



Hochschule Reutlingen
Reutlingen University

T-Systems

Big Data Report | Teil 1

Perspektiven von Big Data für IT, Marketing und Vertrieb

Eine Studie des Research Lab for Digital Business an der Hochschule Reutlingen
in Kooperation mit der T-Systems Multimedia Solutions GmbH



Impressum

Herausgeber: Reutlingen University
Research Lab for Digital Business
Alteburgstraße 150 | 72762 Reutlingen
Tel.: +49 7121 271 4100 | Fax: +49 7121 271 90 4100
www.reutlingen-university.de

Ansprechpartner: Prof. Dr. Alexander Rossmann
Tel.: +49 7121 271 4100
alexander.rossmann@reutlingen-university.de

Autoren: Prof. Dr. Alexander Rossmann
Heiko Bonhorst
Tim Kornherr

Kooperationspartner: T-Systems Multimedia Solutions GmbH
Julia Kunert
Tel. +49 351 2820 2543

Bilder: © fotolia.com

Vorwort

„Big Data – kaum ein anderes Schlagwort in Management, Marketing und IT hat in den letzten Jahren so viel Aufmerksamkeit erhalten. In Forschung und Praxis wird intensiv über die Nutzung großer Datenmengen für das Business der Zukunft diskutiert.

Dabei zeigen sich spannende Herausforderungen: Es gibt kaum Standardlösungen und Big Data verlangt eine deutlich intensivere Zusammenarbeit zwischen Business und IT. Gleichzeitig befinden sich die Technologien für Big Data selbst im Umbruch – und die digitale Transformation schreitet mit erheblicher Dynamik fort. Beispielsweise werden durch das Internet der Dinge völlig neue Geschäftsmodelle und Services ermöglicht. Der Druck mit Hinblick auf eine rasche und wirkungsvolle Integration dieser Möglichkeiten nimmt weiter zu.

Kurz und gut: Big Data bietet viele Chancen. Die Umsetzung dieser Chancen ist jedoch mit erheblichen Herausforderungen verbunden. Daher sind weitere Analysen zur aktuellen Wissensbasis rund um Big Data aus theoretischer und praktischer Sicht erforderlich. In dieser Hinsicht leistet der Big Data Report einen wertvollen Beitrag. Die vorliegende Studie reflektiert nicht nur den aktuellen Stand der Forschung. Sie gibt auch Einblicke in die Sichtweisen und Anwendungsfelder der Unternehmenspraxis aus IT-, Marketing- und Vertriebsperspektive. Die empirischen Daten zur Umsetzung von Big Data werden durch qualitative Interviews und fundierte Fallstudien ergänzt.

Sicher lassen sich damit nicht alle Fragen beantworten. Der Report bietet jedoch einen hervorragenden Überblick zum Stand der Dinge und inspiriert den erforderlichen Dialog zu Big Data Lösungen der Zukunft. In diesem Sinne wünschen wir den Lesern eine spannende Lektüre, viele Impulse und Erfolg für eigene Projekte und Lösungen in diesem spannenden Themenbereich“.

Prof. Dr. Alexander Rossmann

Research Lab for Digital Business
Reutlingen University, September 2015

Inhalt

1. Big Data Report	7
1.1. Forschungsfragen	8
1.2. Methodischer Ansatz	9
2. Big Data: Aktueller Stand der Forschung	10
2.1. Literaturanalyse	12
2.2. Treiber für Big Data	14
2.3. Begriffliche Grundlagen	16
2.4. Forschungsprogramme und Frameworks	17
3. Erste Teilstudie:	
Ergebnisse der Befragung von IT-Executives	24
3.1. Begriff und strategische Bedeutung von Big Data	26
3.2. Ziele und Strategien für Big Data	32
3.3. Daten und Technologien	36
3.4. Herausforderungen für die IT	43
3.5. Erforderliche Ressourcen und Kompetenzen	46
3.6. Zusammenfassung der Ergebnisse	50
4. Fallstudien:	
Einblicke in Big Data Lösungen der Unternehmenspraxis	54
4.1. SBB Serviceoptimierung	56
4.2. Planung und Risikosteuerung von Ölplattformen	59
4.3. Smart Energy Stadt Berlin	61
4.4. Smart City Monaco	63
4.5. Predictive Maintenance in der Automobilindustrie	65
4.6. Industrie 4.0 für Chemie und Pharma	67
4.7. Connected Car eCall Service	69
5. Fallstudienvergleich und Implikationen	71
5.1. Cross-Case Analyse	72
5.2. Theorie, Befragung und Fallstudien im Vergleich	75
5.3. Implikationen für Big Data Lösungen	77
6. Fazit und Ausblick	80
7. Literatur	81

Management Summary

Der vorliegende Report bietet wertvolle Erkenntnisse, Hintergründe und Fallbeispiele zur Diskussion von Big Data in Forschung und Praxis. Dabei werden zunächst die theoretischen Grundlagen für Big Data analysiert. Inzwischen ist der Dialog rund um Big Data selbst durch große Datenmengen geprägt. Global wurden bisher über 700.000 Texte zum Thema verfasst. Alleine in 2015 kamen knapp 100.000 weitere Texte hinzu.

In der Forschung ist die Diskussion rund um Big Data v.a. durch die Gestaltung von Referenzmodellen geprägt. Diese zeigen auf, über welche Themenbereiche bei der Gestaltung von Big Data zu sprechen ist und wie diese in Bezug zueinander stehen. In Deutschland ist inzwischen v.a. die Architektur des Branchenverbandes bitkom bekannt. Diese definiert ein Referenzmodell mit sechs Bausteinen. Hierzu zählen die Teilbereiche Datenhaltung, Zugriff, Integration, Analyse, Verarbeitung, Darstellung, Governance und das Thema Datensicherheit. Global werden jedoch auch andere Referenzmodelle diskutiert, die einen stärkeren Fokus auf Services und die Entwicklung von Geschäftsmodellen legen.

Die Diskussion zu Referenzmodellen wird an dieser Stelle durch eine empirische Befragung von IT-Executives ergänzt. Hier geben 87,5% der befragten Unternehmen an, dass man sich weniger als drei Jahre mit Big Data auseinandersetzt. Dennoch ist Big Data bereits für rund die Hälfte der Befragten ein wichtiges Thema der IT-Strategie. Für 79,2% der befragten Executives wird Big Data noch weiter an Bedeutung gewinnen. Während knapp 25% der Unternehmen bereits Big Data Projekte gestartet haben, sind diese erst bei 8,3% der Befragten abgeschlossen. Und fast zwei Drittel der Unternehmen sind weiterhin unentschlossen.

Die empirischen Ergebnisse weisen darauf hin, dass sich Big Data noch in der Startphase befindet. Eine zentrale Frage des Big Data Reports ist daher, wie sich Unternehmen auf die Entwicklungen im Big Data Umfeld vorbereiten, welche Fallbeispiele für erfolgreiche Projekte bereits vorhanden sind und was Unternehmen daraus lernen können.

Aus heutiger Sicht lässt sich ein schwacher Reifegrad bei der Entwicklung und Umsetzung von Big Data Strategien ausmachen. Die Anforderungen an die IT sind hoch, da sich parallel zur Umsetzung businessrelevanter Strategien die IT Plattformen als solche verändern. Dies erfordert einen Wandel auf mehreren Ebenen sowie umfassende Ressourcen zur Gestaltung der digitalen Transformation.

