



Visualization of Cultural Heritage Collection Data

MASTER'S THESIS

submitted in partial fulfillment of the requirements for the degree of

Master of Science

in

Media and Human-Centered Computing

by

Michael Fischer, BSc

Registration Number 0540696

to the Faculty of Informatics

at the TU Wien

Advisor: Univ.Prof. Mag.rer.soc.oec. Dr.rer.soc.oec. Silvia Miksch

Assistance: Davide Ceneda, MSc
Dr. Paolo Federico

Vienna, 20th August, 2020

Michael Fischer

Silvia Miksch



Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar.
The approved original version of this thesis is available in print at TU Wien Bibliothek.

Erklärung zur Verfassung der Arbeit

Michael Fischer, 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, 20. August 2020

Michael Fischer



Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar.
The approved original version of this thesis is available in print at TU Wien Bibliothek.

Acknowledgements

I want to thank my advisors Silvia Miksch, Paolo Federico, and especially Davide Ceneda, for supporting me throughout the whole process of this thesis.

Furthermore, I want to thank all the user study participants for taking the time to help me during the evaluation and for their valuable feedback.

Finally, my thanks go out to my family, my friends, and my wife for the emotional support and constant input, which was vital in order for me to finish this project.



Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar.
The approved original version of this thesis is available in print at TU Wien Bibliothek.

Kurzfassung

Museen und andere kulturelle Einrichtungen haben vermehrt Interesse daran, ihre Sammlungen öffentlich zugänglich zu machen. Dazu bieten sie oftmals web-basierte Interfaces an, die es Benutzer*innen erlauben, ihre Sammlungen ("Cultural Heritage Data Collections") zu durchsuchen. Leider erfordern diese Interfaces in vielen Fällen Spezialwissen über die gespeicherten Daten, ohne das die Sammlung nicht effizient nutzbar ist. Das macht das Verwenden dieser digitalen Sammlungen für Gelegenheitsbenutzer*innen unattraktiv.

Informationsvisualisierung kann dieses Problem lösen und komplexe Daten auch für Gelegenheitsbenutzer*innen verständlich und zugänglich machen. Überraschenderweise werden, trotz ihrer Effektivität, Visualisierungen aber nur selten dafür eingesetzt, die Zugänglichkeit zu kulturellen Sammlungen zu verbessern.

In dieser Masterarbeit gebe ich einen Überblick über den aktuellen Stand von Visualisierungen von kulturellen Sammlungen. Dabei betrachte ich sowohl den Einsatzzweck, als auch deren Effektivität.

Im Anschluss präsentiere ich einen Entwurf für ein interaktives, web-basiertes Interface, welches es erlaubt die Daten einer kulturellen Sammlung anhand verschiedener Visualisierungen zu erforschen. Dabei liegt der Fokus auf räumlichen Daten (Geolocation) und *multiple coordinated views*.

Ich stelle eine prototypische Implementierung vor, welche diese Konzepte umsetzt. Dabei handelt es sich um eine Web-Applikation, welche historische Bilder und Aufnahmen von Wien zwischen 1147 und 1995 zugänglich macht. Der Prototyp verwendet die bestehenden Metadaten, angenäherte Geolocation, historisches Kartenmaterial und eine Vielzahl von Visualisierungstechniken um die Sammlung in interaktiver Weise zu präsentieren.

Anhand einer qualitativen Nutzer*innenstudie wird die Zugänglichkeit der Sammlung durch den Prototypen bewertet. Die Ergebnisse zeigen, dass Informationsvisualisierung helfen kann, sowohl die Zugänglichkeit als auch die Verständlichkeit von kulturellen Sammlungen zu verbessern.



Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar.
The approved original version of this thesis is available in print at TU Wien Bibliothek.

Abstract

Museums and institutions that house cultural collections are increasingly interested in making their cultural heritage data publicly accessible. They often provide web-based interfaces, allowing users to search their cultural heritage data. Unfortunately, in many cases, these interfaces require domain knowledge about the data to use the collections effectively. This makes browsing through these digital collections unattractive for casual users.

Appropriate Information Visualizations can solve this issue and make complex data understandable and accessible to non-experts. Surprisingly, visualizations are rarely used for supporting the accessibility of digital cultural heritage data collections, despite their effectiveness.

In this thesis, I first review the current state of the art of visualizations for digital cultural heritage collections and the effectiveness of Information Visualization in this context.

Subsequently, I propose an interactive web-based interface for exploring cultural heritage data through Information Visualizations, focusing on spatial data and multiple coordinated views.

I present a prototypical implementation of this interface, a *history map* for images depicting scenes in Vienna between 1147 and 1995, owned by the Austrian National Library. The prototype uses existing image metadata, approximated geolocation, historical maps, and various visualization techniques to showcase the collection of images interactively.

To assess the data's accessibility based on this tool, I conducted a qualitative user study. My results indicate that visualizations can help with user engagement and can increase accessibility and understandability of the provided cultural heritage data.



Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar.
The approved original version of this thesis is available in print at TU Wien Bibliothek.

Contents

Kurzfassung	vii
Abstract	ix
Contents	xi
1 Introduction	1
1.1 Motivation	1
1.2 Research questions	2
1.3 Definition of Cultural Heritage Data	2
1.4 Outline	3
2 Problem Description	5
2.1 Information seeking: structured vs. exploratory	5
2.2 Classical cultural collections	7
2.3 Digital cultural collections	8
2.4 Information Visualization	9
2.5 Visualizations of cultural data collections	12
3 State of the Art Report	15
3.1 Related Work	15
3.2 Method	16
3.3 Approaches to visualizations of digital cultural heritage data collections	16
3.4 Categorization of existing approaches	34
3.5 What is missing?	36
4 Prototype Design	39
4.1 Approach description	39
4.2 Scenario and Methodology	42
4.3 User interface design	47
4.4 Design Methodology	47
5 Prototype Implementation	57
5.1 Prototype Characteristics	57
	xi

5.2	Architecture	57
5.3	Technologies	58
5.4	Data Analysis and Usage Permission	59
5.5	Data Model	60
5.6	Data Extraction and Migration	60
5.7	Source Code and Project Structure	60
5.8	Screens	62
6	Prototype Evaluation	71
6.1	Purpose and Method	71
6.2	Interview Guidelines	72
6.3	Participants	75
6.4	Interview Notes	76
6.5	Findings	76
7	Conclusion	81
7.1	Conclusion and main contributions	81
7.2	Future work	82
	List of Figures	83
	List of Tables	85
	List of Algorithms	87
	Bibliography	89
	Appendices	99
A	Interview Notes	101
A.1	Interviewee 1	101
A.2	Interviewee 2	104
A.3	Interviewee 3	107
A.4	Interviewee 4	109
A.5	Interviewee 5	112
A.6	Interviewee 6	115
A.7	Interviewee 7	118
A.8	Interviewee 8	120
A.9	Interviewee 9	123

Introduction

1.1 Motivation

The amount of data in today's information society is growing steadily. The way we handle this vastly expanding data landscape has a crucial impact on our way of life and the distribution of knowledge. Notably, the Internet has become a tool that allows billions of users access to large amounts of information. Even data that was previously both hard to comprehend and hard to access is becoming increasingly available, including cultural heritage data.

Generally speaking, cultural heritage data can include any data about objects, artifacts, or even non-physical entities created by society and passed from one generation to the next, including paintings, sculptures, music, photographs, and many more. Cultural heritage data traditionally was not provided in digital form. Nonetheless, this is changing rapidly as more and more museums and other institutions digitalize their collections and make them accessible via the internet. This increased access to cultural heritage data collections by a broader audience is a massive chance for cultural development and educational purposes. However, it also introduces challenges.

There is a big difference in how different user groups access data. Up until recently, cultural heritage data was mainly used by experts with extensive previous knowledge. Nevertheless, this group of users broadens more and more to include non-expert users as well. That means access to data has to be adapted and designed in a way that also casual users can access data.

The way these casual users seek and use information is often different from the expert user. The data access is less structured and more exploratory. This casual data access to cultural heritage data collections is the topic of this thesis. The design of interfaces for data access is a challenge, and visualizations have been proven to be a useful tool for making complex data more accessible and understandable. Hence I propose using

Information Visualization to enable casual users to explore cultural heritage collections effectively. I will describe the design and implementation of a web-based Information Visualization interface and illustrate the procedure and the results of an evaluation I conducted to test the efficacy of the proposed solution.

1.2 Research questions

In this thesis, I want to describe what work has been done already and what kind of visualization approaches exist in the context of cultural heritage data. My main focus is describing how visualizations can support a non-expert and casual exploration of cultural heritage data. Specifically, I want to answer the following research question:

RQ Can visualization support interactive exploration of cultural heritage collections across several data dimensions in a meaningful way?

The main research question can be further articulated into two sub-questions:

SQ1 Which visual approaches already exist, either documented in academic literature or publicly available as web-based applications?

SQ1 Is it possible to improve existing approaches in terms of interface design and implementation, and how?

1.3 Definition of Cultural Heritage Data

The United Nations Educational, Scientific and Cultural Organization (UNESCO) defines cultural heritage in the following way[UNE16]:

- Tangible cultural heritage:
 - movable cultural heritage (e.g., paintings, sculptures, coins, manuscripts)
 - immovable cultural heritage (e.g., monuments, archaeological sites)
 - underwater cultural heritage (e.g., shipwrecks, underwater ruins and cities)
- Intangible cultural heritage: oral traditions, performing arts, rituals

In this thesis, I use this definition as a reference, concentrating primarily on *tangible, movable cultural heritage*. Nevertheless, the concepts and approaches discussed are also applicable to data dealing with other types of cultural heritage. Hence, I will include those concepts as a reference and for illustration purposes.

1.4 Outline

This thesis is organized as follows: Chapter 2 will be the main entry point to the topic of cultural heritage data collections. Giving background about what defines information seeking in Section 2.1, I will discuss then aspects of classical cultural heritage collections and how these concepts transfer to the digital realm (Sections 2.2 and 2.3). Section 2.5 explores what part data visualizations can play in this context.

Chapter 3 aims to give an overview of existing approaches of visualizations in the context of cultural heritage collections and compiles a State of the Art Report, identifying which aspects need further attention.

Based on the State of the Art Report, a design for a prototype is presented in Chapter 4. The implementation details of this prototype are described in Chapter 5.

Chapter 6 summarizes the evaluation of the implemented prototype, including a description of the conducting interviews and a list of findings.

Finally, in Chapter 7, a conclusion is presented, referencing the original problem description and offering an outlook over potential future work.



Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar.
The approved original version of this thesis is available in print at TU Wien Bibliothek.

Problem Description

The subject of visualizations of cultural heritage data collections is a multifaceted topic depending on factors originating in classical cultural collections (like museums) and the way humans seek information. In this chapter, I want to cover these background areas and set the stage for the actual problem of visualizations in the context of cultural heritage data.

2.1 Information seeking: structured vs. exploratory

The process of information-seeking can be simplified as follows:

- a) the desire to acquire knowledge (information-need)
- b) the actual process of doing so [Mar89]

Both of these aspects can have different motivations and characteristics. These motivations can range from scientific research to boredom. The characteristics can vary from very structured, target-oriented information access (*I know what I am looking for*) to casual, exploratory browsing (*I do not know what I am looking for*). Depending on the motivation, it is vital to choose the right characteristics for the search process. Surprisingly this is often, especially in digital interfaces, not the case. Whitelaw illustrates this paradoxical situation with the following scenario:

"Imagine yourself outside an art gallery in a far-off city, with a collection you don't know well. You enter the building to find a small, drab lobby with an attendant at a desk. The attendant asks you to enter your query on a small slip of paper. Not knowing the collection, and not seeking anything in particular, you write down something arbitrary, and pass it over. The attendant disappears for a moment before returning with a line of artworks sitting on trolleys. These are paraded, ten at a time, through the lobby. You

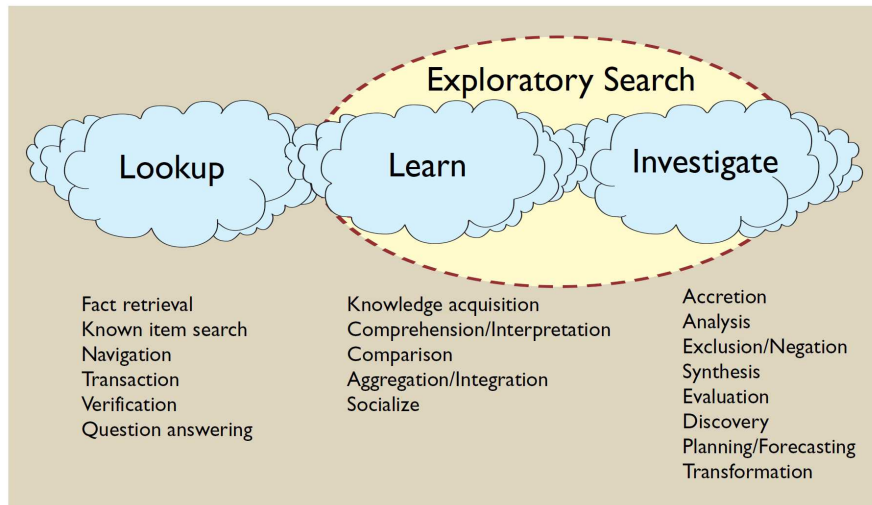


Figure 2.1: Search activities: Explorative Search focuses on *learning* and *investigating* whereas a *Lookup* generally requires specialized domain knowledge. [Mar06]

can submit another query at any time, calling forth more trolleys, but there seems to be no way to explore the gallery beyond this small lobby."[Whi15, p. 1]

In this case, the characteristics of the provided interface do not fit the motivations of the visitor. Therefore the interface is not very useful in retrieving meaningful information. On the other hand, it would be perfect for someone knowing the collection already (expert-knowledge). Someone like this would be able to retrieve what he is looking for nearly instantly without having to go through the hassle of browsing the entire (or even just parts of the) collection.

This suggests that there is no need to replace existing interfaces but rather complement them. A study by the Indianapolis Museum of Art shows that this is also the case in the context of cultural heritage data. The study shows that the motivations and backgrounds of visitors to online museums are very diverse and that, amongst others, 21% of visitors merely want to "engage in casual browsing" when visiting an online collection[INT16].

As already mentioned, I want to concentrate on this user group that has no previous knowledge (non-expert users) and where a simple search and retrieve approach is not sufficient. We need ways to cope with this more exploratory approach to information retrieval, where users do not have a clear picture in their mind of what they are seeking.

Marchionini defines three search activities describing different motivations and approaches to information seeking: *lookup*, *learn*, and *investigate*. According to Marchionini, exploratory search mainly focuses on learning and investigating (see Figure 2.1). This definition is consistent with my assumptions and the work of Whitelaw, who states that the activity of a *Lookup* requires at least basic previous knowledge [Whi15].

Dork et al. define the exploratory type of data access as the *information flaneur*[DCW11].

They compare this exploratory information seeking process with a person walking through an unknown city and experiencing the city with an open mind, trying to understand mechanics without previous theoretical knowledge. Today's nearly universal data availability "can be viewed similarly, as both overwhelming information overload and fascinating information access"[DCW11, p. 1215]. In this regard, understanding how humans search and browse information is crucial. It is necessary to look at all different ways information and data are accessed. The way Dork et al. and Marchionini describe exploratory search can be interpreted as an iterative process. One has to get an overview before being able to refine its search and define a specific direction. Walsh and Hall model exploratory search as a process with iterative steps where the user adapts, refines and specializes the search after each iteration [WH15]. Unfortunately, this casual way of seeking information is rarely considered in scientific literature[DCW11], and Walsh and Hall state that "there remains a gap in our understanding around that first, casual interaction between the user and the collection." [WH15, p. 1].

Nonetheless, it has been shown already that explorative search approaches can be implemented effectively and that especially the introduction of relevance feedback in support of the user can drastically improve retrieval and user engagement [ML13, Mar06]

2.2 Classical cultural collections

To understand digital cultural heritage data collections, I first want to talk about their analog counterpart. A *classical cultural collection* is a collection of cultural artifacts that are often also on display for the public, e.g., a library or a museum. The aspect of public access is especially interesting for understanding how information seeking is implemented in a traditional, conventional way.

Even in a classical setting, we can see different ways data is organized and presented. Libraries, for example, often follow a very structured approach to data presentation. Books are categorized into genres and ordered alphabetically (see figure 2.2). By this, it is possible to display many objects, likely the whole available collection, at once. This kind of organization promotes lookup activities rather than exploratory search (see figure 2.1).

In contrast to this, museums usually do not present the whole collection at once. In some cultural contexts, a traditional organization like alphabetical order or order by creation date often make little sense by themselves. They can be useful as metadata but cannot represent the entire collection in a meaningful and understandable way. A significant role I want to talk about briefly is the *curator*. The curator's objectives are to select, organize, and suitably categorize a collection for public display[BKD]. The curator only presents a logical selection that can be easily browsed by visitors. The concept of the *curator* is also essential in the digital realm, as we will explore further in this thesis.

The phenomenon of *Museum fatigue* is an excellent example of how essential novel and effective data presentation is to keep the interest of visitors. The term Museum fatigue

2. PROBLEM DESCRIPTION

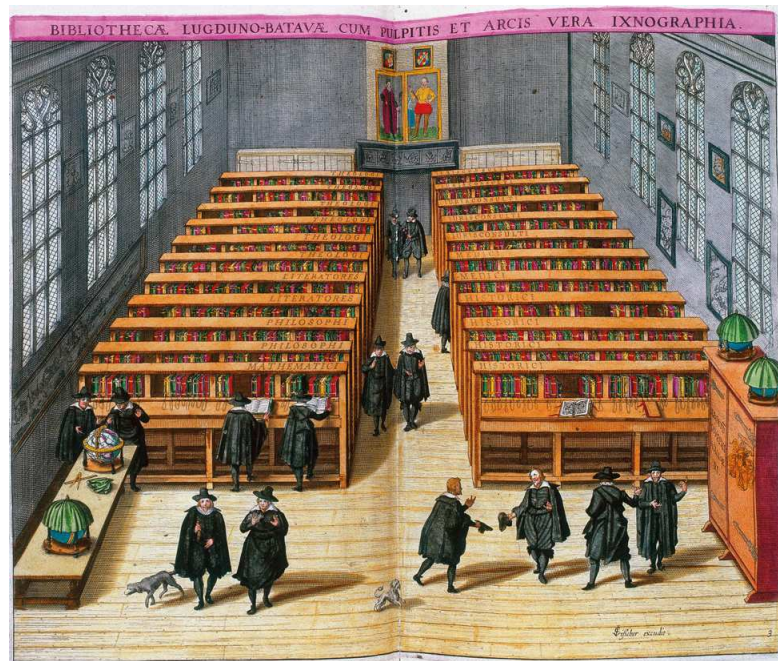


Figure 2.2: Classical library organization [Wik16]

"characterizes decreased visitor interest towards exhibits as visits progress" [Dav05, p. 17]. Additionally to data selection and presentation (via a curator), museums use a wide range of actions to involve visitors in a more exploratory way (e.g., audio guides, video documentation, interactive installations)[DvL05]. These measures enable casual users to discover information in new and unfamiliar areas.

2.3 Digital cultural collections

Like classical cultural collections, digital cultural collections can be a comfortable way for users to browse and access data. However, challenges arise as well: How does a very specific lookup work? How can the collection be browsed without searching for something specific?

Although there are approaches in classical cultural collections on how exploratory search can be facilitated, not many of these concepts have been translated into their digital counterparts[Whi15]. Especially the concept of a *curator* is also fundamental in the digital realm as the vast amount of data available needs some organization and presentation to facilitate the average user. Although important, it is surprising that this concept is often overlooked [Whi15]. Many data retrieval interfaces for cultural heritage data mimic basic database query tools that allow searching the collection depending on various search terms and parameters. These tools are a low-level and often non-intuitive way of accessing data, not facilitating access for a large user group. Especially people without

previous knowledge are left-behind and have no way of effectively using the provided data. Technically the data is available to them; practically, there is no way of making sense without proper support by the providing interface. Why are there not many solutions for the non-expert user, although there is a clear indication that casual browsing is valued highly by visitors?[LBLM13].

Whitelaw introduces the concept of the *generous interface* as a way to *escape the search box*, where he understands the search box as an analogy for the entirety of insufficient query tools in a cultural heritage context[Whi15]. These tools assume that the user can provide a query, neglecting the fact that "a user who is unfamiliar with the collection's scope, contents, or structure may not be in a position to query it effectively"[HW10, p. 52]. Whitelaw explains that "a more generous interface would do more to represent the scale and richness of its collection. It would open the doors, tear down the drab lobby; instead of demanding a query, it would offer multiple ways in, and support exploration as well as the focused enquiry where search excels." [Whi15, p. 1]

The fact that it seems evident that these interfaces are required for exploratory search access makes it even more surprising that they are not yet universally implemented and in use. A possible explanation is that the interface requirements are hard to define and highly subjective[LBLM13]. This constitutes a major challenge to interface designers and requires a deep understanding of the visitors' expectations and the structure and content of a collection.

2.4 Information Visualization

As Information Visualization, we understand the visual approach to help people build knowledge and understand data [KS14]. This is achieved by applying a multitude of interactive visualization techniques ranging from simple charts to complex 3D visualizations. Information visualization is used by public institutions, companies, and individuals to make complex data accessible or deliver a message supported by data. For example, Daniel Lima's collection of network visualizations show in a compelling way how complex data can be communicated effectively through visualizations[Lim13].

But to understand what exactly is Information Visualization, we have to first define what *information* is. Schreiber et al. define the relationship between data, information, and knowledge as follows:

- **Data** are input signals to sensory and cognitive processes.
- **Information** is data with an associated meaning.
- **Knowledge** is the whole body of data and information together with cognitive machinery that people are able to exploit to decide how to act, to carry out tasks and to create new information. [SAA⁺01]

Based on this definition, we can understand visualization as "the process of transforming data, information, and knowledge into visual form, making use of humans' natural visual capabilities." [GEC98, p. 9]

Information Visualization is going a bit further in that it uses "computer-supported, interactive, visual representations of abstract data in order to amplify cognition"[CMS99, p. 1]. In the process, the main focus lies on gaining insights, amplifying cognition and aiding decision making rather than on the presentation of raw data values[Poh13, AJ15].

Information Visualization is strongly related to and interconnected with scientific visualization. Nevertheless, it is essential to discriminate between the two. A significant difference is the type of data they are dealing with. Scientific visualization is based on concrete data with real physical properties (e.g., measured values of the human body, molecules, weather phenomena). This data can be displayed directly using a realistic representation of its physical properties or at least using an estimation of the real physical thing. There can be a certain inclusion of abstract data in scientific visualization, but the focus always lies in the visualization of geometrical data[Poh13, Spe01].

In contrast, Information Visualization focuses on abstract data, which has no apparent physical or geometrical component. Examples can be abstract quantities like connections between human beings, financial data, or sports scores. There are many classes of information falling into this category, having no natural and obvious physical representation, but still conveying highly exciting data. Information visualization deals with the questions of how this abstract data can be made visually accessible for the viewer, and what kind of representation makes the data most understandable. Despite the abstractness of the data, it still can be visualized but because abstract data is not automatically mapped to the physical world, some mapping logic has to be applied. To find a valid mapping between abstract entities and a spatial representation, is one of the major challenges of Information Visualization[Spe01, Poh13, GEC98, Car07, Lan07].

To give an example: The development of the unemployment rate has no apparent physical properties. Hence it needs mapping before it can be visualized with spatial references. By mapping the temporal (months and years) to one geometric axis and the rate values to another, we can create a visual representation of the abstract dataset as figure 2.3 illustrates.

The Visualization Reference Model

The mechanisms behind effective Information Visualizations can be formalized as the mapping from data to visual form. Card calls this formalized concept *The Visualization Reference Model*. It describes the transformation steps from raw data to a visual view, which can be perceived by a human user [Car07]. Figure 2.4 illustrates this process.

1. **Data Transformations:** The first step when building Information Visualizations is to transform *raw data* into structured *data tables* containing all necessary

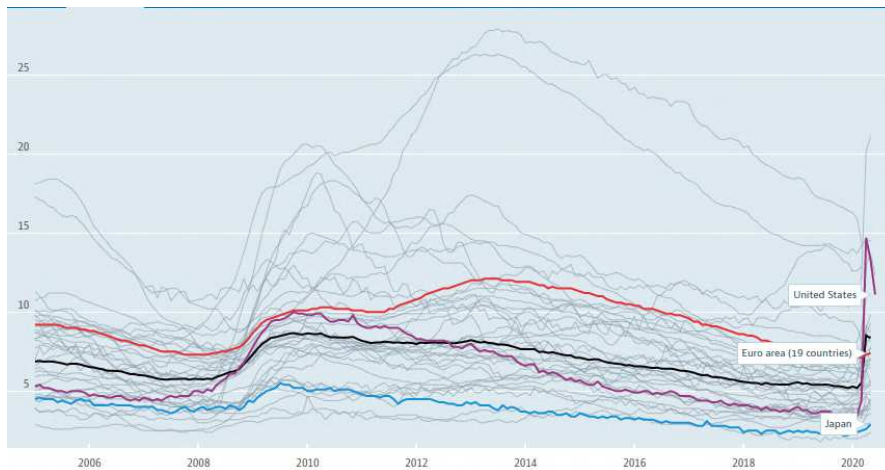


Figure 2.3: Development of the unemployment rate of OECD countries [web17]

information used for creating a visual representation of the data. This includes object relations and other meta-data.

2. **Visual Mappings:** *Data tables* are then mapped into *visual structures*. These are elements representing the original data, but prepared in a way that they can be used for a visual scene (e.g., lines, axis, circles, and others).
3. **View Transformations:** The final transformation maps *visual structures* to final *views* which can actually be perceived by users. These transformations include, for example, visual distortions or the application of a certain 3D viewing angle. [Car07]

Information Visualization often contains a certain element of interactivity in which users can control and manipulate the elements contained in the visualization. These adjustments by the user typically have a direct effect on the visualization and hence the data itself. In the Visualization Reference Model, this is illustrated by the arrow flowing from right to left (see figure 2.4), potentially affecting all transformation steps. E.g., suppose the user applies a filter on the visualization. In that case, this filter might affect the data transformation resulting in a different composition of data tables, which consequently results in a different set of visible elements on the view.

Additionally, the Visualization Reference Model defines data structures, visual structures, and quality criteria. It combines those properties into a detailed *Taxonomy of Information Visualizations*, which can serve as a reference when discussing Information Visualizations. Discussing those concepts would go beyond the scope of this thesis.

In summary, the *Visualization Reference Model* is a capable framework that can be beneficial when evaluating existing visualizations or designing new Information Visualizations.

2. PROBLEM DESCRIPTION

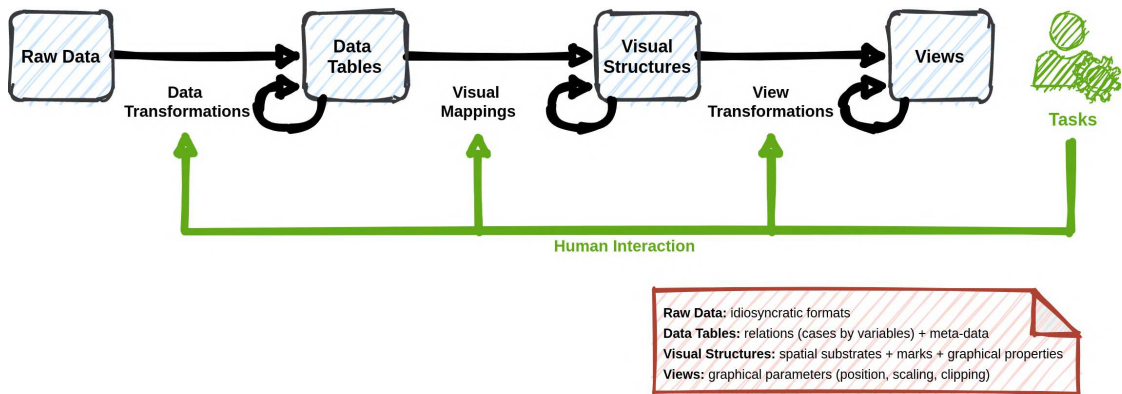


Figure 2.4: Reference model for visualizations [CMS99]

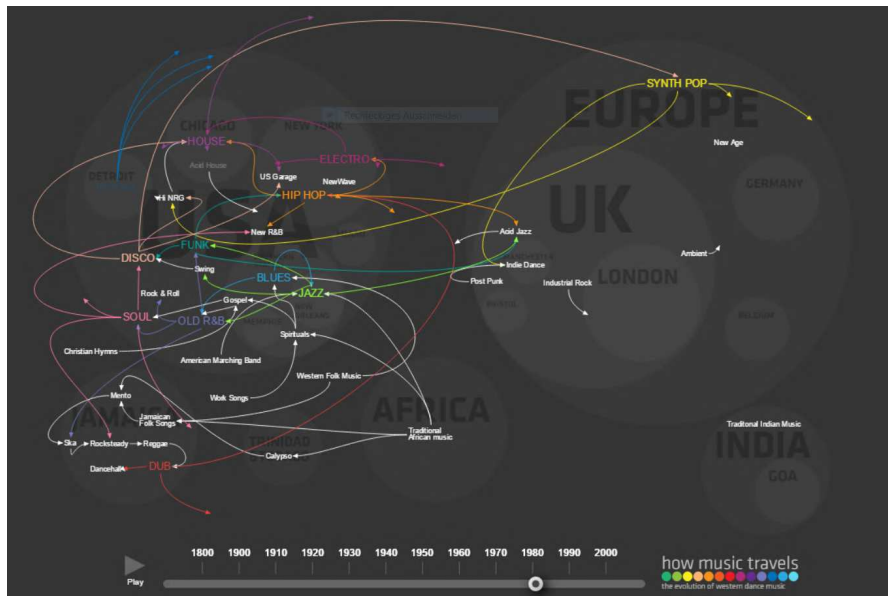


Figure 2.5: How music travels - example of a multidimensional (temporal, spatial, categorial) Information Visualization[Man16]

2.5 Visualizations of cultural data collections

As explained in section 2.3, the design of data-access-interfaces in the context of cultural heritage collections is challenging. Information Visualization can, however, support and advance the development of such interfaces. Visualizations have proven to be a useful tool for making complex data accessible. They can help focussing accessibility to anyone with a general interest in the selected topics, even if they do not have any prior experience

or expert knowledge [Lim13]. Lima's visualizations suggest that this can be applied to cultural data as well (e.g. see figure 2.5).

