

M. Koch, A. Butz & J. Schlichter (Hrsg.): Mensch und Computer 2014 Workshopband, München: Oldenbourg Wissenschaftsverlag, 2014, S. 55-62.

# Visual Analytics für Smart Data

Ariane Sutor

Corporate Technology, Business Analytics and Monitoring, Siemens AG

## Zusammenfassung

Die Verfügbarkeit großer Mengen an Daten und die technischen Möglichkeiten zur Speicherung und Verarbeitung ermöglichen neue Geschäftsmodelle im industriellen Umfeld. Wichtig ist hier nicht nur die Menge der Daten, vielmehr wird durch das Zusammenbringen von Domänenwissen aus den industriellen Anwendungen sowie das richtige Datenanalyse Know-How eine intelligente Auswertung der Daten im Sinne von „Smart Data“ ermöglicht zur Entscheidungsunterstützung und zur Verbesserung der Systeme etwa hinsichtlich Betrieb, Wartung oder Effizienz. Entscheidender Erfolgsfaktor sind neben Geschäftsmodell und Daten die richtigen Technologien, zusammengefasst in einem modularen, workflow-basierten Data Analytics Framework. Dies beinhaltet insbesondere auch die geeignete Visualisierung von Analyse-Ergebnissen, um diese für den Menschen effektiv nutzbar zu machen.

## 1 Smart Data für industrielle Datenanalyse

In den vergangenen Jahren ist die Menge verfügbarer Daten förmlich explodiert; inzwischen sind wir bei etwa 3 Zettabyte angekommen und das Volumen verdoppelt sich alle zwei Jahre. Möglich wird dies durch neue Software und Hardware Architekturen, die massive Datenverarbeitung ermöglichen, etwa Cloud Installationen, massiv verteilte Berechnungen mit Hadoop und NoSQL Datenbanken. Big Data ist hier das meistverbreitete Schlagwort, es beschreibt Technologien zur Erfassung, Speicherung und Analyse großer Datenmengen. Ein weiterer wichtiger Aspekt in diesem Zusammenhang ist außerdem die Visualisierung der Analyse-Ergebnisse, die die Information für den menschlichen Nutzer erst verständlich macht.

### 1.1 Neue Geschäftsmodelle durch Smart Data

Oft wird Big Data im Kontext von Daten aus dem Internet, aus Suchmaschinen und sozialen Medien betrachtet. Aber auch immer mehr Unternehmen investieren in Big Data Technologien. Konkrete Zahlen finden sich in einer Gartner Studie (Gartner 2013) die somit zeigen, dass es sich nicht um ein leeres Schlagwort handelt. Big Data wird dabei nicht mehr nur als technologischer Trend wahrgenommen, vielmehr werden Daten auch als Wettbewerbsfaktor

erkannt. Gemäß einer Studie von BITKOM (BITKOM 2012) wird der Wert von Daten verglichen mit anderen Produktionsfaktoren wie Kapital, Arbeit, Ressourcen.

Darüber hinaus ist Big Data auch ein wichtiger Motor für Innovation und Wachstum für viele industrielle Anwendungen geworden, u.a. für industrielle Anlagen oder Kraftwerke, für Energiesysteme, Gebäude oder ganze Städte. Die Daten für industrielle Big Data Anwendungen liefern intelligente Sensoren, Geräte, Apps, oft auch bezeichnet als das Internet der Dinge. Ziel ist es letztendlich, aus den Daten Geschäftsinnovation und geschäftlichen Mehrwert zu generieren, durch verbesserte Performance, Energie-Einsparungen oder Kostenreduktionen. Um dieses Ziel zu erreichen ist es allerdings erforderlich, die Daten in ihrem Gesamt-Kontext zu verstehen. Den entscheidenden Vorteil hat hier, wer neben den Daten auch das Domänenwissen über die Anlagen und ihren Betrieb besitzt. Wichtig ist somit nicht unbedingt die Menge der Daten, sondern das umfassende Domänen und Device Know How gepaart mit Analytics. Statt Big Data kann man besser von Smart Data sprechen.