Dabei zeigt der vorliegende Report anhand von sieben Fallstudien auf, dass Erfolg mit Big Data bereits heute möglich ist. Dafür sind jedoch neue Strukturen der Zusammenarbeit zwischen unterschiedlichen Wertschöpfungspartnern erforderlich. Big Data Projekte werden dabei häufig auf einer neuen und spezifisch für das Projekt aufgesetzten IT-Plattform entwickelt. Eine Integration in das Kerngeschäft steht noch aus. Darüber hinaus lässt sich aus den Fallstudien ableiten, dass Big Data eine deutlich tiefere Zusammenarbeit zwischen unterschiedlichen Businessseinheiten und der IT erfordert. Big Data ist in dieser Hinsicht eine unternehmensweite Herausforderung.

Durch die Kombination aus Theorie, Praxisbefragung und Fallstudien werden acht Prinzipien für die weitere Entwicklung von Big Data formuliert. Diese greifen sowohl technologische als auch betriebswirtschaftliche Fragestellungen auf. Rund um die Themengebiete Serviceorientierung, Mehrwerte, Architektur, Ökosysteme, Integration, Strategie, Sicherheit und Kompetenz werden damit wesentliche Leitfragen für die weitere Entwicklung skizziert – und anhand von Einzelfallbeispielen zumindest teilweise beantwortet.

1. Big Data Report

Herman Hollerith gilt als Wegbereiter der modernen Datenverarbeitung. Er ist als Sohn deutscher Einwanderer 1860 in Buffalo geboren. Nach seinem Ingenieurstudium arbeitete er bei der amerikanischen Volkszählungsbehörde, wo er der ständigen Herausforderung der Auswertung einer großen Datenmenge gegenüberstand. Als Erfinder und Ingenieur reifte seine Vision, dies durch eine Maschine zu vereinfachen und deutlich schneller durchführen zu können. Im Jahre 1880 entwickelte Herman Hollerith eine elektrische Zähl- und Sortiermaschine, die Lochkartenmaschine. Mit ihr war es nicht nur möglich, große Datenvolumen zu erfassen, sondern diese auch langfristig speichern zu können. Ein neues Zeitalter der automatischen Datenauswertung und -speicherung war geboren.

Big Data ist heute als Schlagwort omnipräsent. Die Wertschöpfung aus Daten gewinnt zunehmend an Bedeutung. Dabei stellen der Übergang vom Industrie- zum Informationszeitalter und die damit verbundene digitale Transformation Unternehmen vor komplexe und weitreichende Herausforderungen. Die wirtschaftliche Bedeutung von Daten ist unstrittig und drängt die immaterielle Wertschöpfung immer weiter in den Fokus. Dabei nimmt die Relevanz von Datenverarbeitungsprozessen stetig zu. Vor dem Hintergrund der steigenden Datenvolumina wird das Management von Daten zum zentralen Thema. Hier stehen Schlagworte wie Big Data oder Smart Data für die Verdichtung von großen unstrukturierten Daten zu wertvollen Informationen.

Die Treiber für diese Entwicklung sind vielfältig. Sie liegen u.a. in der verstärkten Internetnutzung, der Anwendung mobiler Endgeräte, der Evolution in der Nutzung des Internet sowie der zunehmenden Digitalisierung und Vernetzung von Produkten und Dienstleistungen.

Daneben sind technologische Innovationen relevant, die eine Verarbeitung großer Datenmengen in Echtzeit möglich machen. Diese umfassen beispielsweise Systeme und Modelle wie Hadoop oder MapReduce, NoSQL Datenbanken sowie Business Intelligence Systeme in der Cloud. Parallel zu den Geschäftsmodellen und -services wandeln sich die technologischen Grundlagen aus Software- und Hardwareperspektive sowie aus Sicht der gesamten IT-Architektur.

Das Leitbild für die Zukunft bezieht sich auf eine zunehmende Vernetzung aller Lebensbereiche und die damit verbundene Entstehung neuer Services und Wertschöpfungskonzepte. Diese erlauben durch schnelle Rechner, leistungsfähige Breitbandverbindungen und neuartige Datenbankmodelle eine Auswertung von Daten in Echtzeit und die Schaffung von individuellen Produkt- und Serviceerlebnissen in nie dagewesener Form.

1.1. Forschungsfragen

In der Praxis trifft die skizzierte Vision für Big und Smart Data auf die Realität. Daher stellt sich die Frage, wie weit die mit Big Data verbundenen Erwartungen bereits umgesetzt sind. Die Kernfragen der vorliegenden Studie beziehen sich entsprechend auf den Entwicklungsstand von Big Data und die Anwendung in der Unternehmenspraxis. Damit ist auch eine fundierte Analyse der bislang verfügbaren theoretischen Grundlagen für Big Data verbunden.

Eine Untersuchung dieser Leitfragen verfügt über eine hohe Relevanz, denn in der Unternehmenspraxis stellen sich erhebliche Fragen:

Ist Big Data im deutschen Sprachraum wirklich auf dem Vormarsch? Mit welchen Herausforderungen sehen sich Unternehmen und Führungskräfte konfrontiert? Wie lässt sich Big Data in unterschiedlichen Fachbereichen nutzen, um konkreten Mehrwert aus den Daten zu ziehen?

Der Big Data Report untersucht diese Fragen aus empirischer Sicht. Ziel der Forschung ist es, fundierte Gestaltungsmöglichkeiten zu geben. Diese sind erforderlich, damit Unternehmen das volle Potenzial von Big Data in unterschiedlichen Funktionsbereichen nutzen und aus den verfügbaren Daten einen konkreten Mehrwert schaffen können.

Anhand einer breit angelegten Studie wird die Nutzung von Big Data in der Praxis näher untersucht. Dies gibt eine Momentaufnahme der aktuellen Situation wieder und zeigt Perspektiven für die Zukunft auf.

Konkret lassen sich für den Big Data Report die folgenden Forschungsfragen formulieren:

- Wie ist der Begriff Big Data aus theoretischer und praktischer Sicht konzeptualisiert?
- Welche strategische Bedeutung hat Big Data in der Unternehmenspraxis?
- Welche Ziele und Strategien sind für Unternehmen bei der Gestaltung von Big Data relevant?
- Welchen Reifegrad haben diese Ziele und Strategien bereits erreicht?
- Welche Daten werden für Big Data Analysen vorwiegend genutzt?
- Welche Technologien werden für die Analyse vorwiegend eingesetzt?
- Welche Herausforderungen erzeugt Big Data für die Unternehmenspraxis?
- Welche Ressourcen und Kompetenzfelder sind für die Lösung dieser Herausforderungen aufzubauen?

1.2. Methodischer Ansatz

Die skizzierten Forschungsfragen prägen den Fokus des Big Data Reports. Die Studie wurde durch das Research Lab for Digital Business an der Hochschule Reutlingen in Kooperation mit T-Systems Multimedia Solutions umgesetzt und basiert auf einer dreistufigen Vorgehensweise.

Zunächst wurden Executives aus unterschiedlichen Funktionsbereichen im Rahmen einer breit angelegten Onlinestudie befragt. Die standardisierte Befragung bezieht sich auf die Unternehmensbereiche IT, Marketing und Vertrieb. Entsprechend wurden drei Stichproben mit Führungskräften aus den entsprechenden Funktionsbereichen gebildet. Der vorliegende erste Teil der Studie bezieht sich auf die Befragung von CIOs und IT-Executives.