Although visualizations are already extensively used to convey data in an easily understandable way, they are, however, still rare in the context of cultural heritage data. I want to contribute to filling this gap.

Visualizations not only allow us to interact with an existing collection more interactively but actually can provide information that was previously not accessible. By creating visualizations covering a whole collection of artifacts, structures and patterns become visible, e.g., empty spots in the collection ("visualizing the un-seen"[GMD15]).

The basis for my contribution will be a state of the art report in Chapter 3 where I want to summarize and structure previous work in the area of visualizations of cultural heritage data collections. The identified patterns and concepts will be evaluated and incorporated into a new prototype implementation of a visualization interface for a cultural heritage data collection for which a similar interface does not exist yet. The focus should lie on Information Visualization concepts which show potential, but are not yet widely used in the context of cultural heritage data collections. The usage of a data set that is not yet incorporated into visualizations gives me the chance to evaluate the usefulness of such a new interface, which uses visualizations (see Chapter 6). I want to compare the two interfaces (traditional "search-box" interface vs. newly implemented prototype using visualizations) through a user study and prove that visualizations can, in fact, support interactive exploration of existing cultural heritage data collections.



Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar.
The approved original version of this thesis is available in print at TU Wien Bibliothek.

State of the Art Report

3.1 Related Work

Various literature exists on the topics of cultural heritage, visualization, and information-seeking for casual users. There has also been quite some work done connecting these topics [GD15, ML13, GMD15, BKD, Whi15, DLB13, DWC12, WH15].

Nevertheless, overview literature about these topics, especially including actual approaches, is rare. Windhager et al. [WFM⁺16] published a review of visualization approaches in the context of cultural heritage data collections. The review covers a wide range of interfaces, tools, and concepts and organizes them into the following categories:

- spatial encoding
- temporal encoding
- focus on cultural objects or actors
- type of project

The review by Windhager et al. gives an excellent overview of what kind of interfaces and tools already exist and where there is still a potential room for improvement. Windhager et al. published a follow-up report in 2019, comparing a wide range of interfaces and approaches to the visualization of cultural heritage collection data [WFS⁺19]. This State of the art report classifies existing approaches and describes the techniques used. I used the work of Windhager et al. as a starting point for my research. Though I am using a different categorization, the results are still comparable and related to each other.

Hinton and Whitelaw [HW10] explore how visualizations can help to support the comprehensibility of digital museum collections without going into detail about specific approaches.

Glinka and Dörk [GD15] create a relation between research in the area of data visualization and the reception of cultural collections. They do so by giving a few practical examples of interfaces.

3.2 Method

Various methods were used to obtain relevant literature about the topics. The main focus has been on scientific work exploring the interconnection between visualization, cultural heritage data, and exploratory search (casual user access). Besides this crossover literature, I also explored literature dealing with the relevant topics individually. This was especially important in getting familiar with the topics and discovering potential areas that are not yet covered by scientific literature. To find scientific work combining visualization, cultural heritage data, and exploratory search was a quite challenging task.

The main starting point for the scientific research was the paper *A Review of Information Visualization Approaches and Interfaces to Digital Cultural Heritage Collections* by Windhager et al. [WFM⁺16] which gives an excellent overview of most aspects of visualizations in the context of cultural heritage data collections. The paper also references many valuable resources that were, in turn, used for creating and referenced in this document. Different scientific databases were used during the research, primarily *IEEE Xplore Digital Library*[IEE16] and *ACM DL Digital Library*[Ass16]. Furthermore, a general search was done using *Google Scholar*[Goo16b] with few but very useful results. At last, I also want to stress our research done via the Internet-Searchengine *Google*[Goo16a], which was more fertile than expected.

Keywords used while researching on the topics *visualization, cultural heritage data collections, exploratory search*: digital museums, visualization cultural heritage, cultural heritage data collections, visualization cultural data, visualization of museums, exploratory search, casual search, non-expert, iterative search, relevance feedback, generous interface, digital cultural heritage

I found around 20 relevant scientific publications and a wide range of publicly available web-interfaces that helped compose this report. It is essential to bear in mind that not all relevant papers are referenced. Some of them were used to get familiar with the general topic and to broaden my understanding without actually citing or using materials from them. The research was stopped when the number of different papers continued to increase, and the most relevant contributions were already discovered.

3.3 Approaches to visualizations of digital cultural heritage data collections

In this chapter, I want to introduce a range of interfaces, tools, prototypes, and concepts that exist in the context of visualizations of cultural heritage data collections. The research approach for this summary is a more qualitative one than the report of Windhager et

al. [WFS⁺19], which covers a broader range of interfaces. I chose this approach to develop a prototypical interface myself, based on the findings of this State of the Art Report. A more detailed description of individual publications and interfaces allows me to position my work in relation to similar approaches. For my state of the Art Report, I included both scientific publications and publicly accessible interfaces. I would have preferred to concentrate solely on academic publications. However, because most of the relevant implementations are not mentioned in scientific literature, consideration of publicly available interfaces was necessary, even though they were not mentioned in scientific literature.

For being able to analyze the existing approaches to visualizations of digital cultural heritage data collections, I developed a set of tags in order to classify and organize the references, helping me to position my work accordingly (see Chapter 4). Below I briefly want to describe the set of tags used for classification. In the following sections, the tags are used directly below each approach. In Section 3.4, I am using these tags to organize all presented approaches and to create an overview of the broad topic and to identify aspects that still need more focus in scientific literature.

List of Tags:

[interactive] An interactive interface where users can change parameters, viewpoints, and more (in contrast to a static image).

[temporal] Visualization captures a temporal dimension, e.g., the creation date of the artifacts.

[spatial] Visualization captures a spatial or geographical dimension, e.g., place of origin of the artifacts.

[categorical] Visualization captures a categorial dimension, e.g., the specific type of the artifact (e.g., material, size)

[explorative] The interface actively promotes explorative search relevant for casual users.

[visual representation] Their visual appearance structures artifacts within the collection (e.g., similar colored artifacts are placed next to each other).

[multiple coordinated views] Using multiple coordinated views, the interface presents the collection through at least two different data dimensions and puts them in relation to each other (e.g., changing the time period will affect the available list of keywords).

[traditional search] The interface features a traditional, search-box centered interface

[artistic] The interface itself not only should enable the accessibility of a collection, but should become an independent piece of art itself.

[scientific] The interface stems from scientific research and is mentioned in scientific literature.

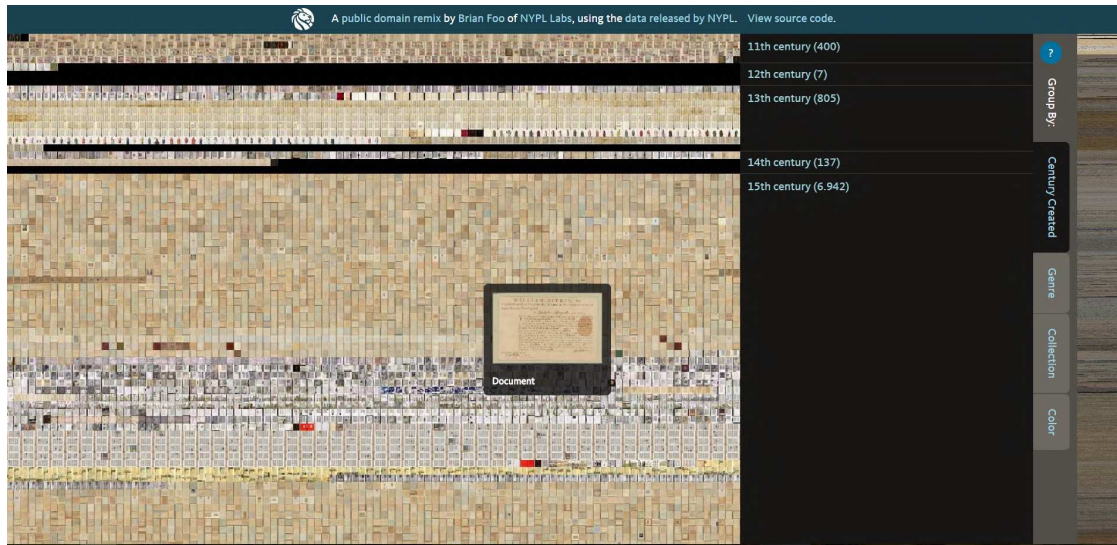


Figure 3.1: NYPL Public Domain Release 2016 Visualization [New16]

3.3.1 NYPL Public Domain Release 2016 Visualization

Tags: [interactive] [visual representation] [temporal] [categorical] [artistic]

The New York Public Library (NYPL) Labs implemented a web interface (see figure 3.1) in 2016 that enables users to explore 187K digital items. The main focus of the visualizations lies in an overview, showing the whole collection at once, at which each item is displayed as a thumbnail. Additionally, there is the possibility to change the organization of the overview by adding temporal or categorial criteria. The functionality is limited. The primary motivation for the online tool seems to be the overview showing the scope of the collection without providing too many navigational options. This implies that the overview by itself gains importance in a functional and an artistic sense, similar to the work of Brendan Dawes[Bre16] who combines thumbnails of movies to create an independent piece of art (see figure 3.2).

3.3.2 Tyne and Wear Archives and Museums - Collections dive

Tags: [interactive] [categorical] [explorative] [traditional search]

The web-interface called *Collections Dive*[Tyn16] provided by Tyne and Wear Archives and Museums is a thumbnails based exploration tool. The provider makes a clear distinction between structured search (Lookup activity) and exploratory search and asks for the visitor's motivation (*I know what I am looking for* vs. *I do not know what I am looking for*). The explore tool is using thumbnails and infinite scrolling. Additionally, relationships between objects (based on metadata) can be inspected in a separate view. The search tool features several search boxes and filters, providing extensive controls for expert users.



Figure 3.2: Cinema Redux by Brendan Dawes [Bre16]

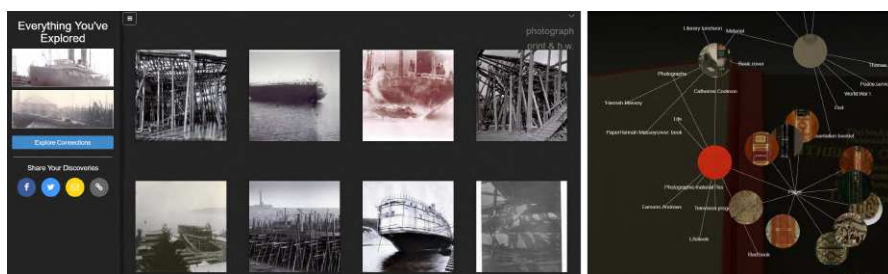


Figure 3.3: Tyne and Wear Archives and Museums: Collectionsdive. Left: Thumbnail explore mode; Right: Relationships [Tyn16]

3.3.3 Cooper Hewitt, Smithsonian Design Museum - A Timeline of Event Horizons

Tags: [interactive] [temporal] [categorial] [explorative] [traditional search]

The Cooper Hewitt, Smithsonian Design Museum has a large online collection that is publicly accessible. They use a search-based interface but have added a temporal visualization to the individual objects. This timeline (see figure 3.4) shows the moments when the object was created and when it became part of the Cooper-Hewitt collection. This puts the individual object in a larger context regarding the whole collection. So although the interface is traditional and keyword-search-centered, these visualizations help the visitor build understanding about the collection and the context a specific object exists in[Coo16].

3.3.4 San Francisco Museum of Modern Art - SFMOMA ArtScope

Tags: [interactive] [explorative] [visual representation] [artistic]

Similar to the New York Public Library (see Section 3.3.1), the San Francisco Museum

3. STATE OF THE ART REPORT

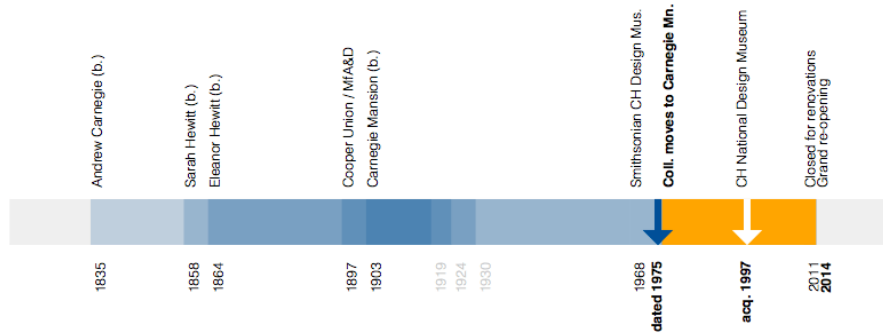


Figure 3.4: Cooper-Hewitt: A Timeline of Event Horizons [Coo16]

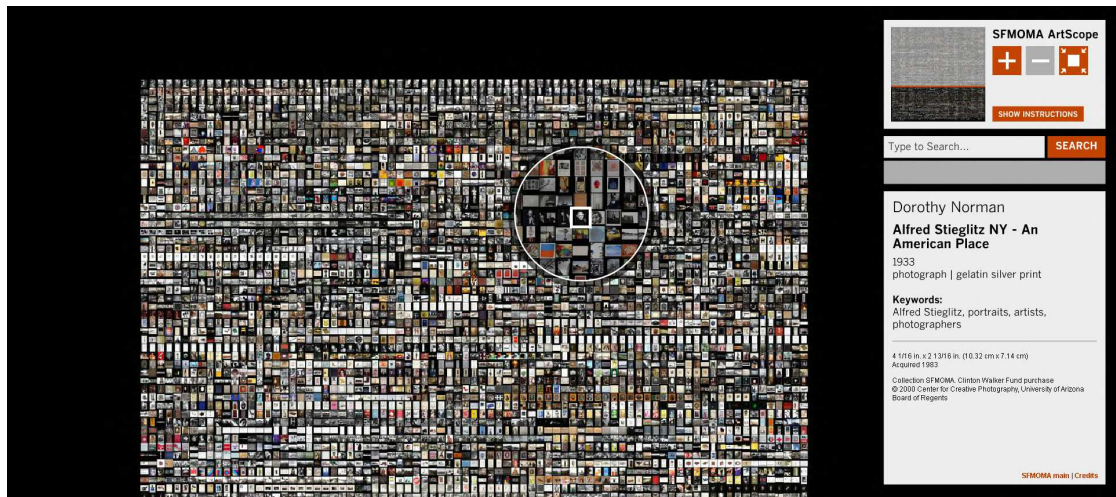


Figure 3.5: SFMOMA ArtScope [San16]

of Modern Art provides visitors with a web interface to explore the entire collection (see figure 3.5). The interface shows the entire collection at once, each object represented by a thumbnail. The collection can be explored by using a virtual magnifying lens. Details about the current object, centered by the lense, are shown to the right.

3.3.5 Florian Kräutli - Visualising Cultural Data - Exploring Digital Collections Through Timeline Visualisations

Tags: [temporal] [spatial] [explorative] [scientific]

Kräutli focuses his work on time-based visualizations "to enable knowledge discovery in digital collections"[Krä16, p. 4]. Figure 3.6 shows one of his visualizations, which contains the artists in Tate's collection plotted in a horizontal timeline. The size of each bubble represents the number of the artist's works in the collection.

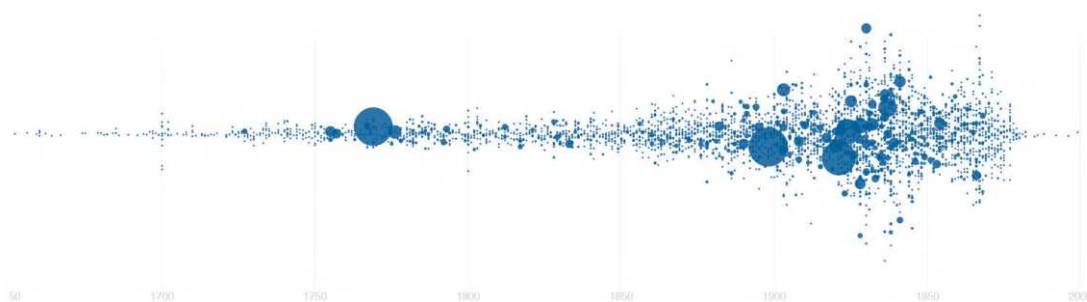


Figure 3.6: Kräutli's visualization of artists in Tate's collection [Krä16]

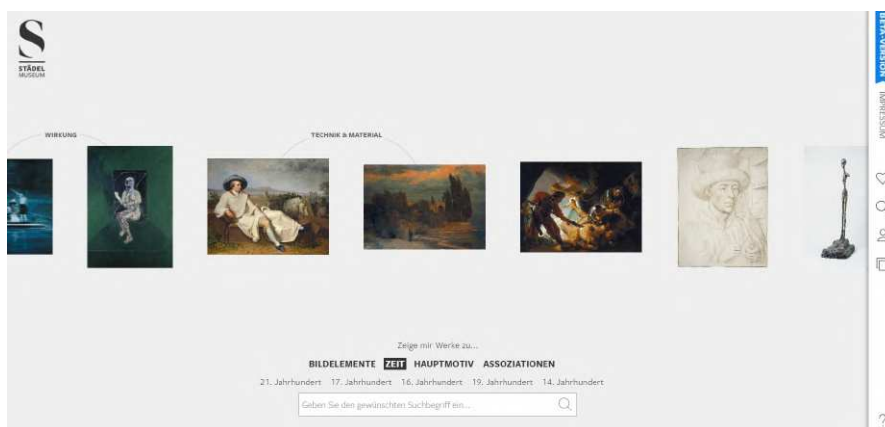


Figure 3.7: Categorical enriched searchbox interface in the digital collection of the Städel museum [MUS16]

3.3.6 Städel Digitale Sammlung

Tags: [interactive] [explorative] [traditional search]

The *Städel Digitale Sammlung*, the digital collection of the Städel Museum in Frankfurt, Germany, is an example of how visitors can still use a traditional search centered interface without previous knowledge about the collections. The search functionality is enriched by a categorial suggestion mechanism (see figure 3.7) that allows new visitors to run queries immediately. Although basic exploration becomes possible, it is still hard to grasp the entire collection's complexity and scope. The focus lies on individual items matching specific metadata rather than on the structure and the composition of the collection as a whole.

3.3.7 metaLAB (at) Harvard: Color Lens

Tags: [interactive] [explorative] [visual representation]

The project *Color Lens* by the metaLab at Harvard University is a very visual approach

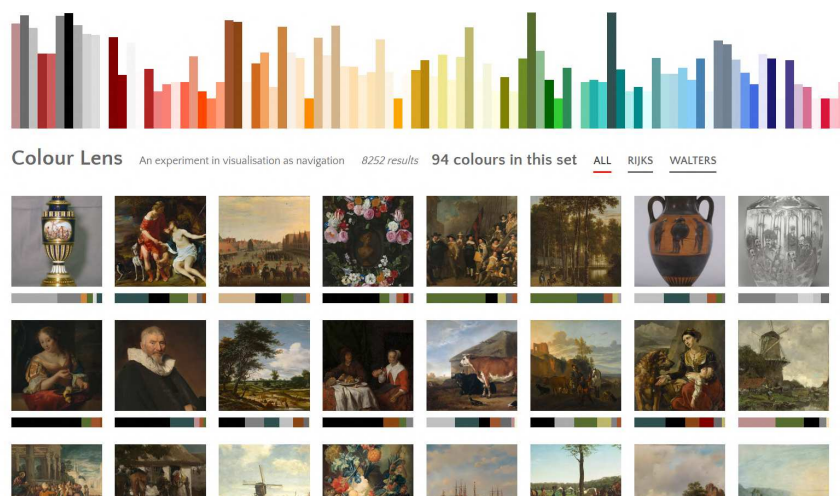


Figure 3.8: Color-based search of the web-interface *Color Lens* [met16]

to digital cultural heritage representation. The interface[met16] combines data from the Rijksmuseum and the Walters Art Museum into one collection. For every individual object in this collection, a color spectrum is defined based on the original artifact's colors. Visitors can then run a color-based search on the collection and explore matching objects.

3.3.8 Getty Research Institute: Auction Market Visualization

Tags: [interactive] [categorical] [explorative]

The Getty Research Institute provides a visualization of the British, Belgian, Dutch, and French auction markets from 1801–1820 (see figure 3.9). The visualization includes individuals and places (represented by differently shaped nodes) involved and active in the auction markets. Because of these two different node types, the visualization includes a spatial and a relationship component. Together they form a complex but still an understandable network of "relationships, patterns and trends"[Get16].

3.3.9 Het Scheepvaartmuseum: Streete View

Tags: [interactive] [spatial] [explorative]

The visualization *Streete View* by the Het Scheepvaartmuseum is a spatial centered visualization of old maps and architectural drawings (see figure 3.10). The visualization allows visitors to navigate interactively through Amsterdam on an old map. Particular points of interest can be clicked and open a detail page of the place, containing a historical illustration of the building or place. The map-based visualization is an excellent example of how metadata, which generally would be quite hard to understand (geolocation), can be prepared to drive exploration.

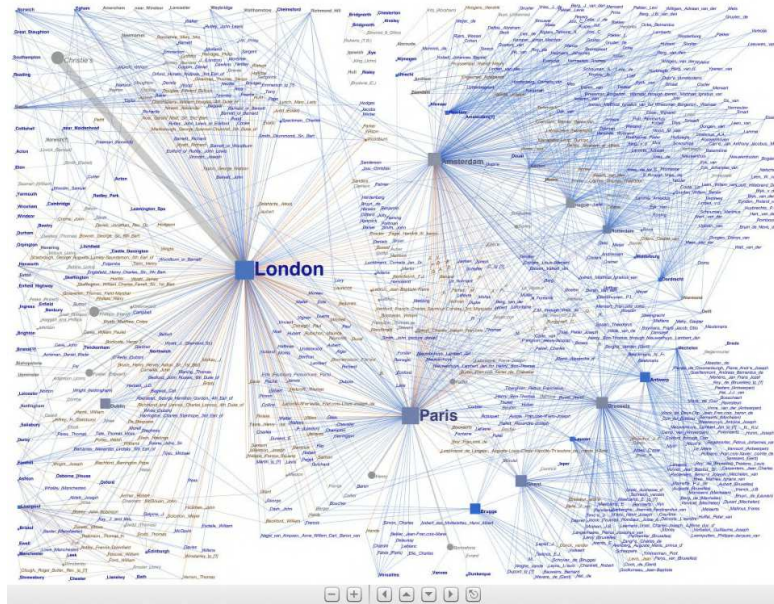


Figure 3.9: Getty Research Institute: Auction Market Visualization [Get16]



Figure 3.10: Spatial visualization *Streete View* by the Het Scheepvaartmuseum using old maps and illustrations. Left: Map overview, Right: detailed illustration of a selected point of interest [Het16]

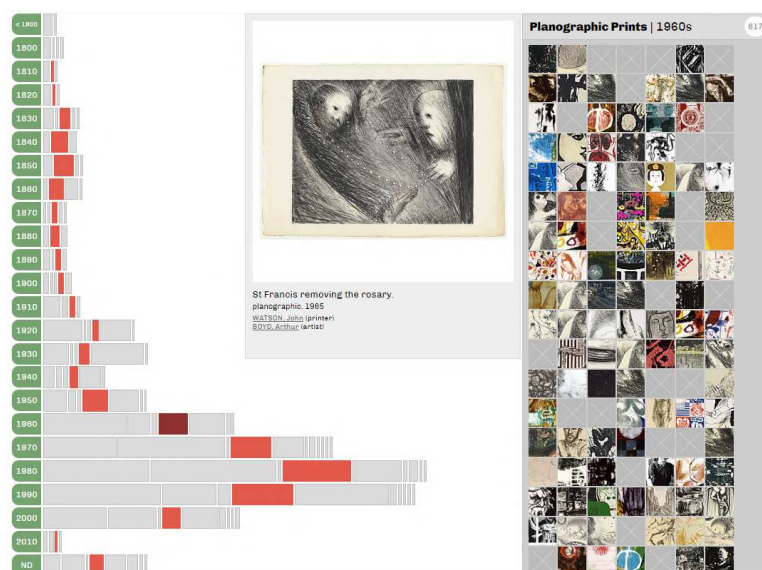


Figure 3.11: Visualization *Decade Summary* of the platform *Prints and printmaking Australia*. Encodes temporal and categorial data dimensions and also allows conclusions about the scope and structure of the collection[Pri16].

3.3.10 Prints and printmaking Australia: Explore

Tags: [interactive] [temporal] [categorial] [explorative]

The platform *Prints and printmaking Australia* offers an extensive interface to explore their collection of printed images (see figure 3.11). Most notable is their visualization *Decade Summary*, a temporal visualization that also displays categories (e.g., type of print) in an understandable way. All objects matching the current filters are shown as a thumbnail-grid and can be opened in a detail-view.

3.3.11 Hampson Archeological Museum State Park: Virtual Museum

Tags: [interactive] [spatial] [categorial] [explorative]

Hampson Archeological Museum State Park has implemented an online interface providing easy access to a series of archaeological artifacts (see figure 3.12). The web interface visually supports users with their browsing and lets them choose between a spatial and a form-dependent explore mode. Although the offered visualizations do not disclose much about the collection's structure and composition, they motivate and enable casual browsing.

3.3.12 Met Museum: 82nd & Fifth

Tags: [interactive] [temporal] [spatial] [categorial] [explorative] [multiple coordinated views] [traditional search]

3.3. Approaches to visualizations of digital cultural heritage data collections

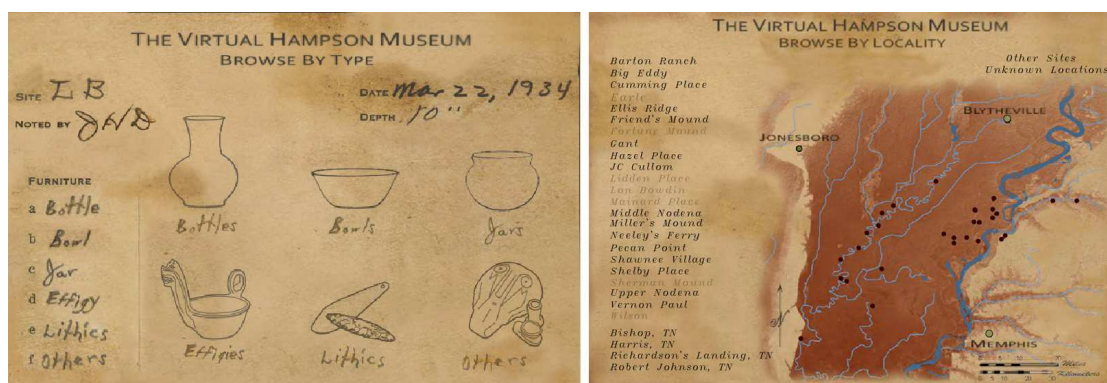


Figure 3.12: The two different browsing-modes of the *Virtual Hampson Museum*[Ham16]



Figure 3.13: Spatial 3D visualization of the 82nd & Fifth collection [The16]

82nd & Fifth is an online platform by the Metropolitan Museum of Art. It is the combined work of 100 curators, each choosing part of the final collection, which is available online. Different ways to explore the collection are offered to the user. A simple thumbnail based overview visualization is available as well as a temporal (see figure 3.14) and a complex spatial 3d visualization (see figure 3.13).

3.3.13 Hinrichs et al.: The Speculative W@nderverse

Tags: [interactive] [temporal] [categorical] [explorative] [multiple coordinated views] [traditional search] [scientific]

The Speculative W@nderverse is an interactive prototype visualizing a subset of the Bob Gibson Collection of Speculative Fiction. The goal was to make the anthologies of the collection accessible and explorable by general users while enabling the use of the prototype in an academic setting.

The core of the prototype is a categorization of the collection in a radial tree of keywords and a keyword cloud (See figure 3.15). Both elements in the word cloud and the radial tree are clickable and will filter the results. Additionally, a high-level categorization of

3. STATE OF THE ART REPORT

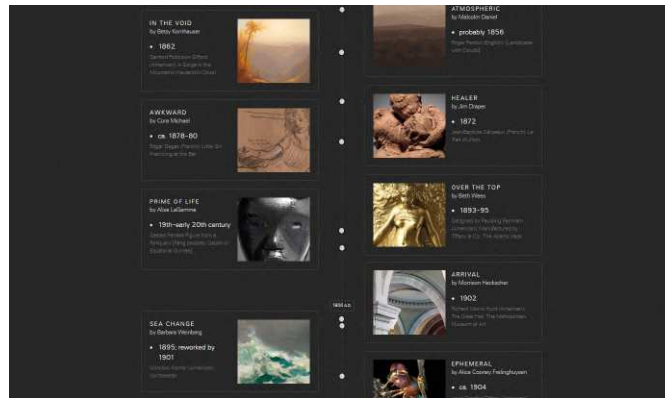


Figure 3.14: Temporal visualization (vertical timeline) of the 82nd & Fifth collection[The16]

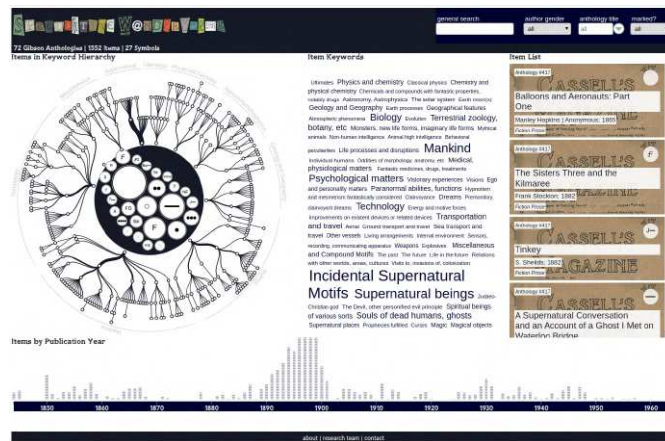


Figure 3.15: *The Speculative W@anderverse*[Uta20]

the stories contained in the collection is located in the center of the tree. Likewise, those elements are clickable and will filter the results. A temporal chart on the bottom of the screen shows the distribution of elements over time, which matches the current filtering. Additionally, a classical search via a search box is provided on the top of the screen. [UHM16]

3.3.14 Kramer-Smyth et al.: ArchivesZ

Tags: [interactive] [temporal] [categorical] [multiple coordinated views] [traditional search] [scientific]

Kramer-Smyth et al. created *ArchivesZ*, a tool for visualizing and exploring archival and manuscript collections. *ArchivesZ* is an early example of an interactive interface for multi-dimensional collection exploration. Figure 3.16 shows the multi-dimensional

Figure 3.16: *ArchivesZ* interface [KSNA07]

interface of ArchivesZ. It is possible to combine a traditional search with a visual filter approach. Search results can be filtered using a temporal graph and inspected via the *Subjects* view, which allows drawing conclusions about the relation of metadata across the whole collection. [KSNA07]

3.3.15 FH Potsdam: Past Visions

Tags: [interactive] [temporal] [categorical] [explorative] [multiple coordinated views] [artistic] [scientific]

Past Visions is a visualization that "shows the collection of historical drawings by Frederick William IV alongside a thematic and temporal arrangement" [FH 20b]. Glinka et al. 's approach is described in a scientific publication [GPDD16] and is available publicly as a web application [FH 20b]. The visualization allows the user to filter the collection via keywords and explore the (filtered) collection on a temporal axis. The displayed collection can be zoomed in seamlessly until individual artifacts can be inspected. The interface structure can be compared to the tool created by Hinrichs et al. (see Section 3.3.13) as both structure the collection by keywords and on a temporal axis by their creation year.

