann irgendwas machen könnte oder dann eben über einen Rechtsklick.
- A: Diese Hilfe, also die, was ist denn das für eine Farbe, Rot, Rosa, Hilfe-Icons rechts
- 291 unten jeweils in den Visualisierungen sind Dir nicht aufgefallen?
- B: Jetzt, wo Du es sagst, ehrlicherweise sind mir die erst mal, da habe ich komplett.
- 293 A: OK.
- B: Echt drübergeschaut, wobei sie ja fett und prominent da sind. (lacht)

#### C: Ja, das ist interessant. (lacht)

A: Genau, ich glaube, wir haben, also Du kannst, glaube ich, auf das Linke der beiden draufklicken.

B: Das ist aber cool, (lacht) für jemand, der den Plot nicht selber gemacht hat, also ich komme meistens ja aus der Richtung, dass ich die Plots selber mache.

#### A: Ja.

B: Und es kommt einfach auch dem geschuldet, da wir, ich sage mal, dieses, wie drücke ich es aus, ja, so ein Plattform-Tool bisher noch nicht haben, ja, oder mehrere Leute gemeinsam an so einem Projekt arbeiten und da finde ich es dann spannend, wenn ich dann wirklich in was reingehe, was jemand anderer angelegt hat, um einfach mal kurz zu sehen, was hat denn der da überhaupt gemacht?

#### A: Hmhm.

B: Aber den Anwendungsfall, den benutzen wir so ja eigentlich nicht, ich meine, Du kennst unsere Software-Plattform.

#### A: Ja.

B: Und da kommt es sehr selten vor, dass man dann sowas macht, ich meine, man reicht die Ergebnisse weiter und der eine arbeitet dann mit den Ergebnissen weiter und je nach Fall fragt er dann halt nach, was hast Du denn da genau gemacht oder wie hast Du das gemacht? Oh ja, das finde ich auch gut, das wäre nämlich jetzt wahrscheinlich auch was gewesen, wo ich gar nicht darauf geachtet hätte, wenn ich mich da drinnen bewege.

#### A: Hmhm.

B: Und wenn dann hier nochmal so ein Kasten kommt, da ist ja noch mal was, kuck mal da rein, finde ich super. OK. OK. Nee, finde ich, also die drei Sachen, die ich da jetzt gesehen habe, finde ich gut, aber ehrlicherweise, ich weiß nicht, warum ich das hier gar nicht gesehen habe, weil ich wahrscheinlich so biased bin.

#### A: Hmhm.

نت3∰4 0 333

B: Auf irgendwelche bekannten Mechanismen, wie Hilfestellung kommt, dass ich da echt nicht darauf geachtet habe.

A: Hmhm. Das heißt, Du hättest einen Rechtsklick erwartet? Rechtsklick und da Hilfe oder?

B: Ja, das ist das, was ich immer intuitiv mache. Wahrscheinlich hätte ich dann aber, wenn ich jetzt den Rechtsklick gemacht hätte und da wäre nicht Hilfe irgendwo gewesen, dann hätte ich natürlich nochmal genauer gekuckt, wo finde ich denn jetzt Hilfe.

#### A: Hmhm.

B: Mein zweiter Blick wäre wahrscheinlich jetzt oben in die Menüleiste gegangen und dann hätte ich, ich sage mal, an der Stelle war es jetzt einfach unerwartet und mein Gehirn hat es wirklich ausgeblendet.

- A: Sehr spannend, na ist gut zu wissen für uns, wo, es ist immer spannend, Benutzern 334
- über die Schulter zu sehen, weil die in der Regel ganz andere Pfade in der Applikation 335
- nehmen, als man selbst als Entwickler das vorgesehen hat. 336

#### 337 B: (lacht)

A: OK, super. Vielen Dank. Vielen Dank für, dass wir Dir über die Schulter schauen 338 339 durften. Wir wären eigentlich schon am Ende des Interviews. Die Frage ist, gibt es von deiner Seite aus noch etwas, das Du zu dem Thema hinzufügen möchtest? Einfach 340 eine offene Frage am Schluss. Noch ein Kommentar? Noch ein Input von Deiner 341

342 Seite?

B: Input weniger, ich, meine Frage wäre, wollt ihr das anwenden so als allgemeines 343 Hilfe-Tool für, in Anführungszeichen, alle Applikationen oder? 344

- A: Genau, also unser Ziel als (?) ist ein gefördertes Forschungsprojekt, auch 345 gemeinsam mit der TU Wien und der FH St. Pölten, wo auch ein bisschen, wie soll ich 346 sagen, die psychologischen Aspekte von Onboarding ergründet werden. Für uns das 347 Ziel, als Data Wish(?), ist natürlich die Library, die wir jetzt, also eine Library zu 348 schreiben, dass wir zu allen unseren Visualisierungen diese Onboarding-Hilfe 349 hinzufügen können. 350
- B: Hmhm. 351
- A: Und ich glaube auch im Triage, die nächsten Schritte, die wir machen in Triage, sind 352
- wahrscheinlich keine Development-Schritte, sondern eher, wir müssen schauen, die 353
- User an Bord zu bekommen und auch das schlägt in dieselbe Kerbe, also ja, wir wollen 354
- diese Library im Rahmen des Forschungsprojekts, die Library entwickeln, die uns das 355
- erlaubt, diese Hilfe, wie Du hier gesehen hast, wirklich in Triage und Hidden und so 356
- einzubauen. 357

B: Ja, nee, finde ich gut, weil oft ist es ja so, eine Hilfestellung, einfach eine Bibliothek 358

- von irgendwelchen Textblöcken, die du dir dann, wo du dann lesen musst und wenn 359
- du hier sowas hast, wenn der User auch das rote Fragezeichen sieht (lacht), finde ich 360
- spannend. Aber by the way, Dominik, die Elisabeth wird sich noch bei Dir melden, 361
- nächste Woche. 362
- A: OK. 363
- B: Sie ist im Urlaub, aber sie meldet sich nochmal in Bezug auf Triage, ja. 364
- A: Hmhm. Sehr gut. 365
- B: Der Mike hat nochmal nachgefragt, dass wir da ja irgendwas planen, wir haben das 366 367 jetzt mal bei uns in der Bereichsbesprechung vorgestellt.
- A: OK. 368
- B: Aber nur als kleinen Teaser, so 15 Minuten lang, und haben dann aber angekündigt, 369
- dass noch was kommt, wo wir das Ganze dann in einem größeren User-Kreis mal 370
- vorstellen und da wäre natürlich Deine Hilfe sehr willkommen. 371
- A: Hmhm. Sehr gerne, sehr, sehr gerne, habe ich eh mit Wiebke schon abgesprochen, 372
- dass wir hier sehr gerne unterstützen natürlich. 373

B: Genau, das noch offline, zu unserem eigentlichen Interview.

A: OK, super, vielen Dank. Dann schalte ich jetzt das Recording wieder aus.



#### Interview 2

#### B: Recording.

A: OK. Ihr müsstet jetzt meinen Screen sehen und zwar den Interviewleitfaden, Stefan, wir, der Dominik hat Dir, glaube ich, den Interviewleitfaden schon zugeschickt und den würden wir einfach durchgehen und die erste Frage wäre auch schon: Welchen Bezug hast Du zu interaktiven Visualisierungen bzw. wofür setzt Du sie ein?

B: Hmhm. Ja. Also interaktive Visualisierung ist immer nötig, um nach bestimmten Sachen zu suchen. Irgendwas ist schwierig, kompliziert, unverstanden und deshalb braucht man interaktive Visualisierung. Man muss den Zoom-Ausschnitt verändern, man will bestimmte Filter anwenden und man sucht eigentlich nach irgendeiner Ursache, also zum Beispiel nach bestimmten Untermengen in der, das ist die zweiet Sache, man sucht nach bestimmten Ursachen, irgendwas sieht komisch aus, in Anführungsstrichen, verhält sich komisch, wo kommt das her, das heißt also immer, man hat multidimensionale oder vieldimensionale Informationen, man kann sich nur einfach nur zweidimensionale Sachen visualisieren und aus kombinierten, zweidimensionalen Visualisierungen versucht man eben nach bestimmten versteckten Informationen zu suchen. Multidimensional ist jetzt 10 oder so, 5, ja, sowas in der Gegend, typischerweise.

A: Ja.

B: Das ist so mein Bezug.

A: Alles klar. Welchen Background hast Du, also welche Ausbildung und welche Aufgabe im Unternehmen hast Du?

B: Ausbildung bin ich klassisch Diplombiologe.

A: Hmhm.

B: Und habe dann Nebenfach Informatik gemacht.

#### A: OK.

B: Vor relativ langer Zeit (lacht), das ist mit der Informatik alles nur noch sehr Mathematik-lastig. (lacht) Nebenfach, Nebenfach, Fach Informatik ist gut, aber (lacht), also in meinem Diplom steht zum Beispiel drinnen, Nebenfach Informatik. (lacht)

A: OK und was ist Deine derzeitige Aufgabe?

B: Meine derzeitige Aufgabe ist, hmm, was ist meine derzeitige Aufgabe genau? (lacht) Bei Bayer sich um frühe Forschungsprojekt zu kümmern, mit dem Schwerpunkt Chemo-Informatik und Compound-Logistik. Das sind immer so Schlagwörter, aber ich weiß nicht, wie man es besser ausdrücken kann, ja, frühe Forschung, Screening, Schwerpunkt Chemo-Informatik, Compound-Logistik.

A: OK. Dann, hier haben wir einige Visualisierungen vorbereitet und da würde ich Dich bitten, selbst zu beurteilen, wie gut Du dich mit den einzelnen Visualisierungen auskennst. Also die Skala geht von 1, Du kennst Dich überhaupt nicht aus, bis 5, dass Du Dich sehr gut auskennst.

- 40 B: Hmhm. Hmhm. Das ist jetzt ja sehr prototypisch gemeint, nicht, was kann man mit
- 41 einem Bar Chart machen, was kann man mit einem Blind Chart machen, nicht?
- 42 A: Genau.
- B: das ist jetzt nicht, wie oft benutze ich es, sondern wüsste ich, was ich damitvisualisieren will, nicht?
- 45 A: Genau. Oder weißt Du, wie Du es liest?
- B: OK. OK. Dann fangen wir an. Bar Chart ist eine 5, habt ihr noch Histogramme, ich
   muss gestehen, ich habe jetzt nicht mehr genau?
- 48 A: Histogramme haben wir nicht.
- B: Dann erweitere das bitte noch um einen Kommentar, das für mich das Wichtigste
   Bar Chart eigentlich auch ein Histogramm ist.
- 51 A: OK.
- 52 B: In meinem, also für meine Fragen, nicht.
- 53 A: Hmhm.
- 54 A: Ein Line Chart?
- 55 B: Auch eine 5.
- 56 A: Das Pie Chart.
- 57 B: Auch eine 5.
- 58 A: Ein Scatter Chart?
- 59 B: Auch eine 5.
- 60 A: Box Plots?
- 61 B: Eine 4.
- 62 A: OK. Violin Plots?
- B: Auch eine 4.
- 64 A: Radar bzw. Sunburst Charts?
- B: Ja, das ist glaube ich auch eine 5.
- 66 A: OK. Heat Maps?
- B: Heat Map auch eine 5.
- 68 A: (?) Programme?
- 69 B: Auch eine 5.
- 70 A: Parallel Coordinates?
- 71 B: Was war jetzt genau die Frage, nach der ihr (lacht) fragt?
- A: Wie gut Du die Visualisierungen kennst, also von 1 bis überhaupt nicht.

**C** 32

Sibliot な後 knew 後 gend

#### B: Ja. Ja.

Bibliothek verfügbar. Bibliothek.

ist an der TU Wien

Dissertation

dieser 185

Originalversion

gedruckte

z5

₹6

କ୍ଷୁailୟୁoleଖ୍ରୀ pgintat

.83 S8

84

**B**ral

0 87

89

-60 -

Bigingel ver

97 Die

98

**0**99

A: Bis 5, Du kennst sie sehr gut.

B: Da würde ich eine 3 machen, weil das sehr schwierig ist, für mich, damit zu arbeiten. (lacht)

A: OK. Eine Tree Map?

B: Also für meine Daten, nicht? Schwierig für meine Daten, mit meinen Daten zu arbeiten, nicht.

- A: Eigentlich, wie gut Du die Visualisierung kennst.
- B: genau.
- A: Also, wenn Du jetzt Deine Tree Map vor Dir liegen hast.

B: Deshalb benutze ich sie theoretisch oder mal ganz selten und deshalb gebe ich da eine 3. (lacht)

A: OK, ein Sankey?

B: Tree Map würde ich auch gerne höher gehen, also eine 5, also kommt irgendwie an Dein Programm ran. Auch eine 3.

- A: Und zum Schluss ein Node-Link-Diagramm?
- B: Auch eine 3.

A: Und kennst Du noch andere Visualisierungstypen, die wir jetzt hier nicht genannt haben, die Du oft verwendest? Mit denen Du oft zu tun hast?

B: Nee, ich glaube, ihr sei ganz gut in der Liste da.

A: Sehr gut. (lacht) Gut. Dann haben wir hier noch einmal Visualisierungstypen und jetzt geht es darum, ob Du die bereits schon selbst erstellt hast?

- A: Einfach Ja oder Nein.
- approbierte ( agprevereo

A: Bar Chart, ja. Line Charts?

- B: Ja.
- A: Pie Charts?
- B: Ja.
- A: Scatter Charts?
- B: Ja.
  - A: Box Plots?
    - B: Ja.

- A: Violin Plots? 106
- B: Nein. 107
- A: Sunburst Diagramme? 108
- 109 B: Ja.
- A: Heat Maps? 110
- B: Ja. 111
- A: (?) Programme? 112
- B: Ja. 113
- A: Parallel Coordinates? 114
- B: Ah, pff, ja, also selten, nicht, ich überlege geradem was jetzt selbstständig erstellt 115
- heißt, aber, OK, ja, ja. (lacht) 116
- A: OK, Tree Maps. 117
- B: Ja. 118
- A: Sankey Diagramme? 119
- 120 B: Nein.
- A: Node Link-Diagramme. 121
- B: Ja. 122
- A: OK. Hier haben wir eine Liste von Visualisierungstools und möchten gerne wissen, 123
- ob Du sie bereist verwendet hast? 124
- B: Ja. Also Tableau, ja. 125
- A: OK. Spotfire? 126
- 127 B: Ja, auch. (?) nur mal angekuckt, nicht selber verwendet, nein.
- A: OK. Power (?). 128
- B: (?) Nein. 129
- 130 A: Excel?
- B: Ja. 131
- A: Quick View? 132
- B: Nee. 133
- A: (?) 134
- B: Nee. Inhouse Tools, ja. 135
- A: So wie Triage zum Beispiel? 136

B: Hmhm. B: Ja.

B: Ja, zum Beispiel, OK. Es gibt noch einen anderes, was ich sehr gut finde, das ist IGOR Pro, ich weiß nicht, ob das bei euch vorkommt, in der Liste?

A: Nein, also unsere Liste endet hier. (lacht)

B: Ja, OK.

A: OK. Wie, also wenn Du Visualisierungstools verwendest, wie hast Du den Umgang damit gelernt, also Selbststudium oder Hilfe durch KollegInnen, Lesen von Tutorials, Videos oder in einem Kurs?

B: Sind ja sehr subsummiert bei euch (lacht), ja.

A: Hmhm.

B: Ich glaube, es ist eigentlich überall ja.

A: OK. In dem strukturierten Kurs, mit welchen Materialien wurde da gearbeitet?

B: Also es ist zum Beispiel die Erfahrung Spotfire gewesen, da war ein Trainer, der hat Trainingsmaterial dabeigehabt und wir konnten auch mit eigenen Sachen was ausprobieren und Fragen und Antwort interaktiv mehrere Stunden, also drei Stunden oder so, so ein interaktiver Kurs, vorgeführtes Trainingsmaterial und dann anschließend interaktives Probieren mit eigenen Daten.

A: Dieses Trainingsmaterial, sind das dann Daten, die ihr mitgebracht habt?

B: Genau, genau.

A: OK. Hmhm. Welche Probleme hast Du erfahren, bei diesen verschiedenen Lernmethoden?

B: Hmm, ja, das ist auch ein bisschen schwierig zu beantworten, was da jetzt, also Probleme, ahm, heuet findet man eigentlich alles durch, ich sage jetzt mal blöd, Google, ja.

A: Aha.

B: Man findet ganz viele Tutorials, Videos, wo freiwillige Leute, wo Leute freiwillig ihr Wissen teilen.

- A: Hmhm.
- B: Suchen im Internet, nicht,
- A: OK.



B: Und wenn die Themen schwieriger werden oder das Problem schwieriger wird, ist es schwierig, den richtigen Ring zu finden (lacht).

A: OK. Gibt es auch irgendwelche Probleme, wenn Du Videos anschaust oder Probleme mit Kollegen, wenn Du Kollegen fragst, wie ein Tool funktioniert oder zum Beispiel, wenn Du versuchst, Dir das selber beizubringen, indem Du einfach mal durchklickst? Wenn es keine gibt?

B: Ja, klar gibt es Probleme, also das ist ja, hmm, ja, ist jetzt ein bisschen schwierig, wie kann ich das verbalisieren, was da passiert, ja. Der Dominik kennt das

- wahrscheinlich auch von anderen Triage-Sessions, ich habe eigentlich immer ein 174
- Gefühl dafür, wie was, oder ich habe irgendwie eine Intuition oder irgendwas, wie ich 175
- denke, wie es funktioniert. Meist hat das aber dann nichts mit dem einzelnen Tool zu 176
- tun und das ist sozusagen dieses, das (lacht), was man immer finden muss, indem 177
- man eben Erwartung und tatsächliches Verhalten irgendwie in Einklang bringen muss. 178
- 179 Das ist das Dilemma, sich selber etwas beizubringen.

#### A: OK. 180

B: (lacht) Und das gilt natürlich für alle, alle Möglichkeiten, etwas Neues zu lernen, 181

man versucht etwas mit Bekannten einzuordnen, auf dass man, mit dem man startet 182

und mit dem sich etwas Neues zu erarbeiten und das ist, glaube ich, eine relativ, ja, 183

fast schon archetypische Technik (lacht), um zu lernen. 184

A: OK, sehr gut. Dann die nächste Frage. Welche Ressourcen verwendest Du, wenn 185

- 186 Du Hilfe brauchst? Also Du hast schon angemerkt. Du verwendest Google?
- B: Ja, ich, Google heißt ja, man sucht, ja, klar, also es ist das Suchen. 187
- A: Genau. 188
- B: Hmhm. Ja. 189
- A: Verwendest Du auch In-App-Hilfen, wenn diese angeboten werden? 190
- B: Ja, ja. 191
- A: Und Videos bzw. Bild und Text, Tutorials zum Beispiel, Text-Tutorials? 192
- 193 B: Text-Tutorials ja, schreiben wir auch oft selber welche.
- A: OK. 194
- B: Also auch Selbstentwickelte sozusagen. Hmhm. Die sind nicht mitgeliefert, aber 195
- weil man immer wieder dieselbe Sache braucht oder nicht weiterweiß, schreiben sie 196
- es auf, haben wir ein WIKI zum Beispiel oder. 197
- A: Hmhm. 198

B: Eine One-Note-Sammlung, genau. Videos gibt es sehr selten, ja, und dann sind sie, 199 wenn es welche gibt, sind sie oft sehr allgemein und helfen, also jetzt mal abgesehen 200 von denen, die man on demand sucht und findet, also wenn man sagt, keine Ahnung, 201 202 Wasserfalldiagramm in Excel, dann findet man ja in Google relativ schnell was und das sind irgendwie 5-Minuten-Videos oder sowas, das ist Google, Aber wenn ihr jetzt sagt, 203 generell, generelle Videos, vielleicht gerade zum Einstieg, um die Philosophie des 204 Programms zu verstehen, aber für konkrete Fragen ist irgendwie, müsste man 205 sozusagen, findet man nicht die genaue Information in kurzer Zeit, wenn man sagt, ich 206 will zwei Minuten Video sehen und mir das, will meine Frage beantwortet haben, das 207 sind dann oft sehr lange Videos, 15 Minuten plus, und da werden nur generelle Sachen 208 gezeigt, das ist gut, wenn man überhaupt ganz bei null anfängt, aber wenn man eben 209 210 mit einem Problem schon unterwegs ist, dann sind die Videos eher schwierig.

A: OK, also nicht so hilfreich. Hmhm. OK. 211

B: Also für den generellen Einstieg sehr, sehr gut, klar, also wenn man gar nichts weiß und will erst mal kucken, was ist da los, dann finde ich ein Video gut, aber dann eben danach, das kucke ich mir zehn Minuten an, aber auch nicht länger (lacht) und dann für konkrete Sachen findet man ja sehr viele Sachen auch in Videos das einfach erklärt. Oder eben so typischerweise was findet man Computer Bild ein Excel-Problem oder irgendein Office-Problem auf einer Seite Computer Bild wird das dann in vier fünf Bullet Points erklärt, was man tun soll, also das ist sozusagen der, die andere Seiet davon Man hat ein Problem, man findet es durch Googlen, entweder man findet zwei bis fünf Minuten Video oder irgendwie ein Fünf-Punkte-Anleitung irgendwie als Text.

A: Alles klar. Dann haben wir da drei Websites vorbereitet. Kennst Du die? Also Datavizcatalogue, From Data to Viz oder Data Viz Project?

- B: Hmm, ich glaube, ich kenne Dat Viz Project oder?
- A: OK, ich kann es mal aufmachen.
- B: Ja.
- A: Verwendest Du die Seite öfters?
- B: Nicht oft, also wirklich für Ideen geben, ja.

A: OK. Und auch zum Nachlesen, wenn Du irgendeine Visualisierung nicht kennst zum Beispiel?

B: Genau, genau, ja, ja.

A: Hmhm. OK. Cool, Datavizcatalogue ist so ähnlich, da kannst Du Dir auch Visualisierungstypen aussuchen, die dann erklärt werden, mit dem Aufbau.

#### B: Hmhm.

A: Und das Dritte, From Data to Viz, da hast Du so einen Entscheidungsbaum, da kannst Du Dir aussuchen, welche Daten, mit welchen Daten Du es zu tun hast, und kannst dann eben durch verschiedene Wege Dir die Visualisierungen suchen, die zu Deinen Daten passen würden.

B: Hmhm.

- A: Findest Du das auch hilfreich?
- B: Ja. Ist, ihr fragt jetzt nicht nach dem Zweck, nicht, wann man das machen würde?

A: Würdest Du es verwenden, wenn Du eine Visualisierung anschaust oder eher, wenn Du es wahrscheinlich erstellst?

- B: Na schwierige Frage, also wenn, beim Erstellen.
- A: Hmhm.
- B: Euch fehlt natürlich noch eine Visualisierung, jetzt.
- A: OK.
  - B: Jetzt, wo wir es gerade haben.
- 248 A: Hmhm.