Die Ergebnisse der Onlinebefragung wurden durch qualitative Interviews vertieft. Diese beziehen sich auf Teilnehmer der Onlinebefragung mit langjähriger Erfahrung in Big Data Projekten. Die qualitativen Interviews wurden telefonisch durchgeführt und durch eine qualitative Datenauswertung analysiert. Diese bildet die Grundlage für die Vertiefung der standardisierten Befragungsdaten.

In einer dritten Teilstufe wurden Fallstudien für die Anwendung von Big Data Strategien in unterschiedlichen Branchen und Anwendungsfeldern umgesetzt. Diese basieren auf qualitativen Tiefeninterviews mit Beteiligten an dezidierten Projekten in der Unternehmenspraxis. Die Ergebnisse der Fallstudien wurden vergleichend analysiert.

Insgesamt bietet der Big Data Report damit auf der Grundlage von Literaturanalyse, standardisierter Befragung, qualitativer Datenanalyse und Fallstudien einen fundierten Einblick in den aktuellen Entwicklungsstand von Big Data aus Sicht von Theorie und Praxis.

Der methodische Ansatz der Studie lässt sich entsprechend durch folgende Abbildung darstellen.

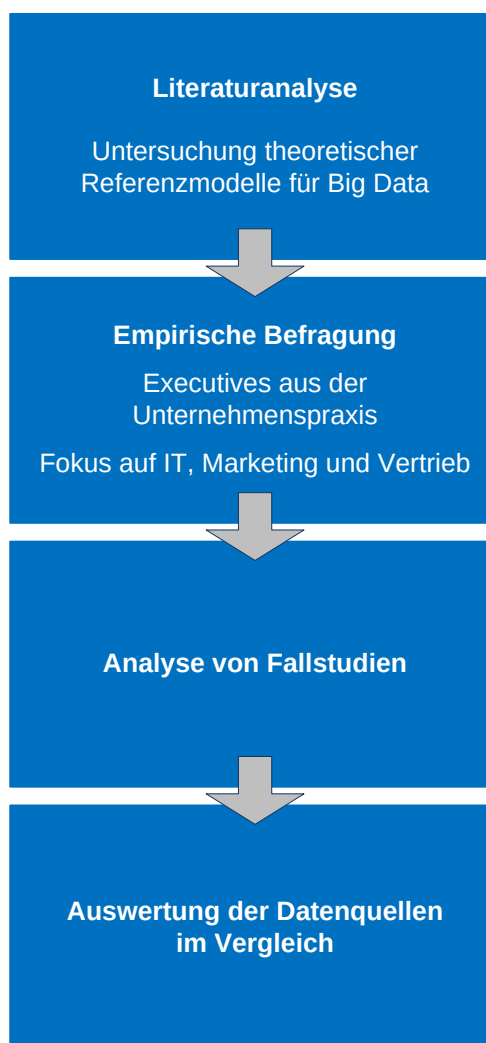


Abbildung 1: Methodischer Ansatz

2. Big Data: Aktueller Stand der Forschung

Aus historischer Sicht durchzieht die Fragestellung nach der Verarbeitung eines hohen Datenvolumens die gesamte Geschichte der Datenverarbeitung. Die aktuellen Entwicklungen rund um Big Data wurden bereits Anfang der 1990er Jahre wissenschaftlich beschrieben.

Erste Publikationen wie „*Saving all the Bits*“ von Denning (1990) oder „*The Evolution of Storage Systems*“ von Morris und Truskowski (2003) legten den wissenschaftlichen Kontext für die Analyse der Verarbeitung großer Datenmengen.

Laney veröffentlichte 2001 einen wegweisenden Beitrag zur Definition von Big Data. Unter dem Titel „*3D Data Management: Controlling Data Volume, Velocity and Variety*“ beschreibt er das sich verändernde Handelsumfeld, welches mit den bisherigen Prinzipien der Datenverarbeitung an seine Grenzen stößt.

Laneys Definition bildet die Grundlage vieler heute noch gängiger Auslegungen von Big Data. Etliche Software- und Beratungsunternehmen wie Gartner, SAS, Strategic Bulletin oder IBM, aber auch Brancheninstitutionen wie die Gesellschaft für Informatik, der Branchenverband bitkom sowie der bitkom Arbeitskreis Big Data setzen auf der 3V-Formel nach Laney auf. Darüber hinaus findet sich das Denkmodell von Laney in vielen praktischen Fallbeispielen wieder.

Volume (Datenmenge)

Schon 1965 erkannte Moore, dass sich die Komplexität integrierter Schaltkreise mit minimalen Komponentenkosten regelmäßig verdoppelt. Einige Jahre später ergänzte Moore die Aussage um eine regelmäßige Verdopplung im zweijährigen Rhythmus. Auch Laney sieht sinkende IT- und Speicherkosten als Treiber für Big Data Anwendungen. Geprägt von großen Datenmengen transaktionsbasierter Daten aus Social Media sowie Sensordaten aus dem Internet der Dinge werden Unternehmen und die verwendeten Datenbanksysteme vor große Herausforderungen gestellt.

Velocity (Geschwindigkeit)

Laney sieht weiterhin eine rasant zunehmende Geschwindigkeit in der Verarbeitung der Daten. Dies gilt beispielsweise für den eCommerce oder für die Unterstützung von Supportprozessen. Nicht nur die Verarbeitung von Daten muss in Echtzeit stattfinden, ebenso die Generierung, Übertragung und Auswertung bedarf hoher Geschwindigkeiten. Im eCommerce-Umfeld sind Antwortzeiten, Warenbestandsanalysen, Transaktionszeiten oder Trackingservices in hoher Geschwindigkeit ein differenzierender Wettbewerbsfaktor. Dies lässt sich analog auf andere Anwendungsbereiche von Big Data übertragen. Daher sind neue technologische Lösungen gefragt, die eine schnelle Verarbeitung und Auswertung großer Datenmengen erlauben.

Variety (Datenvielfalt)

Der Aspekt Variety bezieht sich auf die Datenvielfalt. Dabei bilden die Vernetzung einzelner Sensoren, die existierenden, unterschiedlichen Systeme und ihre inkompatiblen Datenformate, die nicht abgestimmten Datenstrukturen, die inkonsistenten Semantiken sowie die Integration mehrerer Managementsysteme eine wesentliche Herausforderung für die IT. Die Komplexität steigert sich durch interne wie externe Quellen, multimediale Inhalte wie Video, Audio oder Bilder sowie durch strukturierte, semistrukturierte wie auch unstrukturierte Inhalte.

Wie werden Daten smart?

In der Frage steckt die Chance, sich durch datenbasierte Geschäftsmodelle zu differenzieren. Unter dem Schlagwort „Big Data Analytics“ werden methodische Ansätze zur automatisierten Erkennung und Nutzung von Mustern, Zusammenhängen und Bedeutungen zusammengefasst. Möglich wird dies durch die Verarbeitung sehr großer Datenmengen innerhalb statistischer Verfahren, Vorhersagemodellen, Optimierungsalgorithmen, Data Mining, Text- sowie Bildanalytik. Ein wesentlicher Faktor ist hierbei die Verarbeitungsgeschwindigkeit der Daten, die es ermöglicht, die gewonnenen Erkenntnisse zeitgleich in Handlungen umzuformen.

Der Wert von Big Data ergibt sich nicht nur aus der Tatsache, dass es sich bei entsprechenden Anwendungen um hohe Datenvolumina handelt. Wesentlich ist vielmehr die Vernetzung der Daten sowie der durch Analytics für einen spezifischen Anwendungsfall gewonnene Mehrwert.

