



FAKULTÄT FÜR **INFORMATIK**

Visualisierung massenhafter zeitbezogener Daten

DIPLOMARBEIT

zur Erlangung des akademischen Grades

Diplom-Ingenieur

im Rahmen des Studiums

Wirtschaftsingenieurwesen Informatik

eingereicht von

Robert Klausner

Matrikelnummer 0200904

an der
Fakultät für Informatik der Technischen Universität Wien

Betreuung:
Betreuer: Univ. Doz. Dipl.-Ing. Dr. techn. Johannes Gärtner

Wien, 2.12.2008

(Unterschrift Verfasser)

(Unterschrift Betreuer)

ERKLÄRUNG

Robert Klausner

Kirchenstraße 1

4532 Rohr im Kremstal

„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.“

Ort, Datum, Unterschrift

ABSTRACT

Data analysts are facing a fast and steadily growing amount of data in various application domains. There are many approaches to gain deep insight and understanding of the data such as classical statistics or data mining. Visual Exploration focuses on human's remarkable perceptual abilities. Users can scan, recognize and recall images rapidly, and can detect changes in size, color, shape, movement or texture.

Time-series are a data type of utmost importance. In many domains, huge amounts of complex, time-related data are collected and analyzed. Many of them relate to human behavior; social rhythms (e.g. the week) leave a footprint in data stored.

This paper proposes both a method and a tool to systematically visualize patterns, trends, regularities and anomalies in time-oriented data. It shows an approach to develop an understanding of time-oriented data and uses a matrix-based visualization technique.

The design of the visualization and the graphical user interface were evaluated in a survey. The findings of the survey are presented in this work and outline the shapes of future work. (E.g. the demand for more interactivity concerning zoom and filter operations or the wish for predefined color sets)

KURZFASSUNG

Die technologische Entwicklung der letzten Jahre und Jahrzehnte ermöglicht die Speicherung und Analyse immer größerer Datenbestände. Neben den klassischen Methoden der Statistik und des Data Mining gewinnt die visuelle Exploration an Bedeutung, die die visuellen Fähigkeiten von Menschen bei dieser Aufgabe nutzt.

Ein wichtiges Analysefeld sind massenhafte Daten mit Zeitbezügen, die sozialen Rhythmen folgen. Soziale Rhythmen wie z.B. die Woche sind Rahmenbedingungen und Ergebnis menschlichen Handelns.

Diese Arbeit zeigt einen Weg und ein Werkzeug, Muster, Trends, Regelmäßigkeiten und Unregelmäßigkeiten im System derartiger zeitbezogener Daten zu visualisieren und damit besser fassbar zu machen und bedient sich dabei der Möglichkeiten einer Matrix-basierten Darstellung.

Zur Überprüfung der Visualisierung und des Bedienkonzeptes wurde eine Benutzerstudie durchgeführt. Dabei wurden Experteninterviews mit der Methode der qualitativen Inhaltsanalyse ausgewertet. Aus den Ergebnissen wurden Handlungsempfehlungen für die weitere Entwicklung formuliert. Ein Beispiel dafür ist das Verlangen nach mehr Interaktivität bei den Zoom und Filter Operationen oder der Wunsch nach vordefinierten Farbsets.

INHALTSVERZEICHNIS

Erklärung	2
Abstract	3
Kurzfassung	4
Inhaltsverzeichnis	5
Einleitung	8
Motivation und Problemstellung	8
Aufgabenstellung und Überblick	11
Kapitel 1 Visualisierung	12
Visualisierung von Zeit	12
Definition	12
Visualisierung Allgemein	14
Überblick	18
Zeitreihen	18
Karten.....	26
Zusammenfassung	31
andere Darstellungen	31
Visualisierung grosser Datenmengen.....	34
Visualisierung großer Mengen zeitbezogener Daten	37
Kapitel 2 Realisierung	41
Architektur	41
Allgemein.....	42
Vorgehensmodell	43
Extreme Programming	45
Kritik an XP	48

Anwendungsfälle/Stories	50
Szenario 1 - Pivot.....	51
Szenario 2 - Datenlisten.....	52
Mapping der Daten.....	54
Beispiel.....	54
Farbgebung.....	58
Grundlagen.....	58
Farben in der Visualisierung.....	58
Farben für Karten.....	62
Kapitel 3 Benutzerstudie.....	67
Ausgangssituation und Fragestellung	67
Methodisches Design	67
Vorbereitung „Anwendungsfälle“	69
Auswahl der Teilnehmer.....	70
Setting und Kommunikation	70
Acrobat.com.....	70
Skype	71
CAMTASIA.....	71
Der Zeitrahmen der Gespräche	71
Arbeitsanweisungen „Anwendungsfälle“	72
Bemerkungen	72
Anwendungsfall 1 „Operator Liste + Einstellungen + Farbendialog“.....	72
Anwendungsfall 2 „Operator Liste+ Farbendialog“	73
Anwendungsfall 3 “Operator Pivot”.....	73
Anwendungsfall 4 „Operator Pivot + Farbe Perzentil“	74
Leitfragen während/nach den Anwendungsfällen.....	75

Auswertung Interviews	76
Interview A	76
Interview B	77
Interview C	79
Interview D	80
Interview E.....	81
Interview F.....	81
Kategorienbildung	83
Auswertung der Kategorien	85
Allgemeines	85
Berechnung	85
Einstellungen Farbe Allgemeines	86
Einstellungen Farbe Farben	89
Einstellungen Farbe Markierungen bearbeiten	90
Einstellungen Diagramm	92
Legende.....	92
Zeichenfläche.....	93
Zusammenfassung.....	94
Zusammenfassung und Ausblick	96
Abbildungsverzeichnis	99
Literaturverzeichnis	101

EINLEITUNG

MOTIVATION UND PROBLEMSTELLUNG

Das Mooresche Gesetz besagt, dass sich die Anzahl der Transistoren auf einem Prozessor alle 18 Monate verdoppelt (vgl. [Moo98]). Diese in den 60er Jahren formulierte Aussage hat sich bis heute bestätigt und es ist kein Ende dieser technischen Entwicklung in Sicht. Ein anderer Indikator für die Leistungsfähigkeit aktueller Computersysteme kann beim Marktführer bei Suchmaschinen gefunden werden [Goo08]. Während 1998 noch 26 Millionen Seiten indiziert wurden, schafften es Google-Rechner im Juli 2008 etwa eine Billion Seiten im Index zu halten. Diese Zahl gibt nicht nur Auskunft über die im Web gefundenen Seiten, sondern zeigt auch, dass Computersysteme in der Lage sind, unglaubliche große Datenmengen zu speichern und Abfragen über diese Daten auszuführen.

Das Finden relevanter Informationen in großen Ansammlungen von Daten ist eine nicht-triviale Aufgabe. Ein möglicher Ansatz ist die Visualisierung; d.h. abstrakte Daten und Zusammenhänge in eine graphische bzw. visuell fassbare Form zu bringen [Wik084]. Für die Analyse bedeutet dies, Menschen in den Prozess der Datenexploration einzubinden und die in diesem Zusammenhang wichtigen Fähigkeiten von Menschen – Flexibilität, Kreativität und das Allgemeinverständnis – mit der Rechenkraft moderner Rechensysteme zu kombinieren [Kei02], [Kei01]. An grundlegenden Konzepten der Visualisierung wird seit langem gearbeitet (siehe [Ber74], [Ber82], [Tuf83], [Shn96], [Sch00] und [Che08]).

Die Visualisierung von zeitbezogenen Daten birgt einige Besonderheiten, die in der Entwicklung und Anwendung zu berücksichtigen sind [Aig07], [Mül03]. Bei der Suche nach Mustern, Strukturen und Zusammenhängen müssen die Daten oft ohne eine vorab formulierte Hypothese durchsucht werden [Hao07]. Zu diesem Zweck sind Pixelorientierte (auch als Matrix-basiert bezeichnet) Techniken besonders gut geeignet. Sie ermöglichen die Darstellung besonders vieler Daten auf einer Bildschirmseite und können auf eine Aggregation verzichten. Es gibt keine Probleme mit Überlappungen. Auch multivariate Datensätze können abgebildet werden [Kei00].

In dieser Arbeit wird der Begriff „Zeit“ nicht als mechanische Abfolge von Sekunden verstanden. Unser soziales Leben wird von Rhythmen bestimmt, die mit der Zeit in Zusammenhang stehen (vgl. [Zer81], [Zer85], [Zer03] und [Gär04]). Die oben genannten Muster, Strukturen und Zusammenhänge können solche Rhythmen sein, die bei der Analyse von zeitbezogenen Daten eine wichtige Rolle spielen. Die Kenntnis über diese Rhythmen bestimmt das Verständnis eines Datensatzes.

Der wahrscheinlich fundamentalste Rhythmus ist die Sieben-Tage-Woche. Sie wurde durch das Judentum, das Christentum und den Islam praktisch in der gesamten Welt verbreitet. Die Woche bildet einen Grundbaustein des zeitlichen Ordnungssystems. In der Praxis ist die Woche ein bestimmender Faktor für die Anrufstatistik oder den Einsätzen bei der Polizei, für Aufnahmen, Behandlungen und Operationen im Krankenhaus oder für die Anwesenheit der Belegschaft in einem produzierenden Betrieb.

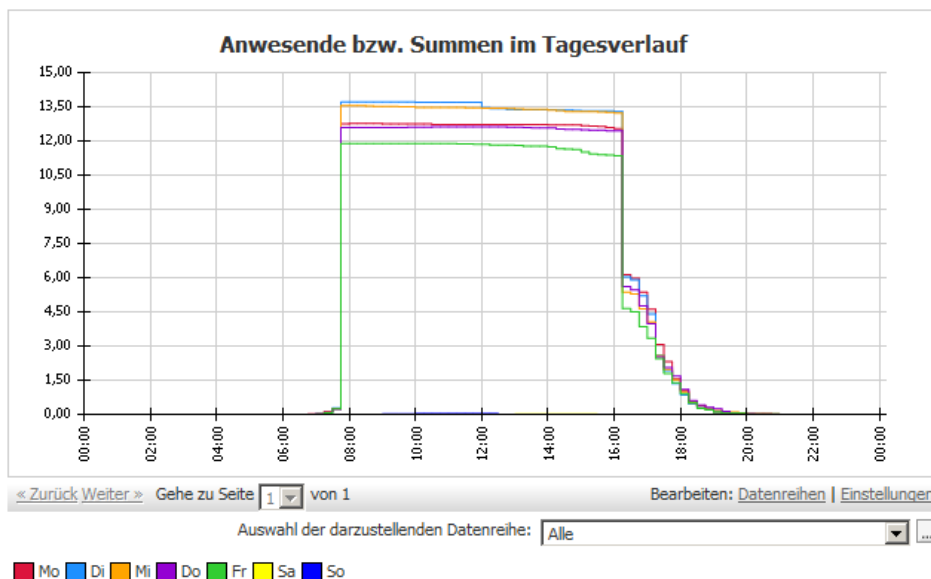


Abbildung 1 Anwesende Mitarbeiter

Diese Grafik zeigt, wie wichtig eine Analyse auf Wochenebene ist. Es ist zu erkennen, dass die Anwesenheit der Mitarbeiter immer um sechs Uhr morgens beginnt, um 16 Uhr stark abfällt und dass nach 20 Uhr keine Mitarbeiter im Betrieb anwesend sind. Sehr leicht sieht man, dass am Freitag die geringste und am Dienstag sowie am Mittwoch die höchste Anwesenheit vorzufinden ist. Nun kann im nächsten Schritt untersucht werden, was die Ursache dafür sein könnte (im konkreten Beispiel sind es nicht häufige Krank-

heiten, sondern unterschiedliche Häufigkeiten von Feiertagen, die Durchschnitte mit geringer Anwesenheit am Feiertag entsprechend verschieben).

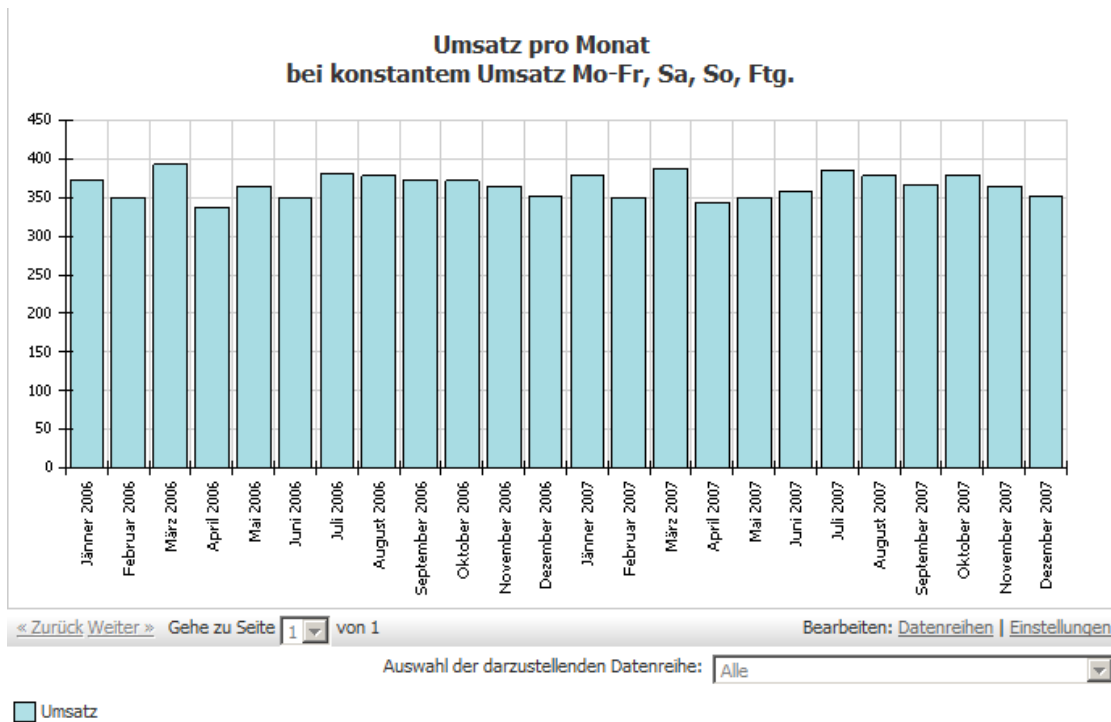


Abbildung 2 Umsatz pro Monat

Die Grafik in Abbildung 2 Umsatz pro Monat zeigt, wie problematisch ein nicht optimal gewählter Auswertungszeitraum ist. Es werden die Monate der Jahre 2006 und 2007 dargestellt. Die Umsätze sind jeden Tag konstant. Diese Darstellung vermittelt jedoch ein komplett anderes Bild. Ein Betrachter würde sofort mit der Interpretation beginnen und sich fragen, woher diese Schwankungen kommen. Diese Darstellung erfüllt das in [Sch00] formulierte Kriterium der Expressivität nicht, sie vermittelt einen falschen Eindruck der Wirklichkeit und birgt das Risiko falscher Interpretationen und Schlussfolgerungen.

Diese Beispiele bieten einen kleinen Einblick in die Problemstellungen und Implikationen bei der Arbeit mit zeitbezogenen Daten, die in dieser Arbeit behandelt werden.

AUFGABENSTELLUNG UND ÜBERBLICK

Ziel dieser Arbeit ist, eine Visualisierung und eine entsprechende Software zur Analyse massenhafter zeitbezogener Daten zu entwickeln. Im ersten Kapitel wird ein Überblick über die Visualisierung von Zeit gegeben. Dazu werden zentrale Begriffe definiert und wichtige Techniken der Visualisierung vorgestellt und diskutiert.

Die Diskussion der sozialen Rhythmen von Zeit ist in diesem Kapitel von immanenter Bedeutung. Sie impliziert einige Eigenschaften einer Visualisierung, die auf diese Besonderheiten Rücksicht nehmen will.

Im nächsten Kapitel wird auf die Realisierung der vorgestellten Visualisierung eingegangen. Dabei wird erläutert, in welcher technischen Umgebung die Entwicklung stattfand und warum das Vorgehensmodell „Extreme Programming“ zur Erstellung der Software gewählt wurde. Ein weiterer wichtiger Bereich ist die Transformation der Eingabedaten hin zur Visualisierung. Dabei spielte auch die Auswahl der Farben der Darstellung eine besondere Rolle.

Zur Überprüfung der Eigenschaften, Möglichkeiten und Grenzen der entwickelten Visualisierung wurde eine Benutzerstudie durchgeführt. Dabei wurden in Experteninterviews Anwendungsfälle abgearbeitet, die mit Hilfe einer qualitativen Inhaltsanalyse ausgewertet wurden. Die Methode, Durchführung und die Ergebnisse werden im Kapitel „Benutzerstudie“ vorgestellt.

Kapitel 1 VISUALISIERUNG

VISUALISIERUNG VON ZEIT

DEFINITION

Die Analyse von Zeitreihen und zeitbezogenen Daten ist für viele Bereiche der Wissenschaft, Technik und Wirtschaft eine wichtige Aufgabe. Die Visualisierung dieser Daten zielt häufig darauf ab, Strukturen und Zusammenhänge der Daten zu erkennen. In einer zeitbezogenen Visualisierung können verschiedene Aspekte in Zusammenhang mit der Zeit kodiert werden:

- Existence of an entity¹
Existiert eine Entität² zu einer spezifischen Zeit? Die zeitliche, wie auch die räumliche Existenz hängen vom Kontext ab. Auf einer Karte etwa scheint eine Entität nur dann auf, wenn sie eine ausreichende räumliche und zeitliche Größe bei der Darstellung einnimmt.
- Temporal location
Wann existiert eine gewisse Entität? Hier kann die Zeit als ein Zeitpunkt oder ein Zeitintervall gegeben sein (z.B. 15:30 oder Juli 2008). Es kann auch eine Ansammlung von Zeitpunkten oder Zeitintervallen gegeben sein (z.B. die Temperatur zu Mittag oder die Niederschlagsmenge eines Monats)
- Time interval
Wie lange ist die Zeitspanne vom Beginn bis zum Ende einer Entität?
- Temporal texture
Wie oft tritt eine Entität auf? Dies ist eine Funktion des Zeitintervalls zwischen zwei Entitäten in Bezug zu einer Zeiteinheit. (z.B. fünf mal pro Stunde, einmal in 100 Jahren)

¹ alle nach [Mac95] p. 316

² Das englische Wort „entity“ wird im Bereich der Datenmodellierung mit „Entität“ übersetzt. und ist hier als „Informationsobjekt“ oder „Datenpunkt“ zu verstehen. Siehe dazu auch [Wik08]

- Rate of change
Wie schnell ändert sich eine Entität oder wie groß ist die Differenz von einer Entität zur nächsten?
- Sequence
In welcher Reihenfolge treten die Entitäten auf?
- Synchronization
Treten Entitäten gemeinsam auf?

Eine gute Visualisierung kann Antworten auf viele oder alle in der Aufzählung genannten Punkte geben. Eine Definition zu zeitbezogenen Daten kann bei Wolfgang Müller und Heidrun Schumann gefunden werden [Mül03]:

„Therefore, we can define that time dependent data is characterized by data elements being a function of time. In general, data takes the following form:

$$d = f(t)$$

For data defined at discrete time stamps t_i , this relation can be represented as:

$$D = \{(t_1, d_1), (t_2, d_2), \dots (t_n, d_n), \}$$

where

$$d_i = f(t_i)$$

“

Eine idente Definition, die sich nur in der Benennung der Variablen unterscheidet, geben Marc Weber et al. [Web01]:

„Time series data is characterized by data elements being a function of time. In general, this data takes the following form:

$$D = \{(t_1, y_1), (t_2, y_2), \dots (t_n, y_n)\}$$

with

$$y_i = f\langle t_i \rangle$$

“

Die Datenelemente d_i oder y_i können verschiedene Datentypen darstellen. Eine gängige Unterteilung erfolgt nach nominalen, ordinalen oder quantitativen Daten. Neben den Daten können auch die Zeitachsen genauer beschrieben werden.

VISUALISIERUNG ALLGEMEIN

Nach Schuhmann und Müller hat die wissenschaftlich-technische Visualisierung die Aufgabe, geeignete visuelle Repräsentationen einer Menge von Daten zu erzeugen, um eine effektive Auswertung zu ermöglichen.³ Die Visualisierung dient der Analyse, dem Verständnis und der Kommunikation von Modellen, Konzepten und Daten. Tufté betont, dass die Visualisierung oft der beste Weg ist, um eine Menge von Zahlen zu beschreiben und zu verstehen. Auch große Datensätze lassen sich meist am besten verstehen, wenn man Abbilder davon betrachtet. Colin Ware zeichnet in [War04] ein ähnliches Bild: Visualisierungen schaffen die Möglichkeit, große Mengen an Daten zu verstehen. Sie ermöglichen es, Eigenschaften von Daten wahrzunehmen, die während der Betrachtung sichtbar werden und nicht vorhersehbar waren. Die Entdeckung eines Musters kann eine neue Erkenntnis über wohlbekanntes Daten hervorrufen. Die Visualisierung kann Fehler ans Tageslicht bringen, die bei der Sammlung des Datenbestands passiert sind und kann erklären, wie die Daten gesammelt wurden. Das bedeutet, dass eine Visualisierung etwa systematische Fehler bei der Erhebung ans Tageslicht bringen kann.

Schuhmann und Müller unterteilen die Visualisierung in drei Stufen:

³ nach [Sch00] p. 5 ss.

- die explorative Analyse
- die konfirmative Analyse
- die Präsentation⁴

Die erste Stufe sucht nach Strukturen und Eigenschaften in Datensätzen, über die noch keine Hypothesen vorliegen. Die Suche ist ungerichtet, die Daten allein bilden den Ausgangspunkt. Hier kann eine Visualisierung Hinweise zur Formulierung von Hypothesen über die Daten liefern.

Die konfirmative Analyse basiert auf Daten, zu denen bereits eine Hypothese existiert. Diese soll mit Hilfe einer geeigneten Visualisierung auf ihre Tauglichkeit überprüft werden. Analog zu einer statistischen Überprüfung mit Tests wird die These im Anschluss angenommen oder verworfen.

Der letzte Schritt im Prozess der Analyse stellt die Kommunikation und Präsentation dar. Bestätigte Fakten und Aussagen sollen nun erkennbar dargestellt werden. Das Publikum soll die Ergebnisse identifizieren und verstehen können.

Im Gegensatz zu Schuhmann unterteilt Ware in [War04] den Prozess der Visualisierung von Daten in vier Stufen:

- The collection and storage of data itself
- The preprocessing designed to transform the data into something we can understand
- The display hardware and the graphics algorithms that produce an image on the screen
- The human perceptual and cognitive system (the perceiver)⁵

⁴ *ibid.*

⁵ [War04] p. 5

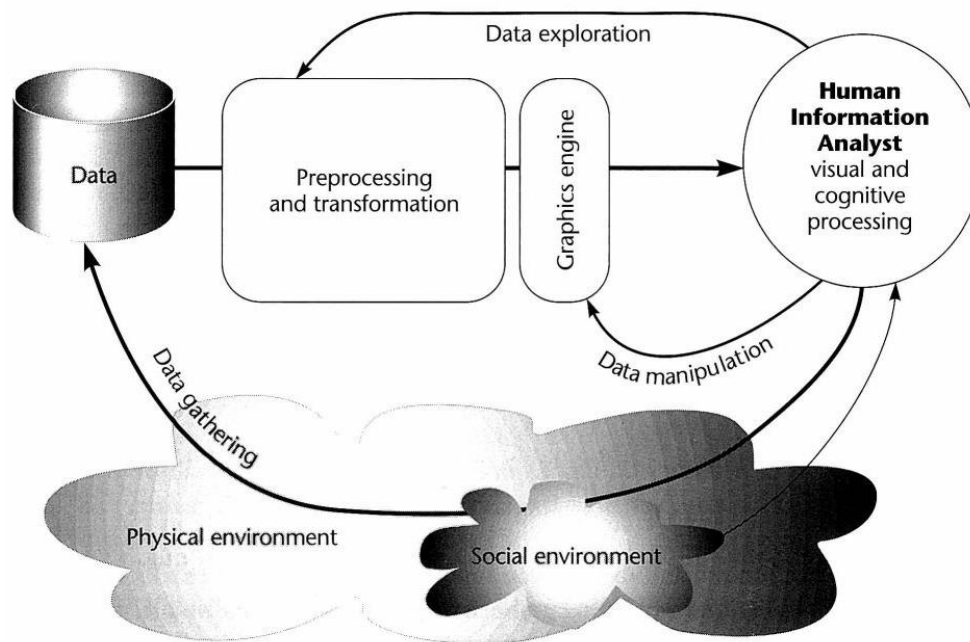


Abbildung 3 Schematische Darstellung des Visualisierungsprozesses [War04]

Bei Ware beginnt der Prozess der Visualisierung bereits beim Sammeln und Speichern der Daten. Dieser Schritt ist genau genommen eine Schleife, die immer und immer wiederholt wird, da sich das Interesse an den Daten bei jeder Betrachtung ändern kann. Das bedeutet, dass zum Beispiel ein Wissenschaftler die Entscheidung trifft, neue oder mehr Daten zu sammeln, wenn er im vorhandenen Material interessierende Dinge entdeckt, die er genauer untersuchen möchte. Aus der Betrachtung kann sich der Bedarf für mehr Daten oder mehr Details ergeben oder die Transformation vor der Visualisierung kann geändert werden. Der Prozess kann auch interaktiv sein, das heißt die Parameter werden während der visuellen Analyse festgelegt. Moderne Softwarepakete bieten viele Möglichkeiten zur interaktiven Exploration eines Datenbestandes. Während dieser Schleife werden die Daten in zweierlei Hinsicht untersucht. Ware unterscheidet in einen physikalischen und einen sozialen Rahmen. Der physikalische Rahmen ist die Datenquelle, während der soziale Rahmen subtil und komplex bestimmt, wie die Daten gesammelt und interpretiert werden. Der nächste und wichtigste Schritt ist, die Daten in eine Form zu bringen, die von der menschlichen Wahrnehmung leicht aufgenommen werden kann:

„The critical question is how best to transform the data into something that people can understand for optimal decision making.“⁶

Für Schuhmann und Müller bedeutet dies, dass eine Visualisierung expressiv, effektiv und angemessen sein muss. Expressivität bedeutet:

„Nur die in den Daten enthaltenen Informationen und nur diese sollen durch die Visualisierung dargestellt werden. Diese Fähigkeit einer Visualisierung wird als Expressivität oder Ausdrucksfähigkeit einer visuellen Darstellung bezeichnet.“⁷

Das heißt, dass eine Visualisierung unter keinen Umständen Informationen enthalten darf, die nicht direkt aus den Daten kommen. Dieses Kriterium ist eine Grundvoraussetzung zur Erreichung einer guten Qualität der Darstellung.

Die Effektivität stellt die Frage, welche Visualisierungstechniken die Daten am besten präsentiert, natürlich unter der Voraussetzung, dass das Expressivitätskriterium erfüllt ist:

„Es wird also die Visualisierung gesucht, die bei der Darstellung die (visuellen) Fähigkeiten des Betrachters und die charakteristischen Eigenschaften des Ausgabegeräts unter Berücksichtigung der Zielsetzung und des Anwendungskontexts optimal ausnutzt.“⁸

Hier ist entscheidend, dass hinter einer Visualisierung eine Zielsetzung steckt, an der die Umsetzung gemessen werden kann. Eine Visualisierung kann meist nicht allen möglichen Bearbeitungszielen dienen.

Das Kriterium der Angemessenheit betrachtet die Kosten, die mit der Erstellung der Visualisierung verbunden sind. Aufwand und Nutzen sind gegenüber zu stellen und zu bewerten. Die Angemessenheit beschränkt sich aber nicht nur auf die Erstellung, sondern auch auf den Aufwand zur Interpretation durch den Betrachter. In der Praxis sind die Effektivität und die Angemessenheit eng miteinander verknüpft.