Smart Data ermöglicht in vielen Branchen neue Geschäftsmodelle. Neben der Überwachung von Anlagen und Systemen rückt die Diagnose in den Fokus, also die Analyse von Ursachen beispielsweise von Fehlern. Ferndiagnose in Echtzeit kann den Service für Anlagen komplett verändern. Ein weiterer Innovationssprung liegt in der Möglichkeit, Analytics nicht nur rückwärtsgerichtet zu nutzen. Noch wichtiger erscheinen Antworten auf die Fragen „was wird passieren?“ oder sogar „was sollen wir tun?“. Wer hier den Zugriff auf die technischen Systeme hat kann Analytics nutzen für Entscheidungsunterstützung und damit für eine verbesserte Steuerung und Betrieb der Systeme.

Eine wichtige Klasse von Beispielen solcher neuen Geschäftsmodelle sind industrielle datengetriebene Services zur Überwachung von Industrieanlagen, für vorausschauende Wartung oder zur Optimierung der Energieeffizienz. Anlagenbetreiber nutzen Analytics-as-a-Service und profitieren von Infrastruktur und Algorithmen. Die Daten bleiben dabei Eigentum des Kunden. Im Bereich Verkehr und Mobilität sind eine Vielzahl innovativer Anwendungsfälle möglich, z.B. eine Integration verschiedener Verkehrssysteme und vereinfachte Schnittstellen für Planung oder Bezahlung. Möglich wird dies durch eine genaue Echtzeit-Erfassung der Verkehrslage sowie gute kurzfristige Verkehrsprognosen. Weiter können städtische Infrastrukturen optimiert werden, z.B. durch die gemeinsame Optimierung intelligenter Stromnetze und intelligenter Gebäude mit dem Ziel der Energieeffizienz, aber auch einer guten Positionierung am Strommarkt.

Damit solche innovativen Geschäftsmodelle realisiert werden können, ist eine entsprechende Infrastruktur erforderlich sowie das richtige Set an Analytics Methoden. Im nächsten Abschnitt wird beschrieben, welche Elemente dabei wichtig sind.

## 1.2 Smart Data Analytics Framework

Damit Smart Data im industriellen Kontext ein Erfolg wird, müssen verschiedene Faktoren zusammenkommen (Zillner et al. 2014):

- Geschäftsmodelle: Untersuchung des Geschäftspotenzials bekannter und noch unbekannter Kundenbedürfnisse

- Daten: sukzessiver Aufbau von Daten und Einsicht in ihren Wert
- Technologie: Implementierung eines kosteneffizienten, flexiblen und skalierbaren Technologie-Stacks

Damit ein Smart Data Ansatz sich in der Praxis bewährt ist es sinnvoll, schrittweise vorzugehen. Der Ausbau kann erfolgen durch die Implementierung weiterer Referenzbeispiele wie im letzten Abschnitt beschrieben. Darüber hinaus gibt es weitere wichtige Rahmenbedingungen, etwa das Smart Data Ökosystem, in dem mit Partnern Daten oder Technologien fallweise geteilt werden. Gerade die Fusion von Daten verschiedener Owner kann einen Mehrwert darstellen, beispielsweise im Falle Mobilität die Kombination von städtischen Infrastrukturdaten wie Ampelanlagen mit Floating Car Daten.

Fundamental sind außerdem rechtliche Fragen, vor allem hinsichtlich Privacy der Daten; hier können u.a. Anonymisierungsmethoden zum Einsatz kommen. Ferner ist integrierte Security sowie Schutz der Daten, Algorithmen und Modelle wesentlich. Schließlich ist es erforderlich, die richtigen Kompetenzen zusammenzubringen, insbesondere aus Computer Science, Mathematik, Ingenieurwesen und Wirtschaftswissenschaften.