Damit geht es um den Einsatz großer Datenmengen zur Erzeugung wirtschaftlichen Nutzens. Nach *Gualtieri* (2012) liegt das Ziel von Big Data darin, effektiv zu operieren, Entscheidungen zu fällen, Risiken zu minimieren und Kundennutzen zu erzeugen.

Neuere wissenschaftliche Betrachtungen fügen die Perspektive der Korrektheit bzw. Richtigkeit, die Vollständigkeit sowie die Verlässlichkeit der Daten hinzu. Diese Herausforderung ergibt sich aus der Konsolidierung verschiedener Datenquellen bzw. Datenhierarchien. Gerade im Zeitalter von Social Media sind oftmals der Wahrheitsgehalt externer Informationen sowie deren Autor ungewiss. Je mehr Quellen verwendet werden und je höher die Daten aggregiert werden, desto höher ist das Risiko für statistische Fehler.

Ein weiterer wesentlicher Aspekt liegt im Faktor Variabilität. Neben verschiedenen Quellen oder der hohen Verarbeitungsgeschwindigkeit ergibt sich beispielsweise bei der Trendanalyse in Social Media Netzwerken eine hohe Varianz des Datenaufkommens. Tägliche, wöchentliche, jährliche Steigerungen, aber auch saisonale oder eventbasierte Veränderungen sind möglich. Dies kann die IT kurzfristig überlasten. Somit muss ein weiterer Fokus des Managements auf der Beherrschung von Datenmengen mit hoher Varianz liegen.

2.1. Literaturanalyse

Die theoretische Grundlage für die vorliegende Studie bildet eine umfassende Analyse der bislang verfügbaren Literatur zum Themenfeld Big Data. Auf Grund der umfangreichen Ergebnisse werden hier nur die wesentlichen Teilbereiche der Analyse skizziert. Eine umfassende Literaturliste kann am Research Lab for Digital Business angefordert werden.

Die Literaturanalyse basiert auf mehrstufigen Suchprozessen. Dabei werden klassische Internetquellen und wissenschaftlichen Datenbanken herangezogen. Da das Themengebiet sowohl in der Informatik, als auch in der Betriebswirtschaftslehre verankert ist, werden die Quellen der Datenbanken IEEE, ACM, Emerald und Ebsco genutzt.

Big Data wird häufig mit unterschiedlichen Begrifflichkeiten assoziiert und mit Geschäftsmodellen, Technologien oder Datenschutzaspekten gleichgesetzt. Daher werden für die vorliegende Fragestellung umfangreiche deutsche und englische Stichworte genutzt. Die damit verbundenen Suchbegriffe sind in der Spalte rechts oben abgebildet.

Die wesentlichen Beiträge werden in den folgenden Abschnitten zusammengefasst und gegliedert. Dabei werden relevante Referenzmodelle für Big Data sowie insbesondere das Architekturmodell des Branchenverbandes bitkom näher erläutert.

Suchbegriffe

- Big Data
- Big Data Reference Model / Big Data Referenzmodell
- Big Data Reference Architecture / Big Data Referenzarchitektur
- Big Data Framework
- IoT-Reference Model / IoT-Referenzmodell
- IoT-Data Reference Architecture / IoT-Referenzarchitektur
- IoT-Data Framework
- Internet of Things Reference Model / Referenzmodell
- Internet of Things Reference Architecture / Referenzarchitektur
- Internet of Things Framework
- Big Data Business Model / Geschäftsmodell

Datenbanken

- IEEE
- ACM
- Emerald
- Ebsco

Publikationen im Themenfeld Big Data

Schon weit vor der Aufnahme von Big Data in den Gartner Hype Cycle beschäftigte sich die Forschung mit der theoretischen Beschreibung der Thematik. Die unten skizzierte Tabelle zeigt, dass bisher nur etwas weniger als drei Prozent aller Publikationen zum Thema Big Data in 2015 entstanden sind.

Die skizzierten Quellen wurden jeweils in einzelne Teilbereiche segmentiert und durch die Anwendung semantischer Suchalgorithmen strukturiert. Die Publikationen lassen sich entsprechend in die Kernbereiche Gesellschaft/Recht, Ökonomie sowie IT/Technik gliedern. Innerhalb dieser Kernbereiche ist eine tiefere Untergliederung in einzelne Teilfragen möglich.

Im Anhang der vorliegenden Studie sind die gefundenen und analysierten Frameworks und konzeptionellen Modelle für Big Data zusammengefasst. Diese wurden durch eine manuelle Inhaltsanalyse näher untersucht. Damit liegen inhaltliche Indikationen für die Betrachtung und Entwicklung eines integrierten Big Data Konzepts vor.

Auf Grund der Spezialisierung der Einzelpublikationen werden jedoch nicht alle Referenzarchitekturen für den Big Data Report herangezogen. Eine Auswahl an besonders relevanten Modellen wird in Kapitel 2.4. näher skizziert.

Quellen	Publikationen Treffer Gesamt	Publikationen Treffer 2015
Ebsco	41.911	2.373
IEEE	9.526	420
Emerald	39.319	2.973
ACM	36.300	997
Google	~ 713.000.000	
Google Scholar	~ 3.970.000	~ 92.700

Abbildung 2: Tabellarische Darstellung der Suchtreffer in verschiedenen Datenbanken

2.2. Treiber für Big Data

In den 1990er Jahren legte Tim Berners-Lee die Grundlagen für das so genannte Web 1.0. Er charakterisierte damit eine Kommunikationsplattform, bei deren Inhalte von Unternehmen angelegt und gepflegt werden. Der Kunde hat lediglich die Möglichkeit, diese statischen und nicht-interaktiven Inhalte zu lesen (read-only) bzw. Daten auszutauschen. Im Jahr 1996 erreichte das Web 1.0 bereits 45 Millionen Menschen weltweit. Die ersten Anbieter für Speicherung und Hosting der Daten im Internet etablierten sich. Die Datenmengen wuchsen langsam heran.

Mit dem Web 2.0 setzte eine regelrechte Explosion an erzeugten und zu verarbeitenden Daten ein. Über virtuelle Communities generieren User den Inhalt, teilen ihn mit Freunden und tauschen eine hohe Anzahl an digitalen Inhalten wie Bildern oder Videos über das Netz aus. Die Digitalisierung des Handels setzt ein, primär geprägt durch Unternehmen wie Amazon. Auch die Unterhaltungsindustrie erfindet sich über Plattformen wie YouTube und durch e-Books neu. Im Jahre 2005 steigen die Nutzerzahlen des Internets auf eine Milliarde Menschen weltweit.

Das Internet sozialisiert sich. Große, soziale Gruppen finden sich in Communities. Die digitale Kommunikation rund um Facebook, Twitter und Co. ist ein wesentliches Kommunikationsmedium der modernen Gesellschaft. Ständig werden Informationen geteilt und multimediale Inhalte, wie Bilder und Videos über das Internet ausgetauscht.

Neue digitale Plattformen stehen vor großen Herausforderungen. Alleine Twitter verarbeitet pro Minute über 98.000 Tweets, bei Facebook werden rund 695.000 Status Updates pro Minute im Netz verbreitet.

Die skizzierten Entwicklungen werden durch das so genannte Web 3.0 und das Internet der Dinge weiter beschleunigt. Die informationstechnische Voraussetzung für die Umsetzung dieser technologischen Innovation sind Ontologien, maschinell verarbeitbare Beschreibungen von Domänenmodellen. Im industriellen Kontext wird das Web 3.0 als Internet der Dinge titulierte.