⁶ [War04] p. 5

⁷ [Sch00] p. 9

⁸ [Sch00] p.11

Zusammenfassend kann man feststellen, dass eine gute Visualisierung immer das Kriterium der Expressivität erfüllen muss, um das Risiko falscher Interpretation und Schlussfolgerungen zu minimieren und auch eine möglichst unverfälschte Darstellung zu erreichen. Zusätzlich soll eine Visualisierung die beiden Kriterien der Angemessenheit und Effektivität erfüllen. Eine weitere Diskussion vieler Aspekte von Visualisierungen finden sich in [Cle85], [Few04] oder [Nie97], um hier nur einige Beispiele zu nennen.

ÜBERBLICK

Die Visualisierung ist seit langem eine beliebte Methode, zeitbezogene Daten zu repräsentieren. Einige frühe Beispiele aus der Geschichte sind William Playfair, Etienne-Jules Marey oder Charles Joseph Minard, die im späten 18. respektive frühen 19. Jahrhundert anzusiedeln sind. Tufte stellt fest, dass die Darstellung von Zeit meist als Karte oder Zeitreihe passiert:

„Many information displays report on the world’s workaday reality of three-space and time. Painting four-variable narrations of space-time onto flatland combines two familiar designs, the map and the time series.“⁹

In diesem Abschnitt wird ein Überblick über die Entwicklung der Visualisierung von Zeit als Zeitreihe, Karte und anderen Darstellungsformen gegeben.

ZEITREIHEN

Die Darstellung von Zeit als Zeitreihe ist eine der beliebtesten Formen, um den abstrakten Begriff Zeit darzustellen:

„The time-series plot is the most frequently used form of graphic design.“¹⁰

Bei einer Zeitreihe wird auf einer Dimension die Zeit aufgetragen. Diese Achse stellt die Zeitachse dar:

⁹ [Tuf90] p.97

¹⁰ [Tuf83] p.28

„A time-series graph displays quantitative values in relation to multiple, sequential points in time. Consequently, one axis of the graph needs to serve as the time scale, with labels for each unit of time (years, quarters, etc.). Values of time have a natural order.“¹¹

Praktisch jede Zeitreihe verwendet die horizontale Achse zur Darstellung der zeitlichen Dimension und die vertikale Achse zur Darstellung der quantitativen Daten.

Nach [Tuf83] stammt die früheste Darstellung einer Zeitreihe in der bekannten Literatur von einem unbekanntem Mönch aus dem 10. oder 11. Jahrhundert, der verschiedene Planetenbewegungen in ein Gitternetz einzeichnete. Es sollten weitere 800 Jahre vergehen, bis dass die nächsten gezeichneten Zeitreihen auftauchen. Tufte bezeichnet dies als rätselhaftes und isoliertes Wunder in der Geschichte der Visualisierung. Auch bei C. Chen et al. findet sich in [Che08] die unten abgebildete Visualisierung, die als Pionierleistung gewürdigt wird und bis ins 17. Jahrhundert einzigartig bleibt. Dennoch hinterlässt die eingehende Betrachtung dieser Visualisierung einige Fragen. Die eingezeichneten Gitterlinien können als Koordinatensystem oder einfach nur als Lesehilfe gesehen werden. Genau genommen ist die Darstellung unten aber keine Zeitreihe im Sinne einer systematischen Visualisierung, die den Kriterien von Schuhmann oder Ware genügt. Vielmehr entstand diese Darstellung aus der direkten Übertragung der Wirklichkeit (d.h. der Planetenbahnen) auf einem Stück Papier in Ermangelung anderer Repräsentationsformen. Offen bleibt die Frage, ob diese Darstellung tatsächlich eine wissenschaftlich-technische Visualisierung zur effektiven Analyse oder bloß eine Notationsform zur visuellen Repräsentation beobachteter Phänomene ist.

¹¹ [Few04] p. 70

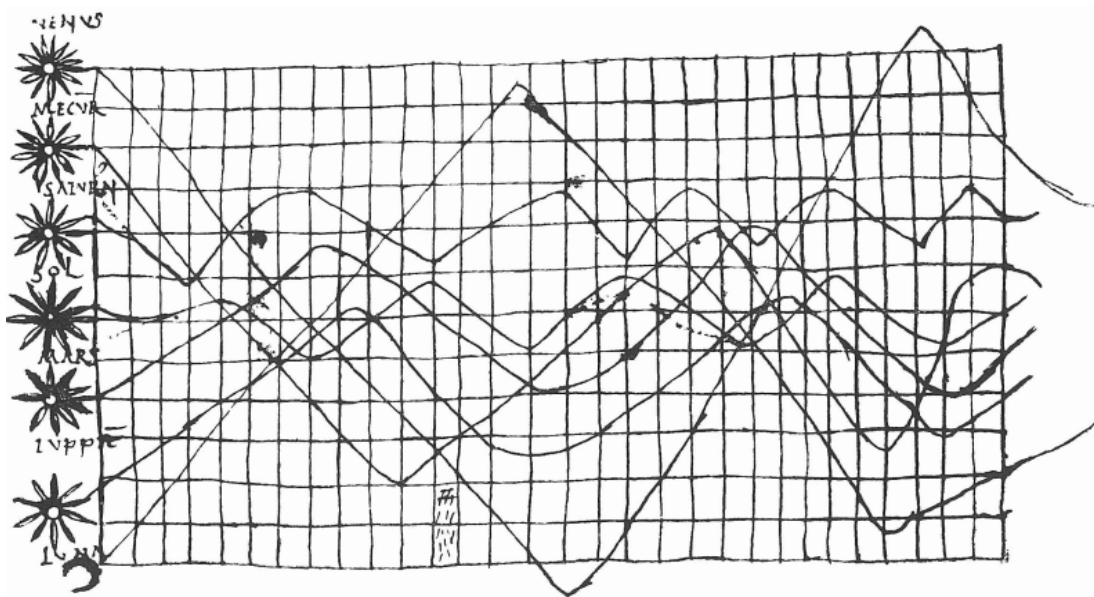


Abbildung 4 Planetenbewegungen 10./11. Jdt [Tuf83]

Erst im späten 18. Jahrhundert tauchen Zeitreihen wieder in den Schriften der Wissenschaft auf. Tufte verweist auf [Lam79], der in seinem Werk „Pyrometrie“ mehrere Zeitreihen darstellt. Eine davon ist die „Jährliche Erwärmung der Erde“ an verschiedenen geographischen Breiten.

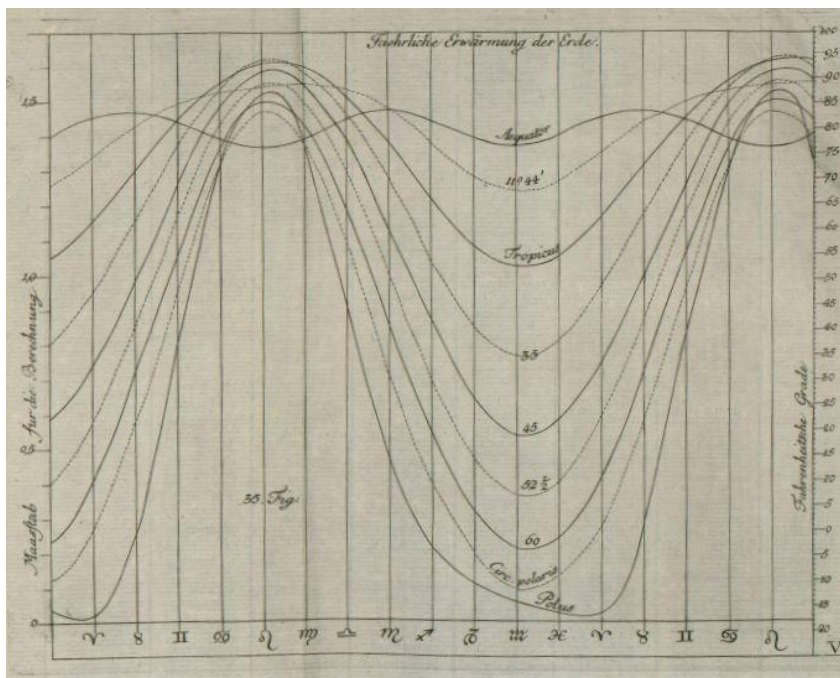


Abbildung 5 Jährliche Erwärmung der Erde von Lambert [Lam79]

Neben Johann H. Lambert gehört William Playfair zu den Erfindern moderner Informationsgrafik. Ihm wird die Erfindung des Balken- und Kreisdiagrammes zugeschrieben. In [Pla86] erstellte er eine Menge von Zeitreihen, die Wirtschaftsdaten repräsentieren:

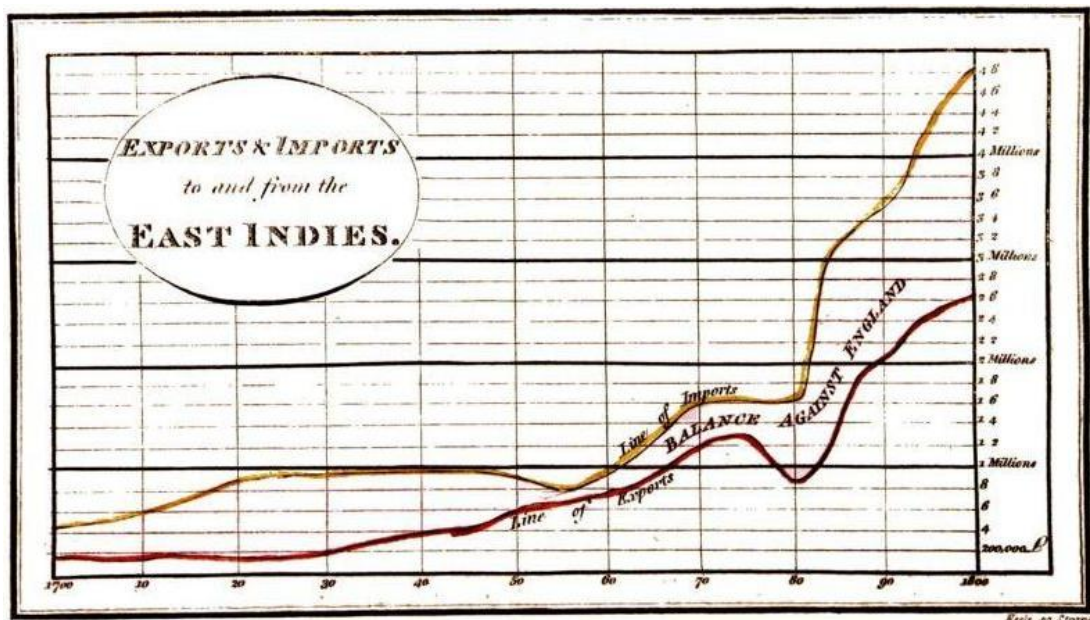


Abbildung 6 Playfair Exporte und Importe von/nach Indien [Pla86]

Das Bild unten zeigt die Kombination aus einer Zeitreihe und einem Balkendiagramm. Playfair sammelte den Preis von Weizen über 250 Jahre, genau so wie die Löhne von Handwerkern. Gleichzeitig sieht man die Regierungszeiten von Königen und Königinnen des britischen Empires. Ziel der Präsentation war die Analyse der Verhältnisse zwischen Preis und Einkommen, also der Reallöhne.

Entlang der X-Achse wurde die Zeit von 1565 bis 1821 aufgetragen. Die untere Kurve zeigt die Entwicklung der Löhne, die Balken visualisieren den Weizenpreis, der ab 1795 geradezu explodiert, während die Löhne in der Entwicklung nachhinken. Auch in [Che08] wird auf diese Visualisierung verwiesen. Besonders wichtig dabei ist, dass Playfair diese Visualisierung schuf, um eine Hypothese argumentativ zu festigen. Diese Darstellung ist eine typisch wirtschaftswissenschaftliche Visualisierung, die in abgewandelten Formen auch heute oft noch in der Literatur zu finden ist.

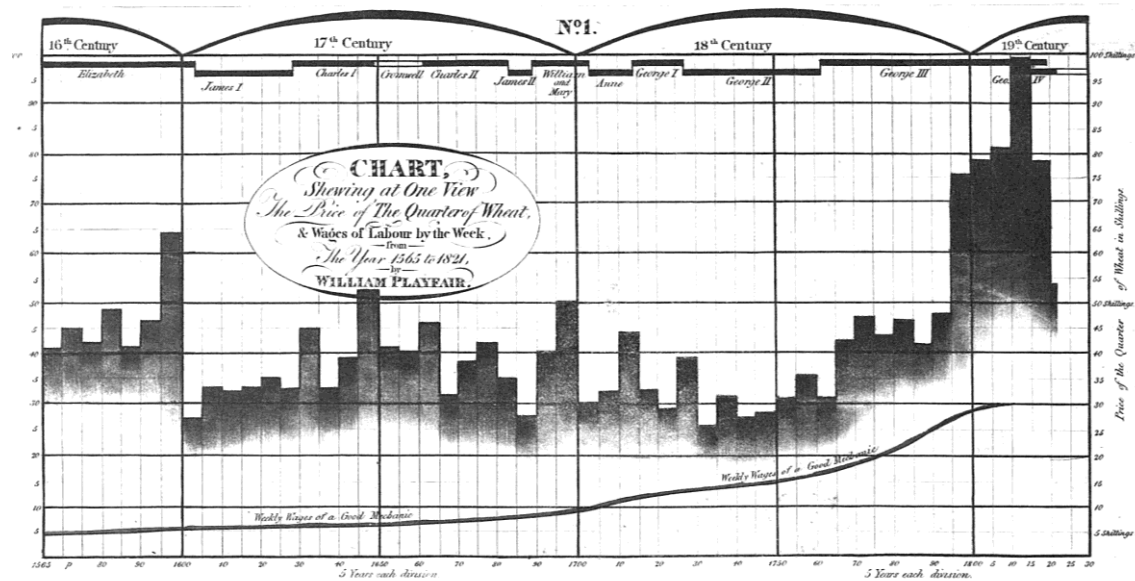


Abbildung 7 Playfair, Preise, Löhne und Regierungszeiten [Tuf83]

Im Vergleich zu der Abbildung aus dem 10. Jahrhundert basieren diese Zeitreihen auf einer seriellen Zuordnung einer Ausprägung einer oder mehrerer Variablen zu einem Datum, während das Beispiel der Planetenbewegung die Visualisierung eines Prozesses mit einer zeitlichen Dimension darstellt. Die Darstellungen von Playfair geben der Visualisierung auch eine soziale Konnotation, da sie Zusammenhänge ans Licht bringt, die den täglichen Lebensalltag der Menschen betreffen.

Darstellungen dieser Form blieben mangels quantitativer Daten noch selten, während Daten aus der Astronomie oder Physik schon in sehr umfangreichen Maß zur Verfügung standen.

Anregungen zu dieser Darstellung kommen dann auch aus der wirtschaftswissenschaftlichen Konjunkturforschung des 19. Jahrhunderts, die durch die historische Demographie seit den fünfziger Jahren des 20. Jahrhunderts einen neuen Schub erhielt. (vgl. dazu [Che08] oder [Jev24])

Bei E.J. Marey (aus [Tuf83]) sehen wir, dass die Zeitachse nicht immer horizontal verlaufen muss. Die Abbildung stammt aus 1895 und zeigt die Schrittfolgen der verschiedenen Gangarten eines Pferdes:

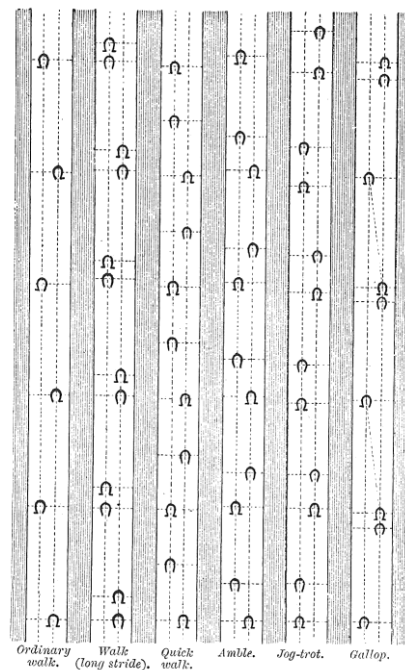


Abbildung 8 Marey, Gangarten des Pferdes [Tuf83]

Marey interessierte sich auch für die Darstellung historischer Zusammenhänge. Folgendes Bild zeigt Regierungszeiten britischer Könige und Königinnen, die auch Informationen über Kriegs- und Friedenszeiten enthält.

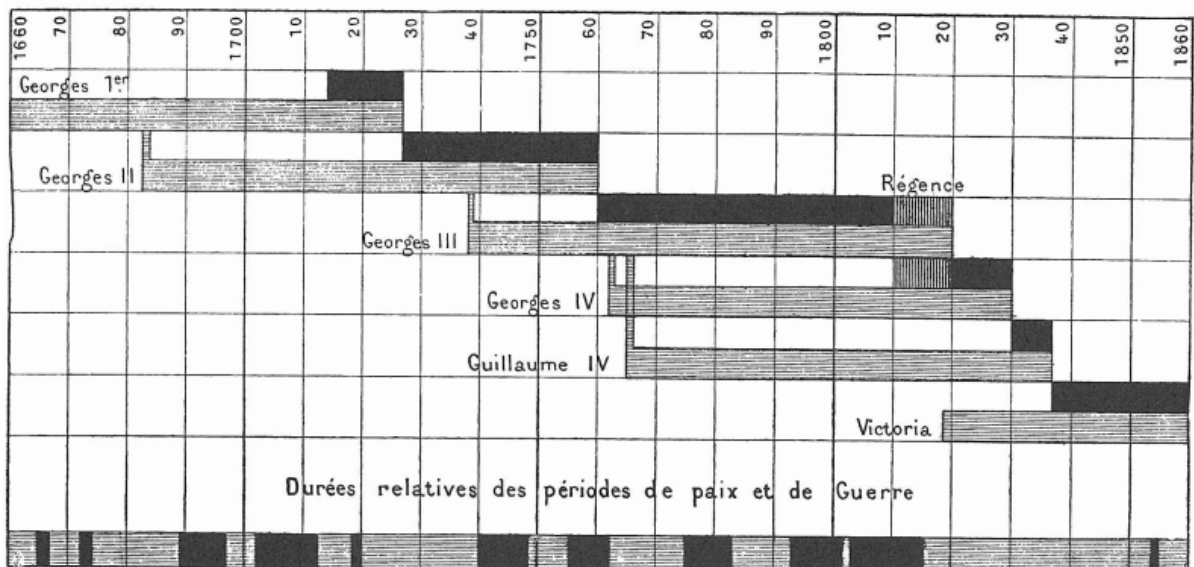


Abbildung 9 Marey [Tuf83]

Eine noch effektivere Methode zur Verbesserung der Zeitreihen ist die Hinzufügung räumlicher Dimension. So bewegen sich die Daten gleichzeitig über Raum und Zeit. Wenn die räumliche Dimension sehr geschickt integriert wird, fällt sie dem Betrachter kaum auf und wird doch sehr gut wahrgenommen. Ein Beispiel wird bei Tufte besonders hervorgehoben:

„It may well be the best statistical graphic ever drawn.“¹²

Gemeint ist die Karte von Charles Joseph Minard, der den Marsch der napoleonischen Armee nach Russland beschreibt und ihren Untergang nachzeichnet.

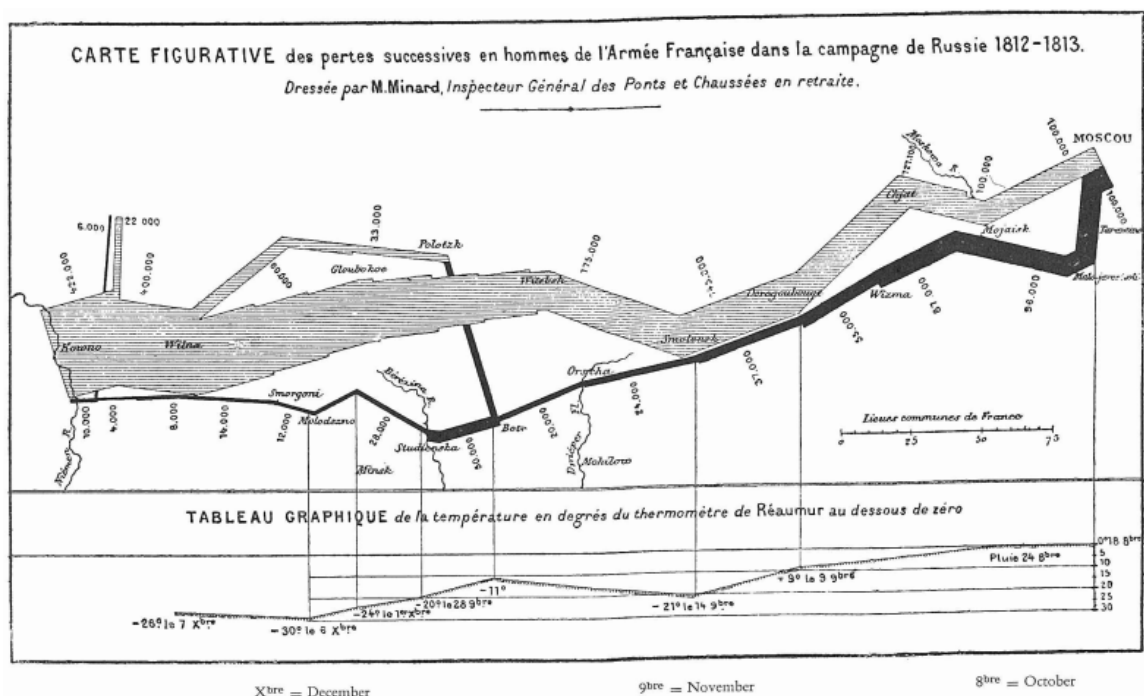


Abbildung 10 Minard, Russlandexpedition der napoleonischen Armee [Tuf83]

Diese Visualisierung kombiniert eine Karte mit einer Zeitreihe, die die schrecklichen Verluste der Armee auf eine beeindruckend einfache Weise nachzeichnet. Sechs verschiedene Variablen werden auf zwei Dimension dargestellt: Die Zeit, die Position auf einer zweidimensionalen Karte, die Richtung, die Temperatur und die Breite der Linie zeigt die Größe der Armee. Napoleon beginnt den Feldzug mit 422.000 Mann, von denen 100.000 in Moskau ankommen. Dort wechselt die Farbe der Darstellung von grau

¹² [Tuf83] p. 40

auf schwarz, was die Richtungsänderung der Armee signalisiert. Die schwindende Breite der Linie zeichnet die starken Verluste nach, die auch ohne Feindeinwirkung durch die tiefen Temperaturen verursacht wurden, die wir am unteren Rand der Visualisierung sehen.

Noch ein Beispiel zeigt uns bei Tufte, wie auf vortreffliche Art und Weise Zeit und Raum in einer Visualisierung kombiniert werden können.

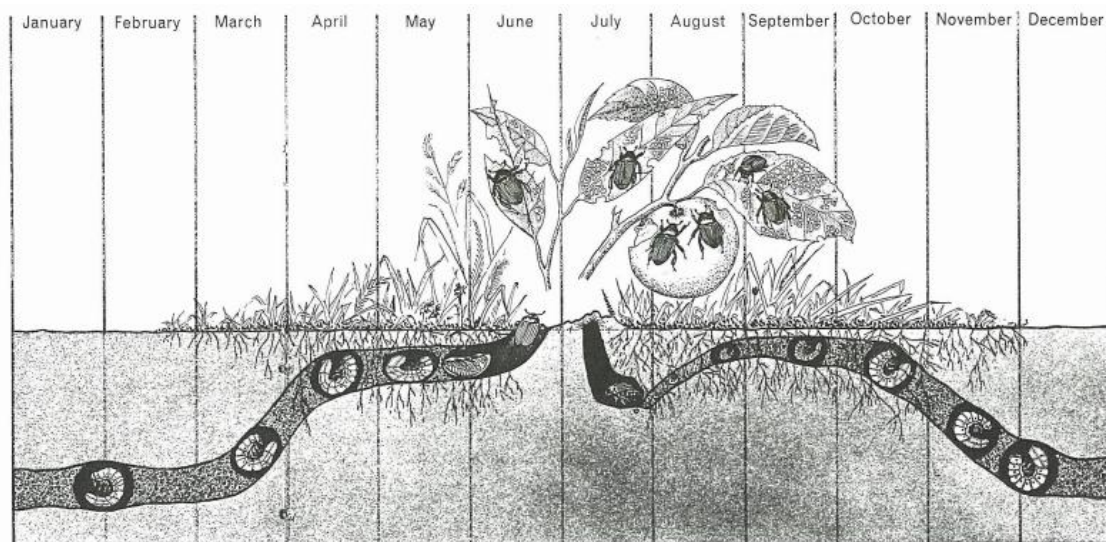


Abbildung 11 Newman, Man and Insects [Tuf83]

Auf der horizontalen Achse verläuft ein Jahr von Januar bis Dezember. Die vertikale Achse zeigt uns die Position der Larve im Erdboden respektive an der Erdoberfläche. Die Darstellung zeigt also den Lebenszyklus des „Japanese beetle“.

Ein letztes Beispiel zu Zeitreihen ist die Darstellung der Geschichte des Pop und Rock. Auf der vertikalen Achse ist auf einem Bereich von 0% bis 85% der relative Anteil am Gesamtmarkt der verkauften Platten eingetragen. Auf der horizontalen Achse verlaufen die Jahre von 1955 bis 1974. Fett formatierte Wörter sind Kategorien (z.B. „Instrumental R+R“, „Uptown R+B“ oder „Jazz/Rock“). Diese parallelen Verläufe vermessen Künstler in zwei Dimensionen – in einer Eltern-Kind Relation, die Einflüsse untereinander ausdrückt sowie ihre Zeitgenossen für das jeweilige Jahr mit ihrem Anteil am Gesamtmarkt. Diese Grafik ist ein detailliertes und gleichzeitig auch faszinierendes Beispiel für eine Zeitreihe.

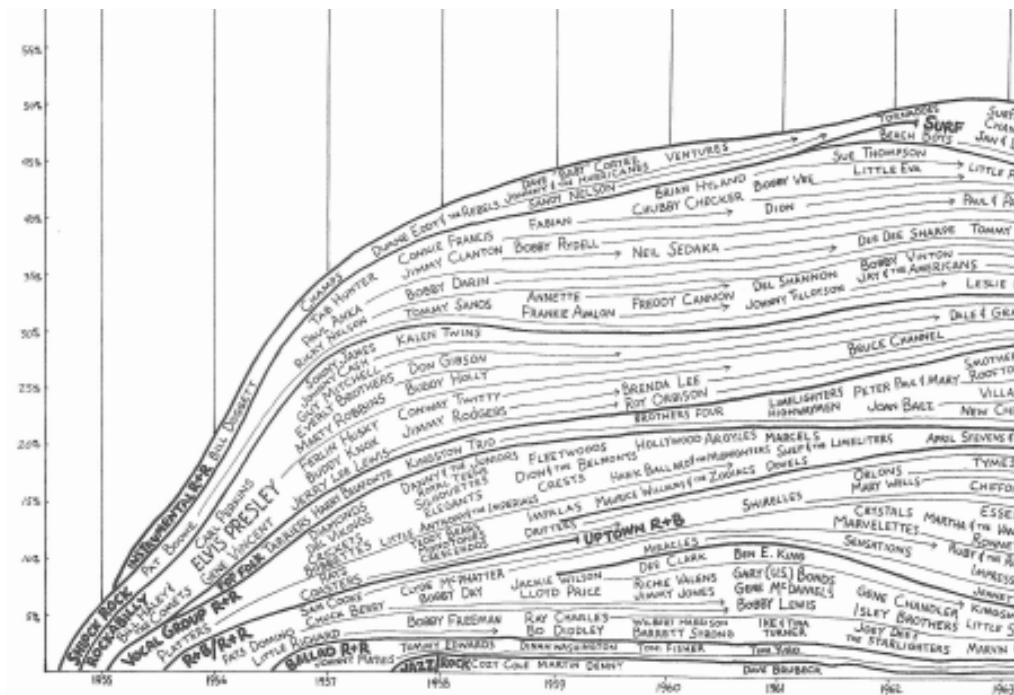


Abbildung 12 The History and Politics of the Music Industry (Ausschnitt, [Tuf05])

Es wurde gezeigt, dass sich Zeitreihen für die Darstellung von Daten mit zeitlichem Bezug verschiedenster Herkunft eignen und ein gern verwendetes Konzept zur graphischen Repräsentation ist.