3.3.16 Mitchell Whitelaw: Manly Images

Tags: [interactive] [temporal] [categorical] [explorative] [visual representation]

Mitchel Whitelaw created the interface *Manly Images* as part of a research project in cooperation with the State Library of New South Wales and supported by the Manly Library [Whi12]. The core of the interface is a categorization of an image library by title and by decade. Figure 3.18 shows the view grouped by titles. Grouping by decade produces a similar view. The interface organizes the library into tiles, each representing a

3.3. Approaches to visualizations of digital cultural heritage data collections

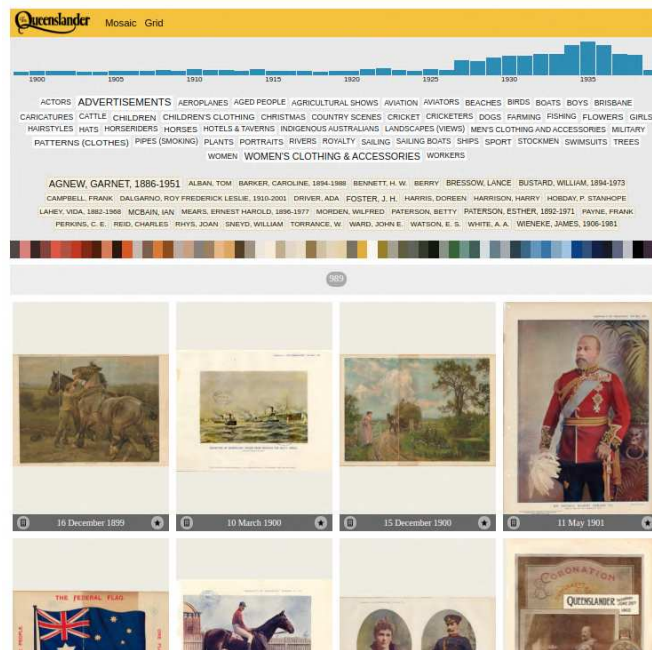


Figure 3.19: Discover the Queenslander[Sta20b]

(now The Courier-Mail) and was published from 1866 – 1939"[Sta20b]. The state library of Queensland implemented an interactive interface on its website to explore a collection of around one thousand artifacts containing covers and illustrated pages from The Queenslander Magazine. Similar to *Past Visions* (see Section 3.3.15) and *The Speculative W@anderverse* (see Section 3.3.13) it features a temporal visualization and a keyword cloud. Respectively the collection can be filtered by selecting a certain time period or keywords. Additionally, filtering is possible by selecting the name of an individual (artist, motif) or a color. Results will be displayed accordingly.[Sta20b]

3.3.18 InnoVis Group of the University of Calgary: Bohemian Bookshelf

Tags: [interactive] [temporal] [categorical] [explorative] [visual representation] [multiple coordinated views] [scientific]

Thudt et al. from the University of Calgary suggest, that explorative search and serendipity in cultural research can be facilitated through visualizations. They introduce a tool to visualize and explore library collections among several data dimensions, including time, cover design, scope (number of pages), keywords (see figure 3.20). The different visualizations included in the tool are interconnected, and changing the parameters in one will directly affect the other visualizations[THC12].

3. STATE OF THE ART REPORT

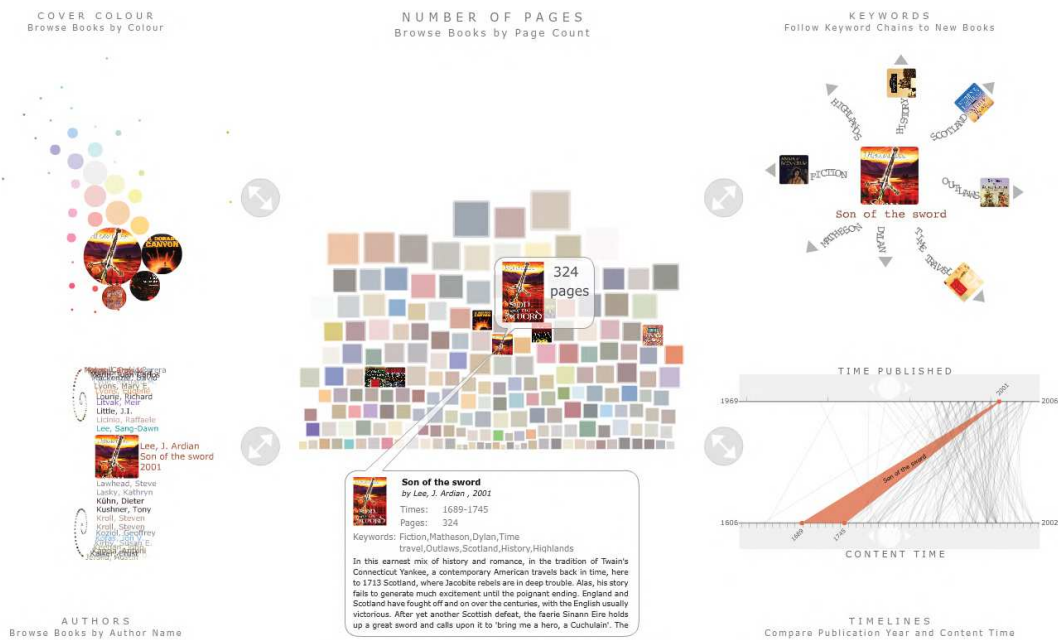


Figure 3.20: The Bohemian Bookshelf[THC12]

3.3.19 Rothko Viz: Visualizing 201 Paintings by Mark Rothko

[temporal] [explorative] [visual representation] [artistic]

Everardo Reyes is following a more experimental approach when visualizing cultural data. His work *Rothko Viz* is a good example of that, in which he experimented with a collection of 201 images of paintings by Mark Rothko. One example of that series of experiments is the "360 Cartesian Transformation" seen in figure 3.21. The images of the collection were "sliced and organized orthogonally, then processed with 360 Cartesian transformation"[Eve20]. The temporal organization allows insights into the artistic development of the artist and gives an overview of the scope of the collection.[Eve20]

3.3.20 Deutsche Digitale Bibliothek Visualized

Tags: [interactive] [temporal] [categorical] [explorative] [multiple coordinated views] [scientific]

The interactive tool *Deutsche Digitale Bibliothek Visualized* is a design study of researchers of FH Potsdam[FH 20a]. Dörk et al. created a tool with 4 different visualizations, each "focusing on different facets of the collection"[DPC17, p. 39]. The four visualizations are timeline, keywords, places, and networks. The visualization *Timeline* shows the number of artifacts of a certain category through time (see Figure 3.22). Each timeline represents a specific category of the collection (e.g., archives, library, media). The visualization *Keywords* shows a keyword cloud displaying the relative frequency of each word (the often a keyword appears in the collection, the bigger it is displayed). *Places* is

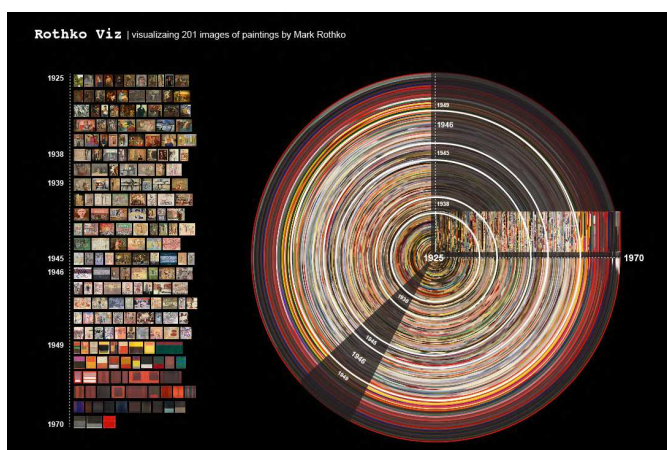


Figure 3.21: Rothko Viz[Eve20]



Figure 3.22: Deutsche Digitale Bibliothek Visualized[FH 20a]

a visualization consisting of several doughnut charts representing a city and showing the relative composition of artifacts in each specific city. Finally, the visualization *Networks* visualizes relationships between entities (e.g., artists or museums) and the number of artifacts connected with each entity[FH 20a].

3.3.21 British Museum - The Museum of the World

Tags: [interactive] [temporal] [spatial] [categorical] [explorative] [multiple coordinated views]

The Museum of the World is a web application created by the British Museum in cooperation with the Google Cultural Institute. It can be described as "interactive experience through time, continents and cultures"[Bri20]. In contrast to most other interfaces presented in this thesis, the Museum of the World features not an existing

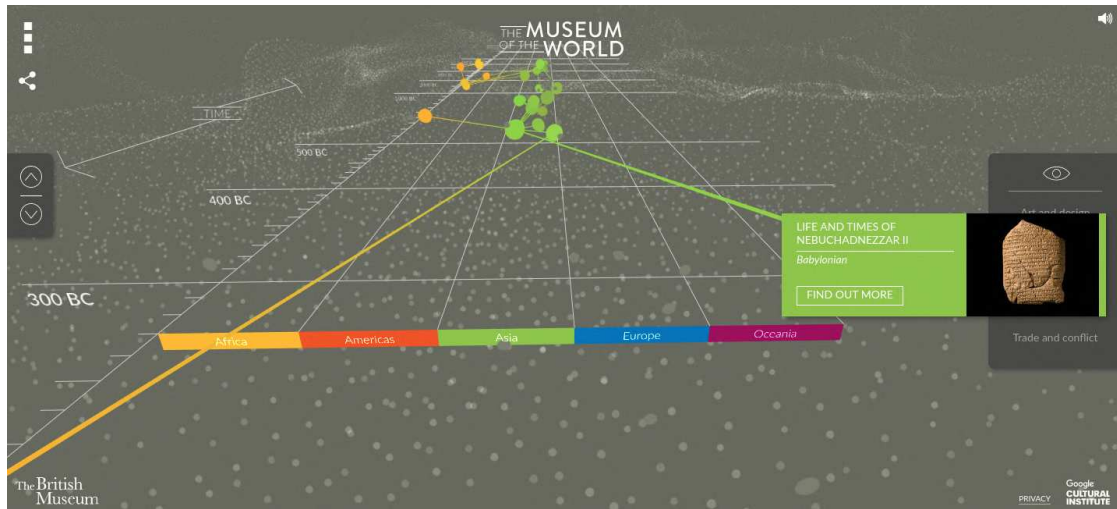


Figure 3.23: British Museum: The Museum of the World[Bri20]

collection but rather a specific selection prepared for the visualization. The main element of the visualization is a vertical, scrollable timeline that allows the user to interactively "travel through time" (see figure 3.23). Every artifact in the collection is marked as a dot on the timeline. Additionally, the artifacts are divided into columns representing their geographical location. Relationships between objects are visualized through lines between the dots. Details to each artifact can be requested by clicking on specific dots. Additionally the interface provides a basic filter consisting of thematic options ("art and design", "living and dying", "power and identity", "religion and belief", "trade and conflict") [Bri20].

3.3.22 Corley Explorer

Tags: [interactive] [spatial] [explorative] [traditional search]

Corley Explorer is a tool for exploring a collection of photographs of private homes in Queensland, Australia. They were taken by Frank and Eunice Corley from the 1960s to the 1970s to sell them to the homeowners. Those that were not sold were donated to the State Library of Queensland and are now accessible via Corley Explorer. The tool allows users to explore the images but also contribute by describing or tagging the images. A map overview shows the spatial distribution of the images in the collection. Another view allows for exploring the ratio of different tags among the collection. These tags can be enriched by users to help improve the collection [Sta20a]. Corley Explorer is a perfect example of how visualizations and explorative design can be accompanied by crowdsourcing-features to enhance the quality of the collection. The explorative design in itself promotes user engagement, and the elaborate tagging and documentation features ensure that this engagement is utilized constructively. Surprisingly, there are not more tools following this approach, as lack of quality is a recurring issue in cultural heritage

3.3. Approaches to visualizations of digital cultural heritage data collections

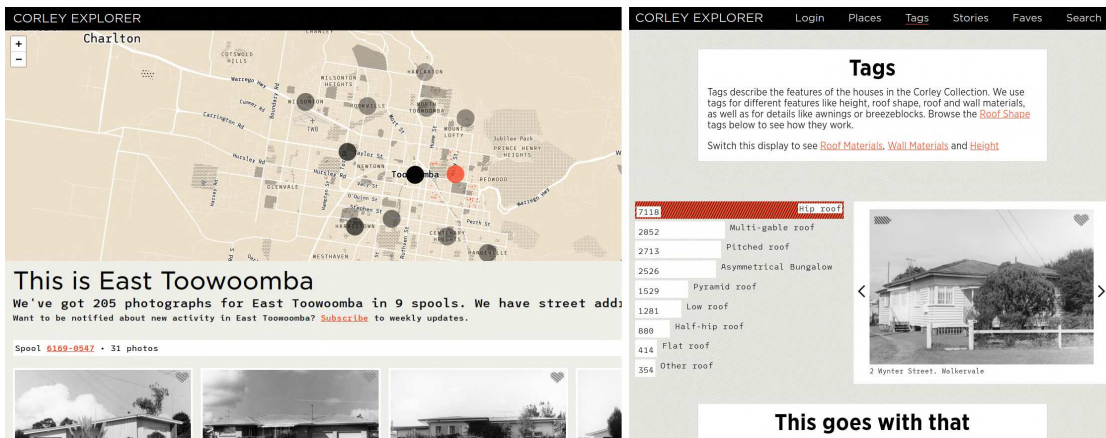


Figure 3.24: Corley Explorer: Map overview (left) and Tag-view(right)[Sta20a]

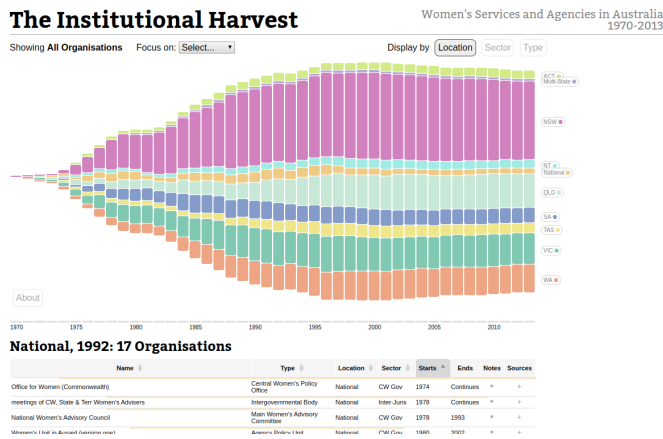


Figure 3.25: The Institutional Harvest[And20]

data [BN13].

3.3.23 The Institutional Harvest

Tags: [interactive] [temporal] [categorical] [explorative] [multiple coordinated views] [scientific]

The Institutional Harvest, developed by Andrew Merrindahl and Mitchell Whitelaw, is an online tool created as part of a larger project led by Marian Sawer and Sarah Maddison[And20, MS13]. It shows "Australian women's agencies and services over the period 1970-2013. It reveals the growth and diversity of the organizations emerging from the Australian women's movement"[And20]. The visualization is a centered, stacked bar graph showing the development over time. Several filter options like *Location*, *Sector* or *Type* exist and have impact on the visualization.

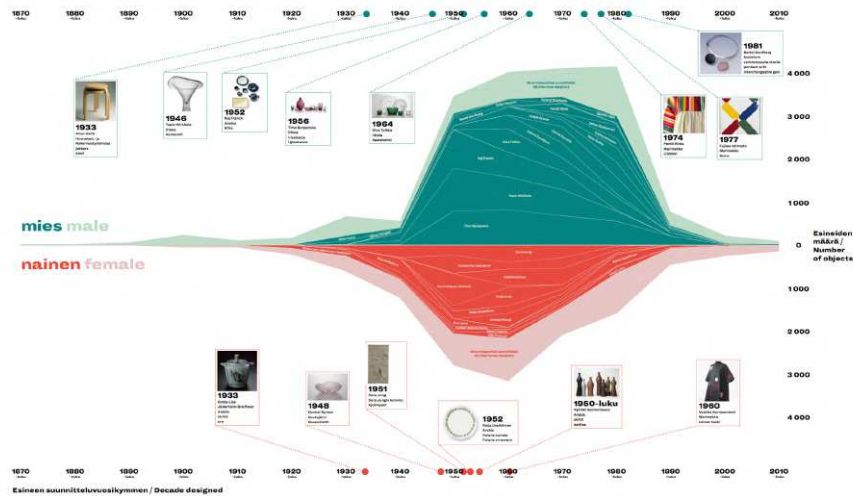


Figure 3.26: Behind The Utopia[Lil18]

3.3.24 Behind the Utopia - Aalto ARTS

Tags: [temporal] [categorical] [explorative] [artistic]

Although the project *Behind the Utopia* by students of Aalto University did not result in an interactive interface, I still want to include their work because of their illustrative quality for describing a museum collection. A central piece of their work (see Figure 3.26) shows the composition of the museum’s collection over time, including featured artists, the total number of exhibits, sample exhibits, and the proportion between male and female exhibitors[Lil18].

3.4 Categorization of existing approaches

In this section, I want to structure the different approaches introduced in Section 3.3 regarding shared concepts and visualization techniques. As a central guideline for the categorisation, I will use the tags listed at the beginning of Section 3.3.

Table 3.1 shows an overview of all mentioned approaches and interfaces and lists which tags (and therefore which techniques) they use.

Using the data from table 3.1, figure 3.27 shows which concepts are frequently used and are which not yet utilized very often. It’s clearly visible that many explorative interactive interfaces visualizing temporal and categorial data exist (tags [explorative], [interactive], [temporal], [categorical]). When considering the type of data that typically exists in cultural heritage data collections, these numbers are sound. Most cultural collections displaying artifacts to the public provide explicit metadata about an object’s creation date and its basic properties (style, shape, technique) [PWM⁺05].

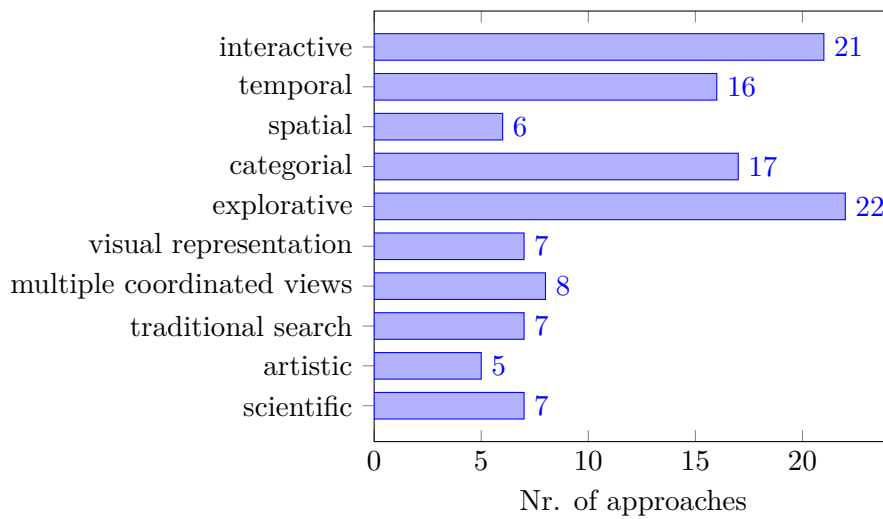


Figure 3.27: Number of approaches per tag

On the other hand, we see that interfaces exploring cultural collections in an artistic sense are rare. A possible explanation for this low number is that this is a specific use case mainly targeted towards other artists or people specifically interested in experimental art. These experimental characteristics of these artistic interfaces might prevent a broader application (e.g., for historians, people interested in history or classical arts). Similar, interfaces trying to structure the visual representations of artifacts (tag [visual representation]) are maybe more likely to be used by a very specialized user group and less useful for a broad audience, hence the moderate number of interfaces implementing this concept.

Likewise, only a few interfaces feature some spatial dimensions in their visualizations. This is surprising considering the [spatial] tag features all interfaces containing some spatial data, even if the spatial data is used similar to categorial data (e.g., names of countries used as categories instead of the geographical location on a map). Furthermore, the fact that spatial data is increasingly available in the context of cultural data collections [CS10] makes the lack of more respective interfaces striking.

A moderate amount of interfaces implement the concepts of [multiple coordinated views], [traditional search] and [scientific]. It is less surprising that the mentions of these types of interfaces in scientific literature are rare, as similar studies have shown [WFM⁺16, WFS⁺19].

In contrast, a traditional search interface and a multidimensional filter are features that would typically greatly benefit a digital collection. The lack of those concepts in our sample can maybe be attributed to some selection bias as we were actively concentrating on more experimental interfaces that do not focus on traditional concepts.

3. STATE OF THE ART REPORT

Approach	int	temp	spat	cat	expl	vis	mult	trad	art	scie
NYPL Public Domain (3.3.1)	X	X		X		X			X	
Collections dive (3.3.2)	X			X	X			X		
A Timeline of Event Horizons (3.3.3)	X	X		X	X			X		
FSFMOMA ArtScope (3.3.4)	X				X	X			X	
Kräutli - Timeline Visualisations (3.3.5)		X	X		X					X
Städel Digitale Sammlung (3.3.6)	X				X			X		
metaLAB (at) Harvard: Color Lens (3.3.7)	X				X	X				
Auction Market Visualization (3.3.8)	X			X	X					
Streete View (3.3.9)	X		X		X					
Prints and printmaking Australia (3.3.10)	X	X		X	X					
Virtual Hampson Museum (3.3.11)	X		X	X	X					
Met Museum: 82nd & Fifth (3.3.12)	X	X	X	X	X		X	X		
The Speculative W@nderverse (3.3.13)	X	X		X	X		X	X		X
ArchivesZ (3.3.14)	X	X		X	X		X	X		X
FH Potsdam: Past Visions (3.3.15)	X	X		X	X		X		X	X
Manly Images (3.3.16)	X	X		X	X	X				
Discover the Queenslander (3.3.17)	X	X		X	X	X				
Andean Pots (??)										
Bohemian Bookshelf (3.3.18)	X	X		X	X	X	X			X
Rothko Viz (3.3.19)		X			X	X			X	
Deutsche Digitale Bibliothek (3.3.20)	X	X		X	X		X			X
The Museum of the World (3.3.21)	X	X	X	X	X		X			
Corley Explorer (3.3.22)	X		X		X			X		
The Institutional Harvest (3.3.23)	X	X		X	X		X			X
Behind the Utopia (3.3.24)		X		X	X				X	

Table 3.1: Overview of tags of each individual approach.

Abbreviations: **[int]**eractive, **[temp]**oral, **[spat]**ial, **[cat]**egorial, **[expl]**orative, **[vis]**ual representation, **[mult]**iple coordinated views, **[trad]**itional search, **[art]**istic, **[scie]**ntific

3.5 What is missing?

In the previous section, we discussed the various concepts frequently used throughout our sample and those who lack a broader application. In this section, we want to summarize which concepts are not yet implemented comprehensively and set the stage for my visual prototype design (see 4).

The conclusion we can draw from our sample and its categorization is that although the structure of cultural collections suggests that spatial data can be utilized effectively in cultural collection visualization interfaces, there is still a lack of actual implementations available to the public. The reason for this is unclear. Nevertheless, I will try to find answers through implementing the visual prototype and observing the challenges that might arise during the design and implementation.

Furthermore, only a few interfaces try to combine an experimental, visual approach with a traditional, search-box-centered approach. However, during our research, such combined interfaces proved to be highly effective (e.g., The Speculative W@nderverse, see Section 3.3.13).

To summarize, the following concepts need further attention, and I will include them in the prototypical implementation and its documentation:

- **Visualization of spatial data:** The prototype should have a map-centered design focusing on spatial data.
- **Combination of traditional, search-box-centered design with an experimental visual interface:** The prototype should feature a search function to facilitate traditional lookup activities additionally to the visual approach.
- **Multiple coordinated views:** The prototype should display several data dimensions at the same time and connect them with each other through connected filters (e.g., changing the filter on one data dimension, will affect what is displayed in another).
- **Academic documentation:** Only few approaches listed in Section 3.3 were part of an academic publication. The prototype will be thoroughly documented in this master's thesis.

Additionally, throughout the design process, I want to focus on the concept of explorative search and enable the easy usage of the interface by casual users.



Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar.
The approved original version of this thesis is available in print at TU Wien Bibliothek.

Prototype Design

In this chapter, I describe the process of designing the prototype. The State of the Art Report's results in Chapter 3, especially the areas of improvements summarized in Section 3.5, will be the basis for most design decisions. I will discuss the general structure of the approach and describe specific design decisions in detail. I will also briefly discuss the utilized dataset and the general technical features of the software. Detailed information about the technical implementation can be found in Chapter 5.

4.1 Approach description

This section briefly covers the point of departure and the general objective of the prototype. The prototype should be based on an existing interface and fulfill all requirements identified in Section 3.5 in order to enhance exploration and user engagement. Hence it was important to find an existing cultural heritage collection for which a digital interface exists, and which contains enough meaningful data to derive the prototype from it. This interface needed to be publicly available and fall into the category of a typical *search-based interface* (as discussed in Chapter 2). I was looking at various interfaces for cultural collections to find a suitable candidate, which would be the reference for my own prototype as I will elaborate further in this section. Firstly, I will describe the chosen existing interface in Section 4.1.1. Secondly I will describe in Section 4.1.2 the general characteristics of the visual prototype based on the existing interface.

4.1.1 Search-based interface [S-INTF] - <https://www.bildarchivaustria.at>

After exploring various digital collections, I decided to use the sub-collection *Wien/Bezirke* (Vienna/districts) of the Bildarchiv Austria (part of the Austrian National Library)[Ös20] for my prototype. It features a vast compilation of historical pictures and photographs,

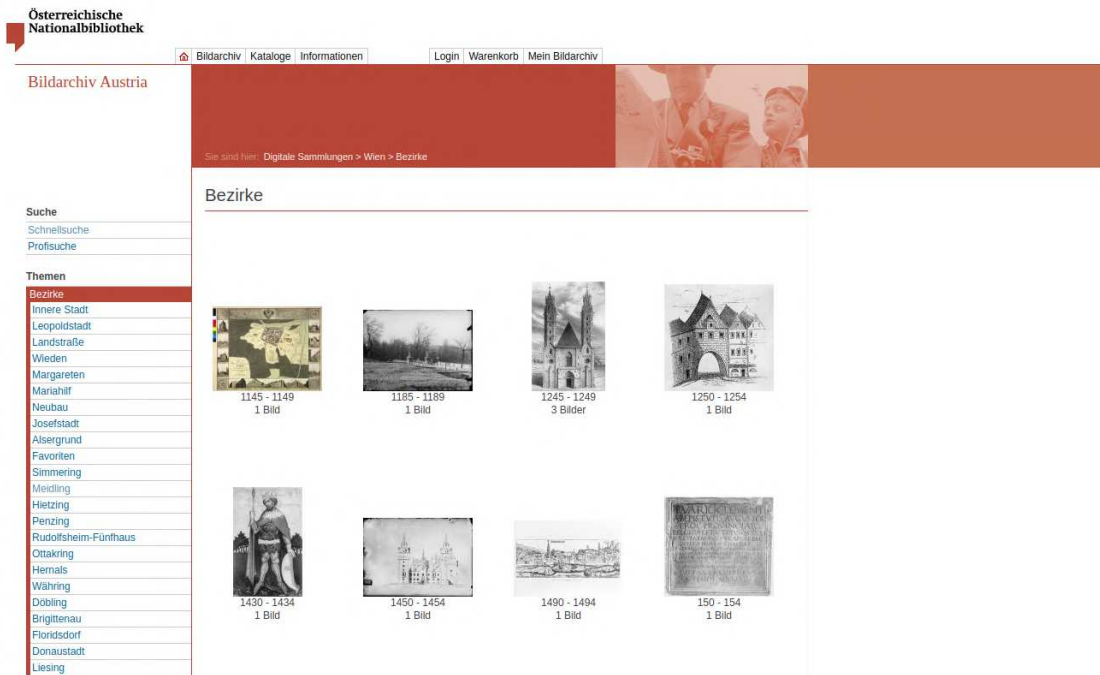


Figure 4.1: Bildarchiv Austria - Current Search-Based Web Interface [Ös20]

organized by their district or the featured motif. For each district, the contained images are organized into folders, each representing a time period of five years. Although the data is far from complete or perfect, it conveys much information and has the potential for extensive data exploration and visualization. The structure of the sub-collection *Wien/Bezirke* allows for detailed geocoding of the artifacts. For details about the contained data, see Section 4.2.3. Apart from the technical requirements regarding the dataset, I saw much potential in terms of explorative search and user engagement. I believe most people are interested in the history of their city and especially their immediate surroundings. The sub-collection *Wien/Bezirke* potentially allows users to explore historical images of their own neighborhood and to see how it has changed over the years. This is one main reason why I wanted to use this collection, as I saw a real possibility to create an interesting, interactive interface which could interest a lot of people living in Vienna or visiting the city. The Bildarchiv Austria hosts the collection on the web themselves. Figure 4.1 shows the current interface for browsing the collection.