Bei der Eröffnungsfeier der Cebit 2015 beschreibt Martin Winterkorn, Vorstandsvorsitzender der Volkswagen AG, die Auswirkungen des Megatrends wie folgt:

„Die Informationstechnologie ist längst fester Bestandteil der Automobilindustrie. Unsere Autos sind heute schon rollende Rechenzentren mit 1,5 Kilometer Kabeln, mehr als 50 Steuergeräten und der Rechnerleistung von 20 hochmodernen PCs. Jetzt geht es um die große Aufgabe, gemeinsam mit der IT-Branche die Mobilität noch intelligenter und noch vernetzter zu gestalten“.

Geprägt durch moderne Sensorik, Telematik sowie dauerhaften Internetzugang hat die Entwicklung des Internet der Dinge gerade erst begonnen. Hierbei werden physische Produkte, wie Fahrzeuge jeglicher Art, Häuser, Fertigungsfabriken, Stromnetze, Städte bis hin zu Uhren oder Zahnbürsten in einer virtuellen Welt abgebildet. Dadurch eröffnen sich vor allem die Möglichkeiten neuer Geschäftsmodelle.

Die Unternehmensberatung Gartner (2013) schätzt, dass bis zum Jahre 2020 rund 26 Milliarden vernetzte Devices auf dem Markt sein werden. Das Marktvolumen soll sich dann auf rund 1,9 Trillionen USD belaufen.

Besonders die Fertigungsindustrie, welche eine steigende Produkt- und Prozesskomplexität zu verwalten und auf volatilen Märkten mit sich verkürzenden Produkt-, Markt-, Technologie- und Innovationszyklen zu kämpfen hat, wird innerhalb kurzer Zeit große Technologiesprünge erfahren. In diesem Zusammenhang ist unter „Industrie 4.0“ die vierte Evolutionsstufe – die Digitalisierung - der Industrie zuzuordnen. Dabei ist mit einer intelligenten Vernetzung von Produkten sowie zugehörigen Prozessen in der industriellen Wertschöpfung zu rechnen.

Big Data wird darüber hinaus durch einen exponentiellen Anstieg der Nutzerzahlen und des damit verbundenen Datenvolumens getrieben. In 2014 hat bereits rund 40% der gesamten Weltbevölkerung Zugang zum Internet (im Vergleich zu 1995: 1%). Allein zwischen 1999 und 2013 verzehnfachte sich die Anzahl der Nutzer.

Bereits in 2005 wurde die erste Milliardenmarke erreicht, nur fünf Jahre später die Zweite. Seit 2014 gibt es schließlich drei Milliarden Internetnutzer. Dies öffnet Unternehmen innerhalb kurzer Zeit noch nie dagewesene Absatzpotentiale.

Ein bedeutender Anteil des gesamten, weltweiten Datenvolumens wird durch mobile Endgeräte verursacht. Dies zeigt, dass die Verbreitung von Smartphones, Tablets und Wearables einen wesentlichen Beschleunigungsfaktor darstellt. Allein in 2014 kamen weltweit 497 Millionen neue Geräte hinzu. Wachstumstreiber sind nicht mehr nur die Industrienationen: Der Mittlere Osten und Afrika verzeichnen mit einer CAGR (Compound Annual Growth Rate) von 72 % die stärksten Zuwächse bei der mobilen Nutzung von Diensten über das Internet.

Global führen die skizzierten Entwicklungen zu einer Verdopplung der zu speichernden und verarbeitenden Daten im Rhythmus von zwei Jahren. Analysten erwarten in 2020 eine Datenmenge von 40 Zettabytes – das wären 50 mal so viele Daten wie drei Jahre zuvor.

Damit wandelt sich die individuelle und unternehmerische Art und Weise wie wir leben, arbeiten, kaufen und Geschäfte machen fundamental. Das Geschäftsmodell von gestern wird mit hoher Wahrscheinlichkeit nicht mehr jenes von morgen sein. Entsprechend schlägt sich die digitale Transformation inzwischen in allen Branchen und unternehmerischen Größensegmenten nieder – und der Wandel gewinnt weiter an Dynamik.

2.3. Begriffliche Grundlagen

Die Analyse der zum Thema Big Data verfügbaren Literatur macht deutlich, dass bislang kein einheitliches Begriffsverständnis vorliegt. Die Forschung hat sich mit den Auswirkungen von Big Data auf unterschiedliche Lebensbereiche befasst.

Die gesellschafts- und rechtspolitische Perspektive beschreibt die veränderten Lebensbedingungen sowie das Konsumverhalten, wobei die Privatsphäre durch staatliche Rechtsvorgaben zu schützen würdig und notwendig ist. Ökonomische Implikationen beziehen sich auf die Fragestellung, wie sich Daten als Produkt gestalten lassen und die Innovations- sowie Wettbewerbsfähigkeit von Unternehmen stärken. Technisch gesehen bezieht sich Big Data auf Methoden und Technologien zur Speicherung, Analyse und Prozessieren schnell wachsender Datenmengen.

Daher liegen heute sehr viele Beispiele für Definitionen zum Thema Big Data vor, siehe Forbes (2014), Opentracker (2015), MIT Technology Review (2013) oder Berkeley (2014) – eine einheitliche Definition ist jedoch nicht in Sicht. Neuere Publikationen scheinen den Umfang der Diskussion rund um Big Data sogar noch stetig zu erweitern. Daher wird für die vorliegende Studie die oben skizzierte Definition als Arbeitsgrundlage herangezogen.

Big Data bezieht sich auf die schnelle Erzeugung, Speicherung, Analyse, Auswertung und Darstellung großer, vielfältiger und variabel aufkommender Datenmengen, um für Unternehmen und Nutzer einen Mehrwert zu erzeugen.

Implikationen von Big Data zeigen sich in unterschiedlichen Gesellschaftsbereichen, z.B. in Wirtschaft, Industrie, Handel, Rechtsprechung und Technologie.

2.4. Forschungsprogramme und Frameworks

Für Big Data stehen global bereits unterschiedliche Forschungsprogramme und Frameworks zur Verfügung. Das Ziel dieser Konzepte liegt in der Regel in der Strukturierung des Themengebiets und der Darstellung eines Referenz- bzw. Metamodells für Lösungen in der Unternehmenspraxis. Wesentliche Modelle für die Strukturierung von Big Data Anwendungen wurden durch die Literaturanalyse in Kapitel 2.1. erfasst. Dabei zeigt sich, dass Big Data offenkundig aus unterschiedlichen Perspektiven betrachtet werden kann. Die Auswahl relevanter Referenzmodelle muss sich daher an geeigneten Kriterien orientieren. Für den Big Data Report wurden v.a. die Kriterien Aktualität der Publikationen, Abdeckung wesentlicher Themengebiete sowie Bezug zum Geschäftsmodell genutzt. Darüber hinaus muss die Logik der Erzeugung, Verarbeitung, Speicherung und Analyse großer Datenmengen in den Modellen enthalten sein. Eine Selektion auf Basis dieser Kriterien führt zur Konzentration der Analyse auf folgende sechs Referenzmodelle:

- bitkom Big Data Architektur
- Big Data Framework,
 - University of Manchester
 - Tekinger et al.
- IoT Architectural Reference Model,
 - Bassi et al.
- IoT Infrastruktur,
 - Harvard Business School, Porter et al.
- 6C Framework, Ke et al.
- Hybride Geschäftsmodelle,
 - Universität St.Gallen, Fleisch et al.

bitkom

Big Data Architekturmodell

Der Branchenverband bitkom (2014) definiert durch das Big Data Referenzmodell sechs verschiedene Bausteine einer Plattformarchitektur. Hierzu gehören die Datenhaltung, der Zugriff, die Integration, die Analyse, die Verarbeitung, die Darstellung, die Governance und Aspekte der Datensicherheit. Das zugehörige Referenzmodell ist in Abbildung 3 dargestellt.

