

# Informationsvisualisierung und Visual Analytics

**Zusammenfassung**

Fabian Damken

9. März 2022



TECHNISCHE  
UNIVERSITÄT  
DARMSTADT

---

# Inhaltsverzeichnis

---

<b>1</b>	<b>Einleitung</b>	<b>7</b>
1.1	Anwendungen von Visualisierungen . . . . .	8
1.2	Identifizierung der Visualisierungsaufgabe . . . . .	8
1.3	Negativbeispiele . . . . .	8
<b>2</b>	<b>Der Informationsvisualisierungsprozess</b>	<b>9</b>
2.1	Daten und Datenvorverarbeitung . . . . .	9
2.1.1	Datentypen und Datenstrukturen . . . . .	9
2.1.2	Vorverarbeitung . . . . .	9
2.2	Die Visuelle Abbildung . . . . .	10
2.2.1	Beispiele . . . . .	10
2.2.2	Visuelle Strukturen . . . . .	10
2.2.3	Bildunterschriften . . . . .	11
2.3	Wahrnehmung, Position und Layout . . . . .	11
2.3.1	Wahrnehmungsmodelle von Ware . . . . .	11
2.3.2	Elementare Visuelle Aufgaben . . . . .	11
2.3.3	Eigenschaften Verschiedener Visueller Channels . . . . .	11
2.3.4	(Ungewollte) Einflüsse . . . . .	11
2.3.5	Position, Layout und Komposition . . . . .	12
2.4	Interaktion . . . . .	12
2.4.1	Benutzungsschnittstellen . . . . .	12
2.4.2	Interaktionstechniken . . . . .	12
2.4.3	Design . . . . .	13
<b>3</b>	<b>Spezialisierte Visualisierungstechniken</b>	<b>14</b>
3.1	Hochdimensionale Daten . . . . .	14
3.1.1	Quantitative Daten . . . . .	14
3.1.2	Kategorische Daten . . . . .	14
3.2	Große Datenmengen . . . . .	14
3.2.1	Ordnen . . . . .	14
3.2.2	Aggregieren . . . . .	15
3.3	Zeitbasierte Daten . . . . .	15
3.3.1	Viele Zeitreihen . . . . .	15
3.3.2	Periodische Zeitreihen . . . . .	15
3.3.3	Diskrete Ereignisse . . . . .	15
3.4	Graphen und Bäume . . . . .	15
3.4.1	Bäume . . . . .	15
3.4.2	Allgemeine Graphen . . . . .	16
3.4.3	„Search, Show Context, Expand on Demand“ . . . . .	16

---

3.5	Geobasierte Daten und Karten . . . . .	16
3.5.1	Karten als Metapher . . . . .	16
3.5.2	Geobezogene Daten . . . . .	16
3.5.3	Nicht-Geobezogene Daten . . . . .	17
3.5.4	Raum-Zeit Daten . . . . .	17



---

# Abbildungsverzeichnis

---

1.1	Von Daten zu Entscheidungen . . . . .	7
-----	---------------------------------------	---



---

## Tabellenverzeichnis

---



---

## Liste der Algorithmen

---

---

# 1 Einleitung

---

In dieser Zusammenfassung werden zwei Themen behandelt: Informationsvisualisierung und Visual Analytics. Dabei sollen die Frage beantwortet werden, wie verschiedene Daten *gut* visualisiert werden können und wie Visualisierung die Analyse unterstützen können. Viele Ideen der folgenden Kapitel und grundlegender Techniken bauen dabei auf dem Verständnis grundlegender Prozesse des Gehirns ab. Denn: Eine Visualisierung soll dem Gehirn des\*der Nutzer\*in Arbeit abnehmen. Dabei sollen für jede Visualisierung die folgenden drei Fragen beantwortet werden:

- Was wird wie dargestellt?: Formale Beschreibung einer Visualisierung.
- Was ist gut und (möglichst) ohne Anstrengung sichtbar?: Prinzipien der Wahrnehmung kennen und anwenden.
- Was hilft dem\*der Nutzer\*in bei der Aufgabe?: Beschreibung und Wahrnehmungsprinzipien im Kontext einer Aufgabe bewerten.

Ein relevantes, bisher noch nicht erwähntes, Wort in den obigen Fragen ist die *Aufgabe*. Zu Beginn jeder Visualisierung muss zunächst die *Aufgabe* der Visualisierung identifiziert werden. Dies sind oftmals Entscheidungen, die (objektiv) durch Daten und Informationen getroffen werden (sollten). Dies ist in Abbildung 1.1 dargestellt. In der Praxis werden jedoch häufig einfach Visualisierungskataloge (große Datenbanken mit Visualisierungstechniken) nach einer „schönen“ Visualisierung durchsucht. Dadurch wird das Dasein der Visualisierung als Werkzeug jedoch zu dem Zweck gemacht, d. h. die Visualisierung wird ein Selbstzweck. Dies sollte eigentlich nie der Fall sein!