The interface qualifies as a traditional search-box-centered interface. It features detailed search functionality (see Figure 4.2) and structures images into temporal buckets (see Figure 4.1). These buckets allow the user limited explorative search characteristics (e.g., if a user wants to explore old pictures, he can click on the appropriate time-categories). Based on the findings in Chapter 3, I believe the interface of the Bildarchiv Austria can be a very descriptive example of showing that visualizations can support the exploration of cultural heritage collections as proposed by my research questions (see Section 1.2).

Österreichische Nationalbibliothek

Bildarchiv | Kataloge | Informationen | Login | Warenkorb | Mein Bildarchiv

Bildarchiv Austria

Sie sind hier: Suche > Profisuche

Profisuche

Suchbegriff(e)	<input type="text"/>
Titel	<input type="text"/>
Datierung	Vor <input type="text"/>
Bildnummer	<input type="text"/>
Inventarnummer	<input type="text"/>
Digitale Sammlung	Wien <input type="text"/>
Thema	-- keine Auswahl -- <input type="text"/>
Subthema	-- keine Auswahl -- <input type="text"/>
Schlagwort	<input type="text"/> Lookup <input type="button" value=""/>
Ort	<input type="text"/> Lookup <input type="button" value=""/>
Person	<input type="text"/> Lookup <input type="button" value=""/>
Körperschaft	<input type="text"/> Lookup <input type="button" value=""/>

Bei der Profisuche kann in jedes Eingabefeld ein Begriff eingegeben werden. Die Begriffe werden automatisch mit UND verknüpft. Die Lookup Buttons ermöglichen Ihnen, die zur Auswahl stehenden Begriffe aufzufinden und durch Klick für die Suche auszuwählen.

last update 26.06.2019

Über uns - Kontakt - Impressum
Copyright © 2019 Österreichische Nationalbibliothek

Figure 4.2: Bildarchiv Austria - Search interface

The existing interface is not utilizing yet any visualization concepts. Hence it might be suitable for showing that its functionality can be extended and adapted to facilitate additional requirements (e.g., explorative search).

In the rest of the thesis, I will use the term *search-based interface* [*S-INTF*] to refer to the interface of the Bildarchiv Austria.

4.1.2 Visual Prototype [P-INTF] - <https://historymap.wien>

Based on the list of topics summarized in Section 3.5 I decided to create a new interface using the same data provided by the Bildarchiv Austria in the Collection *Wien/Bezirke*. As identified in section 3.4, not many other interfaces use spatial visualizations. Hence I want to focus on this aspect during the prototype implementation. The new interface should center around a map conveying the spatial artifact data of the collection. Other visualizations shall complete the prototype's explorative character and allow the casual user to discover additional aspects of the collection and filter accordingly. Nevertheless, the map will always stay in the center of the interface, and the goal is to convey as much information directly on the map, without visually overloading the interface. The

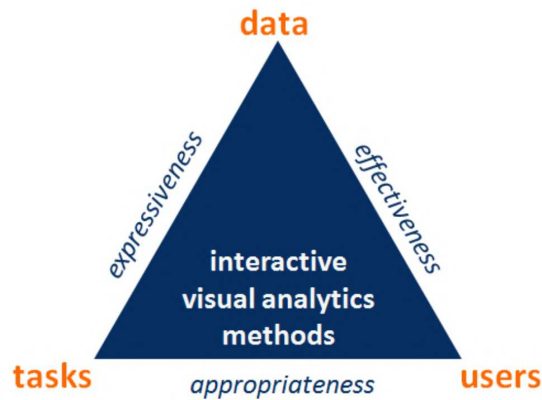


Figure 4.3: Design Triangle[MA14]

prototype is using the same dataset as the *search-based interface* [*S-INTF*], with the exception of some minor additional data processing (see Section 4.2.3). The prototype will be publicly available at the following URL: <https://historymap.wien>.

In the rest of the thesis, I will use the term *visual prototype* [*P-INTF*] to refer to the prototype I created in the scope of this thesis, which focuses on Information Visualization.

4.2 Scenario and Methodology

In order to design a useful interface, it is essential to focus on the specific requirements it is supposed to meet. For designing the visualization techniques and concepts, I want to use the so-called *design triangle* introduced by Miksch et al. [MA14]. It describes a process for designing visualizations concerning the following three requirement categories:

- What kinds of data are the users working with? (data)
- Who are the users of the VA solution(s)? (users)
- What are the (general) tasks of the users? (tasks)

Additionally, Miksch and Aigner introduce a set of qualities that should be met by data visualizations. They can be seen as the edges of the design triangle (see figure 4.3). These include expressiveness ("The requirement of showing exactly the information contained in the data"), effectiveness ("The degree to which visualization addresses the cognitive capabilities of the human visual system") and appropriateness ("The cost-value ratio in order to assess the benefit of the visualization process")[MA14, p. 286-287].

In the following sections, I want to detail the specifics for the requirements (users, tasks, data) and summarize how they can be facilitated to create meaningful visualizations that satisfy all desired qualities (expressiveness, effectiveness, and appropriateness).

4.2.1 Users

At the beginning of Chapter 2, I discussed the qualities of a casual user and why interfaces have to facilitate the special requirements presented by this user group. Unlike expert users, they do not know what they are looking for, although a general interest for the presented topic might exist. The existing interface can be considered an interface targeted towards expert users. The advanced search interface supports a knowledgeable user in various ways to enable him to find results as fast as possible. However, the advanced search requires understanding about the topic and the structure of the collection. The goal of the *visual prototype [P-INTF]* is to enable casual users to explore the collection in a meaningful way. The required domain expertise can, therefore, be defined as *novice* (following the classification by Miksch and Aigner [MA14]). Additionally, the interface should ideally engage users with limited interest in the broader topic of Viennese history. Nevertheless, I will focus on casual users who tend to show some interest in the topic, without having specific domain knowledge.

From a technical perspective, the focus should lie on users accessing the interface via a web browser on a desktop computer. Nonetheless, the interface should also be usable on mobile devices, and the map-centric-design could potentially promote mobile usage (e.g., people want to wander around the city and explore old pictures of the buildings around them).

4.2.2 Tasks

The primary supported task is the exploration of the cultural heritage collection, and the ability to search the spatial data dimension via the map. A secondary task is the inspection of the collection structure. Another secondary task is the visualization of collection details, which were previously not visible by merely accessing the interface. E.g. the relationship between artifacts is hard to identify using the *search-based interface [S-INTF]*.

Brehmer and Munzner call these exploratory tasks *enjoy*, *browse* and *explore*. *Enjoy* refers to "casual encounters with visualization" without being driven by a clear need to find specific information [BM13, p. 2378]. *Browse* and *Explore* refer to search tasks to find unknown targets matching certain characteristics (e.g., finding old images of Vienna without being familiar with the exact images).

A critical aspect of all those tasks is not to overwhelm the user. The tasks should be simple to execute, and the basic feature set should be easy to understand. Expert features that allow the user to explore particular aspects of the collection should not be shown by default. They might confuse the target user (casual users without domain knowledge). At the same time, tasks should be executable effectively. E.g., temporal and spatial visualizations should work together to allow for easy and fast querying of the contained data (e.g., "Which artifacts exist in a certain location for a selected timeframe?").

Wien 12, Schönbrunner Allee 66



#1883715

Titel	Wien 12, Schönbrunner Allee 66
Beschreibung	Hotel-Café-Meierei Siller. Frontansicht mit Gartengitter.
Datierung	1940
Orte	Wien 📍
Schlagworte	Meidling, Wienansichten
Digitale Sammlung	Wien
Inventarnummer	Pk 3227, 218C
Bildnachweis	ÖNB

Figure 4.4: Bildarchiv Austria - Collection Data Entry

4.2.3 Data

As mentioned in Section 4.1.1, the chosen dataset is contained in the sub-collection *Wien/Bezirke* (Vienna/districts) of the Bildarchiv Austria (part of the Austrian national library)[Ös20]. The dataset contains a total of 13,958 artifacts, including historical drawings, paintings, and photographs. Figure 4.4 shows an example entry of the original collection.

As we can see, the dataset contains multiple data dimension:

- **Temporal Data:** Every Artifact in the collection has an assigned creation date or year. This data is mostly accurate, and only a few artifacts are missing this information.
- **Spatial Data:** Most artifacts in the collection have an assigned address or a named point of interest. This data is less accurate than the temporal one but precise enough to utilize it
- **Categorical Data:** There is a range of categorial data assigned to the artifacts in the collection. This includes data about authors, keywords, and creation techniques. This type of data is not complete throughout the whole collection. Some artifacts include a data category that others don't, but in general, the majority of the categorial data is complete.

In the following paragraphs, I want to mention more specifics about each type of data dimension.

Temporal Data:

Although temporal data should not be the main focus (unlike spatial data), the general good quality of the temporal data justifies a more intense utilization of the temporal data compared to other data dimensions. In general, all artifacts in the collection contain either an exact creation date or at least the creation year, which allows a meaningful temporal distribution of the data. The few artifacts which do not contain the creation year were filtered out and not used in the new *visual prototype [P-INTF]*. This resulted in 475 artifacts from the original collection, which do not contain a precise creation year, being excluded from the dataset used in the new *visual prototype [P-INTF]*.

The resulting dataset has a minimum year-value of 150 and a maximum value of 1995 and can, therefore, be displayed on a linear temporal scale. Because of the large number of possible values, grouping the values by periods of ten years makes the visualization of the temporal data easier and comprehensible.

Categorical Data:

The dataset includes a range of additional metadata, which qualifies as categorical data. This includes keywords, creation techniques, and authors. The assigned keywords and authors of the artifacts are exact and easy to use. In total, there are 991 unique authors and 636 unique keywords used throughout the collection.

In contrast, the creation techniques vary greatly throughout the collection and are not usable unambiguously. There are 53 techniques used throughout the collection, but as the data type is a simple string, there are many duplicates that are hard to clean up (e.g., combinations of other techniques with a slightly different spelling). Furthermore, the categorization of the original collection varies wildly. E.g., some photographs are categorized by their capture technique (what type of photography was used), and other photographs are categorized by their printing technique (how they were printed and what kind of paper was used). To utilize the technique data more effectively, I classified all available techniques in 3 broad categories (Druck|Print, Fotografie|Photography, Malerei/Zeichnung|Painting/Drawing) like shown in algorithm 4.1.

Unfortunately, the raw data is too vague. Therefore, the final result using categorial technique data must be considered a proof-of-concept only and not a production-ready solution yet.

Spatial Data:

As described in Section 4.1.2, the new interface should be centered around a map and concentrate on the spatial data dimension (enriched by the other data dimensions). As the collection does not provide comprehensive spatial data, some data transformation had

Algorithm 4.1: Assigning Technique Category

```

UPDATE technique SET category_id =
  (SELECT id FROM technique_category WHERE name = 'Druck')
WHERE name IN
  ('Lithografie', 'Kupferstich', 'Druck', 'Buchdruck', 'Kupferstich,_handkoloriert',
  'Kupferstich_Radierung', 'Schabkunst', 'Punktierstich', 'Stahlstich', 'Kupferstich,_koloriert',
  'Xylografie_(Holzstich)', 'Salzpapier', 'Zinkotypie', 'Lichtdruck', 'Heliogravure',
  'Chromolithographie', 'Edeldruckverfahren', 'Farbdruck', 'Offsetdruck', 'Farblithografie',
  'Brom\'lumdruck_Edeldruckverfahren_Vintage_Print');

UPDATE technique SET category_id =
  (SELECT id FROM technique_category WHERE name = 'Fotografie')
WHERE name IN
  ('Schwarz-Weiss-Negativ', 'Silbergelatineabzug',
  'Schwarz-Weiss-Negativ,_Glasplatte', 'Farbnegativ_(Colornegativ)', 'Glasplatte',
  'Albuminabzug', 'Albuminpapier', 'Schwarz-Weiss-Diapositiv',
  'Schwarz-Weiss-Negativ_(Film)', 'Schwarz-Weiss-Abzug', 'Kollodiumpapier', 'Fotografie',
  'Kopierverfahren', 'Schwarz-Weiss-Negativ_Schwarz-Weiss-Negativ_(Film)',
  'Farbdiapositiv', 'Autochrome', 'Schwarz-Weiss-Abzug_Schwarz-Weiss-Abzug',
  'Autochrome_Platten');

UPDATE technique SET category_id =
  (SELECT id FROM technique_category WHERE name = 'Malerei/Zeichnung')
WHERE name IN
  ('Federzeichnung', 'Federzeichnung,_Aquarell', 'Radierung', 'Aquarell', 'Graphit',
  'Mischtechnik', 'Bleistiftzeichnung', 'Pastell', 'Deckfarben', 'Handzeichnung',
  'Pinselzeichnung', 'Aquatinta', 'Gouache', 'Aquatinta_Radierung');

```

to occur. To display artifacts on a precise location on a map, I used Google's Geocoding API[Goo20a] to lookup an artifact's location (coordinates) via its address or description, respectively.

This process by itself was not yet exact as some queries returned multiple possible locations (e.g., if the address is not distinct). For these cases, I developed an algorithm handling those events. See algorithms 4.2, 4.3 and 4.4 for the actual implementation. The structure of the collection (see figure 4.1) allows us to determine the exact district an artifact is located in. In case we receive multiple possible hits from the Geocoding API, we filter out all hits not located in the correct district. If multiple hits are located in the correct district, we estimate the relative location to each other and assume the first located coordinates as the correct ones, if all other possible hits are very close (< 300m) to that location (see algorithm 4.4).

In case of a single hit, we check if the district of the found location matches the one provided through the collection structure (see algorithm 4.3).

This process allowed for detailed geocoding of most artifacts. Nevertheless, 2.902 artifacts (out of 13,958) of the original collection could not be successfully geocoded unambiguously. Those were excluded from the dataset of the new *visual prototype [P-INTF]*. As the geocoding was fully automated, some errors regarding the geocoding might still be present in the resulting spatial data.

Algorithm 4.2: Location Parsing using Google Geocoding API

```

public Location parseLocationForArtifact(Artifact artifact) throws NoResultsException {
    String address = extractAddress(artifact);

    try {
        GeoApiContext context = new GeoApiContext.Builder()
            .apiKey(googleGeocodeApiKey)
            .build();
        GeocodingResult[] results = GeocodingApi.geocode(context, address).await();

        if (results.length == 1) {
            log.debug("Address_found_for_" + artifact.getOnbImageId() + ":" +
                results[0].formattedAddress + "_raw_address:" + address);
            return handleSingleResult(artifact, results[0]);
        }
        if (results.length > 1) {
            return handleMultiResult(artifact, results);
        }
    } catch (Exception e) {
        log.debug(e.getMessage() + "_" + artifact.getOnbImageId());
        return null;
    }
    throw new NoResultsException();
}

```

Algorithm 4.3: Handling of single results from Geocoding API

```

private Location handleSingleResult(Artifact artifact, GeocodingResult result) throws Exception {
    String postalCode = extractPostalCodeFromResult(result);

    if (postalCode.equals(getFormattedPostalCode(artifact.getDistrict()))) {
        return createAndSaveLocation(result);
    }
    throw new Exception(artifact.getOnbImageId() +
        ":_single_result_found_but_postal_code_was_wrong._artifact's_postal_code_is_" +
        getFormattedPostalCode(artifact.getDistrict()) + "_but_location's_is_" + postalCode);
}

```

4.3 User interface design

The basic user-interface structure is a side-by-side view of the map and the various visualizations. Figure 4.5 shows an early sketch of the basic structure. The individual UI-Elements are following Material Design Guidelines[Goo20b] as it provides consistent UI concepts and components which help to create appealing web applications even without the specific knowledge of a professional UI-Designer.

4.4 Design Methodology

Based on the definition of Users, Tasks, and Data in the previous section, I want to use this section to illustrate the design process and explain the reasoning behind individual design decisions.

Algorithm 4.4: Handling of multiple results from Geocoding API

```

private Location handleMultiResult(Artifact artifact, GeocodingResult[] results) {
    List<GeocodingResult> correctPostalCode = Arrays.stream(results)
        .filter(geocodingResult -> containsPostalCode(
            getFormattedPostalCode(artifact.getDistrict()), geocodingResult))
        .collect(Collectors.toList());

    if (correctPostalCode.size() == 1) {
        return createAndSaveLocation(correctPostalCode.get(0));
    } else if (correctPostalCode.size() > 1) {

        double maxDistance = 0D;

        for (int i = 0; i < correctPostalCode.size(); i++) {
            for (int j = i + 1; j < correctPostalCode.size(); j++) {
                double calculatedDistance = distance(
                    correctPostalCode.get(i).geometry.location.lat,
                    correctPostalCode.get(j).geometry.location.lat,
                    correctPostalCode.get(i).geometry.location.lng,
                    correctPostalCode.get(j).geometry.location.lng
                );
                if (calculatedDistance > maxDistance)
                    maxDistance = calculatedDistance;
            }
        }

        if (maxDistance < 300D) {
            log.debug("multiple_locations_with_correct_postalcode_found_but_all_closer_than" +
                "300m_too_each_other,_return_the_first_one.");
            return createAndSaveLocation(correctPostalCode.get(0));
        }
    }
    throw new IllegalArgumentException("no_correct_postalcode_or_too_many_matches");
}

```

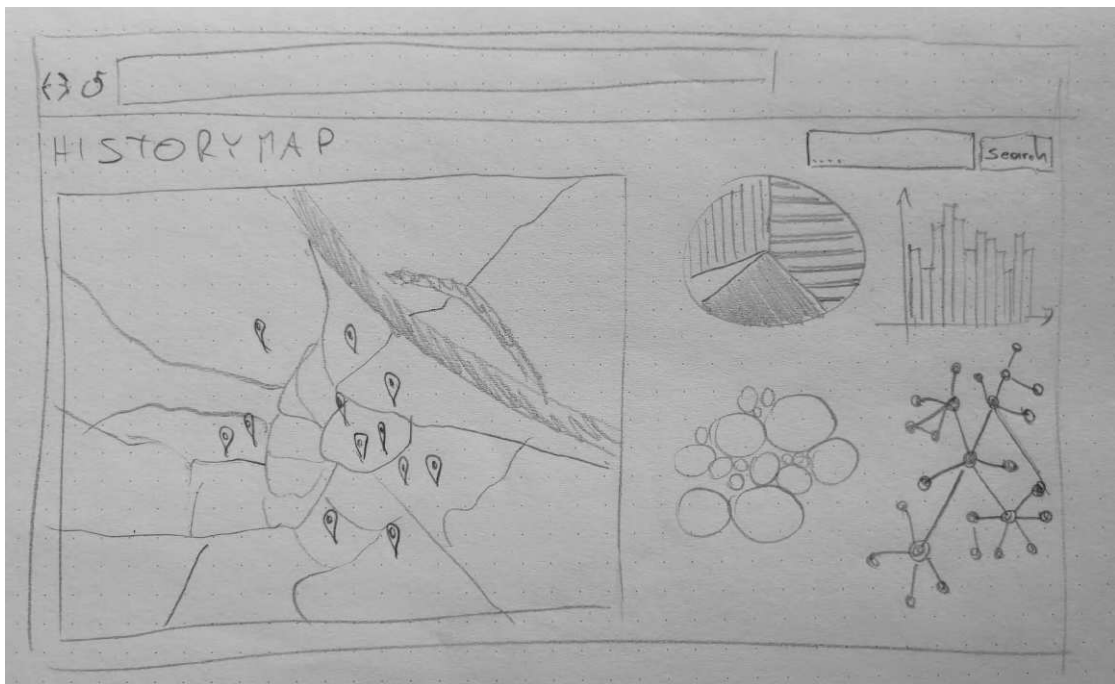


Figure 4.5: User Interface Design - Sketch

4.4.1 Multiple coordinated views

All visualizations show a different aspect of the same dataset and allow it to filter this dataset. As they all deal with the same dataset, it is important to recognize the relationship between the individual visualizations. E.g., what happens if I modify or filter one visualization? Will this affect the other visualizations representing the same dataset? To allow the user to use the relationship between the visualizations while exploring the connection, I decided to implement multiple connected views where filtering one view (or visualization) will directly affect the display of another visualization. E.g., filtering the time period using the temporal bar chart will result in fewer authors being displayed in the author bubble chart. Using multiple views in data visualizations is a common practice that can be highly useful[WBWK00]. Therefore I decided to include this concept in my prototype, although it also might increase the complexity, both programmatically and in regards to user experience and cognitive overhead[WBWK00].

4.4.2 Color Encoded Map Icons

To allow for a quick association between an artifact and its age, every icon on the map has a color encoded border reflecting its approximate age. The temporal bar chart (see Section 4.4.5) acts as a reference, allowing the user to understand the meaning behind the colors. This decision was made as the map is the central visualization of the interface and should convey as much information as possible. Of course, with the requirement that the information is not overwhelming the user. The icon borders aim to be a non-invasive way to convey additional information.

In general, color can be a potentially problematic tool when conveying numerical data. Nowell and Hix summarize this dilemma but propose that under certain circumstances, color encoding can be useful[NH97]. Acknowledging this fact, the moderate application of color encoding in the new *visual prototype [P-INTF]* aims to enhance user understanding. Nonetheless, no information is solely encoded through color (e.g., the age of artifacts can also be inspected in the detailed view when clicking on an icon or by filtering the temporal bar chart) to account for the possible downsides of color encoding. See figure 4.6 for an example of how the final color encoded icons (including the temporal bar chart acting as an explanation) look like in the *visual prototype [P-INTF]*.

4.4.3 Pictograms for conveying additional information on map icons

As part of the design process, I explored the idea of conveying additional information directly on the map icons as pictograms. Siti and Swasti showed that pictograms could be especially effective for conveying information in the context of cultural heritage [CS17].

My approach was to encode the information about the technique category (see Section 4.2.3) as small pictograms on the top right corner of the map icons, additionally to the color encoded icon-border. See Figure 4.7 for an example implementation. Although the implementation is feasible, the concept of pictograms was not included in the final

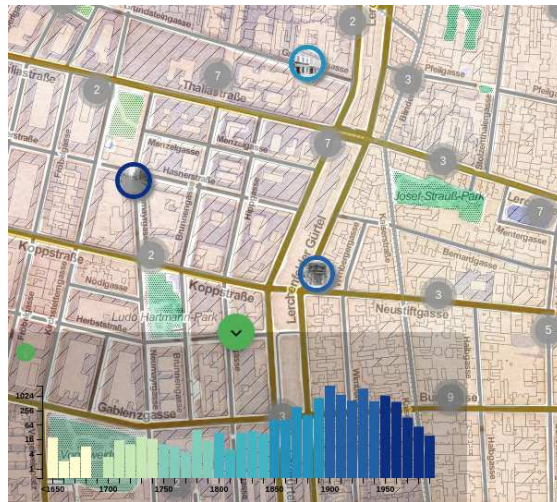


Figure 4.6: Color Encoded Icons



Figure 4.7: Pictograms on map icons

prototype as it would have worked against one of the central goals of the study: not to overload the interface to allow casual users to explore the collection. The icon-pictograms would add a dimension that would make it harder for the user to comprehend its meaning. Additionally, other visualizations convey the same information (e.g., see Section 4.4.9), hence omitting this information in the map seems justifiable.

4.4.4 Map Clustering

For being able to display a large number of elements on the map, I decided to utilize Map Clustering, a popular technique when displaying many elements on a single map. It is utilized by many Mapping libraries (e.g. Leaflet)[Vla20]. Map Clustering allows a large number of elements to be displayed on the map without deteriorating performance

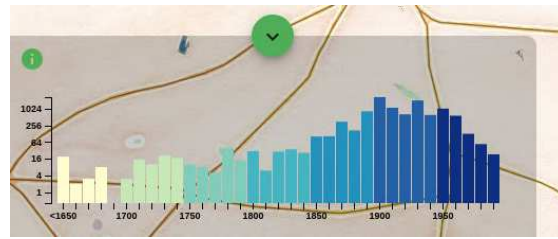


Figure 4.8: Temporal bar chart

(more than 10.000 in our case). Elements close to each other are combined to a cluster that displays the number of elements inside the cluster. When zooming the map closer to a cluster, the individual elements are loaded instead of the cluster.

4.4.5 Bar Chart for Time series

The temporal data dimension is, besides the spatial dimension, especially important for the collection (see Section 4.2.3). The data quality regarding the creation date is especially good compared to other parameters (e.g., creation technique). Thus, I decided to put the chart representing the temporal data dimension in a prominent place in the visual prototype and choose a visualization type that is both easy to understand and expressive. Bar charts are generally easier to read than radial designs and are very popular for a wide range of applications [WDG⁺20]. Furthermore, they are especially useful for time series [Szo82]. My assumption was that bar charts being the most common graph [TSA14], most people will have no problems reading a simple bar chart, so I chose this type of chart. Besides the bars' length, the color of the bars helps build a direct correlation between an artifact on the map and a specific time period (see 4.4.2). Schloss et al. showed that color could help to structure the data conveyed in a chart [SGS⁺19].

As the number of artifacts is distributed very irregularly along the time axis (most pictures created between 1900 and 1950 and very few pictures in other periods), I chose to utilize a logarithmic scale with base two to allow a better relative comparison between neighboring decades.

See Figure 4.8 for the final implementation of the temporal bar chart.

4.4.6 Author Bubble chart

The *search-based interface [S-INTF]* does list the authors for most artifacts on their details page. However, it is hard to judge for a user how significant each author's contribution to the collection is. For solving this, I introduced a bubble chart displaying each author's contribution to the collection. Each bubble represents an author's contribution (number of artifacts created in the collection), the bigger the bubble, the more significant the contribution. According to Bertin, size is an ordered visual variable and can be used to visualize relative proportions of numbers; in our case, the number of artifacts per author

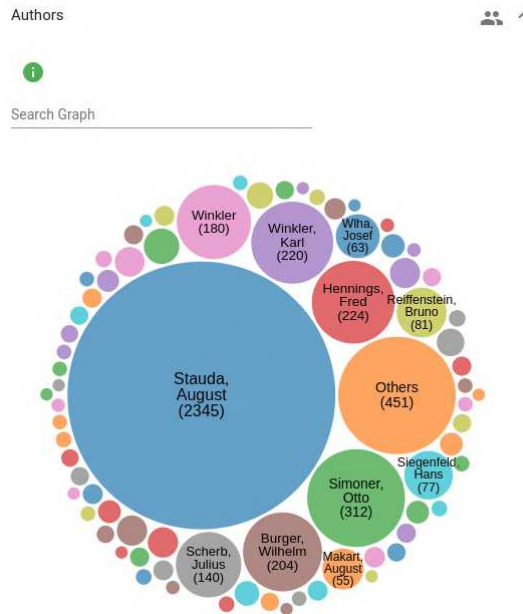


Figure 4.9: Author Bubble Chart

[Ber10, Rot17]. Although not entirely precise, bubble charts are useful for getting a quick overview of relative sizes[PAW12].

Because of the large number of different authors in the collection, authors with less than five contributions in the collections are consolidated into one bubble *others*. For easier access to small bubbles, an additional search function is available, which allows highlighting of authors.

See figure 4.9 for the example implementation of the author bubble chart.

4.4.7 Keyword Cloud

Keywords are critical in the *search-based interface [S-INTF]* as each artifact has them assigned, and both quick search and expert search allow searching for images via assigned keywords. Thus, they allow logical segregation of the collection into different topics by filtering via those keywords. As the assigned keywords are already present in the original data, no further data processing was necessary. I chose the technique of *Keyword Clouds* as an appropriate visualization for the available keywords in the collection. Word Clouds are a popular technique to visualize texts or word collection. Although they cannot be considered accurate analysis tools, they can serve as a suitable starting point for a more in-depth analysis of a given data set [WXQW14]. The word cloud displays all available keywords and allows the filtering of the displayed artifacts by clicking on the keywords. See figure 4.10 for the implementation of the keyword cloud.

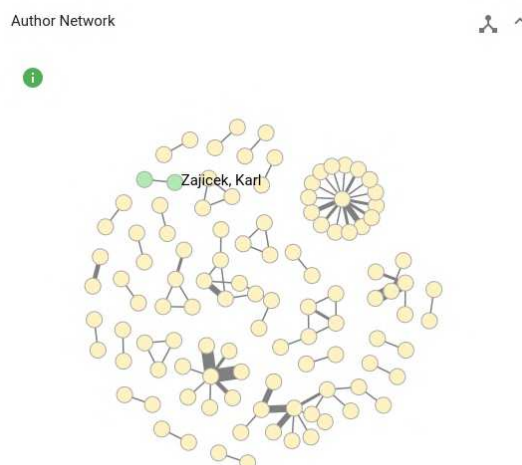


Figure 4.11: Author Network Graph

data structure. Stasko and Zhang propose the *Sunburst* visualization for hierarchical data. It is a visualization similar to a Treemap but organized radially. According to Stasko and Zhang, there is a subjective user-preference for the Sunburst visualization compared to the Treemap, although the performance of the two is comparable [SZ00]. Because user engagement is an important goal, I chose the Sunburst visualization for visualizing the artifact creation techniques. Because of the large number of creation technique sub-types, I chose a zoomable Sunburst chart to visualize even proportionally small leaves. Zooming a Sunburst chart allows us to see even low-level details, which otherwise would be hard to inspect [AHAJ15]. See figure 4.12 for the sample implementation of the Technique Sunburst Chart.

4.4.10 Map Overlays

To improve the user experience and to enrich the encrypted information of the spatial visualization, I added the optional feature to add historical map overlays to the displayed map. They are allowing the user to inspect the artifacts of the collections in an additional historical context. Additionally to improved user engagement, this feature should allow casual users and experts to draw additional conclusions and discover new information patterns. Historical Map Overlays are a useful tool for those purposes [WT92]. The used data material is available through public accessible Map Tile Services provided by the city of Vienna [Mag20] and the Austrian Academy of Sciences [Dig20]. See figure 4.13 for a sample of an active historical map overlay.