Die Aufgabe von Big Data Lösungen liegt in der Analyse von Daten zur Erzeugung geschäftsrelevanter Informationen und zur Entscheidungsunterstützung. Die Vielfalt an Anwendungsszenarien erfordert vielseitige Werkzeuge auf jeder Schicht einer Taxonomie. Die in Abbildung 3 dargestellte Architektur beinhaltet nur kommerziell verfügbare Bausteine, welche für eine Big Data Komplettlösung benötigt werden und für den Einsatz in Unternehmen Marktreife erlangt haben – ohne jedoch einen Anspruch auf Vollständigkeit zu erheben. Die Taxonomie ist darüber hinaus nur als modularer Baukasten zu sehen und nicht als präskriptive Referenzarchitektur.

Datenhaltung als Basis

Die Basis des Modells bildet die Datenhaltung und hier das Hadoop Ökosystem. Das Open Source Framework Hadoop bietet eine fundierte Datenspeicherung und erlaubt es mittlerweile Entscheidungen durch verschiedene Analysemethoden nahezu in Echtzeit zu treffen.

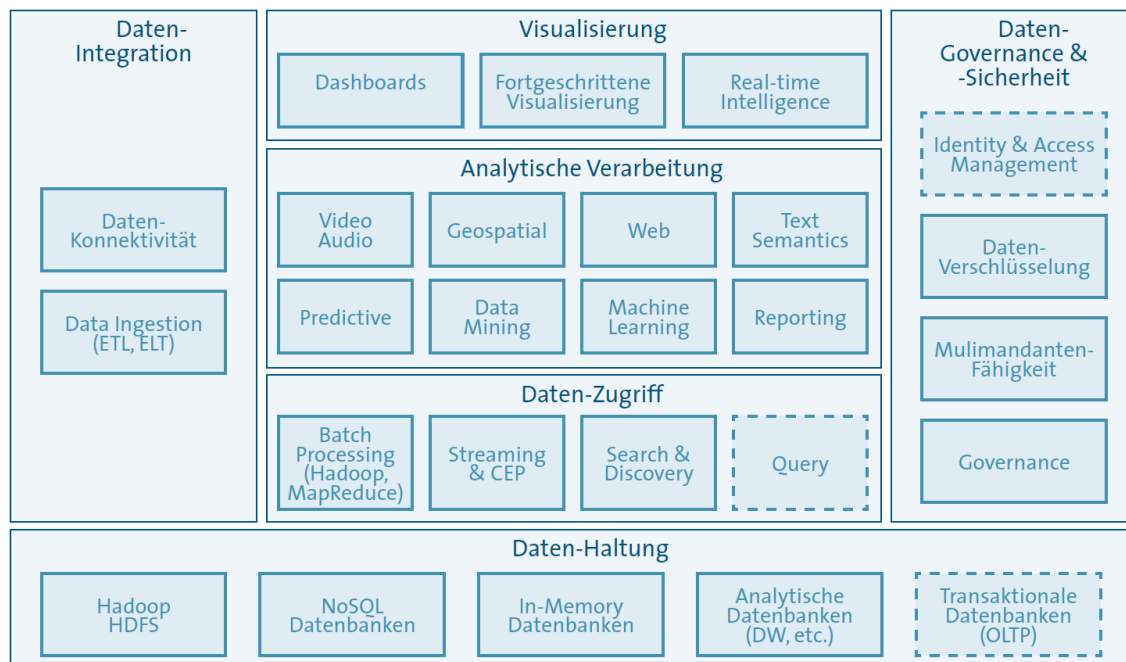


Abbildung 3: Big Data Architekturmodell der bitkom (2014)

Dies basiert auf einer hochgradig parallelen Architektur zur Datenspeicherung und Datenverarbeitung. Dabei stellt HDFS die Hochverfügbarkeit und Redundanz der Daten sicher. Neben Hadoop sind auch weitere Datenbanksysteme dargestellt.

NoSQL Datenbanken kommen z.B. zum Einsatz, wenn Daten keinem relationalen Modell entsprechen. Sie eignen sich besonders dazu, nicht-relationale Daten zu speichern und zu verarbeiten. Um auf Informationen in Echtzeit zuzugreifen, bieten sich In-Memory-Datenbanken an, welche die notwendigen Daten von bis zu mehreren Hundert Terabyte zur Analyse im Arbeitsspeicher vorhalten. Analytische Datenbanken, auch als Data Warehouse bekannt, befassen sich mit relationalen Datenbanken, die für den Einsatz von OLAP optimiert sind und eine umfangreiche Unterstützung für SQL Abfragen bieten.

Dieses Konstrukt ist nicht neu und wird oft in Verbindung mit Hadoop eingesetzt. Der Vollständigkeit halber sind auch transaktionale Datenbanken erwähnt. Diese gehören zur Gruppe der relationalen Datenbanken und sind für den Einsatz von OLTP optimiert.

Datenzugriff als zweite Stufe

Die Schicht des Datenzugriffs befasst sich mit verschiedenen Technologien, die es unterschiedlichen Analyseanwendungen ermöglicht, auf den Datenpool zuzugreifen. MapReduce dient zur Stapelverarbeitung und Planung der parallelen Ausführung von Jobs auf einen Datenknoten im Hadoop Cluster und zur anschließenden Zusammenführung der Ergebnisse. Ergänzt wird der MapReduce Ansatz u.a. durch Hive, welcher es ermöglicht, auf Daten im HDFS mittels der Abfragesprache SQL zuzugreifen.

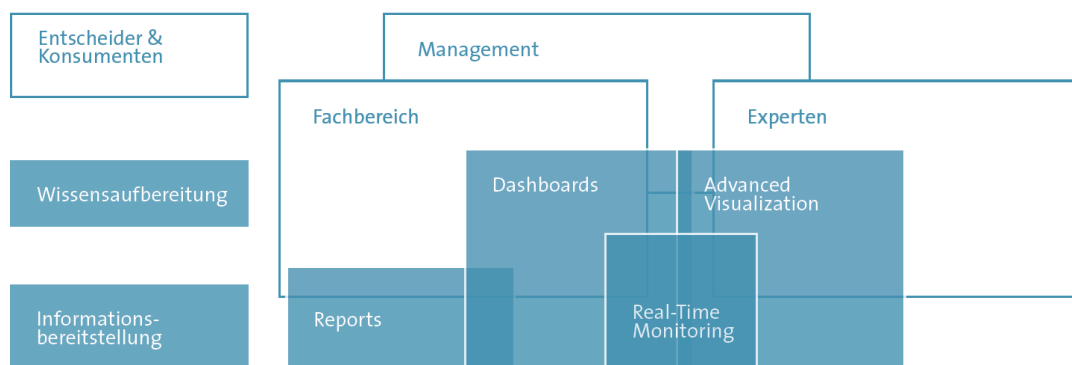


Abbildung 4: Ebenen der Visualisierung gemäß bitkom (2014)

Analysekomponenten als Kern der Big Data Lösung

Die Kernaufgabe der Schicht „analytische Verarbeitung“ besteht in der Generierung und Lieferung von geschäftsrelevantem Wissen und ist dabei stark vom jeweiligen Einsatzgebiet und Datentyp abhängig. Diese Schicht stellt den eigentlichen Kern der Big Data Anwendung dar und befasst sich mit einer Menge von Technologien sowie den Themen OLAP, Data Mining, Predictive Analytics und Machine Learning.