KARTEN

Die bei [Ber74] rein geographische Definition der Karte wird bei [Mac95] ausgedehnt. Der Begriff der Karte wird in dieser Arbeit ausgedehnt und wird nicht nur geographisch verstanden. Aus der Geographie lassen sie viele gut erforschte Konzepte wie die Chloprolethenkarte oder Schemata zur Bestimmung der Farben wie aus [Bre031] und [Bre032] übernehmen. Die unten vorgestellten Konzepte Heatmap und Pixel-Based Charts werden bei [Che08] als Matrix Visualization bezeichnet

HEATMAP

Eine gängige Definition gibt Stephen Few:

„In general, the term heatmap refers to any display that uses color to represent quantitative data.“¹³

Diese sehr allgemein gehaltene Formulierung meint jede Art der Darstellung von Daten, in der Farben verwendet werden, quantitative Informationen darzustellen. Im alltäglichen Leben sehen wir zum Beispiel oft Wetterkarten, die Temperaturen in dieser Form zeigen.

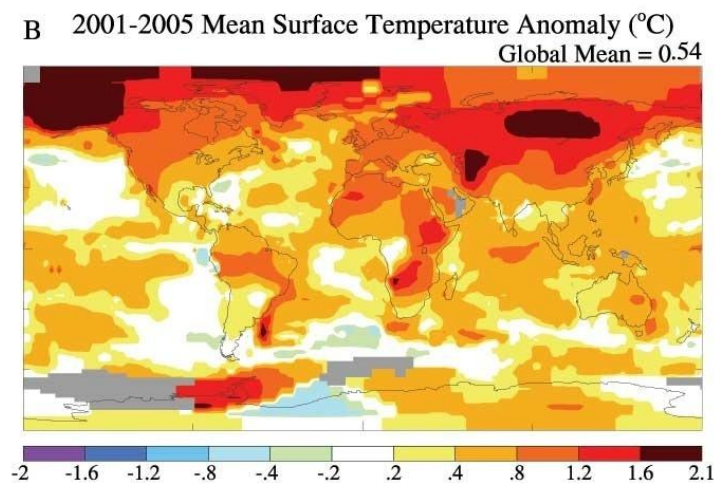


Abbildung 13 Oberflächentemperatur der Erde

In der multivariaten Datenanalyse werden Daten üblicherweise in einer Matrix dargestellt. Die folgende Abbildung zeigt ein Beispiel für diese Darstellung. Sie visualisiert das Ergebnis einer Umfrage unter Angestellten nach der Arbeitszufriedenheit. In jeder Zeile wird ein Mitarbeiter dargestellt, während in den Spalten die Variablen zu sehen sind. In diesem Beispiel sind dies „Working Conditions“, „Salary“, „Average Hours Per Week“, „Average Percent Travel“, „Supervisors’s Ranking“, „Self Ranking“. Diese Variablen wurden in einer Umfrage unter Angestellten einer Firma erhoben, jeder Mitarbeiter beurteilte die Arbeitsbedingungen in Hinsicht auf die Arbeitszufriedenheit mit „Very Poor“ (Weiß), „Poor“, „Acceptable“, „Good“ oder „Very Good“ (Schwarz)¹⁴. Die Sortierung erfolgte nach der ersten Spalte. Nun kann man diese Graphik zeilenweise lesen, um sich pro Angestellten einen Überblick über die Arbeitszufriedenheit zu verschaffen. Natürlich ist es in diesem Fall nicht sehr einfach, auf den ersten Blick Zu-

¹³ [Few06] p. 1

¹⁴ nach [Few06] p. 2

sammenhänge zwischen den verschiedenen Variablen herzustellen. Wenn man sich aber ein wenig Zeit nimmt, die Graphik zu lesen, kann man einige Implikationen herauslesen. Besonders interessant scheinen die Variablen „average hours per week“, „average percent travel“ und „self rating“ zu sein. Je mehr pro Woche gearbeitet wird, desto schlechter werden die Arbeitsbedingungen bewertet. Die Mitarbeiter, welche sehr viel für ihren Job reisen müssen, beurteilen ihr Arbeitsumfeld auch eher schlecht. Diejenigen, die ihre Leistung als gut einschätzen, bewerten auch die Arbeitsbedingungen besser. Alle Aussagen ließen sich durch Filterung oder Sortierung der Daten nach den übrigen Variablen besser untermauern.

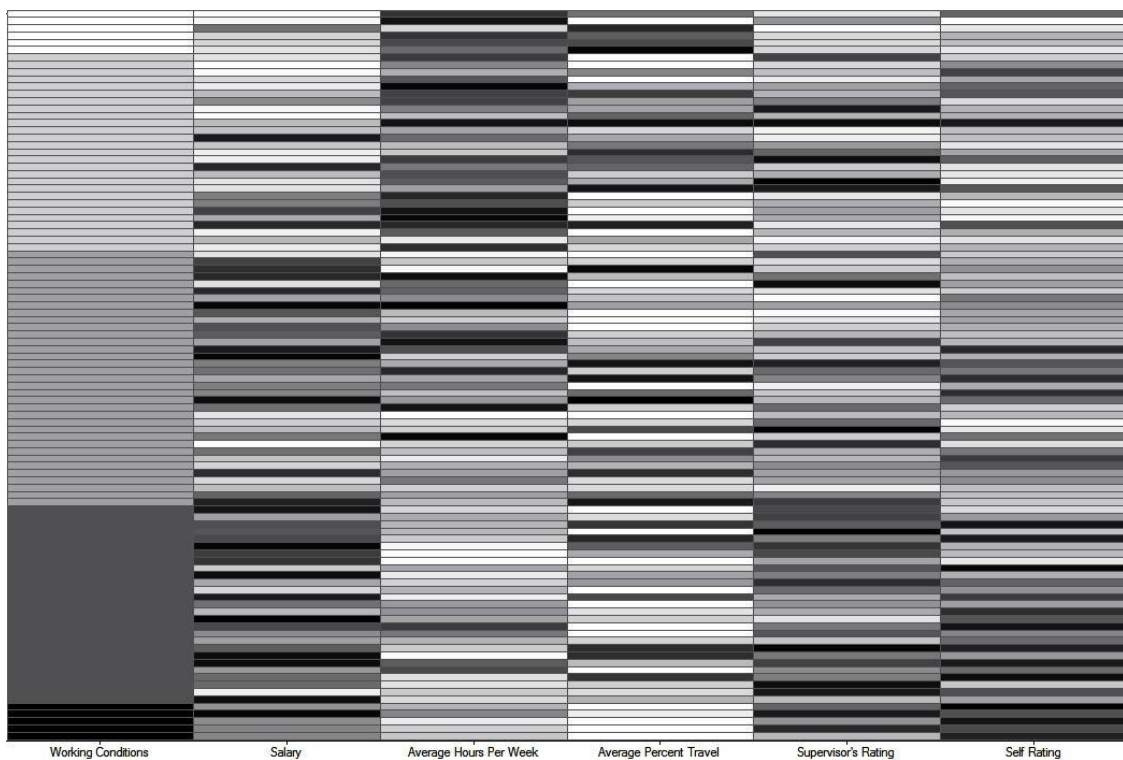


Abbildung 14 Eine typische multivariate Heatmap Matrix¹⁵

Wichtig ist die Feststellung, dass diese Darstellung keine Daten weglässt, summiert oder aggregiert. Jeder Datensatz bleibt vorhanden und lässt sich auch visuell erfassen. Durch die Abstraktion in der eigenen Wahrnehmung lassen sich sehr gut Zusammenhänge herstellen bzw. kann man wie hier Abhängigkeiten der Variablen feststellen. Hier

¹⁵ aus [Few06] p. 2

zeigt sich die Stärke der Visualisierung zur visuellen Analyse eines Datenbestandes, die mit statistischen Verfahren ungleich aufwändiger und bei Weitem nicht so intuitiv wäre.

PIXEL-BASED CHARTS

Die von mir als Landkarte bezeichnete Visualisierung wird in der Literatur in die Kategorie „Pixel-based Charts“ eingeordnet. Ziel dieser Darstellung ist es, große Datenmengen in einem Diagramm darzustellen. Ein einzelnes Datenelement kann bis zur Größe eines Pixels geschrumpft werden, um möglichst viel Information in ein Diagramm zu bringen. Bei einer angenommenen Auflösung von 1.000x1.000 Pixel lassen sich bereits 1.000.000 Datenpunkte direkt darstellen. So kann vorerst auf eine Aggregation oder Filterung der Daten verzichtet werden. Der einzelne Punkt wird mit einer Farbinformation versehen, die dem Wert entspricht. Ziel ist es, die Möglichkeit zu schaffen, Trends und Muster in großen Datenbeständen zu erkennen. Dabei kommt der Aufteilung und Wahl der Farben eine große Bedeutung zu, da sie uns erst die Grundlage zur erwünschten Erkennung der Trends und Muster liefert. Die Forschungsgruppe CAVIS (Center for Analysis and Visualization of large Information Spaces) an der Universität Konstanz beschäftigt sich besonders mit dem Thema. [Kei00] beschreibt wichtige grundlegende Überlegungen, die auch für die vorliegende Anwendung von großer Relevanz sind:

„To achieve that objective a number of design problems have to be solved. The first problem is the mapping of data values to colors. A good mapping is obviously very important, but has to be [sic] carefully engineered to be intuitive. A second very important question is how pixels are arranged inside the subwindows.“¹⁶

Dieser Ausschnitt beschreibt sehr treffend die Kernfelder meiner Forschungsfrage. Eine Herausforderung ist die Farbgebung der Datenpunkte, die andere ist die Anordnung der Pixel. Die Anordnung hängt von der Art der Daten und dem Zweck der Visualisierung ab, in dieser Arbeit konzentriere ich mich auf zeitbezogene Daten.

In [Hao07] wird die Idee der *Pixel-Matrix* vorgeschlagen. Dieses Paper beschreibt, dass Diagramme oft nur aggregierte Daten enthalten und dadurch ein Informationsverlust stattfindet, der die Ausgangsdaten verfälschen kann. Die Darstellung in Form der Pixel-

¹⁶ [Kei00] p. 3

Matrix ermöglicht es, wichtige Bereiche mit mehr Information auszustatten. Die beiden Abbildungen zeigen sehr gut, was mit dieser Aussage gemeint ist:

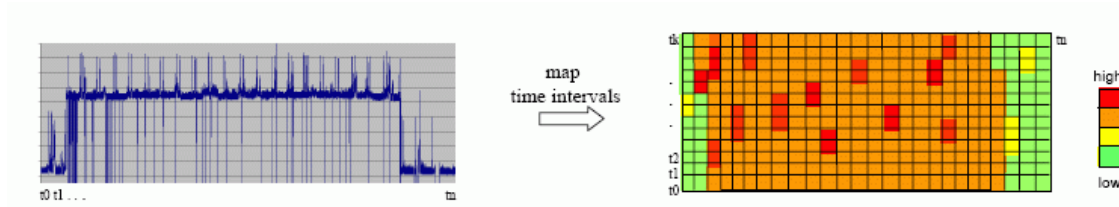


Abbildung 15 Zeitreihe eines Wartungsvertrags, die Ausschläge stellen Verletzungen der Antwortzeit dar¹⁷.

Diese Abbildung zeigt eine Zeitreihe, ein Ausschlag nach oben (entlang der y-Achse) weist auf eine Verletzung des Wartungsvertrages hin. Je weiter die Kurve nach oben ausschlägt, desto schwerer wiegt die Verletzung. Die Kurven sind schwer zu lesen, die kleinen Ausreißer nach unten und oben sind nicht auf einen Blick zu interpretieren. Der Darstellung auf der rechten Seite liegen die gleichen Daten zu Grunde, hier wurde der Grad der Verletzung des Wartungsvertrags in vier Farbbereiche unterteilt. Grün signalisiert niedrig, rot steht für hoch. In der Pixel-Matrix Darstellung sieht man auf einen Blick den großen orangen Bereich in der Mitte, in dem einige rote Rechtecke auftauchen, während zu Beginn und am Ende der Zeitreihe alles in Ordnung ist.

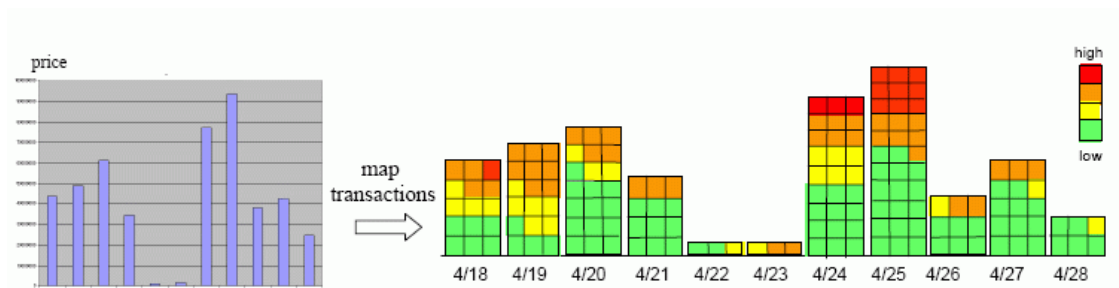


Abbildung 16 Darstellung der Umsätze pro Tag¹⁸

Die zweite Graphik verdeutlicht den Unterschied bei der Visualisierung von Umsatzdaten und zeigt, wie diese Darstellung eine zusätzliche Dimension darstellen kann. Das Balkendiagramm auf der linken Seite zeigt nackte Umsatzdaten nach Tagen. Die Umsetzung als Pixel-Matrix bringt eine zusätzliche Aussage, nämlich die über den erzielten Preis einzelner Transaktionen. Die Höhe spiegelt wie in der linken Seite die erzielten

¹⁷ aus [Hao07] p. 2

¹⁸ aus [Hao07] p. 2

Umsätze wider, während die Farbe etwas über die Qualität der Umsätze aussagt. Rot steht hier für hohe Preise. Der Leser des Diagramms kann leicht erkennen, dass z.B. am Tag „4/25“ neben dem höchsten Umsatz auch sehr hohe Preise erzielt wurden, da die Rechtecke mit den Farben Rot und Orange fast die Hälfte der Fläche einnehmen, während am Tag „4/28“ fast nur Umsatz mit niedrigpreisigen Artikeln erzielt wurde. Weiters kann man erkennen, dass es wohl eine gewisse Korrelation „hohe Preise“ → „hohe Umsätze“ gibt.

ZUSAMMENFASSUNG

Die wesentlichen Vor- und Nachteile der Visualisierung als Landkarte sind:

- + Leichtes Erkennen von Mustern, Korrelationen, Abhängigkeiten, Wiederholungen
- + Unverfälschte Darstellung der Daten (keine Aggregation, keine Überdeckung)

Neben diesen sehr guten Eigenschaften gibt es auch noch Folgendes zu berücksichtigen

- Intuitiver Zugang für Non-Professionals kann schwierig sein

Wie weiter oben bereits festgestellt, liegen die größten Herausforderungen im „Mapping“, also der Transformation der Daten in die tatsächliche Darstellung am Bildschirm und in der Farbgebung der Datenpunkte.

ANDERE DARSTELLUNGEN

Neben den oft verwendeten Formen zur Darstellung von Zeit, die auf Zeitreihen und Karten zurückgehen, möchte ich noch einige Techniken vorstellen, die sich nicht in diese Kategorien einordnen lassen.

SONNENUHR

Die Sonnenuhr ist das früheste Instrument zur Bestimmung der Tageszeit. Sie diente den alten Ägyptern ab etwa dem 3. Jahrtausend v. Chr. zur Einteilung des Tages in mehrere Abschnitte. Ein Punkt oder Stab hinterlässt abhängig vom Stand der Sonne am Himmel einen Schatten auf einem Ziffernblatt, von dem die Tageszeit abgelesen werden kann. (Nach [Wik081])

Die Projektion des Schattens auf eine Skala stellt damit die erste Visualisierung von Zeit dar.

Sonnenuhren blieben bis zum Ende des Römischen Reiches in Verwendung, gerieten dann in Vergessenheit und kamen über die Kreuzzüge wieder zurück nach Europa, wo sie erst später durch das Aufkommen mechanischer Uhren ersetzt wurden.

WASSERUHR

Die Geschichte der Wasseruhr geht ins 2. Jahrtausend v. Chr. zurück. Das Prinzip war den Sumerern, Ägyptern und Chinesen bekannt. Ein Gefäß mit einem bestimmten Volumen wird mit Wasser gefüllt. Nach einer festen Zeiteinheit ist der Inhalt durch ein Loch ausgelaufen und kann so zur Zeitbestimmung verwendet werden. Diese Methode ist im Gegensatz zur Sonnenuhr unabhängig vom Wetter. Die Gefäße müssen nur vor Frost geschützt sein. (Nach [Wik083])

Die Konstruktion legt eine Metapher vom Fluss der Zeit bzw. „Verrinnen“ der Zeit nahe.

SANDUHR

Die Sanduhr ist erst seit dem 13. oder 14. Jahrhundert bekannt und ist eng mit der Wasseruhr verwandt. Zwei Glaskolben sind durch ein dünnes Rohr verbunden, wobei einer der Kolben mit Sand gefüllt ist. Mit dem Umdrehen der Sanduhr rinnt der Inhalt von oben nach unten. Die Menge des eingefüllten Sandes ist so festgelegt, dass die Dauer des Durchrieselns einem festen, definierten Zeitraum von zum Beispiel einer Stunde entspricht. (Nach [Wik082])

Auch diese Konstruktion ist neben ihrer inhärenten Funktion eine Visualisierung von Zeit, die uns als Metapher für das „Verrinnen“ von Zeit ständig begleitet – die meisten der populären Betriebssysteme und Anwendungen verändern den Cursor der Maus vom Pfeil zur Sanduhr, wenn eine längere Berechnung ansteht.

LEGO-POWERED TIME TRACKING

Michael Hunger beschreibt in seinem Blog [Hun08] eine sehr unkonventionelle Methode, seine Arbeitszeit zu dokumentieren. Er verwendet Lego-Bausteine, um seine Arbeitszeit an verschiedenen Projekten abzubilden. Die Länge des Bausteines repräsentiert

die Dauer. Es gibt Bausteine mit der Länge eins, zwei, drei und vier. Das bedeutet dann 15, 30, 45 und 60 Minuten Arbeitszeit. Die Bausteine werden dann übereinander geschichtet um den Arbeitstag abzubilden. Die verschiedenen Farben der Bausteine repräsentieren die Arbeit an den unterschiedlichen Projekten. Ein Arbeitstag sieht dann so aus:

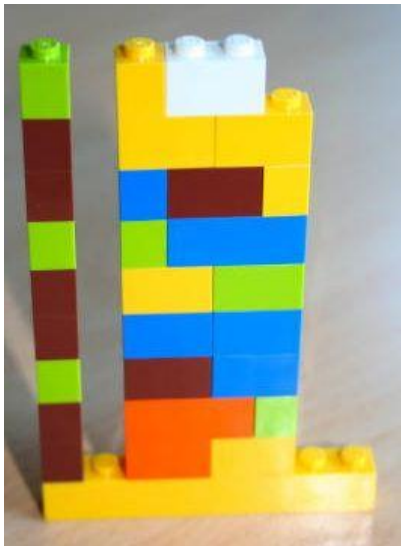


Abbildung 17 Lego Powered Time Tracking – ein Arbeitstag

Der Stapel links ist ein Maßstab für die Arbeitsstunden von zehn Uhr bis 18 Uhr. Das Bild zeigt, dass von zehn Uhr an eine halbe Stunde für das Projekt orange und eine halbe Stunde für das Projekt gelb gearbeitet wurde. Für eine ganze Arbeitswoche sieht das dann so aus:



Abbildung 18 Lego Powered Time Tracking Wochenansicht

Die Farben, auf die die Bausteine gestapelt werden stehen für die Wochentage Montag bis Freitag. Das Ergebnis ist eine gute Übersicht über eine ganze Arbeitswoche. Es ist sehr gut abschätzbar, wie viel relative Zeit an einem Projekt gearbeitet wurde. Auch die Planung einer Arbeitswoche funktioniert perfekt mit diesem Werkzeug. Man braucht sich nur eine Menge an Lego-Steinen vor Beginn einer Woche zurechtlegen – natürlich in den richtigen Farben für die verschiedenen Projekte – und merkt dann sofort wo die Planung fehlgeschlagen ist, wenn die Lego-Steine zur Neige gehen. Michael Hunger arbeitet auch einer Applikation, die ein Foto der Lego-Steine auswertet und automatisch ein eine Kalender-Anwendung überträgt. Diese Methode macht Zeit zu einem haptischen Erlebnis und damit zu einer wirklich greifbaren Abbildung eines abstrakten Konzepts.

VISUALISIERUNG GROSSER DATENMENGEN

Die technologische Entwicklung der letzten Jahrzehnte ermöglicht die Speicherung und Bearbeitung riesiger Datenmengen. Ein billiger Standardrechner ist heute mit Arbeits- und Festplattenspeicher so bestückt wie ein Supercomputer vor zehn Jahren. Neben der sich nach dem Mooreschen Gesetz entwickelnden Prozessorleistung steigt die Menge der oft völlig automatisch aufgezeichneten Daten enorm an:

„Forscher an der Universität Berkeley haben berechnet, dass jedes Jahr ca. 1 Exabyte (= 1 Million Terabyte) Daten generiert werden – ein großer Teil davon in digitaler Form. Das bedeutet aber, dass in den nächsten 3 Jahren mehr Daten generiert werden als in der gesamten menschlichen Entwicklung zuvor.“¹⁹

Die größte Herausforderung stellt sich an das Finden wertvoller Information in großen Datenmengen. Wenn diese Aufgabe nicht sinnvoll bewältigt wird, verkommen diese Datenmengen zu den Müllhalden der Informationsgesellschaft. Für Daniel Keim liegen in [Kei02] die Vorteile der visuellen Datenexploration auf der Hand. Während bei der Ausgabe als Text auf dem Bildschirm oder Drucker nur einige wenige Datensätze auf eine Seite passen, kann die visuelle Datenexploration die Daten in eine geeignete Form bringen, die es dem Menschen erlaubt, einen Eindruck in die Struktur der Daten zu bekommen, Schlussfolgerungen daraus zu ziehen und mit den Daten zu interagieren.

„Visuelle Data-Mining-Verfahren haben in den letzten Jahren einen hohen Stellenwert innerhalb des Forschungsbereichs Data Mining erhalten. Ihr Einsatz ist immer dann sinnvoll, wenn wenig über die Daten bekannt ist und die Explorationsziele nicht genau spezifiziert sind. Dadurch dass der Mensch direkt am Explorationsprozess beteiligt ist, können die Explorationsziele bei Bedarf verändert und angepasst werden.“²⁰

Diese Vorgehensweise entspricht auch dem von [War04] beschriebenen Prozess der Visualisierung. Gerade bei großen Datenmengen, die einer Analyse unterzogen werden, ist es oft nicht möglich, ex ante Hypothesen über die Daten zu formulieren. Erst die Visualisierung ermöglicht ein tieferes Verständnis für die Daten, wodurch ein Betrachter Hypothesen aufstellen kann, die mit visuellen Datenexplorationsverfahren untersucht und verifiziert respektive falsifiziert werden können. In dieser Phase können auch Techniken aus der Statistik zu Hilfe genommen werden.

[Shn96] unterteilt die visuelle Datenexploration in drei Schritte:

- overview
- zoom and filter

¹⁹ [Kei02] p. 1

²⁰ *ibid.*

- details-on-demand

Diesen Dreischritt bezeichnet [Shn96] als das Visual Information-Seeking Mantra. Zu Beginn verschafft sich ein Benutzer einen Überblick über die Daten (overview), um dann Muster, Zusammenhänge und Strukturen zu identifizieren. Mit Hilfe von Zoom- und Selektionsoperationen (zoom and filter) werden die gewonnen Erkenntnisse genauer untersucht werden. Für die genaue Analyse der Muster möchte ein Benutzer auf Details der Daten eingehen (details-on-demand). Für alle drei Schritte können Visualisierungstechniken eingesetzt werden.

„Es ist in diesem Zusammenhang wichtig, dass Visualisierungstechniken nicht nur grundlegende Verfahren für alle 3 Schritte bereitstellen, sondern auch die Schwierigkeiten bei den Übergängen zwischen diesen Schritten überbrücken helfen.“²¹

In [Kei00] zeigt Keim wie sehr sich pixel-orientierte Techniken für die Bewältigung großer Datenmengen eignen:

		Clustering	multivariate hot spot	no. of variates	no. of data items	categorical data	visual overlap	learning curve
Geometric Techniques	Scatterplot Matrices	++	++	+	+	-	o	++
	Landscapes	+	+	-	o	o	+	+
	Prosection Views	++	++	+	+	-	o	+
	Hyperslice	+	+	+	+	-	o	o
	Parallel Coordinates	o	++	++	-	o	--	o
Icon-based Techniques	Stick Figures	o	o	+	-	-	-	o
	Shape Coding	o	-	++	+	-	+	-
	Color Icon	o	-	++	+	-	+	-
Pixel-oriented Techniques	Query-Independent	+	+	++	++	-	++	+
	Query-Dependent	+	+	++	++	-	++	-
Hierarchical Techniques	Dimensional Stacking	+	+	o	o	++	o	o
	Treemap	+	o		o	++		o
	Cone Trees	+	+	o	+	o	+	+
Graph-based Techniques	Basic Graphs	o	o	-	+	o	o	+
	Specific Graphs	++	+	-	+	o	+	+

Abbildung 19 Vergleich multidimensionaler Visualisierungstechniken [Kei00]

In der Spalte „no. of data items“ findet sich bei den Pixel-orientierten Techniken der einzige Eintrag für sehr gut. (++: sehr gut, +: gut, o: neutral, -: schlecht, --: sehr schlecht) Die grundlegende Idee ist es, jedem Datenwert ein gefärbtes Pixel zuzuordnen

²¹ [Kei02] p. 1

und so ein Attribut zu visualisieren. Wenn also nur ein Pixel pro Wert benötigt wird, erlaubt es diese Technik, die größte Anzahl an Werten auf einen Bildschirm zu bringen.

VISUALISIERUNG GROßER MENGEN ZEITBEZOGENER DATEN

Neben den speziellen technischen Anforderungen an die Achsenbeschriftung, Skalierung etc. wird dem sozialen Rahmen bei der Interpretation zeitbezogener Daten eine besondere Bedeutung beigemessen. Wie schon oben erwähnt, meint der soziale Rahmen bei [War04] die subtilen und komplexen Wege, wie Daten gesammelt und interpretiert werden. [Zer81] stellt fest, dass einer der fundamentalsten Parameter der Ordnung der Welt die Zeit ist. Es gibt viele Muster und Strukturen, die die Zeit betreffen. Die Zeit hilft uns, Ereignisse im Sinne einer sequentiellen Struktur zu ordnen, eine Dauer zu bestimmen oder einzuordnen, wann sie stattgefunden haben. Ein vierter Parameter ist die periodische Wiederkehr bestimmter Ereignisse. Der Kalender wurde erfunden, um eine Regelmäßigkeit zeitlicher Ereignisse festzulegen und dann aufrechtzuerhalten.

Nach [Zer81] ist die Etablierung einer täglichen Regelmäßigkeit eine alte Idee, die aus der Religion entwickelt wurde. Zur Verehrung eines höheren Wesens wurden von Priestern regelmäßige Gottesdienste vorgeschrieben. Die Grundlage der westlichen Einteilung des Tages bis auf die Ebene einer Stunde geht auf die benediktinischen Klöster zurück. Die Regel des Hl. Benedikt ist der früheste Beweis für die Einführung der zeitlichen Regelmäßigkeit, die so charakteristisch für das moderne Leben ist. Sie stellt eine der fundamentalsten soziokulturellen Institutionen des modernen Westens dar.²² Die Einteilung des Tages durch eine Stundentafel, den *Horarien* wurde im Sommer und Winter unterschieden. Alle Ereignisse im monastischen Leben wiederholten sich in 24-Stunden Intervallen.

