

Abstract

Titel: Big Data Identifikation, Klassifikation und Nutzenbewertung von Anwendungsfällen

Kurzzusammenfassung: In den letzten Jahren hat sich ein regelrechter Hype um das Thema Big Data gebildet. Hauptaugenmerk der Arbeit ist die praktische Anwendung von Big Data im Unternehmen. Dabei wird aufgezeigt, welchen Nutzen Big Data im Unternehmen stiften kann. Weiter werden die wichtigsten Voraussetzungen im Unternehmen beleuchtet. Einige Anwendungsfälle werden dargelegt und anschliessend klassifiziert. Weiter wird eine mögliche Vorgehensweise und Strategie für Big Data Projekte vorgestellt. Experteninterviews sollen weitere Erkenntnisse der praktischen Anwendung von Big Data liefern.

Verfasser: Mattia Betti

Herausgeberin: lic. oec. HSG Pascale Baer-Baldauf, IPM-FHS

Publikationsformat:

- BATH
- MATH
- Semesterarbeit
- Forschungsbericht
- Anderes

Veröffentlichung (Jahr): 2015

Sprache: Deutsch

Zitation: Betti, M. (2015). *Big Data: Identifikation, Klassifikation und Nutzenbewertung von Anwendungsfällen*. FHS St.Gallen Hochschule für angewandte Wissenschaften: Bachelor Thesis.

Schlagwörter (3-5 Tags): Big Data, Anwendungsfälle, Klassifikation

Ausgangslage

Um das Thema Big Data hat sich ein regelrechter Hype gebildet. Immer wieder tauchen Artikel über Big Data in den Medien auf. Viele Unternehmen suchen deshalb Rat, ob und wie Sie nun Big Data nutzen sollen.

Ziel

Das Ziel dieser Bachelor Thesis ist es, einen Überblick über Big Data zu erhalten. Dabei sollen praktische Anwendungsfälle vorgestellt und kategorisiert werden. Verschiedene Aspekte von Big Data im Unternehmen sind aufzuzeigen unter anderem: Voraussetzungen, Herangehensweisen, Strategien, Vorteile und Risiken, Stand der Technik, Datenakquisition und Datenspeicherung. Im Zusammenhang mit Big Data werden vielfach Begriffe wie Data Mining, Business Intelligence, Data Warehouse und Predictive Analytics erwähnt. Deren Bedeutung und Zusammenhang mit Big Data ist aufzuzeigen. Ein weiteres Ziel ist es, einen Überblick über den Datenschutz im Bereich Big Data zu geben. Zuletzt soll ein persönliches Fazit des Verfassers die Thesis abrunden.

Vorgehen

Eine fundierte Literaturrecherche bildete den ersten Schritt dieser Bachelor Thesis. Die meisten Informationen stammen aus Büchern, Fachzeitschriften, Dissertationen sowie vertrauenswürdigen Webseiten. Zweitens sind möglichst heterogene Anwendungsfälle von Big Data gesucht worden, um die vielfältigen Einsatzmöglichkeiten aufzuzeigen. Eine Auswahl an Anwendungsfällen wurde beschrieben und in einer vom Autor erstellten Klassifikation eingeteilt. Zusätzlich sind mehrere Experteninterviews geführt worden, um in den jeweiligen Big Data Projekten vertiefte Informationen zu erhalten. Die Erstellung eines Forschungsdesigns und darauf aufbauend eines Leitfadens, war für die Experteninterviews essentiell. Aus all den gewonnenen Informationen ist ein Fazit über das Thema Big Data und deren Anwendung im Unternehmen gezogen worden.

Erkenntnisse

Um überhaupt mit Big Data arbeiten zu können, muss ein generell gültiges Verständnis von Big Data bestehen. Es gibt sehr viele verschiedene Definitionen von Big Data. Der Verfasser hat sich für die Definition des Bundesamts für Informationswirtschaft, Telekommunikation und neue Medien e.V. [BITKOM] (2012, S. 7) entschieden. Diese ist seiner Meinung nach äusserst treffend und wird deshalb als Grundlage für die Thesis verwendet:

Big Data bezeichnet die wirtschaftlich sinnvolle Gewinnung und Nutzung entscheidungsrelevanter Erkenntnisse aus qualitativ vielfältigen und unterschiedlich strukturierten Informationen, die einem schnellen Wandel unterliegen und in bisher ungekanntem Umfang anfallen. Big

Data stellt Konzepte, Methoden, Technologien, IT-Architekturen sowie Tools zur Verfügung, um die geradezu exponentiell steigenden Volumina vielfältiger Informationen in besser fundierte und zeitnahe Management-Entscheidungen umzusetzen und so die Innovations- und Wettbewerbsfähigkeit von Unternehmen zu verbessern.