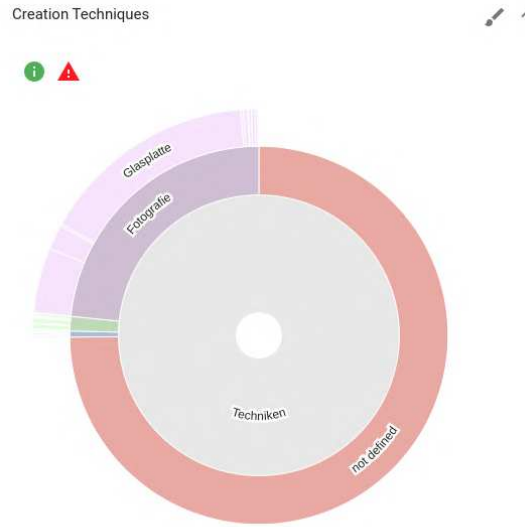


Figure 4.12: Creation Techniques Sunburst Chart

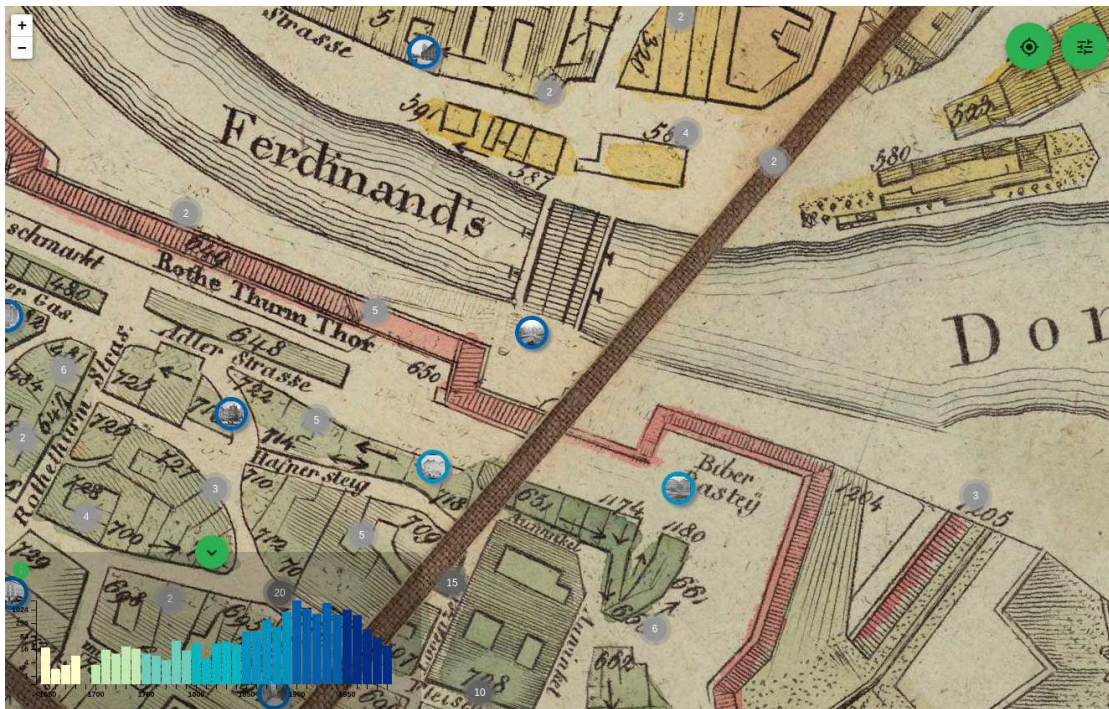


Figure 4.13: Historical Map Overlay on the Map



Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar.
The approved original version of this thesis is available in print at TU Wien Bibliothek.

Prototype Implementation

Based on the design methodology and the definition of users, tasks, and dataset (as described in Chapter 4), an interactive visual prototype based on the same dataset as the existing interface (see Section 4.1.1) was implemented. This chapter summarizes the prototype implementation details, including individual steps of the development process and the documentation of technical details.

5.1 Prototype Characteristics

It was essential to create an accessible, easy-to-use, and interactive prototype. Hence, the decision was made to create a public web application that can be run directly in a modern web-browser. Also, the existing interface (www.bildarchivaustria.at) is accessible via a regular web-browser, thus limiting the usability for the end-user (e.g., by forcing users to install an application on their computer) was not an option.

As interactivity is a major topic of this thesis, the prototype utilizes modern web-technologies (JavaScript/TypeScript, CSS3) to create a modern, interesting interface that aims to be both easy to use and at the same time appealing and entertaining.

The application was deployed on a public web-server is reachable via the domain `https://historymap.wien`.

5.2 Architecture

As the application can be seen as a standalone application without any external services, I decided to use a traditional three-layered Services Architecture. It consists of a Layer for the User Interface, a Layer for Business Services, and a Layer for Data Access and Manipulation (see figure 5.1).

In our case, the three layers are divided as follows:

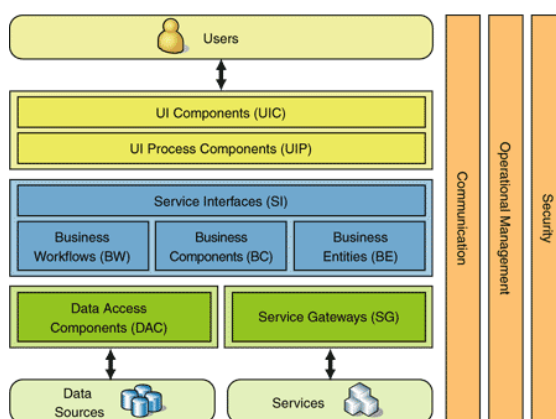


Figure 5.1: Three-Layered Services Application [Mic20]

- **Presentation Layer:** This Layer contains all user-facing elements running in the users' browser. This includes the user interface written in HTML, JavaScript/TypeScript and CSS, the map (using Leaflet), and all visualizations (using D3.js).
- **Service Layer:** This Layer contains REST interfaces, which are called by the UI, and all necessary Service Classes for preparing Data objects to be used in the UI Layer. It also contains Functions for data manipulation, which are currently not called by the UI. The service layer is written in Java using the Spring Framework.
- **Persistence Layer:** This Layer is responsible for the communication with the database (in our case PostgreSQL). It provides functions for querying the database and persisting objects into the database. The persistence layer is implemented using Java, Spring Data, and Hibernate.

5.3 Technologies

Figure 5.2 shows the applied technologies per layer. Java was used as the primary programming language for the persistence and service layer. It is a popular language with a lot of available resources and libraries, and my prior experience allowed me to create a suitable backend codebase quickly. This was especially true in conjunction with the Spring Framework and Hibernate, as I already have experience with those technologies as well. PostgreSQL was chosen as data storage as it is open-source, and I already had some previous experience with it.

In contrast to the backend technologies, I was not very familiar with the applied frontend technologies before starting this thesis. I chose angular as the main frontend framework as it uses TypeScript by default, which is easier to use when used to strongly typed languages. Leaflet is an open-source library for developing mapping applications. I evaluated Google Maps as a mapping solution but decided against it because of its closed

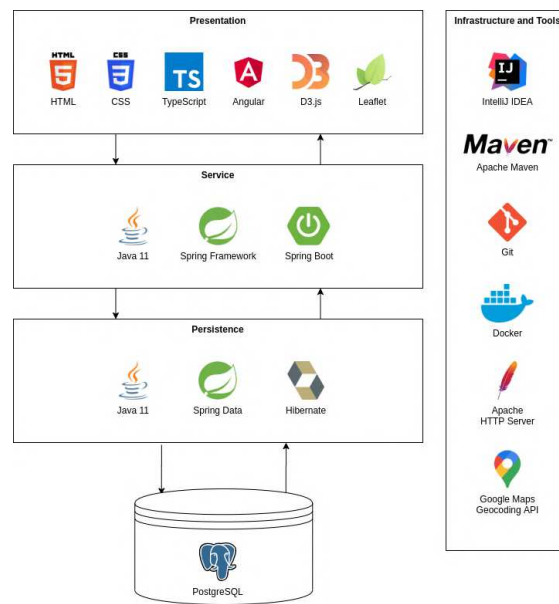


Figure 5.2: Technology Stack

nature (certain features are limited or only available as paid features). I chose Leaflet over Google Maps.

For creating all other data visualizations I used D3.js as it open-source, has a big community, and allows creating performant dynamic data-visualizations that run directly in the browser. As it turned out, both D3.js and Leaflet are not fully TypeScript compatible, so I had to write some of the front end-code in JavaScript.

As development tools, I used IntelliJ IDEA as an integrated development environment as it works well with both Java and JavaScript/TypeScript code. I used Git for version control and Maven as build management tool.

Finally, the application was deployed as Docker-Container on a private Server with an Apache HTTP Server as Reverse Proxy.

5.4 Data Analysis and Usage Permission

As I worked with an existing data set, one of the first tasks was to analyze the dataset and see what kind of data dimensions are available and can be used. This was done by manually looking at the data of the *search-based interface [S-INTF]*. It is essential to mention that this development task overlaps with the design activities described in Section 4.2.3. Please refer to that section for detailed information about the identified data dimensions.

The interface and the contained data is provided by the Austrian National Library, which explicitly "asserts no own proprietary exploitation rights" to their publicly available digital

content[Aus20]. Hence using the provided data in an academic sense is allowed. To make sure, I additionally enquired with Mag. Peter Prokop from the Austrian National Library on August 17th, 2017. He confirmed that my understanding is correct, and the data usage is allowed.

5.5 Data Model

Based on the identified data dimensions, I developed an entity-relationship model that stores the extracted data in a structured way. Figure 5.3 shows the final data model used for the prototype.

5.6 Data Extraction and Migration

Unfortunately, no API access was available, and hence it was necessary to resort to web scraping. This meant extracting data from a webpage using HTTP, which can be resource and time consuming [DSS⁺19]. Fortunately, in the case of the *search-based interface [S-INTF]*, this task was rather easy because of the static nature of the interface. It uses basic HTML and no Scripting, and the Data is displayed in a very structured way, which allows for easy data extraction. For downloading and parsing the HTML elements, I used the jsoup HTML parser[Jon20]. The elements extracted by jsoup were then directly transformed into our data model and stored in the database.

5.7 Source Code and Project Structure

The source code is structured as a multi-module project (see figure 5.4) and is publicly available at <https://github.com/brintal/historymap.wien>. The following modules exist:

- **domain:** [persistence layer] Contains all domain entities (Java classes) which map to the database tables of the data model (see section 5.5).
- **persistence:** [persistence layer] Contains all Spring Data Repositories (Java interfaces) and Hibernate configurations used to read from and write to the database.
- **frontend-ng:** [presentation layer] Contains all frontend specific code and configuration. This includes Angular components (written in Typescript), HTML and CSS files, map specific code (written in Typescript/JavaScript using Leaflet) visualization specific code (written in TypeScript/JavaScript using D3.js).
- **service:** [service layer] Contains all service classes (written in Java) responsible for functionality which is provided by the standard functionality of the Spring Data Repositories (e.g. data transformation, entity parsing and more).

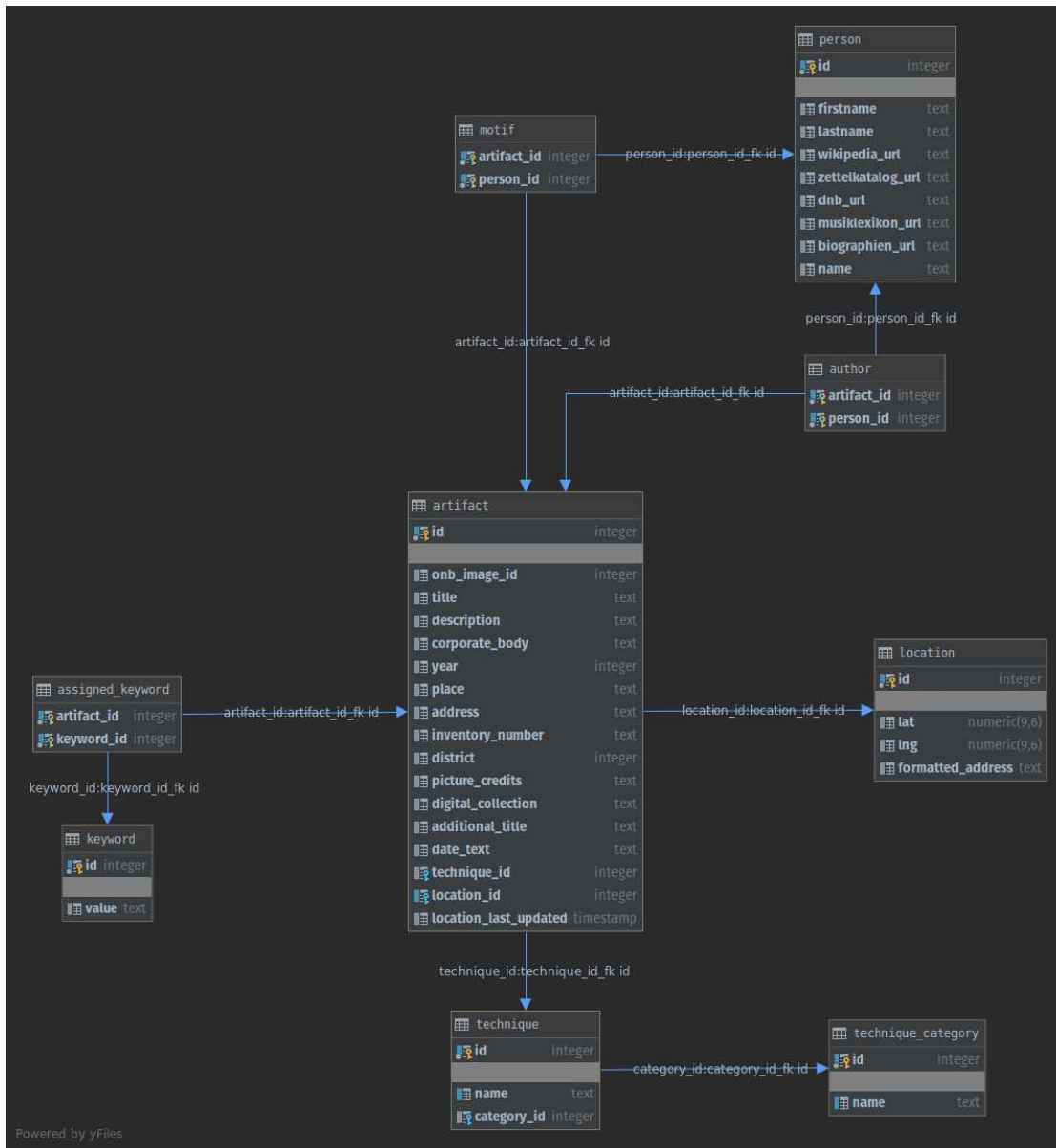


Figure 5.3: Prototype Database Model

5. PROTOTYPE IMPLEMENTATION

- **app:** Contains the Docker-configuration and is used to package and create the final artifact which can be deployed.

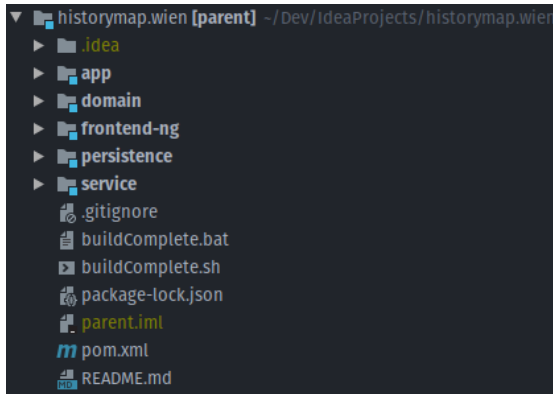


Figure 5.4: Source Code Project Structure

5.8 Screens

The following images illustrate how the actual *visual prototype [P-INTF]* looks like and what features and different functionalities it has. Each screenshot shows a different aspect of the prototype.

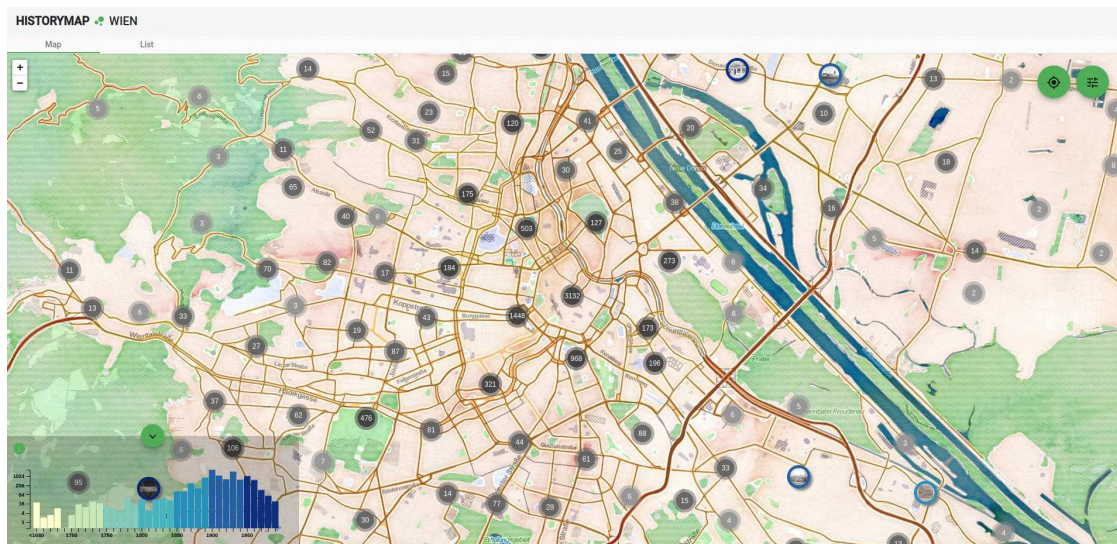


Figure 5.5: Screen: Homescreen

The Map (including the temporal bar chart on the bottom left of the screen) is visible. The Sidebar is hidden by default.

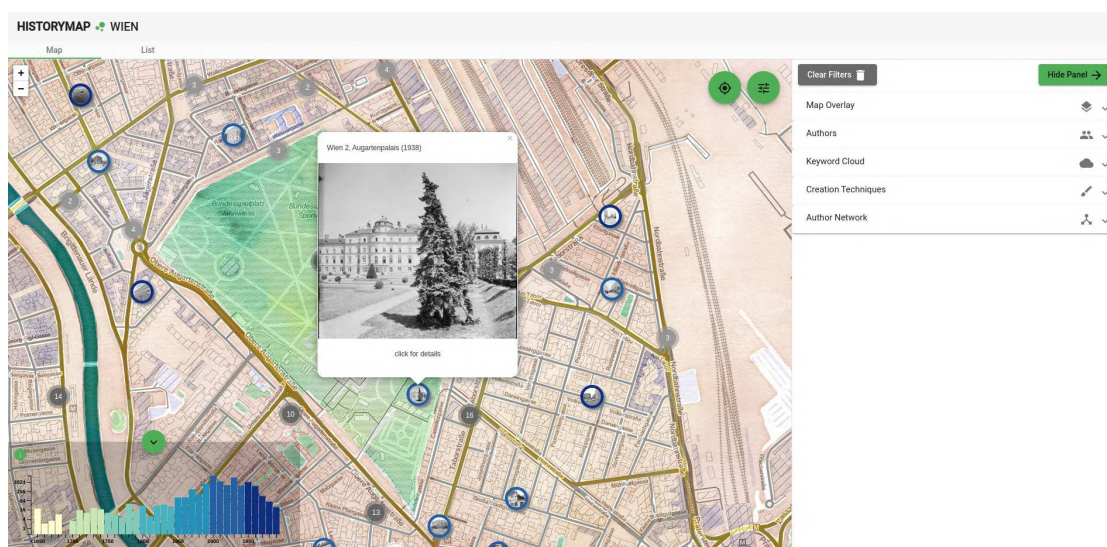


Figure 5.6: Screen: Image Popup

The Image Popup is shown when the user clicks on a map icon.

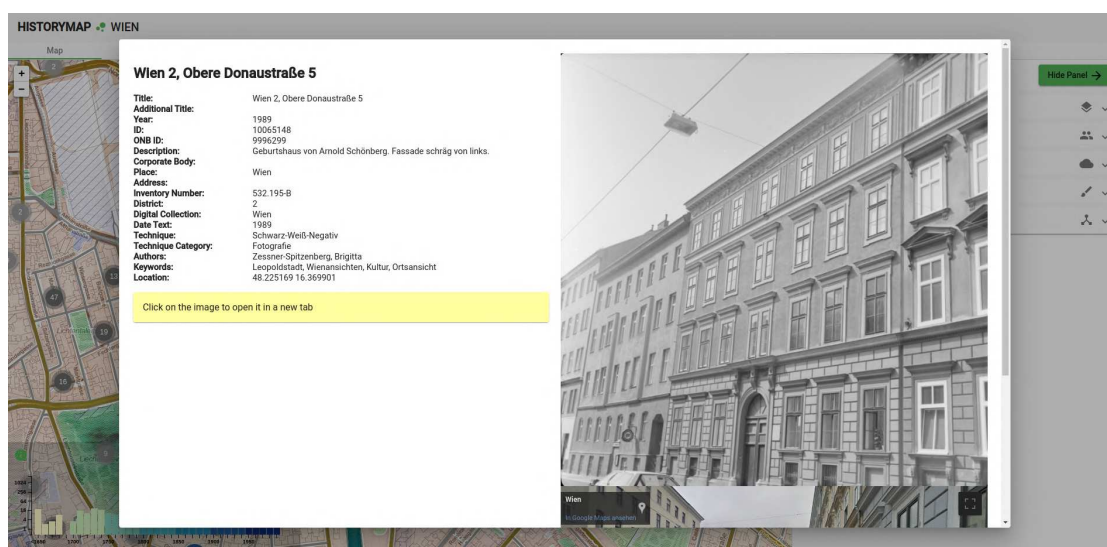


Figure 5.7: Screen: Artifact Detail View 1

Displayed when clicking on "click here for details" inside the image popup(see figure 5.6)

5. PROTOTYPE IMPLEMENTATION

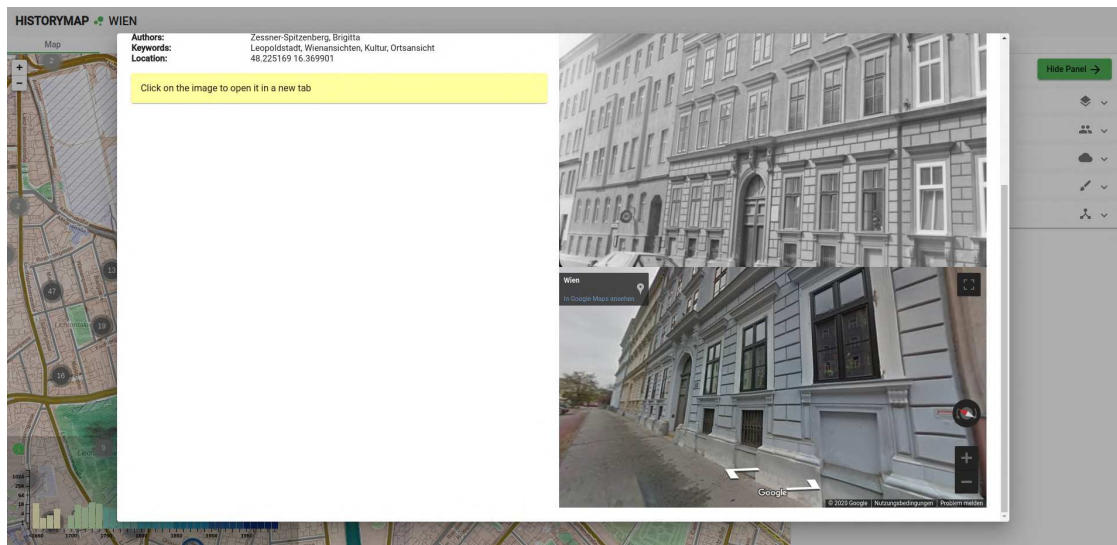


Figure 5.8: Screen: Artifact Detail View 2

When scrolling down the detail view, a Google Street View Image of the same location is shown to illustrate the contrast between the old and new image.

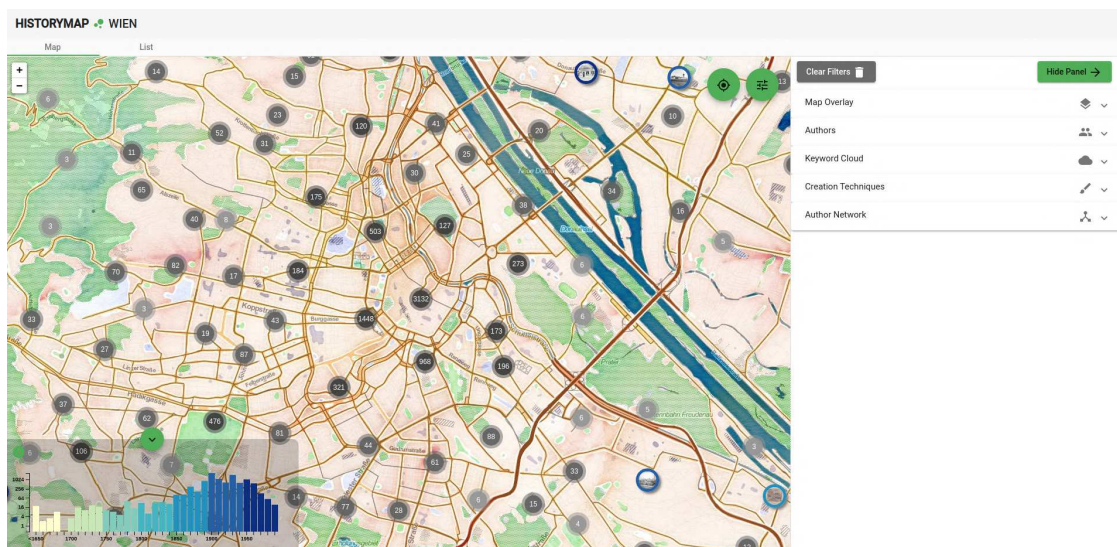


Figure 5.9: Screen: Sidebar

Map View with the open sidebar. The sidebar contains a wide range of visualizations and filter options and can be opened by the green round button located at the top right corner of the map

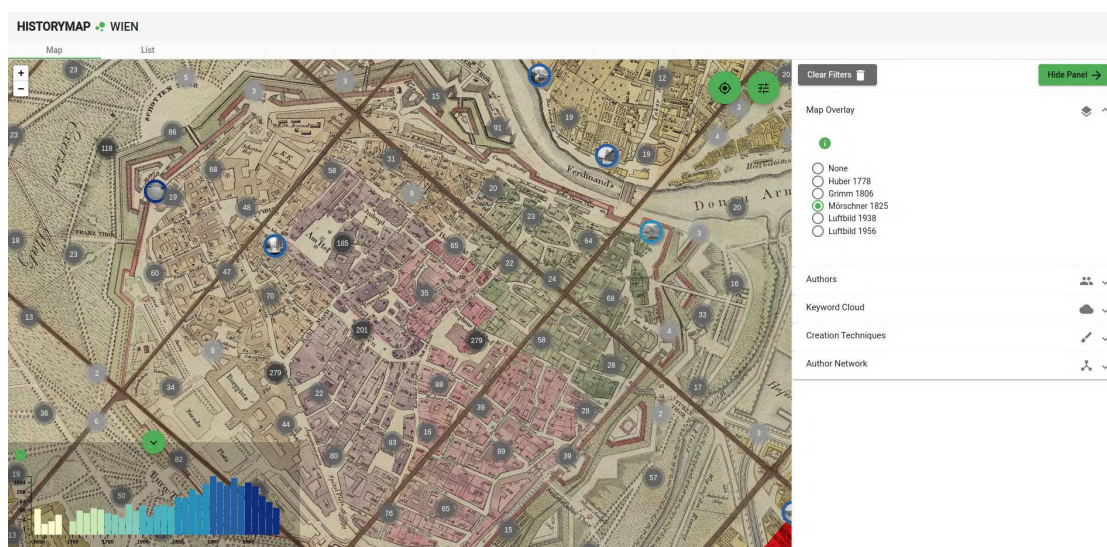


Figure 5.10: Screen: Map Overlays

The Menu option "Map Overlay" inside the Sidebar allows for selecting which historical map should be shown above the normal map. When selecting an option, it will immediately change the appearance of the map as shown here.

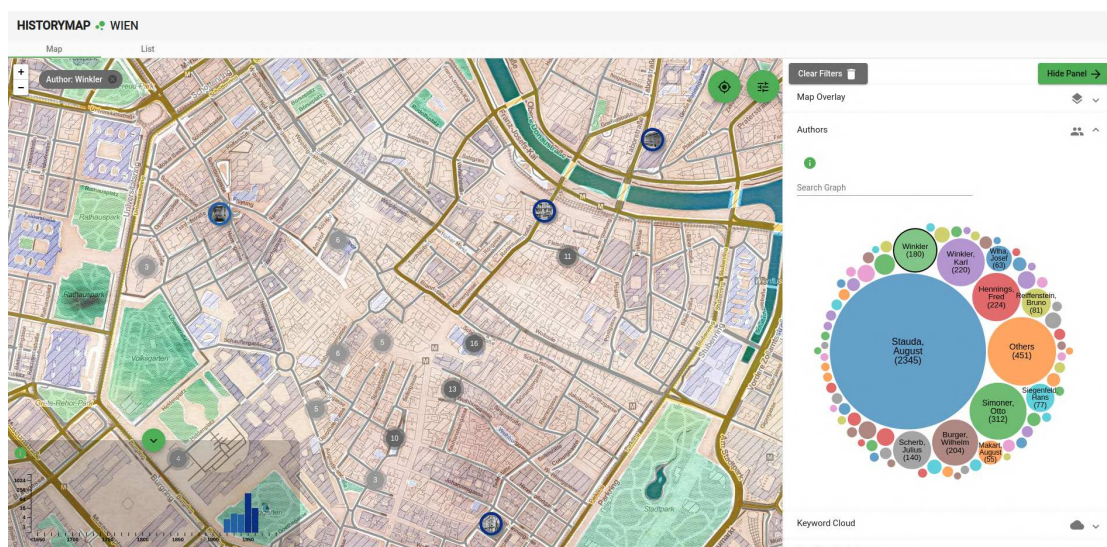


Figure 5.11: Screen: Author Bubble Chart

Selecting the option "Authors" in the sidebar displays a bubble chart visualization that allows filtering the displayed artifacts by their author. An artist is selected by clicking directly on the corresponding circle. To find authors with smaller contributions (identified by a smaller circle), the user can search the author through a search field above the visualization.