- 249 B: Eine Tabelle, nicht. Text.
- 250 A: Das stimmt. Das werden wir da oben dazuschreiben.
- 251 B: Also das klassische Data Grid sozusagen, ja.
- A: Hmhm. Guter Punkt. OK.
- 253 B: Und das wäre für mich so ein Einstieg auch, in diesen Entscheidungsbaum, nicht,
- 254 Du hast eine Spalte in Deinem Datenraum, OK, dann würde es vielleicht angeboten
- bekommen, Du kannst ein Histogramm machen oder Du kannst einen, ja, da bist Du
- ja schon auf deinem, in Deiner Ontologie oder Hierarchie auf bestimmte Sachen festgelegt, willst Du fitten, willst Du ein Hitsogramm einfach so machen, willst Du einen
- 258 Dichte-Plot?
- 259 A: Ja. OK, dann die nächste Frage. Welche zusätzlichen Ressourcen würdest Du Dir
- 260 wünschen, also abgesehen von den Dingen, die wir bereist gesagt haben, wie Video
- 261 oder Tutorials? Also wenn Du es Dir aussuchen könntest, was würde Dir einfallen?
- B: Ja. Eigentlich gute Beispiele, ja.
- 263 A: OK.
- 264 B: Also gute Beispiele heißt, irgendwie wissenschaftliche Veröffentlichung, Du siehst
- irgendwie, oh Mann, die haben das aber super smart jetzt visualisiert, toll. Womit
- 266 haben die das gemacht, wie haben die das gemacht, ja?
- A: Hmhm. OK. Wo und wann würdest Du Dir Hilfe erwarten, wen Du mit einem Tool
- 268 arbeitest? Also direkt im Tool selbst oder zum Beispiel?
- 269 B: Ja, klar, warum nicht F1 drücken, ja, klar. Also kontextmäßig.
- 270 A: OK.
- 271 B: Kontextabhängig.
- A: Also würdest Du sagen, Du möchtest es on demand haben oder möchtest Du, dass
- 273 die Hilfe gleich von vornherein angezeigt wird?
- 274 B: Nee, möchte ich selber sagen, ich brauche Hilfe und dann eben kontextuell.
- 275 A: OK. Gut. Unser Projekt dreht sich ja um Onboarding und jetzt haben wir hier eine kurze Definition, wie wir Onboarding verstehen. Und das wäre, 276 Visualisierungsonboarding ist der Prozess, der dem BenutzerInnen beim Lesen, 277 Interpretieren und Extrahieren von Informationen aus visuellen Darstellungen von 278 Daten unterstützt. Und wir haben einen kleinen Prototypen vorbereitet, bei dem wir 279 Dich jetzt bitten würden, dass Du den durchklickst, ich würde Dir den Link über Zoom 280 schicken und Du einfach Deine Gedanken offen aussprichst, was Du Dir denkst, wo 281 Du nicht weiterweißt zum Beispiel, wo Du Dir Hilfe erwartest und Deinen Screen sharst. 282 bitte. Ich würde den Link in den Chat posten. 283
- B: Ja, ich kucke schon, ich warte schon. (lacht)
- A: Ja, ist schon gesendet.
- 286 B: OK.

#### A: Hmhm.

#### B: Sehr ihr was?

A: Das sind, genau, wir sehen Deinen Screen.

B: Da muss ich erst mal agreen. OK, ich sehe jetzt hier zwei Scatter Plots, mit verschiedenen Achsen, verschiedene Achsendimensionen, verschiedene Achsennamen. Ich habe hier die Möglichkeit, relativ wenig zu tun, ich kann hier eigentlich nur auf das, was ich gar nicht, ach so, hier kann ich, bekomme ich Mouseover-Informationen, wo irgendwie ein Link versteckt ist. Hier sehe ich, ich kann irgendwas am Hardboard zeigen, ich kann es hier zumachen, es ist so ein bisschen das Problem, für mich jetzt eigentlich gerade im Moment, was ist Prototyp, was muss ich mir jetzt vorstellen, was sozusagen interaktiv ist und was sozusagen einfach nur deskriptiv sein soll dabei?

C: Das ist ein Screenshot, ich glaube Stefan, Du wirst es kennen, es ist von Hidden.

B: Ja. Hmhm.

C: Genau und die Annahme ist jetzt, Du bist jetzt ein Hidden-User und suchst Hilfe.

B: OK. OK.

C: Was würdest Du, also die Applikation im Hintergrund ist natürlich nur ein Screenshot.

B: Und das gehört zur Hintergrundapplikation Close und?

C: Das gehört zu dieser Sketch-Applikation, genau.

B: Alles klar, alles klar, OK. OK, also mit dem Background ist es Hidden, ich sehe zwei Plots, einmal habe ich den Run Level, Scatter Plots, mit zwei verschiedenen oder vier verschiedenen Dimensionen, Latent x und Latent y, Value, Back Value. Und ich frage mich jetzt, was kann ich machen? Typischerweise würde ich mal denken, ja, es sind die Grundfunktionen hier dabei, ich kann das Fenster kleiner, größer machen, zumachen, ich kann, habe hier eine Legende auch, das ist irgendwie, erklärt mir, es ist eine Eigenschaft s, die ich auf beiden Plots sehe. Und in der Farbe ist sozusagen alles homogen, alles ha das Gleiche, ich könnte jetzt wahrscheinlich mich fragen, warum habe ich zwei Plots? Zwei Plots heißt ja eigentlich, ich will irgendwie Zusammenhänge erklären, das verhält sich in dem einen und in dem anderen Plot mit diesen verschiedenen Eigenschaften, wenn ich hier etwas anklicke. Also intuitiv würde ich jetzt loslegen und sagen, ja, was kann ich hier machen, ich müsste irgendwas markieren können und markieren ist wahrscheinlich eine Eigenschaft, die mir bestimmte Sachen hier auch in den verschiedenen Plots zeigt. Um was weiter raus zu bekommen, was ich hier einstellen kann, wie ich die Achsen einstellen kann, wie ich zoomen kann oder so, würde ich erst mal drauflos probieren, durch Mausrad oder Steuerung oder sowas, da würde ich also erst mal kucken, was man da machen kann, wie sich das verhält. Dann gibt es natürlich sehr prominent dieses erste Fragezeichen, wahrscheinlich würde ich hier irgendwie kontextuell irgendwie Hilfe bekommen, das mache ich jetzt mal gerade. OK. Da kann ich irgendwas readen, interaktiv irgendwas machen, irgendwas analysieren und wieder das Ganze zumachen. So tendenziell hätte ich jetzt gesagt, ich würde was Interaktives machen, und da bekomme ich dann

auch die Information, was kann ich interaktiv machen, ich kriege hier nochmal erklärt, 329 oh, da oben, da kannst Du was auswählen, Du kannst wohl ein Bildschirmfoto machen, 330 was zoomen, was verschieben, was mit dem Lasso und mit dem Rechteck markieren 331 und da kriegt man nochmal hier gesagt, dass entsprechend, wenn ich irgendwo 332 rüberfahre, hier kontextuell das bekomme, das ist ja schon mal ganz gut. Das erwarte 333 334 ich eigentlich, erklärt mir, was ich da auch sehe. Dann kucke ich mir mal an, um was bei Reading geht, da bekomme ich die Information, wo kann ich was lesen (lacht). Das 335 hört sich jetzt erst mal ein bisschen trivial an, gibt aber hier noch zusätzliche 336 Hintergrundinformationen, wie viele Substanzen sehe ich hier insgesamt drauf, und 337 was ist der Hintergrund eben da am 3.6, '21 in Berlin getestet und was der Hintergrund 338 oder die Idee des Projektes ist, Similar Cinetic Curves, und irgendwie die Dimensionen 339 werden mir hier erklärt, was ich da eigentlich sehe und verstehe, finde ich auch hilfreich 340 und gut. Das hilft mir also das Verstehen. Analyzing scheint wohl was Neues. Weiteres 341 zu sein. Hmm. Das ist ein bisschen schwierig zu lese jetzt, aber Ähnlichkeit und 342 343 Eigenschaften in diesem Raum. Da habe ich ein bisschen Schwierigkeiten mit der Semantik (lacht) gerade, aber erklärt mir irgendwie, was die Dimensionen wohl hier 344 irgendwie bedeuten. OK. So, kann ich was markieren? Das ist jetzt, glaube ich, da die 345 Stelle, wo der Prototyp oder das Visualisieren hier interaktiv, glaube ich, dann zu Ende 346 347 ist oder? Kann ich noch irgendwas Anderes anklicken? C: Genau, der interaktive Prototyp beschränkt sich wirklich auf die, auf diese 348 Hilfefunktion. 349

- 350 B: OK. Alles klar.
- 351 A: Genau. Hast Du?
- 352 B: Jetzt würde ich natürlich was machen wollen (lacht), da ist natürlich für die Idee des
- 353 Prototypen hier dann das Ende. (lacht)
- 354 A: Genau, ja. Hast Du Feedback für uns und wie findest Du diese Art von Onboarding?
- 355 Würdest Du Dir das in Deinen Applikationen wünschen?
- 356 B: Ja.
- 357 A: OK. Also findest Du es besser als Text-Tutorial zum Beispiel?
- 358 B: Hmm. Nein.
- 359 A: OK. (lachen)
- 360 C: OK. Wenn Du an unsere User-Base denkst, die wir in.
- B: Also ich hätte mir, na vielleicht noch einen kleinen Kommentar, ohne Dich jetzt ganz
- 362 ausbremsen, ich könnte mir auch vorstellen, dass da auch dabei gewesen wäre. Bist
- 363 Du neu, willst Du en Video sehen? Hätte ich mir wahrscheinlich auch angekuckt.
- 364 A: Hmhm. OK.
- 365 C: Das heißt, Du würdest diese interaktive Hilfe auch verwenden, um externe Hilfe
- 366 einzubinden?
- 367 B: Ja.

C: Hmhm. Und wenn Du an unsere User-Base denkst, die wir in Triage und Hidden haben, ganz ehrlich, macht das Sinn, in diese Richtung, das in diesen Tools auch wirklich einzubauen, jetzt nicht nur auf diesen Screenshot draufzupacken, sondern wenn das fertig ist, dort auch zu integrieren oder sind die so firm mit der Applikation und mit Plots, dass die das nicht, dass die da keinen Mehrwert sehen?

B: Ja, jetzt hast Du eine Zwei-Ebenen-Frage gestellt. Also Vertrautheit mit einem Plot und Interagieren mit einem Tool.

#### C: Hmhm.

ek verfügbar. Ønek.

분麗

<u>i 37</u>0

B: Für das Interagieren mit dem Tool, glaube ich, ist so eine Hilfe, so ein Onboarding immer gut, weil dieses Tool, jedes Tool ist ja doch immer ein bisschen speziell.

#### C: Genau.

B: Und wer noch nicht mit diesem, mit dieser Implementierung dieses Plots gearbeitet hat, der hat sozusagen immer das Beginners-Dilemma, ja. Wie kann ich denn das machen, was ich machen will? Und letztendlich entscheidet das immer ich über die Akzeptanz des Tools und auch die Effizienz, mit der man es letztendlich bedienen kann. Also das Wie mache ich es mit, in dieser Implementierung, hilft glaube ich, immer.

#### C: Hmhm. OK.

B: Und so ein bisschen losgelöst davon ist fast schon wissenschaftlich, Du hast zweidimensionale Daten, wie kannst Du die am besten visualisieren, wie verhalten die sich, das ist ja schon fast so was, wie, ja, so ein Data Science-Kurs oder Data Analyst-Tutorial oder irgendwie sowas, ja, das ist ja ein bisschen losgelöst davon. Da, glaube ich, sind die meisten eher ignorant (lacht) und nicht so offen für die Investition von Zeit für Grundsätzliches. (lachen)

C: OK.

B: So wie in der typischen Arbeit, ja. (lacht)

C: Ja, die Grenze zur Guidance ist natürlich sehr, sehr fließend. OK.

B: Also das ist was, was ich oft feststelle, nicht, auch gerade, wenn wir jetzt diskutieren über was, welche Möglichkeiten habe ich mit Tool XY, und, ja, ich habe meine Daten, das mache ich immer so, warum soll ich mir noch was Anderes ankucken oder das verhält sich anders als ich das kenne, das ganze Tool ist schlecht oder (lacht), müssen wir das da nicht auch so immer machen, wie wir es immer gemacht haben? Das ist eine, glaube ich, sehr konservative, aber auch sehr weit verbreitete Sache, die man in diesem Umfeld, man arbeitet in der Firma an einem einzelnen Problem und die Visualisierung ist das Mittel zum Zweck, glaube ich, irgendwie doch prominent hat. (lacht)

#### C: OK.

A: OK, sehr interessant. Hast Du vielleicht sonst noch irgendwelche Anmerkungen oder liegt Dir noch was am Herzen?

408 A: OK. In wie fern barrierefrei?

409 B: Erst mal die Älteren sollten lesen können, ohne Vergrößerungsgläser einsetzen zu

- 410 müssen (lacht).
- 411 A: Hmhm.

412 B: Das ist das eine. Das andere ist natürlich auch die, das betrifft mich jetzt nicht selber,

- aber wir haben Kollegen, die damit zu tun haben, Farbfehlsichtigkeiten, sind, glaube
   ich, auch immer so ein bisschen unterschätzt, ja.
- 415 C: Hmhm. Guter Punkt.
- 416 A: Ja.

B: Also die typische Python Library Visualisierung ist schon ganz gut, aber hat das alle
nicht, die ist so klein, man kann die Achsen nicht lesen, die Achsen sind, na ja, jetzt
sage ich mal inhaltlich auch nicht beschriftet, de Beschriftungen sind zu klein, dass
man sie nicht lesen kann, die Punkte sind, überdecken alle oder sind so klein, dass
man sie nicht lesen kann, die Farben, wenn man mehrere Farben hat, ja, auch
schwierig, also da gibt es genug Raum nach oben (lacht). Kann man alles einstellen,
aber sozusagen die (?) sind dann meistens so, na ja, es ist ein Kompromiss, aber der

424 Kompromiss ist auch sehr groß oft. (lacht)

425 A: OK. Gibt es sonst noch Fragen oder Anmerkungen von allen Beteiligten?

B: Hmm. Ja, ich glaube, ich weiß nicht, irgendwie habe ich, glaube ich, schon eineMenge erzählt. (lacht)

- 428 A: Ja.
- 429 B: Also Visualisierung ist total wichtig, glaube ich, das ist so ein Resümee von mir und
- 430 nie ist es so gut geworden wie jetzt, aber es richtig zu benutzen, das ist immer die
- 431 Herausforderung. (lacht)
- 432 C: Hmhm.
- 433 A: OK.
- 434 B: Ja.
- 435 A: Das war es dann auch von unserer Seite. Danke jedenfalls, dass Du Dir Zeit
- 436 genommen hast, Ich glaube, wir können die Aufnahme auch schon stoppen, Dominik.
- 437 C: Hmhm. Hmhm.

B: Barrierefreie Visualisierung. (lacht)

#### Interview 3

A: Dann starten, genau und ich (?) meinen Screen. So. Ihr müsstet jetzt alle den Fragenbogen sehen. OK, dann.

#### B: Ja.

A: Ok, dann gleich zur ersten Frage: Welchen Bezug hast Du, Barbara, zu interaktiven Visualisierungen bzw. wofür setzt Du die in Deinem Alltag in der Arbeit ein?

B: Bezug? Relativ regelmäßige Verwendung für Projektplanung, Experimentplanung, Datenanalyse und, ja, also in erster Linie explorative Datenanalyse, also weniger, wie soll man sagen, weniger Figure-Preparation für Publikationen oder so, sondern wirklich, ja, explorative Datenanalyse, was ändert sich, wenn ich Parameter X ändere und welche unterschiedlichen Zelllinien haben Feature Y und wie schaut das aus im Vergleich, also eher so Fragestellungen. Aber grundsätzlich, ja, würde ich schon oder dafür sind vor allem die interaktiven Visualisierungen geeignet, weil man einfach schnell und einfach Dinge ändern kann, die vor allem als wenig oder als Command Line User mit wenig Erfahrung wesentlich länger brauchen würde.

A: Voll. OK. Hmhm. Zur nächsten Frage haben wir schon ein bisschen mitgeschrieben. Welchen educational background hast Du bzw. welche Aufgaben bei Boehringer Ingelheim (?)?

B: Hmhm. Also das ist meine, meine (?) job description sozusagen ist Target ID, also wir sind mit unserem Team sozusagen in der frühesten Phase der Medikamentenentwicklung, wobei man das in dem Fall fast noch gar nicht sagen kann, weil da sind überhaupt noch keine Substanzen oder Wirkstoffe involviert, da geht es wirklich darum, neue Targets, neue mögliche Angriffspunkte zu finden und die dann mit experimentellen Mitteln, also genetisch zum Beispiel, zu validieren und den biologischen Hintergrund da herauszufinden, um dann festzustellen oder, ja, eine Analyse abgeben zu können, ob das denn ein gutes Target wäre, gegen dass man eine Drug dann entwickeln könnte. Also das ist so der Haupthintergrund und, ja, also experimentell, wir machen eben viel, hauptsächlich Zellkulturarbeit und arbeiten da halt mit verschiedenen, ja, Knockdowns, Knockouts, molekularbiologischen Methoden aller Art, aber eben hauptsächlich in der Zellkultur.

A: Sehr cool. OK. (lacht) Gut. Die nächste Frage. Wir haben da unterschiedliche Chart-Typen vorbereitet und würden gerne wissen, wie gut Du die Visualisierungen schon kennst, wobei 1 bedeutet, Du kennst sie gar nicht und 5 bedeutet, Du kennst sie sehr gut.

B: Gibt es da noch genauere Definitionen, zum Beispiel zwischen sehr und äußerst oder so, also selber verwendet und so weiter oder?

- A: Zur Verwendung kommen wir in der nächsten Frage.
- B: Hmhm. OK. Ja.

A: Hier geht es wirklich nur darum, ob Du sie selbst interpretieren könntest.

B: OK.

- 41 A: 1 würde sagen, ja, Du hast sie schon einmal gesehen, weißt aber nicht, wie Du das
- 42 lesen könntest.
- 43 B: OK. Ja.
- 44 C: Mira, kannst Du ein bisschen reinzoomen vielleicht?
- 45 A: Sicher, ja, Entschuldigung. Ist es so besser?
- 46 B: Hmhm.
- 47 C: Ja.

48 B: OK, also ich glaube, ich würde jetzt da zumindest bei den ersten drei eigentlich mal

49 5 sagen, das sind einfach Dinge, die ich im Alltag eigentlich täglich anschaue. Da

vielleicht 4, sagen wir mal, etwas weniger häufif4, Box Plot, sagen wir mal auch 4 oder5.

- A: Hmhm. 4 bis 5 ist auch OK. (lacht)
- 53 B: OK, das kriegt dann auch 4 bis 5. Das kriegt, sagen wir mal 3, etwas weniger häufig,
- 4, 3, 2, 2, Wählerstromanalyse kriegt, sagen wir mal 3 (lachen). Ahm. Ja, das ist auch
   relativ häufig, aber sagen wir mal 3.

A: OK. Kennst Du noch andere Visualisierungen, mit denen Du häufiger zu tun hast,die da jetzt nicht aufgelistet worden sind?

- 58 B: Mal kurz nachdenken. Ich glaube, das meiste ist.
- 59 A: Haben wir alles abgedeckt? (lacht)
- 60 B: Ist schon da, ja, also, ja, wir haben so, oder Variationen davon, also so Radial Bar
- 61 Charts haben wir manchmal verwendet oder Histogramme, über, also auch mit
- 62 Overlay Density Plots, so in, alles, was in die Richtung geht.
- 63 A: OK. Hmhm.

B: Was haben wir sonst noch? Ja, das fällt eh so ähnlich unter die Bubble Charts so,
ja, das läuft bei uns halt auch oft unter Manhattan Plots irgendwie so, oder, ja, wobei
das fällt wahrscheinlich, also alle anderen kann man dann irgendwo zwischen den
Bubble Charts und Heat Maps wahrscheinlich einkategorisieren, also das,
verschiedene Ableitungen davon, mit Heatmaps mit verschiedenen Covariates und so
weiter, aber ich glaube, das fällt in den Bereich schon rein.

- A: Genau. Gut. Also, wie vorher schon gesagt, die Frage ist jetzt, zielt jetzt darauf ab,
- ob Du die jetzt selbst schon mal erstellt hast, also einfach Ja oder Nein.
- 72 B: Hmhm. Hmhm. Ja. Ja. Ja. Ja. Ja. Ja. Indirekt (lacht), ahm.
- 73 A: Ja?
- 74 B: Also sagen wir mal Ja.
- 75 A: (?) (lacht)
- B: Sagen wir mal Ja, aber wie gesagt, eher indirekt.
- 77 A: OK.

B: Ja. Ja. Nein. Nein. Das ist auch indirekt, aber Ja.
A: OK. Sehr gut. Und möchtest noch andere hinzufügen?
B: Ja, ähnlich, wie oben, wahrscheinlich Density Plots, Histogramme fallen da noch rein.
A: OK.
B: Ja, der Rest ist wahrscheinlich mehr oder weniger inkludiert da.
A: Sehr gut. Welche Visualisierungstools verwendest Du in Deiner Arbeit?
B: Ja. also Tableau gar nicht, kenne ich auch gar nicht.
A: OK.
B: Spotfire, eher passiv, zum Anschauen von anderen, also von Datensätzen, die von anderen erstellt worden sind, aber da, ja, kenne ich mich selber weniger gut aus.
A: OK.
B: Nein.
A: Nein, OK. Power (?)
B: Sehr selten und auch nicht, nicht als aktiver User, also eher so al passiver Konsument.
A: OK.
B: Ja, Excel regelmäßig.
А: ОК.
B: Ja, nein.
A: OK.
B: Nein.
A: Nein?
B: Ja, genau.
A: OK.
B: Inhouse Tools, ja.
A: Welche wären das zum Beispiel?
B: Major Lab in erster Linie und die 360.
A: Schreibt man das so?
B: Ja.
A: OK. OK. Und sonst noch andere?
B: Graphpad Prism und R.

- 110 A: So? (lacht)
- 111 B: Graphpad Prism.
- 112 A: Ah, klar. (lacht)
- 113 B: (?)
- 114 A: So.
- 115 B: Ja, genau und dann.
- 116 A: OK. Wie hast Du diese Tools, die Du verwendest, gelernt? War vieles Selbststudium
- 117 oder Trial & Error?
- 118 B: Ja, eigentlich alle aufgeführten Methoden in unterschiedlicher Gewichtung, aber.
- 119 A: Hmhm. Hast Du da auch Hilfe durch KollegInnen?
- 120 B: Ja, durchaus, also jetzt vor allem bei R.
- 121 A: OK. Liest Du gerne oder oft Tutorials?
- 122 B: Durchaus, also das trifft auch hauptsächlich auf R und Graphpad zu.
- 123 A: OK. Videos?
- B: Zum Teil, also hauptsächlich eben für, entweder für eine sehr bestimmte, konkrete
- 125 Fragestellung, wenn man, keine Ahnung, gerade in der neuesten Excel-Version die
- 126 bestimmte Funktion irgendwo nimmer findet oder so irgendwie oder im Rahmen von,
- 127 ich meine, das ist dann so ein bisschen integriert mit dem letzten Punkt, für so Kurse.
- 128 Also ich habe da einiges gerade für R in Data Camp gemacht, also.
- 129 A: Ah. OK. Hmhm.
- 130 B: Ja.
- 131 A: OK. Zu den strukturierten Kursen, mit welchen Materialien wurde da gearbeitet?
- 132 Also hauptsächlich wahrscheinlich Videos, wie Du schon gesagt hast?
- 133 B: Video, Text. Ja und es war durchaus, also ich habe auch einen R-Kurs mal auf der
- 134 Uni gemacht, während meinem PHD.
- 135 A: Ah. OK.
- 136 B: Also das war durchaus eine In-Person-Lehrveranstaltung.
- 137 A: OK. Welche Probleme sind Dir da untergekommen, wie Du diese Techniken 138 verwendet hast?
- 139 B: Beim, also für alle Methoden oder für?
- 140 A: Genau.
- 141 B: Für den strukturierten Kurs jetzt?
- 142 A: Für alles.
- 143 B: Ja, Trial & Error führt nicht immer zum gewünschten Erfolg, manchmal, wenn man,
- ja, und Ähnliches von Tutorials und Videos, also man muss halt dann, man findet nicht

immer genau das, was man tatsächlich braucht. Die Formulierung der Fragestellung ist oft ein Problem, um an wirklich an die Information zu kommen, die man möchte.

A: Hmhm.

B: Und es ist zeitaufwändig.

A: OK.

B: Hilfe durch KollegInnen, ja, nicht natürlich dann die Zeit der KollegInnen in Anspruch und.