Es überrascht nicht, dass Instrumente zur Zeitmessung zuallererst in Klöstern ihre Verbreitung fanden. Über die Turmuhr wurde die Bevölkerung in Kenntnis über die aktuelle Stunde gesetzt. Auch hier lässt sich wieder die Bedeutung der Stunde erkennen: Minutenzeiger wurden – obwohl technisch machbar – viel später eingebaut weil sie in der von Stunden geprägten Einteilung des Tages einfach keine Bedeutung haben. Sie

²² [Zer81] p. 32

waren nur dazu gebaut, die Stunde anzuzeigen und nicht eine Zeitdauer zu messen. Die Frage nach der Zeit lautet zwar im Deutschen „Wie spät ist es?“ – in Sprachen wie Französisch, Spanisch oder Hebräisch fragt man „Welche Stunde ist es?“.

Ein weiterer bestimmender Rhythmus im Leben der Benediktiner ist die Woche. In der Liturgie werden bestimmte Psalmen im Wochenzyklus gewechselt. Das Konzept der Woche ist viel älter. Nach [Zer85] existieren dafür zwei Erklärungsmuster: Eines geht auf die Schöpfungsgeschichte im 1. Buch Mose zurück, wonach Gott in sechs Tagen die Welt schuf und am siebten Tage ruhte. Das andere basiert auf den sieben Planeten der antiken Astrologie. Das Konzept der sieben Tage wurde durch das Judentum, Christentum, den Hinduismus und den Islam auf der gesamten Welt verbreitet, sodass praktisch jeder erwachsene Mensch mit diesem Zyklus vertraut ist.

In vielen Sprachen ist das Wort „Woche“ ident mit dem Wort „sieben“ oder leitet sich direkt davon ab:²³ Hebräisch, Arabisch, Griechisch, Persisch, Armenisch, Serbokroatisch, Ungarisch, Bretonisch, Gälisch. Bei romanischen Sprachen wird das Wort für Woche aus dem Lateinischen „*septimana*“ abgeleitet: *settimana*, *semana*, *semaine*.

Die Bedeutung der Woche in der Sprache ist nur die Spitze des Eisbergs in der Bedeutung des Sieben-Tage-Zyklus. In der Etablierung und Ausbreitung dieses Zyklus spielten die Religionen die zentrale Rolle. Dabei ist es auch nebensächlich, dass der Höhepunkt der Woche entweder am Freitag, Samstag („Sabbat“) oder Sonntag angesetzt ist. Die Religion ist, so scheint es, der stärkste Unterbau einer Tradition. Der Versuch, im Zuge der französischen Revolution auf eine Zehn-Tage-Woche umzusteigen scheiterte genau so wie der Versuch der russischen Revolutionäre, eine Fünf-Tage-Woche zu etablieren. Beide hatten zum Ziel, den Einfluss der Religion zurückzudrängen.

„In both France and the Soviet Union, some desperate attempts were made by two of the most ruthless totalitarian regimes in history to completely destroy the Judeo-Christian seven-day week. In both societies, to this day, it still remains the dominant “beat” of social life.“²⁴

²³ [Zer85] p. 5

²⁴ *ibid.* p. 43

Nur aus diesem Verständnis heraus lassen sich folgende Fragen beantworten: Warum sind wir an Freitagnachmittagen unproduktiv? Warum besuchen wir Verwandte an Sonntagen und nicht alle 16 Tage? Warum ist die Suizidrate am Montag am Höchsten und am Wochenende am Niedrigsten? Warum „zählen“ Abende an einem Mittwoch weniger als die an einem Samstag?

Die Feststellung des Wochentages ist eines der ersten Dinge, die man nach dem Aufwachen tut. Es ist viel weniger wichtig, das genaue Datum zu kennen. Die Unkenntnis des Wochentages ist der erste Schritt zur Verbannung aus dem sozialen Leben.

Die Vereinbarungen im Zusammenhang mit der Zeit haben für die Gruppenbildung eine wichtige Bedeutung. Eine zeitliche Ordnung, die sich eine Gruppe teilt, hat eine Bedeutung hinsichtlich der Exklusion und Inklusion ihrer Mitglieder. Außenseiter werden klar abgetrennt, während die zeitliche Ordnung ein identitätsstiftendes Merkmal für Mitglieder ist.

Nach Max Weber ist die Beachtung des Sabbats das zentrale Merkmal der Trennung zwischen Juden und Nicht-Juden. Das Judentum betrachtete den Sabbat immer als den einen Tag, den nur sie verpflichtet waren zu ehren. Mit dem Entstehen des Christentums war es wieder wichtig, für dieses Gruppe einen Tag zu schaffen, der sich vom Sabbat unterscheidet:

„Choosing one particular day of the week as a day on which they could congregate on a regular basis was among the very first things the early Christians did as a group.“²⁵

Die Woche bildet den wichtigsten Grundbaustein des zeitlichen Ordnungssystems. Parallel dazu existieren Monate und Jahre, die mit der Woche synchronisiert und koordiniert werden müssen.

Die simple Division eines Kalendermonats oder Jahres hinterlässt immer einen Rest. Monate weichen immer zwei oder drei Tage von Wochen ab, während Jahre immer ein oder zwei Tage von 52 Wochen abweichen. (Auch ein Kartenspiel besteht aus 52 Karten und einem „Joker“, die vier Farben bilden, es könnte sein, dass sie ihren Ursprung in

²⁵ *ibid.* p. 72

einem symbolischen Kalender haben. Auch die Zahlenwerte der 52 Karten ergeben gemeinsam 364)²⁶

Jedes Kalenderjahr hat zwölf Monate und beginnt immer mit einem neuen Monat, dem Januar. Mit jedem neuen Jahr beginnt auch ein neuer Tag, eine neue Minute, eine neue Sekunde usw. Der Beginn einer neuen Woche fällt nur etwa jedes fünfte Jahr mit dem Beginn des neuen Jahres zusammen. Daraus ergeben sich einige Ungereimtheiten: Ein April, der an einem Donnerstag beginnt, hat 22 Werktage (Mo-Fr), wenn er an einem Sonntag beginnt, fallen nur 20 Werktage in den April. (nach [Zer85]) Dies ist neben den beweglichen Feiertagen ein weiterer Grund, warum eine Auswertung nach Monaten prinzipiell abzulehnen ist.

Diese zeitlichen Strukturen beeinflussen soziales Verhalten und werden gleichzeitig durch sie (re-)produziert. Sie stellen dadurch eine Möglichkeit der Analyse dar sowohl bezüglich ihrer konkreten Form und Ausprägung wie auch ihrer Veränderung. Diese sozialen Strukturen der Zeit in der visuellen Analyse nutzbar zu machen, ist das Ziel des nächsten Schrittes.

²⁶ *ibid.* p. 61

Kapitel 2 REALISIERUNG

ARCHITEKTUR

Das Software Engineering Institute der Carnegie Mellon Universität versucht den Begriff „Software-Architektur“ exakt zu definieren. Unter [Car08] finden sich dazu mehrere Dutzend Definitionsversuche. [Sta05] versuchte keine möglichst knappe Beschreibung, sondern erläutert den Begriff in mehreren Absätzen. Demnach ist Architektur Struktur. Sie definiert die Komponenten eines Systems, ihre Schnittstellen und ihre Beziehungen untereinander. Die Software-Architektur bildet den Übergang von der Analysephase zur konkreten technischen Realisierung. Sie beschreibt das Softwaresystem in unterschiedlichen Sichten. Die wichtigsten Sichten sind die Bausteinsicht, die Laufzeitsicht sowie die Zuordnungssicht, die eine Abbildung von Laufzeitbausteinen auf technische Infrastruktur (Hardware) beschreibt. Die Architektur macht die Komplexität beherrschbar und verständlich, in dem sie komplexe Anforderungen in geordnete Strukturen übersetzt und diese Strukturen übersichtlich dokumentiert. Die Architektur ist auch eine Abstraktion. Das Weglassen von Information ermöglicht das leichte Verständnis, die gefilterte Information macht eine spezifische Darstellung der Architektur lesbar.

Für den Entwurf einer Architektur müssen die Komponenten, Schnittstellen und Strukturen konstruiert werden. Komponenten können komplexe Subsysteme sein. Die Architektur bestimmt die Verantwortlichkeiten von Komponenten. Die Schnittstellen sind Kommunikationspfade zur Außenwelt. Sie legen einen Vertrag fest, der bestimmt, wie diese Kommunikation abläuft. Durch Komponenten und ihr Zusammenspiel entwerfen Architekten sowohl statische als auch dynamische Strukturen.

Zur Dokumentation der Architektur schlägt [Sta05] vier Sichten vor: Kontext-, Baustein-, Laufzeit- und Verteilungssicht. Die Kontextsicht zeigt den Zusammenhang des Systems mit seiner Umgebung aus einer Vogelperspektive.

Diese Sichten beantworten folgende Fragen:²⁷

²⁷ nach [Sta05] p. 87

Frage	Antwort
In welchem Kontext steht das System, welche Zusammenhänge gibt es zu Nachbarsystemen, Stakeholdern und der technischen Umgebung	Kontextsicht
Wie ist das System als Menge von Implementierungseinheiten strukturiert und welche Beziehungen bestehen zwischen denen?	Bausteinsicht
Wie verhalten sich die Bausteine des Systems als Laufzeitelemente und wie arbeiten sie zusammen?	Laufzeitsicht
Wie sieht die Verteilung des Systems aus (geografisch, auf Prozessoren, Prozesse, Abteilungen oder ähnliches), insbesondere auf seine technische Infrastruktur	Verteilungsschicht (Infrastruktursicht)

ALLGEMEIN

Am Beginn der Programmierung standen mehrere Optionen zur Umsetzung offen. Eine Idee war eine Stand-alone Lösung zu wählen und die komplette Visualisierung inklusive Renderer selbst zu programmieren. Weiter stand zur Diskussion, auf einen Renderer aufzusetzen und mit diesem die Darstellung zu machen. Bald zeigte sich, dass es viel Aufwand ist, Daten einzulesen, umzuformen, in das richtige Format zu bringen etc. Da ich für diese Dinge möglichst wenig Zeit vergeuden wollte, um mich auf die Visualisierung zu konzentrieren, entschloss ich mich nach einer Software zu suchen, die es ermöglicht, meine Visualisierung als Plugin einzubinden. Eine weitere Anforderung war, dass diese Software das Handling mit Daten und Datenbanken weitgehend übernehmen sollte.

So fiel die Wahl auf das Produkt Time Intelligence Solutions (TIS)²⁸ der Firma XIMES GmbH. Diese Software ist eine Web-basierte Lösung zur Arbeit mit zeitbezogenen Daten. Die Software ist eine mit C# programmierte ASP.NET Seite, der Datenabstraktionsslayer unterstützt die Datenbanken Oracle und Microsoft SQL Server. Ein Plugin-Konzept ermöglichte es mir, auf sehr einfache Art und Weise den Code in das System einzubinden.

²⁸ <http://www.ximes.com/de/software/products/tis/index.php>

Der besondere Vorteil dieser Software liegt im flexiblen Konzept mit Daten umzugehen. TIS ist in Datenknoten organisiert, auf die Operatoren angewendet werden. Daten können mit Hilfe von Operatoren aus verschiedenen Quellen (ODBC, Excel, CSV, XML, ...) eingelesen werden. Nach dem Einlesen können die Daten mit anderen Operatoren weiter bearbeitet werden, bis sie dann zur Visualisierung gelangen. Die Aufgabe bestand darin, einen Operator zu implementieren, der die Daten übernimmt und dann in der richtigen Form zur Anzeige bringt. Für diesen Operator definierte ich einige Anwendungsfälle („Stories“), um klarzustellen, was vom System erwartet wird.

VORGEHENSMODELL

Die Wahl des Vorgehensmodells bei der Softwareentwicklung ist am Beginn des Projekts eine wichtige Frage. Um ein systematisches, methodisches Vorgehen zu garantieren, war es mein Ziel, die Arbeit nach einem Vorgehensmodell zu strukturieren. Nach [Lud07] spricht man von einem Vorgehens- oder Projektmodell, wenn ausdrücklich vorgegeben ist, wie ein Projekt ablaufen soll. Bei den meisten Autoren werden beide Begriffe nicht unterschieden. Ludewig meint, dass das Wort „Prozess“ mehr Aspekte einschließt als das „Vorgehen“. Für ihn bildet das Vorgehensmodell den Kern des Projektmodells, das dann um Organisationsstrukturen, die Vorgaben für das Projektmanagement und die Qualitätssicherung, die Dokumentation und die Konfigurationsverwaltung ergänzt wird. Diese Unterscheidung macht durchaus Sinn, da sie die Einteilung und Klassifizierung der Modelle erleichtert und damit die Entscheidung für oder gegen eine Methode leichter macht.

Eine andere Unterteilung der Modelle ist die Unterscheidung in *leichte* und *schwere* Prozessmodelle, die aber alles andere als präzise ist.

„In Zusammenhang mit den so genannten agilen Prozessen ist die Unterscheidung zwischen leichten und schweren Prozessmodellen aufgekommen. Leider ist unklar, worauf sich diese Attribute beziehen. Ganz offensichtlich sind alle Prozessmodelle, die von ihren Verfassern als leicht eingestuft werden, durch einen geringen Aufwand für die Dokumentation gekennzeichnet. Das ist natürlich kein Wert an sich.“²⁹

Schwere Prozesse müssen auch nicht unbedingt unflexibel sein oder mehr Regeln beinhalten als leichte Prozesse. Für Ludewig macht es keinen Sinn, Prozesse in das Schema „leicht – schwer“ einzuordnen. Es hängt natürlich primär vom Projekt und der Einstellung des Kunden ab, ob ein Projekt viel Dokumentation und detaillierte Planung oder mehr Flexibilität braucht. Diese Entscheidung ist abhängig von der konkreten Situation zu treffen. Für Softwareprojekte, bei denen die Fähigkeit, rasch auf Anforderungen der Kunden oder des Marktes zu reagieren, im Vordergrund steht, entstand Ende der Neunzigerjahre die agile Bewegung, die einen Gegenpol zu bürokratischen Prozessen darstellt.

„Es entstand eine Reihe von „Anti-Prozessen“ (Extreme Programming, Adaptive Software Development, Crystal, SCRUM), die unter der Sammelbezeichnung „agile Prozesse“ bekannt wurden.“³⁰

Eine griffige Formulierung bieten [Abr02] im Abstract:

„Agile – denoting „the quality of being agile; readiness for motion; nimbleness, activity, dexterity in motion“ – software development methods are attempting to offer an answer to the eager business community asking for lighter weight along with faster and nimbler software development processes. This is especially the case with the rapidly growing and volatile Internet software industry as well as for the emerging mobile application environment.“

Auch hier wird wieder von einem „leichten“ Prozess gesprochen, ohne genauer zu spezifizieren, was damit gemeint ist. Oft wird leicht in einer Bedeutung wie „neu“, „innovativ“ oder Ähnlichem verwendet, während das Gegenteil „schwer“ eher Bedeutungen

²⁹ [Lud07] p. 173

³⁰ [Lud07] p. 10

wie „bürokratisch“, oder „veraltet“ annimmt. Diese Zuschreibungen sind nicht beweisbar und nicht ausreichend durch Studien gesichert. Damit ist diese Einteilung, wie oben erwähnt, sehr problematisch. Dazu kommt, dass in der Diskussion über Prozessmodelle die rein wissenschaftliche Auseinandersetzung oft ein wenig in den Hintergrund gerät:

„In der öffentlichen Diskussion ist das Thema der Prozesse (...) durch Heilslehren, Glaubenskriege und Lagerdenken geprägt. Das macht eine rationale Auseinandersetzung mit den Konzepten schwierig und eine Synthese nahezu unmöglich.“³¹

Die Softwareentwicklungs-Philosophie des Extreme Programming und das damit verbundene Prozessmodell wurden gewählt, da sie nach einer offenen Suche vielversprechend erschien.

EXTREME PROGRAMMING

Extreme Programming wurde von Kent Beck, Ward Cunningham und Ron Jeffries Ende der Neunziger entwickelt und erfreut sich seither großer Beliebtheit. Da es für XP keine Standardisierung oder Zertifizierung gibt und auch die Methode einer Veränderung unterliegt, stütze ich mich in meinen Aussagen auf [Bec04]. XP versucht viel mehr als andere Modelle, sich Veränderungen und Risiken zu stellen und aktiv auf sie zu reagieren. Eine Kernaussage ist, Veränderung zu leben:

„This is the paradigm for XP. Stay Aware. Adapt. Change.“³²

Die Methode versucht auch, soziale Belange zu berücksichtigen und bleibt damit nicht nur auf einer technischen Ebene.

XP besteht aus drei Teilen: Werte, Prinzipien und Praktiken. Prinzipien bilden die Brücke zwischen den Werten und Praktiken. Nach [Bec04] könnte man die Geschichte etwa so erzählen: Ein guter Freund, Paul, ist ein guter Gärtner. Er kennt gute Techniken für die Arbeit im Garten, die alleine aber noch nicht einen guten Gärtner ausmachen. Das sind die Praktiken, Dinge die man tatsächlich macht. Wenn sich Paul einen Garten ansieht, hat er eine gewisse Vorstellung, wie der Garten aussehen soll, was für das gesam-

³¹ ibid. p. 171

³² [Bec04] p. 11

te Umfeld gut oder schlecht ist und wo es notwendig ist, Dinge zu verändern. Diesen Bereich bildet der Begriff Werte ab. Prinzipien bauen die Brücke. Prinzipien sind anwendungsbezogene Richtlinien, das würde in der Gartenmetapher heißen, dass gute Gärtner wissen, welche Pflanzen zueinander passen, um sich gegenseitig zu ergänzen.

Für meine Interpretation der XP-Methode gab es eine starke Einschränkung. Ich musste die Arbeit natürlich prinzipiell selbst machen, da ich die Programmierarbeit für die Diplomarbeit alleine machte. Die Bezüge zu einem Team waren insofern gegeben, als das TIS-Entwicklerteam auch nach XP-Ansätzen arbeitet und ich in verschiedenen Phasen der Programmierung mit dem Team zusammen arbeitete.

WERTE

XP definiert folgende fünf Werte: *Kommunikation, Einfachheit, Feedback, Mut und Respekt*. Diese Werte bilden die Grundlage, um nach diesem Modell erfolgreich Software entwickeln zu können. Ein Programmierer muss sich zu den Werten bekennen. Kommunikation steht nicht zufällig ganz am Anfang, sie wird als Mittel zur Kooperation im Team betrachtet. Sehr oft kommt eine Information im Entwicklungsprozess einfach nicht bei der Person an, die eine Veränderung bewirken könnte. Ein Faustegel lautet, wenn ein Problem besteht, danach zu fragen ob es durch zu wenig Kommunikation verursacht wurde und was man dagegen unternehmen kann, um dies in Zukunft zu verhindern. Die Einfachheit bezieht sich vor allem auf die Frage „Was ist die simpelste Lösung, die noch funktionieren könnte?“. Eine technisch brillante Lösung hilft nicht, wenn sie im Team nur von einer Person verstanden wird. Einfachheit kann auch helfen, überflüssige Dokumentation zu verhindern. Feedback ist eine spezielle Form der Kommunikation, Feedback ist wichtig um sich in Schritten den Zielen anzunähern. Mut ist da, um offen Probleme anzusprechen und ehrliche Rückmeldungen zu geben. Mut ist die Grundlage für echtes, konkretes Feedback. Respekt im Sinne von XP heißt, dass jedes Teammitglied gleich viel wert ist, alle Beiträge sind gleich wichtig.

PRINZIPIEN

Die 14 Prinzipien in der Brückenfunktion sind folgende: *Menschlichkeit, Wirtschaftlichkeit, Beidseitiger Vorteil, Selbstgleichheit, Verbesserungen, Vielfältigkeit, Reflexion,*

Lauf, Gelegenheiten, Redundanzen, Fehlschläge, Qualität, Kleine Schritte und Akzeptierte Verantwortung.

PRAKTIKEN

Praktiken sind Dinge, die in XP-Teams jeden Tag passieren und je nach spezifischer Situation eingesetzt werden können. Es sind Wege, die zeigen, wie man an das Ziel kommt, XP zu leben. Beck teilt die Praktiken in zwei Bereiche: „primary practices“ und „corollary practices“. Erstere sind immer nützlich und anwendbar, zweitere sind nicht so leicht zu beherrschen und sollen erst eingesetzt werden, wenn die traditionellen Praktiken beherrscht werden. Ich möchte nur auf die „primary practices“ genauer eingehen, die für mich wichtig im Entwicklungsprozess waren und nach denen ich meine Arbeit ausgerichtet habe. Da die deutschen Übersetzungen den Sinn nicht so präzise wiedergeben können, verwende ich die englischen Originalbegriffe.

1. Sit Together
2. Whole Team
3. Informative Workspace
4. Energized Work
5. Pair Programming
6. Stories
7. Weekly Cycle
8. Quarterly Cycle
9. Slack
10. Ten-Minute-Build
11. Continuous Integration
12. Test-First Programming
13. Incremental Design

Gleich die erste XP-Praktik wurde von mir nicht immer eingehalten. Aus verschiedenen Gründen konnte ich nicht immer im gleichen Raum wie die Entwickler des TIS-Projekts sitzen. Für wichtige Entscheidungen, Integrations- und Testarbeit war ich natürlich immer wieder einige Tage vor Ort. Die zweite Praktik konnte ich beobachten und auch ich war für die Zeit meiner Entwicklungstätigkeit ins Team integriert und konnte so von

anderen lernen und der guten Arbeit profitieren. Die Praktik des „Pair Programming“ wurde pragmatisch interpretiert und nur in wichtigen Codeteilen praktiziert.

Alle anderen Praktiken wurden so eingesetzt, wie sie in der Literatur vorgeschlagen wurden.

Eine Beschreibung der Praktiken und eine allgemeine Einführung in Extreme Programming findet man bei [Wik085].

KRITIK AN XP

Mitte der 90 Jahre begann die IT-Branche über den bevorstehenden Wechsel von 1999 auf das Jahr 2000 nachzudenken. In vielen Unternehmen machte sich die Angst breit, dass alte Systeme aus den 60er und 70er Jahren nicht mehr funktionieren würden. Grund dafür war, dass für das Jahr im Speicher nur zwei Stellen reserviert waren, das bedeutet, man ging davon aus, dass die Programme die interne Uhr von 1999 auf 1900 stellen würde. Kein Programmierer dachte wohl, dass seine Software 30 oder 40 Jahre lang laufen würde. So wurde im Jahr 1995 bei DaimlerChrysler das Projekt C3 (Chrysler Comprehensive Compensation System) gestartet. Nach einem Jahr ohne große Fortschritte kam Kent Beck, der dann auch Ron Jeffries und Ward Cunningham in das Team brachte. In dieser Zeit entstand XP im Sinne einer „best practice“-Methode. Das C3 Projekt scheiterte jedoch und wurde im Jahr 2000 eingestellt. Die Gründe des Scheiterns wurden auf der C2 Wiki Seite „CThreeProjectTerminated“³³ intensiv diskutiert. Gegner der XP-Methode interpretieren das Scheitern im Sinne einer Unbrauchbarkeit des Entwicklungsmodells während Befürworter andere Gründe für den Abbruch des Projekt sehen. Wie auch immer, es lassen sich doch einige Kritikpunkte an diesem Entwicklungsmodell feststellen. Die einfache Formulierung der Anforderungen als Stories ist in großen Projekten einfach nicht ausreichend:

³³ <http://c2.com/cgi/wiki?CthreeProjectTerminated>

„Forty years of software engineering experience has taught us that major problems when you don't write the requirements down [...]. It is sheer lunacy to ignore this experience“³⁴

Ein weiterer Grund für das Scheitern von XP-Projekten kann beim Vertreter des Kunden liegen. Grob gesprochen ist der „on-site customer“ für die Formulierung der Anforderungen zuständig. Darüber hinaus soll er aber noch technisches Verständnis für die Programmierer haben, Test für die Stories formulieren und auch überprüfen, bei der Planung der Releases mitwirken, mit dem Management in Kontakt stehen und natürlich das Hintergrundwissen zu den Anforderungen haben. Dieser „on-site customer“ ist ein wenig der „Wunderwuzzi“ des Projekts und ist darüber hinaus ein single point of failure. Bei größeren Projekten wird ein Mensch kaum die Möglichkeit haben, all diese Tätigkeiten sinnvoll auszuüben. Für das Projekt ergeben sich folgende bedrohliche Potenziale: (nach [Ste03]) Der „on-site customer“ braucht nur krank zu werden und das Projekt steht, da nur er die genauen Anforderungen kennt und nichts oder zu wenig in schriftlicher Form vorliegt. Ähnlich gelagert ist auch der Fall, wenn sich der „on-site customer“ nicht mehr genau an seine Anforderungen erinnert, die in der user story nur sehr unscharf formuliert sind. Dies führt unweigerlich zu Konfusion. Nicht jeder Mensch kann sich immer an alles erinnern, es würde ausreichen dass der „on-site customer“ hin und wieder eine Sachlage anders formuliert, um Verwirrung zu stiften, die zu Chaos führen kann.

³⁴ [Ste03] p. 48

ANWENDUNGSFÄLLE/STORIES

Vor dem Beginn der Entwicklung legte ich die zu entwickelnde Funktionalität in Anwendungsfällen fest. Die Anwendungsfälle zeigen, was von dem Plug-In erwartet wird und welche Eingabedaten verarbeitet werden sollen. Die Formulierung im XP-Stil als Stories wollte ich nicht wählen, da sie sehr oberflächlich wirken. Die Anwendungsfälle legen im Sinn einer Spezifikation die Transformation der erwarteten Eingabedaten in die Ausgabedaten fest.

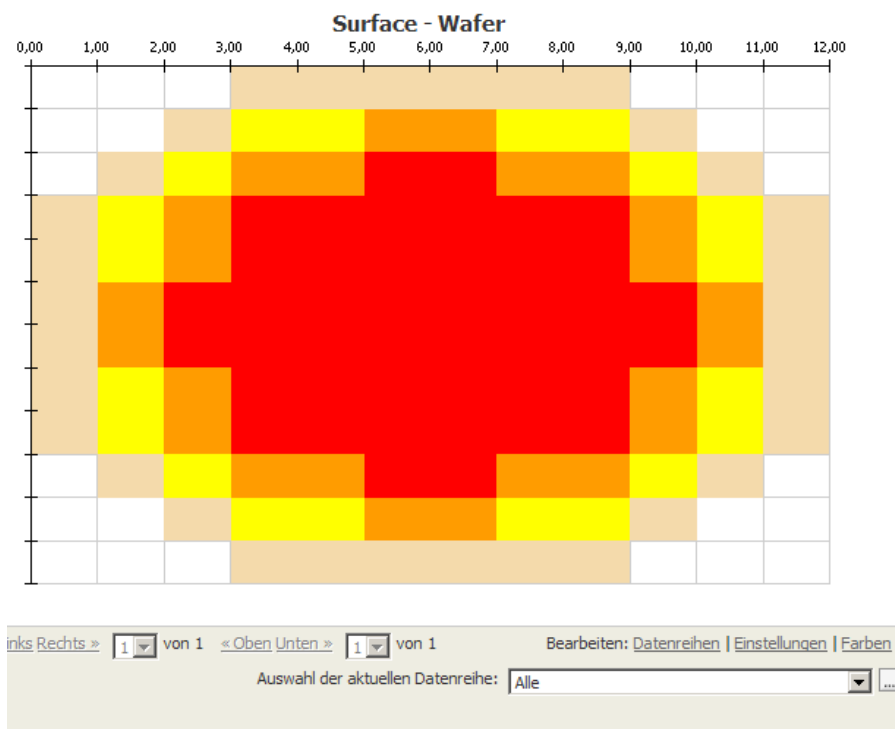


Abbildung 20 erster Entwurf der Visualisierung

Der Screenshot zeigt die Visualisierung als ersten Entwurf in einer frühen Designphase. Die Abbildung ist das Produkt des ersten Prototypen, der zeigte, wie ein Ergebnis aussehen konnte. Sehr deutlich sind hier die Rechtecke zu erkennen, die in diesem Beispiel sehr groß gewählt sind und in vier Farben eingefärbt sind. Die Skalierung der Felder kann bis zu einer Größe von einem Pixel gehen.

