

# #11 Big Data

## Datenvisualisierung zwischen Big Data und Standardreporting

*Der Begriff Datenvisualisierung oder Data Visualization bezeichnet die Aufbereitung von Daten in grafischer Form. Dies klingt simpel, ist aber hochkomplex. Data Visualization vereint u. a. Wissen aus Statistik, Mathematik, Biologie, Psychologie und Design und will die kognitiven Fähigkeiten des Menschen mit seinem Wissen in der Datenexploration unterstützen.*

### Informationen sichtbar machen

Im betrieblichen Alltag kennt man häufig das Problem: Viele Daten sind grundsätzlich verfügbar, werden aber nur unzureichend genutzt. Schuld daran sind oftmals nicht etwa die fehlende Auswertung oder die unzureichende Datenqualität. Auch eine ungeeignete Visualisierung der Informationen kann verwirren und falsche Eindrücke vermitteln, obwohl die eigentliche Datenbasis korrekt ist. Dem Empfänger der Daten ist es deshalb nicht möglich, die wesentlichen Aussagen der Daten und somit das zu Entscheidende wahrzunehmen. Entwicklungen, Muster oder Abweichungen sind nicht erkennbar; Chancen und Risiken werden leicht übersehen.

### Möglichkeiten des Information Designs

Auch im Zeitalter von Big Data erfolgt die Informationsvermittlung häufig über Excel-Reports, Unternehmenspräsentationen in PowerPoint und gedruckte Geschäftsberichte.

Die zentralen Ziele der Visualisierung sind hier das Hervorheben der relevanten Informationen ohne Verzerrung der Daten und die Erhöhung der Lesbarkeit. Die auf diese grundlegenden Anforderungen ausgerichtete Teildisziplin der Datenvisualisierung wird als Information Design bezeichnet. Information Design setzt sich vorrangig mit statischen Vi-

sualisierungen auseinander und gibt Empfehlungen zu Visualisierungsformen, Layout, Farbgebung und Notation.

Die Frage nach Skalierung, Schriftgröße oder Farbgebung mag trivial klingen. Aber von einer Rot-Grün-Sehschwäche oder -Blindheit sind ca. 1 Prozent aller Frauen und ca. 9 Prozent der Männer betroffen. Insbesondere durch die Auswahl der Skalierung und der Achsenabschnitte können Grafiken in ihrer Wirkung manipuliert werden.

### Die optimale Informationsdichte

Ein weiterer Aspekt ist die Dichte der Informationen. Jeder von uns kennt umfangreiche PowerPoint-Präsentation, die in wenigen Grafiken oder Textausagen zusammengefasst werden könnten. Diese Negativbeispiele zeigen das Problem: Die Informationsdichte ist so zu gestalten, dass der Adressat nicht überfordert ist, aber sich auch nicht angesichts einer „Informationsleere“ langweilt.

Zahlreiche Unternehmen vereinheitlichen auf der Basis dieser Prinzipien ihr Berichtswesen und führen entsprechende Vorgaben und Notationskonzepte ein. Die Vorteile eines einheitlichen Information Designs sind nicht zu unterschätzen:

Zeitersparnis in der Informationsvermittlung und den Entscheidungsprozessen  
transparente und aussagekräftige

Berichtsinhalte & hohe Informationsqualität  
Kostensparnis durch Standardisierung.

Wie aber sieht eine sinnvolle Visualisierung in einem Informationssystem oder für Big Data aus? Je nach Art, Struktur und Volumen der Daten und der Zielsetzung der Berichte und Analysen sind andere Visualisierungsformen notwendig. Im Big Data-Umfeld werden verstärkt neue Analyse- und Visualisierungstools entwickelt, die die Interaktion mit dem Analysten unterstützen und seine Fähigkeiten bewusst in die Informationsermittlung einbeziehen.

*Autor: Sonja Lehmann*

### Die wichtigsten Grundsätze eines guten Information Designs sind:

- Botschaften klar zeigen und Wichtiges hervorheben
- die Wahrnehmung des Adressaten zielgerichtet auf das Wesentliche lenken
- Informationen verdichten
- gleiche Inhalte gleich darstellen
- Inhalte sinnvoll gliedern
- Darstellungsformen so einfach wie möglich wählen
- keine visuellen Elemente ohne Bedeutung
- Inhalte am Informationsbedürfnis der Adressaten ausrichten

# Informationen für Visual Thinkers

*Viele Unternehmen stehen heute vor dem Kraftakt, mit rasant wachsenden Datenmengen umzugehen. Die Antwort suchen sie vielfach in leistungsfähiger Hardware, hoch entwickelten Analysetools und effizienten Algorithmen.*