In dieser Thesis wird Big Data durch vier Merkmale charakterisiert: Volume, Velocity, Variety und Veracity. Volume steht für das Sammeln, die Verarbeitung und die Analyse von enormen Mengen an Daten. Das zweite Merkmal Velocity bezieht sich auf die rasante Geschwindigkeit, in der die Daten kreiert, verarbeitet und analysiert werden können. Variety ist die dritte Charakteristik und steht für die Verarbeitung verschiedenster Datentypen und Datenquellen. Das letzte Merkmal Veracity zielt auf die inhärente Ungewissheit von Datentypen bezüglich ihrer Verlässlichkeit. (IBM, 2012, S. 4–5)

Es hat sich herausgestellt, dass die diversen Trendthemen wie Business Intelligence, Data Analytics, Data Warehouse, Data Mining und Predictive Analytics miteinander im Zusammenhang stehen und je nach Autor verschieden eingegrenzt und definiert werden. Innerhalb der Thesis wurden die Begriffe deshalb präzisiert und abgegrenzt. So ist zum Beispiel Data Warehouse, Predictive Analytics und Data Mining ein Bestandteil von Business Intelligence. Data Mining und Predictive Analytics sind gleichzeitig auch Teile von Big Data. Eine grundsätzliche Unterscheidung zwischen Business Intelligence und Big Data ist durchgeführt worden. Die wesentlichsten Unterschiede sind in der Tabelle 1 zu sehen.

	Business Intelligence	Big Data
Verwendung unstrukturierter Daten	Nein	Ja
Datenmenge	Klein	Hoch
Verarbeitung	Batch	Bis zu Realtime
Zeitaspekt	Vergangenheit	Zukunft
Datenhaltung	Zentralisiert (Z.B. Data-warehouse)	Dezentral
Probleme mit dem Datenschutz	Kaum	Häufig
Ziel	Verstehen was passiert ist	Neue Möglichkeiten schaffen

Tab. 1: Business Intelligence vs. Big Data.
Quelle: angelehnt an diverse Quellen (BITKOM, 2012, S. 23–27; Revolutionanalytics, 2013).

Jedes Projekt hat seine eigenen Vor- und Nachteile, die mit der Nutzung von Big Data entstehen können. Im Allgemeinen kann aber von sechs Hauptvorteile für ein Unternehmen gesprochen werden:

- Optimierung und Erschaffung von neuen Geschäftsprozessen
- massive Aufwertung der Kundenkenntnisse
- schnellere und exaktere Vorhersagen verbessern die Entscheidungsfindung
- frühere Erkennung und Vorhersage von Marktveränderungen
- positive Auswirkungen auf die Time-to-Market Dauer
- Verbesserung von Produkten und Dienstleistungen oder komplette Neuentwicklungen
- Vereinfachung der IT-Infrastruktur

Big Data birgt auch einige Risiken für Unternehmen. Das sind zum einen Data-Compliance-Risiken. Diese sind vor allem bei internationalen Projekten eine grosse Herausforderung. In die Compliance wurde auch das Sozialrisiko eingeordnet. Das Sozialrisiko ist wohl das Risiko, das den meisten Unternehmen am ehesten in den Sinn kommt, wenn von Big Data die Rede ist. Dabei handelt es sich um die Nutzung von Personen und Social Media Daten. Dies wird von der Allgemeinheit öfters kritisch beobachtet und muss deshalb genau geklärt werden. Eine weitere Herausforderung ist das Datenrisiko, darunter ist der Erhalt der Datenqualität und Datensicherheit zu verstehen. Zuletzt sind Risiken, wie beispielsweise Fehlinterpretationen, falsche Modellauswahlen, etc. zu beachten. (BITKOM, 2012, S. 14)