Die folgenden Fälle zeigen, welche Eingabedaten entgegengenommen werden können. Die Szenarien 1 und 2 unterscheiden, ob die Daten in Form einer Pivot-Tabelle oder einfache Liste entgegengenommen werden. In beiden Szenarien können die Daten je-

weils einer oder mehreren Datenreihen zugeordnet werden. Aus diesen Anwendungsfällen wurden nach der Test-First Praktik die Testfälle für die Entwicklung erstellt.

Die Zuordnung einer Datenreihe als Identifizierer ist die Benennung einer Datenreihe mit einem Namen. Ein Identifizierer muss nicht eindeutig sein, ist auf keinen Datentyp festgelegt und kann auch wie im ersten Beispiel eine Von-Bis Spalte sein. Diese Zuordnung ermöglicht, dass über Steuerelemente auf der Zeichenfläche gewisse Bereiche ausgeblendet werden können.

SZENARIO 1 - PIVOT

(a) Pivot Kategorie-Kategorie

Achse	Daten	Zusatz
X	Identifizierer (1 oder mehrere Spalten, werden zusammengefasst)	Kategorienachse
Y	Eingabe Pivot (1 oder mehrere Spalten, Spaltennamen sind die Beschriftung der Kategorien)	Kategorienachse Beispiel 1.1.2010-2.1.2010-3.1.2010
Farbe	Intervalle	Parameter (Werte ergeben sich aus Pivot-Tabelle)


Uhrzeit	01.01.2010	02.01.2010	03.01.2010
00:00-01:00	5	6	7
01:00-02:00	4	6	3

Abbildung 21 Eingabedaten Pivot

In diesem Anwendungsfall liegen Daten in Form einer Pivot-Tabelle vor. Die Tabelle wird aufgelöst, die Werte der ersten Spalte werden entlang der X-Achse aufgetragen, die Werte der ersten Zeile kommen auf die Y-Achse. Die Abstände zwischen den Y-Werten ist immer 1, die Daten haben keinen Zusammenhang und können entlang dieser

Achsen nicht skaliert werden, das heißt die Abstände zwischen den Werten werden nie verändert. In den Einstellungen kann man nur anpassen, wie viele Werte pro Seite gezeichnet werden beziehungsweise ob nur jeder n-te Werte dargestellt wird.

(b) Pivot Kategorie-Kategorie

Achse	Daten	Zusatz
X	Identifizierer	Kategorienachse
---	Identifizierer	Dropdown für Teile
		
<p>Abbildung 22 Auswahlelement für einen zusätzlichen Identifizierer</p>		
Y	Eingabe Pivot (1 oder mehrere Spalten, Spaltennamen sind die Beschriftung der Kategorien)	Kategorienachse Beispiel 1.1.2010-2.1.2010-3.1.2010
Farbe	Intervalle	Parameter (Werte ergeben sich aus Pivot-Tabelle)

In diesem Anwendungsfall können die Datenpunkte einem zweiten Identifizierer zugeordnet werden, der in einem Dropdown-Feld unter dem Zeichenbereich dargestellt wird. So können die Daten einer zweiten Kategorie zugeordnet werden, die dynamisch bei der Anzeige ein- oder ausgeblendet werden kann. Mögliche Beispiele dafür sind eine Abteilung, Produktgruppe oder ein Land.

SZENARIO 2 - DATENLISTEN

(a)

Achse	Daten	Zusatz
X	Identifizierer	Kategorie
Y	von	Von-Grenze Y Achse
Y	bis	Bis-Grenze Y Achse
Farbe	Intervalle	Parameter

Filiale	von	bis	Frequenz
A	07:00	09:00	10000
A	09:00	10:00	12000
A	10:00	11:00	14000
B	08:00	09:00	900
B	09:00	10:00	400
B	10:00	12:00	300
C	08:00	09:00	7000
C	09:00	10:00	460
C	10:00	11:00	1000

Abbildung 23 Eingabedaten Liste

Im Szenario 2 werden Datenlisten als Eingabedaten entgegengenommen. Eine Spalte dient als Identifizierer („Kategorienachse“) an der X-Achse, zwei weitere Spalten enthalten Zeitangaben, die als Von-Bis-Angaben interpretiert werden. Diese Datumswerte werden an der Y-Achse aufgetragen und können später beliebig skaliert werden. Eine dritte Spalte enthält die Werte für die Farbkodierung.

(b)

Achse	Daten	Zusatz
X	Identifizierer	Kategorie
Y	Identifizierer	Kategorie
Z	Werte, Abstände Y Achse	Spalte für Farbinformation
Farbe	Intervalle	Parameter

Filiale	KW	Wert
a	1	23
a	2	15
b	1	66
b	3	44

Abbildung 24 Eingabedaten Liste mit zwei Identifizierern

Wie bei Szenario 1 Fall (b) wird auch hier ein weitere Spalte angegeben, um die Zuordnung der Datenpunkte zu einem zweiten Identifizierer zu ermöglichen, der in der Visualisierung dynamisch selektiert werden kann.

MAPPING DER DATEN

Das Mapping der Daten in die Visualisierung lässt sich am Besten anhand eines Beispiels zeigen.

BEISPIEL

Die Abbildungen unten zeigen eine Anwendung des Operators zur Analyse von Umsätzen eines Handelsbetriebs. Dabei werden nicht die Umsätze, sondern die Abweichungen zum Durchschnitt gezeigt.

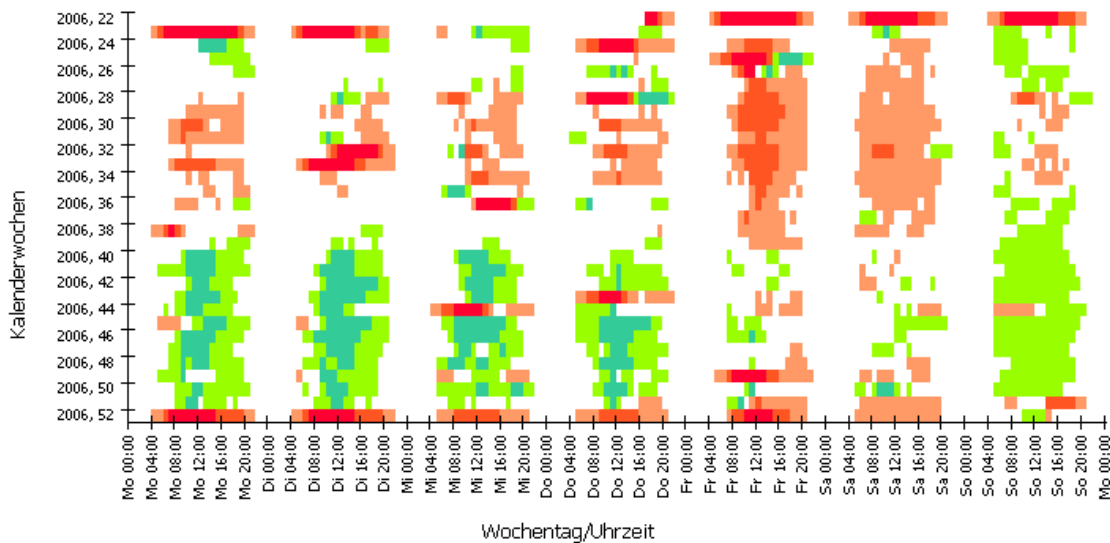


Abbildung 25 Visualisierung Landkarte der Abweichungen

Die vertikale Achse zeigt den Verlauf eines Halbjahres skaliert nach Kalenderwochen. Entlang der horizontalen Achse ist der Verlauf einer Woche von Montag bis Sonntag zu sehen. Die Skalierung erfolgt in Einheiten von vier Stunden. Pro Tag sind so sechs Markierungen zu sehen, was dann insgesamt 42 Intervalle pro Seite ergibt. Auf der Zeichenfläche wurden die Umsatzdaten farblich kodiert. Werte, die negativ zum Durchschnitt sind, werden in Rottönen dargestellt, während Werte, die positiv zum Durchschnitt abweichen grün und blau markiert sind.

So werden Muster, Strukturen und Zusammenhänge klar ersichtlich und können gut erfasst werden.

Das vorliegende Beispiel könnte so interpretiert werden, dass es im Wochenverlauf zwischen den Kalenderwochen 24 und 38 starke negative Abweichungen zum Durchschnitt gibt. Besonders stark sind die Abweichungen an Freitagen zwischen 9 und 15 Uhr. Etwa ab Woche 40 sind an montags, dienstags, mittwochs und donnerstags starke positive Abweichungen, besonders zur Mittagszeit zu sehen. Man könnte hier Schulferien vermuten, die in der Woche 39 enden. Die letzte Kalenderwoche im Jahresverlauf weist wieder starke negative Abweichungen auf.

Hier zeigt sich, wie einfach sich ein Betrachter einen Überblick über den Datenbestand eines Halbjahres verschaffen kann.

Die Benutzeroberfläche erlaubt die Zuordnung der Werte zu einer Identifizierervariablen. So können etwa die Daten verschiedener Abteilungen, Länder, Regionen etc. direkt miteinander verglichen werden. Für n -verschiedene Ausprägungen dieser Variable werden n Subsets gezeichnet, die untereinander gestellt werden.

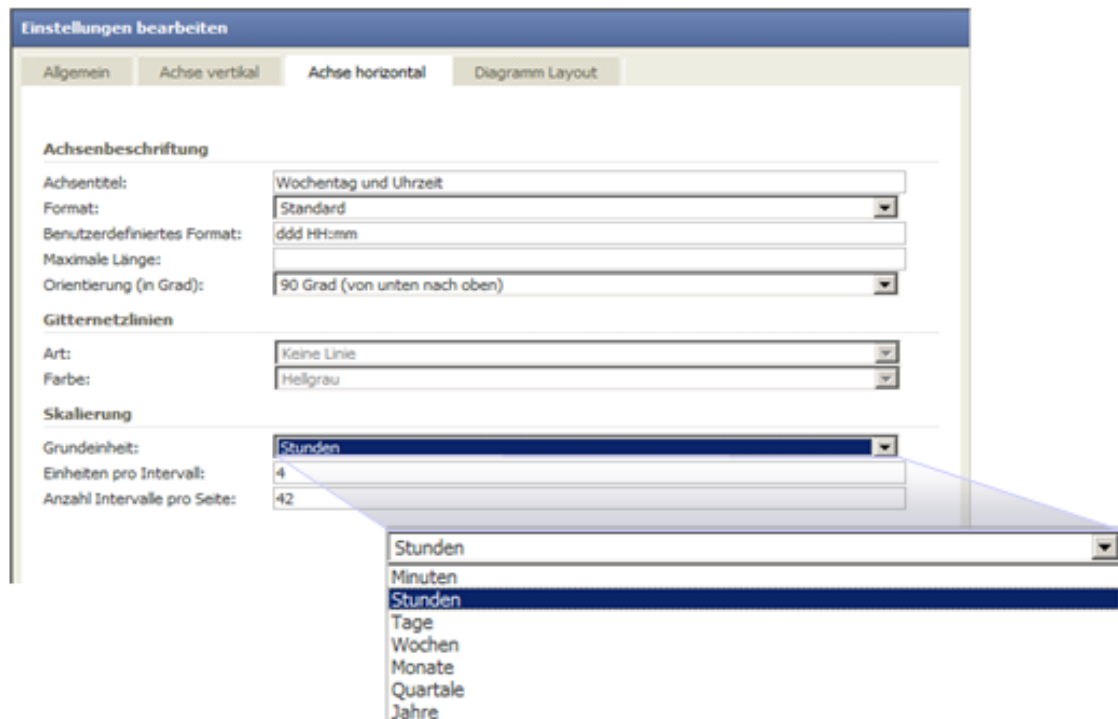


Abbildung 26 Einstellungsdialog Visualisierung

Der Einstellungsdialog zur Visualisierung ermöglicht die Bearbeitung des Achsentitel, eine Drehung der Intervallbeschriftungen und eine beliebige benutzerdefinierte Formattierung der Intervallbeschriftungen.

Die Skalierung bietet viele Einheiten der Zeit zur Auswahl (Minuten, Stunden, Tage, Wochen, Monate, Quartale und Jahre).

Ein Abstand zwischen den Datenpunkten ist nicht vorgesehen, sie werden kontinuierlich aneinander gezeichnet. In der höchsten Auflösung hat ein Datenwert die Größe eines Pixels am Bildschirm.

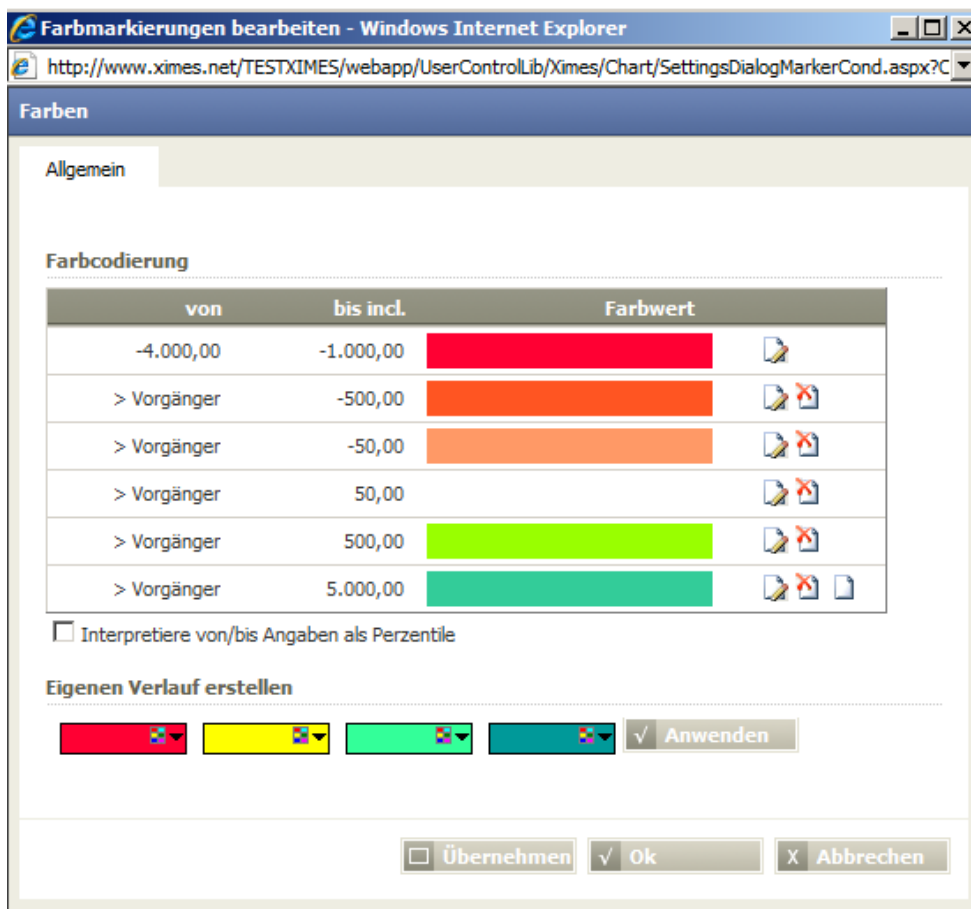


Abbildung 27 Einstellungsdialog Farbmarkierungen

Der Einstellungsdialog zu den Farbmarkierungen legt Farbskalen und Verläufe fest. Hier wird die Zahl der zu unterscheidenden Farbbereiche festgelegt. An dieser Stelle kann auch eingestellt werden, ob die Bereichsgrenzen absolut oder nach Perzentilen ausgewertet werden.

In diesem Beispiel werden Datenwerte von -4.000 bis -1.000 mit einem dunklen Rot eingefärbt, Werte von 50 bis 500 mit Hellgrün. Hier wurden sechs verschiedene Farbbereiche gewählt. Der Benutzer kann hier beliebig viele Grenzen festlegen und mit jeweils verschiedenen Farben versehen.

Die Funktion „Eigenen Verlauf erstellen“ eignet sich besonders zur Festlegung der Farben bei besonders vielen Farbbereichen. Hier werden vier verschiedene Farben ausgewählt, die dann gleichmäßig verlaufend über die Farbmarkierungen verteilt werden.

Wenn die Grenzen nicht so wie in diesem Beispiel mit absoluten Zahlen markiert werden, sondern das Häkchen „Interpretiere von/bis Angaben als Perzentile“ gesetzt wird, können bei den Eingabefeldern „von“ und „bis incl.“ Zahlen zwischen 0 und 100 eingegeben werden. Eine Eingabe von 90-100 und die Auswahl der Farbe Rot würde bedeuten, dass die oberen 10% der Werte mit der Farbe Rot markiert werden.

FARBGEBUNG

Neben dem Mapping der Daten spielt die Farbgebung der Datenpunkte eine große Rolle bei der erfolgreichen Umsetzung. Der Benutzer soll im User-Interface alle Möglichkeiten zur Auswahl und Festlegung seiner Wünsche vorfinden, dabei sollen aber Vorlagen helfen, eine rasche und gute Vorauswahl zu treffen.

GRUNDLAGEN

In unserer Wahrnehmung werden Farben durch Rezeptoren auf der Netzhaut wahrgenommen. Der Mensch hat drei verschiedene Arten dieser Zapfen, die auf bestimmte Wellenbereiche des Lichtes reagieren. Die Unterteilung wird in S (Small)-, M (Medium)- und L (Large)-Zapfen vorgenommen, die jeweils einen bestimmte Wellenlänge des Lichtes aufnehmen können. Diese drei Bereiche sind blau, grün und rot. Aus diesen Grundlagen ergibt sich die dreidimensionale Aufteilung des Farbraumes.

Ungefähr zehn Prozent der männlichen und circa ein Prozent der weiblichen Bevölkerung haben Sehschwächen. Die meisten Sehschwächen lassen sich durch einen Mangel der M- oder L-Zapfen erklären. Dies bedeutet in der Praxis, dass diese Personen Rot und Grün nicht oder nicht sehr gut unterscheiden können. Der dreidimensionale Farbraum wird so zu einem zweidimensionalen.³⁵

FARBEN IN DER VISUALISIERUNG

ALLGEMEIN

[Few08] gibt in einem anwendungsorientierten Artikel eine sehr gute Übersicht über die Verwendung von Farben für Visualisierungen. Er formuliert neun Regeln, an denen ich mich auch orientieren möchte.

³⁵ nach [War04] p. 99

FARBE IM KONTEXT

Ein weit bekannter Effekt ist, dass eine Farbe je nach Hintergrund verschiedenartig auf den Betrachter wirkt. Ein heller oder dunkler Hintergrund verändert die Wahrnehmung der Vordergrundfarbe, besonders wenn der Hintergrund nicht gleichmäßig mit einer Farbe ausgefüllt wird. Daraus ergeben sich die ersten beiden Regeln:

Regel 1: Wenn verschiedene Objekte mit der gleichen Farbe in einer Tabelle oder einem Diagramm gleich wirken sollen, muss der umgebende Hintergrund konsistent sein.

Für meine Visualisierung bedeutet dies, dass der Hintergrund weiß oder mit einem hellen Grau eingefärbt sein muss. Wenn ein Benutzer Farbmarkierungen für Daten auswählt, die auch weiß sind, wird die Unterscheidung der Datenpunkte unmöglich.

Regel 2: Wenn Objekte in einer Tabelle oder einem Diagramm deutlich erkennbar sein sollen, muss eine Hintergrundfarbe mit ausreichendem Kontrast gewählt werden.

Diese Regel ist besonders wichtig bei der Repräsentation von Daten in Form der Landkarte oder allgemein bei pixelbasierten Visualisierungen. Gerne werden die Ampelfarben zur Kodierung der Farbinformation herangezogen. Grün signalisiert positive Trends, Wachstum oder Ähnliches, Gelb signalisiert einen neutralen Verlauf und Rot wird mit Verlusten oder negativen Entwicklungen gleichgesetzt.

Hier können bereits erste Schwierigkeiten in der Darstellung auftreten. Bleibt man bei der Konvention Grün-Gelb-Rot und hinterlegt man in den einzelnen Zellen der Darstellung noch Informationen in Textform wie zum Beispiel den konkreten Umsatz als Zahl, so ergeben sich in den Zellen von besonderem Interesse, nämlich Dunkelrot und Dunkelgrün, Schwierigkeiten bei der Lesbarkeit.

Als Lösung empfiehlt es sich, die Textinformation einfach wegzulassen. Der Grund liegt nicht nur in der Schwierigkeit mit den Kontrasten, sondern auch darin, dass man beim Auffinden interessanter Zellen durch das ständige Lesen der Zahlen stark abgelenkt wird. In der vorliegenden Implementierung wird der Wert als Tooltip angezeigt, das heißt man muss mit der Maus eine kurze Zeit über einem Datenpunkt verweilen, um herauszufinden welcher Zahlenwert sich dahinter verbirgt.

EINSATZ VON FARBE MIT ZURÜCKHALTUNG

Farbe sollte immer möglichst zurückhaltend eingesetzt werden und nur dort, wo die Farbe eine tatsächliche Bedeutung hat und den Einsatz rechtfertigt.

Regel 3: Benütze Farbe nur dort, wo sie zur Kommunikation eines spezifischen Zweckes oder Zieles dient.

Diese Aussage meint, dass Farbe nicht zur Ausschmückung oder Zierde dienen soll, sondern nur dann verwendet werden soll, wenn es der konkreten Aussage dient. Farbe sollte den Betrachter niemals vom eigentlichen Inhalt, den Daten, ablenken.

Regel 4: Verwende verschiedene Farben nur, wenn sie zu verschiedenen Bedeutungen korrespondieren.

Ein Beispiel ist, bei Säulendiagrammen, die schon auf der X-Achse beschriftet sind, noch einmal jede Säule mit einer anderen Farbe zu versehen. Die Information, wodurch sich diese Säulen unterscheiden, ist schon durch die Beschriftung gegeben. Die Farbe hat also keine zusätzliche Bedeutung oder einen Wert. Dies trägt eher zur Verwirrung bei, da der Betrachter meinen könnte, das dies so wäre und vielleicht viel Zeit verschwendet, darüber nachzudenken, was ihm die Farbe wohl vermitteln möchte.

DEFINIERE FARBPALETTEN FÜR STANDARDAUFGABEN

So wie Blumen leuchtende Farben einsetzen um Bienen zur Bestäubung anzulocken, sollten wir Farben einsetzen, um Aufmerksamkeit auf gewisse Daten zu ziehen. Dies kann vor allem in drei Bereichen passieren:

- Hervorhebung bestimmter Daten
- Gruppierung von Elementen
- Kodierung quantitativer Werte

Regel 5: Verwende weiche, natürliche Farben zur Darstellung der meisten Daten und leuchtende oder dunkle Farben zur Hervorhebung von Informationen, die weiterer Aufmerksamkeit bedürfen.

Anstatt sich jedes mal wieder den Kopf über einer großen Liste verschiedener Farben zu zerbrechen, ist es von Vorteil, sich einen Satz standardisierter, guter Farben zurechtzu-

legen. Ein Vorschlag ist eine Palette mit kraftvollen, dunklen Farben, eine mit mittleren Schattierungen und eine mit hellen, schwachen Farben. Diese Paletten können dann je nach Zweck zum Einsatz gebracht werden. Die erste Palette mit den kräftigen Farben eignet sich sehr gut zur Hervorhebung von Daten in Diagrammen. Wenn die dargestellten Objekte besonders klein oder dünn sind, muss die Farbe oft intensiver oder dunkler gewählt werden. Ein anderer Weg ist, die kleinen Objekte schlicht größer oder dicker zu machen. Die Palette mit den hellen, schwachen Farben sollte für Teile herangezogen werden, die keine wichtige Rolle spielen, wie die Achsen oder Gitterlinien. Wenn man Farben verwendet, um Daten zu gruppieren, sollten sich die Farben klar unterscheiden, um auch die Gruppen deutlich voneinander abzugrenzen. Die Farben sollen sich hinsichtlich des Farbtons unterscheiden, aber die gleiche Intensität haben.

Regel 6: Wenn man Daten quantitativer Natur mit Farbinformation versehen will, sollte man bei einem Farbton bleiben und diesen nur in der Intensität variieren. Helle Farben für niedrige Werte, dunkle Farben für hohe Werte.

Der Begründung dieser Regel liegt darin, dass man Farben, die nur über die Intensität abgestuft sind, intuitiv ordnen kann, während das bei verschiedenen Farbtönen nicht möglich ist.

DIE BESTE FARBE FÜR JEDE KOMPONENTE

Tabellen und Diagramme bestehen aus Bereichen, die die tatsächlichen Daten enthalten und anderen Bereichen, die eine unterstützende Funktion einnehmen. Der unterstützende Bereich meint Text und Beschriftungen von Achsen, Legenden oder Markierungen auf den Achsen. In Diagrammen sind die typischen Vertreter dieses Bereichs die Achsen, Hintergrundfarbe(n) und Grenzlinien um die Legende oder den Zeichenbereich.

Regel 7: Bereiche, die keine Daten enthalten, sollten nur gerade so weit sichtbar sein, als dass sie ihre Funktion erfüllen, sie dürfen aber nicht von den Daten ablenken.

Hier sind einige Vorschläge angeführt:

Komponente	Farbe
------------	-------

Achsen	Graue Linien mittlerer Intensität
Grenzlinien	Meist gar nicht benötigt, sonst graue Linien mittlerer Intensität
Hintergrund	Weiß
Säulen, Balken	Deutlicher Farbton mittlerer Intensität für jede Datenserie
Linien	Dünne Linien: deutlicher Farbton mit hoher Intensität für jede Linie Dicke Linien: deutlicher Farbton mittlerer Intensität
Datenpunkte	Kleine Datenpunkte: deutlicher Farbton mit hoher Intensität Große Datenpunkte: deutlicher Farbton mittlerer Intensität

Regel 8: Rot-Grün Kombinationen in Visualisierungen sind zu vermeiden, um Personen mit Sehschwächen nicht zu benachteiligen.

Ein sehr guter Ersatz in einer heatmap für rot-grün ist rot-blau.

Viele Programme, mit denen man Diagramme erstellen kann, bieten eine Vielzahl grafischer Effekte wie zum Beispiel eine Art dreidimensionaler Darstellung. Diese Effekte haben keinen Wert an sich und tragen höchstens zur Verwirrung bei.

Regel 9: Vermeide visuelle Effekte in Diagrammen

FARBEN FÜR KARTEN

Die von mir vorgestellte Visualisierung funktioniert nach einer Karten-Metapher. Nichts liegt also näher, als diese Metapher in der Farbgebung fortzusetzen. Das Flash-Programm ColorBrewer von [Bre032] zeigt einen verblüffend einfachen und innovativen Weg, Farben für Karten auszusuchen.

„We set out to make the tool simple to use and specific in its function. We sought to guide novice map makers by offering them varied color scheme options categorized using a few basic cartographic design principles that they should consider in order to make effective choices for representing their data.“³⁶

Diese Applikation wendet sich an Ersteller von Karten, die nicht unbedingt aus dem Fachbereich der Kartographie kommen müssen – und genau das ist der Benutzer der vorgestellten Visualisierung. Es braucht nicht weiter erläutert zu werden, dass die wenigsten Benutzer erfahrene Hersteller von Karten sind und noch dazu intuitiv gute Farben auswählen. Wie die Benutzerstudie zeigt, waren viele Benutzer beim Einstellen und Auswählen der Farben überfordert. Auch diese Feststellung wird von den Autoren des ColorBrewer geteilt.