#### A: Hmhm.

B: Haben dann halt nicht immer genau dann Zeit, wenn man gerade selber daran arbeiten möchte.

#### A: OK.

B: Und beim strukturierten Kurs, also bei dem Data Camp finde ich das sehr gut, also die machen, meiner Meinung nach, da sehr viel richtig, also guter Mix aus Videos und dann auch praktischer Anwendungen, direkt auf der Website und so, also das, für meinen persönlichen Geschmack sehr gut gemacht.

#### A: Hmhm.

B: Der Uni-Kurs damals war eher so, die erste Stunde war R kann auch als Taschenrechner verwendet werden. Die Zweite war, das sind die Basics, das sind die R-Basics und die dritte Stunde war dann, ja, jetzt machen wir eine differentielle Genanalyse mit diesen sieben Packages. Also vom Aufbau her einfach nicht ideal, bzw. die Lernkurse war dann nicht, ja, nicht dementsprechend, was die Leute können, zu dem Zeitpunkt, glaube ich, also ja.

A: Ja. OK. Alles klar. Welche Ressourcen bzw. welche Webseiten oder Plattformen verwendest Du, wenn Du Hilfe brauchst?

B: Ja, für Videos, wie schon erwähnt Data Camp und hin und wieder YouTube oder was auch immer kommt, wenn man nach einer bestimmten Frage googelt.

#### A: OK.

B: Bild und Text, ja, die eingebauten Tutorials zum Beispiel für Excel, Graphpad hat eigentlich ein sehr gutes Tutorial.

A: OK.

B: Dann gibt es, ja, die verschiedenen Foren, vor allem für R, Stack, Overflow und so weiter, die üblichen.

- A: Hmhm.
- B: Interaktiv in App, ja, also vor allem auch Data Camp in dem Fall.

A: OK. Kennst Du noch andere Ressourcen, die da nicht aufgelistet sind? Eine Google-Suche machst Du oft, hast Du gesagt, oder? 182 A: Hmhm.

183 B: Ich habe tatsächlich auch ein R-Buch in meinem Regal stehen.

184 A: Hmmm. (lachen) OK. Sehr cool. OK, wir haben jetzt drei Webseiten aufgelistet, die

- ich Dir gerne kurz zeigen würde bzw. wo Du uns dann vielleicht kurz sagst, ob Du diekennst bzw. ob Du die vielleicht sogar verwendest?
- B: Vom Namen her nicht, aber wenn man hin und wieder nach Plots sucht, kann esschon sein, dass ich da mal war.
- 189 A: (lacht) Der Data Viz Catalogue listet als erstes einmal alle Visualisierungs- oder
- 190 viele Visualisierungsarten auf, wenn man dann zum Beispiel ein Bar Chart auswählt.
- 191 B: Ja. Ja.
- 192 A: Bekommt man auch die Skripten, dann wie es ausschaut, genau. Ähnliche Charts
- 193 und Tools, mit denen man die Charts erstellen kann.
- 194 B: Hmhm.
- 195 A: So wie Beispiele dazu.
- 196 B: Könnte schon sein, dass ich da mal war, aber.
- 197 A: Ja. (lacht) OK. Dann From Data to Viz hat so einen hierarchischen Aufbau, das
- 198 heißt, Du kannst zuerst einmal auswählen, welche Daten Du hast.
- 199 B: Hmhm.
- 200 A: Wenn Du zum Beispiel numerische Daten hast, hast Du da eben so einen Baum,
- 201 den Du verfolgen kannst und dann kommen halt verschiedene Arten von
- 202 Visualisierungen, die für Deine Daten am besten geeignet sind. Genau und da kriegst
- 203 Du, glaube ich, sogar, genau, da kriegst Du auch noch zusätzliche Informationen,
- 204 wenn Du dann draufklickst.
- B: Hmhm. Also die R Graph Gallery kenne ich zum Beispiel, die da jetzt als Link gelistet
   ist, die Seite direkt jetzt nicht, aber.
- 207 A: OK. OK. Dann das Dritte ist Data to Viz, da wo, genau so wie bei der ersten einfach
- 208 die verschiedenen Arten von Charts aufgelistet werden, wenn Du dann eines anklickst,
- 209 kriegst Du wieder eine kurze Beschreibung, so wie Beispiele.
- B: Hmhm.
- 211 A: Wie Du die zum Beispiel verwenden könntest.
- B: Kommt mir jetzt auch nicht bekannt vor, zumindest vom?
- 213 A: OK. Genau, verwendet hast Du die dann wahrscheinlich auch nicht?
- B: Nicht aktiv zumindest, also ich schaue immer wieder mal, also gerade, vor allem im
- 215 Zusammenhang mit R, immer wieder einmal gesucht nach bestimmten Plots und
- 216 Beispiel-Code dafür.
- 217 A: Hmhm.

1 B: Ja.

B: Aber nicht, sozusagen nicht mit dem Zweck oder mit dem Zweck, herauszufinden, welche Visualisierung ich denn jetzt für meine Daten verwenden werde, also.

#### A: OK.

B: Sondern eher umgekehrt, ich wähle eine bestimmte Visualisierung und wie mache ich das praktisch.

A: OK. Welche zusätzlichen Ressourcen, die wir oben schon gelistet haben, würdest Du Dir wünschen, für Visualisierungen? Also was fehlt Dir? Oder gibt es irgendwas, was Dir fehlt oder das?

B: Generell in meiner täglichen Anwendung oder?

A: Genau oder auch darüber hinaus, wenn Dir etwas einfällt?

B: Also grundsätzlich glaube ich, dass die Weiten des Internets für die meisten Probleme tatsächlich ein Lösung finden oder bereitstellen, manchmal ist es eben bisschen schwierig oder, ja, man braucht einfach, glaube ich, ein bisschen Übung und das Vokabular, um nach den richtigen Dingen zu suchen, also tatsächlich auch zu wissen, wie denn diese Plots alle heißen oder, ja, das, was man halt selbervielleicht unter Plot-Typ X bezeichnet, lauft aber online unter einem anderen, geläufigeren Begriff oder so, also am ehesten vielleicht noch, ja, standardisiertes Vokabular oder so, in die Richtung, aber sonst eigentlich, glaube ich, gut versorgt.

A: OK. Wann bzw. wo würdest Du Dir gerne Hilfestellung erwarten, wenn Du zum Beispiel Deine Tools verwendest, Visualisierungen verwendest? Würdest Du die gerne on demand erhalten, also wenn Du es gerade brauchst oder on first usage, wenn Du es das erste Mal verwendest?

B: Also on demand, grundsätzlich, glaube ich, besser, und first usage ist manchmal, ja, verwendet man die simpelste Anwendung, die dann eh selbst recht intuitiv ist, kompliziert wird es meistens dann, wenn man halt irgendeine spezielle Funktion braucht und dann aber das Tutorial nimmer da ist oder nicht mehr so einfach aufrufbar ist oder so, also on demand würde ich bevorzugen. Im Tool selbst, kann sein oder eventuell auch verlinkt.

#### A: Hmhm.

B: Ja. Mit einer internen Suchfunktion, also keine große Präferenz von meiner Seite her, glaube ich.

A: OK. OK. Also wir haben uns dann im Zuge von unserem Projekt die Frage gestellt, was ist Onboarding und haben das dann so definiert: Visualization Onboarding ist der Prozess, der BenutzerInnen beim Lesen, Interpretieren und Extrahieren von Informationen aus visuellen Darstellungen von Daten unterstützt. Und wir haben dann versucht, ein Onboarding-Konzept zu erstellen und haben in dem Zuge auch einen Prototypen erstellt, den Link würde ich Dir jetzt gerne schicken.

B: Hmhm.

A: Und wenn Du dann de Screen sharen könntest und Dich einfach kurz durchklicken, schauen, ein bisschen umschauen und einfach laut Deine Gedanken aussprechen könntest, was Du Dir dazu denkst und uns ein bisschen Feedback geben könntest, ob

- 259 das für Dich vielleicht auch in Frage kommen würde, sowas? Wenn ich es schaffe,
- 260 dass ich den Chat irgendwie öffne.
- 261 B: Sonst notfalls per E-Mail?
- 262 A: Nein.
- B: Bei mir geht er.
- A: Jetzt habe ich es gefunden. (lacht) Ja, Chat. OK.
- B: OK. OK, Screen Sharing funktioniert mal oder?
- A: Hmhm. Ja. Und jetzt (?). Danke.
- B: OK. Schaue mal. OK, also das nimmt einmal mit durch den Graph und erklärt mal,
- was wir genau sehen. Das ist der Graph selbst, das ist interaktiv. Ups. OK, also die
- sind dann alle da gemeinsam in dem Button.
- 270 A: Hmhm.
- B: Und das finde ich interessant, finde ich hilfreich.
- 272 A: OK.
- 273 B: Die verschiedenen Gruppierungen.
- 274 A: Hmhm.
- 275 B: Genau, der recte, das ist nur der linke Teil, der funktioniert?
- 276 A: Ja.
- 277 B: OK, ja.
- 278 A: Findest Du sowas hilfreich, wenn es direkt im Chart ist?
- B: So lange man es wieder ausblenden kann, durchaus, ja, also ich weiß nicht, ob man
- es herumschieben kann theoretisch, funktioniert zumindest bei mir nicht.
- A: Ja, nein, im Prototypen funktioniert es nicht.
- 282 B: Ja.
- 283 A: Hmhm. OK. Also Du würdest Dir wünschen?
- B: Also, dass man zum Beispiel, keine Ahnung, wen ich jetzt gerade mir da den rechten
- 285 oberen Teil der Punkte anschauen will und die Erklärung gleichzeitig offen haben will.
- 286 A: Ja.
- 287 B: Dass man das Fenster da irgendwo herumschieben kann, sozusagen.
- 288 A: Hmhm.
- 289 B: Wie häufig das dann tatsächlich dann vorkommt, weiß ich nicht.
- 290 A: Ja.

B: Aber theoretisch. Vielleicht, ich meine, es kommt dann noch wahrscheinlich natürlich immer darauf an, was das jeweilige Experiment dann genau darstellt und in dem Fall ist es nur One Compound, aber ob man da dann, ja, Hovering, den Compound-Namen oder was dieser Compound tut oder so, in dem Fall dann, also guasi zusätzliche Informationen einblenden kann, on Hover. Und wenn man das dann auswählen würde, OK, das funktioniert jetzt wahrscheinlich im Prototyp auch nicht oder?

A: Nein, genau, das funktioniert nicht.

B: Dass man dann eventuell in dem ausgewählten Fenster de Compound-Namen dann angezeigt kriegt oder so, je nach dem.

- A: Hmhm.
- C: Darf ich mal was fragen, Barbara?
- B: Hmhm, Ja.

C: Und zwar, ist dieser, also ist das OK für Dich, dass die Hilfe, der Hilfemodus guasi jetzt die Visualisierung überlagert? Also wir haben uns gedacht, dass die, der Hilfemodus guasi den Explorationsmodus sozusagen überlagert und Du kannst jetzt nicht gleichzeitig explorieren, jedenfalls ist das mindestens im Prototypen so und die Frage ist: Würde Dich das stören? Weil Du hattest jetzt, glaube ich, zwei Mal angemerkt, dass Du gerne noch die anderen Punkte sehen möchtest, also würdest Du erwarten, dass die Hilfe und Dein Explorationsmodus gleichzeitig gehen?

B: Wäre wahrscheinlich in meinem Fall besser, weil es, also intuitiv würde ich ietzt guasi anfangen, die Daten zu explorieren und eventuell halt dann vielleicht noch einmal nachschauen wollen, Moment mal, was war denn das nochmal gleich, ohne dass ich dann quasi die explorative Analyse von Neuem anfangen muss, also wenn das parallel möglich wäre, würde ich jetzt mal das mit Ja beantworten, glaube ich.

C: Und ist es für Dich ausreichend oder OK. dass es sozusagen die Hilfe normalerweise in diesem Prototyp mal eine Visualisierung beschreibt? Das heißt, wir haben jetzt wirklich sozusagen eingegrenzt, weil Du hast den Fokus auf die (?)-Visualisierung und es werden quasi jetzt nicht sozusagen das gesamte Dashboard oder alle Visualisierungen gleichzeitig erklärt, das wäre auch Ok für Dich oder?

- B: Im Sinne von, wenn wir jetzt mehrere Plots da hätten gleichzeig oder?
- C. Genau, so wie jetzt, genau.
- B: Ja.

C: Jetzt hast Du ja links einen Plot und rechts einen Plot und dann müsstest Du halt die Hilfe vom linken Plot aufrufen und die später vom rechten Plot auch noch mal.

B: Ja, ich glaube, also das macht schon durchaus Sinn, weil sich ja dann, was weiß ich, die Bedeutung eines einzelnen Punktes auch ändern könnte.

C: Hmhm. Genau.

was die spezifische Anwendung für diesen Scatter Plot ist und was mir der genau sagt. 331

- Insofern macht das für mich schon Sinn, das auch wirklich dann spezifisch für den 332
- jeweiligen Plot zu haben. 333
- C: Hmhm. Danke. Daniela, dann kannst Du gerne noch? 334
- A: Ja. OK. Wir haben dann noch eine letzte Frage in unserem Fragebogen. 335
- B: Hmhm. 336

A: Und das ist eine ganz offene Frage. Möchtest Du noch irgendwas hinzufügen, zu 337

- unserem Interview? Möchtest Du uns noch irgendein Feedback geben, zum 338
- Prototypen oder zum Interview oder möchtest Du sonst noch irgendwas loswerden? 339

B: Ich bin schon gespannt, wie die Sachen dann in, die noch implementiert werden, 340 eventuell oder auch in anderen Tools. 341

- A: Ja. 342
- 343 B: Sonst, spontan, ja.
- A: OK. 344

B: Was ist das Dissertationsthema konkret oder was ist das Hauptziel des Projekts, 345

- sozusagen oder was erhofft ihr euch, rauszukriegen von den Interviews? 346
- A: Also Dissertationsthema ist eben Onboarding-Konzepte zu entwickeln, die quasi in 347
- App sind, so wie Du s jetzt gesehen hast und guasi, wie das gestaltet werden kann, 348
- welche Dinge müssen wir erklären in einer Visualisierung, dass zum Beispiel einer, der 349
- was mit dem Tool nicht so viel arbeitet, auch wirklich was herausfinden kann. Und wir 350
- haben eben diese drei Gruppierungen, die Du angesprochen hast, dass die super sind, 351
- also die haben wir guasi uns überlegt, wie können wir das machen, dass man den 352
- Reading-Part quasi hat, dass man das Chart dann erklärt, dann aber auch, wie man 353
- damit interagieren kann, weil die meisten Charts ja interaktiv sind. Von dem Interview 354
- einfach einmal wissen, welche Materialien werden sonst verwendet, können wir uns 355 vielleicht da noch was ableiten für unseren Prototypen, weil ja viele auch so In-App-356
- Hilfen oder halt auch nicht verwenden, einfach mal zu schauen, wie wird es in der 357
- Domäne gemacht quasi, was, wie machen as vielleicht auch andere Tools, vielleicht 358
- 359 können wir uns von denen was abschauen, also guasi so in die Richtung.
- B: Ja, vielleicht also so als, sagen wir mal, für mich zumindest, also mit Scatter Plots 360
- 361 sind die meisten Menschen, die mit Daten zu tun haben, relativ gut vertraut, insofern
- würde ich da jetzt, unter Anführungszeichen, wahrscheinlich nicht wahnsinnig viel 362
- Erklärung brauchen. Die Frage ist, ob man, wenn man, ob man mehr rauskriegt aus 363
- Leuten, wenn man ein Beispiel-Chart verwendet, das wenigen Leuten geläufig ist. 364
- 365 A: Hmhm.

366 B: Also, dass man dann einfach halt tatsächlich irgendwie ein bisschen mehr

- herumklicken und sich informieren muss, was da dann wirklich jetzt genau steht, 367
- sozusagen, und was die einzelnen Dinge für eine Bedeutung haben, also das wäre 368
- das Einzige, was mir sonst noch einfällt dazu. 369
- A: OK. Hmhm. 370

5 295 ≥ 296 0 298 ଳ 2 ଥିରୁ ti 300 ertati Sevai Dissel Sigs Sissel ŝ ⊡ 3<u>6</u>4 0 355 <sup>----</sup> 306 0367 <u>5</u> 308 10121/0 0 03<u>9</u>1 0.312 ¥313 n\_3₽4 ĕ3∰ erte 346 .≝ 3≵7 2388 <u>ठ</u>ू3₽ പ 320 ΠÈ 321 **3**22 323 **S \_**∃₹ **M**a \_ ≥ 330

verfügbar. ek. ने **2**हें।

ਚ ਤੋਂ ਤੋਂ ਤੋਂ

293

B: Und die Erklärung, was jetzt ein Scatter Plot im Allgemeinen ist, die fände ich ja auch anderswo und in diesem, in der Visualisierung hier will ich ja dann guasi wissen,

B: Ja, nein, aber wie gesagt, ich finde das, diese Visualisierungsdinge grundsätzlich sehr spannend und auch gerade in der Biologie und in den Bereichen, wo ich, ja, viel gearbeitet habe und immer noch arbeite, also große Datensätze einfach zu visualisieren auch für jemanden, der diese oder komplexeren Plots nicht unbedingt selber generieren kann oder nicht ewig viel Zeit dafür verwenden will, und trotzdem einfach aus den Daten das meiste dann rausholen kann, in einem, ja, Graphical User Interfaced Base mit bisschen Klicken, ohne wirklich Codes schreiben zu müssen oder so, das hilft bestimmt. Und ja, spart definitiv auch viel Zeit und auch einfach die Möglichkeit, dann die Plots oder die Visualisierungen irgendwie ein bisschen zu customizen, also, keine Ahnung, die signifikantesten Hits aus dem Plot jetzt dann rot einzufärben oder so, für interaktive Dinge, dass man dann auch für, ja, jetzt vielleicht nicht unbedingt für eine Publikation, aber interne Präsentation, wenn man die Daten diskutieren mag, schnell einen Graphen zur Verfügung hat, also diese Dinge. Zum Exportieren wäre dann natürlich PDF-Format super, weil dann kann man es auch wirklich für Publikationen eventuell später umformatieren und oder halt andere vektorbasierte Grafiken, die man dann in Illustrator und Co noch nachbearbeiten kann, aber das ist eher, ja, das ist dann schon weniger das Onboarding als die Weiterverwendung der Daten, vermutlich. Aber ja, von daher finde ich das sehr wichtig und glaube ich, wäre auch gesamtgesellschaftlich hilfreich, wenn Leute besser lernen, mit Daten umzugehen und Daten zu verstehen.

A: Hmhm. Voll. OK. Dann sage ich mal Danke für Deine Zeit und dass Du Dich bereiterklärt hast, mit uns das Interview zu machen.

B: Danke euch.

C: Danke. Waren sehr interessante Einblicke. Danke auch von.

#### Interview 4

A: Starten, OK, Danke schön, Dann wechsle ich jetzt, müsstet ihr jetzt meinen Browser sehen, ja, den Fragebogen.

#### B: Ja.

A: Genau, OK, Dann Manfred, dann würde ich direkt anfangen mit der ersten Frage und die ist sozusagen, das hast Du jetzt schon angefangen zu erzählen, aber jetzt vielleicht für das Recording nochmal. Welchen Bezug hast Du oder haben Sie zu interaktiven Visualisierungen und wofür setzen Sie diese Visualisierungen ein?

B: Puh, also OK, (?) ist eine Sache, andere Sachen ist, wenn man mehrere Parameter hat und versucht, dann mit Hilfe dieser Parameter Muster zu erkennen in Punktwolken. Also ich rede von, wie nennt man das, wo jeder Punkt ist eine Kombination aus zwei Parametern, da hat man so einen 2D-Plot.

#### C: Scatter Plot?

B: Scatter Plot, genau, ja, genau. Also das in weiterem Sinne, ja. Ja. Und klar.

A: OK. Und dann vielleicht noch allgemein erst mal, welchen educational background und Aufgaben haben Sie, also mikrobiologische Forschung?

B: Ja, ich bin in Biologie promoviert.

A: OK.

B: Also, ia.

A: OK. Und die Gruppenleitung ist sozusagen von dieser Forschungsgruppe, machst Du sozusagen und dann?

B: Ja, genau, genau. Also wir versuchen neue Angriffspunkte für die Medikamentenentwicklung zu finden, unter anderem habe ich manchmal mit Datenbanken und Visualisierungen aller Art zu tun. Ich meine, alles, was zum Beispiel Krebsmutationen, Häufigkeiten, Muster, Überlappung und Nicht-Überlappung und so weiter, betrifft, kenne ich halt aus den entsprechenden Software- Visualisierungen, die da angeboten werden, zum Teil im Netz, zum Teil von unseren eigenen Informatikern.

A: Hmhm. OK.

B: Oder auch sowas wie 2D-Plots, nichts, also wo man (?) Cluster in zwei Dimensionen macht, das ist auch eine typische Sache.

A: Hmhm. Genau. Da kommen wir jetzt guasi direkt bei der nächsten Frage dazu.

B: Bei (?) Plots oder was auch immer.

A: Hmhm. Also die nächste, im nächsten Abschnitt geht es darum, sozusagen, wie gut Du die Visualisierung kennst, ob Du die schon mal gesehen hast, ob Du sie vielleicht sogar selber verwendest? Und die Frage ist ietzt, hast Du Bar Charts, arbeitest, hast Du die schon mal gesehen?

B: Ja, also würde ich sagen, kommt sehr oft vor, also ist ziemlich das einfachste, bei der Nummer 5.

- A: Genau, also Nummer 5. 39
- B: Line. 40
- A: Line Charts auch? 41
- B: 5. Pie Chart. 5. Scatter Plot, Bubble Chart, ja, 5, habe ich auch schon gemacht. Box 42 Plot kenne ich auch, ist auch sehr häufig. Violin Plot auch, 5. Radar Plot ist, kenne ich 43 auch, ja, kommt vor, aber seltener, da würde ich mal eine 4 geben. Heat Map würde 44 ich auch eine 5 geben, Dendrogramme würde ich eine 4 geben, Parallel Coordinates 45 46 ist eigentlich nicht sehr schwierig zu verstehen, würde ich auch eine 5 geben. Tee Map kam bei mir nicht vor, kommt vor, aber ich habe nie damit gearbeitet, also habe ich 47 mal gesehen, aber das war schon mal. Und Sankey, das gibt es, das ist intuitiv sehr 48 leicht verständlich, weil wenn man einen Primärtumor hat und dann die Metastasen 49 und schaut, welche Anteile welche Mutationen haben, das kommt vor, aber ist 50 51 überhaupt nicht erstellt, also kann ich wahrscheinlich, wenn das das ist, was ich glaube, kann ich es wahrscheinlich verstehen und interpretieren, aber ich habe es 52 noch nie selbst erstellt. 53 A: So 3 vermutlich? 54

B: Node Link Network Plot, das habe ich sogar selber schon in (?) erstellt, da würde 55 ich sagen 5. 56

A: Hmhm, OK, Und hattest Du noch andere Visualisierungen? Du hattest gesagt, 2D 57 oder (?) Clustering oder 2D? 58

- B: Ja. ich meine, keine Ahnung sowas wie Dimension Reduction Plots, (?) oder so. 59
- A: Ja, das sind ja am Ende wahrscheinlich auch bloß Scatter Plots, die Du da ansiehst? 60
- B: Das sind eigentlich auch nur Scatter Plots, ja, die haben halt einen anderen 61 Hintergrund-. 62
- A: Genau. 63
- B: Was kommt noch vor? Na ich glaube, das war es, also mir fällt jetzt nichts ein. 64
- A: OK. Genau. Jetzt haben wir das guasi schon gemischt sozusagen erstellt, selber 65
- erstellt für Präsentationen oder Management-Ebene, ich glaube. Bar Charts hast Du 66
- gesagt, Pie Charts. 67
- B: Bar, ja, Pie, ja, Scatter, ja, Box Plot, ja. 68
- A: Box, ja. 69
- B: Violin ja. Radar nein, Heat Map ja, Dendrogramme auch, ich meine, die kommen 70
- bei den Heat Maps automatisch. Parallel Coordinates Plot, wo man links und rechts 71
- zwei verschiedene Koordinaten hat, nein, 72
- A: Genau. 73
- B: Meinst Du das? 74

A: Es sind zwei Achsen, genau, und dann dazwischen ist sozusagen immer, also hier sind die Werte aufgetragen und man hat dann eine Linie für einen Datensatz sozusagen, der die einzelnen Attribute dann verbindet.