In der Thesis haben sich gewisse Grundvoraussetzungen herauskristallisiert, die nötig sind, damit Big Data im Unternehmen sinnvoll eingesetzt werden kann. Das womöglich wichtigste Kriterium ist die vorhandene Datenmenge; ohne genügend Daten gibt es kein Big Data. Weiter muss ein klares Ziel mit Big Data anvisiert werden. Big Data ist kein Selbstzweck. Die Managementunterstützung ist von kritischer Relevanz für eine erfolgreiche Projektumsetzung. Weiter muss die bestehende Infrastruktur die Ressourcen besitzen, um mit den Anforderungen an die Datenmenge, die Geschwindigkeit und die Zuverlässigkeit schritthalten zu können. Es sollte sichergestellt sein, dass das nötige Knowhow vorhanden ist, um Big Data betreiben zu können. Theoretisch könnte in jedem Unternehmen Big Data eingesetzt werden, aber ob das Unternehmen in der Lage ist, davon zu profitieren, ist fraglich. (BITKOM, 2014, S. 136-140)

Der Autor hat mit fünf Experten leitfadengestützte Interviews durchgeführt. Namentlich mit der Suva, der Swisslos Landeslotterie, der Kantonspolizei St. Gallen, SRF Bewegt und Nestlé. In den durchgeführten Interviews gaben alle Experten an, dass Ihnen durch die Nutzung von Big Data keine eigentlichen Nachteile erwachsen sind und sie von den Projekten überzeugt sind. Vielfach besteht sogar der Wunsch nach einem Ausbau des Big Data Projektes. Dies spricht klar für Big Data. Die meisten befragten Experten haben im

Gegensatz zu dem in der Literatur empfohlenen Projektvorgehen eher zu weniger detaillierten Projektabläufen tendiert. Bei allen befragten Experten ist für Big Data nicht explizit eine Strategie entwickelt worden, sondern das Projekt wurde in die bestehende IT-Strategie eingebettet.

Bestehende Klassifikationen von Anwendungsfällen waren nach der Meinung des Verfassers zum Teil zu wenig trennscharf. Ausserdem bieten viele Klassifikationen keine nennenswerten Vorteile. So hat sich der Verfasser für eine eigene Einteilung entschieden, welche nach dem Einsatzzweck gegliedert ist – wie in der Abbildung 1 zu sehen ist. Jedes Big Data Projekt begründet seinen Hauptnutzen auf einem dieser Faktoren und nach diesem Nutzen soll es eingeordnet werden.

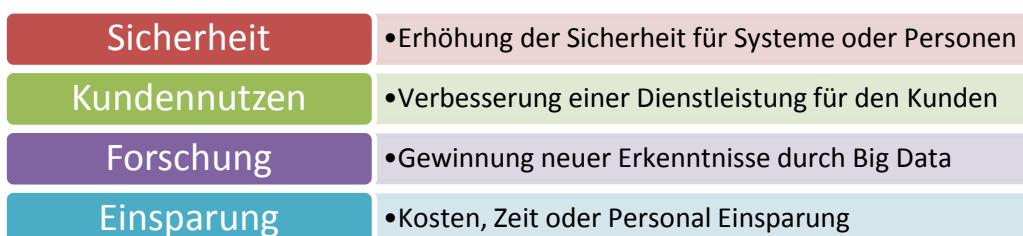


Abb. 1: Eigene Klassifikation der Anwendungsfälle.
Quelle: Eigene Darstellung.

Das Fazit dieser Thesis ist, dass Big Data für viele Unternehmen, in verschiedensten Bereichen, einen grossen Nutzen stiften könnte. Big Data findet heute bereits in diversen Branchen Einzug und erfüllt dort eine Vielfalt an Aufgaben. Bevor ein Unternehmen Big Data verwenden kann, soll sich das Unternehmen vertieft mit der Thematik befassen und sich gut vorbereiten, um mögliche Risiken zu minimieren. Durch die hohe Komplexität von Big Data Projekten ist es ratsam, Experten zu Rate zu ziehen. Dies kann die Wahrscheinlichkeit eines Fehlschlags erheblich reduzieren.

Literaturquellen

Bundesamt für Informationswirtschaft, Telekommunikation und neue Medien e.V. [BIT-KOM]. (2014). *Big-Data-Technologien – Wissen für Entscheider*. Berlin: Autor.

Bundesamt für Informationswirtschaft, Telekommunikation und neue Medien e.V. [BIT-KOM]. (2012). *Big Data im Praxiseinsatz – Szenarien, Beispiele, Effekte*. Berlin: Autor.

IBM. (2012). *Analytics: The real-world use of big data. How innovative enterprises extract value from uncertain data*. New York: Autor.

Revolutionanalytics. (15.05.2013). *Statistics vs Data Science vs BI*. Gefunden am 17.05.2015 unter <http://blog.revolutionanalytics.com/2013/05/statistics-vs-data-science-vs-bi.html>