„ [...], I am hearing from GIS analysts and developers, remote sensing consultants, computer programmers, statisticians, [...]. They are making maps and visualizing data in contexts where they do not have contact with cartographers. They express their frustration about their own ability to design color schemes and to convince clients to use effective schemes.“³⁷

³⁶ [Bre031] p. 159

³⁷ [Bre031] p. 161

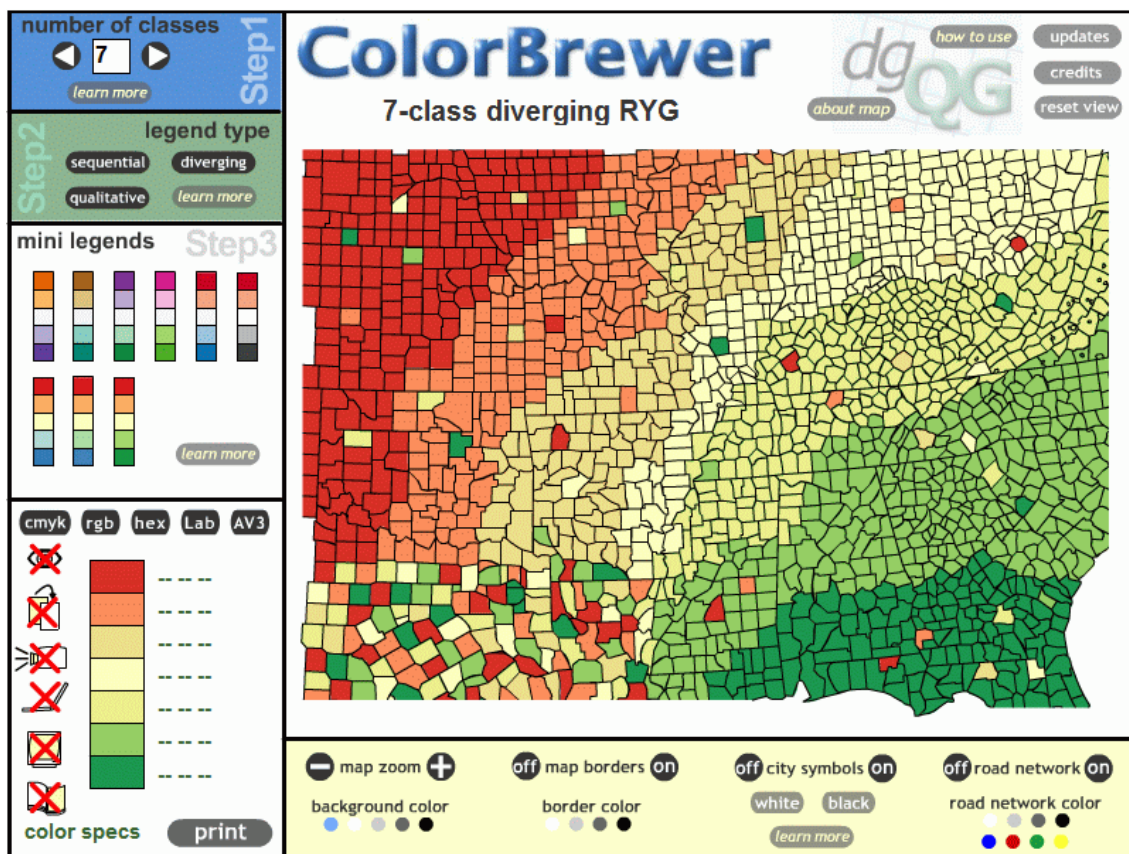


Abbildung 28 Benutzeroberfläche ColorBrewer

Die Flash-Oberfläche besteht aus drei Teilen. Zuerst legt man die Anzahl an verschiedenen Klassen fest, für die man Farben zur Auswahl bekommen möchte. Nachdem man den Typ der Farbschemen gesetzt hat, bekommt man verschiedene Farbschemen zur Auswahl, für die auch die Farbdefinitionen in RGB, CMYK und Hex vorliegen. Alle Farben wurden so ausgewählt, dass sie sowohl am Bildschirm als auch am Ausdruck gut lesbar sind.

[Bre03] unterscheidet drei Typen von Farbschemen:

- sequential
- diverging
- qualitative

Die folgenden Abbildungen zeigen Beispiele der Farbschemen für sieben Farbklassen.

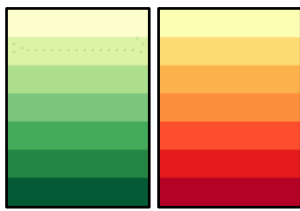


Abbildung 29 Sequential

Eine Übersetzung für den ersten Typ „Sequential“ könnte sequentiell oder verlaufend bzw. Verlauf sein. Diese Farbpalette eignet sich für Daten, die sich von einem Minimum zu einem Maximum erstrecken. Helle Farben werden für niedrige Werte genommen und dunkle, kräftige Töne für hohe Werte.

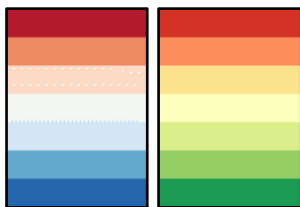


Abbildung 30 Diverging

Das Diverging-Farbschema eignet sich für Daten, die um einen Referenz- oder Nullpunkt angesammelt sind und einen negativen sowie positiven Bereich haben. Die Farbe in der Mitte ist hell und verläuft in zwei Richtungen zu zwei verschiedenen Farbtönen.

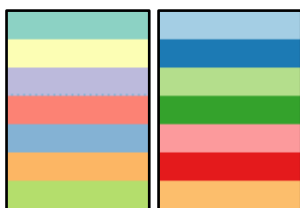


Abbildung 31 Qualitative

Das „Qualitative“-Farbschema wird verwendet, um Klassen zu kennzeichnen, die keinen Wertunterschied aufweisen. Dieses Farbschema eignet sich sehr gut für Daten, die nominal skaliert sind oder Kategorien darstellen.

Die in [Bre03] vorgestellten Farbschemen und die Benutzeroberfläche des ColorBrewer zeigen eine sehr gute Methode zur Auswahl passender Farben, die es auch nicht-professionellen Benutzern ermöglicht, richtige Farben zu wählen. Der Auswahlbereich

aus „Step 2“ und „Step 3“ wird in den Dialog zur Farbauswahl der vorgeschlagenen Visualisierung integriert.

Kapitel 3 BENUTZERSTUDIE

AUSGANGSSITUATION UND FRAGESTELLUNG

Im Laufe der Entwicklung der Visualisierung gab es viele Entscheidungen bezüglich der Benutzeroberfläche und des Bedienkonzeptes zu treffen. Das Ergebnis als vorliegendes Produkt sollte von einer Öffentlichkeit auf die Tauglichkeit im Einsatz überprüft werden.

Am Ende des Entwicklungsprozesses stand neben eigenen Tests eine Benutzerstudie. Ziel der Studie war es, die Benutzbarkeit der Bedienoberfläche zu verbessern und die getroffenen Designentscheidungen durch Experten kritisch hinterfragen zu lassen und neue Ideen für die Benutzerschnittstelle zu sammeln.

Die Benutzerstudie wurde im Aufbau, in der Auswertung und Analyse nach den unten dargestellten, aus den Sozialwissenschaften stammenden Methoden durchgeführt, die im folgenden Abschnitt erläutert werden.

METHODISCHES DESIGN

Anhand der verschiedenen Komponenten der Benutzeroberfläche wurde ein Leitfaden für die Befragung erstellt. Der Leitfaden besteht aus vier Anwendungsfällen, die die Benutzer im Rahmen des Interviews bearbeiteten. Das Experten-Interview (vgl. [Fli02] 139 ss.) erwies sich dafür als vorteilhaft. Der Experte steht als Repräsentant für ein bestimmtes Handlungsfeld im Mittelpunkt, er wird als Repräsentant einer Gruppe in die Untersuchung einbezogen. Dies schränkt die Bandbreite an potentiell relevanter Information, die der Befragte liefern soll, ein. Deshalb kommt hier dem Leitfaden eine besondere Bedeutung zu, da er die Steuerungsfunktion im Hinblick auf das Thema übernimmt. Es muss gelingen, den Experten im Gespräch auf das interessierende Expertentum zu begrenzen. Die Steuerung des Interviews ist von zentraler Bedeutung, da weniger die Person, sondern vielmehr eine gewisse Eigenschaft der Person von Interesse ist. Die Anwendungsfälle zwängten den Befragten in ein Korsett und übernahmen die Steuerungsfunktion.

Die Auswertung und Interpretation der Daten wird nach der von Flick und Mayring beschriebenen qualitativen Inhaltsanalyse durchgeführt (vgl. [Fli02] 279 ss. und [May02] 114 ss.). Ziel ist die Reduktion der vorliegenden Aufzeichnungen durch Zusammenfassung und Kategorisierung, die die Grundlage für die Theorieentwicklung bilden.

„Die Stärke der Inhaltsanalyse ist, dass sie streng methodisch kontrolliert das Material schrittweise analysiert. Sie zerlegt ihr Material in Einheiten, die sie nacheinander bearbeitet. Im Zentrum steht dabei ein theoriegeleitet am Material entwickeltes Kategoriensystem; durch dieses Kategoriensystem werden diejenigen Aspekte festgelegt, die aus dem Material herausgefiltert werden sollen.“³⁸

Die Inhaltsanalyse ist eine Vorgehensweise zur Analyse von Textmaterial verschiedener Herkunft. Wesentlich dafür ist die Verwendung von Kategorien, die an das Material herangetragen und auch aus dem Material heraus entwickelt werden können.

1. Festlegung des Materials, Auswahl der für die Fragestellung interessierenden Teile
2. Analyse der Erhebungssituation
3. formale Charakterisierung
4. Fragestellung der Analyse muss vorher geklärt sein, theoretisch an die Forschung angebunden, differenziert in Unterfragestellungen

Techniken der Inhaltsanalyse

- zusammenfassend
- explizierend
- strukturierend

³⁸ [May02] p. 114

„In der zusammenfassenden Inhaltsanalyse wird das Material paraphrasiert, wobei weniger relevante Passagen und bedeutungsgleiche Paraphrasen gestrichen (erste Reduktion) und ähnliche Paraphrasen gebündelt und zusammengefasst werden (zweite Reduktion). Dies stellt eine Kombination der Reduktion des Materials durch Streichungen mit einer Generalisierung im Sinne der Zusammenfassung auf einem höheren Abstraktionsniveau dar.“³⁹

[May02] spricht ähnlich wie [Fli02] von drei Grundformen qualitativer Inhaltsanalyse:⁴⁰

- Zusammenfassung
Ziel der Analyse ist es, das Material so zu reduzieren, dass die wesentlichen Inhalte erhalten bleiben, durch Abstraktion ein überschaubares Korpus zu schaffen, das immer noch ein Abbild des Grundmaterials ist.
- Explikation
Ziel der Analyse ist es, zu einzeln fraglichen Textteilen (Begriffen, Sätzen, ...) zusätzliches Material heranzutragen, das das Verständnis erweitert, das die Textstelle erläutert, erklärt, ausdeutet.
- Strukturierung
Ziel der Analyse ist es, bestimmte Aspekte aus dem Material herauszufiltern, unter vorher festgelegten Ordnungskriterien einen Querschnitt durch das Material zu legen oder das Material auf Grund bestimmter Kriterien einzuschätzen.

In der Auswertung meiner aufgezeichneten Interviews verzichte ich auf eine Transkription, das Material wird direkt aus der Aufzeichnung heraus paraphrasiert, gebündelt und soweit als möglich an die Kategorien angenähert. In der abschließenden Analyse werden die Kategorien neu geordnet, überarbeitet und mit Beispielen hinterlegt. Anschließend werden Vorschläge und Konzepte zur Überarbeitung der Visualisierung und des Bedienkonzepts erarbeitet.

VORBEREITUNG „ANWENDUNGSFÄLLE“

Vor der Durchführung der Gespräche wurden der Inhalt, Ablauf und der Zeitrahmen festgesetzt.

³⁹ [Fli02] p. 280

⁴⁰ nach [May02] p. 115

Alle Gespräche wurden als Video mit Bild und Ton aufgezeichnet. Das Bild lieferte der Bildschirm, auf dem gearbeitet wurde, der Ton ist der Mitschnitt des Gespräches. Ein schriftliches Protokoll wurde nicht geführt, da dadurch die Konzentration im Interview gestört werden würde.

Die Struktur wird durch die Anwendungsfälle vorgegeben, muss aber nicht sklavisch eingehalten werden. Die Leitfragen werden während des Interviews eingebaut sofern dies möglich ist. Fragen, die nicht in diesem Zug beantwortet werden, werden am Ende des Gesprächs gestellt.

AUSWAHL DER TEILNEHMER

Für die Studie wurden Experten aus dem Bereich Arbeitszeit und Visualisierung ausgewählt. Ich versendete Emails an acht Personen mit dem Hinweis, dass ich sie in den nächsten Tagen kontaktieren würde. So konnte ich acht Experten aus Aachen, Erlangen, Hamburg, Nürnberg und Wien gewinnen, die Interviews gemeinsam mit mir durchzuführen.

SETTING UND KOMMUNIKATION

Der Gesprächspartner und ich sitzen gemeinsam am Computer oder sind über eine Internet-Konferenz verbunden. Grundbedingung war für mich immer, dass ich das Gespräch und die Aktionen am Bildschirm aufzeichnen konnte. Dies war in den meisten Fällen einfach, da die Interviewpartner direkt neben mir sitzen konnte, für räumlich weit entfernte Experten musste ich eine Lösung finden, die über das Internet funktioniert.

ACROBAT.COM

Die Firma Adobe betreibt unter der Webadresse Acrobat.com eine Plattform zur webbasierten Gruppenarbeit. Eine Applikation ist der Dienst ConnectNow, mit dem man auf sehr einfache Art und Weise seinen Bildschirm freigeben und sogar seinen Computer fernsteuern lassen kann. Die Applikation wird aus dem Browser gestartet und benötigt nur die Softwaren Flash 9, die auf den meisten Rechnern ohnehin installiert ist. Ich nutzte den Dienst, um meinen Desktop zur Fernsteuerung durch den Interviewpartner freizugeben, da ich so das Interview auch aufzeichnen konnte. Der Dienst ConnectNow

war eine gut funktionierende technische Basis für die Interviews, die ich über das Internet durchführte.

SKYPE

Die Software Skype ist eine mittlerweile weit bekannte und viel genutzte Voice-over-IP (VOIP) Software, die es ermöglicht, Telefonate direkt übers Internet zu führen. Das Program ist für alle gängigen Betriebssysteme erhältlich. Es nutzt ein proprietäres Protokoll zur Kommunikation, welches sehr gut quer durch Router und Firewalls funktioniert. Aus diesem Grund werden oft auch Sicherheitsbedenken gegen den Einsatz von Skype geäußert. In der Praxis überwiegt aber der Vorteil der einfachen Nutzung. Das Programm benötigt keine physische Telefonleitung, lediglich eine Internetverbindung ist von Nöten. Über das Service Skype Out können auch Teilnehmer im klassischen Festnetz oder Mobilfunknutzer erreicht werden. Diesen Dienst verwendete ich, um mit meinen Partnern vom Computer aus sprechen zu können. Gleichzeitig war es mir so möglich, das Gesagte aufzuzeichnen.

CAMTASIA

Das Programm Camtasia der Firma TechSmith ermöglicht komfortable Bildschirmaufnahmen. Die aufgenommenen Inhalte lassen sich auf sehr einfache Art und Weise für die Weitergabe in verschiedenen Formaten produzieren.

DER ZEITRAHMEN DER GESPRÄCHE

Zeit	Aktion	Bemerkungen
00:00	Begrüßung, Vorstellen, Ziel erklären, Anonymität, Anwendungsfälle erklären	Kleines Geschenk überreichen, angenehme Atmosphäre schaffen
00:05	TIS hochfahren, Projekt öffnen, Basisdaten erheben	
00:05	Anwendungsfälle durcharbeiten	Ca. 10 Minuten pro Anwendungsfall
00:40	Leitfragen wenn noch nicht gestellt	

00:40 - 00:45	Gespräch beenden	
---------------	------------------	--

ARBEITSANWEISUNGEN „ANWENDUNGSFÄLLE“

BEMERKUNGEN

Den Anwendungsfällen liegen zwei verschiedene Datensätze zu Grunde. Für die beiden ersten Anwendungsfälle mit Zeitbezug gibt es den Datensatz Handel Umsatz, der Umsatzdaten mit einem Zeitstempel enthält. Der zweite Datensatz stammt aus der OECD Pisa-Studie und enthält die Daten vier verschiedener Länder. Pro Schüler wurde die Lese-, Mathematik- und Wissenschaftskompetenz erhoben.

ANWENDUNGSFALL 1 „OPERATOR LISTE + EINSTELLUNGEN + FARBENDIALOG“

Fügen Sie im Knoten C00 den Operator Landkarte 2D Liste hinzu. Die Einstellungen sind: Identifizierer 1: D Woche, Von: E Von Hilf, F Bis Hilf, Werte C Daten (Default passt außer von, bis)

Achse vertikal: Achsentitel „Kalenderwochen“, zeige nur jede x-te Kategorie 4 (Damit wird erreicht, dass nur jede 4. Kalenderwoche in der Beschriftung aufscheint)

Achse horizontal: Achsenbeschriftung: Achsentitel „Wochentag und Uhrzeit“, Benutzerdefiniertes Format „ddd HH:mm“ (Wochentag kurz + Stunde und Minute), Skalierung: Grundeinheit Stunden, Einheiten pro Intervall 4, Anzahl Intervalle pro Seite: 42
Diagramm Layout: Höhe 100%, Breite 200%

Dialog Farben:

Erstellen Sie nun einen eigenen Verlauf mit den Farben wie unten.



Bestätigen Sie mit „Anwenden“ und dann „OK“.

Besprechen was auf X-Achse, Y-Achse zu sehen ist, was sind die Werte

Legende! 0-Werte ansprechen.

Vorteile hier: Lücken, Extremwerte, alle Datensätze in einer Grafik, Tagesverlauf, Wochenverlauf

ANWENDUNGSFALL 2 „OPERATOR LISTE+ FARBENDIALOG“



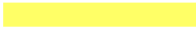






Öffnen Sie den Knoten C10.

Hier liegt das gleiche Diagramm wie oben vor.

Öffnen Sie den Dialog „Farben“. Stellen Sie die Intervalle 0-10, 10-90, 90-100 ein. Nehmen Sie für das erste Perzentil ein Rot, für das zweite ein Gelb und für das dritte ein Grün.

Allgemein

Farbcodierung

von	bis incl.	Farbwert	
0,00	10,00		
> Vorgänger	90,00		 
> Vorgänger	100,00		  

Interpretiere von/bis Angaben als Perzentile

Diskutieren, was man sieht.

Schwäche, dass auch undefinierte Werte wie der Hintergrund rot eingefärbt werden.

ANWENDUNGSFALL 3 “OPERATOR PIVOT”

Fügen Sie im Knoten C20 den Operator Landkarte 2D Pivot hinzu.

Die Einstellungen sind: Identifizierer 1: A Id, Identifizierer 2: B Land, Werte: D-F

Einstellungen: Achse vertikal: Skalierung: Anzahl Kategorien pro Seiten: 14142. Zeige nur jede 4714. Kategorie

Achsenbeschriftung

Achsentitel:

Format:

Benutzerdefiniertes Format:

Maximale Länge:

Orientierung (in Grad):

Gitternetzlinien

Art:

Farbe:

Skalierung

Anzahl Kategorien pro Seite:

Zeige nur jede xte Kategorie

Bitte betrachten Sie das Ergebnis. Erklären, was auf welchen Achsen zu sehen ist. Eine Zeile = ein individueller Schüler.

Können Sie einen Effekt erkennen. Wenn ja, welchen?

Korrelation Lesen->Mathematik

Wie breit sind die einzelnen Balken?

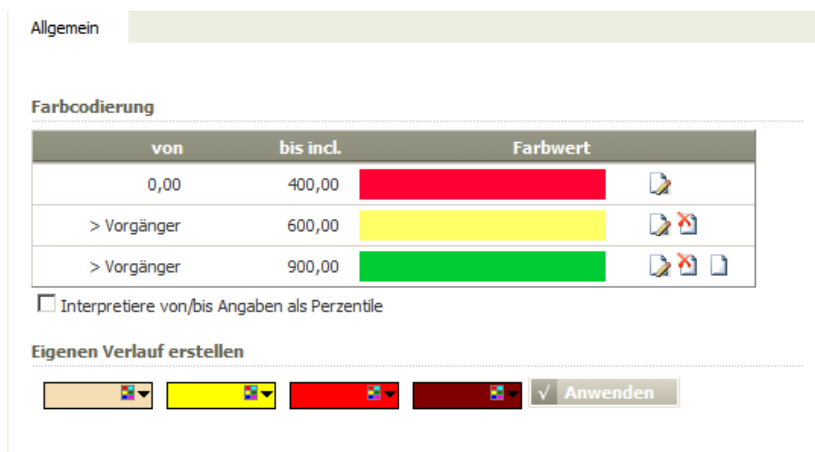
ANWENDUNGSFALL 4 „OPERATOR PIVOT + FARBE PERZENTIL“

Fügen Sie im Knoten C30 den Operator Landkarte 2D Pivot hinzu. Die Einstellungen sind: Identifizierer 1: A Id, Identifizierer 2: B Land, Werte: G-I

Einstellungen: Achse vertikal: Skalierung: Anzahl Kategorien pro Seiten: 14142. Zeige nur jede 4714. Kategorie (wie oben)

Öffnen Sie den Einstellungsdialog „Farbe“. Ziel ist es nun, Farbmarkierungen und Markierungsgrenzen zu ändern.

Deaktivieren Sie das Häkchen „Interpretiere von/bis Angaben als Perzentile“. Löschen Sie nun alle Markierungen. Stellen Sie die Grenzen auf 0-400, 400-600 und 600-900. Verwenden Sie Ampelfarben für den Farbwert.



Was ist zu erkennen, wo liegt der Unterschied zu Scatterplots, Liniendiagramm.

LEITFRAGEN WÄHREND/NACH DEN ANWENDUNGSFÄLLEN

- Sind Ihnen Schwächen/Stärken des Operators aufgefallen, die Sie erwähnen möchten?
Diese Frage soll feststellen, was dem Benutzer besonders gut oder schlecht in Erinnerung geblieben ist.
- Liefern die Einstellungen die erwarteten Ergebnisse?
Sind die Einstellungsmöglichkeiten ausreichend beschriftet und wurden klare Bezeichnungen gewählt.
- Wo lagen die Herausforderungen bei der Anwendung/den Einstellungen?
Der Dialog zur Einstellung der Farbkodierung könnte für die Benutzer Tücken beinhalten. Die automatische Sortierung nach „bis incl.“ ist nicht intuitiv, aber praktisch.
- Fehlen Ihnen Einstellungsmöglichkeiten?
Gibt es Einstellungen, die die Benutzer erwartet, die aber von mir nicht berücksichtigt wurden.
- Zeichenfläche
Sind die Möglichkeiten zur Formatierung der Zeichenfläche ausreichend?
- Legende
Ist die Legende gut lesbar und verständlich, ist die geänderte Anzeige bei der Auswahl von Perzentilen nachvollziehbar.
- Andere relative Bezüge als Perzentil denkbar?
Im Einstellungsdialog kann man momentan nur Perzentile als relatives Maß definieren. Sind auch andere Bezüge gewünscht. Denkbar sind hier zum Beispiel Abweichungen vom arithmetischen Mittel, vom Median, Minimum oder Maximum.
- Interaktivität

KATEGORIEN

Bereits vor der Auswertung und Durchführung legte ich die Kategorien fest, zu denen ich die Antworten zuordnete. Nach dem ersten Durchlauf ist eine Anpassung zu erwar-

ten. Die Kategorien für die erste Zuordnung lauten: Zeichenfläche, Legende, Einstellungen Farbe, Einstellungen Diagramm, Auswahldialog Operator und Allgemeines.

AUSWERTUNG INTERVIEWS

INTERVIEW A

Wien, am 23.7.2008. Eine Person. Vor Ort.

Erfahrung mit Visualisierung: keine, hat aber diese Visualisierung schon gesehen.

Nr.	Zeit	Aussage	Kategorie
1	3:30	Achsenbeschriftung horizontal/vertikal ersetzen durch X-Achse, Y-Achse	Einstellungen Diagramm
2	7:11	Viele gleichartige Werte z.B.: 0 „verziehen“ den Bereich der Perzentile	Allgemeines
3	10:40	Daten entlang der vertikalen Achse mit Zeitbezug sollten von unten nach oben angeordnet werden. d.h. niedrigstes Datum unten	Allgemeines
4	14:01	0-Werte sollten automatisch einer Farbe zugeordnet werden, ein Wert sollte automatisch nur zu einer Farbe zugeordnet werden können	Farbe Allgemeines
5	14:28	Checkbox „0-Werte zusammenfassen“	Farbe Allgemeines
6	14:52	Andere Farbpalette, es fehlt ein schönes Rot, es fehlen schöne Grundfarben, es fehlen klare Farben mit schönen Abstufungen. So wie Weiß-Schwarz Abstufungen in den Primärfarben	Farben
7	15:58	Farbdarstellung ist unlogisch, sollte ein Farbkreis sein	Farben
8	17:05	Weniger bunte Farben	Farben
9	19:28	Schöne Verläufe z.B. Rot-Schwarz	Farben
10	21:30	Vordefinierte Farbverläufe in n Abstufungen	Farben
11	23:16	Viel Aufwand für das Nachbearbeiten der Farbmarkierungsgrenzen erforderlich, es sollte mehr automatisieren	Markierungen bearbeiten

		tisch gehen oder zumindest benutzerfreundlicher	
12	23:44	Farbstufen festlegen können	Farben
13	24:07	Farbschema vorschlagen	Farben
14	24:38	Farbwerte direkt in der Legende verändern	Legende
15	24:57	Ausblenden von Farben in der Legende	Legende
16	26:12	Nur eine Farbmarkung anzeigen, alles andere ausblenden	Zeichenfläche
17	26:46	Farben invertieren	Zeichenfläche
18	33:54	Durchschnitt als Vergleichsmaß	Berechnung
19	35:35	Referenzwert eingeben, wo bin ich besser oder schlechter als ein Referenzwert	Farbe Allgemeines
20	38:7	Möglichkeit zum setzen von Markierungen an der X- und Y-Achse z.B. wo beginnt die Urlaubszeit	Zeichenfläche
21	53:21	Sortierung nach „bis“ im Farbeinstellungsdialog ist unklar	Markierungen bearbeiten
22	53:34	„Anwenden“ im Farbeinstellungsdialog ist sehr gemein, weil es die Farben überschreibt	Farbe Allgemeines
23	57:34	Ausblenden von gewissen Wertebereichen ermöglichen	Zeichenfläche

INTERVIEW B

Nürnberg, am 24.7.2008. Zwei Personen. Vor Ort.

Erfahrungen mit Visualisierung: keine.