B: Hmhm. Es ist ein, das ist ein Line Plot mit zwei Y-Achsen.

A: Genau, genau oder mehreren, kann sein.

B: Ja, also mit zwei habe ich es gemacht, mit mehr nicht. Tree Map habe ich nicht gemacht. Sankey habe ich nicht gemacht, Node Link habe ich gemacht.

A: OK. OK. Sehr schön. Danke. Dann die nächste Frage ist, welche Visualisierungstools verwenden Sie?

B: Spotfire.

A: Ja, KNIME.

B: Excel natürlich, SAS nicht. Inhouse-Tools, ahm, na ja, Inhouse-Tools, ich meine, die machen das, also sowas, wie (?).

A: Ja.

B: Sowas, wie TIF, ich weiß nicht, ob ihr das kennt. Sowas, wie Mega Lab, sowas wie 360, die habe ich alle verwendet.

A: OK. Die anderen hast Du dementsprechend nicht verwendet?

B: Die anderen habe ich nicht verwendet.

A: Nur der Vollständigkeit halber. OK.

B: Habe ich nicht verwendet. Na ja, bei (?) könnte man schreiben noch, was relativ oft verwendet wird in der Biologie, ist Graph Pad.

A: Graph Pad. So?

B: Das ist für die Dosis-Wirkungsbeziehungen einfach eine sehr praktische Software, die so ausgelegt ist, dass sie für Biologen intuitiv zugänglich ist, daher ist sie sehr weit verbreitet. Also sie geht nicht nur für Dosis-Wirkungsbeziehungen, aber dafür, wenn man die Wirkung einer Pille auf eine Zelle darstellt, in Dosisabhängigkeit, dann würde man das sehr oft verwenden. Was ich gemacht habe, ist noch, ich habe in R meine Tabellen so aufbereitet, dass ich sie dann visualisieren kann, aber ich habe sie meistens dann mit Spotfire visualisiert, weil ich mit GGPlot mich nicht einlesen wollte.

A: Dann mit R also die Visualisierung mit Spotfire hast Du gemeint?

B: Ja, genau.

A: Spotfire. Hmhm. Das klingt gut. Super, danke. Genau. Und dann, wenn Du jetzt mal zurückdenkst, vielleicht an Dein Studium oder jetzt bei Deiner Arbeit bei Boehringer Ingelheim, das ist ja nun auch schon eine Weile, wie hast Du diese Tools sozusagen gelernt, also wie wurdest Du quasi ge-onboarded zu diesen Tools und da haben wir halt mehrere Sachen vorgegeben, aber Du kannst natürlich auch gerne noch, wenn Du noch was anderes hast natürlich.

12 B: Also das Wichtigste war Selbststudium, Trial & Error.

- 113 A: Hmhm. Und genau, da hast Du einfach das Programm genommen und Dir selber.
- 114 B: Genau.
- 115 A: Das durchgenommen. Und.
- 116 B: (?), ich meine, Du hast ja dann im Programm immer die Hilfefunktion.
- 117 A: Hmhm.

B: Und dann hast Du Google offen, zur Not, und wenn Du dann irgendwo was weiter, also wenn das nicht reicht, ist eh schwierig. Die einzige Ausnahme ist R, dann habe

- ich tatsächlich (?), Book, über dings gelesen, über (?).
- 121 A: Hmhm.
- 122 B: Weil das sonst praktisch nicht intuitiv zugänglich ist, eher dass, da klappt das nicht,
- aber das ist dann Buch und da geht es halt auch, Google ist permanent offen, wenn
- 124 ich R verwende, weil ich halt keine Routine habe.
- A: Hmhm. OK. Ja. Und Google-Suche für R-Programmierung, oder, kann man so sagen?
- 127 B: Genau.
- 128 A: Genau. Fragst Du Kollegen oder hast Du Kollegen gefragt?
- B: Ja, ja, doch, also gerade, wenn ich nicht mehr weiterweiß, wenn ich bei Googlestecke, dann frage ich Kollegen.
- 131 A: OK. Bei Google nicht mehr weiterweiß, schreibe ich auf. (lacht) Genau. Und
- 132 Tutorials eigentlich dann sozusagen, von Google kommst Du dann wahrscheinlich auf
- 133 irgendwelche Tutorials, die Du dann halt durchgehst, nehme ich mal an, oder?
- 134 B: Ja.
- 135 A: Also.
- 136 B: Ja. Also ja, aber Lesen von Tutorials ist tatsächlich sehr begrenzt.
- 137 A: OK.

138 B: Also selbst bei Spotfire habe ich nie ein Tutorial gelesen. Pff. Einfach meine

- 139 Tabellen da rein (?) und schaut, was passiert und wenn ich irgendwo gesehen habe,
- dass ein Kollege was gemacht hat, was ich nicht kann, habe ich ihn gefragt, aber
   Videos habe ich für diese Dinge auch nicht verwendet und wie gesagt, einen wirklich
- 142 einen strukturierten Kurs habe ich eigentlich nicht gemacht.

A: Hmhm. OK. Genau, da fällt das sozusagen weg, weil da hat es keinen Kurs. Und
hattest Du sozusagen, also Du hattest es ja jetzt schon angesprochen, Probleme, was
halt schwierig war oder was gut funktioniert hat. Also teilweise hast Du wahrscheinlich
das Problem gehabt, dass Du Sachen nicht so gefunden hast, die Du wahrscheinlich
gesucht hast, oder?

- B: Genau, genau, genau und dann habe ich eben Kollegen gefragt, die sich auskennen, das funktioniert meistens gut. Ich, mir, also was mir manchmal Probleme
- 150 bereitet hat, zum Beispiel bei Spotfire, ist, wenn ich mich an diese ganze Oberfläche

gewöhnt habe und dann navigieren konnte, in dem Teil des Universums, wo ich mich auskannte und dann haben sie versucht, eine neue, benutzerfreundliche Version herauszubringen, wo das Ganze auf einmal versteckt war, sodass ich ewig gesucht habe, und sozusagen so eine Anfängernutzungsguidance gab, die ich überhaupt nicht wollte. Das hat mich sehr geärgert bei Spotfire und auch waren Visualisierungen zunächst möglich, und dann haben sie das eingeschränkt, dass, wo ich gewusst habe, das ging früher, geht jetzt nicht mehr.
A: Hmhm.
B: Sowas führt, ohne dass ich es kontrollieren kann, zu einer emotionalen Reaktion der negativen Art.

A: (lacht) Ja. Kann ich mir gut vorstellen (lacht), dass da problematisch ist.

B: Das passiert einfach, also das, bei R ist das Problem der Variantenreichtum, meiner Meinung nach, da gibt es verschiedene Pakete, die zum Teil den wörtlich gleichen Befehl mit was Anderem oder leicht Unterschiedlichem belegen.

#### A: Hmhm.

B: Und was mich in R nervt, (?) das Tidyverse verwende, ist, aber das hat hiermit eigentlich nichts zu tun, also das ist halt, da gibt es eine, es gibt eine Matrix und es gibt einen Data Frame und eigentlich ist alles dasselbe und dann wieder doch nicht und dann gibt es Zeilenbeschriftungen, das ist dann eine besondere, und da denke ich mir, also das hätte man sich so überlegen können, dass es zugänglicher, einfacher und nachvollziehbarer ist und das hat eben, das hat das Tidyverse dann für mich gelöst, weil dort funktioniert das so, dass ich es verstehen kann. Anyway.

#### A: Hmhm.

B: Da war das, das war, wo wir noch ein Problem mit so Visualisierungen, ja.

A: OK. Danke schön.

B: Ja.

A: Dann würde ich zur nächsten Frage gehen. Du hast jetzt schon angesprochen wahrscheinlich, so welche Ressourcen, Webseiten, Plattformen verwenden Sie, wenn Sie Hilfe brauchen? Videos, haben wir gesagt, nicht.

# B: (?), ja.

A: Ja, genau, das war eigentlich eher so, wenn Du über die Google-Suche kommst, vermutlich?

B: Genau, über Google-(?) kommt man dann meistens auf diese eine Website, also gerade beim Programmieren, wo dann steht, wer hatte schon mal dieses Problem und dann kommen sieben verschiedene Lösungen.

A: Ja.

B: Na, wie heißt denn das, da gibt es so eine.

A: Stack Overflow.

- 189 B: Stack Overflow, genau.
- 190 A: Stack Overflow, ja.
- 191 B: Stack Overflow ist eine, ja.
- 192 A: Stack Overflow.
- B: Also diese forenartigen Lösungen, die sind super. Interaktiv In App ist auch super,wenn es das gibt
- 195 A: Gibt es das bei Spotfire, nutzt Du das da zum Beispiel?
- B: Also bei Spotfire war das Problem, dass diese Foren so komplex sind, also die, das sind die Leute, die dann in Spotfire Zeug hineinprogrammieren.
- 198 A: OK.
- B: Und das tu ich ja nicht.
- A: Genau. Und interaktive InApp-Hilfe, also weil Du hast gesagt, Du liest auch die Hilfe Dokumentation oder die Hilfe der Programme?
- 202 B: Ja, genau, also vor allem bei Excel mache ich das, wenn ich irgendwie, wenn ich
- 203 einen Excel-Befehl nicht verstehe, was der tut, wenn ich irgendwelche
- 204 Zeilenumformungen mache oder so, dann lese ich nach. Bei Spotfire mache ich das
- 205 eigentlich nicht, nein, also da versuche ich, mich durchzuklicken, bei Spotfire.
- A: Hmhm. Ich habe auch Spotfire, ja, Try & Error.
- 207 B: Genau.
- A: Error. OK. Hast Du sonst noch andere Ressourcen, die Du nutzt eigentlich, mir wäre jetzt nichts aufgefallen, aber falls Dir noch was einfällt?
- B: Nein, nein, das waren hier die Kollegen halt.
- B: Nein, nein, das waren hier die
- A: Ja, genau, ja.
- B: Genau, aber ich meine, man versucht halt immer, deren Zeit zu schonen, nicht.
- A: Ja. (lacht) Das ist klar.
- B: Das ist immer das Letzte, was man macht.
- 215 A: Genau. Jetzt haben wir eine Frage sozusagen, wir haben verschiedene Webseiten
- 216 recherchiert natürlich und sind auf die folgenden Webseiten gestoßen, da wollten ich
- 217 noch fragen, ob Du die kennst, bzw. verwendest?
- B: Nein.
- 219 A: Die kennst Du alle nicht, also?
- B: Kenne ich alle nicht, nein.
- 221 A: Soll ich sie kurz aufmachen, dass Du vielleicht mal siehst, was es ist.
- 222 B: Ja.

A: Weil manche fanden das immer ganz hilfreich, Also der Data Viz oder Data Vizualisation Catalogue, der listet guasi alle Visualisierungen auf, die so bekannt sind oder die üblichen Visualisierungen und das Nette ist hier, wenn ich jetzt hier zum Beispiel, nehmen wir mal Box Plot anklicke, dann wird halt erklärt, wie sozusagen, was ein Box Plot ist, wie der aufgebaut ist und wann man das sozusagen am besten verwendet und für was man das verwenden kann, für Comparison und für Patterns und welche ähnlichen Visualisierungen gibt es dazu sozusagen und dann kommt man direkt sozusagen auf eine Erklärung und sogar einzelne Tools, womit man das generieren kann, also wenn man jetzt in R, also wie Du das zum Bespiel Box Plot generieren möchte, dann kommt man direkt zu der Library, das wäre sozusagen eine Übersicht, sozusagen, um schnell die Visualisierung zu finden, die man haben möchte. Das Zweite wäre From Data to Viz, das ist so ein Entscheidungsbaum, da heißt, hier kann man sozusagen sich durchklicken und sagen, ich habe numerische Daten und ich habe jetzt nur eine numerische Variable oder ich habe zwei, drei oder mehrere und dementsprechend legt er dann hier vor, welche möglichen Visualisierungsformen für welche Daten bzw. für welchen Anwendungsfall möglich sind und kann so die richtige Visualisierung für seinen Anwendungsfall sozusagen rausfinden. Und wenn ich jetzt auf die Visualisierung gehe, dann auch hier wieder kriege ich eine kleine Beschreibung, was das ist, was die meisten Fehler sind, die bei auftrete und auch hier wieder kriege ich halt (?)-Beispiele, für diese Visualisierungen und wie ich sie anwenden und verwenden kann.

#### B: Hmhm.

A: Und das letzte Projekt, was wir gefunden haben, ist auch wieder so die Visualisierungsform und wenn ich jetzt hier, jetzt habe ich, glaube ich, keinen Box Plot hier, um das zu vergleichen, aber wenn ich jetzt hier das Sankey Diagramm aufmache, dann habe ich auch hier wieder die Visualisierung und die haben halt mehr fokussiert auf Visualisierungen, auf Beispiele, das heißt, hier kann ich direkt sehen, wie sehen verschiedene Formen von Sankey Diagrammen aus? Wie, die sehen ja nicht immer gleich aus sozusagen oder künstlerisch teilweise schon angehaucht, dass man dann sieht sozusagen, was ist da möglich an der Stelle, genau.

#### B: Hmhm.

A: Das sind so drei Seiten, wir können Dir die Links auch gerne später noch zuschicken, dann kannst Du vielleicht später selber für Deen Arbeit auch nochmal schauen, ob da was dabei ist.

#### B: Hmhm. OK.

A: Dann trage ich mal Nein ein. Genau. Und jetzt sozusagen noch Deine Wünsche, also hast Du irgendwas, was Dir fehlt, wenn Du jetzt mit, wenn Du eine neue Visualisierung hast, wo Du sagen würdest, wow, das hätte ich mir schon immer gewünscht, dass das irgendwie erklärt wird, dass ich die richtige Visualisierung finde oder ich habe eine Visualisierung gesehen, gefunden oder jemand hat mir die gegeben und ich weiß nicht, wie ich die lesen soll, ist das vorgekommen oder was würdest Du Dir da wünschen?

B: Also es sind ja doch immer die Gleichen. Also wenn ich eine neue Visualisierung habe und ich kenne sie nicht, dann ist es ein einmaliger, kurzer Effekt und also ist sehr,

- 267 sehr selten, dass das passiert, weil der Sinn einer Visualisierung in einer biologischen
- 268 Darstellung ist, dass die anderen Biologen ja auch kapieren, was man da darstellt. Das
- 269 heißt, man kann auch selbst jetzt da nicht irgendwelche vollkommen exotischen
- 270 Visualisierungen verwenden, weil dann erfüllt das Ding seinen Zweck nicht, ja.
- 271 A: Hmhm.
- 272 B: Das heißt, eigentlich sollte man das durch Hinschauen verstehen und das ist
- eigentlich auch in 99 Prozent der Fälle so.
- A: Hmhm. OK.
- B: Wenn ich was ganz Deppertes sagen darf?
- 276 A: Sicher, ja.
- 277 B: Die meisten, bis auf GraphPad, sind Excel und Spotfire schrecklich darin,
- 278 Standardabweichungen anzuzeigen. Standardabweichungen in Excel einfügen, ist
- 279 eine 20 Klick-Operation und das sollte selbstverständlich sein, dass ein
- 280 Visualisierungsprogramm einem erlaubt, sowas anzuzeigen, also da, ja.
- A: Genau.
- B: Nicht nur Excel, ja.
- 283 A: Ach so, das ist in anderen auch so? Also in Spotfire auch oder?
- B: Ich finde es in Spotfire ähnlich, da habe ich nicht so viel mit Histogrammen oder sogearbeitet, weil Spotfire zeigt mir meistens alle Daten dann.
- A: Hmhm.
- B: Ja, anyway, das ist nur en Ärgernis. Hmhm.
- 288 A: OK. Genau und jetzt sozusagen die Frage, wo und wann wollen Sie Hilfestellung
- erhalten, also das heißt, soll das direkt im Tool, wäre das Dir am liebsten oder so, wie
- 290 Du es jetzt machst, dass Du einen Browser normalerweise öffnest und dann bei
- 291 Google suchst? Und die weitere Frage wäre sozusagen, soll das, würdest Du am
- liebsten haben, wenn man jetzt die Visualisierung hat, dass Dir das beim ersten Mal,
- 293 wenn Du das öffnest, erklärst, oder möchtest Du es lieber on demand mit
- 294 irgendwelchen Hilfe-Icons oder sowas haben, dass man dann die Erklärung bekommt?
- B: Also direkt eine gute Hilfe im Tool, finde ich gut.
- 296 A: Hmhm.

B: Also was nicht geklappt hat bei Excel, ist diese Büroklammer, ich weiß nicht, ob die,ob ihr die noch erlebt habt?

299 A: Ja.

B: Die war, die sich immer aufdrängt, das wurde, das wurde als so belehrend und

- aufdringlich befunden, also so muss es nicht sein, also was bei den Hilfen im Netz
- 302 manchmal das Problem ist, dass diese Hilfen aus Foren stammen von vor vier Jahren,
- 303 jetzt ist aber schon wieder die nächste Version und es funktioniert nicht mehr, ja. Das
- heißt, wenn es im Tool ist, dann muss es halt ajour sein, das muss wirklich stimmen,

(lacht) was da kommt und nicht veraltet sein und wenn die Bibliotheken sehr, sehr groß sind, frage ich mich, wie der Softwarehersteller das wirklich aktuell hält. Das wäre halt wichtig. On First.

A: Hmhm.

B: On first usage.

A: Das heißt.

B: Da kommt das, man startet, sie wollen zum ersten Mal einen Tree Chart machen, wollen Sie Hilfe dabei?

A: Genau, dann würde vielleicht sowas aufpoppen, wie bei dem Data Viz Catalogue, wo steht, wann verwendest Du Deine Bar Chart, Deinen Tree Chart, was auch immer, und wie, was brauchst Du für Daten sozusagen, dass man damit gleich die Einführung bekommt.

B: Ja, ich meine, ich finde das schon nicht schlecht, on demand, on first usage, ist schon, das ist OK, das kann man ja dann ignorieren, wenn man glaubt, man weiß es eh besser.

A: Hmhm.

B: Das ist zumindest ein freundliches Angebot.

A: Hmhm. So. Aber einfach überspringen können?

B: Ja.

A: Hmhm. OK. So und jetzt kommen wir noch, wir haben nicht mehr viel vor uns.

#### B: Hmhm.

A: Gleich kommt noch ein kleiner, interaktiver Teil für Dich. Die Frage, also oder wir habe uns sozusagen die Frage gestellt, was ist Vizualisation Onboarding und was ist Onboarding? I Rahmen von unserem Projekt. Und wir haben halt Vizualisation Onboarding definiert, dass Vizualisation Onboarding is the process of supporting users in reading, interpreting und expecting information from visual representation from the data. Das heißt sozusagen, dass wir schauen, wie kann man letztendlich, wie kann der User das Chart lesen, auf verschiedenen Ebenen und wie kann er damit dann sozusagen umgehen, um halt das Ganze zu verstehen. Und wir haben einen kleinen Prototypen gebaut, wo wie wir uns denken, wie das Onboarding ausschauen könnte. Das heißt, da ist noch nicht sehr viel Funktionalität. Du kannst ein bisschen durchklicken und Dir vielleicht laut Gedanken machen, sozusagen.

B: OK.

A: Wie Du da vorgehen würdest. Ich würde Dir jetzt diese URL in den Chat kopieren hier, von Zoom.

B: Hmhm.

A: Und vielleicht kannst Du Deinen Screen dann sharen, sodass wir das aufnehmen können oder sehen können, was, oder wie Du damit umgehst? Vielleicht, ich weiß nicht, ob Du das siehst oder?

- B: Ja, es geht gerade auf. Ahm. Ich share mal meinen Screen.
- 345 A: Hmhm.
- B: Wie share ich denn da?
- 347 D: Unten ist so ein grüner Button, bei mir in der.
- B: Ja, nein, ich bin schon dabei, ich frage mich nur, egal, wo ich den ganzen Bildschirm
- teile, jetzt teile ich den ganzen Bildschirm mit euch, was immer gefährlich ist. So. Und
- 350 was ist jetzt meine Aufgabe? Ich schau mir das an?
- 351 A: Moment, wir sehen Deinen Bildschirm noch nicht, Du hast wahrscheinlich noch nicht
- die Bestätigung, Du musst noch irgendwo den Bildschirm bestätigen da oder den Tab
- 353 oder das Window oder so.
- B: Hmhm. Sekunde. Ah. (?) Seht ihr was?
- A: Es lädt, ja, jetzt sehen wir Deinen Bildschirm, ja. Hmhm.
- B: Seht ihr, steht da was von Projektverbesserung des Onboardings oder sind da die Scatter Plots zu sehen?
- A: Nein, wir sehen die Scatter Plots, das schaut gut aus.
- 359 B: OK. Gut.

A: Genau, also Du siehst jetzt hier einen Ausschnitt von einer Applikation, mit
 dementsprechend zwei Visualisierungen und die Frage ist jetzt, Du möchtest jetzt
 wahrscheinlich wissen, was dieser Scatter Plot bedeutet und was man dort sieht und
 wie würdest Du jetzt vorgehen, um Hilfe zu holen oder wie würdest Du jetzt sozusagen
 (?)?

B: OK, also zunächst steht da Latent X, Latent Y, XY-Achse, sehe ich, ich nehme an,
jeder Punkt ist eine Messung und das ist die Verteilung der Messungen über diese
zwei Dimensionen, insofern glaube ich, dass ich schon verstehe, wenn ich ehrlich bin,
ich meine, es ist ja nur ein Scatter Plot. Wenn jetzt dieser Scatter Plot, zum Beispiel
jeder Punkt eine Aggregation ist und so weiter, würde ich mir erwarten, dass mir das
jemand sagt. (lacht) Weil wenn ich es nicht selber gemacht habe, woher sollte ich das
wissen, nicht. Wenn oder? Oder sehe ich das zu einfach?

- 372 A: Also jein, natürlich sind jeder Punkt eine Messung, die Frag eist natürlich jetzt,
- 373 warum, was ist zum Beispiel die, was bedeuten die Achsen in dem Fall oder und da ist 374 die Frage, wie würdest Du jetzt vorgehen, wenn Du das wissen möchtest, wenn Du Dir
- 375 Hilfe holen würdest?
- 376 B: Also ich finde, die Situation ist für mich sehr ungewöhnlich, wen ich in einem
- 377 Programm bin und das ein Visualisierungsprogramm ist und ich stoße auf so etwas,
- 378 wie soll das gehen, da müsste, also entweder ich habe es gemacht, dann weiß ich,
- 379 was die Achsen sind oder ich habe es nicht gemacht, dann muss mir das jemand
- 380 erklären. Wenn diese Visualisierung oder?

381 A: Ja. Jetzt nehmen wir aber mal an, also wer würde Dir das erklären, in dem Fall?

C

verfügbar. ek. ₩352

분麗

B: Na der, der sie gemacht hat, also wenn ich es nicht gemacht habe und jemand zeigt mir das, dann sollte die Achsenbeschriftung so erweitern, dass ich verstehe, was das ist. Wenn ich jetzt, wenn aber jetzt die Situation wäre, ich werde in dieses Programm geworfen, dann hätte ich mir gedacht, aha, da liegt irgendeine Tabelle drunter, dann würde ich anfangen, nach dieser Tabelle zu suchen.