Visualisierung für den Anwender

Zur zielgruppengerechten Präsentation und effizienten Kommunikation von Ergebnissen werden Visualisierungstechniken im Analyseprozess eingesetzt. Das bitkom Modell unterteilt die verschiedenen Techniken zur Visualisierung grob in drei Kategorien: (1) Visualisierung zur Informationsbereitstellung versus visuelle Analyse zur Wissensaufbereitung, (2) Visualisierungsnutzung durch den Fachbereich versus Visualisierungsnutzung als Analyst sowie (3) Analyse ortsgebunden versus mobil.

Unterschiedliche Ebenen der Visualisierung

In Abbildung 4 werden die Zusammenhänge der einzelnen Techniken dargestellt.

Vor dem Hintergrund der Skalierbarkeit sind im Big Data Reporting neue Herausforderungen zu beachten. Bei der Erstellung von interaktiven Darstellungen erfordert das Reporting sehr große Datenmengen sowie eine effiziente Integration der Anwendung zur Visualisierung wie auch zur Analyse.

Das Datenmanagement muss die Aggregation sowie Verdichtung der unterschiedlichen Datenströme und Datenquellen auf der Präsentationsschicht sichtbar machen. Dabei rückt die Echtzeitanalyse immer weiter in den Fokus der Unternehmen, welches vollständig neue Ansätze für Analysten bedingt. Die Präsentationsschicht unterteilt entsprechend drei Bereiche:

- Dashboard und Reporting
- Fortgeschrittene Visualisierung
- Realtime Intelligence

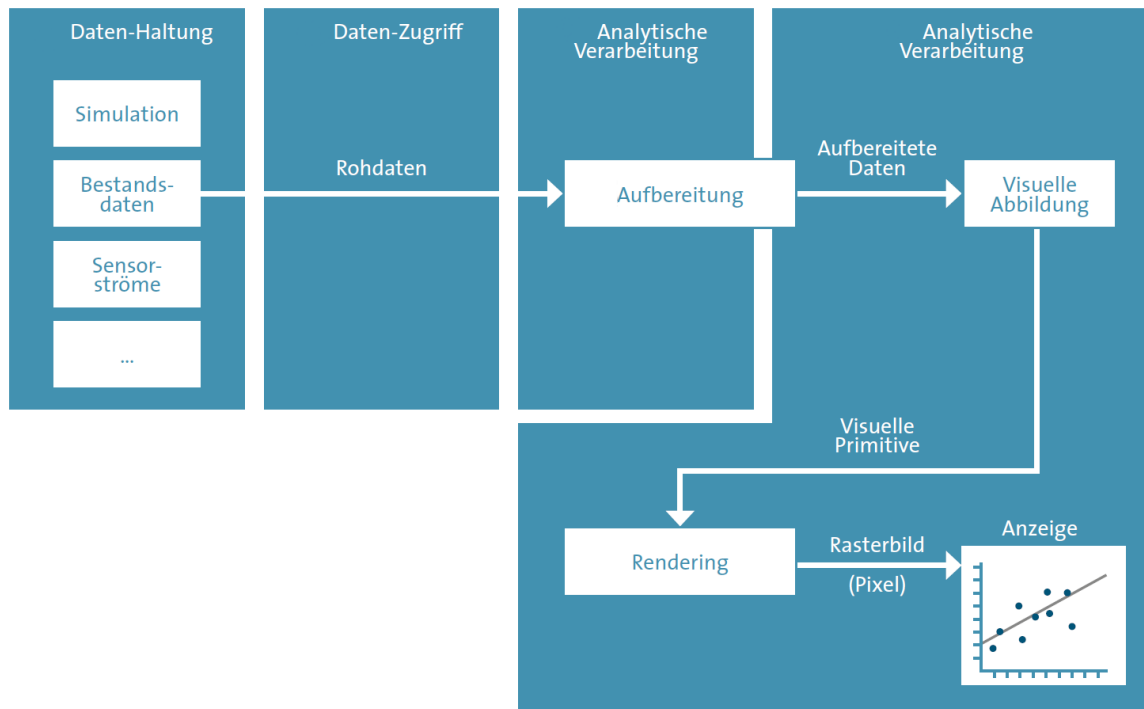


Abbildung 5: Big Data Architekturmodell der bitkom (2014)

Abbildung 5 skizziert die Darstellung des Visualisierungsprozesses und somit den direkten Weg von den Rohdaten hin zu entscheidungsrelevantem Wissen.

Die Bereiche Datenintegration und Datengovernance umrahmen dabei den Prozess. Damit wird sichergestellt, dass sich ein harmonisches Zusammenspiel aus den verwendeten Technologien, Prozessen und Complaincerichtlinien ergibt.

Die Stärken des Modells liegen in der ausführlichen Darstellung einer Vielzahl von technischen Bausteinen und der logischen Zuordnung dieser Elemente. Die Architektur bildet durch die Dimension der Datenintegration einen indirekten Bezug zu Partnern ab. Jedoch ist der Ansatz insgesamt stark nach innen fokussiert.

Ebenso bleiben ökonomische Aspekte völlig unberücksichtigt. Derartige Fragestellungen werden in anderen Modellen deutlich stärker gewichtet.

Dennoch bietet das Referenzmodell der bitkom (2014) unterschiedliche Vorteile. Die Darstellung ist einfach adaptierbar und hat daher eine hohe Praxisrelevanz. Zur weiteren Verwendung dienen vor allem die gut dargestellten Bausteine für die Beschreibung der technischen Anforderungen.

Nachfolgend werden weitere Referenzmodelle für Big Data im Überblick skizziert und im Modellvergleich bewertet. Dies bietet eine fundierte theoretische Grundlage für die empirische Analyse und die Darstellung von Fallstudien.

Big Data Framework, University of Manchester, Tekinger et al.

Tekinger et al. (2013) definieren einen dreistufigen Prozess innerhalb eines siebendimensionalen Frameworks zur Entwicklung von Big Data Anwendungen. Sie beginnen prozessual bei der Datensammlung, folgend mit der Prozessierung bzw. Verarbeitung der Daten. Danach schließt sich die Modellierung und Analyse der Informationen an. Daraufhin werden der Zugriffslayer und die Präsentation bzw. Visualisierung beschrieben. Eingebettet ist das Modell in einer Systemlandschaft sowie einem Datenbankmanagementsystem für strukturierte und unstrukturierte Daten.

Das Framework nach Tekinger et al. (2013) zeigt einen umfassenden Überblick über die technische Infrastruktur. Auf ökonomischer Seite wird lediglich innerhalb des Visualisierungslayers auf die Storytelling-Methode aufmerksam gemacht. Weitere Aspekte eines Geschäftsmodells werden nicht berücksichtigt.

Das Modell ist einfach zu verstehen und bietet durch seine umfassende Darstellung einen hohen Nutzen in der Anwendung. Bis auf den Layer der Datensammlung gibt es keinen Bezug zu externen Partnern und deren Systemen. Außerdem bleiben einige Grundkomponenten unberücksichtigt, z.B. die Datensicherheit. An diesen Stellen weist das Framework Schwächen auf. Die Stärken des Referenzmodells liegen in der einfachen, aber umfassenden Übersicht der IT-Architektur.