In technischer Hinsicht umfasst Big Data alle Technologien, die die Basis schaffen zur Verarbeitung großer und komplexer Datenmengen in vernünftiger Zeit, einschließlich Erfassung, Verarbeitung, Speicherung, Analyse und Visualisierung der Daten (NESSI, 2012). Wesentliche Basis für Smart Data im industriellen Kontext sind ebenfalls die technologischen Bausteine. Wichtige Merkmale sind hier ein modularer und Service-orientierter, Workflow-basierter Aufbau. Außerdem müssen verschiedene Betriebsarten unterstützt werden, insbesondere Cloud-basiert und On-Premise. Die folgende Abbildung zeigt wesentliche Elemente eines industriellen Data Analytics Frameworks.

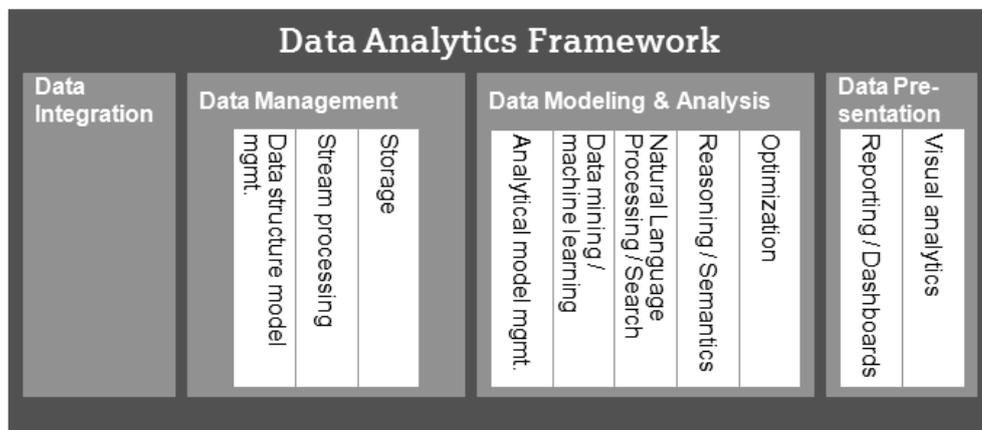


Abbildung 1: Data Analytics Framework

Erfahrungen in der Umsetzung von Referenzbeispielen zeigen, dass Anwendungsfälle und Geschäftsmodelle es erfordern, die Datenanalyse im Einzelfall speziell anzupassen aufgrund

der verschiedenen Datentypen, Menge, Qualität, Quellen, Modelle. Eine einzige Standardlösung ist nicht ausreichend. Dies kann es auch erforderlich machen, neue Technologien zu entwickeln, sowohl hinsichtlich Infrastrukturen als auch hinsichtlich vertikaler Algorithmen, die die Sensor- und Analytics-Ebene eng verknüpfen.

In diesem Beitrag werden nicht alle Elemente des Frameworks erläutert, sondern wir konzentrieren uns auf die Datenpräsentation und Visual Analytics.

## 2 Mensch-Computer-Schnittstellen und Visual Analytics für Smart Data

Die Visualisierung der Daten und Analyseergebnisse ist von fundamentaler Bedeutung, damit Menschen Smart Data effektiv nutzen können. Eine geeignete Visualisierung kann deshalb der Schlüssel sein, um große Datenmengen handhabbar zu machen. Wesentliche Aufgaben der Visualisierung sind gemäß (Shneiderman, 1996) Überblick, Zoom, Filter, Details-on-Demand, Zusammenhänge, Historie und Extraktion. Allerdings können die Visualisierungen abhängig von ihrer Komplexität aufwendig in der Berechnung sein mit entsprechenden Einschränkungen für eine interaktive Nutzung.

Visual Analytics umfasst im weiteren Sinn den kompletten Smart Data Workflow inklusive Datenerfassung, Speicherung und Analyse. In (Keim et al., 2010) wird Visual Analytics definiert als multidisziplinäres Gebiet, das verschiedene Forschungsbereiche kombiniert, insbesondere Visualisierung, Mensch-Computer-Interaktion, Datenanalyse, Datenmanagement, räumlich-geographische und zeitliche Analyse, Entscheidungsunterstützung und Statistik. Hier wird in einem engeren Sinn insbesondere die interaktive Kombination von Datenanalyse und Visualisierung damit bezeichnet.