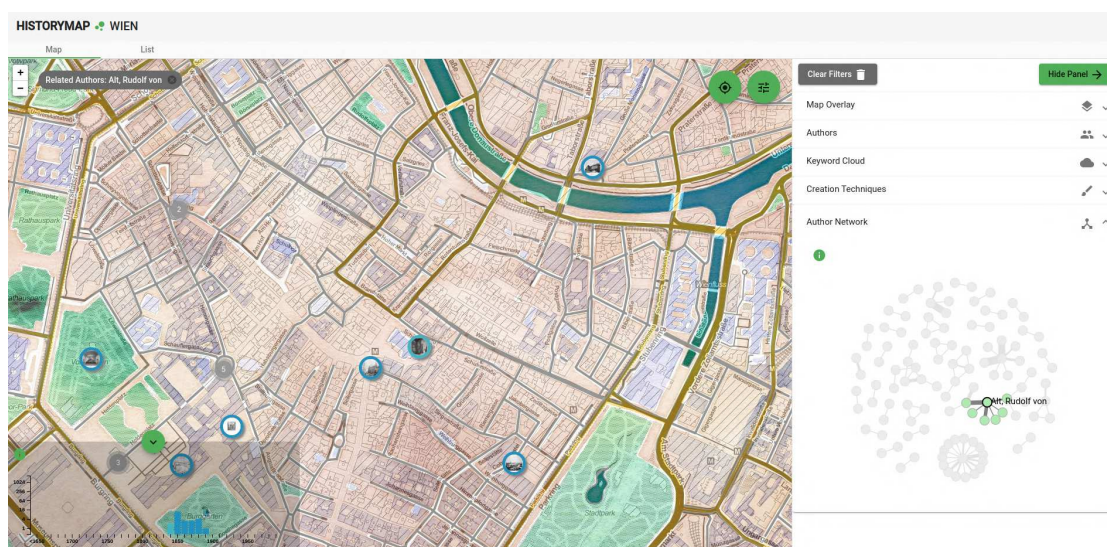


Figure 5.14: Screen: Author Network Visualization

By clicking on the option "Author Network" in the sidebar, users can inspect the work relationships between authors. Every author is represented by a node in the graph. If two authors were working together (if at least one artifact exists which has those two authors assigned), their nodes are connected. The connection between nodes is thicker, the more often the connected authors were working with each other. The graph can be used to filter the artifacts as well. Clicking on a node selects an author and disables all unconnected nodes in the graph. Clicking on another node selects the other related author, which results in only pictures created by those two authors together being displayed on the map.

5. PROTOTYPE IMPLEMENTATION

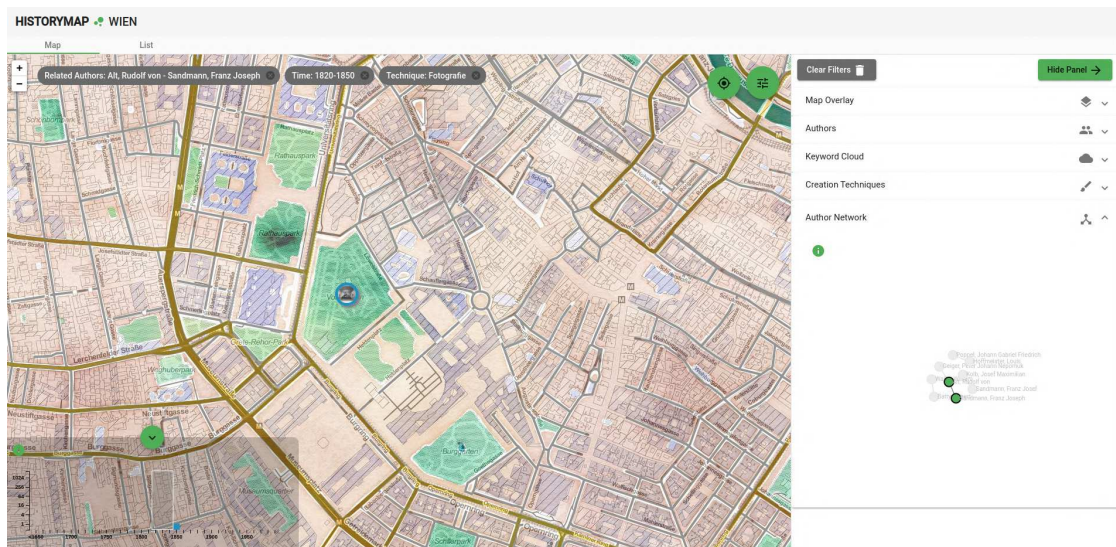


Figure 5.15: Screen: Multiple Active Filters

Any filter that is applied through any visualization is visible through grey filter indicators on the top left of the screen. Every filter can be individually removed by clicking the x-button directly on the filter indicator.

Icon	Id	Title	Year	District	Authors	Longitude	Latitude
	13763720	Wien, Plan	1147	1	Vasquez, Karl Graf	16.373921	48.208265
	13769695	Wien 1, Stephansdom	1245	1		16.373471	48.208411
	13769702	Wien 1, Stephansdom	1245	1		16.373471	48.208411
	13769600	Wien 1, Stephansdom	1245	1		16.373471	48.208411
	11466726	Wien 1, Erzbischöfliches Palais	1430	1		16.373402	48.209333
	10527129	Wien 1, Hofburg	1458	1		16.366054	48.207658
	17770554	Panorama Wien von Norden	1550	1	Stimmer, Tobias, Frig, Ludwig, Comesina, Albert	16.375001	48.205742
	9587804	Wien 1, Hofburg	1559	1		16.366054	48.207658
	10130489	Wien 11, Schloss Neugebäude	1600	11		16.443912	48.160485
	9587811	Wien 1, Hofburg	1600	1	Wiesböck, Karl	16.366054	48.207658

Items per page: 10 | 1 - 10 of 10889

Figure 5.16: Screen: Artifact List View

Displayed when clicking the Tab "List" on top of the screen, the list view allows for a textual search of the collection. The result list can be sorted by clicking on the column header and browsed by clicking on the arrows located on the bottom right of the screen.



Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar.
The approved original version of this thesis is available in print at TU Wien Bibliothek.

Prototype Evaluation

6.1 Purpose and Method

The goal of the evaluation is to **assess the *visual prototype [P-INTF]* in terms of usability, explorative nature and user engagement**. The *search-based interface [S-INTF]* will be used as a reference for the participants and they will be asked to evaluate both interfaces. The assessment of the *search-based interface [S-INTF]* mainly is intended for creating awareness about certain concepts (e.g. expert-search) that should help the evaluation of the *visual prototype [P-INTF]*. **A direct comparison** of the two interfaces in terms of efficiency **is not the goal**.

As one central theme of this thesis is exploratory search and how visualizations can help user engagement, I concluded a semi-structured interview would be the best way to assess user experience and understand how users are exploring and using the *visual prototype [P-INTF]*. The subjective nature of visualizations [FCB06] also requires to not only look at *how* users evaluate visualizations but also to understand *why* they evaluate the way they do. Because of this reasoning, together with the fact that qualitative studies have been known to be effective when evaluating visualizations [TSF08], I concluded a qualitative study could be beneficial in my case as well. Lam et al. suggest a set of qualitative evaluation techniques to assess user experience in visualization tools [LBI⁺12]. These include:

- informal evaluation
- usability test
- field observation
- laboratory questionnaire

Those techniques are the foundation for the actual interview design. The following section describes the actual interview guidelines derived from these concepts.

6.2 Interview Guidelines

This section describes the general interview structure for the interviews. The listed points should be understood as general interview guidelines and not a strict plan for conducting the interviews. Some questions might have been omitted or explored in detail, depending on the participant and the observed behavior during the interview. The interview guidelines also contain a set of tasks that each participant is asked to complete. The tasks intend to enforce an interaction with a wide range of features provided by the *visual prototype [P-INTF]*.

Because of the research goal focusing on experimental browsing of digital cultural heritage collections, **it is more important to identify how the user study participants use the interfaces rather than if they were able to complete a certain task.** E.g. it is possible that participants take a long time to complete a certain task, but use the interface in an exploratory fashion and are eager to discover new aspects of the interface. In that case the explorative nature of the interface would weight more than the absolute time users take to finish certain tasks.

Additionally, to the semi-structured interview, every participant was given a small questionnaire for evaluating how understandable and useful each visualization and feature of the prototype is. The questionnaire was designed as a simple multiple-choice survey (see Section 6.2.3).

Please refer to Appendix A for the individual results and interview notes of each user study participant. For the summary of the evaluation as a whole, please refer to Section 6.5.

6.2.1 Introduction

- **Motivation:** Short explanation by the Interviewer of what a cultural heritage collection is and how this thesis tries to complement existing digital collections.
- **Interviewee characteristics:**
 - What is your previous experience with (digital) cultural heritage collections?
 - What is your previous experience with data visualizations?
 - Would you consider yourself interested in history and arts?
- **Introduction of the two interfaces:** Introduction of the two interfaces by the Interviewer:
 - <http://www.bildarchivaustria.at>
 - <https://historymap.wien>

6.2.2 Interview Tasks and Questions

Search-Based Interface:

The following tasks and questions refer to the *search-based interface [S-INTF]* (<http://www.bildarchivaustria.at>).

- Please take a moment to browse around the collection
- What is your first impression of the interface?
- How would you assess the general usability of the interface?
- Find a picture of Schönbrunn
- Find an old picture (created before 1750) of Schönbrunn
- Find a picture of a church close to your home
- Which decade (10 years period) has the most pictures in the collection?

Visual Prototype:

The following tasks and questions refer to the *visual prototype [P-INTF]* (<https://historymap.wien>).

- Please take a moment to browse around the collection
- What is your first impression of the interface?
- How would you assess the general usability of the interface?
- Find a picture of Schönbrunn
- Find an old picture (created before 1750) of Schönbrunn
- Find a picture of a church close to your home
- Which decade (10 years period) has the most pictures in the collection?
- Which author created most pictures in the collection?
- Which author created most pictures dated 1800 or earlier?
- Find 2 authors which were working often with each other and open a picture created by them.
- List a few keywords used often throughout the collection
- Find a few pictures of buildings standing in the place of the old vienna city wall.

6.2.3 Questionnaire: Evaluation of Visualizations

Map including artifact Icons

How easy to understand is this feature?

not understandable at all difficult to understand somewhat understandable very easy to understand

How useful is this feature?

not useful at all not particularly useful somewhat useful very useful

Temporal Bar Chart

How easy to understand is this feature?

not understandable at all difficult to understand somewhat understandable very easy to understand

How useful is this feature?

not useful at all not particularly useful somewhat useful very useful

Author Bubble Chart

How easy to understand is this feature?

not understandable at all difficult to understand somewhat understandable very easy to understand

How useful is this feature?

not useful at all not particularly useful somewhat useful very useful

Map Overlays

How easy to understand is this feature?

not understandable at all difficult to understand somewhat understandable very easy to understand

How useful is this feature?

not useful at all not particularly useful somewhat useful very useful

Keyword Cloud

How easy to understand is this feature?

not understandable at all difficult to understand somewhat understandable very easy to understand

How useful is this feature?

not useful at all not particularly useful somewhat useful very useful

Creation Techniques Chart

How easy to understand is this feature?

not understandable at all difficult to understand somewhat understandable very easy to understand

How useful is this feature?

not useful at all not particularly useful somewhat useful very useful

Author Network Chart

How easy to understand is this feature?

not understandable at all difficult to understand somewhat understandable very easy to understand

How useful is this feature?

not useful at all not particularly useful somewhat useful very useful

Artifact List

How easy to understand is this feature?

not understandable at all difficult to understand somewhat understandable very easy to understand

How useful is this feature?

not useful at all not particularly useful somewhat useful very useful

6.2.4 Final Questions and Conclusion

- In the future, which tool are you more likely to use if you want to search / browse old pictures of vienna? Why?
- Pro/Contra of the *search-based interface [S-INTF]*?
- Pro/Contra of the new interface?
- What features/improvements are you missing in the new interface?

6.3 Participants

My thesis and the visual prototype's main target group are casual users with no expert domain knowledge. General interest in history and cultural heritage collections is desirable but also not necessary.

Therefore I chose the group of participants matching these criteria. In total nine participants took part in the evaluation. The participants include software engineers, mathematicians, artists, researchers, and students. Most of them stated they have some interest in a collection of historical images of Vienna though none of them was extraordinarily interested in the topic. This was important for me as I wanted to evaluate user engagement, and a more pronounced interest in the topic might have falsified the results.

The interview notes (see next section) also contain a short background description for each participant.

6. PROTOTYPE EVALUATION

Feature	Participant	Usefulness				Understandability			
		0	1	2	3	0	1	2	3
Map including artifact icons	1				x				x
	2				x				x
	3				x				x
	4			x				x	
	5				x				x
	6				x				x
	7				x			x	
	8				x				x
	9				x				x
Temporal Bar Chart	1				x				x
	2				x				x
	3			x					x
	4			x			x		
	5				x			x	
	6				x				x
	7				x			x	
	8				x				x
	9				x			x	
Author Bubble Chart	1				x				x
	2				x				x
	3				x				x
	4				x				x
	5		x					x	
	6				x				x
	7				x				x
	8				x				x
	9				x				x
Map Overlays	1				x				x
	2				x				x
	3				x				x
	4				x				x
	5				x				x
	6				x				x
	7				x				x
	8				x				x
	9				x				x
Keyword Cloud	1							x	
	2								x
	3								x
	4								x
	5								x
	6								x
	7								x
	8								x
	9								x
Creation Techniques Chart	1								x
	2								x
	3								x
	4								x
	5								x
	6								x
	7								x
	8								x
	9								x
Author Network	1								x
	2								x
	3								x
	4								x
	5								x
	6								x
	7								x
	8								x
	9								x
Artifact List	1								x
	2								x
	3								x
	4								x
	5								x
	6								x
	7								x
	8								x
	9								x

Table 6.1: Prototype Evaluation - Questionnaire Results

6.4 Interview Notes

Please see Appendix A for the interview notes for each participant. Additionally, table 6.1 summarizes the results of the multiple choice questionnaire (see Section 6.2.3) given to each interviewee.

6.5 Findings

This section summarizes the findings of the prototype evaluation and lists identified potential improvements.

6.5.1 Observations

In general, the participants reacted mostly positively towards the *visual prototype [P-INTF]*. Nearly all participants describe the interface as modern, well-arranged, and easy

to use. The participants were visibly motivated by the interface and tend to engage in casual, explorative browsing.

When talking about positive aspects of the interface, most participants mentioned the map as a prominent quality that draws attention to the interface and promotes usage. The provided map overlays were also well liked by the participants. The general positive notion towards the spatial aspects of the interface is also reflected by the answers the participants gave in the questionnaire (see Table 6.1).

When looking at the other visualizations, results were mixed. Some visualizations perform better than others in regarding usefulness and understandability (see Table 6.1). The author bubble chart was apparently both easy to use and easy to understand. Some participants (interviewee 3 and 4) even stated that they think the bubble chart visualization should be used in place of other visualizations, e.g., the keyword cloud. Because other interviewees actually enjoyed using the keyword cloud very much (e.g., interviewee 7) and proposed using the word cloud visualization more, it is inconclusive if one of the two visualizations performs better than the other.

Both the questionnaire (see Table 6.1) and the rest of the interview showed that the temporal bar chart was mostly understood and liked by the interviewees. Nevertheless, there were some apparent usability issues identified during the interviews. E.g., the selection functionality is not intuitive enough, and the labeling is not explicit enough.

The evaluation clearly showed that people enjoyed exploring the data visually. E.g., interviewee 1 mentioned explicitly that through the visualizations, he suddenly considered a particular aspect (e.g., the relationship between authors) interesting, which he would never even look at when using the *search-based interface [S-INTF]*. Nevertheless, the interviewees' interest and professional field seemed to affect the evaluation of visualizations considerably. E.g., the creation techniques sunburst visualization mostly performed mediocre at best in regards to usefulness but was well-liked and described as very useful by interviewee 5, who has a strong background in arts and is working in that field.

Especially while observing the interviewees completing the given tasks (see Section 6.2.2), one could see that users could complete the tasks very efficiently. Nevertheless, participants needed some preparation time for each visualization before being able to use it effectively. E.g., interviewees 8 and 9 were both familiar with these types of data visualizations and could complete the given tasks much faster than most other participants. Because of the wide range of available visualizations and filter options, it took most users some time to understand all the concepts necessary to use the interface effectively. This might be one reason why some tasks being completed faster on the *search-based interface [S-INTF]*.

Furthermore, many participants seemed to prefer using the search function, which is currently very limited in the *visual prototype [P-INTF]*. The search function was also one of the most criticized aspects of the *visual prototype [P-INTF]*. It is only possible to search the collection on the list view but not on the map view. Additionally, the search function only allows for basic text search but not for more sophisticated queries. In this

regard, the *search-based interface* [*S-INTF*] is more mature and capable than the *visual prototype* [*P-INTF*]. The observation during the interviews clearly showed that most users want an easy way to search for what they are looking for. In the *visual prototype* [*P-INTF*] implementation, the search function is only accessible via the artifact list and, therefore, hard to find for the users. Table 6.1 also shows that participants generally respond positively towards the list view (which embeds the search function) but had some problems using it.

Another significant aspect identified during the evaluation were issues with the logical UI-structure. In general, the navigational concepts were understood right away by the users. However, some users (e.g., interviewees 3 and 6) tried to apply a filter through visualizations on the map view and then switched over to the list view realizing that the filter is not affecting the list. The same happens when the user tries to search in the list view and switches to the map view. This implies that there remain some improvements regarding user interface design and UI-component-hierarchy.

As mentioned before, the map view was well received by nearly all of the participants. Nevertheless, some artifacts of the collection were placed imprecisely on the map, which sparked some discussion during the interviews. In Section 4.2.3, I presented the process for geocoding all artifact locations, which proved to work well but still produces a few erroneous results. Some participants (e.g., interviewees 8 and 9) wish they could report those imprecise locations or correct them themselves.

6.5.2 Summary of Findings

Based on the observations (see Section 6.5.1) and the questionnaire results (see Table 6.1) we can summarize the findings as follows:

- Participants generally value the interactivity and the modern user interface design. The interface visibly promotes casual exploration and can engage the users through its interactive interface and visualizations.
- The spatial aspects (map, map overlays) were received highly positively by all of the participants.
- The prototype still has some minor usability issues and software bugs.
- The search function of the prototype should have more features to support more complex queries.
- The visualizations require basic knowledge about data visualization techniques. Participants without appropriate knowledge needed a bit longer to understand the meaning behind the visualizations in order to know how to use them.
- With the appropriate knowledge (or after some time of usage and building this knowledge), the visualizations are understandable and usable for nearly all the participants.

- The visualizations help the users to understand certain aspects about the collection (e.g., collection composition, relationships between entities).

6.5.3 Identified Improvements

The *visual prototype [P-INTF]* was generally received very positively. Nevertheless, some improvements could be identified during the interviews which are listed below:

- **Temporal Bar Chart:** The temporal bar chart should be clickable (selecting a single bar by clicking on it), and the drag-and-drop functionality should be reworked. It is often not clear why a particular boundary is selected when clicking and dragging. Additionally, the labeling should be more concise, and the selected timeframe should be immediately visible directly above the chart. Furthermore, a tooltip when hovering a bar would be helpful
- **Search Function:** The search function should be more sophisticated and more comfortable to find (located somewhere on the top affecting both list view and map view). It should support both quick search queries with multiple parameters and allow for detailed queries on parameter level similar to the expert search of the *search-based interface [S-INTF]*.
- **Different UI-Hierarchy:** The map view and list view should not be completely independent of each other. Changing a filter in the sidebar should equally affect the map and the list, and one should be able to seamlessly switch between the two views without any reloading or refreshing of the page.
- **Easier Filtering:** The active filters should be shown above each respective visualization with an option to clear this one filter. Additionally, the filter overview (currently on the top left of the screen) should be more visually noticeable.
- **More Map Overlays:** The map style was well-liked, but participants sometimes preferred a cleaner map similar to the default street map of Google Maps. Adding additional map overlays could help.
- **District Filter:** Adding visualization and filter options for the district the artifacts are located in, could help the users quickly filter for relevant images on the map. It also might be interesting to see the distribution of artifacts per district.
- **Remove Map-Cluster Boundaries:** Leaflet includes a standard feature when using clusters on the map, which shows on mouseover the area in which all the artifacts of a cluster are located in. This proved to be of little value for the user and is somewhat confusing. This marker should be removed.
- **Author Network: Selecting nodes by clicking on the link:** Some users tried to select two nodes in the author network chart by clicking on the link between them. This feature should be added.

- **Wording Improvements:** The wording of labels and info-texts, in general, is not very precise. This should be reworked and reviewed by a domain expert.
- **Visualization search function:** Like the search-field above the author bubble chart, every other visualization should contain such an individual search function, allowing us to use the visualization more effectively.
- **Additional visualizations:** There is more potential for additional visualization. E.g., a title cloud similar to the keyword cloud visualizes words used in the artifacts' titles. Another potential visualization is a district bubble chart showing the distribution of artifacts across all city districts.
- **Crowd-sourcing and community features:** As mentioned in the last paragraph of Section 6.5.1, some artifacts still have an imprecise location assigned. This could be corrected by integrating crowd-sourcing and community features allowing users to actively participate in the improvement and maintenance of the collection. Allowing users to bookmark pictures and to upload their own pictures could be an excellent addition as well.

Conclusion

7.1 Conclusion and main contributions

In this thesis, I explored the importance and potential of data visualizations in the context of cultural heritage data collections. Apart from a broad literature review and State of the Art Report, I presented a visual prototype illustrating the value of visualizations compared to a more traditional search-based approach. In conclusion, I want to reference my original research questions (see Section 1.2) and discuss them with the thesis's achieved results:

RQ *Can visualization support interactive exploration of cultural heritage collections across several data dimensions in a meaningful way?*

To answer this question, I implemented an interactive prototype for an existing digital cultural heritage collection and evaluated it through a qualitative user study. The findings summarized in Sections 6.5.1 and 6.5.2 clearly show, that utilizing Information Visualization when designing interfaces for browsing digital cultural heritage collections, improve usability, accessibility, and user engagement. The prototype demonstrates how multiple data dimensions can be visualized effectively. Especially visualizing spatial data proved to be a very efficient tool as it facilitates exploration and manages to keep the user's interest much better than a simple search-based interface. Additionally, I compiled a comprehensible State of the Art Report covering other concepts dealing with the same problems. I summarized how they are utilizing visualizations to facilitate the exploration of cultural heritage collections.

SQ1 *Which visual approaches already exist, either documented in academic literature or publicly available as web-based applications?*

Through my thorough State of the Art Report in Chapter 3, I could give a well-informed analysis of the various approaches which exist in the context of visualization of digital cultural heritage data collections. The categorization in section 3.4 shows which concepts are already well explored and what aspects need more attention.

SQ2 *Is it possible to improve existing approaches, in terms of interface design and implementation, and how?*

Based on the State of the Art Report's results, I positioned my work complementing existing approaches by filling the gap, especially in dealing with spatial data. I demonstrated that it is, in fact, possible to utilize visualizations to improve existing approaches when dealing with multiple data dimensions, including spatial, temporal, and categorial data. Through my evaluation, I showed that by doing so, it is possible to create additional value for users by promoting exploration and interactivity.

7.2 Future work

The evaluation clearly showed that people are engaged by the use of Information Visualization and the results are indeed promising. Especially the visualization of the spatial data has proven to be a precious approach when trying to increase accessibility and promote exploration of a digital cultural heritage collection. This should be explored further and applied to different datasets and contexts. Some spatial data is often present in cultural heritage collections (e.g., findspot and place of origin of archaeological artifacts). Creating visual interfaces focusing on this spatial aspect and providing a map to navigate the collection is an effective way to promote explorative search like it was shown in this thesis. These interfaces could help to make topics like Archaeology, Anthropology, and others more accessible and relatable to the general public.

Additionally, an extensive quantitative study with more participants could help complement the qualitative study presented in this thesis to understand more about the characteristics and strength of specific visualizations.

Although the dataset considered in this thesis (see Section 4.2.3) is generally of good quality, it still was noticeable that some aspects of the data are missing or hard to work with. It can only be expected that when working with different cultural heritage data collections, challenges regarding data quality and structure will arise. *Visual Analytics* has proven to be a useful concept when dealing with large, heterogeneous, and incomplete data sets. Hence, utilizing those concepts (statistics and data mining combined with Information Visualization) could also be beneficial in the context of cultural heritage data collections[DCK13, TC05]. This thesis illustrates how interactive Information Visualization interfaces can help to make structured cultural heritage data accessible. Visual Analytics could help to increase accessibility further and to gain insights into very complex data sets.

List of Figures

2.1	Search activities: Explorative Search focuses on <i>learning</i> and <i>investigating</i> whereas a <i>Lookup</i> generally requires specialized domain knowledge. [Mar06]	6
2.2	Classical library organization [Wik16]	8
2.3	Development of the unemployment rate of OECD countries [web17]	11
2.4	Reference model for visualizations [CMS99]	12
2.5	How music travels - example of a multidimensional (temporal, spatial, categorical) Information Visualization[Man16]	12
3.1	NYPL Public Domain Release 2016 Visualization [New16]	18
3.2	Cinema Redux by Brendan Dawes [Bre16]	19
3.3	Tyne and Wear Archives and Museums: Collectionsdive. Left: Thumbnail explore mode; Right: Relationships [Tyn16]	19
3.4	Cooper-Hewitt: A Timeline of Event Horizons [Coo16]	20
3.5	SFMOMA ArtScope [San16]	20
3.6	Kräutli's visualization of artists in Tate's collection[Krä16]	21
3.7	Categorical enriched searchbox interface in the digital collection of the Städel museum [MUS16]	21
3.8	Color-based search of the web-interface <i>Color Lens</i> [met16]	22
3.9	Getty Research Institute: Auction Market Visualization [Get16]	23
3.10	Spatial visualization <i>Streete View</i> by the Het Scheepvaartmuseum using old maps and illustrations. Left: Map overview, Right: detailed illustration of a selected point of interest [Het16]	23
3.11	Visualization <i>Decade Summary</i> of the platform <i>Prints and printmaking Australia</i> . Encodes temporal and categorical data dimensions and also allows conclusions about the scope and structure of the collection[Pri16].	24
3.12	The two different browsing-modes of the <i>Virtual Hampson Museum</i> [Ham16]	25
3.13	Spatial 3D visualization of the 82nd & Fifth collection [The16]	25
3.14	Temporal visualization (vertical timeline) of the 82nd & Fifth collection[The16]	26
3.15	<i>The Speculative W@anderverse</i> [Uta20]	26
3.16	<i>ArchivesZ</i> interface[KSNA07]	27
3.17	Past Visions[FH 20b]	28
3.18	Manly Images[Whi12]	28
3.19	Discover the Queenslander[Sta20b]	29
3.20	The Bohemian Bookshelf[THC12]	30
		83

3.21 Rothko Viz[Eve20]	31
3.22 Deutsche Digitale Bibliothek Visualized[FH 20a]	31
3.23 British Museum: The Museum of the World[Bri20]	32
3.24 Corley Explorer: Map overview (left) and Tag-view(right)[Sta20a]	33
3.25 The Institutional Harvest[And20]	33
3.26 Behind The Utopia[Lil18]	34
3.27 Number of approaches per tag	35
4.1 Bildarchiv Austria - Current Search-Based Web Interface [Ös20]	40
4.2 Bildarchiv Austria - Search interface	41
4.3 Design Triangle[MA14]	42
4.4 Bildarchiv Austria - Collection Data Entry	44
4.5 User Interface Design - Sketch	48
4.6 Color Encoded Icons	50
4.7 Pictograms on map icons	50
4.8 Temporal bar chart	51
4.9 Author Bubble Chart	52
4.10 Keyword Cloud	53
4.11 Author Network Graph	54
4.12 Creation Techniques Sunburst Chart	55
4.13 Historical Map Overlay on the Map	55
5.1 Three-Layered Services Application [Mic20]	58
5.2 Technology Stack	59
5.3 Prototype Database Model	61
5.4 Source Code Project Structure	62
5.5 Screen: Homescreen	62
5.6 Screen: Image Popup	63
5.7 Screen: Artifact Detail View 1	63
5.8 Screen: Artifact Detail View 2	64
5.9 Screen: Sidebar	64
5.10 Screen: Map Overlays	65
5.11 Screen: Author Bubble Chart	65
5.12 Screen: Keyword Cloud Visualization	66
5.13 Screen: Creation Techniques Visualization	66
5.14 Screen: Author Network Visualization	67
5.15 Screen: Multiple Active Filters	68
5.16 Screen: Artifact List View	68
5.17 Screen: Mobile Map View	69
5.18 Screen: Mobile Image Popup	69
5.19 Screen: Mobile Details View	69

List of Tables

3.1	Overview of tags of each individual approach.	36
6.1	Prototype Evaluation - Questionnaire Results	76



Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar.
The approved original version of this thesis is available in print at TU Wien Bibliothek.

List of Algorithms

4.1	Assigning Technique Category	46
4.2	Location Parsing using Google Geocoding API	47
4.3	Handling of single results from Geocoding API	47
4.4	Handling of multiple results from Geocoding API	48



Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar.
The approved original version of this thesis is available in print at TU Wien Bibliothek.

Bibliography

- [AHAJ15] Ragaad AlTarawneh, Shah Rukh Humayoun, and Abdel-karim Al-Jaafreh. Towards optimizing the sunburst visualization for smart mobile devices. In *INTERACT 2015 Adjunct Proceedings: 15th IFIP TC. 13 International Conference on Human-Computer Interaction 14-18 September 2015, Bamberg, Germany*, volume 22, page 2015323. University of Bamberg Press, 2015.
- [AJ15] S. A. Akanmu and Z. Jamaludin. Measuring infovis' decision support effectiveness: From theory to practice. In *2015 Science and Information Conference (SAI)*, pages 560–564, 2015.
- [And20] Andrew Merrindahl, Mitchell Whitelaw. The institutional harvest. <http://institutionalharvest.net/>, April 2020.
- [Ass16] Association for Computing Machinery. Acm dl digital library. <http://dl.acm.org/>, December 2016.
- [Aus20] Austrian National Library. Use - Österreichische Nationalbibliothek. <https://www.onb.ac.at/en/use>, July 2020.
- [Ber10] Jacques Bertin. *Semiology of Graphics: Diagrams, Networks, Maps*. Esri Press, Redlands, 2010. ISBN: 978-1-589-48261-6.
- [BKD] Viktoria Brüggemann, Sarah Kreisler, and Marian Dörk. Museale bestände im web: Eine untersuchung von acht digitalen sammlungen. *EVA Berlin: Proceedings of the Electronic Media and Visual Arts conference*, pages 227–236.
- [BM13] M. Brehmer and T. Munzner. A multi-level typology of abstract visualization tasks. *IEEE Transactions on Visualization and Computer Graphics*, 19(12):2376–2385, 2013.
- [BN13] Emanuele Bellini and Paolo Nesi. Metadata quality assessment tool for open access cultural heritage institutional repositories. In *International Conference on Information Technologies for Performing Arts, Media Access, and Entertainment*, pages 90–103. Springer, 2013.