Nr.	Zeit	Aussage	Kategorie
1	5:22	Daten entlang der vertikalen Achse mit Zeitbezug sollten von unten nach oben angeordnet werden. d.h. niedrigstes Datum unten	Allgemeines
2	10:28	Link zum Einstellungsdialog Farbe sollte anders heißen	Farbe Allgemeines

3	11:03	Dialog zur Farbeinstellung könnte unter den Einstellungen integriert werden	Farbe Allgemeines
4	13:07	Bei „Eigenen Verlauf erstellen“ sind vier Farbbereiche zu viel, zwei würden reichen	Farben
5	14:06	Drei Farbbereiche beim Verlauf würden das Optimum sein, da in der Mitte in neutraler Wert dargestellt werden könnte	Farben
6	15:47	Eigener Verlauf müsste auch mit zwei Farben gehen, das hängt von der Anwendung ab	Farben
7	16:55	Absolute Werte, die aus den Perzentilen berechnet werden, sollte man gleich im Einstellungsdialog anzeigen	Markierungen bearbeiten
8	19:10	Vor der Aufteilung nach Perzentilen sollte es möglich sein, Grenzwerte (Min, Max) zu definieren, zwischen denen dann die Aufteilung in 10 Prozent Schritten erfolgt	Markierungen bearbeiten
9	21:08	„Lupenfunktion“ wäre toll, oberen und unteren Bereich wegnehmen und dazwischen Bereiche generieren	Markierungen bearbeiten
10	22:30	Möglichkeit zur Eingabe von-bis und dann zum Beispiel 10 Intervalle generieren das heißt Eingabefelder Minimum – Maximum und Anzahl Intervalle	Markierungen bearbeiten
11	23:43	Wichtig ist eine schnelle Eingabe und nicht die Möglichkeit alles selbst definieren zu können	Farbe Allgemeines
12	23:56	Dialog nach der Vorgehensweise ausrichten: Ganz oben die Grenzen, dann die Anzahl der Intervalle, dann unten den Farbverlauf	Farbe Allgemeines
13	24:14	Auch beim „Eigenen Verlauf erstellen“ sollte bei jeder Farbe gezeigt werden, welche Werte damit eingefärbt werden würden	Farben
14	25:56	Eingabemöglichkeit, um eigene Farbe für Ausreißer	Markierungen

		zu definieren	bearbeiten
15	27:00	Der Auswahldialog Farbe sollte nicht mit zu viel Möglichkeiten überladen werden, wichtig ist die Möglichkeit zur schnellen Auswahl von Farbwerten und Intervallen	Farbe Allgemeines
16	27:26	Perzentile sollte reichen, alles andere wäre zu aufblasen	Berechnung
17	28:11	In einem eigenen Tab-Reiter könnten andere Berechnungsmöglichkeiten verpackt werden	Berechnung
18	29:55	Gitternetzlinien sollten horizontal und vertikal auswählbar sein	Zeichenfläche
19	30:50	Einzelne Werte sollten beschriftet werden können (mit dem Datenwert)	Zeichenfläche
20	47:50	Der Button „Anwenden“ kann in seiner Wirkung mit „Übernehmen“ oder „OK“ verwechselt werden und kann einen mühevoll erstellten benutzerdefinierten Verlauf zerstören	Farbe Allgemeines
21	49:08	Checkbox „Interpretiere von/bis Angaben als Perzentile“ sollte kürzer beschriftet werden, die Bezeichnung „in Prozent“ sollte reichen	Farbe Allgemeines

INTERVIEW C

Erlangen, am 25.7.2008. Eine Person. Vor Ort.

Erfahrungen mit Visualisierung: ja, hat schon eigene Auswertungen damit gemacht.

Nr.	Zeit	Aussage	Kategorie
1	8:58	Einstellungsdialog Farbe würde hinter dem Link Einstellungen vermutet werden.	Farbe Allgemeines
2	9:18	Vorschlag: Farben und Bereiche	Farbe Allgemeines

3	12:17	Dynamische Gitternetzlinien, zum Beispiel jede n-te Kategorie, jede m-te Minute, Stunde, Woche	Zeichenfläche
4	15:05	Im Einstellungsdialog Farben sollte der Button „Neu“ auch dann bleiben, wenn man darauf geklickt hat	Markierungen bearbeiten
5	17:17	Standard-Farbverläufe wie Rot-Grün anbieten	Farben
6	17:52	Abstufungen mit zwei oder drei Farben	Farben
7	21:06	Nicht nur Verlauf sollte automatisch erstellt werden, sondern auch die Farbmarkierungen	Markierungen bearbeiten
8	21:19	Von-Bis angeben und Farbmarkierungen in gleichen Abständen ausrechnen	Markierungen bearbeiten
9	22:18	Eigener Verlauf erstellen sollte nach oben, dann Eingabefelder für Von, Bis und Schrittweite	Markierungen bearbeiten
10	22:45	Einstellungsdialog Farbe schneidet Wertetabelle ab, Scrollfeld muss angezeigt werden	Farbe Allgemeines
11	41:25	Farbdialog könnte bei den anderen Einstellungen integriert werden	Farbe Allgemeines
12	42:13	Logarithmische Skalierung auswählen können	Farbe Allgemeines
13	43:03	Generator von-bis mit bestimmten Intervallen	Markierungen bearbeiten
14	47:06	Tooltips mit dem Datenwert sollen angezeigt werden	Zeichenfläche

INTERVIEW D

Erlangen, am 25.7.2008. Eine Person. Vor Ort.

Erfahrungen mit Visualisierung: keine.

Der Interviewpartner betonte, dass er erst mehr mit der Visualisierung arbeiten müsste, um genauere Aussagen treffen zu können.

Nr.	Zeit	Aussage	Kategorie
-----	------	---------	-----------

1	15:10	Tooltip mit Datenwert muss immer funktionieren	Zeichenfläche
2	17:18	Nullwerte rausnehmen, mit weiß einfärben	Farbe Allgemeines
3	24:43	Einfache Filtermöglichkeit für Datenwerte	Zeichenfläche

INTERVIEW E

Erlangen, am 25.7.2008. Eine Person. Vor Ort.

Erfahrungen mit Visualisierung: sehr wenig, schon gesehen.

Nr.	Zeit	Aussage	Kategorie
1	10:23	Link „Farben“ sollte „Farbeinstellungen“ heißen	Farbe Allgemeines
2	18:54	Link sollte eigentlich „Farb- und Intervalleinstellungen“ heißen	Farbe Allgemeines
3	20:07	Kategorien, die sich widersprechen, sollten automatisch rausfallen	Markierungen bearbeiten
4	23:05	Durchschnitt als Referenzwert automatisch berechnen	Berechnung
5	24:45	Vordefinierte Farbverläufe wären schön zum Beispiel Weiß nach Braun über Gelb	Farben
6	25:51	Möglichkeit, um eigene Vorlagen speichern zu können, die man später wieder aufrufen kann	Farbe Allgemeines
7	26:32	Hier sollen auch schon vordefinierte Vorlagen enthalten sein	Farbe Allgemeines

INTERVIEW F

Hamburg/Wien, am 25.7.2008. Eine Person. Webkonferenz.

Erfahrungen mit Visualisierung: keine.

Nr.	Zeit	Aussage	Kategorie
-----	------	---------	-----------

1	16:01	Manche Bereiche sollten noch markanter markiert werden können (abseits der Farbgebung)	Farbe Allgemeines
2	16:15	Filterfunktionen für Daten, die interaktiv ausgelöst werden können	Zeichenfläche
3	16:33	Markierungen über Drag and Drop treffen	Zeichenfläche
4	17:15	Interaktivität ist sehr wichtig	Zeichenfläche
5	28:42	Im Dialog zur Farbauswahl sollten vordefinierte Skalen(=Markierungen) zur Verfügung stehen	Markierungen bearbeiten
6	30:12	Vorgabe, um nicht einzeln die Farbintensitäten eingeben zu müssen	Farben
7	30:20	Skalierung 1-x, dann soll der Fokus auf bestimmte Intervalle (z.B. Extremwerte) gelegt werden	Markierungen bearbeiten
8	30:55	Bessere Orientierung auf der Zeichnungsfläche durch Gitternetzlinien	Zeichenfläche
9	31:23	Raster über der Grafik	Zeichenfläche
10	31:57	Experte ist Rot-Grün blind, mögliche vordefinierte Farbvorlagen sollten darauf Rücksicht nehmen	Farben
11	32:14	Auch der Ausdruck sollte darauf Rücksicht nehmen	Farben
12	32:19	Zahlen andrucken, in bestimmten Raster Zahlen von 1-10 drucken, weg von der Farbgebung und mehr Richtung Zahlen	Allgemeines
13	33:34	Tooltip sollte immer den Wert des Datenpunktes zeigen	Zeichenfläche
14	35:00	Bereiche an den Achsen mit Mausklick aktivieren/deaktivieren	Zeichenfläche
15	35:28	Auswahldialog zur Selektion von Identifizierern sollte Mehrfachselektion zulassen	Einstellungen Diagramm
16	36:46	„Farben setzen“ als Bezeichnung des Links für den Farbauswahldialog	Farbe Allgemeines
17	51:04	Sondertage in die Visualisierung aufnehmen, zum Beispiel offene Samstage im Advent	Allgemeines

18	51:40	Bestimmte Kalendertage vormarkieren, die anders behandelt werden	Allgemeines
19	52:39	Tagestypen, die aus einer Kalenderfunktion extrahiert werden	Allgemeines
20	52:58	Kalender auf die Daten legen, Mo-So und bestimmte Tagestypen extra definieren	Allgemeines

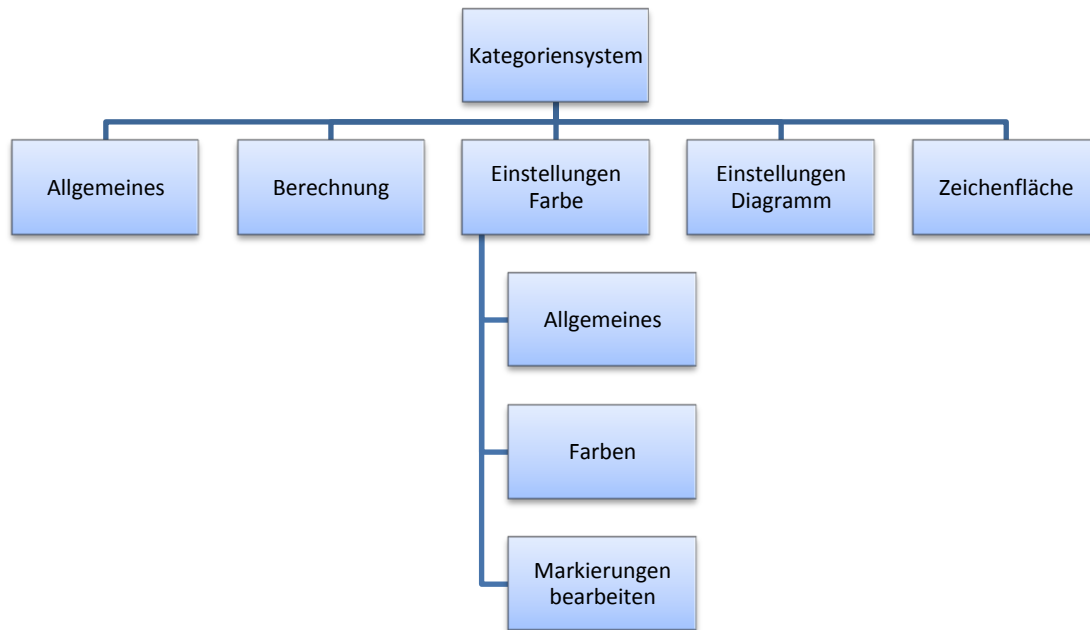
KATEGORIENBILDUNG

Die Einteilung der Aussagen aus den Interviews ergibt folgendes Bild:

Kategorie	Interview-Nr.	Anzahl
Zeichenfläche	A16, A17, A20, A23, B18, B19, C3, C14, D1, D3, F2, F3, F4, F8, F9, F13, F14	17
Legende	A14, A15	2
Einstellungen Farbe	A4, A5, A6, A7, A8, A9, A10, A11, A12, A13, A19, A21, A22, B2, B3, B4, B5, B6, B7, B8, B9, B10, B11, B12, B13, B14, B15, B20, B21, C1, C2, C4, C5, C6, C7, C8, C9, C10, C11, C12, C13, D2, E1, E2, E3, E5, E6, E7, F1, F5, F6, F7, F10, F11, F16	55
Einstellungen Diagramm	A1, F15	2
Auswahldialog Operator	-	0
Allgemeines	A2, A3, B1, F12, F17, F18, F19, F20	8
Berechnung	A18, B16, B17, E4	4

Die meisten Fundstellen in den Interviews ließen sich der Kategorie „Einstellungen Farbe“ zuordnen. Aufgrund der vielen Einträge in der Kategorie ist eine Revision notwendig. Daher teile ich diese Kategorie weiter in die Bereiche „Farben“, „Markierungen bearbeiten“ und „Allgemeines“ auf. Die Kategorie „Auswahldialog Operator“ konnte

fallen gelassen werden, da es keine Aussagen dazu gab. Die Kategorie „Berechnung“ wurde neu eingeführt. So ergibt sich das endgültige Kategoriensystem:



Nach der wiederholten Einteilung ergibt sich die endgültige Zuordnung, die in der folgenden Tabelle und in den Protokollen oben eingetragen wird. Nun erfolgt die Analyse der Kategorien und die Ableitung weiterer Schritte, die für eine Weiterentwicklung des Programmteils von Nöten sind.

Kategorie	Interview-Nr.	Anzahl
Allgemeines	A2, A3, B1, F12, F17, F18, F19, F20	8
Berechnung	A18, B16, B17, E4	4
Einstellungen Farbe Allgemeines	A4, A5, A19, A22, B2, B3, B11, B12, B15, B20, B21, C1, C2, C10, C11, C12, D2, E1, E2, E6, E7, F1, F16	23
Einstellungen Farbe Farben	A6, A7, A8, A9, A10, A12, A13, B4, B5, B6, B13, C5, C6, E5, F6, F10, F11	17
Einstellungen Farbe Markierungen bearbeiten	A11, A21, B7, B8, B9, B10, B14, C4, C7, C8, C9, C13, E3, F5, F7	15
Einstellungen Diagramm	A1, F15	2

Legende	A14, A15	2
Zeichenfläche	A16, A17, A20, A23, B18, B19, C3, C14, D1, D3, F2, F3, F4, F8, F9, F13, F14	17

AUSWERTUNG DER KATEGORIEN

ALLGEMEINES

Diese Kategorie fasst Aussagen zusammen, die sehr allgemeine Feststellungen oder Anmerkungen beinhalten oder sich nicht eindeutig in andere Bereiche einordnen lassen. Ein Beispiel, das zweimal erwähnt wurde, ist die Sortierung der vertikalen Achse in Anwendungsfall 1 und 2, die umgedreht werden sollten (A3, B1). Eine Ansicht der Visualisierung wie in F12 vorgeschlagen, die nicht auf Farben basiert, sondern die Zahlen direkt andruckt, könnte als eigenständiges Projekt vorgeschlagen werden. Die Vorschläge von F17-F20 beschreiben eine sehr sinnvolle Erweiterung. Die Visualisierung sollte die Möglichkeit bieten, mit einem Kalender verknüpft zu werden. Hier sollten verschiedene Tagestypen definiert werden können, die auf der Zeichenfläche gesondert hervorgehoben werden. Denkbar sind hier Tage wie Einkaufssamstag in der Zeit vor Weihnachten, Vorfeiertage oder Tage, an denen von Unternehmen spezielle Aktionen durchgeführt werden. Dieses spezielle Merkmal wurde aber nur in einem Interview angesprochen und wird daher von mir als weniger dringlich angesehen.

BERECHNUNG

Die Kategorie Berechnung sammelt Aussagen, die zur Art der Berechnung der Grenzen der Farbmarkierungen gehören. Momentan sind hier nur zwei Arten möglich. Die Standardeinstellung interpretiert die Grenzen als Perzentilgrenzen. Hier werden die angegebenen Prozentwerte über den Wertebereich gerechnet und das Ergebnis als Grenze gesetzt. Die andere Möglichkeit ist, die Grenzen direkt als absolute Werte anzugeben. In A18 wurde der Durchschnitt als Vergleichsmaß gefordert, während in B18 gesagt wird, dass hier keine weiteren Einstellmöglichkeiten geschaffen werden sollten. Nach B17 sollten wenn überhaupt andere Berechnungsmöglichkeiten in einen eigenen Tab-Reiter verpackt werden. E4 sagt, dass der Durchschnitt der Werte automatisch gerechnet und angezeigt werden sollte. Alle Vorschläge wirken nicht sehr überzeugend und da es nur vier Aussagen zu dieser Kategorie gibt, sieht man auch, dass für Endbenutzer in diesem

Bereich offensichtlich wenig Handlungsbedarf besteht. Die Frage nach alternativen Berechnungsmöglichkeiten relativer Maße sollte in einem Hilfetext geklärt werden. In der Software TIS gibt es schon jetzt die Möglichkeit, mit Hilfe des Operators „In Prozent“ automatisch verschiedene relative Maße basierend auf Durchschnitt, Medien, Minimum, Maximum, Rang etc. zu rechnen. Das heißt, dass man die Daten bereits vor der Visualisierung auf einfache Art und Weise aufbereiten kann, um etwa Abweichungen vom Durchschnitt zu zeigen.

EINSTELLUNGEN FARBE ALLGEMEINES

Diese Kategorie enthält Aussagen, die sich allgemein auf den Dialog beziehen, allerdings nicht den Bereich zur Festlegung der Farbmarkierung betreffen. Die Abbildung zeigt diesen Dialog.

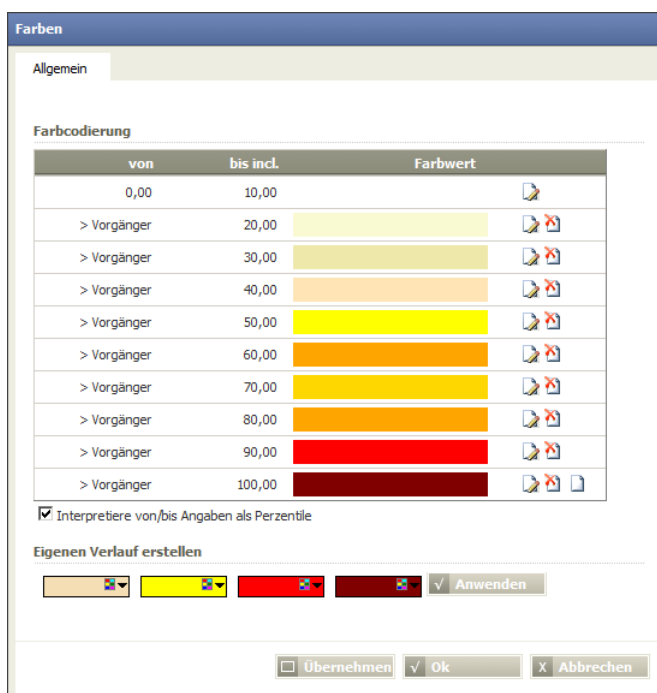


Abbildung 32 Auswahldialog Farbe

A4 meint hier, dass 0-Werte automatisch einer Farbe zugeordnet werden sollten. Dies könnte nach A5 etwa mit Hilfe einer Checkbox passieren. Diese Checkbox ist eigentlich nicht notwendig. Im Bereich „Farbcodierung“ könnte ein eigener Eintrag von 0 bis 0 stehen, dem die Farbe weiß zugewiesen wird. Dies wäre eine sinnvolle Ergänzung der bereits bestehenden Standardeinstellungen. Auch D2 wünscht sich eine Standardeinstel-

lung, die 0-Werte weiß einfärbt. A19 bringt den Vorschlag, einen Referenzwert eingeben zu können, um dann sehen zu können, wo man besser oder schlechter als dieser liegt. Dies ermöglicht die Oberfläche bereits jetzt, da man für die Farbmarkierungen beliebige Werte und Farben wählen kann.

A22 bezieht sich auf den Button „Anwenden“, der einen Verlauf erstellt und die vorhandenen Farbwerte überschreibt. B20 formuliert dies ähnlich und meint zusätzlich, dass die Buttons von „Anwenden“, „Übernehmen“ und „OK“ leicht verwechselt werden können. Der Klick auf „Anwenden“ kann in der Tat ungewünschte Auswirkungen haben, da die standardmäßigen Farbwerte nicht wiederhergestellt werden können und dann mühevoll rekonstruiert werden müssen. Dieser Mangel soll durch die Möglichkeit, Vorlagen zu speichern behoben werden. Neben eigenen Vorlagen werden auch Standardvorlagen zur Verfügung stehen, die den ursprünglichen Zustand wiederherstellen können. Eine zusätzliche Entschärfung des Buttons „Anwenden“ kann die Umbenennung in „Eigenen Verlauf Anwenden“ bringen.

B2 meint, dass der Link zum Einstellungsdialog anders heißen sollte. Auch C1 findet, dass dieser Dialog nicht hinter dieser Bezeichnung vermutet werden würde. C1 hätte diesen Dialog bei den Einstellungen vermutet bzw. dort nach Einstellungsmöglichkeiten der Farbe gesucht. Der Vorschlag von C2 zur Beschriftung lautet „Farben und Bereiche“. E1 und E2 schlagen anstatt der Bezeichnung „Farben“ den Text „Farbeinstellungen“ oder „Farb- und Intervalleinstellungen“ vor, während F16 den Link mit „Farben setzen“ beschriften würde. Hier besteht Handlungsbedarf, die lapidare Bezeichnung mit dem einfachen Substantiv „Farben“ ließ kaum einen Benutzer die Möglichkeiten dieses Dialogs vermuten. Das heißt, dass eine zusätzliche Beschriftung notwendig ist. Mit gefiel die Bezeichnung „Farb- und Bereichseinstellungen“ am Besten, da hier explizit auf die Möglichkeit zur Bearbeitung der Farbbereiche hingewiesen wird. Das Wort „Bereiche“ ist dem Wort „Intervalle“ vorzuziehen, da es geringfügig kürzer ist und weniger kompliziert ist. Wenn noch ausreichend Platz vorhanden ist, könnte der Text noch um das Verb „bearbeiten“ ergänzt werden, um noch zusätzlich auf die Möglichkeit der Interaktion hinzuweisen. Viele Benutzer vermuteten vorerst, dass eine Veränderung der Intervallgrenzen nicht möglich sei und lediglich die Farben verändert werden könnten.

Daher klickten einige Benutzer nicht einmal auf den Link, um nachzusehen, was sich dahinter verbirgt.

Nur von B3 kommt der Vorschlag, diesen Dialog zur Farbauswahl bei den allgemeinen Einstellungen zu integrieren. Auch C11 würde eine Integration bevorzugen. Der Dialog wurde bewusst in ein eigenes Fenster abseits der anderen Einstellungen verpackt, da er sich deutlich von den allgemeinen Möglichkeiten zur Bearbeitung der Einstellungen eines Diagramms abhebt. Daher bleibt dieser Dialog auch eigenständig.

B11 enthält den Ratschlag, beim Design des Dialogs darauf zu achten, dass der Benutzer sehr schnell möglichst vordefinierte Einstellungen treffen kann. Zu diesem Zweck könnte auch die Möglichkeit geopfert werden, die Farbmarkierungen in allen Details selbst zu bearbeiten. Ich denke, dass es nicht von Nöten ist, bestehende Features wegzulassen, da sie von anderen Usern gut angenommen wurden. Sehr wichtig ist es jedoch, einige vordefinierte Vorlagen zu schaffen, die man jederzeit einfach aufrufen und anwenden kann. B15 bekräftigt die Forderung zur schnellen Auswahl von Farbwerten und Intervallen.

B21 stellt fest, dass die Checkbox „Interpretiere von/bis Angaben als Perzentile“ umständlich und lang beschriftet ist. Ein Text „In Prozent“ sollte reichen. Diese Umbenennung kann tatsächlich durchgeführt werden, da einiges an Text eingespart wird. Auch mit dieser kurzen Beschriftung sollte klar sein, was gemeint ist. Besonders sinnvoll ist diese Änderung in Kombination mit der Anforderung, dass auch schon im Auswahldialog die absoluten Werte aus den prozentuellen Angaben berechnet und ähnlich wie in der Legende angezeigt werden.

C10 weist auf den Fehler hin, dass bei der Eingabe zu vieler Markierungen kein Scrollbalken angezeigt wird.

C12 würde gerne automatisch eine logarithmische Skalierung auswählen können. Diese Möglichkeit möchte ich nicht direkt in den Dialog integrieren, da der Benutzer ja auch die Möglichkeit hat, diese Skalierung selbst einzugeben und nachzubilden. Diese Anforderung wurde von keinem anderen Benutzer formuliert und auch vom Interviewpartner C relativiert, sodass ich hier keinen Handlungsbedarf sehe.

E6 und E7 wünschen sich die Möglichkeit, Vorlagen abspeichern und wieder aufrufen zu können. Es sollten auch schon vordefinierte Vorlagen zur Verfügung stehen. Dieser Vorschlag ist sehr sinnvoll und auf jeden Fall umzusetzen.

F1 regt an, dass es gut wäre, bestimmte Bereiche unabhängig von der Farbe noch mehr hervorheben zu können. Dies könnte durch das Zeichnen eigener Linien passieren. Leider ist dies nur sehr schwer umzusetzen und daher kaum machbar.

EINSTELLUNGEN FARBE FARBEN

Diese Kategorie sammelt Aussagen, die sich ausschließlich auf die Farben und den Umgang mit der Farbauswahl im Dialog beziehen.



Abbildung 33 Dialog zur Farbauswahl

A6 meint, dass schöne Grundfarben wie Rot fehlen. Es sollte verschiedene Grundfarben geben, die man in einzelnen Abstufungen auswählen kann. Der Vorschlag ist, so wie für den zurzeit existierenden Verlauf von Weiß bis Schwarz ähnliche Verläufe in den Primärfarben zu schaffen. A7 meint, dass es weniger bunte Farben geben müsste, sondern eine Reduktion mehr bringen würde. Dies könnte nach A8 zum Beispiel ein Verlauf von Rot nach Schwarz sein. Für A12 und A13 wäre es wichtig, wenn die Software ein Farbschema vorschlagen könnte, aus dem sich der Benutzer dann seine gewünschten Farben auswählen könnte. Auch C5 und C6 fordern Standard-Farbverläufe wie etwa Rot-Grün, um die Farbauswahl zu erleichtern. Abstufungen mit zwei oder drei Farben würden hier reichen. Für E5 sind vordefinierte Farbverläufe genau so von Bedeutung, hier wird etwa

Weiß nach Braun über Gelb vorgeschlagen. F6 findet es auch erwähnenswert, dass es bei den Farben Vorgaben geben muss, um nicht selbst die einzelnen Farben eintragen zu müssen. Es ist evident, dass die Auswahl der Farben für die Benutzer ein Problem darstellt. Die aktuelle Form ist wenig zufrieden stellend und nicht praktikabel. Daher werden die Farbverläufe reduziert und weniger, dafür bessere Einstellungen zur Auswahl gebracht. Siehe dazu den Abschnitt weiter unten, in dem ich versuche bessere Farbverläufe anhand der in der Literatur gebrachten Vorschläge zu erstellen.

B4 und B5 finden, dass vier Farben für den eigenen Verlauf zu viel sind und dass drei Farbbereiche auf jeden Fall reichen würden. Für manche Anwendungen würden sogar zwei Farbbereiche reichen, findet B6. Beim Erstellen des eigenen Verlaufes wäre es für B13 wichtig, zu sehen welche Werte von diesen Farben eingefärbt werden würden.

Ein weiterer wichtiger Hinweis kommt von F10 und F11. Dieser Benutzer ist Rot-Grün blind und weist darauf hin, dass mögliche Farbpaletten darauf Rücksicht nehmen müssten. Dies gilt sowohl für den Bildschirm als auch für den Ausdruck.

EINSTELLUNGEN FARBE MARKIERUNGEN BEARBEITEN

In dieser Kategorie finden sich die Aussagen zur Bearbeitung der Markierungen mit ihren Grenzen (Von-Bis) und der dazugehörigen Farbe.

A11 stellt fest, dass im Allgemeinen sehr viel Aufwand für das Nachbearbeiten der Farbmarkierungsgrenzen erforderlich ist und dass Wege gefunden werden müssen, die eine leichtere Bearbeitung ermöglichen. A21 findet, dass die Sortierung nach „bis“ unklar ist. Alle anderen Benutzer kamen mit dieser Einstellung gut zurecht, sodass hier kein Handlungsbedarf besteht.

B7 bringt den sehr guten Hinweis, dass man die absoluten Werte, welche aus den Perzentilen berechnet werden, auch gleich bei der Eingabe anzeigen könnte. Es ist aber zu befürchten, dass die Performance und damit die Antwortzeit etwas leiden könnte, wenn bei jeder Änderung der Markierungsgrenzen die Perzentile neu berechnet werden müssen.