Oft die Aussage getroffen, dass „ein Bild mehr sagt als tausend Worte.“ Im Allgemeinen sollte allerdings eher gesagt werden, dass ein Bild etwas *anderes* als tausend Worte sagt: Sprachliche Artefakte (wozu auch Zahlen gehören), werden von dem Gehirn *bewusst* und verarbeitet und oftmals in eine zeitliche Reihenfolge gebracht. Eine Visualisierung der selben Daten hingegen ist ein bildliches Artefakt und erlaubt eine *unbewusste* Verarbeitung, bei der die Informationen im Raum strukturiert werden. Dadurch können komplexe Zusammenhänge sehr schnell vermittelt und erfasst werden.



Abbildung 1.1: Eine Visualisierung dient immer der Erfüllung einer Aufgabe und soll zu einem Erkenntnisgewinn führen. Oftmals steht am Ende davon eine Entscheidung, die objektiv durch Daten getroffen werden soll. Bei der Erstellung einer Visualisierung sollte dieser Prozess also rückwärts durchgeführt werden, d. h. ausgehend von der Frage, welche Entscheidung getroffen werden soll.

---

## 1.1 Anwendungen von Visualisierungen

---

Die Anwendung einer Visualisierung, lässt sich in zwei Kategorien einteilen: *Erfassen* und *Produzieren* von Informationen, wobei erstere durch Informationsvisualisierung und letztere durch Visual Analytics „bearbeitet“ werden.

Innerhalb der Informationsvisualisierung werden die folgenden Gruppen unterschieden:

- *Explain*: Es sollen bekannte Informationen an andere vermittelt werden (möglicherweise, aber nicht immer, interaktiv).
- *Explore*: Es sollen neue Information auf Basis von Daten gefunden oder unsichere Informationen bestätigt werden (üblicherweise sehr interaktiv; das Ziel ist nicht immer bekannt).
- *Enjoy*: Zwanglose und durch Neugier getriebene Begegnung mit den Daten; dabei ist die Aufgabe selten bekannt.

Von diesen drei Arten der Informationsvisualisierung werden in dieser Zusammenfassung vor allem die ersten beiden behandelt.

Innerhalb der Visual Analytics werden die folgenden Gruppen unterschieden:

- *Annotate*
- *Record*
- *Derive*

---

## 1.2 Identifizierung der Visualisierungsaufgabe

---

Bei der Identifizierung der Aufgabe sollte auch mit einbezogen werden,

- welche Informationen als bekannt vorausgesetzt werden,
- welche Informationen gesucht werden, und
- was mit den neuen Informationen gemacht wird oder gemacht werden soll.

Das Design der Visualisierung bestimmt damit essentiell, wie gut mit der Visualisierung gearbeitet werden kann durch Wiedererkennung bekannter Informationen und Erkennung neuer Informationen. Die erste Regel ist dabei, wie bereits erwähnt, dass die Visualisierung ein Werkzeug und kein Selbstzweck ist. Das lässt sich wie folgt zusammenfassen:

Have something to tell or something to ask!

---

## 1.3 Negativbeispiele

---



---

## 2 Der Informationsvisualisierungsprozess

---

---

### 2.1 Daten und Datenvorverarbeitung

---

---

#### 2.1.1 Datentypen und Datenstrukturen

---

Datentypen

---

Datenstrukturen

---

Datentabellen

Zeitbezogene Daten

Ortsbezogene Daten

Bewegungsdaten

Graphen und Netzwerke

Bäume und Hierarchien

---

#### 2.1.2 Vorverarbeitung

---

Metadaten und Statistik

---

Fehlende Werte und Datenbereinigung

---

Ausreißer (-detektion)