#### A: Hmhm.

B: Das heißt, ich würde mich mal durch die Header da klicken, auf Essay oder Run oder Plate, oder sonst was, und schauen, was da an Datengrundlagen zu finden ist.

#### A: Hmhm.

B: Oder ich würde auf diesen Pfeil da klicken und schauen, ob das Ding da weggeht und dahinter dann eine Tabelle zum Vorschein kommt.

A: Genau. Hast Du den, das Hilfe-Icon gesehen, in dem Fall?

B: Ja, das habe ich gesehen, ja.

A: OK.

B: Dann klicke, da würde ich, kann man da auch drauf?

A: Genau, genau, da kannst u draufklicken. Und was siehst Du jetzt? Beziehungsweise kannst Du da noch weiter durchklicken.

B: Genau, was sehr ich hier? Ich sehe hier verschiedene Möglichkeiten, ich kann Daten analysieren, ich kann Interacting, wüste ich nicht, was das ist und Reading auch nicht, also klicke ich mal drauf. Reading. Ah, jetzt bekomme ich eine Erklärung. Ja. OK, und da kann ich jetzt dann wahrscheinlich da durchgehen zum nächsten Punkt, wo ich.

A: Ja, Du bist knapp danebengetroffen, aber auf den Pfeil war schon richtig, aber der Klickbereich ist sehr klein.

B: Aber das ist sehr hilfreich so, das ist super. Und dann bin ich wieder bei Punkt 5, Ok, gut. Und dann mache ich das Reading wieder zu und dann probiere ich mal aus Interacting, da kann ich wahrscheinlich auf die Punkte klicken. Genau. Und dann kann ich hier mir noch einen zweiten Punkt erklären lassen. Hmhm. Jetzt kann ich wieder zumachen, ia, das ist super. Interacting, das heißt, ich könnte jetzt theoretisch, nein. so weit sind wir jetzt nicht.

A: genau, also den Klick-Prototypen gibt es jetzt nicht, genau.

B: Das ist nur eine Demo, aber ich könnte da halt all diese Dinge machen, genau. Wenn das nicht kommen würde und ich würde das hier links oben sehen, würde ich da einfach ausprobieren, ich hatte das schon in Visualisierungen, dass diese genau die gleichen Symbole waren und die habe ich natürlich ausprobiert und dann ist es relativ klar zu sehen, was passiert, wenn man das (?).

#### A: Hmhm.

B: Und Analysing schaue ich mir jetzt auch an. Was kommt dann? Closer (?), this is the dot XX close to the origin in 00. Ahm, ahm. OK, verstehe ich. Similar to the negative

control. Und dann kann ich da wieder durchklicken, da gibt es auch noch was (?), das 420

ist schön, ja. OK. Ja. 421

#### A: Hmhm. 422

B: Ich meine, das ist hilfreich, doch, das ist hilfreich, also wenn, ich frage mich nur, in 423 welcher Situation das passieren soll, nicht, also es ist meistens so, dass diese 424

425 Interaktionen in Form von, wenn es in Form von wissenschaftlichen Publikationen sind,

wenn es entweder immer noch das, was früher einfach ein Druck auf einem Stück 426

- 427 Papier war, also nicht interaktiv.
- A: Hmhm. 428

429 B: Ohne Erklärungen und da muss ich dann die Legende lesen oder es ist eine PowerPoint, die mir jemand vorträgt. 430

A: Hmhm. 431

B: Ich weiß nicht, in welcher Situation mir das passieren sollte, dass so eine 432 Visualisierung ich sehe, aber sie mir niemand zeigt und ich sie auch nicht selbst 433

gemacht habe, die Situation ist mir nicht begegnet noch, in meinem Leben. 434

A: Hmhm. OK. Das heißt, wahrscheinlich würdest Du, also die Idee wäre sozusagen, 435

dass Du, das ist ja ein Scatter Plot, den hast natürlich schon mal gesehen, aber es war 436

jetzt sozusagen, das hast Du auch selber gesagt, der Kontext sozusagen vielleicht 437

- schon das Interessante und normalerweise, wie Du sagst, würde Dir das ein Kollege 438
- erklären, und die Idee wäre halt, dass wenn Du jetzt einen Tool bekommst, sozusagen, 439
- und warum auch immer der Kollege (lacht) vielleicht nicht so viel zeit hat, oder Du halt 440
- nicht alles vom Programm erklärt bekommst, und dann direkt in die Visualisierung 441
- kommst, dass Du dann Dir die Informationen noch holen kannst, ohne, dass Du jetzt, 442
- was Du auch vorher gesagt hast, den Kollegen nicht nochmal stören musst oder willst 443
- wenn Du dann sozusagen das sehen möchtest. 444
- B: Ja. 445

A: Das wäre ein use case, den wir sozusagen auch in Betracht gezogen hätten. 446

B: Ja, also ein use case könnte zum Beispiel sein, was mir jetzt auch einfällt, dass 447 448 unsere Informatiker ein Analyseprogramm programmieren, was sie ja gemacht haben und ich mache diese Analyse und kriege dann das Ergebnis als Plot präsentiert, das 449 heißt, ich habe diesen Plot sozusagen machen lassen, aber die Informatiker haben 450 sich überlegt, wie sie es darstellen und ich bekomme das einfach konfrontiert und in 451 452 dem Fall wären solche Erklärungen schon sinnvoll. Ja, doch, ja, doch, kann es geben, 453 stimmt.

A: Hmhm. 454

B: Oder bei so Webseiten, wie cBioPortal, ich weiß nicht, ob ihr das kennt? 455

- A: Ja. 456
- 457 B: Das hat einen Haufen Visualisierungen von Krebsmutationsdaten, da mag das dem

einen oder anderen helfen, der sowas noch nie gesehen hat, aber das ist mit 458

Mouseover, ist es dort eigentlich getan, muss ich sagen, also. Ja oder es gibt auch 459

noch andere, FoundationNet sind wir, haben auch eine Visualisierung von großen Zahlen von Krebsdaten, die haben ihre eigene gemacht, da passiert was Ähnliches, ja, auch da wäre es.

A: OK, ich verstehe. Hmhm. OK. Ja, super, danke. Dann sind wir eigentlich schon am Ende von dem Interview. Wir haben noch zwei abschließende Fragen. Oder eine nur noch, genau. Ob es noch irgendwas gibt, was Du generell, entweder zum Onboarding oder zum Interview noch hinzufügen möchtest, was Du uns noch mitteilen möchtest? Noch irgendwas sozusagen was Du uns vielleicht noch fragen möchtest, irgendwelche Wünsche, Anregungen?

B: Nein, weiß nicht, ich meine, wie gesagt.

A: Ja. Ist OK. War so eine abschließende Frage. Passt schon, Also dann sind wir auch am Ende des Interviews angekommen. Vielen, herzlichen Dank für Deine Zeit und.

B: Mir hat es Spaß gemacht.

A: Und auch für Deine Eindrücke und für Deen Erfahrungen, das ist immer sehr hilfreich, auch für uns, weil, also wir lernen auch sehr viel durch diese Experteninterviews dazu, weil wir am anderen Ende sozusagen sitzen und manchmal, aber nicht sehr häufig, mit den End-Usern in Kontakt kommen und dann ist es halt schon interessant, wie die am Ende oder wie ihr die Visualisierungen seht und benutzt.

B: Alles klar.

A: Hmhm. Gut.

C: Danke schön.

A: Dann, danke schön.

D: Danke.

B: Viel Erfolg. Ciao.

A: Schönen Tag. Danke.

C: Danke.

A: Tschüss.

C: Tschüss.
#### Interview 5

A: OK, dann teile ich auch meinen Bildschirm mit dem Fragebogen. So. Elisabeth, kannst Du den Bildschirm schon sehen?

#### B: Hmhm. Ja, kann ich.

A: Passt. Schriftgröße passt auch. Wir fangen mal an mit einer ganz lockeren Einstiegsfrage: Welchen Bezug zur interaktiven Visualisierung oder Datenvisualisierung müsste man genau sagen hier, hast Du, wofür setzt Du diese ein?

B: Ich setze sie hauptsächlich ein, um meine Ergebnisse visuell sichtbar zu machen, ob das jetzt aus Essays ist oder aus Screens, ja.

#### A: Hmhm.

B: Sprich, die Ergebnisse, die wir täglich im Labor erzeugen, werden im Allgemeinen visualisiert.

#### A: Hmhm.

B: Um sie, ich kann es weiter ausführen, um sie besser vergleichen zu können und sie besser einschätzen zu können, teilweise auch, um vielleicht Artefakte erkennen zu können.

A: Hmhm. OK. Cool. Vielen Dank. Vielleicht ganz kurz noch demographisches Detail, Deinen educational background und offizielle Aufgabe bei Dir?

B: Ich bin promovierte Biologin. Meine Berufsbezeichnung ist Senior Scienist bei Bayer. Und mein Arbeitsfeld ist die Entwicklung und Durchführung von High (?) Essays, um es mal ganz minimalistisch zu sagen, es stimmt natürlich nicht ganz, ich kann es noch weiter ausbauen, aber ich kann es auch dabei belassen.

A: OK. Das heißt, Planung und Durchführung von HTS?

B: Nein, Durchführung mache ich nicht, also Aufbau von Essays, Durchführung von Essays, teilweise im High (?), wobei das ja eher Michaels Aufgabe ist.

A: Hmhm.

B: Und Substanztestungen vom HTS bis hin zur klinischen Substanz, Substanztestungen und Charakterisierungen kann man sagen.

A: Hmhm. Cool. Damit bis Du genau unsere Zielgruppe.

#### B: Toll.

A: Genau, genau, Daniela und ich, wir schreiben uns gerade gegenseitig die Mitschrift, aber das werden wir im Laufe des Interviews hoffentlich.

B: Das ist ja, ich habe am Anfang schon gedacht, Wahnsinn, wie schnell die sind, aber.

A: (lacht)

B: Rechte Hand, linke Hand, nicht, ja, genau.

A: Ja, genau.

- 37 B: (lacht)
- 38 A: Genau, wir schreiben zu zweit und
- 39 B: Ja.
- 40 A: Der Multiplayer funktioniert nicht so gut. OK, Ich habe jetzt hier eine Reihe von
- 41 kleinen Bildern und Chart-Typen vorbereitet, ich würde Dich bitten, selbst zu
- 42 beurteilen, wie gut Du damit vertraut bist, von 1 überhaupt nicht vertraut, über nur
- 43 bedingt vertraut, 2, etwas vertraut, sehr vertraut 4 und äußerst vertraut, heißt, Du weißt
- 44 alles, was es zu wissen gibt darüber, wäre dann der höchste Wert, 5.
- 45 B: Also, wenn ich es verstehe, dann kann ich auch 5 geben, ja?
- A: Genau, wenn Du sagst, über bar Charts könnt ihr mir nichts mehr erzählen, dann
   gib eine 5 her.
- 48 B: OK, ich sage mal 5.
- 49 A: Bar Chart 5.
- 50 B: Nein, wir machen 4, wir machen 4, wir machen 4. Vielleicht verstehe ich ja 51 irgendwas daran nicht, keine Ahnung. OK.
- 52 A: Line Chart?
- 53 B: 5.
- 54 A: Pie Charts.
- 55 B: 5.
- 56 A: Scatter Charts, Bubble Charts?
- 57 B: Ahm, 5.
- 58 A: Hätte ich jetzt auch gesagt. Box Plots?
- 59 B: Ach, nehmen wir 4, die habe ich nicht so oft, aber eigentlich, eigentlich habe ich sie
- 60 verstanden, aber vielleicht gibt es irgendwas, was ich nicht weiß, ich weiß es, also
- eigentlich ist alles 5, eigentlich ist alles 5, weil das ist Quatsch, ich weiß auch ganz
- 62 oben kannst Du auch den, was sollte ich daran nicht verstehen können?
- 63 A: Hmhm. Hmhm.
- 64 B: Ich benutze sie nur nicht so oft, aber ist egal.
- 65 A: OK. Violin Plots?
- 66 B: Ja, das ist eher, da würde ich mal sagen, das ist 3, weil die hatte ich, bevor Du sie 67 uns gezeigt hast, tatsächlich noch nicht gesehen. Ja.
- 68 A: OK, Da gibt es ein ganz spannendes Video, wo dieselben Daten mit Violin Plots
- 69 und Box Plots animiert werden, also die Daten werden verändert und du siehst unten
- 70 die Violin Plots und die Box Plots und die Box Plots bleiben gleich, während sich die
- 71 Violin Plots mit den Daten verändern, kann ich Dir mal zeigen.

B: Ja. Ja, na ja, also die sind schon echt cool, aber ja, tatsächlich war es für mich neu. Und Sunburst kann ich auch noch nicht so lange.

A: Hmhm.

B: Da würde ich echt sogar mal 2 sagen, 2 war was nochmal.

- A: 2, nur bedingt vertraut.
- B: Ja, das ist, das ist OK, das ist OK, ja.
- A: Genau, Heatmaps?
- B: Pff. 4.
- A: Hmhm. Dendrogramme?
- B: 5.
- A: Parallele Koordinatensysteme?
- B: Pff, 4, eigentlich dachte ich, ich kenne es, aber who knows, ja.
- A: Treemaps?
- B: Hmm, so wie das da aussieht, würde ich sagen 3.
- A: Hmhm. Send Key Diagramme? So wie man es auch aus der Wählerstromanalyse kennt.
- B: Hmhm. Ja, auch 3, würde ich sagen.
- A: Hmhm. Und Node Link-Netzwerk oder Netzwerkvisualisierungen, wie wir sie hier skizziert haben?
- B: Ahm, ja, da würde ich, hmm, 3, wenn überhaupt. Wir machen 2, schon mal gesehen. aber kenne ich mich nicht gut aus.
- A: Hmhm. Gibt es andere Visualisierungstypen, die Du aktiv verwendest, die hier nicht vorgekommen sind?

B: Nee, ich bin da relativ einseitig. Ich glaube, ich, fällt mir jetzt gar nichts ein, also

A: OK, na uns wäre auch nichts eingefallen, also ist kein Beinbruch. OK. Dann bitte ich Dich, welche Visualisierungen, die Du bereist selbst erstellt hast, sei es jetzt für

Präsentation, auf Konferenzen, für Dich persönlich zur Analyse, also was Du selbst

- B: Nee, Scatter ist ja bei den Bubbles und so weiter mit drin, nicht?
- A: Genau, richtig, genau, richtig.

spontan.

B: Ja.

auch erzeugt hast.

- A: Mit den Dir zur Verfügung stehenden Daten.
- B: Ja. Bar Charts habe ich gemacht.
- A: Ja oder Nein, genau, ja. Line Charts? 106

- B: Ja. 107
- A: Pies? 108
- B: Ja. Ja. 109
- A: Kann ich schon ein Ja machen, genau, Box Plots? 110
- B: Ahm, pff, nee, also selber gemacht, wenn ich ganz ehrlich bin, glaube ich sogar 111
- 112 noch, ja, vielleicht ein Mal, ach schreib mal Nein, das ist auf jeden Fall nicht meine
- tägliche Arbeit. 113
- A: Hmhm. Violin Plots? 114
- B: Nein, habe ich selber nicht gemacht, nur gesehen und die Sunburst auch und Heat 115
- Maps, Ja, Trees auch, Ja. Ja. Tree Map, wie ich schon gesagt habe, weiß ich nicht, 116
- was, so richtig, was damit gemeint ist, weil ich habe das Gefühl, das, was ich da sehe, 117
- kenne ich nicht, also Nein. 118
- 119 A: Hmhm.
- B: Und die habe ich auch noch nicht selbst gemacht, nur angekuckt. 120
- A: Hmhm. 121
- B: Und die habe ich jetzt tatsächlich mal selber erzeugt, ein Mal, aber egal, nicht. 122
- A: Hmhm. OK. Dann haben wir noch eine Frage über Tools, die verwendet werden, 123 einfach Ja, Nein? Tableau. 124
- B: Also ob ich sie regelmäßig verwende? 125
- A: Genau, ob Du sie im täglichen? 126
- B: Nein, nein, kenne ich, verwende ich nicht, kenne ich nicht, 127
- A: Spotfire? 128
- B: Spotfire, evaluieren wir gerade, aber regelmäßig nein, also eigentlich ist es echt ein 129
- Jein, kennen Ja, verwenden Nein. 130
- A: Ja. Ich schreibe mal, genau. 131
- B: OK, genau. 132
- 133 A: KNIME?
- B: Nein. 134
- A: Power PI? 135
- B: Nein. 136
- A: Excel? 137
- B: Ja. 138
- A: Der Klassiker, Quick View? 139
- 140 B: Nein.

## A: SAS?

B: Nein.

A: Inhouse Tools, das heißt?

B: Ja, ganz viel.

A: Genau, ganz spannend. So. Eines machen wir, eines machen ja wir. Genau, andere Tools, die Du verwendest, die jetzt nicht bei euch Inhouse entwickelt sind, aber ihr trotzdem als zugekauftes Tool verwendet?

B: GraphPad Prism.

A: Wie schreibe ich das? So?

B: Ja, es kann auch, ja, ja, ja, genau, ich glaube, das wird auseinander geschrieben. Dann habe ich früher, ach, jetzt habe ich es vergessen, wie es hieß, ich wollte das, ach, eben hatte ich es doch. Sigma Plot.

### A: Hmhm.

B: Hmm. Genau, ja, ich glaube, das war es, es geht ja jetzt um Datenvisualisierung, nicht.

## A: Genau.

B: Dass Visualisierung, ja, alles andere ist Inhouse, fällt mir jetzt gerade nichts anderes ein.

A: Hmhm.

B: Genau.

A: Sehr gut. Danke. Wie wurden diese Tools, die Du verwendest, gelernt? Das heißt, egal, ob jetzt gekaut oder Inhouse Development oder wie wurde das gelernt? Selbststudium, Trial & Error?

B: Trial & Error, Hilfe durch Kollegen.

A: Hmhm. Liest Du Tutorials, wenn Du welche zur Verfügung hast?

B: Wenn es sein muss? Videos auch, aber mag ich nicht.

## A: Hmhm.

- B: Weil ich finde Tutorials eigentlich besser.
- A: Hmhm.

B: Also intuitiv (?) finde ich am besten, Hilfe durch Kollegen finde ich auch toll.

A: Hmhm.

B: Tutorials dauert eben einfach länger, ja.

A: Hmhm.

- 174 B: Und deswegen, ich bin ja faul, und meistens auch relativ ungeduldig und deswegen,
- 175 sowas, wie IPhone, einfach drücken, wischen und irgendwas passiert, das ist immer
- 176 schick, nicht.
- 177 A: Hmhm.

B: Also ja und Videos finden ich eben wirklich, also das Lesen von Tutorials ist auch Ja, nicht, übrigens.

- 180 A: Danke, ja.
- 181 B: Das mache ich noch eher und das mache ich lieber, als mir Videos anzukucken,
- habe ich festgestellt, weil man bei Videos einfach länger suchen muss, wo was ist, ich
- 183 habe nicht diese Kapitel und manchmal geht es zu schnell, da muss man sowieso
- 184 wieder stoppen, da lese ich es mir lieber selber durch.
- 185 A: Hmhm.
- 186 B: Strukturierte Kurse habe ich eigentlich fast gar nichts gehabt, außer vielleicht einmal
- 187 ein Inhouse-Tool, die sind eigentlich natürlich auch schön, ja, weil man natürlich auch
- 188 Fragen stellen kann, sehr gezielt seine Probleme adressieren kann, das ist dann halt
- 189 was, wenn man schon ein bisschen fortgeschritten ist und dann eben spezifische
- 190 Fragen hat, dann macht das Sinn.
- 191 A: Hmhm. Wie sind diese Schulungen und Kurse du vielleicht, Christine, für Dich zur
- 192 Ergänzung, Elisabeth und ich planen gerade einen so einen Kurs, für das Tool, das
- 193 wir entwickeln, darum ist die Frage jetzt eigentlich hier extrem passend. Welche
- 194 Materialien wurden das eingesetzt, wenn Du an die Inhouse-Schulungen denkst?
- B: Ja, ist (?) Inhouse-Schulung, also von daher wurde Screenshots, ja, genau, könnteman sagen.
- 197 A: Hmhm.
- 198 B: Also ich kann mich tatsächlich nur an eine echte Schulung erinnern, das war eine
- 199 PICS-Schulung und das wurde hands on gemacht und wahrscheinlich wurden auch
- 200 ein paar Screenshots gezeigt, das ist schon ewig her.
- 201 A: Hmhm.
- B: Ansonsten muss ich mal ein bisschen aus dem Nähkästchen plaudern, werden wir
   hier bei Bayer einfach, uns werden Programme vorgesetzt und do it, ja.
- A: Hmhm.
- B: Das ist, keiner erklärt dir was, vielleicht kannst du dir ja irgendwo in den Tiefen des
  Netzwerkes irgendwo in Tutorial finden, aber im Grunde genommen wirst du damit
  alleine gelassen.
- 208 A: Hmhm. OK.
- 209 B: Eher unschön.
- 210 A: Genau. Das heißt, die Probleme bei der, also diese Frage bezieht sich jetzt auf
- diese, die Methoden, die Du bist jetzt angewandt hast, also alle.

Bibliothek verfügbar. BiblioBiek. Wien 14 14 14 14 2145 146 1 tan de 141 141 141 141 en Ist rtation i 感ail哉ol 0,151 152 5 dieser Mai tige 5155 ersio 156 inalve Pithis Origii 128 0 jedruckte ( ⊕ 162 approbiente Supprovente Supprovente Ð Φ ⊡ 1<del>6</del>6 **0**68 oth deegub **.** m

#### B: Smartkey. Hmhm.

#### A: Was sind, aus Deiner Sicht, die Probleme?

B: Hmm. OK. Beim strukturierten Kurs? Welche Probleme sind aufgetreten? Ja, das war schwierige. Ja, wenn ich mal einen gehabt hätte, verflixt. Jetzt muss ich mal kurz nachdenken, also. Hmm. Was sind, genau, was für Probleme auftauchen können, ja, genau, wir hatten natürlich letztens auch, muss ich sagen, also so eine Einführung in StarDrop, teilweise ging es zu schnell, man konnte nicht folgen.

#### A: Hmhm.

B: Hmm. Probleme können auch sein, dass die falschen, ich sage mal, Arbeitsvorgänge gezeigt werden, die die eigene Arbeit nicht tatsächlich adressieren und abbilden. Gut ist dann natürlich en Kurs, in dem ich intervenieren kann und sagen kann, ich möchte gerne dieses oder jenes wissen oder gezeigt haben, das ist dann immer toll. Wenn es nicht geht, ist es schade, dann kann es eben sein, dass der ganz Kurs für die Katze ist, weil mir Dinge gezeigt werden, die ich sowieso nie anwende, vom Programm.

#### A: Hmhm.

B: Deswegen sind interaktive Kurse, in denen man grob das Programm anreißt und dann auf die speziellen Bedürfnisse der User eingeht, denke ich, das Sinnvollste.

#### A: Hmhm.

B: Dazu muss man natürlich so ein bisschen darauf achten, das bezieht sich ja letztendlich auch auf das, was wir planen, dass man alle mit ins Boot holt, da heißt, dass die Gruppen oder die Menschen, die da zusammen geschult werden, auch nicht zu heterogen sind.

#### A: Hmhm.

B: Sonst kann das für den einen oder anderen schnell mal langweilig werden, das heißt, ich muss es letztendlich, ich kann es einteilen in verschiedene Zeiten oder in verschiedene Tage oder in verschieden, ja, Kursabschnitte für verschiedene Themen quasi, ganz klar umreißen, damit nicht eine Gruppe dasitzt und Däumchen dreht, während die andere dann stundenlang über irgendeinen Fokus diskutiert.

A: Hmhm. OK. Gut.

B: Ja, wenn das reicht, ist es gut.