Beim Versuch, komplexe Informationen darzustellen, multiplizieren sich rasch die Schwierigkeiten. Deshalb ist es wichtig, dass solche Darstellungen menschlich sind in dem Sinn, dass sie sich an menschliche Bedürfnisse und Schwächen anpassen. Kontextbezug, zeitliche Anforderungen und der richtige Informationsumfang sind hier wesentlich. Visual Analytics hat das Ziel, die Stärken von Menschen und Computern zu verbinden, nämlich Geschwindigkeit und Genauigkeit auf der einen Seite, Intuition und Intelligenz auf der anderen, um damit Einsichten in große Mengen von Daten zu ermöglichen und richtige Entscheidungen abzuleiten. Insbesondere menschliche Faktoren wie Wahrnehmung, Darstellung, Interaktion, Auswertung und Erkenntnis spielen eine wichtige Rolle in der Kommunikation zwischen Mensch und Computer sowie im Entscheidungsprozess.

Die folgende Abbildung zeigt den interaktiven Ansatz von Visual Analytics. Dabei geht es nicht nur um die Darstellung bekannter Sachverhalte aus den Daten. Vielmehr werden die Daten explorativ ausgewertet und über die Visualisierung wird die Modellbildung direkt beeinflusst.

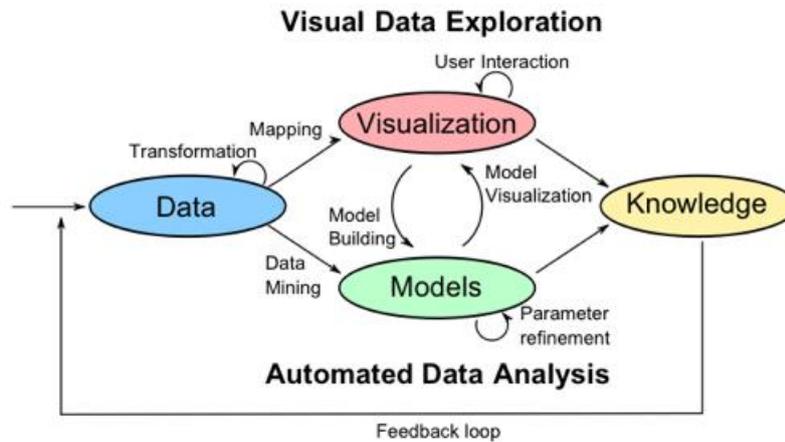


Abbildung 2: Visual Analytics Ansatz (Keim et al., 2010)

Folgende Elemente sind wesentlich für einen effektiven Visual Analytics Ansatz im Kontext von Smart Data:

- Ein Modell für die Analyse von Ursache und Auswirkungen sowie die Beschreibung der bekannten und unbekannt Einflüsse; dieses wird entwickelt mit Hilfe von Simulation und Optimierung unter Verwendung von Echtzeit-Daten des Systems
- Korrelation von Daten aus verschiedenen Quellen, die die statischen und dynamischen Eigenschaften des Systems sowie die Umgebung erfassen; hierdurch werden weitere Erkenntnisse ermöglicht durch Machine Learning, Data Mining und Data Discovery Analyse
- Kombination menschlicher Erkenntnis und Rechnerleistung mithilfe von Visual Analytics

## 2.1 Visual Analytics Use Cases

Visual Analytics muss Antworten geben auf die wesentlichen Herausforderungen im Kontext von Smart Data. Große Mengen von Daten erfordern die Visualisierung und Analyse hochdimensionaler Daten und die Darstellung vielschichtiger und verknüpfter Graphen, in wichtigen Anwendungsfällen interaktiv und dynamisch verknüpft. Auch die Geschwindigkeit der Daten erhöht die Anforderungen an dynamische Aktualisierungen der Darstellung im Vergleich zu herkömmlichen statischen Visualisierungen. Die Vielfalt der Daten schließlich erfordert vielfältige Charts für Cockpits und Dashboards.