B8 und B10 beschreiben eine sehr gute Idee und sinnvolle Ergänzung, die unbedingt in den Einstellungsdialog integriert werden muss. Anstatt die Farbmarkierungen einzeln zu

erfassen sollte es möglich sein einen Von- respektive Bis-Wert zu erfassen und eine gewünschte Anzahl an Unterteilungen. Die Farbmarkierungen würden dann automatisch generiert werden. So würden die Eingabe Von-Bis von 0 bis 100 und die Schrittweite 10 eine Aufteilung in 10 Farbmarkierungen ergeben. C7, C8, C9 und C13 beschreiben die gleiche Idee, um erstellen Farbmarkierungen zu können. Hier werden Eingabefelder für Von, Bis und Schrittweite vorgeschlagen. In Kombination mit voreingestellten Farblayouts könnte man so in kürzester Zeit tolle Ergebnisse bei der Auswertung erreichen. Eine Erweiterung dazu ist die In B9 vorgeschlagene Lupenfunktion, die vorhandene Farbmarkierungen in zwei oder mehr weitere Bereiche aufteilen kann.

E3 möchte, dass Farbmarkierungen, die sich widersprechen, automatisch rausfallen. Dies kann aber dazu führen, dass die Reaktionen der Software nicht mehr nachvollziehbar werden. Der Benutzer kann dann nicht mehr erkennen, warum der Eintrag gelöscht wurde. Vielleicht sollte im Falle einer Fehleingabe ein Hinweiszeichen eingeblendet werden.

B14 fordert eine Möglichkeit, bestimmte Farben für Ausreißer zu vergeben. Dieser Vorschlag lässt sich nicht umsetzen, da eine Definition für Ausreißer oft nur anwendungsspezifisch Sinn macht und nicht generisch zu lösen ist. Darüber hinaus steht es jedem Anwender frei, Farbbereiche so zu wählen, dass Daten, die für ihn Ausreißer darstellen, individuell eingefärbt werden können.

EINSTELLUNGEN DIAGRAMM

Der Bereich zum Einstellungsdialog Diagramm brachte nur zwei Aussagen. A1 formuliert den Wunsch, die Beschriftung „Achse horizontal/vertikal“ durch „Achse X/Y“ zu ersetzen. Dieser Schritt ist ein guter Vorschlag und auch logisch hinsichtlich der Konsistenz innerhalb der Software. Die zweite Aussage F15 betrifft den Auswahldialog für Identifizierer.

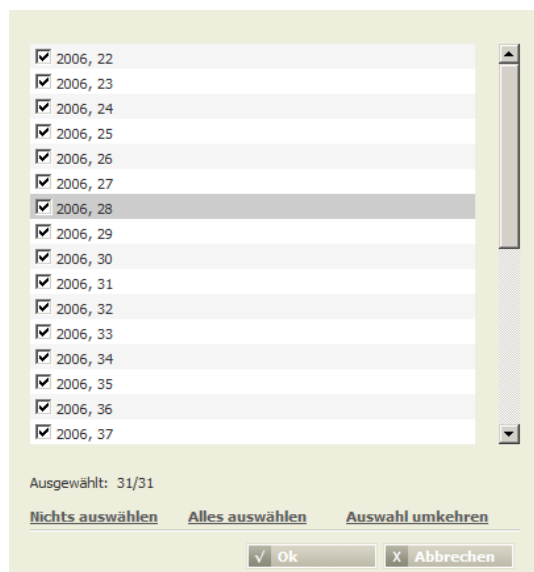


Abbildung 34 Auswahldialog Identifizierer

Die Abbildung zeigt ein Beispiel für diesen Auswahldialog. Hier werden die Kalenderwochen des Jahres 2006 angezeigt. Obwohl der Dialog ein Element zur Auswahl aller oder keiner Elemente bzw. ein Element zur Inversion der Selektion enthält, kann es sehr mühsam sein einen Bereich von etwa 15 Elementen oder mehr auszuwählen. Dafür sind n Klicks für n Elemente von Nöten. Der Vorschlag von F15 dazu ist, die einzelnen Checkboxes durch ein Listenelement zu ersetzen. Ein Listenelement ermöglicht im Gegensatz zu dieser Form der Darstellung eine Mehrfachselektion über Bereiche und den Einsatz der Strg-Taste zur Selektion einzelner Einträge. Durch eine Anpassung in diesem Sinne wird der Bedienkomfort entscheidend erhöht.

LEGENDE

Zur Kategorie Legende kamen nur zwei Aussagen. Es besteht der Wunsch, Farbwerte in der Legende direkt zu ändern (A14) bzw. Farbwerte durch Klick auf die Legende aus-

blenden zu können (A15). A14 könnte durch einen Link realisiert werden, der auf den Einstellungsdialog Farbe verweist. Der zweite Vorschlag ist aufwändig in der Umsetzung und wurde nur von einem Benutzer gewünscht. Diese Änderung ist daher abzulehnen.

ZEICHENFLÄCHE

Die Kategorie Zeichenfläche bündelt alle Aussagen, die sich direkt auf die Zeichenfläche des Diagramms beziehen. Eine sehr allgemeine Feststellung dazu kommt von F4, es wird betont dass Interaktivität an der Zeichenfläche sehr wichtig ist. Dies bezieht sich vor allem auf die Möglichkeit, Filter und Markierungen zu ändern oder andere Einstellungen zu treffen. Jede Aktion sollte es möglichst direkt eine Reaktion an der Zeichenfläche auslösen.

A16 hätte gerne die Möglichkeit, alle Farbmarkierungen bis auf eine auf einfache Art und Weise (nicht über den Dialog „Farbe“) wegschalten zu können. Ähnlich dazu wünscht sich A17 die Möglichkeit, Farben direkt über die Zeichenfläche invertieren zu können. Andere Experten äußerten keine ähnlichen Wünsche. A23 schlägt vor, dass es eine Möglichkeit zum Ausblenden von gewissen Wertebereichen direkt an der Zeichenfläche geben müsste. Dies ist aber auch jetzt bereits sehr einfach möglich, da man im Dialog „Farbe“ jederzeit gewisse Bereiche mit der gleichen Farbe wie die Hintergrundfarbe einfärben kann und so den Effekt erzielt, dass gewisse Bereiche nicht sichtbar sind. Da der Dialog parallel, das heißt neben der Zeichnungsfläche geöffnet sein kann, ist diese Einstellungsmöglichkeit recht komfortabel zu benutzen.

B18 und F8 fordern horizontale und vertikale Gitternetzlinien, die momentan noch nicht existieren. F9 meint, dass ein einfaches Raster über die Grafik reichen würde. A20 würde auch gerne direkt an der Zeichenfläche Markierungen setzen können. Dies könnte zum Beispiel eine Linie sein, die den Beginn der Urlaubs- oder Ferienzeit markiert. Diese Erweiterung wäre für die Visualisierung gewiss sehr hilfreich, da oft große Zeiträume dargestellt werden und diese Hilfslinien die Orientierung erleichtern. C3 ergänzt diesen Wunsch und hätte gerne dynamische Gitternetzlinien, die jede n-te Kategorie oder jede m-te Minute, Stunde, Woche oder Monat gezeichnet werden. A20 würde auch gerne direkt an der Zeichenfläche Markierungen setzen können. Dies könnte zum Bei-

spiel eine Linie sein, die den Beginn der Urlaubs- oder Ferienzeit markiert. Diese Erweiterung wäre für die Visualisierung gewiss sehr hilfreich, da oft große Zeiträume dargestellt werden und diese Hilfslinien die Orientierung erleichtern. F3 würde gerne über Drag and Drop Markierungen setzen.

B19, C14, D1 und F13 wünschen sich, dass die einzelnen Punkte mit dem Datenwert beschriftet werden. Dies ist momentan auch der Fall, aber nur wenn weniger als 5000 Datenpunkte vorhanden sind. Wenn man mit der Maus über einem Punkt verweilt, wird der Wert als Tooltip angezeigt. Bei mehr als 5000 Datenpunkten leidet die Geschwindigkeit des Seitenaufbaus zu stark. Es wäre angebracht, hier eine andere technische Lösung zu finden, sodass die Datenwerte immer als Tooltip angezeigt werden.

D3 und F2 wünschen sich direkt an der Zeichenfläche einfache Filtermöglichkeiten für Datenwerte. Diese sollten möglichst interaktiv ausgelöst werden können. Auch F14 würde gerne Bereiche entlang der Achsen mit Mausclick aktivieren und wieder deaktivieren.

ZUSAMMENFASSUNG

Die Auszählung der Nennungen bei den Kategorien ergibt, dass bei den Kategorien „Einstellungen Farbe Allgemeines“, „Einstellungen Farbe Farben“ und der Kategorie „Zeichenfläche“ der meiste Handlungsbedarf besteht.

Die detaillierte Beschreibung der Möglichkeiten zur Verbesserung findet sich in den Abschnitten weiter oben.

Bei den „Einstellungen Farbe Allgemeines“ ist es am wichtigsten, die Möglichkeit zur Generierung der Bereichsgrenzen in einem definierten Intervall zu schaffen. Außerdem sollen Vorlagen angeboten werden und Benutzern die Möglichkeit geboten werden, Vorlagen abzuspeichern und wieder zu laden.

Bei den „Einstellungen Farben Farben“ müssen vordefinierte Farbskalen in Anlehnung an den ColorBrewer (vgl. [Bre032]) zur Verfügung stehen.

Bei der Zeichenfläche ist es wichtig, über mehr Interaktivität, z.B. durch einen dynamischen Filter, zur Verfügung zu stellen. Tooltips, die den genauen Datenwert beim länge-

ren Verweilen des Mauszeigers anbieten, müssen integriert werden. Der Bereich der Gitternetzlinien muss überarbeitet werden. Auch hier sind dynamische Eigenschaften wichtig, wie zum Beispiel eine Linie nach jeder dritten. Woche zu zeichnen.

ZUSAMMENFASSUNG UND AUSBLICK

Die Exploration großer Datenmengen ist ein sehr wichtiges, aber schwieriges Problem. Techniken der Visualisierung können helfen, dieses Problem zu lösen. An den grundlegenden Techniken der Visualisierung wird seit langem gearbeitet. Ein fundamentales Konzept ist das von [Shn96] vorgeschlagene „Information Seeking Mantra“. Es fasst die wichtigsten Schritte der Visualisierung sehr prägnant zusammen: „Overview first, zoom and filter, then details-on-demand“.

Bereits die Forderung, sich einen Überblick zu verschaffen, stellt ein grundlegendes Problem dar. Ein Datensatz mit mehreren Hunderttausend Datensätzen passt in Form einer Tabelle nicht auf eine Bildschirmseite.

Ein möglicher Weg, einen Überblick über Daten zu bekommen ist die Aggregation. Dazu können Summen, Durchschnitte, Standardabweichungen etc. berechnet werden. Wenn Daten mit Zeitbezug vorliegen, können solche Operationen problematisch sein und den „Blick“ auf die Inhalte verstellen. Zum Beispiel kann eine Auswertung über die Monate eines Jahres kann sehr problematisch sein und zu falschen Ergebnissen führen. Wenn ein April mit einem Donnerstag beginnt, hat er 22 Werktage (Mo-Fr), beginnt er jedoch an einem Sonntag, bleiben nur 20 Werktage übrig. Werden nun konstante Umsätze über diese beiden Monate zweier Jahre verglichen, so bringt nur diese Summenbildung eine Schwankung von 10%!

Die Einteilung der Zeit in Form des Kalenders ist eine der ältesten Errungenschaften der Menschheit.

„The calendar is primarily responsible for the creation of most of the temporally regular patterns through which nearly all societies, social institutions, and social groups manage to introduce some orderliness into their lives“⁴¹

So wird deutlich, dass die Visualisierung von Zeit den Blick auf soziale Rhythmen ermöglichen muss. Die Woche mit sieben Tagen ist ein zentrales Element unserer Organi-

⁴¹ [Zer81] p. 31

sation der Zeit. Die Kenntnis des Wochentages lässt uns am sozialen Leben teilnehmen. Menschen verabreden sich z.B. für einen Samstagabend. Viele Menschen mögen Freitage lieber als Montage. Sehr viele Menschen mögen Samstag und Sonntag („Papa, warum ist das Wochenende so kurz und die Woche so lang?“⁴²). Das genaue Datum ist dabei nicht entscheidend.

Bei der Synchronisation der Wochen mit Monaten und Jahren treten Probleme auf, da ein neues Kalenderjahr nicht immer mit einer neuen Woche beginnt. Daher ist es häufig sinnvoll, zeitbezogene Daten im Zusammenhang mit der Woche zu visualisieren.

Dies kann die vorliegende Arbeit leisten. Die Darstellung von massenhaften zeitbezogenen Daten in Form der Landkarte kann einen kompletten Jahresverlauf auf einer Bildschirmseite abbilden und trotzdem die Genauigkeit des Wochenverlaufs bis mindestens einer Stunde abbilden. Die Matrix-basierte Darstellung ermöglicht auf einer angenommenen Auflösung von 1000x1000 die Anzeige von einer Million Datenpunkten. Die Darstellung kann auf eine Aggregation verzichten. Es werden keine Summen gebildet. Es gibt keine Überdeckung von Datenpunkten. Im Wesentlichen wird das „Information Seeking Mantra“ in einer Darstellung erfüllt. Die Visualisierung ist geradezu ideal, sich einen Überblick zu verschaffen. Die Benutzeroberfläche ermöglicht Zoom und Filter Operationen. Es ist nicht notwendig, in dieser Visualisierung nach mehr Details zu suchen, da stets alle Daten direkt zur Verfügung stehen.

Die vorliegende Benutzerstudie gibt einen wichtigen Einblick, wie Experten mit dieser Visualisierung arbeiten. Dabei wurden die Vorteile dieser Form der Visualisierung identifiziert, die Landkarte wird in vielen Analysen als unverzichtbarer erster Schritt eingesetzt, um einen Überblick über Daten zu gewinnen. Die vorliegenden Ergebnisse umreißen klar den Verlauf der nächsten Schritte, die gesetzt werden müssen, die vorgeschlagene Methode der visuellen Datenexploration weiter zu verbessern.

Die von [Aig07] geforderte Orientierung zu den Aufgaben muss konsequent umgesetzt werden. Das bedeutet, dass die Visualisierung automatisch Methoden zur Interaktion

⁴² Verbale Kommunikation: Mikko Miksch, 5 Jahre

zur Verfügung stellt, die mögliche Aufgaben eines Benutzers unterstützen, und diese weitgehend automatisch parametrisiert.

Das meint im Bereich der Farben auf die Arbeit von [Bre031] aufzusetzen und vordefinierte Sets von Farben anzubieten, da die Benutzern bei der Auswahl guter Farben, die die Analyse unterstützen, zur Hand gehen.

Auch in anderen Bereichen der Visualisierung wie bei den Gitternetzlinien können dynamische Elemente zur Verfügung gestellt werden. Das kann zum Beispiel eine Markierung von Wochenenden, Feiertagen, Ferienzeiten oder der Adventzeit sein. Die Auswahl von Zeitbereichen (z.B. Sommerwochen, aber auch bestimmter Tageszeiten (z.B. Vormittag) als Zoom In stellen hier weitere interessante Möglichkeiten dar.

Zusammenfassend ist festzustellen, dass alle Operationen, die die Visualisierung betreffen (Zoom, Filter, Details, etc.) einen möglichst hohen Grad der Interaktivität bieten und in künftige Versionen integriert werden sollen.

ABBILDUNGSVERZEICHNIS

Abbildung 1 Anwesende Mitarbeiter	9
Abbildung 2 Umsatz pro Monat	10
Abbildung 3 Schematische Darstellung des Visualisierungsprozesses [War04]	16
Abbildung 4 Planetenbewegungen 10./11. Jdt [Tuf83]	20
Abbildung 5 Jährliche Erwärmung der Erde von Lambert [Lam79]	20
Abbildung 6 Playfair Exporte und Importe von/nach Indien [Pla86]	21
Abbildung 7 Playfair, Preise, Löhne und Regierungszeiten [Tuf83]	22
Abbildung 8 Marey, Gangarten des Pferdes [Tuf83]	23
Abbildung 9 Marey [Tuf83]	23
Abbildung 10 Minard, Russlandexpedition der napoleonischen Armee [Tuf83]	24
Abbildung 11 Newman, Man and Insects [Tuf83]	25
Abbildung 12 The History and Politics of the Music Industry (Ausschnitt, [Tuf05]) ...	26
Abbildung 13 Oberflächentemperatur der Erde	27
Abbildung 14 Eine typische multivariate Heatmap Matrix	28
Abbildung 15 Zeitreihe eines Wartungsvertrags, die Ausschläge stellen Verletzungen der Antwortzeit dar	30
Abbildung 16 Darstellung der Umsätze pro Tag	30
Abbildung 17 Lego Powered Time Tracking – ein Arbeitstag	33
Abbildung 18 Lego Powered Time Tracking Wochenansicht	34
Abbildung 19 Vergleich multidimensionaler Visualisierungstechniken [Kei00]	36
Abbildung 20 erster Entwurf der Visualisierung	50
Abbildung 21 Eingabedaten Pivot	51
Abbildung 22 Auswahlelement für einen zusätzlichen Identifizierer	52
Abbildung 23 Eingabedaten Liste	53
Abbildung 24 Eingabedaten Liste mit zwei Identifizierern	54
Abbildung 25 Visualisierung Landkarte der Abweichungen	54
Abbildung 26 Einstellungsdialog Visualisierung	56
Abbildung 27 Einstellungsdialog Farbmarkierungen	57
Abbildung 28 Benutzeroberfläche ColorBrewer	64

Abbildung 29 Sequential	65
Abbildung 30 Diverging.....	65
Abbildung 31 Qualitative	65
Abbildung 32 Auswahldialog Farbe.....	86
Abbildung 33 Dialog zur Farbauswahl.....	89
Abbildung 34 Auswahldialog Identifizierer	92

LITERATURVERZEICHNIS

- [Abr02] **Abrahamsson, Pekka, Salo, Outi und Ronkainen, Jussi. 2002.** Agile software development methods. *VTT Publications*. 2002, 478.
- [Aig07] **Aigner, Wolfgang, et al. 2007.** Visualizing time-oriented data - A systematic view. *Computers and Graphics*. 2007, Bd. 31, 3, S. 401-409.
- [Bec04] **Beck, Kent. 2004.** *Extreme Programming Explained*. Boston : Addison-Wesley, 2004.
- [Ber82] **Bertin, Jacques. 1982.** *Graphische Darstellungen und die graphische Weiterverarbeitung der Information*. [Übers.] Wolfgang Scharfe. Berlin New York : Walter de Gruyter, 1982.
- [Ber74] **Bertin, Jaques. 1974.** *Graphische Semiologie: Diagramme, Netze, Karten*. Berlin : de Gruyter, 1974.
- [Bre032] **Brewer, Cynthia A. 2003.** *ColorBrewer*. [Online] 2003. [Zitat vom: 6. 9 2008.] www.colorbrewer.org.
- [Bre031] **Brewer, Cynthia A. 2003.** A Transition in Improving Maps: The ColorBrewer Example. *Cartography and Geographic Information Science*. 30, 2003, Bd. 2, S. 159-162.
- [Bre03] **Brewer, Cynthia A., Hatchard, Geoffrey W. und Harrower, Mark A. 2003.** ColorBrewer in Print: A Catalog of Color Schemes for Maps. *Cartography and Geographic Information Science*. 2003, Bd. 30, 1, S. 5-32.
- [Car08] **Carnegie Mellon. 2008.** Software Architecture for Software-Intensive Systems. [Online] 2008. [Zitat vom: 12. 10 2008.] <http://www.sei.cmu.edu/architecture/definitions.html>.
- [Che08] **Chen, Chun-hou, Härdle, Wolfgang und Unwin, Antony. 2008.** *Handbook of Data Visualization*. Berlin Heidelberg : Springer, 2008.

- [Cle85] Cleveland, William S. 1985. *The Elements Of Graphinc Data*. Monterey : Wadsworth, 1985.
- [Few081] Few, Stephen. 2008. Dual-Scaled Axes in Graphs. *Perceptual Edge*. [Online] März 2008. [Zitat vom: 13. April 2008.] <http://www.perceptualedge.com/library.php>.
- [Few06] Few, Stephen. 2006. Multivariate Analysis Using Heatmaps. *Perceptual Edge*. [Online] 10. Oktober 2006. [Zitat vom: 12. April 2008.] <http://www.perceptualedge.com/library.php>.
- [Few08] Few, Stephen. 2008. Practical Rules for Using Color in Charts. *Perceptual Edge*. [Online] Februar 2008. [Zitat vom: 12. April 2008.] <http://www.perceptualedge.com/library.php>.
- [Few04] Few, Stephen. 2004. *Show Me the Numbers*. Oakland : Analytics Press, 2004.
- [Fli02] Flick, Uwe. 2002. *Qualitative Sozialforschung*. Hamburg : rowohlt, 2002.
- [Gär02] Gärtner, J., S. Miksch, and S. Carl-McGrath. 2002. ViCo: A Metric for the Complexity of Information Visualizations. *Second International Conference on Theory and Application of Diagrams (Diagrams 2002)*. 2002.
- [Gär04] Gärtner, J., S. Popkin, W. Leitner, S. Wahl, T. Åkerstedt, and S. Folkard. 2004. Analyzing Irregular Working Hours - Lessons Learned in the Development of RAS 1.0 - The Representation & Analysis Software. *Chronobiology International*. 2004, Bd. 21, 16, S. 1025-1036.
- [Goo08] 2008. Google - Official Google Blog. *We knew the web was big...* [Online] 25. 2 2008. [Zitat vom: 25. 11 2008.] <http://googleblog.blogspot.com/2008/07/we-knew-web-was-big.html>.
- [Hao07] Hao, Ming C., et al. 2007. A Visual Analysis of Multi-Attribute Data Using Pixel Matrix Displays. *Visualization and Data Analysis 2007*. Proceedings of the SPIE, 2007, Volume 6495.
- [Hea96] Healey, Christopher G. 1996. *Choosing Effective Colours for Data Visualization*. San Francisco : Proceedings IEEE Visualization '96, 1996. S. 263-270.

- [Hun08] **Hunger, Michael. 2008.** *Better Software Development*. [Online] 1. August 2008. [Zitat vom: 25. 9 2008.] <http://jexp.de/blog/archives/16-On-LEGO-Powered-Time-Tracking;-My-Daily-Column.html>.
- [Jev24] **Jevons, Stanley W. 1924.** *Die Theorie der Politischen Ökonomie*. Jena : Gustav Fischer, 1924.
- [Kei02] **Keim, Daniel A. 2002.** Datenvisualisierung und Data Mining. *Datenbank-Spektrum*. 2002, Bd. 2, 2, S. 30-39.
- [Kei00] **Keim, Daniel A. 2000.** Designing Pixel-Oriented Visualization Techniques: Theory and Applications. *IEEE Transactions on Visualization and Computer Graphics*. 2000, Bd. Volume 6, Issue 1.
- [Kei01] **Keim, Daniel A. 2001.** Visual Exploration of Large Data Sets. *Communications of the ACM*. 2001, Bd. 44, 8, S. 38-44.
- [Lam79] **Lambert, Johann Heinrich. 1779.** *Pyrometrie*. Berlin : s.n., 1779.
- [Lud07] **Ludewig, Jochen und Lichter, Horst. 2007.** *Software Engineering*. Heidelberg : dpunkt, 2007.
- [Mac95] **MacEachren, Alan. M. 1995 .** *How Maps Work*. New York : Guilford Press, 1995 .
- [May02] **Mayring, Philipp. 2002.** *Einführung in die Qualitative Sozialforschung*. Weinheim und Basel : Beltz, 2002.
- [Moo98] **Moore, Gordon E. 1998.** Cramming more components onto integrated circuits. *Proceedings of the IEEE*. 1998, Bd. 86, 1, S. 82-85.
- [Mül03] **Müller, Wolfgang und Schumann, Heidrun. 2003.** Visualization Methods for Time-dependent Data - an Overview. [Hrsg.] S Chick, et al. *Proceedings of the 2003 Winter Simulation Conference*. 2003.

- [Nie97] Nielson, Gregory M., Hagen, Hans und Müller, Heinrich. 1997. *Scientific Visualization: Overviews, Methodologies, and Techniques* . Los Alamos : Institute of Electrical & Electronics Engineering, 1997.
- [Pla86] Playfair, William. 1786. *The Commercial and Political Atlas*. London : s.n., 1786.
- [Sch00] Schumann, Heidrun und Müller, Wolfgang. 2000. *Visualisierung. Grundlagen und allgemeine Methoden*. Berlin Heidelberg : Springer, 2000.
- [Shn96] Shneiderman, Ben. 1996. The Eyes Have It: A Task by Data Type Taxonomy for Information Visualizations. *Visual Languages*. 1996.
- [Sta05] Starke, Gernot. 2005. *Effektive Softwarearchitekturen*. München : Hanser, 2005.
- [Ste03] Stephens, Matt und Rosenberg, Doug. 2003. *Extreme Programming Refactored: The Case Against XP*. Berkeley : Apress, 2003.
- [Sto07] Stone, Maureen. 2007. Expert Color Choices for Presenting Data. *StoneSoup Consulting*. [Online] 1. März 2007. [Zitat vom: 12. April 2008.] <http://www.stonesc.com/pubs/Expert%20Color%20Choices.pdf>.
- [Tuf90] Tufte, Edward R. 1990. *Envisioning Information*. Cheshire : Graphics Press, 1990.
- [Tuf83] Tufte, Edward R. 1983. *The Visual Display of Quantitative Information*. Cheshire : Graphics Press, 1983.
- [Tuf05] Tufte, Edward R. 2005. *Visual Explanations*. 7. Ausgabe. Cheshire : Graphics Press, 2005.
- [Vog05] Vogel, Oliver. 2005. *Software-Architektur: Grundlagen, Konzepte, Praxis* . München : Elsevier, 2005.
- [War04] Ware, Colin. 2004. *Information Visualization*. San Francisco : Morgan Kaufmann, 2004.

[Web01] Weber, Marc, Alexa, Marc und Müller, Wolfgang. 2001. Visualizing Time-Series on Spirals. *Proceedings of the IEEE Symposium on Information Visualization 2001*. 2001, S. 7-14.

[Wik08] 2008. Wikipedia. *Entität (Informatik)*. [Online] 5. November 2008. [Zitat vom: 5. November 2008.] [http://de.wikipedia.org/wiki/Entit%C3%A4t_\(Informatik\)](http://de.wikipedia.org/wiki/Entit%C3%A4t_(Informatik)).

[Wik081] 2008. Wikipedia. *Sonnenuhr*. [Online] 8. November 2008. [Zitat vom: 9. November 2008.] <http://de.wikipedia.org/wiki/Sonnenuhr>.

[Wik082] 2008. Wikipedia. *Sanduhr*. [Online] 21. Oktober 2008. [Zitat vom: 9. November 2008.] <http://de.wikipedia.org/wiki/Sanduhr>.

[Wik083] 2008. Wikipedia. *Wasseruhr*. [Online] 2008. Oktober 2008. [Zitat vom: 9. November 2008.] <http://de.wikipedia.org/wiki/Wasseruhr>.

[Wik084] 2008. Wikipedia. *Visualisierung*. [Online] 22. 11 2008. [Zitat vom: 28. 11 2008.] <http://de.wikipedia.org/wiki/Visualisierung>.

[Wik085] 2008. Wikipedia. *Extreme Programming*. [Online] 8. 10 2008. [Zitat vom: 10. 10 2008.] http://de.wikipedia.org/wiki/Extreme_Programming.

[Zer81] Zerubavel, Eviatar. 1981. *Hidden Rythms. Schedules and Calendars in Social Life*. Berkeley, Los Angeles, London : University of California Press, 1981.

[Zer85] Zerubavel, Eviatar. 1985. *The Seven Day Cycle. The History and Meaning of the Week*. New York : The Free Press, 1985.

[Zer03] Zerubavel, Eviatar. 2003. *Time Maps. Collective Memory and the Social Shape of the Past*. Chicago, London : The University of Chicago Press, 2003.