A: Ja, sollten wir auch uns zu Herzen nehmen, wenn wir unseren Kurs planen. Bei den anderen Methoden, die Du angewandt hast, was stört Dich da, ich meine, Video hast Du schon mal angesprochen, es gibt keine Übersicht, das Tempo ist manchmal zu schnell, manchmal zu langsam, man findet nichts. Selbststudium, welche Probleme oder Tutorials, was sind da?

B: OK, also, ja, genau, was beim Selbststudium natürlich einfach sein kann, wenn ich Dinge versuche, dass ich einfachste Handgriffe nicht kenne, dass ich vielleicht das Prinzip eines Programmes nicht kenne und dann komme ich irgendwann nicht weiter. Ich probiere Sachen aus und bleibe dann quasi an einer Stelle hängen, weil mir

- 252 vielleicht da eine bestimmte Taste fehlt oder ich vielleicht irgendwie was falsch
- angelegt habe und deswegen keine Grafik erzeugen kann, was auch immer, das ist
- 254 natürlich das Problem.
- 255 A: Hmhm.

B: Bei Kolleginnen und Kollegen ist es eigentlich immer super, da gibt es keinen
Kritikpunkt. Der einzige wäre, ich frage und der- oder diejenige kann mir nicht helfen,
das ist dann schade, aber ist egal, das ist immer toll, das mit Kollegen gemeinsam zu
erarbeiten oder durchzugehen, genau. Bei den strukturierten Kursen ist es eben so,
die sind zeitaufwändig, also ich lerne zwar gut in der Gruppe, aber ich muss viel Zeit
mitbringen und wenn ich Pech habe, ist nur ein Drittel der Zeit für mich wirklich gut
investierte Zeit.

263 A: Hmhm.

B: Ja, also das Zeitthema ist eigentlich da ganz wichtig Deswegen sind ja im Grunde
genommen solche Tutorials auch ganz nett, wo ich gezielt nachlesen kann, was ich,
wenn die gut gemacht sind, sind die eigentlich toll.

A: Hmhm. Hmhm. OK. Eine Frage, die sehr eng verwandt ist, von der Modalität der
Hilfe, die Du suchst, Videos, hast Du ja schon gesagt, aus schon genannten Gründen
eher nicht so gut, Bild und Text oder auch interaktiv bzw. in the app, suchst Du Dir
auch dort Hilfe?

- 271 B: Ja.
- A: Hast Du schon?
- 273 B: Genau.
- A: Also auch das?
- 275 B: Ja, Du meinst so eine Hilfefunktion in dem jeweiligen Programm, nicht?
- 276 A: Zum Beispiel, ja.
- B: Ja, klar, absolut.

A: Hmhm. OK. Kennst Du eine der folgenden drei Ressourcen, die wir hier genannt haben? der Data Viz Catalogue, From Data to Viz oder Data Viz Project?

B: Nein. Kenne ich alle nicht, alles unbekannt.

A: Dann darf ich Dir kurz mal zeigen. Na. Das sind drei Webseiten, wo man
Informationen über Visualisierungen bekommt. Gier aufgelistet nach der
Visualisierungsart, wo man dann draufklicken kann, ein Haufen Werbung, aber dann
die Erklärung dazu bekommt, einerseits aus Visualisierungssicht, andererseits aber
auch nach der Funktion, das heißt, ich kann auch so einsteigen, ich will eine Proportion
darstellen, welche Visualisierungstypen habe ich zur Verfügung.

- 287 B: Hmhm.
- A: Und ähnlich sind die anderen Kataloge aufgebaut, wo ich hier sagen kann, OK, wie
- 289 sind meine Daten aufgebaut, Time Serious Networks, kategorischem numerische

Daten, heterogene Daten und dann so einen Entscheidungsbaum bekomme, welche Visualisierungsart für welches Problem sich eignen würde.

B: Hmhm. OK. Aber mal ganz ehrlich, sowas wäre ja auch eher für die Entwickler von Software wichtig zu wissen.

#### A: Hmhm.

B: Ich wende es ja quasi nur an, nicht, ich entwickle ja nicht selber etwas. Aber sagen wir mal, es ist natürlich für uns schön, weil wir dann, also gerade, wenn wir so ein Projekt mit euch zusammen entwickeln, da vielleicht noch eine Anregung bekommen könnten, wie man was besser visualisieren kann, ich sage mal, das macht ihr zwar auch, aber keiner kennt unsere Daten und unsere, sage ich mal, unsere Ansprüche so gut wie wir selber, nicht.

A: Genau, genau. Welche zusätzlichen Ressourcen würdest Du Dir wünschen für das Onboarding oder für das Lernen von Visualisierungen und Visualisierungstools? Gibt es irgendwas, was fehlt, aus Deiner Sicht?

B: Hier bei Bayer jetzt oder?

A: Grundsätzlich, in Deiner Arbeitswelt, so wie Du sie bei (?) kennengelernt hast, ja.

B: Hmm. Ich sage mal, ich sage, das ist ein bisschen, ist jetzt echt sehr allgemein, weil es kommt ja sehr auf das Programm darauf an, also zum Beispiel in Excel gibt es Hilfen, die kann man durchaus verwenden, ja.

#### A: Hmhm.

B: Also, wenn ich es einfach mal so als Beispiel nehme, wie mache ich da irgendwas, ich suche erst in Excel, wenn ich es nicht finde, suche ich in Google.

#### A: Hmhm.

B: Das heißt, welche, ich brauche keine zusätzlichen Ressourcen, ich brauche vielleicht optimalere Ressourcen, ja, aber es ist ja für fast jedes Programm was da, was man braucht.

#### A: Hmhm.

B: Also von daher, weil, sorry, wir sind doch alles durchgegangen, also entweder es gingt Videos, es gibt Tutorials, es gibt irgendwo in jedem Programm eine Hilfefunktion, es kann natürlich trotzdem sein, dass man nicht weiterkommt, dann liegt es aber meistens an einem selber oder weil diese Hilfen einfach schlecht gemacht sind, nicht.

#### A: Ja.

B: Wenn ich zum Beispiel in einem Programm, zum Beispiel mit GraphPad Prism arbeite ich relativ selten, kenne ich mich eigentlich sehr schlecht mit aus, nicht und da muss ich sagen, sind auch die Hilfen, wenn ich mal wieder wild drin rumklicke, einfach schlecht gemacht. Welche zusätzlichen Ressourcen würde ich mir wünschen? Gar keine, außer dass die Hilfen mal ein bisschen besser sind.

A: Hmhm. Hmhm.

C: Darf ich da was nachfragen?

329 B: Ja.

C: Wie sind die Hilfen bei dem Tool, also wie, warum sind sie schlecht oder was könnteman da verbessern?

B: Ich habe es jetzt so lange nicht verwendet, aber wenn die Hilfen schlecht sind, dann
liegt es daran, dass sie entweder, ich sage mal, zu schlecht verständlich sind, dann
langer Text, wenig Bild, ja, ah, eine gute Hilfe ist eigentlich eine, in der Bild und Text
gut zusammenpassen und mich dann durch diese Applikation durchführen. Und was
ich eben auch brauche, das ist eine vernünftige Benennung der Arbeitsschritte, ja,
dass ich eben das, was ich suche, auch leicht finden kann und mich dann nicht
mühselig durch scrollen muss.

339 A: Hmhm

- 340 C: Hmhm. OK.
- 341 B: OK. Ja.

A: OK, bezogen jetzt auf die Hilfe, würdest Du es bevorzugen, direkt in der App zu
 haben und wenn ja, soll es immer, also soll es beim ersten Mal aufpoppen, soll es
 immer aufpoppen oder willst Du bewusst die Hilfe antriggern oder willst Du überhaupt,

345 bevorzugst Du überhaupt Hilfe, die außerhalb des Tools ist?

B: Hmhm. Nein, ich denke, wenn die Hilfe gut gemacht ist, dann ist es gut, wenn sie innerhalb des Tools ist. Ich denke, man sollte sie aktiv anwählen müssen.

348 A: Hmhm.

349 B: Also ich muss jetzt nicht immer, überall, zwei Dinge, also wenn ich zum Beispiel,

- 350 was ich sehr schön finde, ist, wenn ich zum Beispiel über eine Funktion drüber (?),
- also drüber streiche und dann eingeblendet wird, wofür ist das da. Finde ich im Grunde
- genommen sehr schön, braucht man irgendwann nicht mehr, aber es ist nice to have,
   es stört auch nicht.
- 354 A: Hmhm.

B: Genau. Und ich sage mal, Hilfe außerhalb, also auf einer anderen Webseite zu
Beispiel, mache ich auch, mache ich teilweise auch gerne, weil bestimmte Funktionen
sind dann auf der einen Webseite vielleicht besser beschrieben als auf der anderen,
ja.

359 A: Hmhm.

B: Gerade, wenn ich so an Excel denke, da kann einem zum Beispiel der Bayer-interne
Helpdesk meistens auch nicht helfen, das ist echt unglaublich, da bin ich schneller,
wenn ich einfach im Web suche und klicke da verschiedene, ich habe da so meine
bevorzugten Seiten, wobei ich euch die nicht nennen kann, aber wenn ich sie sehe,
erkenne ich sie immer wieder. Auf denen dann gewisse Vorgänge erklärt werden und
manche Webseiten machen das richtig gut, ja, und die kucke ich mir dann lieber an
als dass ich da in Excel selber herum suche und surfe.

A: OK. Cool. Dann kommt jetzt eine kurze Definition, was wir unter Onboarding
 verstehen, wir verstehen unter Onboarding den Prozess, die BenutzerInnen beim

Lesen interpretieren und extrahieren von Informationen aus visuellen Darstellungen von Daten zu unterstützen. Ich schicke Dir jetzt diesen Link hier, das ist so ein Sketch, so ein interaktiver Klick Prototyp, wo ich Dich bitten würde, Deinen Bildschirm zu teilen und einfach mal laut zu denken, was Du von dem hältst und wo Du in dieser Visualisierung die Hilfe suche würdest?

B: Hmhm.

- A: Ich schicke Dir jetzt hier im Chat diesen Link.
- B: Ja.

A: Und würde Dich dann bitten, Deinen Bildschirm zu teilen und einfach mal laut zu denken.

B: OK. Könnt ihr meinen Bildschirm, nee, ihr könnt ihn noch nicht sehen oder?

A: Ja, wir können ihn schon sehen.

B: Sh ja, ihr könnt ihn sehen, OK. Alles klar. Ich denke laut, ich denke laut, dass ich zwei Satter Plots sehe. Was mir auffällt, ist. Das jeweils nur eine Achse deutlich angezeigt wird, einmal die X-, einmal die Y-Achse.

#### A: Hmhm.

B: Dass die Daten, das ist eben bei uns nicht so typisch über beide Achsen quasi hinweg gehen, na ja, gut, klar, sie sind natürlich auch beide bei 0, das kann man so machen, muss man nicht so machen.

#### A: Ja.

B: Ich, wenn ich einen Graphen anlege, dann würde ich eher die Achse dann eben bei dem negativen Wert enden lassen, ich würde es nicht auf 0 setzen.

#### A: Hmhm.

B: Aber das ist, glaube ich, Geschmacksache, das hat mich jetzt im ersten Augenblick so ein bisschen verwirrt, weil ich dachte, huch, warum gehen denn die Punkte jetzt über die Achse, macht man doch gar nicht. Doch, kann man natürlich machen.

A: Wenn Du mehr erfahren würden wolltest, über dieses Chart, wo würdest Du Hilfe suchen?

B: Hmm. Ich sehe es gerade, back value, kommt das von uns, wo ich, ich will mehr wissen.

#### A: Ja.

B: OK. Also ich würde wahrscheinlich, könnt ihr meine Mouse sehen?

A: Ja. Ja, sehen wir.

B: OK, dann tu ich jetzt mal so als würde ich, ich würde es mir erst mal, ich würde hier bei (?) kucken oder ich würde bei Run, ich würde erst mal hier oben alles anklicken, würde erst mal kucken, wo kann ich überhaupt was anklicken?

5 A: Hmhm.

B: Also eigentlich, genau, jetzt erwarte ich, wenn ich irgendwo drüber gehe, dass sich 406 eben da die Farbe ändert oder ein Fenster aufpoppt, sodass ich weiß, kann ich hier 407 was anklicken, kann ich eine Information kriegen oder ist das hier einfach nur 408 irgendwas, weil Scatter Plot ist die Überschrift, da erwarte ich mir natürlich gar keine 409 Hilfe oder auch bei der Achsenbeschriftung, das ist normalerweise nicht interaktiv, 410 411 aber ich würde jetzt erst mal kucken, was hier auf diesen Seiten ist überhaupt interaktiv, wo kann ich klicken, kann ich über die Punkte gehen, passiert dann 412 irgendwas, kann ich die, als allererstes würde ich, wahrscheinlich als aller allererstes 413 die Punkte markieren und kucken, kann ich was markieren und passiert dann was in 414 beiden Plots, haben die irgendwas miteinander zu tun, sind die miteinander verlinkt? 415 Dann würde ich, dann sehe ich, da oben ist ein Feld, da steht Run3, dann gibt es 416 bestimmt auch einen Run1 und 2, und würde das aufklappen und kucken, was kann 417 ich da machen? Genau. Ich sehe, dass ich hier irgendwo vor und zurück gehen kann, 418 das ist auch toll. Ich würde halt überall draufklicken, wo man mal klicken kann, um zu 419 420 kucken, was tut sich.

A: OK. Und vielleicht Gegenfrage: Dieses rote Fragezeichen rechts unten in den Plotsist Dir aufgefallen?

B: Das ist mir aufgefallen, ja, aber das Fragezeichen sieht irgendwie ein bisschen
seltsam aus, das habe ich eher als Fremdkörper tatsächlich interpretiert.

425 A: OK.

426 B: Gute Frage, weil es ist so ziemlich das Auffälligste, was einem sofort ins Auge sticht

- 427 und deswegen habe ich es auch bei mir, weil es nicht in mein Weltbild passt, sofort428 erst mal weggestrichen.
- 429 A: OK.
- 430 B: Tatsächlich hätte ich da nicht als erstes draufgeklickt.
- 431 A: Wirklich spannend, weil Michael hat es zum Beispiel gar nicht gesehen.
- 432 B: (lacht) Doch, ich habe es gesehen, weil ich mir gedacht habe, das sieht so komisch
- 433 aus, das Fragezeichen, ist das umgedreht und dann habe ich irgendwie gesehen, nee,
- 434 es ist nicht umgedreht, es ist einfach ganz komisch gezeichnet. Also aber das war
- alles, also mit einem Fragzeichen (lacht) in einem Scatter Plot kann ich echt nichtsanfangen.
- 437 A: OK. Das heißt, diese Diskussion mit dem Icon hatten wir auch, weil wir auch nicht
- 438 wussten, ob es ein richtiges Fragezeichen ist oder nicht. Darf ich Dich mal bitten, im
- 439 linken Scatter Plot auf das Fragezeichen zu klicken?
- B: Ah, das geht ja, also ist es ja doch interaktiv, das ist ja toll.
- A: Genau, nur dieser Teil, das Tool im Hintergrund, das Du siehst, ist übrigens einScreenshot von Hidden.
- B: Ah, von, ah, OK. Ah, oh. (lacht) Woher, OK.
- 444 A: Genau.
- 445 B: Ja.

B: Ja, genau, jetzt würde ich erst mal, also auf das Kreuz gehe ich jetzt nicht, doch, wobei, warum denn nicht, ja klar, alles klar, da kann ich es an und aus, da kann ich analyzing, closest to origin, das ist. Autsch. Die verflixte Kiste. OK, relativ viel geschrieben, the closer the origin, the more similar is it to negative sample intensed to be less interesting. Na wer soll das denn verstehen? Entschuldigung, also da bin ich, da muss ich mich echt konzentrieren, um da zu überlegen, wie soll es überhaupt gehen. Was ist denn das? OK. (?) OK. Ja, alles klar, also die Punkte sind näher zusammen, wenn die Kinetik irgendwie zusammengehören. Würde ich gerne mal, verflixt, nee, also.

A: Genau, also der Scatter Plot wäre im Tool dann natürlich interaktiv.

B: Interaktiv, OK, ich kann jetzt hier nichts anklicken, hier wird nur was erklärt. Jetzt gehe ich hier mal wieder raus, genau, jetzt kann ich hier zurück. Interacting. Ah, ja, when hovering over scatter plot. OK. Kann ich das auch aktiv machen? Nee, kann ich nicht. Hier ist nur eine Erklärung, alles klar. Und wenn ich dann selber rübergehe, ist auch noch eine Erklärung.

A: Genau, also wie gesagt, die App im Hintergrund ist jetzt wirklich nur ein Screenshot, nur dieses Onboardingding ist interaktiv.

#### B: Hmhm.

A: In der richtigen App wäre natürlich die App weiterhin interaktiv, also.

B: Ja.

A: Da könntest Du die Dinge ausprobieren, die hier beschrieben sind.

B: Genau, was ich jetzt hier noch nicht verstanden habe, ist tatsächlich, ach so, hier gehe ich auf diese zwei Unter-, (?) OK, ja, also ich muss mal ein bisschen geduldiger sein, nicht so schnell durchgehen und das vielleicht in Ruhe durchlesen, dann kann ich es auch möglicherweise verstehen. Hmm. Was heißt (?) substance sample. OK. So. OK. (?) a view to the right, ja, das ist, es ist ein bisschen blöd, wenn man hier was hat, also ich sehe es hier, verschieden Icons, ja, ich kann zoomen, ich kann einen Screenshot machen, ich kann selecten und ich kann (?), OK, ja, das ist nett. Da kann ich das ja versuchen, ach so, das kann ich jetzt leider alles gar nicht machen.

A: Genau. Genau.

B: Das ist jetzt ein bisschen schade, weil normalerweise, wenn jetzt irgendwie so ein Fenster auftaucht, dann würde ich natürlich da sofort draufklicken und das sofort ausprobieren, um zu sehen, ob das Gelesene auch mit dem, wie ich es verstanden habe, übereinstimmt. Also wenn ich zum Beispiel Zoom lese, dann will ich eben kucken, habe ich das richtig verstanden, ich kann es groß und klein machen, und will das sofort ausprobieren.

A: Genau, also das wäre in der richtigen App natürlich dann sehr wohl der Fall. Also wir könnten das auch in Triage einbauen und dann würden wir es natürlich direkt beim echten Scatter Plot einbauen.

487 A: Hier liegt halt nur ein Screenshot drunter.

B: Also das verstehe ich zum Beispiel hier überhaupt nicht. Negative control? Und wo
ist die Negativkontrolle? Also diese Beschreibung, diese ellenlangen, die find ich jetzt
irgendwie so ein bisschen suboptimal, weil es ist jetzt zwar auch closest und similar
unterstrichen, da ist ja alles schön, X, wieso denn auch, this ist the dot XX, the origin
00, compounds (?) to the negative control. Und wo ist die Negativkontrolle? Sorry, das

493 ist, das ist echt nicht verständlich, also ich verstehe es nicht.

#### 494 A: Hmhm.

B: Ja, weil was ist XX, was ist 00, also 00 kann ich mir vorstellen, liegt hier unten, XX,
da würde ich eher XY erwarten, ich bin ja hier in einem Plot, ich habe zwei Achsen, die
sind unterschiedlich benannt, wieso habe ich XX? Und dass es um Kinetiken geht, gut,
das hast Du mir gesagt, weil es geht um Hidden, das ist schon klar, most similar to

- 499 negative control, klar, da unten ist vielleicht irgendwo bei 0 die Positivkontrolle, da
- 500 könnte hier oben die Negativkontrolle sein, aber die Frage ist, wo ist die
- 501 Negativkontrolle und ist das, das hier oben ist Substanz und hier sind ja noch welche
- 502 noch weiter weg und daher verstehe ich nicht, warum das die nächste zur
- 503 Negativkontrolle ist, nein, absolutes Rätsel.
- A: Hmhm. OK. Wie ist die grundsätzlich, ich meine, abgesehen davon, dass die Texte
- natürlich jetzt nicht die passenden sind, weil es ja auch wir geschrieben haben und
- nicht David und Patrick, also die hätten wahrscheinlich eine andere Qualität für uns.
- 507 B: Die wissen, wie blöd wir sind, genau. (lacht)
- 508 A: So habe ich es jetzt nicht gemeint. Aber auch wir finden hier sicher nicht das richtige
- 509 wording. Aber wie bist Du Dir grundsätzlich mit diesen Punkten und so, wir haben
- 510 gesehen, Du nimmst die Punkte von unten nach oben?
- 511 B: Ja, ich habe tatsächlich, ja, ist interessant, überlegt, ob ich jetzt, also ich will immer
- 512 systematisch vorgehen und systematisch eigentlich von oben nach unten, aber
- tatsächlich habe ich mir jetzt auch bei Zeitungsartikeln angewöhnt, die von hinten zu
- 515 A: (lacht)

lesen.

- 516 B: Weil die interessanten Sachen oft weiter hinten kommen, irgendwie, das ist so.
- 517 A: Ja, am Anfang ist (?).
- 518 B: Man vertut sich oft so die einen, also oft fängt man irgendwo oben an und denkt
- sich, hey, wann passiert denn oder wann kommt denn jetzt endlich was Interessantes,
- 520 ja, und die richtig interessanten Sachen, die sind eigentlich erst da, wo man weiter
- runter geht, deswegen ist es so eine kleine Macke von mir, muss ich sagen, also
- eigentlich gehe ich sonst logisch vor, ich habe auch jetzt kurz überlegt, nimmst Du den
- 523 obersten Punkt, aber ich war erst gerade unten und bin einfach unten geblieben, es ist
- 524 so eine intrinsische Faulheit, mich dann zu bewegen. Und wie schon gesagt, ich 525 versuche tatsächlich, auch längere Artikel, weil ich morgens auch nicht so, insgesamt
- nicht so viel Zeit zum Lesen habe, mir die Bröckchen rauszusuchen, ich fange hinten
- 527 an und suche mir die Bröckchen raus und arbeite mich dann nach vorne, wenn es

interessant ist, komme ich auch ganz vorne an, wenn es nicht interessant ist, komme ich niemals vorne an.

A: (lacht) Eine coole Vorgehensweise, ich weiß nicht, ob wir unser Onboarding Tool auf der Prämisse aufbauen sollen, aber.

B: Nein, das ist von mir so eine Macke, die ich seit einem Jahr oder sowas irgendwie führe, weil ich sonst, weil ich immer alles Mögliche gerne lesen möchte, aber die Zeit nicht habe, ja.

A: Na ich glaube, das gewöhnt man sich an, wenn man wissenschaftliche Papers liest, oder, weil Du liest den Abstract, dann die Conclusion ganz hinten und wenn das schon fertig ist, dann gönnst Du Dir den Rest auch.

B: Ja, weil tatsächlich, bei wissenschaftlichen Papern ist es so, die bauen so ein bisschen auf einander auf, das heißt, die Information, die Du vorne kriegst, die ist hinten auch wichtig, das ist bei Zeitungsartikeln nicht so, da ist so viel Geschwafel drin, dass, das kann kein Mensch ertragen, ja. Also wissenschaftliche Paper lese ich tatsächlich von vorne nach hinten, das ist.

#### A: OK, spannend.

B: Die sind, die sind ja auch vernünftig aufgebaut, aber ich sage nur, deswegen, weil ich oft so Zeitungen lese oder so, da gehe ich dann eben auch mal von unten nach oben, weil ich denke, es könnte einfach interessanter sein und ich habe auch nicht, bin nicht davon ausgegangen, dass die Punkte aufeinander aufbauen, also war es sehr willkürlich.

A: Hätte Dir eine kleine Zahl in den Punkten geholfen, wenn die nummeriert gewesen wären?

B: Pff, nee, hätte ich wahrscheinlich ignoriert.

A: OK. (lacht) OK.

B: Ja, aber weiß ich nicht, kann ich nicht so richtig sagen.

A: Grundsätzlich jetzt noch eine.

B: Aber ich glaube, Analyzing ist natürlich auch interessanter als Reading, nicht, erst mal.