Anwendungsfälle für Visual Analytics finden sich in allen Industrie Sektoren, insbesondere im Bereich Automatisierung und Prozessindustrie, im Bereich intelligenter Städte, Transport und Mobilität sowie im Gesundheitsbereich. Dashboards und Cockpits sind wichtige Ausprägungen, um Visual Analytics Ergebnisse darzustellen.

### **Dashboards für den Betrieb**

Dashboards sind inzwischen weit verbreitet. Ihr Ziel ist es, durch die Visualisierung Monitoring und Steuerung für verschiedene Prozesse zu unterstützen – von der Fertigung bis zu Businessprozessen oder der Darstellung der Zustände technischer Systeme und Anlagen.

Ein Dashboard für die Überwachung einer technischen Anlage umfasst in der Regel vielfältige Aspekte und Darstellungen. Ein Merkmal kann ein geographischer Überblick sein mit Zoom-Möglichkeit in die Anlage und bis zu einzelnen technischen Komponenten. Außerdem sind in der Regel verschiedene Charts enthalten, die eine Darstellung aktueller technischer Werte abbilden. Erweiterungen können Vorhersagen sein für zukünftige Entwicklungen der Anlage. Besonders wertvoll sind außerdem Vorhersagen mit Handlungsempfehlungen, z.B. für einen verbesserten Service durch vorausschauende Wartung. Durch Dashboards für den Betrieb erhalten Anlagenbetreiber einen Überblick über ihre Komponenten und Anlagen. Smart Data unterstützt hier die Schaffung neuer datengetriebener Services.

Für Smart Grids können Dashboards wesentliche Informationen zusammenfassen. Eine geographische Darstellung kann kombiniert werden mit Charts über die verschiedenen Zustandsgrößen des Stromnetzes. Im Falle von Störungen und Ausfällen kann diese Information korreliert werden mit weiteren Datenquellen, z.B. Wetterdaten. Solche Korrelationen erlauben dann, mithilfe von Visual Analytics Root Cause Analysis zu unterstützen. Außerdem können Vorhersagen gebildet werden, um die operative Planung zu verbessern. Durch die Analyse längerfristiger Trends kann dies sogar ein wichtiger Input werden für strategische Planung. Ein wichtiger Trend ist hier das Zusammenführen von operativer und strategischer Planung.

Ein besonderer Aspekt ist die zunehmende Verfügbarkeit räumlicher Daten, die Visual Analytics mit geographischem Fokus ermöglicht. Etwa im Bereich städtischer Mobilität können Datenquellen wie Smartphones, Floating Car Data, Sensoren in der Verkehrsinfrastruktur und aus dem öffentlichen Nahverkehr genutzt werden, um eine genaue Darstellung der aktuellen Verkehrslage für verschiedene Verkehrssysteme zu generieren. Dies wiederum kann die Umsetzung zahlreicher innovativer Use Cases unterstützen, von der inter-modalen Routenplanung bis zur Verkehrssteuerung.

### **Management Cockpits**

Management Cockpits fassen multi-dimensionale Aspekte eines komplexen Systems zusammen. Entscheidend ist hier die Verdichtung und Verknüpfung der wesentlichen Informationen, um einen Überblick über die aktuelle Situation sowie Handlungsempfehlungen zu geben.

Für Smart Cities kann ein Cockpit vielfältige Funktionen erfüllen. Zum einen gibt es einen Überblick über die verschiedenen städtischen Infrastrukturen wie Verkehr, Transport, Stromnetz, Gebäude, Wassernetz. Wichtige übergreifende Indikatoren können ermittelt werden, z.B. CO2 Emissionen. Durch eine Zusammenfassung relevanter Bereiche lassen sich außerdem Investitionsentscheidungen transparent gestalten.