Diese Werkzeuge sind im Umgang mit der zunehmenden Datenflut unverzichtbar. Aber es gibt noch einen zentralen Erfolgsfaktor, der oft vernachlässigt wird: Der Mensch und seine Fähigkeiten wie Flexibilität, Kreativität und Neugier. Seine kognitiven Fähigkeiten ermöglichen es, relevante Informationen zu erfassen und komplexe Zusammenhänge zu analysieren. Das menschliche Auge ist das Wahrnehmungsorgan mit der höchsten Signalübertragung an das menschliche Gehirn. Der Mensch lernt quasi mit dem Auge. Auch wenn wir uns dessen oft nicht bewusst sind, die meisten Menschen sind „Visual Thinker“. Sinnvolle Visualisierungen unterstützen die visuelle Wahrnehmung. Je nach Art, Struktur und Volumen der Daten und der Zielsetzung der Berichte und Analysen sind jedoch andere Visualisierungsformen notwendig.

## Dashboards und Ad-hoc-Abfragen – Navigation durch die Daten

Viele Unternehmen setzen Business Intelligence-Tools oder Data Warehouse-Systeme für die Integration und multidimensionale Analyse von Daten ein. Die Darstellung der Daten erfolgt bei diesen Tools über automatisierte Standardreports, Ad-hoc-Abfragen und Dashboards.

Für die Visualisierung gelten weiterhin die Prinzipien des Information Design. Die Anforderungen an das Layout eines Diagramms ist unabhängig vom Medium – egal ob auf Papier oder am Bildschirm. Der wesentliche Unterschied liegt in der Nutzung der Daten. Der Adressat ist nicht mehr auf eine statische Darstellung beschränkt, sondern kann die Sicht seinen Fragestellungen anpassen.

Die freie Navigation in den Daten ermöglicht einen deutlich höheren Infor-

mationsgewinn, erfordert aber auch ein anderes Vorgehen in der Visualisierung.

Das grundlegende Prinzip lautet:

1. Überblick aller Daten
2. Zoom & Filter-Funktionalitäten
3. Details auf Anforderungen

Eine gute Benutzeroberfläche muss mehr liefern als nur die richtigen Kennzahlen, sondern auch Kennzahlen in Relation zu anderen Kennzahlen sinnvoll darstellen. Die Visualisierung soll einen Mehrwert bieten; im Idealfall wird eine Maßnahme oder Handlung empfohlen. Bei der Realisierung kann – soweit vorhanden – ein bereits bestehendes Notationskonzept für das statische Information Design erweitert werden. Bereits verwendete Darstellungselemente behalten ihr vertrautes Layout und entsprechen den Sehgewohnheiten der Nutzer.

Neue Visualisierungselemente wie Tachometer oder Ampeln sind in kompakter Form zu definieren. Spezielle Präsentationssoftware ermöglicht es, ein Notationskonzept zentral zu hinterlegen und alle Datensichten automatisch anzupassen.

## Visual Analytics – die Antwort auf Big Data

Die Menge der verfügbaren Daten wächst exponentiell. Gleichzeitig steigen die Geschwindigkeit der Datenentstehung und die Vielfalt der Datenquellen und Formate. Big Data eröffnet den Zugang zu bislang nicht verfügbaren oder nicht genutzten Daten, stellt aber auch neue Herausforderungen an die Visualisierung. Es werden keine bekannten Fragestellungen beantwortet, sondern die Daten werden erforscht, um neue Erkenntnisse zu gewinnen. Es gilt nicht nur strukturierte Daten als Grafik

zu präsentieren, sondern z. B. auch Muster in unstrukturierten Daten (z. B. Texten) zu finden.

Sogenannte Visual-Analytics bzw. Visual Intelligence-Lösungen kombinieren die Stärken digitaler und menschlicher Datenanalyse: Der Analyst wird durch automatisierte Ergebnisse und statistisch-mathematische Berechnungen entlastet; umfangreiche Visualisierungen verkürzen den Analyseprozess. Der User bestimmt dabei interaktiv die Richtung der Datenexploration. Eine sinnvolle Datenanalyse ist ohne das fachliche Know-how und die Intuition des Analysten kaum möglich.

Visualisierungen für Big Data müssen neuen Wegen folgen. Der bislang geltende Grundsatz, dem Adressaten zunächst einen Gesamtüberblick zu vermitteln, ist hier unrealistisch. Es gibt keine Möglichkeit, auf dem Bildschirm mehr Informationen als Pixel darzustellen. Ab einem bestimmten Volumen müssen Daten verdichtet, gefiltert oder vorverarbeitet werden. Neue Darstellungsformen unterstützen die interaktive Analyse durch Entscheidungsbäume oder hoch komplexe Grafiken wie Heatmaps, Gantt-Diagramme, Treemaps oder Circos-Diagramme.

Autor: Sonja Lehmann