A: Ja, aber die Art und Weise, die Hilfe so zu integrieren, mit diesem extrem aufdringlichen, komisch gestylten Fragezeichen, das dann aufpoppt und die die Hilfe dann in Teile dann unterteilt und dann in der App an sich die Hilfe direkt anbietet, ist das ein Format, das Du grundsätzlich annehmen würdest oder würdest Du erwarten, dass ein sauber strukturiertes Dokument mit Texten und Screenshots aufgeht?

B: Ahm, weil ich es gewohnt bin, würde ich tatsächlich wahrscheinlich dieses sauber geordnete, also da oben so eine Help-Funktion, die mir dann vielleicht was erklärt, auf der anderen Seite ist das hier ja, sagen wir so, also ich kenne es so nicht, es ist ja viel intuitiver, also wenn ich das jetzt als echte Hilfe erkannt hätte, Fragezeichen impliziert natürlich irgendwie, dass es eine Hilfe ist, aber wenn da Help stehen würde oder irgendwas Anderes, das mir sagt, hier erklären wir Dir, wie es geht, dann hätte ich da

wahrscheinlich auch früher draufgeklickt, also ich finde es eigentlich sehr schön, weil,

ja, ich bin es gewohnt, da oben so eine Hilfetaste zu haben und mich dann dadurch

570 meistens nicht so übersichtliche Kapitel durchzuarbeiten, von daher ist das ja viel

571 intuitiver, also es ist eigentlich sehr schön, es ist nur ungewohnt.

572 A: Hmhm, Hmhm. Ja und auch noch nicht ausgereift, also das ist auch der Grund,

warum wir diese Interviews jetzt machen. Es geht nun mal um die Idee grundsätzlich einzufangen, ob das OK wäre oder ob Du jetzt auch ganz ehrlich sagst, Leute, das

575 nehme ich nicht, das ist ein Scheiß.

576 B: Nee, ich habe es nur tatsächlich nicht als Hilfe erkannt, also das poppt dann hier

577 halt auf und dann ist Reading, Interacting, Analyzing und das müsste für mich einfach

578 sichtbar gemacht werden, nicht, dass das eine Hilfe ist. Sobald es dransteht, mache

579 ich es auch.

580 A: Hmhm. OK. Cool.

581 B: Ja.

A: Vielen Dank, dann sind wir schon bei der letzten Frage und die ist ganz offen. Gibt

es an dieser Stelle noch was, was von Deiner Seite aus zu dem Thema hinzuzufügenist?

585 B: Hmm. Jetzt muss ich ja echt kreativ sein.

586 A: Kannst auch sagen, Du hast alles gesagt, dann ist es auch OK.

587 B: Ja, ich glaube, ich habe alle Fragen, ich habe, glaube ich, auch genug erzählt, alle Fragen soweit beantwortet, ich denke, es geht ja hier hauptsächlich auch um 588 Datenvisualisierung, nee, alles, alles fein. Also ich glaube, man muss eben tatsächlich 589 berücksichtigen, wenn man sowas entwickelt, und das ist so die Schwierigkeit, wenn 590 591 man, eigentlich kann man es gar nicht berücksichtigen, aus was für einem Hintergrund kommen die Leute? Ich habe Dir das ja die letzten Tage schon gesagt, auch wenn ich 592 mit Triage arbeite, da jetzt haben wir eben gerade so ein anderes Programm, dieses 593 StarDrop evaluiert, Du kennst irgendwas von Excel oder Du kennst irgendwas von Pics 594 595 und das versuchst Du dann immer auf die Dinge zu übertragen.

#### 596 A: Ja.

B: Also inzwischen ist es ja so, dass wir von euch, also von Data Vision jetzt einige 597 Programme auch haben, ja, das heißt, so mit eurer Art zu arbeiten, die ist mir jetzt 598 nicht mehr so fremd, ja, die Art der Pfeile, wie sind die gemacht und so, wo kann ich 599 was finden? Mit rechte Mouse komme ich meistens dann doch nicht so weit, wie ich 600 es denke, das weiß ich, sodass man, dass man sich dann so langsam darauf 601 einschießt, aber oft hängt man eben tatsächlich sehr an der Art zu arbeiten, wie man 602 sie kannte und tut sich sehr schwer, neue Klicks zu machen, wie zum Beispiel mit 603 dieser Hilfe-Funktion, weil in jedem Programm ist es so, dass das eigentlich irgendwo 604 oben rechts in der Ecke ist. 605

606 A: Hmhm. Hmhm.

B: Ich glaube, das war auch mal so, als bei den Browsern, ich glaube, irgendwann war

das Menü mal oben, es war aber auch mal an der Seiet gewesen, das war so eine

ganz markante Umstellung vor 10 oder 15 Jahren, ich weiß es gar nicht, da wurde alles umgestellt da wurde alles quasi um 90 Grad gedreht.

#### A: Ja.

B: Und hat viele Leute vor irrsinnige Probleme gestellt, was natürlich, irgendwie rein logisch gesehen, kompletter Quatsch ist, nicht.

A: Ja, man ist eingetreten in den Pfaden, geht mir selbst genauso.

B: Ja, genau, nicht, und da ist man so ein bisschen eingefroren und man muss dann eben bereit sein, sich auf was Neues einzulassen, aber das fällt eben manchmal doch ein bisschen schwer. Und, ja, genau, ich glaube, dass das größte Problem ist, dass man hier mit dem Fragezeichen auch, dass man, dass ich aus keiner Anwendung, die ich kenne, also die jetzt mit Visualisierung zu tun hat, unten, rechts in der Ecke irgendwas Sinnvolles vermute, ja.

#### A: Hmhm.

B: Das ist eigentlich alles, das fängt oben links an und zieht sich dann entweder nach, läuft als Band oben drüber oder zieht sich nach unten, aber unten rechts vermute ich eben einfach nichts.

A: Das heißt, auch die Position?

B: Ja, die Position ist eigentlich das Entscheidende, für mich, wenn ich so darüber nachdenke. Hmhm.

A: OK. Ganz spannender Input für uns.

B: Genau. Ja, dann kann ich dann kann das ja der krönende Abschluss gewesen sein.

A: Cool, dann bedanke ich mich an der Stelle mal und stoppe jetzt das Recording.

B: Ja.



#### Interview 6

A: Ja. Danke, Dominik. So, dann würde ich mal direkt reinspringen in die erste Frage. Welchen Bezug haben Sie oder hast Du zu interaktiven Visualisierungen und wofür setzt Du diese ein, also?

B: Also meinst Du Datenanalyse und natürlich jetzt, um zum Beispiel (?) nutze, um irgendwelche Zelllinien und, ja, Backgroundinfos nachzuschauen?

#### A: Hmhm.

B: Ich weiß jetzt nicht, ich habe mir schon, ich habe mir ja die Fragen vorhin schon durchgelesen, ich bin jetzt nicht ganz sicher, ob zum Beispiel Excel-Grafiken oder GraphPad Prism oder so, als interaktiv zählen? Vielleicht könnten Sie mir kurz?

#### A: Ach so, ja.

B: Interaktiv noch definieren? (lacht)

A: Genau, also interaktiv wäre zum Beispiel, dass Du Sachen selektieren kannst, also zum Beispiel, wenn Du einen Scatter Plot hast, also Punktbalkenplot und Du dann verschiedene Items selektieren kannst oder Du kannst die einzelnen, die Ranges sozusagen anpassen, in Deinem Scatter Plot.

#### B: Hmhm.

A: Und dass sich verschiedene Visualisierungen miteinander vielleicht aktualisieren, das heißt, wenn Du in dem einen was machst, dass im anderen was passiert, das verstehen wir unter interaktiven Visualisierungen.

#### B: Hmhm.

A: Wir haben jetzt natürlich schon mehrere Interviews geführt und viele Nutzer, Nutzerinnen verwenden Excel, das heißt, das ist vollkommen legitim, dass wir das jetzt auch mit anführen, also das ist OK.

B: OK, passt, ja, Excel ist schon meine.

- A: Genau.
- B: Nummer 1 Standardanalyse.
- A: Genau. Excel für Standardanalyse schreibe ich auf.
- B: Ja, genau.

A: Standardanalyse. Hmhm. OK. Und dann noch vielleicht, Du hattest schon kurz gesagt, Du arbeitest seit 2017 bei Boehringer Ingelheim.

#### B: Genau.

A: Du bist Lab Scientist und was hast Du für eine Ausbildung bzw. was sind sonst Deine Aufgaben dort?

B: Ich habe Molecular Biotechnology studiert auf der FH Campus Wien, den Masterabschluss, genau.

A: Wie hieß das? Molecular?

- 38 B: Biotechnology.
- 39 A: Bio.
- 40 B: Genau.
- 41 A: Hmhm.

B: Und meine sonstigen Aufgaben, jetzt bei Boehringer? Also Experimente
durchführen, teilweise Planung mit meiner Chefin zusammen und dann auch
Auswertung und Analysen und, ja, auch Teilinterpretationen, jetzt nicht bis zum Ende,
oder meistens nicht bis zum Ende, weil dann doch die Chefin auch übernimmt, ja,
genau.

47 A: Hmhm. OK. Super, danke. Dann, im nächsten Teil vom Interview geht es jetzt erst

- 48 mal darum, ob Du die verschiedenen Visualisierungen schon mal gesehen hast, bzw.
- 49 damit vertraut bist oder bzw., wie gut Du damit vertraut bist? 5 ist sozusagen, Du
- 50 kennst sie sehr gut, Du hast sie vermutlich auch schon mal verwendet, und 1, Du hast
- sie überhaupt noch nicht gesehen, Du siehst sie jetzt gerade zum ersten Mal.
- 52 B: Hmhm.
- 53 A: Im zweiten Schritt werden wir Dich dann fragen, ob Du sie schon mal verwendet
- hast, das heißt, hier geht es jetzt wirklich erst mal darum, ob Du sie schon mal gesehen
- 55 hast, bzw. ob Du sie kennst?
- 56 B: OK. Ja. Also da erste 5, das, genau.
- 57 A: Ja.
- 58 B: Das Zweite, ja, 4 oder 5, ja.
- 59 A: Hmhm.
- 60 B: Das Pie Chart auch 4. Bubble Chart weniger eigentlich, also 2. Box Plot auch
- 61 weniger, 2 bis 3. Keine Ahnung, wie ihr, ob ihr Bis-Sachen wollt? Genau.
- 62 A: Ja, 2 bis 3 (?), ja.
- B: Violin Plot 2.
- 64 A: Hmhm.
- 65 B: Radar, also jetzt geht es darum, wenn ich es noch nie gesehen habe oder?
- 66 A: Genau.
- 67 B: Ich habe alles schon mal gesehen, prinzipiell.
- 68 A: Aber Du könntest.
- 69 B: Aber ich bin jetzt nicht super vertraut damit.
- A: Du könntest es lesen oder könntest Du es nicht lesen? Also würdest Du esinterpretieren können und verstehen?
- 72 B: Das Radar Chart jetzt zum Beispiel wahrscheinlich nicht.
- 73 A: OK, also dann wahrscheinlich eher eine 1, oder?

B: Dann wäre das eine 1. OK. Hmhm. Ja. Ja.

A: Ja.

- B: Heat Map, ah, pff, 3 bis 4. Tree 2 bis 3. 1.
- A: Hmhm.
- B: Oder keine Ahnung, (lacht) was das ist. Tree Maß 1.
- A: Hmhm.
- B: Und 3, kennt man aus der Politik, politische Bewegung. (lacht)
- A: Richtig, genau, genau. Wählerstromanalyse.
- B: Ja, genau. Und 2.

A: 2. Und haben wir jetzt noch irgendwelche Visualisierungen vergessen, die Du vielleicht in Deinem Alltag verwendest, die Du jetzt hier nicht gesehen hast?

B: Na eigentlich verwende ich die Bart Charts am meisten halt jetzt, ja, natürlich jetzt nicht unbedingt jetzt, wie ihr es gezeigt habt, sondern mit einzelnen, also nicht so Geteilte, aber das ist, glaube ich, egal.

- A: Hmhm. Genau, das ist egal, ja. Hmhm.
- B: Ja. Hmhm. Ja.

A: OK. Dann kommen wir zum nächsten Teil, da geht es jetzt darum, das sind nochmal die gleichen Visualisierungen und ob Du die für, schon mal erstellt hast? Das heißt, selber für Präsentationen für Deine Chefin vielleicht.

## B: Hmhm.

A: Und für Konferenzen oder sowas schon mal aufbereitet hast, also Daten in den Visualisierungen aufbereitet hast.

## B: Ja.

- A: Und da einfach ja oder nein, also (?) durchgeführt haben.
- B: OK. OK. Ja, die ersten drei, ja.
- A: Hmhm.
- B: Dann nein. Das in meinem Studium. (lacht)
- A: Studium. OK.
- B: Das könnte. (lacht)
- A: OK. Ich schreibe mal dazu, im Studium.
- B: Das ist schon sehr lange her.
- A: Du verwendest es, Du verwendest es also nicht auf Deiner Arbeit?
- B: Nein, verwende ich tatsächlich nicht.

- 107 A: OK. Hmhm.
- 108 B: Genau. Die ersten zwei nein, die Heat Map ja. Und den Rest nein.
- 109 A: Passt. Ach so, noch eines, genau.
- 110 B: Ah ja, genau. Hmm. Auch nicht.
- 111 A: Auch nicht. OK. Dann sind wir eigentlich schon direkt bei den Visualisierungstools.
- 112 B: Hmhm.
- 113 A: Und wir haben so die größten Visualisierungstools oder die bekanntesten
- rausgesucht, Du wirst vielleicht einen oder andere kennen und da ist jetzt die Frage,
- 115 hast Du das schon mal verwendet, also oder (?)?
- 116 B: Ich sage Dir, glaube ich, einmal, was ich schon verwendet habe.
- 117 A: Genau.
- 118 B: Ich habe Excel verwendet und Spotfire ein Mal, aber nur, um was anzuschauen,
- 119 was jemand anders ausgewertet hat, also ich bin jetzt nicht vertraut damit.
- 120 A: OK. Einmal zum Anschauen.
- 121 B: Genau und dann Inhouse-Tools wäre jetzt zum Beispiel MegaLab, das haben wir
- zur Visualisierung von unser IC50s und genau, zum Kurvenzeichnen quasi.
- 123 A: Hmhm.
- B: Ja. Das regelmäßig natürlich.
- 125 A: Hmhm. Und Ordino, nehme ich an, oder?
- B: Ja, genau. Ordino, ach so, und, ja, andere Tools aus diesem Pool jetzt auch noch gelegentlich.
- 128 A: Genau. OK.
- 129 B: Aus der Tool Box.
- 130 A: (?) nein, so. OK. Danke. Und wenn Du Dich jetzt daran erinnerst, ist vielleicht auch
- 131 schon ein bisschen länger her, wie hast Du denn diese Tools gelernt, also das heißt,
- 132 hast Du das als Selbststudium zum Beispiel in Trial & Error durchgeführt oder hattest
- 133 Du Kollegen, die Dir das erklärt haben, hast Du Tutorials gelesen, Videos angeschaut
- 134 oder hattest Du sogar einen Kurs an der Uni damals.
- 135 B: Hmhm.
- 136 A: Oder vielleicht auch vom Arbeitgeber bezahlt?
- 137 B: Hmhm.
- 138 A: Um halt diese Tools zu erlernen?
- 139 B: Also von bis. Excel in der Schulzeit noch und dann schon Learning by Doing.
- 140 A: (?)
- 141 B: Und von Kollegen, also dieses eigentlich von bis.

- B: Und einen Kurs habe ich auch schon mal gemacht.
- A: Hast Du auch, ja, Hmhm.
- B: Also bei Boehringer oder über Boehringer, genau.
- A: Hmhm.
- B: Ordino habe ich vom Thomas gelernt.
- A: Das ist (?).
- B: Also in den Inhouse-Kursen quasi.
- A: Das wäre auch Kurs, ja, genau, also.
- B: Ja, wäre auch Kurs. Ahm.
- A: Inhouse. Hmhm.
- B: Genau. Ja, und MegaLab, dieses Inhouse-Tool, ist eigentlich Selbststudium und Hilfe von KollegInnen.
- A: Selbststudium. Das ist MegaLab?
- B: MegaLab, genau.
- A: MegaLab und Hilfe von Kollegen, hast Du gesagt.
- B: Ja, genau.

A: MegaLab, genau. Das heißt, Videos und Tutorials hast Du an sich jetzt nicht geschaut?

- B: Lass mich ganz kurz überlegen?
- A: Ich meine.

B: Nein, eigentlich nicht, ich bin mehr so der Typ, ich gehe tatsächlich mal lieber zu jemandem und spreche da drüber und vor allem ist es dann eher auch anwendungsorientiert, weil dann sage ich, ich habe das und das Experiment, wie machst Du das?

## A: Hmhm.

B: Und wenn ich das google, dann, glaube ich, also ich würde mir ein bisschen schwerer tun, das Spezifische zu finden, deswegen, ja, ich gehe lieber zu Kollegen. (lacht)

A: OK. Ich schreibe das mal noch als Kommentar dazu auf, weil das ist ein sehr wertvoller Kommentar.

B: Hmhm.

A: Spricht lieber mit Kollegen und da anwendungsorientiert.

B: Ja, genau. Ja, ich schätze einfach diesen Austausch, also ich merke es mir dann auch besser und, genau.

- 177 A: OK. Der Austausch, ja, passt.
- 178 B: Ja.

A: Na das klingt sehr gut, guter Kommentar. Dann, wenn Du Dich vielleicht hier bei
den strukturierten Kursen und Lehrveranstaltungen eventuell, daran zurückerinnerst,
welche Materialien wurden dafür eingesetzt? Also was hat jetzt zum Beispiel (?) oder
bei Deinem Excel, (?), bei Deinem Excel-Kurs, was haben sie da eingesetzt, Video,
Text, Screenshots, gab es einen moderierten Workshop zum Beispiel oder wie war
das?

- B: Ja, bei Ordino wird dann eigentlich durchgeführt, also, genau, ein Demo, würde ichmal sagen, simuliert, dem, ja. Demo.
- 187 A: Demo. Hmhm.
- 188 B: Und dann, genau, und sonst, beim Excel-Kurs haben wir Unterlagen, also
- 189 Materialien bekommen und dann auch einfach Fallbeispiele besprochen, kombiniert 190 mit Frontalunterricht.
- 191 A: OK. Frontalunterricht.
- 192 B: Genau.
- 193 A: Ja.

B: Bei MegaLab gibt es dann noch, erstens kann man natürlich den Support
kontaktieren, wenn man was braucht oder es gibt auch auf der Startseite so
Lernvideos, das ist alt für uns sehr hilfreich, weil das MegaLab halt tatsächlich sehr
anwendungsorientiert ist, also da macht es dann auch Sinn, mal ein Video zu schauen,
weil das genau auf uns zugeschnitten ist.

A: Videos und, OK, das heißt, de Lernvideos sind, also das ist ja ein Boehringerinternes Tool?

- 201 B: Hmhm.
- 202 A: Und die wurden dann wahrscheinlich auch von einem, von, ich weiß nicht, ob es die
- 203 gleiche Abteilung ist, die MegaLab programmiert?
- B: Ja, genau, vom Support Team werden die gemacht, genau, das sind dann so.
- 205 A: OK.
- B: 30 Sekunden bis 2-Minuten-Videos oder so.
- 207 A: Hmhm.
- 208 B: Ganz, ganz kurze eigentlich zu, ja.
- A: Hmhm. Genau und ansonsten gibt es noch einen Support, den man anschreibenkann der wie?
- B: Ja, genau.
- 212 A: Also, OK.

B: Und da kann man dann auch recht spontan einfach telefonieren, den Screen sharen und da ist recht praktisch, weil die halt auch alles sehen dürfen, alles, was jetzt confidential ist oder so, weil die eh von Inhouse sind, das heißt, die wissen auch den biologischen Hintergrund teilweise Bescheid und das ist echt super, also man spricht da jetzt nicht mit IT, die dann keinen Plan haben, mein Kombiexperiment, wie das funktioniert de so.

#### A: Hmhm.

B: Genau, das ist, habe ich erst vor Kurzem wieder genutzt und war sehr erfreut darüber. (lacht)

A: Ah. Ja, ist eh gut. Ich glaube, dann hast Du schon einen Teil von dieser anderen Frage hier erklärt. Das heißt, welche Probleme sind dabei aufgetreten, was war schwierig, was hat gut funktioniert, also bei MegaLab hast Du jetzt gerade schon gesagt.

B: Hmhm.

A: Dass halt der direkte Austausch hilfreich ist.

B: Ja, genau. Du meinst jetzt bei den Kursen oder?

A: Genau, also.

- B: Wenn jemand so, ja.
- A: das bezieht sich eigentlich, aber kann auch generell so.

B: Genau. Ja, ich finde es immer ein bisschen problematisch, Frontalunterricht oder einfach viele Beispiel, die jetzt nicht unbedingt anwendungsorientiert sind, auf meine Sparte, also wenn man jetzt irgendwo extern einen Kurs bucht, zum Beispiel zu Excel, dann kommt da halt sehr viel generelles Zeug daher, wo ich mir dann denke pah, wo würde ich das anwenden wollen können (lacht) und dann nehme ich nicht so viel mit, als wie bei Ordino zum Beispiel, wo es sehr spezifisch ist, eben (?)-Suche, Mutationen, etc., oder bei MegaLab, wo ich ja auch muss, also ich muss das ja auch in MegaLab dokumentieren und meine Werte dann da rausziehen, aus der Software, das heißt, ich muss mich damit beschäftigen und, ja, ich bin auch sehr froh, dass das eigentlich eine gute. Übersichtliche Software ist, weil, ja, ist ein wichtiges Tool für uns, das heißt, ein wichtiger Support ist da, also guter Support ist auch sehr wichtig. Genau. Also einfach, es muss anwendungsspezifisch sein, mit Beispielen untermauert sein und, ja.

A: Hmhm. Ich schreibe mal auf, muss anwendungsspezifisch sein.

B: Genau.

A: Ja.

B: Was bei Kursen, finde ich, auch noch wichtig ist, dass man eigentlich mit Leuten, vielleicht aus der gleichen Spate drin sitzt, eben dann ist as Anwendungsorientierte natürlich auch leichter umzusetzen, weil wenn ich da jetzt mit lautet Wirtschaftern drinnen sitze, die können natürlich mit meinen fachspezifischen Experimenten oder so, nichts anfangen und dann möchte ich mich auch nicht jetzt zehn Minuten mit der Vortragenden darüber unterhalten, weil das die anderen einfach langweilt oder die, die

- 253 nichts damit anfangen können. Das heißt, das ist, das habe ich von dem Excel-Kurs
- eigentlich mitgenommen, da war für mich nicht so hilfreich oder so lehrreich, weil das
- sehr generell war und viele, mehr so Buchhaltungsrichtung vielleicht war, genau.
- 256 A: Ja, OK, aber das habe ich aufgeschrieben. Hmhm. Genau.
- 257 B: Hmhm.
- 258 A: Ja, das ist also, das ist, glaube ich auch, dass das natürlich sehr viel spezifischer
- 259 wird und natürlich auch viel anwendungsorientierter, wenn man halt aus der gleichen
- 260 Domäne kommt.
- B: Ja. Ja. Genau.
- 262 A: Als wenn man halt sehr viele verschiedene Domänen hat.
- 263 B: Genau, weil da weiß man schon, OK, ich frage jetzt eine Frage, die betrifft
- 264 wahrscheinlich eh mehrere.
- A: Genau.
- 266 B: Ja.
- A: Genau Hmhm. Super. Genau. Dann generell die Frage, das kann man jetzt, glaube ich auch relativ schnell durchgehen.
- 269 B: Hmhm.
- 270 A: Welche Ressourcen, Webseiten, Plattformen verwenden Sie, wenn Sie Hilfe
- 271 brauchen? Also das heißt, wenn zu sozusagen Hilfe suchst, dann hast Du gesagt, Du
- 272 liest wahrscheinlich Bild und Text, also sprich diese (?)-Dokumentation. Videos hast
- 273 Du Dir vermutlich über MegaLab angeschaut, hast Du gesagt?
- B: Ja, genau. Hmhm.
- A: Und die Frage ist jetzt noch, bei interaktiv oder InLab-Hilfe, da geht es darum, diese
- 276 Programme haben ja meistens so ein Fragezeichen.
- 277 B: Ja.
- 278 A: Wo man draufklicken kann und dann kriegt man halt mehr Informationen, entweder
- 279 generell oder zu einer bestimmten Funktion. Verwendest Du sowas?
- 280 B: Ahm, eigentlich jetzt auch nur bei MegaLab, hmm, sonst eigentlich nicht so.
- 281 A: MegaLab, OK, und ansonsten eher nicht, also dann würdest halt, wie gesagt?
- 282 B: Nein, ich finde, ich finde, da verbringt man ziemlich schnell mal viel Zeit damit und
- 283 dann findet man eh wieder nicht das Abwendungsspezifische, was man braucht.
- 284 A: OK.
- B: Also oft ist das so. (lacht)
- 286 A: Ja. Ja. Na, das ist interessant. Hast Du noch andere Hilferessourcen,
- 287 Hilfematerialien, die Du verwendest? Also Mitarbeiter sozusagen?
- B: Ja, genau, also KollegInnen. Nein, sonst.