- [Bre16] Brendan Dawes. Brendan dawes - cinema redux TM. <http://www.brendandawes.com/projects/cinemaredux>, December 2016.
- [Bri20] British Museum, Google Cultural Institute. The museum of the world. <https://britishmuseum.withgoogle.com/>, March 2020.
- [Car07] Stuart Card. Information visualization. In A. Sears and J.A. Jacko, editors, *The Human-Computer Interaction Handbook: Fundamentals, Evolving Technologies and Emerging Applications, Second Edition*, pages 509–543. CRC Press, 2007.
- [CMS99] Stuart K. Card, Jock D. Mackinlay, and Ben Shneiderman, editors. *Readings in Information Visualization: Using Vision to Think*. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 1999.
- [Coo16] Cooper Hewitt, Smithsonian Design Museum. A timeline of event horizons | cooper hewitt labs. <https://labs.cooperhewitt.org/2013/a-timeline-of-event-horizons/>, December 2016.
- [CS10] A Corns and Robert Shaw. Cultural heritage spatial data infrastructures (sdi)-unlocking the potential of our cultural landscape data. In *30th EARSeL Symposium*, pages 1–8, 2010.
- [CS17] Siti Clara and Wirania Swasty. Pictogram on signage as an effective communication. *Jurnal Sositologi*, 16:166–175, 08 2017.
- [Dav05] Gareth Davey. What is museum fatigue? *Visitor Studies Today*, 8:17–21, 01 2005.
- [DCK13] Robin De Croon and Verbert Katrien. From information visualization to visual analytics. University Lecture - <https://osf.io/mdfwb/download>, 2013.
- [DCW11] Marian Dörk, Sheelagh Carpendale, and Carey Williamson. The information flaneur: A fresh look at information seeking. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '11, pages 1215–1224, New York, NY, USA, 2011. ACM.
- [Dig20] Digital Humanities-Projekt, an der BAS:IS der Österreichischen Akademie der Wissenschaften. Sammlung Woldan Kartenportal. <http://sammlung.woldan.oeaw.ac.at/>, November 2020.
- [DLB13] Marian Dörk, Heidi Lam, and Omar Benjelloun. Accentuating visualization parameters to guide exploration. In *CHI '13 Extended Abstracts on Human Factors in Computing Systems*, CHI EA '13, pages 1755–1760, New York, NY, USA, 2013. ACM.

- [DPC17] Marian Dörk, Christopher Pietsch, and Gabriel Credico. One view is not enough: High-level visualizations of a large cultural collection. *Information Design Journal*, 23(1):39–47, 2017.
- [DSS⁺19] R. Diouf, E. N. Sarr, O. Sall, B. Birregah, M. Bouso, and S. N. Mbaye. Web scraping: State-of-the-art and areas of application. In *2019 IEEE International Conference on Big Data (Big Data)*, pages 6040–6042, 2019.
- [DvL05] Christian Heath Dirk vom Lehn. Accounting for new technology in museum exhibitions. *International Journal of Arts Management*, 7(3):11–21, 2005.
- [DWC12] Marian Dörk, Carey Williamson, and Sheelagh Carpendale. Navigating tomorrow’s web: From searching and browsing to visual exploration. *ACM Trans. Web*, 6(3):13:1–13:28, October 2012.
- [Eve20] Everardo Reyes. Rothko viz: Visualizing 201 paintings by mark rothko. <http://ereyes.net/rothkoviz/>, March 2020.
- [FCB06] Sarah Faisal, Paul A. Cairns, and Ann Blandford. Subjective information visualizations. In *Proc. Workshop on Combining Visualisation and Interaction to Facilitate Scientific Exploration and Discovery*. British HCI 2006 conference, 2006.
- [FH 20a] FH Potsdam. Deutsche digiatle bibliothek visualized. <https://uclab.fh-potsdam.de/ddb/>, March 2020.
- [FH 20b] FH Potsdam. Past visions. <https://uclab.fh-potsdam.de/fw4/en/vis/>, March 2020.
- [GD15] Katrin Glinka and Marian Dörk. Museum im Display. Visualisierung kultureller Sammlungen. In *EVA Berlin 2015 (Electronic Media and Visual Arts)*, pages 21–27, 11 2015.
- [GEC98] Nahum Gershon, Stephen G. Eick, and Stuart Card. Information visualization. *Interactions*, 5(2):9–15, March 1998.
- [Get16] Getty Research Institute. Auction market visualization (getty research institute). <https://www.getty.edu/research/tools/provenance/zoomify/>, December 2016.
- [GMD15] Katrin Glinka, Sebastian Meier, and Marian Dörk. Visualising the » un-seen « : Towards critical approaches and strategies of inclusion in digital cultural heritage interfaces. In *Culture and Computer Science, Berlin, Volume: XIII.*, pages 105–117, 2015.
- [Goo16a] Google Inc. Google. <http://www.google.com>, December 2016.

- [Goo16b] Google Inc. Google scholar. <https://scholar.google.at/>, December 2016.
- [Goo20a] Google Inc. Google geocoding api. <https://developers.google.com/maps/documentation/geocoding/start>, December 2020.
- [Goo20b] Google Inc. Material design. <https://material.io/design>, July 2020.
- [GPDD16] Katrin Glinka, Christopher Pietsch, Carsten Dilba, and Marian Dörk. Linking structure, texture and context in a visualization of historical drawings by frederick william iv (1795-1861). *International Journal for Digital Art History*, 0(2), 2016.
- [Ham16] Hampson Archeological Museum State Park. Browse the hampson collection. <http://hampson.cast.uark.edu/browse.htm>, December 2016.
- [Het16] Het Scheepvaartmuseum. Het scheepvaartmuseum straet view. <http://atlassen.hetscheepvaartmuseum.nl/>, December 2016.
- [HW10] Sam Hinton and Mitchell Whitelaw. Exploring the digital commons: An approach to the visualisation of large heritage datasets. In *Proceedings of the 2010 International Conference on Electronic Visualisation and the Arts, EVA'10*, pages 51–58, Swinton, UK, UK, 2010. British Computer Society.
- [IEE16] IEEE (Institute of Electrical and Electronics Engineers) . Ieee xplore digital library. <http://ieeexplore.ieee.org/>, December 2016.
- [INT16] INTK. Visitor motivation in online museum audiences — intk. <http://www.intk.com/en/ideas/visitor-motivation-in-online-museum-audiences>, December 2016.
- [Jon20] Jonathan Hedley. jsoup html parser. <https://jsoup.org/>, July 2020.
- [Krä16] Florian Kräutli. *Visualising Cultural Data: Exploring Digital Collections Through Timeline Visualisations*. PhD thesis, Royal College of Art, April 2016.
- [KS14] A. Kerren and F. Schreiber. Why integrate infovis and scivis?: An example from systems biology. *IEEE Computer Graphics and Applications*, 34(6):69–73, 2014.
- [KSNA07] Jeanne Kramer-Smyth, Morimichi Nishigaki, and Tim Anglade. Archivesz: Visualizing archival collections. 01 2007.

- [Lan07] Monika Lanzenberger. Grundlagen methodischen Arbeitens - Informationsvisualisierung. University Lecture - http://www.ifs.tuwien.ac.at/~mlanzenberger/teaching/ps/ws07/stuff/2007.10.17.gma.VO01_4-1.pdf, 2007.
- [LBI⁺12] H. Lam, E. Bertini, P. Isenberg, C. Plaisant, and S. Carpendale. Empirical studies in information visualization: Seven scenarios. *IEEE Transactions on Visualization and Computer Graphics*, 18(9):1520–1536, 2012.
- [LBLM13] Irene Lopatovska, Iris Bierlein, Heather Lember, and Eleanor Meyer. Exploring requirements for online art collections. In *Proceedings of the 76th ASIS&T Annual Meeting: Beyond the Cloud: Rethinking Information Boundaries*, ASIST '13, pages 111:1–111:4, Silver Springs, MD, USA, 2013. American Society for Information Science.
- [Lil18] Lilla Tóth, Xiaoyi Xu, Qin Yang, Juuso Koponen. Behind the utopia. <https://www.facebook.com/aaltovisualcommunicationdesign/posts/1091764944315178/>, September 2018.
- [Lim13] Manuel Lima. *Visual Complexity: Mapping Patterns of Information*. Princeton Architectural Press, 2013.
- [MA14] Silvia Miksch and Wolfgang Aigner. Special section on visual analytics: A matter of time: Applying a data-users-tasks design triangle to visual analytics of time-oriented data. *Comput. Graph.*, 38:286–290, February 2014.
- [Mag20] Magistrat Wien - Magistratsabteilung 01 - Wien Digital. Web map tile service - stadt wien. <http://maps.wien.gv.at/wmts/1.0.0/WMTSCapabilities.xml>, July 2020.
- [Man16] Manuel Lima. visualcomplexity.com | how music travels. <http://www.visualcomplexity.com/vc/project.cfm?id=804>, November 2016.
- [Mar89] Gary Marchionini. Information-seeking strategies of novices using a full-text electronic encyclopedia. *Journal of the American Society for Information Science*, 40(1):54–66, 1989.
- [Mar06] Gary Marchionini. Exploratory search: From finding to understanding. *Commun. ACM*, 49(4):41–46, April 2006.
- [met16] metaLAB (at) Harvard. Colour lens. <http://colourlens.org/>, December 2016.
- [Mic20] Microsoft Corporation. Three-layered services application | microsoft docs. [https://docs.microsoft.com/en-us/previous-versions/msp-n-p/ff648105\(v=pandp.10\)](https://docs.microsoft.com/en-us/previous-versions/msp-n-p/ff648105(v=pandp.10)), July 2020.

- [ML13] Diogo Oliveira de Melo and Alneu de Andrade Lopes. Data visualization and relevance feedback applied to information retrieval. In *Proceedings of the Sixth Workshop on Ph.D. Students in Information and Knowledge Management*, PIKM '13, pages 27–32, New York, NY, USA, 2013. ACM.
- [MS13] S. Maddison and M. Sawyer. *The Women's Movement in Protest, Institutions and the Internet: Australia in transnational perspective*. Routledge Research in Gender and Politics. Taylor & Francis, 2013.
- [MUS16] STÄDEL MUSEUM. Städel Digitale Sammlung. <https://digitalesammlung.staedelmuseum.de/index.html#/>, December 2016.
- [New16] New York Public Library Labs. Nypl public domain release 2016 - visualization. <http://publicdomain.nypl.org/pd-visualization/>, December 2016.
- [NH97] Lucille Terry Nowell and Deborah Hix. *Graphical Encoding for Information Visualization: Using Icon Color, Shape, and Size to Convey Nominal and Quantitative Data*. PhD thesis, 1997. AAI9957085.
- [PAW12] Kalpesh Padia, Yasmin AlNoamany, and Michele C. Weigle. Visualizing digital collections at archive-it. In *Proceedings of the 12th ACM/IEEE-CS Joint Conference on Digital Libraries*, JCDL '12, page 15–18, New York, NY, USA, 2012. Association for Computing Machinery.
- [Poh13] Margit Pohl. Psychologische Grundlagen der Visualisierung. University Lecture - https://vowi.fsinf.at/images/2/24/TU_Wien-Psychologische_Grundlagen_der_Visualisierung_VO_%28Pohl%29_-_Foliensatz_SS18.pdf, 2013.
- [Pri16] Prints and printmaking Australia Asia Pacific. Explore · australian prints + printmaking. <http://printsandprintmaking.gov.au/explore/decade-summary/>, December 2016.
- [PWM⁺05] Manjula Patel, Martin White, Nicholaos Mourkoussis, Krzysztof Walczak, Rafal Wojciechowski, and Jacek Chmielewski. Metadata requirements for digital museum environments. *International Journal on Digital Libraries*, 5(3):179–192, 2005.
- [Rot17] Robert E. Roth. *Visual Variables*, pages 1–11. American Cancer Society, 2017.
- [SAA⁺01] Guus Schreiber, Hans Akkermans, Anjo Anjewierden, Robert Hoog, Nigel Shadbolt, Walter Velde, and Bob Wielinga. *Knowledge Engineering and Management - The CommonKADS Methodology*, volume 24. 01 2001.

- [San16] San Francisco Museum of Modern Art. Sfmoma | explore modern art | sfmoma artscope. <https://www.sfmoma.org/experience/artscope/>, December 2016.
- [SGS⁺19] K. B. Schloss, C. C. Gramazio, A. T. Silverman, M. L. Parker, and A. S. Wang. Mapping color to meaning in colormap data visualizations. *IEEE Transactions on Visualization and Computer Graphics*, 25(1):810–819, 2019.
- [Spe01] R. Spence. *Information Visualization*. ACM Press Books. Addison-Wesley, 2001.
- [Sta20a] State Library of Queensland. Corley explorer. <https://explorer.corley.slq.qld.gov.au/>, April 2020.
- [Sta20b] State Library of Queensland. Discover the queenslander. <https://www.slq.qld.gov.au/discover/exhibitions/past-exhibitions/discover-queenslander>, March 2020.
- [SZ00] J. Stasko and E. Zhang. Focus+context display and navigation techniques for enhancing radial, space-filling hierarchy visualizations. In *IEEE Symposium on Information Visualization 2000. INFOVIS 2000. Proceedings*, pages 57–65, 2000.
- [Szo82] K. Szoka. A guide to choosing the right chart type. *IEEE Transactions on Professional Communication*, PC-25(2):98–101, 1982.
- [TC05] James J Thomas and Kristin A Cook. *Illuminating the path: The research and development agenda for visual analytics*. 2005.
- [THC12] Alice Thudt, Uta Hinrichs, and Sheelagh Carpendale. The bohemian bookshelf: Supporting serendipitous book discoveries through information visualization. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI '12*, pages 1461–1470, New York, NY, USA, 2012. ACM.
- [The16] The Metropolitan Museum of Art. 82nd & fifth | the metropolitan museum of art. <http://82nd-and-fifth.metmuseum.org/>, December 2016.
- [TSA14] J. Talbot, V. Setlur, and A. Anand. Four experiments on the perception of bar charts. *IEEE Transactions on Visualization and Computer Graphics*, 20(12):2152–2160, 2014.
- [TSF08] Melanie Tory and Sheryl Staub-French. Qualitative analysis of visualization: A building design field study. In *Proceedings of the 2008 Workshop on BEyond Time and Errors: Novel Evaluation Methods for Information Visualization*, BELIV '08, New York, NY, USA, 2008. Association for Computing Machinery.

- [Tyn16] Tyne and Wear Archives and Museums. Tyne & wear archives & museums collections dive. <http://www.collectionsdivetwmuseums.org.uk/>, December 2016.
- [UHM16] Stefania Forlini Uta Hinrichs and Bridget Moynihan. Speculative practices: Utilizing infovis to explore untapped literary collections. *IEEE Transactions on Visualization and Computer Graphics (Proceedings Visualization / Information Visualization, Oct. 2015)*, 22(1):429–438, 2016.
- [UNE16] UNESCO. Definition of the cultural heritage | united nations educational, scientific and cultural organization. <http://www.unesco.org/new/en/culture/themes/illicit-trafficking-of-cultural-property/unesco-database-of-national-cultural-heritage-laws/frequently-asked-questions/definition-of-the-cultural-heritage/>, December 2016.
- [Uta20] Uta Hinrichs, Stefania Forlini, Bridget Moynihan . The speculative w@anderverse. <http://stuffofsciencefiction.ca/vis/Wonderverse/>, March 2020.
- [VD04] F.B. Viegas and J. Donath. Social network visualization: Can we go beyond the graph. In *Workshop on Social Networks, CSCW*, volume 4, pages 6–10, 2004.
- [Vla20] Vladimir Agafonkin. Leaflet.markercluster. <https://github.com/Leaflet/Leaflet.markercluster>, July 2020.
- [WBWK00] Michelle Q. Wang Baldonado, Allison Woodruff, and Allan Kuchinsky. Guidelines for using multiple views in information visualization. In *Proceedings of the Working Conference on Advanced Visual Interfaces, AVI '00*, page 110–119, New York, NY, USA, 2000. Association for Computing Machinery.
- [WDG⁺20] M. Waldner, A. Diehl, D. Gračanin, R. Splechtna, C. Delrieux, and K. Matković. A comparison of radial and linear charts for visualizing daily patterns. *IEEE Transactions on Visualization and Computer Graphics*, 26(1):1033–1042, 2020.
- [web17] OECD - Harmonised unemployment rate (HUR). <https://doi.org/10.1787/52570002-en>, November 2017.
- [WFM⁺16] Florian Windhager, Paolo Federico, Eva Mayr, Günther Schreder, and Michael Smuc. A review of information visualization approaches and interfaces to digital cultural heritage collections. In *9th Forum Media Technology 2016*, volume 1734, pages 74–81. CEUR-WS, 2016.

- [WFS⁺19] F. Windhager, P. Federico, G. Schreder, K. Glinka, M. Dörk, S. Miksch, and E. Mayr. Visualization of cultural heritage collection data: State of the art and future challenges. *IEEE Transactions on Visualization and Computer Graphics*, 25(6):2311–2330, June 2019.
- [WH15] David Walsh and Mark M. Hall. Just looking around: Supporting casual users initial encounters with digital cultural heritage. In *SCST@ECIR*, 2015.
- [Whi12] Mitchel Whitelaw. Manly images. <http://mtchl.net/manly-images/>, October 2012.
- [Whi15] Mitchell Whitelaw. Generous interfaces for digital cultural collections. *Digital Humanities Quarterly*, 9(1), 2015.
- [Wik16] Wikimedia Foundation Inc. Universitätsbibliothek leiden – wikipedia. https://de.wikipedia.org/wiki/Universit%C3%A4tsbibliothek_Leiden#/media/File:Leiden_1610.jpg, December 2016.
- [WT92] Mats Widgren and Clas Tollin. Historical map overlays - a method for the analysis and planning of the agrarian landscape. In *Papers from the 1990 meeting of the Standing European Conference for the Study of the Rural Landscape*, 01 1992.
- [WXQW14] P. Wei, T. Xu, X. Qin, and C. Wang. Visualization of police intelligence data based on word clouds. In *2014 Tenth International Conference on Computational Intelligence and Security*, pages 539–543, 2014.
- [Ös20] Österreichische Nationalbibliothek, Bildarchiv. Bildarchiv Austria: Digitale Sammlungen > Wien > Bezirke. http://www.bildarchivaustria.at/Pages/Themen.aspx?p_iCollectionID=9846792&p_iKlassifikationID=12257731, June 2020.



Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar.
The approved original version of this thesis is available in print at TU Wien Bibliothek.

Appendices



Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar.
The approved original version of this thesis is available in print at TU Wien Bibliothek.

Interview Notes

A.1 Interviewee 1

A.1.1 Introduction

The interviewee is a scientific assistant in the area of psychology. He has no explicit previous experience with cultural heritage data collections. He is used to use simple data visualizations (graphs) in his scientific work but besides that he has expert knowledge about data visualizations. He considers himself interested in history but not really interested in art.

A.1.2 Interview Tasks and Questions

A.1.3 Search-Based Interface:

- **Please take a moment to browse around the collection**
The interviewee was browsing through the collection exploring different districts and time periods.
- **What is your first impression of the interface? How would you assess the general usability of the interface?**
The interviewee stated that the interface does not look especially inviting but the structure is logical and easy to understand. He was confused that the interface tends to open a lot of new browser tabs and that it took quite a lot of clicks to reach the detail page of an image.
- **Find a picture of Schönbrunn**
The interviewee started by filtering by district but was unsure which district is the correct one. Because of that he switched to the search function and could immediately find a correct picture through it.

- **Find an old picture (created before 1750) of Schönbrunn**
The interviewee immediately used the expert search ("Profisuche") by typing in the date and the keyword and could successfully retrieve a correct image.
- **Find a picture of a church close to your home**
The interviewee recalled a name of a big church close to his home and searched for it using the search function. He could successfully retrieve a correct image.
- **Which decade (10 years period) has the most pictures in the collection?**
The interviewee started by consulting the expert search and tried to type in different time periods but was unable to count the number of images per decade. He then clicked through most of the districts estimating for each which decade had the most images in it. Conveniently the interface shows the number of images contained in each time period which allowed the interviewee a correct estimation.

A.1.4 Visual Prototype:

The following tasks and questions refer to the new *visual prototype [P-INTF]* (<https://historymap.wien>).

- **Please take a moment to browse around the collection**
The interviewee started by exploring the map and clicking on various images on the map. Shortly after he explored different areas of the interface, first the list view including the search function and then the map sidebar and its visualizations which he tried out as well.
- **What is your first impression of the interface? How would you assess the general usability of the interface?**
The interviewee states that the interface seems very inviting and interesting to look at. It's easy to use and understandable.
- **Find a picture of Schönbrunn**
The interviewee changed to the list view and quickly found a correct picture using the search function.
- **Find an old picture (created before 1750) of Schönbrunn**
The interviewee followed the same approach as in the previous task. Using the search function he clicked through the result list and relatively quickly found a correct picture.
- **Find a picture of a church close to your home**
The interviewee utilized the search function again (searching for the name of the church he knows is closed to his home) and immediately found a correct picture. After being asked if he knows another solution he clicked on the keyword "church" in the keyword cloud visualization and found a correct picture on the map.

- **Which decade (10 years period) has the most pictures in the collection?**
The interviewee immediately looked for the longest bar in the temporal bar chart and gave the correct answer.
- **Which author created most pictures in the collection?**
The interviewee immediately opened the author bubble chart but needed a moment for giving the correct answer as there was still another filter active which he was not aware of. After realizing and clearing the filters, he gave the correct answer.
- **Which author created most pictures dated 1800 or earlier?**
Additionally to using the bubble chart, the interviewee immediately filtered the temporal bar chart and gave the correct answer.
- **Find 2 authors which were working often with each other and open a picture created by them.**
The interviewee opened the author network visualization and quickly was able to solve the task. (note: during the first task the interviewee already experimented with the author network visualization and therefore most likely was already a bit familiar with its functionality)
- **List a few keywords used often throughout the collection**
The interviewee opened the keyword cloud visualization and correctly listed a few keywords used often.
- **Find a few pictures of buildings standing in the place of the old vienna city wall.**
The interviewee used his history knowledge about vienna and used the search function to find a picture of the viennese parliament, which is partly correct because it is right next to the old city wall but not directly on the same location.

A.1.5 Questionnaire: Evaluation of Visualizations

A.1.6 Final Questions and Conclusion

The interviewee prefers very much the new *visual prototype [P-INTF]*. He describes the *search-based interface [S-INTF]* as a bit boring and less user friendly. But he thinks that the expert search of the *search-based interface [S-INTF]* is very sophisticated and useful, especially for very specific queries. He states that the new interface is inviting and arouses interest for the pictures. He was motivated to browse through the collection and because of the map he instantly had some kind of personal connection because of places he knows, where he lives or where he works. Through the visualizations aspects of the collection (e.g. relationship of authors with each other - author network) became suddenly interesting for him. The interviewee sees the expert search as something that possibly should be added to the new interface. The existing search function in the new interface is easy to use but not as flexible as the expert search in the *search-based interface [S-INTF]*. Also embedding the search function directly on the map could be helpful.

Note: asked about the meaning of the colors of the map icons and the temporal bar chart, the interviewee answered that he didn't really pay any attention to it but he understand its meaning.

A.2 Interviewee 2

A.2.1 Introduction

The interviewee is currently in education of becoming a software engineer. He has no explicit previous experience with cultural heritage data collections nor with data visualizations. He considers himself interested in both history and arts. In the past he did some part time jobs for museums as security but was not involved in the curation of any collection or exhibition.

A.2.2 Interview Tasks and Questions

A.2.3 Search-Based Interface:

- **Please take a moment to browse around the collection**
The interviewee randomly browsed through the collection exploring districts and time periods.
- **What is your first impression of the interface? How would you assess the general usability of the interface?**
The interviewee states that the interface is very standard and how he would expect an image database to look like. He describes the usability as neither especially good nor especially bad.
- **Find a picture of Schönbrunn**
The interviewee clicked on the correct district and after some browsing through the different time periods was able to find a correct picture.
- **Find an old picture (created before 1750) of Schönbrunn**
The interviewee tried the same approach as for the previous question: he clicked on the district and randomly searched for a correct picture. After a few failed attempts he switched over to the simple search function ("Schnellsuche"), typed in the keyword "Schönbrunn" and browsed through the result list until he found a correct picture.
- **Find a picture of a church close to your home**
The interviewee clicked on his home district and browsed through the pictures until he found a picture of a church that he recognizes.
- **Which decade (10 years period) has the most pictures in the collection?**
The interviewee didn't really know how to complete the task and concluded he

would need to calculate quite a bit to reach the correct result. He clicked through several districts and estimated that the time period between 1900 and 1910 contains a lot of pictures for most districts, which is in fact the correct answer.

A.2.4 Visual Prototype:

The following tasks and questions refer to the new *visual prototype [P-INTF]* (<https://historymap.wien>).

- **Please take a moment to browse around the collection**
The interviewee focused mainly on the map during this task. He was exploring different areas of Vienna on the map and clicked on several images to look at them and read their description. He also inspected the temporal bar chart whose usage was not immediately clear. He read the info text and was then able to successfully filter the temporal bar chart.
- **What is your first impression of the interface? How would you assess the general usability of the interface?**
The interviewee likes the interactive interface and thinks it's fun to use. The structure is clear and understandable. Note: by that point the interviewee did neither use the list function nor did he access any of the other visualizations via the sidepanel
- **Find a picture of Schönbrunn**
The interviewee used the map to locate the general area of the Schönbrunn castle and could immediately open a correct picture on the map.
- **Find an old picture (created before 1750) of Schönbrunn**
The interviewee used the color correlation between the temporal bar chart and the map icon borders to find an old picture amongst the pictures of Schönbrunn castle which he could successfully complete within seconds. Being asked if he knows a different way to do it, he applied filtering on the temporal bar chart and showed me that only the correct pictures are shown on the map.
- **Find a picture of a church close to your home**
The interviewee used only the map to search for a church close to his home. It took him a moment of browsing through various pictures but he was able to find one rather quickly. I asked him if there is a different way and after further exploring he managed to click on the correct keyword in the keyword cloud effectively filtering the map to show only the churches.
- **Which decade (10 years period) has the most pictures in the collection?**
The interviewee immediately could give the correct answer by checking the temporal bar chart.

- **Which author created most pictures in the collection?**
By then the interviewee was aware of the additional visualizations on the sidebar. He opened the author bubble chart and could immediately give the correct answer.
- **Which author created most pictures dated 1800 or earlier?**
The interviewee tried to click on various authors in the bubble chart. He concluded that this would be a complicated (but possible) way to find the correct answer. He then used the temporal bar chart to filter the collection by time period and then gave the correct answer by looking again at the author bubble chart.
- **Find 2 authors which were working often with each other and open a picture created by them.**
The interviewee immediately understood that the author network could give him this information, but it took him a while to understand how the visualization works. In the beginning he was not aware that by clicking on an author, unrelated authors get disabled and related authors stay active and should be clicked in a second step to show related pictures. After experimenting further with the visualization he was eventually able to give the correct answer.
- **List a few keywords used often throughout the collection**
The interviewee immediately opened the keyword cloud visualization and gave the correct answer.
- **Find a few pictures of buildings standing in the place of the old vienna city wall.**
The interviewee selected an appropriate map overlay and searched on the map for images located on the city wall. He additionally applied a filter on the temporal chart to only show recent buildings.

A.2.5 Questionnaire: Evaluation of Visualizations

A.2.6 Final Questions and Conclusion

The interviewee thinks that the *search-based interface [S-INTF]* is a little bit unclear and not interactive which makes the interface a bit boring. The expert search in the *search-based interface [S-INTF]* is very good for advanced searches and is definitely an advantage when compared to the new interface. He thinks that the new *visual prototype [P-INTF]* is fun, interesting and appealing. For general queries and just casual browsing the net interface is very suitable and he thinks it would also work very well as mobile application. The biggest disadvantage of the new interface is that the sidepanel is not really self-explanatory and can easily be missed. Also the application should be available in german. Another possible improvement would be to include a search function (keyword, expert) with the map.

Note: The interviewee had problems with the side bar. He did not understand that additional visualizations and features can be accessed via the "options" icon. He proposes

to use a different icon or opening the side panel by default.

Note: It was not immediately understandable for the interviewee how the clearing of the filters works.

A.3 Interviewee 3

A.3.1 Introduction

The interviewee is a professional software engineer working in the insurance industry. He has no explicit previous experience with cultural heritage data collections nor with data visualizations. He considers himself interested in both history and arts.

A.3.2 Interview Tasks and Questions

A.3.3 Search-Based Interface:

- **Please take a moment to browse around the collection**
The interviewee clicked on several districts and time periods and opened a few images. He was a bit confused about the many tabs which are opened when clicking various links on the interface.
- **What is your first impression of the interface? How would you assess the general usability of the interface?**
The interviewee describes the interface as outdated, not responsive and a bit hard to use as he is used to use "modern software" and the interface does not behave like he is used to nowadays (e.g. by opening so many tabs instead of a single page application). The navigational flow of the interface is not very intuitive.
- **Find a picture of Schönbrunn**
The interviewee clicked on the correct district and then just browsed the resulting images. He could find a correct image very fast.
- **Find an old picture (created before 1750) of Schönbrunn**
Like in the previous task, the interviewee clicked on the correct district and additionally accessed one of the matching time periods. He could find a correct image within seconds.
- **Find a picture of a church close to your home**
The interviewee tried to select the correct district and randomly browsed through the collection but could not find a matching image. He then tried to use the quick search with multiple search terms which failed. He then gave up as he considers the task effectively unsolvable through the given interface.
- **Which decade (10 years period) has the most pictures in the collection?**
The interviewee clicked on the first district (Innere Stadt) as he identified this

district as the one with most pictures. He then inspected the numbers of the different timeframes (5 years each) and estimated that the period 1890 to 1899.

A.3.4 Visual Prototype:

The following tasks and questions refer to the new *visual prototype [P-INTF]* (<https://historymap.wien>).