Ein weiterer Anwendungsfall für ein Management Cockpit ist die Zusammenfassung wesentlicher Kennzahlen z.B. für eine Flotte von Anlagen und Systemen oder für globale kaufmännische Kennzahlen.

In den beschriebenen Anwendungsfällen ist die Visualisierung der entscheidende Schritt der Aufbereitung der Ergebnisse für den menschlichen Nutzer. Visual Analytics ist damit auch oft der Türöffner für neue datengetriebene Geschäftsmodelle.

## 2.2 Visual Analytics im industriellen Einsatz, Zusammenfassung und Trends

Visual Analytics ermöglicht eine interaktive, explorative Datenanalyse. Dabei werden die neuen Möglichkeiten konsequent weiterverfolgt. Trends sind u.a.

- Interaktive, visuell unterstützte Abfragen, visuelle Exploration, multi-modale Interaktion
- Große Bandbreite verschiedener Graphen-Typen, Animationen, mit Bildern angeereicherte Darstellungen
- Personalisierung und Anpassung an den Nutzer
- Semi-Automatisierung, Alarmer, CEP (Complex Event Processing)
- Zeitlich-räumliche Daten, GIS, Near Real-Time Visualisierungen
- Einbindung semantischer Daten und deren Nutzung zur Kontrolle der Visualisierungen

Für die industrielle Anwendung von Visual Analytics ist es insbesondere wichtig, die richtige Methodik und die richtigen Tools und Algorithmen zu identifizieren und zu kombinieren. Dabei müssen die Anforderungen an Visual Analytics vorab evaluiert werden. Visual Analytics kann eine sinnvolle Ergänzung und Alternative sein zu traditioneller Datenanalyse und Visualisierung und idealerweise Zeit und Kosten reduzieren.

Gerade für den industriellen Einsatz ist die richtige Auswahl an Tools entscheidend. Eine Reihe von Tools am Markt bietet fertige Lösungen (Weber et al. 2012). In speziellen Anwendungsfällen können jedoch maßgeschneiderte Ergänzungen erheblichen Mehrwert liefern. Eine Toolbox und wiederverwertbare Module können hier als Best Practice betrachtet werden.

### Literaturverzeichnis

BITKOM (2012). Big Data im Praxiseinsatz – Szenarien, Beispiele, Effekte. Leitfaden des BITKOM AK Big Data. Online: [http://www.t-systems.de/loesungen/leitfaden-bitkom-big-data-2012/1014046\\_1/blobBinary/Leitfaden\\_Bitkom\\_Big\\_Data\\_2012\\_online.pdf](http://www.t-systems.de/loesungen/leitfaden-bitkom-big-data-2012/1014046_1/blobBinary/Leitfaden_Bitkom_Big_Data_2012_online.pdf)

Gartner (2013). Survey Analysis: Big Data Adoption in 2013 Shows Substance Behind the Hype online: <http://www.gartner.com/newsroom/id/2593815>

- 
- Keim D. A., Kohlhammer J., Ellis G. (2010). Mastering the Information Age: Solving Problems with Visual Analytics, Eurographics Association.
- NESSI (2012). Big Data – A new world of opportunities. NESSI White Paper, Online: [http://www.nessi-europe.com/Files/Private/NESSI\\_WhitePaper\\_BigData.pdf](http://www.nessi-europe.com/Files/Private/NESSI_WhitePaper_BigData.pdf)
- Shneiderman, B. (1996). Visual Languages, Proceedings., IEEE Symposium, 336-343.
- Weber S., Mittelstadt S., Stoffel A., Keim D. et Al. (2012). Visual analytics for the Big Data era - A comparative review of state-of-the-art commercial systems, In VAST '12 Proceedings of the 2012 IEEE Conference on Visual Analytics Science and Technology (VAST), pages 173-182, IEEE Computer Society Washington, DC, USA ©2012
- Zillner, S., Rusitschka, S., Skubacz, M. (2014). Whitepaper The Big Data Story. Preprint.