---

Normalisierung und Skalierung

---

Min-Max Normalisierung

Lokale und Globale Skalierung

---

---

**Diskretisierung**

**Sampling, Segmentierung und Untermengen**

**Datenintegration**

## **2.2 Die Visuelle Abbildung**

### **2.2.1 Beispiele**

**Scatterplot-Matrix**

**Stacked Graph**

**Dot Matrix**

### **2.2.2 Visuelle Strukturen**

**Raum**

**Marks**

**Punkt vs. Fläche**

**Channels**

**Position**

**Farbkanäle (Farbton, Helligkeit, Sättigung)**

**Länge**

**Größe und Flächeninhalt**

**Form**

**Orientierung**

**Exoten**

**Glyphen**

**Visualisierungen als Glyphen**

---

### **2.2.3 Bildunterschriften**

---

## **2.3 Wahrnehmung, Position und Layout**

---

### **2.3.1 Wahrnehmungsmodelle von Ware**

---

**Farbe und Farbmodelle**

---

**Color-Mapping**

---

### **2.3.2 Elementare Visuelle Aufgaben**

---

**Anwendersicht**

---

**Beispiel**

---

**Suche**

---

**Queries**

---

### **2.3.3 Eigenschaften Verschiedener Visueller Channels**

---

**Auswahl/Hervorhebung**

---

**Ordnung**

---

**Differenzen**

---

**Zusammenfassen**

---

### **2.3.4 (Ungewollte) Einflüsse**

---

**Kontrast und Perzeptuelle Länge**

---

**Farbnamen und Farbkategorien**

---

**Konsistente Bewegung und 3D aus Bildern**

---

**Kontrast und Kontext**

---

---

---

**Separierende und Integrierende Channels**

---

### **2.3.5 Position, Layout und Komposition**

---

**Beispiele**

---

**Zusammengesetzte Visualisierungen**

---

**Beispiele**

---

## **2.4 Interaktion**

---

### **2.4.1 Benutzungsschnittstellen**

---

**Bedienung und Interaktion nach Norman**

---

**ISO 9241**

---

**Interaktionsmodi nach Spence**

---

**Kontinuierliche Interaktion**

**Schrittweise Interaktion**

**Passive Interaktion**

**Gemischte Interaktion**

---

### **2.4.2 Interaktionstechniken**

---

**Systemnahe Interaktionstechniken**

---

**Selektion**

**Navigation**

**Shneidermans Mantra**

**Fokus und Kontext**

**Überblick und Detail**

**Brushing und Linking**

---

**Kategorien der Interaktion nach Yi et al.**

---

### **2.4.3 Design**

---

**Leitsätze**

---

**Navigation**

**Organisation**

**Erzeugung von Aufmerksamkeit**

**Unterstützung der Dateneingabe**

---

**Prinzipien**

---

**Ermittlung des fachlichen Niveaus des Nutzers**

**Ermittlung der Arbeitsaufgaben**

**Wahl des Interaktionsstils**

**Die Acht Goldenden Regeln der Gestaltung**

---

**Menschliche Reaktionszeit**

---

---

## 3 Spezialisierte Visualisierungstechniken

---

---

### 3.1 Hochdimensionale Daten

---

---

#### 3.1.1 Quantitative Daten

---

Scatterplot-Matrix

---

Starplot

---

Small Multiple Plots

---

Parallele Koordinaten

---

... mit Interaktion

---

RadViz

---

---

#### 3.1.2 Kategorische Daten

---

Parallel Sets

---

... mit Interaktion

---

Mosaic-Plot

---

KV-Map

---

Tabelle

---

---

### 3.2 Große Datenmengen

---

---

#### 3.2.1 Ordnen

---

Dimensionsreduktion und Feature Selektion

---

Principal Component Analysis (PCA)

---

Linear Discriminant Analysis (LDA)

---

Multidimensional Scaling (MDS)

---

---

Self-Organizing Map (SOM)

---

3.2.2 Aggregieren

---

## 3.3 Zeitbasierte Daten

---

3.3.1 Viele Zeitreihen

---

Filtern

---

Heatmaps

---

... mit Aggregation

---

Horizon Plots

---

Small Multiples

---

... mit Aggregation

---

3.3.2 Periodische Zeitreihen

---

Spiral Layouts

---

Matrix-Layout

---

3.3.3 Diskrete Ereignisse

---

Sequenzbaum

---

## 3.4 Graphen und Bäume

---

3.4.1 Bäume

---

Node-Link-Diagramm

---

Radiales Layout

---

TreeMaps

---

Cushions

---

Squarified

---

---

**Icicle Plot und Sunburst**

---

### **3.4.2 Allgemeine Graphen**

---

**Layouts**

---

**Force-Directed**

**Layer-Based: Sugiyama**

**Constraint-Based: Metro-Map**

---

**(Hierarchisches) Edge-Bundling**

---

### **3.4.3 „Search, Show Context, Expand on Demand“**

---

## **3.5 Geobasierte Daten und Karten**

---

### **3.5.1 Karten als Metapher**

---

**Karten und Schematisierungen**

---

### **3.5.2 Geobezogene Daten**

---

**Kartenprojektion**

---

**Plattkarte**

**Mercator Projektion**

**Winkel-Tripel**

---

**Verzerrte Darstellungen**

---

**Metro-Map**

**(Stetige) Kartogramme**

---



---

## **Abstrakte Geovisualisierungen**

---

### **3.5.3 Nicht-Geobezogene Daten**

---

**Wikipedia World Map**

---

**Themescapes**

---

**Themengebiete**

---

**Gmap World of Music**

---

**Metro-Map Immitation**

---

**Rekonstruktion von Terrain aus Knotenattribut**

---

### **3.5.4 Raum-Zeit Daten**

---

**Darstellung von Richtungen**

---

**Darstellung von Geschwindigkeiten**

---

**Darstellung von Vielen Trajektorien**

---