- **Please take a moment to browse around the collection**
The interviewee inspected the map, the list and was able to open the sidebar immediately. He opened each visualization trying to understand its meaning. He also read the information text displayed via the help-icon.
- **What is your first impression of the interface? How would you assess the general usability of the interface?**
The interviewee states that the interface is clean and modern and that he knows what the Layout-Components are used for. In his opinion the interface feels much more modern than the old one.
- **Find a picture of Schönbrunn**
The interviewee clicked on the tab "List" and used the search to filter for images matchin the search term "schönbrunn". He could find an image immediately.
- **Find an old picture (created before 1750) of Schönbrunn**
The interviewee started by filtering the Temporal Bar Chart to the appropriate timeframe and then switched to the List. He was confused that filtering the temporal bar chart on the map does not affect the results shown in the artifact list. He then proceeded to search again on the artifact list and sorted the resulting list by their creation year. He browsed through the pages of the result list until he found a correct image.
- **Find a picture of a church close to your home**
The interviewee used the search of the artifact list to search for the name of a church close to his home. He could find an image immediately.
- **Which decade (10 years period) has the most pictures in the collection?**
The interviewee immediately changed to the map and looked at the temporal bar chart. He had some difficulties selecting the correct bar but succeeded and gave the correct answer eventually.
- **Which author created most pictures in the collection?**
The interviewee opened the author bubble chart and was able to give the correct answer immediately
- **Which author created most pictures dated 1800 or earlier?**
The interviewee filtered the temporal bar chart and then looked at the bubble chart. He was confused as the biggest bubble shown was "others" which combines

various authors with smaller contributions. He eventually gave the correct answer (the second biggest bubble representing only a single author). He did not read the information text above the visualization which would explain the meaning of "others".

- **Find 2 authors which were working often with each other and open a picture created by them.**

The interviewee immediately opened the author network chart but was unsure on how to use it. He tried to click on the line connecting two nodes which is currently not supported by the chart. After some trial and error he clicked on the two nodes and opened a corresponding image on the map.

- **List a few keywords used often throughout the collection**

The interviewee immediately opened the keyword cloud visualization and gave the correct answer.

- **Find a few pictures of buildings standing in the place of the old vienna city wall.**

The interviewee browsed through some images on the map along the Ring-Street as he knew that the street was built in place of the old city wall. As he was not able to give a clear answer, he figured the map overlays could help him. After turning the map overlay on he was able to give the correct answer.

A.3.5 Questionnaire: Evaluation of Visualizations

A.3.6 Final Questions and Conclusion

The interviewee prefers the new interface as it feels modern and more visually pleasing. He is missing a search function directly on the map and would like to have a simpler way to filter the images by district. He thinks the temporal bar chart is especially useful but is a bit hard to use as it is not clickable and it's immediately visible which period is selected. In general it should be a bit easier to see which filters are active and easier to delete a certain filter. The interviewee would prefer a more traditional map style (e.g. like Google Maps). He thinks the map style fits the design but makes it a bit hard to use. He considers the indicator for the map-cluster-boundaries confusing.

A.4 Interviewee 4

A.4.1 Introduction

The interviewee is a professional software engineer working in the telecommunications industry. He studied media informatics at the technical university of vienna and therefore has moderate experience with data visualizations. He considers himself not especially interested in history nor arts.

A.4.2 Interview Tasks and Questions

A.4.3 Search-Based Interface:

- **Please take a moment to browse around the collection**
The interviewee clicked on several districts and time periods and opened a few images without problems.
- **What is your first impression of the interface? How would you assess the general usability of the interface?**
The interviewee states that the interface feels old and is rather slow. He is missing filter options when browsing the collection. From a UX-perspective he thinks the order of the buttons of the search interface is incorrect.
- **Find a picture of Schönbrunn**
The interviewee opened the expert search and selected searched via district and search term "schönbrunn". He was able to find a correct image right away.
- **Find an old picture (created before 1750) of Schönbrunn**
The interviewee tried again the expert search and tried to enter the term "schönbrunn" in the "Location" search field. This attempt failed hence he tried again using the general search term and the time filter which allowed him to retrieve a correct image.
- **Find a picture of a church close to your home**
The interviewee again consulted the expert search trying to search for Location "Leopoldstadt" (his home district) and search term "Kirche" which did not result in any matches. He then removed the Location and tried again. He browsed through the result list until he found a matching image depicting a church close to his home.
- **Which decade (10 years period) has the most pictures in the collection?**
The interviewee clicked on the parent element "Bezirke" and looked at various time periods concluding that most likely the decade 1900 - 1909 has the most pictures in the collection.

A.4.4 Visual Prototype:

The following tasks and questions refer to the new *visual prototype [P-INTF]* (<https://historymap.wien>).

- **Please take a moment to browse around the collection**
The interviewee inspected the map and the list and was also able to open the sidebar inspecting the various visualizations. He spent some time browsing the map and looking at various areas and pictures around the city.
- **What is your first impression of the interface? How would you assess the general usability of the interface?**

The interviewee states that the interface is modern and very visual which he likes. Although he was able to navigate the interface without problems, he is missing a filter option to filter by district.

- **Find a picture of Schönbrunn**

The interviewee opened the artifact list and was immediately able to find a correct image using the search term "schönbrunn".

- **Find an old picture (created before 1750) of Schönbrunn**

Using the result list of the previous task, the interviewee browsed the results until he found a correct one, without applying any additional filters or executing a new search.

- **Find a picture of a church close to your home**

The interviewee tried again to search using the artifact list. He entered the search term "kirche" but was unable to find a church close to his home in the result list. He then proceeded to the map and navigated to a familiar church close to his home and opened an image of it.

- **Which decade (10 years period) has the most pictures in the collection?**

The interviewee immediately understood that the temporal bar chart encodes this information. But he had troubles using the chart as he tried to select the longest bar by clicking on it, which currently is not supported. Eventually he was able to select the correct bar by clicking and dragging which allowed him to give the correct answer.

- **Which author created most pictures in the collection?**

The interviewee opened the author bubble chart and was able to give the correct answer immediately

- **Which author created most pictures dated 1800 or earlier?**

The interviewee filtered the temporal bar chart and then looked again at the bubble chart. He was able to give the correct answer immediately.

- **Find 2 authors which were working often with each other and open a picture created by them.**

The interviewee opened the author network and clicked on a certain author, unaware that he also could select a second author directly on the author network chart. He then navigated to the author bubble chart and inspected how the author bubble chart changed through the applied filtering. Unable to retrieve the correct information from the bubble chart, he randomly browsed the map and clicked on a few images eventually giving a correct answer.

- **List a few keywords used often throughout the collection**

The interviewee immediately opened the keyword cloud visualization and gave the correct answer.

- **Find a few pictures of buildings standing in the place of the old vienna city wall.**

The interviewee browsed around the map and through the visualizations but was not able to make the connection between the map overlays and the information required by the task, hence was unable to complete the task.

A.4.5 Questionnaire: Evaluation of Visualizations

A.4.6 Final Questions and Conclusion

The interviewee prefers the new interface because of the spatial dimension added to the interface. He states it's fun to browse around the map. Still he has some issues with some of the visualizations. E.g. the keyword cloud is confusing and does not provide much information. He would prefer a bubble chart instead of the keyword cloud. The search function generally should be more sophisticated and there should be a search function directly on the map. He also would like to have more navigational possibilities e.g. from the detail page of an artifact it should be able to navigate to the map location.

A.5 Interviewee 5

A.5.1 Introduction

The interviewee is an artist currently working on a film project. During her studies and her work she made various experiences with digital cultural heritage collections. Her experience with data visualizations is limited. She would consider herself both interested in history and arts.

A.5.2 Interview Tasks and Questions

A.5.3 Search-Based Interface:

- **Please take a moment to browse around the collection**
The interviewee explored the collection via the district overview and tried out both quick search and expert search. She tried to find images related to specific topics that came to her mind.
- **What is your first impression of the interface? How would you assess the general usability of the interface?**
The interviewee got bored very fast because the search function is very slow (during this interview it was in fact much slower than normally). She said under normal circumstances she would probably not spend much time using the interface but quit very soon. She would like more textual descriptions on how the collection is organized (e.g. that the artifacts are organized temporally and sorted ascending by their creation year). When clicking on "districts" it's unclear for her if the pictures are shown are from all districts combined or not.

- **Find a picture of Schönbrunn**
The interviewee tried to find the picture via the quick search entering the search term "schönbrunn", but gave up because the search was too slow (slower than usual). She then selected the correct district and browsed through the result list until she found one.
- **Find an old picture (created before 1750) of Schönbrunn**
The interviewee opened again the correct district and then opened a few matching time periods until she found again a picture of schönbrunn.
- **Find a picture of a church close to your home**
Again the interviewee opened the corresponding district and browsed through the images until she found a matching one, which was easy because she is living close to a famous, old church of which many images exist (Karlskirche).
- **Which decade (10 years period) has the most pictures in the collection?**
The interviewee concluded immediately that answering this question would be very time-consuming. She opened again the districts-overview and estimated that most pictures exist between 1910 and 1920.

A.5.4 Visual Prototype:

The following tasks and questions refer to the new *visual prototype [P-INTF]* (<https://historymap.wien>).

- **Please take a moment to browse around the collection**
The interviewee started by using the temporal bar chart. She immediately understood that she can filter the images by dragging and filtering on the temporal bar chart. She had troubles selecting a single decade. She browsed through the collection by exploring the map. She opened the side bar and looked at the different visualizations. She had troubles finding the search function but eventually opened the List view and saw the search function there.
- **What is your first impression of the interface? How would you assess the general usability of the interface?**
The interviewee states that the interface is very well arranged and she likes the spatial aspect in comparison with the *search-based interface [S-INTF]*. She was a bit confused by some of the wording used throughout the interface.
- **Find a picture of Schönbrunn**
The interviewee opened the artifact list and was immediately able to find a correct image using the search term "schönbrunn".
- **Find an old picture (created before 1750) of Schönbrunn**
The interviewee tried to filter the temporal bar chart on the map, switched then to the list view and was confused that the filtering of the bar chart does not affect

which images are shown in the list. She then proceeded the search in the artifact list and browse through the results until she found a matching image in the correct time period.

- **Find a picture of a church close to your home**
The interviewee opened the keyword cloud visualization and selected the keyword church ("Kirche") and selected then an image on the map close to her home.
- **Which decade (10 years period) has the most pictures in the collection?**
The interviewee used the temporal bar chart to select the longest bar and gave the correct answer.
- **Which author created most pictures in the collection?**
The interviewee opened the author bubble chart and was able to give the correct answer immediately
- **Which author created most pictures dated 1800 or earlier?**
The interviewee filtered the temporal bar chart and then looked again at the bubble chart. She was able to give the correct answer immediately.
- **Find 2 authors which were working often with each other and open a picture created by them.**
The interviewee opened the author network and played around a bit with the visualizations. She tried to click on the line between two authors to select their work relationship, which is currently not supported by the interface. After a bit of trial and error she was able to select two authors (by selecting the two nodes after one another) and opening a matching image on the map.
- **List a few keywords used often throughout the collection**
The interviewee immediately opened the keyword cloud visualization and gave the correct answer.
- **Find a few pictures of buildings standing in the place of the old vienna city wall.**
The interviewee tried to use the search function of the artifact list view. Unable to find the correct answer she looked at the different available visualizations for clues and eventually figured out that the map overlays will help her answering the question. She did not immediately understand what the term "map overlays" means and that she can show historic maps through it. Once this was clear for her, she was able to answer the questions without further problems.

A.5.5 Questionnaire: Evaluation of Visualizations

A.5.6 Final Questions and Conclusion

The interviewee in general likes the modern look of the interface and the spatial aspect. If she would search for something specific, she would probably use the *search-based interface*

[S-INTF] as the search function is more sophisticated. She was confused by some of the wording of the interface. E.g. she thinks "authors" should be replaced by "photographers and artists", "map overlays" by "historic map overlays" and "artifacts" by "images". She states that the filter overview is not visible enough. The interviewee proposes to add an additional label "Total images" to the temporal bar chart. She would also like to have a feature "show on map" on the detail page of an artifact. She was confused by the grey color of the map clusters because the description of the temporal bar chart states that the icon border color represents the time period of the artifact. In general she would prefer to have more possible ways to reach the same goals.

A.6 Interviewee 6

A.6.1 Introduction

The interviewee is a software engineer working in the telecommunications industry. He has some experience with data visualizations because of his studies (media informatics) but no experience with digital cultural heritage collections. He would consider himself not especially interested in history nor arts.

A.6.2 Interview Tasks and Questions

A.6.3 Search-Based Interface:

- **Please take a moment to browse around the collection**
The interviewee opened the districts overview and clicked through a few districts. He switched between the list and gallery view and noted that the navigation in general is a bit confusing. He opened a few pictures and also tried out specialized functions like "calculate publication price" and "download layout image". He also explored the two different search functions.
- **What is your first impression of the interface? How would you assess the general usability of the interface?**
The interviewee stated that the navigation is confusing and that the interface opens a lot of browser tabs which need to be closed manually. The quick search function is a bit too simple for him and he wishes a bit more functionality. He prefers the list view but the button to enable the list view is way too hard to find. The expert search is sophisticated but it only allows for basic usage regarding time filtering.
- **Find a picture of Schönbrunn**
The interviewee tried to find the picture via the quick search entering the search term "schönbrunn", and was able to retrieve a correct image immediately.
- **Find an old picture (created before 1750) of Schönbrunn**
The interviewee opened the expert search and tried to enter the term "schönbrunn" in the search field "Location" and the date in the appropriate date field. This didn't

work as the location field only supports few options. He then used "schönbrunn" as general search term and was able to successfully retrieve an image.

- **Find a picture of a church close to your home**

The interviewee continued to use the expert search and searched for the term "schönbrunn" and selecting the correct district as "subtopic" which worked fine and he was able to retrieve a correct image immediately.

- **Which decade (10 years period) has the most pictures in the collection?**

The interviewee opened the districts overview and browsed looked at the number of images per time period (5 years). He stated that he can only estimate that most images were created between 1960 and 1970.

A.6.4 Visual Prototype:

The following tasks and questions refer to the new *visual prototype [P-INTF]* (<https://historymap.wien>).

- **Please take a moment to browse around the collection**

The interviewee inspected the map including the temp chart which he understood immediately and tried out its filtering capabilities. He opened the list view and the sidebar and inspected all visualizations without any problems.

- **What is your first impression of the interface? How would you assess the general usability of the interface?**

The interviewee states that the interface is fun to use as it is interactive and it visualizes relationships between artifacts and allows for more insights into the data. He was confused by the creation technique chart as the chart is refreshed too often.

- **Find a picture of Schönbrunn**

The interviewee opened the artifact list and was immediately able to find a correct image using the search term "schönbrunn".

- **Find an old picture (created before 1750) of Schönbrunn**

The interviewee filtered the temporal bar chart and navigated then to the list view, realizing that the filtering was not applied to the list. He then navigated back to the map, applied the filter via the temporal bar chart again and searched for a picture of schönbrunn directly on the map. This took a while as he was unsure where exactly on the map it is located.

- **Find a picture of a church close to your home**

The interviewee opened the keyword cloud visualization and selected the keyword church ("Kirche"). This took a while as he saw the keyword before but the keyword cloud rebuilds itself when the page is reloaded resulting in new positioning of the individual words. Hence it was difficult for him to find the correct word in the cloud.

- **Which decade (10 years period) has the most pictures in the collection?**
The interviewee used the temporal bar chart to look for the longest bar. Initially he gave the wrong answer as he did not realize that there was still another filter active. After realizing, he was able to give the correct answer.
- **Which author created most pictures in the collection?**
The interviewee opened the author network chart but quickly realized that he will not find the correct information there. He switched to the Author Bubble chart and could quickly give the correct answer.
- **Which author created most pictures dated 1800 or earlier?**
The interviewee filtered the temporal bar chart and then looked again at the bubble chart and was able to give the correct answer.
- **Find 2 authors which were working often with each other and open a picture created by them.**
The interviewee opened the author network, selected two authors and opened an image on the map. Note: He used the network chart before and hence was already familiar with it.
- **List a few keywords used often throughout the collection**
The interviewee immediately opened the keyword cloud visualization and gave the correct answer.
- **Find a few pictures of buildings standing in the place of the old vienna city wall.**
The interviewee first tried to find an appropriate keyword in the keyword cloud. After that he tried a few map overlays and looked for correct pictures on the map. He could complete the task quickly.

A.6.5 Questionnaire: Evaluation of Visualizations

A.6.6 Final Questions and Conclusion

The interviewee prefers the new interface to the old one as it is fun to use and invite for exploration. He would like to have a more sophisticated search function and the search function should affect both map and list view. In general the concept of map vs list is not really clear for him. He really likes the concept of the word cloud and proposes to use it as well for artifact titles (e.g. "Title Cloud") as he thinks it could give some interesting insights. Furthermore he thinks it would be useful to add a search function to each visualization as it already exists for the author bubble chart.

A.7 Interviewee 7

A.7.1 Introduction

The interviewee is a software engineer working in the telecommunications industry. He has little experience with digital cultural heritage data collection but visited a few digital collections of museums. He has no special knowledge about data visualizations (except basic graphs). He is very interested in history and also in arts.

A.7.2 Interview Tasks and Questions

A.7.3 Search-Based Interface:

- **Please take a moment to browse around the collection**
The interviewee explored the districts overview and opened several districts, viewing random images of random time periods. He inspected the quick search function and did a sample search.
- **What is your first impression of the interface? How would you assess the general usability of the interface?**
The interviewee describes the interface as "functional". The interface feels old but he is confident he would be able to find what he is looking for if he is searching for something specific. He is confused by the fact that the interface opens so many tabs which need to be manually closed.
- **Find a picture of Schönbrunn**
The interviewee searched via the quick search for the term "schönbrunn" and found a correct image.
- **Find an old picture (created before 1750) of Schönbrunn**
The interviewee used the expert search to search for an image "before 1750" and containing the term "schönbrunn". He found a correct image immediately.
- **Find a picture of a church close to your home**
Again the interviewee used the expert search by searching for the term "church" ("Kirche") and selecting his home district as sub-topic. He successfully could find an image of a familiar church close to his home.
- **Which decade (10 years period) has the most pictures in the collection?**
The interviewee tried to find an appropriate function in the quick and expert search but quickly realized it is missing. He stated the task is not possible and didn't try further.

A.7.4 Visual Prototype:

The following tasks and questions refer to the new *visual prototype [P-INTF]* (<https://historymap.wien>).

- **Please take a moment to browse around the collection**
 The interviewee browsed through the collection using the map, the list and the side bar. He tried various filters and inspected the results on the map.
- **What is your first impression of the interface? How would you assess the general usability of the interface?**
 The interviewee describes the interface as "inviting" and "playful". He likes the interface but he thinks it probably depends on the target group if the interface is suitable or not (e.g. experts might prefer the traditional interface). The visualizations are interesting but the user has to explore the visualizations first before being efficient when using them.
- **Find a picture of Schönbrunn**
 The interviewee used the search function above the artifact list and was immediately able to find a correct image using the search term "schönbrunn".
- **Find an old picture (created before 1750) of Schönbrunn**
 The interviewee tried to enter a complex search term into the search field above the artifact list. He tried the term "Schönbrunn 1750" which did not succeed as this kind of search is not yet supported. He then proceeded to the map, manually searched for the location on the map and then navigated through all the pictures until he found a correct one.
- **Find a picture of a church close to your home**
 The interviewee tried to search for his district in the artifact list which was unsuccessful. He then opened the map and manually navigated to his district and opened a picture of a church he knew the location of.
- **Which decade (10 years period) has the most pictures in the collection?**
 The interviewee was not really aware of the functionality of the temporal bar chart. After not finding another appropriate visualization or feature, he realized that the bar chart can in fact provide this information. After that he could give the correct answer immediately.
- **Which author created most pictures in the collection?**
 The interviewee opened the author bubble chart and could give the correct answer immediately.
- **Which author created most pictures dated 1800 or earlier?**
 The interviewee tried to use the map overlays to change the time period (as every map overlay contains the creation date as well) which was unsuccessful. After some time he realized that the temporal bar chart can be filtered. He had some troubles with the filter functionality but eventually managed to apply the filter and refer back to the author bubble chart, giving the correct answer.

- **Find 2 authors which were working often with each other and open a picture created by them.**

The interviewee firstly opened the author bubble chart but quickly realized it will not help him. He then opened the author network chart randomly selecting two authors. He was not immediately aware that the thickness of a link between two authors represents the intensity of the work relationship nor did he read the info text. After some trial and error he figured the meaning of the thickness out himself and then could open a correct image hence completing the task.

- **List a few keywords used often throughout the collection**

The interviewee immediately opened the keyword cloud visualization and gave the correct answer.

- **Find a few pictures of buildings standing in the place of the old vienna city wall.**

The interviewee tried to search for the term "city wall" ("Stadtmauer") using the search function of the artifact list. Being unsuccessful he changed to the map view and applied a filter via the temporal bar chart. After more trial and error and more thinking he finally realized about the map overlays (which he did not use until that moment) and applied the map overlay and a temporal filtering at the same time. He then was able to complete the task successfully by showing matching images.

A.7.5 Questionnaire: Evaluation of Visualizations

A.7.6 Final Questions and Conclusion

The interviewee prefers the new interface but clearly states that he thinks certain people (e.g. historians) probably will prefer the existing interface. He is missing a search function directly on the map and generally thinks the search function could be more sophisticated (e.g. directly combining several search variables). It should be clearer in the temporal bar chart which bar refers to which decade and which timeframe is currently selected. He thinks a "time travel" feature would be very interesting which allows to scroll through time and both images and map overlays change on the fly.

A.8 Interviewee 8

A.8.1 Introduction

The interviewee is a researcher working in the field of evolutionary biology. She has a strong background in mathematics and data science and is familiar with data visualizations, especially in a scientific setting. She has no special experience with digital cultural heritage collections but would consider herself interested in both history and arts.

A.8.2 Interview Tasks and Questions

A.8.3 Search-Based Interface:

- Please take a moment to browse around the collection**
The interviewee started exploring her home district, opening various images and trying to figure out where they are located.
- What is your first impression of the interface? How would you assess the general usability of the interface?**
The interviewee's first reaction was, that the collection is very interesting. She describes the interface as plain but structured. She criticises the limited classification of the images as the collection is only divided into districts and then time periods of 5 years. She would like to have more filter options to control the result list.
- Find a picture of Schönbrunn**
The interviewee clicked on the district Hietzing and then randomly browsed the result list until she found a correct image.
- Find an old picture (created before 1750) of Schönbrunn**
As before, she clicked on the district and then opened a time period before 1750 and browsed the result list. She could find a correct image relatively fast.
- Find a picture of a church close to your home**
The interviewee opened the district overview and opened the folder of her home district. She then proceeded to click on the time period with the most pictures and randomly browsed through the result list until she could visually make out a church. She opened the image and verified that the church is in fact close to her home.
- Which decade (10 years period) has the most pictures in the collection?**
The interviewee clicked through a few districts and looked at the number of images per 5-years time-period. She estimated that most images were created between 1900 and 1910.

A.8.4 Visual Prototype:

The following tasks and questions refer to the new *visual prototype [P-INTF]* (<https://historymap.wien>).

- Please take a moment to browse around the collection**
Again the interviewee started to explore her home district, this time navigating on the map. She opened the sidebar and explored various visualizations. She obviously enjoyed navigating around and looking at pictures. She also experimented with the filtering of the temporal bar chart and the map overlays.

- **What is your first impression of the interface? How would you assess the general usability of the interface?**

The interviewee enjoyed the new modern look of the interface. She describes the user experience as neat and smooth. She thinks the interface is self-explanatory.

- **Find a picture of Schönbrunn**

The interviewee immediately navigated to Schönbrunn on the map and opened a correct picture.

- **Find an old picture (created before 1750) of Schönbrunn**

The interviewee opened the list but had troubles finding the search field. After some looking around she was able to find it and used it to search for the term "schönbrunn". She then proceeded to sort the result list by year and opened a correct image.

- **Find a picture of a church close to your home**

The interviewee tried to search for the term "kirche wien 22" but ended up with no search results. She then searched only for the term "Kirche" and sorted the result list by district and browsed through the result list until she found a correct image.

- **Which decade (10 years period) has the most pictures in the collection?**

The interviewee consulted the temporal bar chart and could give the correct answer immediately.

- **Which author created most pictures in the collection?**

The interviewee opened the author bubble chart and could give the correct answer immediately.

- **Which author created most pictures dated 1800 or earlier?**

The interviewee experimented with the temporal bar chart earlier and hence was already familiar with the functionality. She filtered the temporal bar chart and then looked at the author bubble chart and was able to give the correct answer immediately.

- **Find 2 authors which were working often with each other and open a picture created by them.**

The interviewee opened the author network chart, clicked on one author (connected by a thick line) and then immediately on the second author. She opened a correct image on the map.

- **List a few keywords used often throughout the collection**

The interviewee immediately opened the keyword cloud visualization and gave the correct answer.

- **Find a few pictures of buildings standing in the place of the old vienna city wall.**

The interviewee was familiar already with the map overlay function and opened it

now immediately. She additionally applied a filtering on the temporal bar chart and then opened a correct image.

A.8.5 Questionnaire: Evaluation of Visualizations

A.8.6 Final Questions and Conclusion

The interviewee prefers the new interface and states she hopes it will be developed further and publicly available soon. She thinks the temporal bar chart should be labeled better and stated that the current labelling is actually slightly wrong (as the label is in the middle of the bar and not the beginning). She also proposes to add a shorter, more concise help text which is shown as soon as one hovers over a visualization (additionally or instead of the existing help text). In general she proposes the work more on the wording to be more precise and if possible shorter.

Note: The interviewee had troubles finding the search function. Note: The interviewee was much better than most when using new visualizations. When asked about this, she confirmed once again, that she is familiar with those kind of data visualizations and hence knows how to operate them.

A.9 Interviewee 9

A.9.1 Introduction

The interviewee is a mathematician working in the insurance industry. He is used to read data visualizations but has no limited experience with digital cultural heritage collections, e.g. he remembered exploring a digital collection of historic telephone books. He is interested in arts and history.

A.9.2 Interview Tasks and Questions

A.9.3 Search-Based Interface:

- **Please take a moment to browse around the collection**
The interviewee immediately started to explore his home district, looking at various pictures and guessing where they could be located. He seemed very interested in the general topic and the collection. The explored the districts overview and also clicked on a few other districts.
- **What is your first impression of the interface? How would you assess the general usability of the interface?**
The interviewee states that he thinks the interface and the collection is very interested. Still he thinks he might loose interest after some time because of the structure of the collection. He mentions he is a fan of historic maps, hence he is a bit disappointed that it is not included in the collection. He also states that the

navigational concept confuses him as the interface opens new browser tabs all the time which need to be closed manually.

- **Find a picture of Schönbrunn**
The interviewee opened the quick search, entered "schönbrunn" as a search term and could find a correct image immediately.
- **Find an old picture (created before 1750) of Schönbrunn**
The interviewee opened the expert search and entered the search term "schönbrunn" again. Additionally he added the year in the "before year" search field. He could immediately find a correct image.
- **Find a picture of a church close to your home**
The interviewee opened the quick search and searched for the term "kirche donaus-tadt" which again, was successful.
- **Which decade (10 years period) has the most pictures in the collection?**
The interviewee opened the expert search but quickly realized it will not help him solve this task. He then navigated to the districts overview, looked at the folders of the different time periods and concluded, that most likely the decade 1900-1910 contains most images.

A.9.4 Visual Prototype:

The following tasks and questions refer to the new *visual prototype [P-INTF]* (<https://historymap.wien>).

- **Please take a moment to browse around the collection**
The interviewee immediately looked at the temporal bar chart and confirmed his guess about the decade with the most pictures. He browsed around the map, looking at some pictures in his neighbourhood. He opened the sidebar and looked at different visualizations. He was confused by the map overlays as only the very center of the city is visible on the historic maps.
- **What is your first impression of the interface? How would you assess the general usability of the interface?**
The interviewee really likes the interface. He states that because of the different design and the visualizations, it will retain the interest of the users much better than the *search-based interface [S-INTF]*.
- **Find a picture of Schönbrunn**
The interviewee had troubles finding an image on the map as there was still a filter active. Eventually he figured this out, cleared the filter and then quickly was able to find a picture of Schönbrunn on the map.
- **Find an old picture (created before 1750) of Schönbrunn**
The interviewee opened the cluster located at Schönbrunn on the map, and concluded

that the radial oriented icons should be sorted by their creation date. He clicked in the very center (the oldest images) and was able to open a correct image.

- **Find a picture of a church close to your home**
The interviewee tried to search for the term "Kirche 22" using the list view but was unsuccessful. He then continued to change to the map view, open the keyword cloud, clicking on the keyword "Kirche" and then opening one of the resulting images on the map located relatively closely to his home.
- **Which decade (10 years period) has the most pictures in the collection?**
The interviewee answered this question already before but confirmed again using the temporal bar chart.
- **Which author created most pictures in the collection?**
The interviewee opened the author bubble chart and could give the correct answer immediately.
- **Which author created most pictures dated 1800 or earlier?**
The interviewee applied the filter on the temporal bar chart and then looked at the author bubble chart. He was a bit confused about the largest bubble with the label "others" but could give the correct answer.
- **Find 2 authors which were working often with each other and open a picture created by them.**
The interviewee looked at the author bubble chart but realized that it will not help him solve the task. After some more searching he found the author network visualization and immediately understood the meaning of the thickness. He clicked first on one author and then on the second, filtering the images on the map. He opened one of the resulting images.
- **List a few keywords used often throughout the collection**
The interviewee immediately opened the keyword cloud visualization and gave the correct answer.
- **Find a few pictures of buildings standing in the place of the old vienna city wall.**
The interviewee was already familiar with the map overlay feature and therefore immediately could finish the task successfully by applying the map overlay and then searching for a correct image directly on the map.

A.9.5 Questionnaire: Evaluation of Visualizations

A.9.6 Final Questions and Conclusion

The interviewee likes the modern look of the interface and especially enjoys the map feature of the *visual prototype [P-INTF]*. He thinks that it keeps the user's interest much better than the *search-based interface [S-INTF]*. He likes the variety in visualizations

A. INTERVIEW NOTES

(e.g. keyword cloud and bubble chart for the same type of quantitative data). He thinks the detail page should be resized automatically, hence avoiding a scroll bar. He also very much would like to see the addition of user accounts, allowing the users to bookmark images to their account to be able to view them later.