A: Nicht. Dann würde ich Dir jetzt den nächsten Teil, drei Seiten, wahrscheinlich vorstellen.

#### B: Hmhm.

A: Ich frage mal, kennst Du die vom Namen jetzt auf jeden Fall schon mal? Delta Viz Catalogue ist die Erste, From Data to Viz und Dat Viz Project ist die Dritte.

B: Irgendwie kommt es mir bekannt vor, aber ich würde jetzt mal sagen, nein.

A: OK. Ich mache die mal auf und dann können wir die kurz anschaue.

#### B: Ja.

A: Also der Data Viz Catalogue oder Data Vizualisation Catalogue ist eine Liste von allen möglichen Visualisierungen, Du siehst hier Box Plot, alles Kartografie zum Beispiel, Line Graph, Bart Chart und so weiter und hier zum Beispiel auch Parallel Coordinates Plot, da, wo Du gesagt hast, den kennst Du zum Beispiel noch nicht.

#### B: Hmhm.

A: Wenn ich jetzt da draufgehe, dann erhältst Du zum Parallel Coordinates Plot, eine kurze Beschreibung, was der ist.

#### B: Hmhm.

A: Und wie letztendlich der Parallel Coordinates Plot aufgebaut ist und wie die Daten dafür strukturiert sein müssen, damit man halt so einen Parallel Coordinates Plot bauen kann und wie man den am Ende auch lesen kann. Und das Interessante ist hier, dass es Dir die Möglichkeit gibt, zum einen sagt, für was kann man das verwenden, Vergleiche, Comparison, Relationships und welche Visualisierungen sind ähnlich zu diesen Parallel Coordinates und wenn Du das generieren würdest, wenn Du jetzt programmieren würdest, welche Libraries könntest Du verwenden sozusagen oder den Code direkt dann rauskopieren, um halt diese Visualisierung zu erstellen.

#### B: Hmhm.

A: Das ist die eine Seite, die hast Du wahrscheinlich noch nicht gesehen, oder?

B: Nein, kenne ich nicht, aber ist cool.

A: Mache ich hier mal nein. Nein oder, Danke, Daniela, dann schreibst Du mit. Genau, dann haben wir From Data to Viz, ist ähnlich, aber anders strukturiert, das heißt, hier fängt man an mit der, welche Daten man hat, also man sagt sozusagen, ich habe numerische Daten, kategorische Daten und so weitere und dann hat man so einen Entscheidungsbaum, wo man dann sagen kann, ich habe numerische, eine numerische Variable und dann sehen, OK, dann macht sich eine, dann kann ich ein Histogramm verwenden oder einen Density Plot, wenn ich zwei Variablen habe, kann ich, sind die geordnet oder nicht geordnet und dann kann ich dementsprechend mir die anderen Visualisierungen sozusagen anschauen, bzw. komme ich dann direkt zu den Visualisierungen, die für meine Daten passend sind. Und wenn ich jetzt zum Beispiel auf, sagen wir mal hier, auf den Box Plot klicke, dann kriege ich auch hier wieder eine kurze Beschreibung, was sind die typischen Fehler sozusagen, in dem Fall, die beim Box Plot gemacht werden, bzw. auch bei der Interpretation von einem

329 Box Plot und dann komme ich auch hier wieder zu Code-Beispielen, dass ich das halt

direkt in (?) oder R verwenden kann, diese Visualisierungen. Hast Du das schon mal

- 331 gesehen?
- 332 B: Nein.

A: Hmhm. Und das Dritte ist natürlich auch wieder ähnlich. Hier hat man auch wieder 333 334 die verschiedenen Visualisierungstypen und wenn ich jetzt hier, ich nehme jetzt einfach mal den (?) Plot auswähle, dann auch hier wieder eine Einleitung, dann auch 335 336 wieder die Kategorie sozusagen, wie sind die Input-Daten und was man hier sehr schön sieht, ist, dass man verschiedene, also wirkliche, echt Beispiele sieht, sodass 337 man sieht, welche Varianz sozusagen ein Violin Plot haben kann, dass man sieht, OK; 338 der kann zum Beispiel so ausschauen oder er kann zum Beispiel auch so ausschauen, 339 aber das zählt halt alles als Violin Plot, sodass man dann das auch mal als echte 340 Anwendung sozusagen gesehen hat. 341

342 B: Hmhm.

A: Und in verschiedenen Visualisierungsformen. Hast Du das schon mal gesehen oder?

- 345 B: Nein, kenne ich auch nicht.
- 346 A: OK. Wir können Dir die Links auch nach dem Interview gerne noch zuschicken,
- 347 dann kannst Due s Dir vielleicht gerne nochmal selber.

348 B: Ich habe es schon parallel aufgemacht. (lacht)

A: Ach so, gut. Ja, ja, Du hast ja den (?) schon, genau, dann brauchen wir es nicht nochmal schicken. Genau.

351 B: Ja.

A: Na, OK. Sehr interessant. Dann, ich glaube, das haben wir jetzt auch schon zur
 Genüge geredet, aber vielleicht, wenn Du noch was ergänzen möchtest, welche
 zusätzlichen Ressourcen würden Sie sich wünschen, was fehlt zum Beispiel bei Deiner
 tädlichen Arbeit mit Visualisierungen ietzt?

B: Hmm. Hmm. Ja, also nachdem ich sehr viel mit Excel arbeite und jetzt, weil Du mir 356 gerade diese kurzen Beschreibungsboxen gezeigt hast, auf diesen Homepages, das 357 ist schon sehr praktisch, finde ich, und auch mit Beispielen, und in Excel ist jetzt, also 358 359 erst gestern wieder, sitze ich da und dann kriegt es mein Kollege wieder ab, weil ich dann wieder so (lacht), auch, wie visualisiere ich das jetzt, wie macht das denn am 360 meisten Sinn, dass das wirklich gut rüberkommt oder, ja, und dann hänge ich irgendwie 361 und dann guatschen wir kurz darüber und dann geht es auch meistens, aber da wäre 362 363 es vielleicht gar nicht so schlecht, weil Excel halt doch irgendwie, ja, Du kriegst halt oben diese Vorschläge, wie es ausschauen kann, aber ob das jetzt Sinn macht oder, 364 ja, die Gruppierungen, etc. 365

- 366 A: Hmhm.
- 367 B: Hat man nicht so viel Info dazu, genau.
- 368 A: Hmhm.

B: Also diese Beschreibungen, das war jetzt schon sehr anschaulich und schön ich, also optisch ansprechend.

A: Hmhm. Ja, ich glaube, dass, ich meine, das hat Excel jetzt nicht integriert, aber Du kannst ja dann in Zukunft, könntest Du ja auf die Seiten gehen und Dir das anschauen.

B: Hmhm. Zum Beispiel, ja.

A: Und dann die Visualisierung (?). Aber das stimmt, das haben schon mehrere Nutzer angemerkt, dass Excel gerade eben diese kurzen Beschreibungen sozusagen, dass die (?).

B: Ja. Ja. Weil manchmal fehlt einem auch die Vorstellungskraft irgendwie, wenn man jetzt was Neues visualisiert oder irgendwie eine neue Komponente dazukommt, denkt man sich, wie kann das ausschauen, also irgendwie, ja, man kann es sich nicht so ganz vorstellen noch.

A: Genau. Hmhm.

B: Hmhm.

A: Genau und dann ist die Frage noch, das geht jetzt schon in Richtung unser Onboarding. Wann und wo würdest Du Dir Hilfestellung Erhalten wollen? Also würdest Du erwarten, dass das direkt im Tool selber vorhanden ist, so, wie wir jetzt gerade gesagt haben, in Excel oder ist es OK, wenn das so als externe Webseite ist und als zweite Unterfrage sozusagen, soll das on demand, also wenn Du das möchtest auch und Du das aufrufen möchtest, erscheinen, oder soll das ein Mal, wenn Du das erste Mal die Applikationen aufmachst, dann kriegst Du halt die Hilfestellung dann beim nächsten Mal geht das Programm davon aus, dass Du das schon weißt und dann wird das nicht mehr angezeigt?

B: Hmhm. Also ich glaube, on demand macht mehr Sinn.

A: OK.

B: Weil es könnte ja sein, dass man das Tool nicht regelmäßig verwendet und sich das dann eben nicht merkt.

#### A: Hmhm.

B: Und sonst, pah, direkt im Tool oder extern? Hmm. Ich glaube, ich könnte mit beidem gut leben, ich glaube, ich müsste es ausprobieren.

A: Hmhm.

B: Könnte ich jetzt nicht so genau sagen. Weil wenn jetzt dauernd, also wenn im Tool ständig zum Beispiel PowerPoint hat diese neue Funktion, man macht eine Grafik auf eine Seite und auf einmal poppt rechts dieses, wie wollen Sie das visualisieren und gestalten, Design, bla, ultranervig.

A: Hmhm. Hmhm.

B: (lacht) Und so möchte ich es zum Beispiel nicht.

A: (lacht) Das kann ich verstehen, das ist bei mir auch so, wenn ich eine Präsentation mache.

408 B: Ja.

409 A: Dass das dann immer im falschen Moment kommt.

410 B: Genau.

411 A: Ja. Und nervig. OK. Dann danke dafür. So, jetzt kommen wir eigentlich schon zu 412 unserem Onboarding-Teil, den wir in unserem Projekt jetzt gemacht haben.

413 B: Hmhm.

A: Damit sind wir auch schon fast durch mit dem Interview und für Dich vielleicht erst 414 mal, was verstehen wir unter Onboarding, das haben wir noch gar nicht gesagt. Also 415 wir verstehen Onboarding als einen Prozess, der die Benutzer beim Lesen, 416 Interpretieren und Extrahieren von Informationen aus visuellen Darstellungen von 417 Daten unterstützt. Das heißt sozusagen, wir wollen, dass der Benutzer oder die 418 Benutzerin selbst sozusagen die Visualisierung verstehen kann und dann 419 dementsprechend auch die richtigen Schlüsse daraus ziehen kann. Und was wir jetzt 420 gemacht haben, ist, dass wir einen Klick-Prototypen erstellt haben und den würden wir 421 jetzt gerne. Du hast es, den Interviewleitfaden jetzt eh schon, das heißt, wenn Du das 422 aufmachen könntest, und wir jetzt das Screen (?) umdrehen könnten, dass wir Dein 423

- 424 Fenster mit dem Prototypen aufnehmen könnten.
- 425 B: Hmhm.

426 A: Und dann würden wir Dich bitten, einfach mal durch den Klick-Prototypen zu gehen

- 427 und zu erzählen, was Du so siehst und wie Dir die Hilfestellung da gefällt?
- 428 B: OK.
- 429 A: Ich werde mal meinen Screen-Cat stoppen.
- 430 B: Ja. So, ich schaue, dass ich den Screen so hinkriege. (lacht) So da. So. Müsste 431 gehen, oder?
- 432 A: Hmhm. Ja. Wir sehen Deinen Bildschirm. Danke schön. OK. Dann vielleicht kannst
- 433 Du laut denken und mal erzählen, was Du siehst bzw. dann die Hilfestellung?
- 434 B: OK. Also zwei Graphen, zwei Scatter Plots, also Verteilung von Daten. Hmm. Wir
- 435 schauen mal links jetzt, OK, ich habe Achsen, die ins Negative gehen, die x-Achse ist
- 436 überhaupt nur im Negativen. OK. Und rechts fällt mir auf, die y-Achse ist nicht auf 0,
- 437 fängt bei 60 an. Ahm. Gut. Und jetzt hier, gell?
- 438 A: Hmhm.
- 439 B: Soll ich dann einfach mal?
- 440 A: Genau.
- 441 B: Reading, Interacting. Hmhm. OK. Also ich lese das jetzt.
- 442 A: Ja.
- 443 B: Dann denke ich mir, so, was, wieso sind da jetzt quer rüber, OK, da muss ich mich
- 444 durchklicken, verstehe. (lacht) OK, da gibt es keine Beschriftung. Hmhm. OK, ich muss
- 445 mir das jetzt nicht merken oder wiedergeben oder?

#### A: Nein, nein, Du musst nicht.

B: Inhaltlich? (lacht)

A: Außer, Du möchtest gerne was dazu sagen?

B: Nein, danke. (lachen) Hmm. OK. Dann, was passiert hier? Ach so, ich klicke das dann so nochmal durch. Aha. OK. Ist mir jetzt fast bisschen zu viel.

A: Die Information oder zu viel (?).

B: Ja, also quasi hier habe ich eine Option, um mich durchzuklicken und dann kann ich die Achse aber auch nochmal nutzen, also diese hier.

#### A: Hmhm.

B: OK. Dann komme ich in den zweiten Punkt. Das ist jetzt, da müsste ich hier wieder zurück, gell, ah ja, OK. Hmhm.

A: Genau.

B: Interacting, ja. OK. Also ich frage mich, wieso mache ich das und das?

- A: OK. Das heißt, Du würdest Dir nur eine.
- B: Weil es ist eigentlich das Gleiche.
- A: Genau, Du.

B: Im Endeffekt führt es mich zu dem gleichen Ziel, ob ich mich jetzt an den Punkten orientiere oder an den Beschriftungspunkten, also Zahlenpunkten.

- A: Hmhm. Welche Option wäre Dir lieber, also?
- B: Hmm.
- A: Was würdest Du eher?

B: Also vermutlich macht das hier mehr Sinn, wenn man mir im Bild was zeigen will, wenn man meine Aufmerksamkeit auf was lenken möchte.

#### A: Hmhm.

B: Man könnte, aber in dem Fall ist es ja sowieso, dass es dann dort aufpoppt und ich finde, ich glaube, dieses mit den Zahlen durchbeschriften führt Dich ja in einer bestimmten Reihenfolge durch das Bild.

A: Hmhm.

B: Aber jeder User hat ja irgendwie doch einen anderen Blick für die Dinge und fängt vielleicht woanders an.

A: Hmhm. Genau.

B: Also das ist so, wie ich kriege eine Grafik und lese ich zuerst die Legende dazu oder schaue ich zuerst die Grafik an.

A: Hmhm. Genau.

- 480 B: So in der Art und für mich ist das persönlich tatsächlich ein bisschen ein Stress,
- 481 dass ich mich jetzt so durchorientieren muss.
- 482 A: OK. Wie, was würdest Du dann vorschlagen, also wie wäre es Dir lieber?
- B: Hmm. Vielleicht, dass hier die Punkte sind, wo es Erklärungen gibt, aber keineNummerierung festgesetzt ist.
- 485 A: Ah, OK.
- 486 B: Das heißt, da muss ich nicht in der Reihenfolge durchgehen, sondern ich kann mal
- 487 das Bild erfassen, OK. Und wo würde ich jetzt anfangen, mal was nicht zu verstehen,
- dann fange ich dort an zu klicken, zum Beispiel.
- 489 A: Hmhm. Das ist ein sehr guter Hinweis. Hmhm. OK.
- 490 B: Genau. Weil so muss ich mich jetzt ja eben an dieser vorgegebenen Reihenfolge
- 491 eigentlich orientieren und vielleicht käme diese Frage erst später für mich auf oder ist
- 492 jetzt gar nicht so wichtig, genau. Ahm. Ja. Genau. Also hier geht es jetzt darum, wirklich
- 493 ich kann diese Grafik nutzen, einfach mal, um das zu verstehen, was hier geplottet ist
- 494 und dann eben, um drüber zu gehen und das anzuschauen.
- 495 A: Hmhm.
- 496 B: Und dann kann ich noch, ahm, OK, dann geht es mehr ins Detail eigentlich.
- 497 A: Genau. Also bei der Analyse geht es um das Verständnis von den Daten.
- 498 B: Hmhm. OK.
- A: Genau. Macht das, macht diese Kategorisierung in die drei Kategorien Sinn, ausDeiner Sicht?
- 501 B: Hmm. Ja. Reading ist für mich ein bisschen verwirrend, ehrlich gesagt. Reading.
- 502 Aber ich weiß jetzt nicht, vielleicht, weiß ich nicht, ob es ein besseres Vokabel dafür
- 503 gibt, quasi.
- 504 A: OK. Hmhm.
- 505 B: Um das eindeutiger zu machen.
- 506 C: Wäre Reading the Chart für Dich einfacher zu verstehen?
- 507 B: Hmm. Ja, es geht jetzt, Reading ist für mich fast jetzt schon wieder so ein bisschen,
- 508 es schaut schon bisschen detaillierter aus und geht fast ins Analyzing, guasi ich muss
- 509 das jetzt lesen, unter Anführungszeichen, um das auch zu verstehen, gleich,
- 510 gleichzeitig, und nicht nur erfassen, es ist jetzt nur für so ein Mal sehen, wäre für mich,
- aha, OK, es ist jetzt ein Scatter Plot, also (?) die Achsen, das ist türkis, (lacht) das ist
- 512 für mich so, aha, OK, kurz gesehen. Aber vielleicht, ja, ist es jetzt nur meine 513 persönliche.
- 515 personner
- 514 A: Hmhm.
- 515 C: Hmhm. OK.
- 516 B: Weiß ich nicht, ob der Kommentar si wertvoll ist (lacht).

- A: Das auf jeden Fall.
- B: Genau. OK. Ja. Passt oder? Von der Seite noch.
- A: Und noch eine Frage.
- B: Bitte, ja.
- A: Ja, also das passt auf jeden Fall, wenn Du das jetzt noch schließt.
- B: Ja.

A: Die, das, der Hilfebutton sozusagen, ist der prominent genug für Dich und so wie das funktioniert mit dem Aufklappen und so weiter, ist das auch für Dich OK oder hättest Du Dir da was anderes gewünscht, an der Stelle, wenn Du Dir Hilfe suchst oder Hilfe aufmachen möchtest?

B: Nein, ich glaube, das ist gut. Also der ist definitiv sichtbar, wäre jetzt zum Beispiel, nein, glaube ich, ist gut, Rot ist wahrscheinlich, oder Rosa.

A: Ja, genau. (lacht)

B: Auch OK. Ja. Nein, passt. Hmhm.

A: Na dann. Genau. Super, danke. Danke für Dein Feedback, für die Anregung.

#### B: Gerne.

A: Dann haben wir eigentlich nur noch eine Frage, bei unserem Interviewleitfaden, also wir brauchen jetzt nochmal den Bildschirm zu wechseln. Und da ging es jetzt eher darum, ob Du noch irgendwas generell zu unserem Interview oder zu Onboarding oder zur Visualisierung sagen möchtest, also ob Du generell noch irgendwelches Feedback hast, das Du gerne loswerden möchtest?

B: Hmm. Pah, fällt mir jetzt so auf die Gache, also kein Feedback mehr ein, aber was euer, was ist jetzt dann im Endeffekt das Ziel von eurem Interview oder von der Umfrage.

A: Genau, also was wir jetzt gemacht haben, ist ja zum einen erhoben, sozusagen, was die Experten, in dem Fall Du und die anderen, die wir gefragt haben, für Tools verwenden und wie firm, sage ich mal, sie mit Visualisierungen sind.

#### B: Hmhm.

A: Und auf der anderen Seite haben wir unseren Prototypen jetzt evaluiert, also sprich von euch testen lassen und wir werden jetzt den Prototypen mit dem Feedback nochmal überarbeiten und das Ziel ist es sozusagen, dass i die Visualisierungen, die wir auch in Ordino drin haben, mit reinzubringen, sodass man dann als Benutzer, wenn man zum Beispiel Ordino mit Visualisierung verwendet und man weiß nicht mehr, wie man das interpretieren soll oder man ist neu in Ordino, dass man dann eben die Hilfe aufrufen kann und das Chart selbstständig lesen, ah, um halt, ja, genau, das zu verstehen.

B: Hmhm.

- A: Oder richtig zu interpretieren, ja. Genau, also das wäre so das Ziel oder das weitere
- 555 Ziel für dieses Projekt, an dem wir arbeiten.
- 556 B: Ja, einbinden in Ordino. Hmhm.
- 557 A: Genau.
- 558 B: Na super.
- 559 A: Genau, ja, dann danke ich Dir oder danken wir Dir auf jeden Fall für Deine wertvollen
- 560 Kommentare und auch für Deine Zeit, die Du, genau, mit uns erbracht hast.
- 561 B: Sehr gerne zugehört habt.
- 562 A: Und, genau. (lachen) (?) Ja.
- 563

# F.1 Screenshots of VA tool for HTS data analysis



# F.1.1 Hexbin plot onboarding

Figure F.1: Onboarding Message Reading the chart (1)

# F.1.2 Scatterplot onboarding



Figure F.2: Onboarding Message Reading the chart (2)

# F. MATERIAL OF DESIGN STUDY BIOMEDICAL R&D



Figure F.3: Onboarding Message Reading the chart (3)



Figure F.4: Onboarding Message Reading the chart (4)

# F. MATERIAL OF DESIGN STUDY BIOMEDICAL R&D



Figure F.5: Onboarding Message Reading the chart (5)



Figure F.6: Onboarding Message Reading the chart (6)



Figure F.7: Onboarding Message Interacting with the chart (1)



Figure F.8: Onboarding Message Interacting with the chart (2)

# F. MATERIAL OF DESIGN STUDY BIOMEDICAL R&D



Figure F.9: Onboarding Message Interacting with the chart (3)



Figure F.10: Onboarding Message Reading the chart (1)

# F. MATERIAL OF DESIGN STUDY BIOMEDICAL R&D



Figure F.11: Onboarding Message Reading the chart (2)



Figure F.12: Onboarding Message Reading the chart (3)



Figure F.13: Onboarding Message Interacting with the chart (1)



Figure F.14: Onboarding Message Interacting with the chart (2)



Figure F.15: Onboarding Message Interacting with the chart (3)

# Curriculum Vitae





# Personal Data

# Work Experience

Since 2015	Researcher and Lecturer
	at St. Pölten University of Applied Sciences, Austria
	Department Media and Technologies
	$\rm IC \backslash M / T$ Institute of Creative Media Technologies
	Media Computing Research Group
2017 to 2018	Lecturer
	at the University of Applied Sciences Upper Austria
2014 to 2021	Self-employed graphic designer
2013 to 2015	Assistant
	at St. Pölten University of Applied Sciences, Austria
	Department Media and Technologies
	$\rm IC \backslash M / T$ Institute of Creative Media Technologies
	Media Computing Research Group

# Publications

Please see:	<pre>https://icmt.fhstp.ac.at/en/team/christi na-stoiber</pre>
	<pre>https://scholar.google.at/citations?user= 0b65cYsAAAAJ&amp;hl=de</pre>