IoT Architectural Reference Model, Bassi et al.

Das IoT Architectural Reference Model von Bassi et al. (2013) verfügt über drei Ebenen: Gerät, IT-Infrastruktur sowie Applikationen. Es verfolgt die Zielsetzung, ein Referenzmodell für die Entwicklung, Nutzung und Analyse von IoT-Systemen verfügbar zu machen.

Im Vergleich zu Tekinger et al. (2013) fokussiert das Modell ebenfalls auf das Backend. Die Kommunikation, die Serviceorganisation, das Prozessmanagement, die virtuelle Entität sowie der IoT Service selbst sind in einem Plattformansatz zwischen dem Management- und dem Security-Layer integriert. Das Modell sieht keine ökonomischen Prozesse vor.

Durch eine umfangreiche Berücksichtigung relevanter Literatur ist eine hohe Theoriefundierung gegeben. Modelle aus der Praxis sind jedoch nicht wiederzufinden.

Ebenso fehlt eine externe Sichtweise auf notwendige Kooperationspartner. Dennoch bietet das Modell eine gute Grundlage, um interne IoT-Plattformen innerhalb einer Standardarchitektur zu beschreiben. Interessant ist die Ausrichtung des Ansatzes in Bezug auf eine Serviceorientierung. Das Management- sowie das Sicherheitsmodul werden als Element im Arbeitsmodell übernommen.

IoT Infrastruktur, Harvard Business School, Porter et al.

Das IoT Infrastrukturmodell von Porter et al. (2014) beschreibt den Einfluss von Big Data und IoT auf wirtschaftliche Ökosysteme. Neben den drei Hauptebenen (Product, Connectivity, Product Cloud) berücksichtigt das Modell die Infrastrukturumgebung, das Management von Identität und Sicherheit, den Anschluss externer Datenquellen sowie die Integration von geschäftsrelevanten Services.

Porter et al. (2014) haben ein Modell geschaffen, welches sowohl technische als auch kaufmännische Aspekte beleuchtet. Es ist leicht verständlich, beschreibt nahezu alle relevanten Ebenen einer Big Data Architektur und bietet daher einen starken Mehrwert für die Unternehmenspraxis.

Durch die Anbindung externer Datenquellen ist auch die Berücksichtigung von Partnern möglich. Besonders gelungen erscheint die Segmentierung der „Product Cloud Elemente“ mit der Applikation, der Regel- und Analytics-Engine, der Applikationsplattform sowie der Datenbank.

Das Framework bietet eine hervorragende Grundlage für das zu entwickelnde Arbeitsmodell, da es sowohl aus kaufmännischer, als auch aus technischer Sicht die wesentlichen Elemente beinhaltet und zur unternehmerischen Beschreibung eines IoT-Geschäftsmodells geeignet ist.

6C Framework, Ke et al.

Das 6C Framework nach Ke et al. (2015) basiert auf den Vorarbeiten unterschiedlicher Autoren. Er beschreibt nicht nur eine Architektur, sondern bildet die Darstellung eines gesamten Ökosystems ab. In der Kategorie „Context“ beschreibt der Autor die unternehmerischen Umfeldbedingungen und abgeleitete Strategieimplikationen. Ein weiterer Bereich des Modells bezieht sich auf den Aspekt „Cooperations“. Hier werden die Wertschöpfungspartnerschaften zusammengefasst und die Governance definiert. Produktnahe Aspekte lassen sich unter dem Label „Construct“ definieren. An dieser Stelle wird die benötigte Infrastruktur beschrieben. „Configuration“ ergänzt diesen Aspekt durch die Definition der Art und Weise der Anbindung der Partner. Die notwendige Organisation und Kompetenzen in Bezug auf Innovation, Kommunikation, Lernen und Veränderung werden in den beiden letzten Kategorien „Capability“ und „Change“ aufgegriffen.

Das 6C Framework bietet damit eine ökonomisch orientierte Sichtweise auf das Ökosystem und hat seine Stärken in der umfassenden Darstellung relevanter Dimensionen, wie z.B. der Partner. Interne Aspekte (wie die Organisation und das Change Management) werden gleichermaßen wie externe Fragestellungen (Partner, Kontext) aufgegriffen. Aufgrund seines generischen Ansatzes ist das Framework gut in die Praxis integrierbar, stößt bei Diskussionspunkten jedoch rasch an seine Grenzen.

Hybride Geschäftsmodelle, Universität St.Gallen, Fleisch et al.

Der Ansatz hybrider Geschäftsmodelle von Fleisch et al. (2014) beschreibt den Einfluss der digitalen Transformation auf klassische Branchen und Unternehmen. Dabei wird die Hypothese entwickelt, dass die Trennung physischer und digitaler Geschäftsmodelle obsolet ist. Die Zukunft liegt in so genannten hybriden Geschäftsmodellen.

Das Framework bietet einen 5-stufigen Ansatz zur Entwicklung entsprechender Modelle. Die Grundlagen des Konzepts bieten unterschiedliche Geschäftsmodellmuster, die durch die Unterscheidung in physikalische und digitale Dimensionen erweitert werden. Die Wertschöpfungskette ist in fünf Stufen unterteilt.

Der physische Kundennutzen wird direkt durch physische Dinge (Ebene 1) erbracht. Der Nutzen kann dabei nur in der direkten Umgebung des Gerätes erreicht werden. Direkt damit verbunden ist ein Sensor (Ebene 2), der die Umgebung auf diverse Parameter hin prüft, sowie ein Aktuator, der basierend auf den Messergebnissen die Steuerung des Gerätes übernimmt.

Schließlich setzen physische Dinge auf eine Internetverbindung (Ebene 3) auf und sind mit der digitalen Welt vernetzt. Unter dem Aspekt „Analytics“ fassen die Autoren eine wertschaffende Datenverarbeitung auf (Ebene 4), die schließlich die Voraussetzung für den digitalen Service liefert (Ebene 5). Aus der Zusammenführung der physischen und digitalen Produkte und Services ergeben sich neue Wertschöpfungsmöglichkeiten.

Vergleichende Bewertung der Referenzmodelle

Die jeweils nur kurz skizzierten Modelle reflektieren überwiegend eine technische Architektur für Big Data. Nur wenige Autoren beschreiben die geschäftliche Nutzung großer Datenmengen. Dies ist überwiegend in neueren Publikationen der Fall.

Des weiteren unterscheidet sich der Detaillierungsgrad der Modelle. Diese sind zum Teil nur generisch dargestellt, einige Modelle weisen jedoch mehr Informationen bis hin zur praktischen Operationalisierung auf. Das erschwert die Zuordnung von Aspekten bzw. deren Vergleich.

Die wissenschaftliche Analyse und die zuvor genannten Schwächen erzeugen einen Bedarf nach weiteren Referenzarchitekturen. Dabei kann als Grundlage für die Modellierung der IT-Architektur das Modell der bitkom genutzt werden. Dies ist jedoch um weitere Aspekte einer stärkeren Außenorientierung zu ergänzen. Entsprechende Hinweise finden sich in etlichen der an dieser Stelle bereits dargestellten Modelle.

3. Erste Teilstudie:

Ergebnisse der Befragung von IT-Executives

Neben der Analyse vorhandener Frameworks für Big Data in der Forschung baut der Big Data Report auf einer Befragung unterschiedlicher Referenzgruppen aus der Unternehmenspraxis auf. Eine vorliegende erste Teilbefragung bezieht sich auf CIOs bzw. IT-Executives.

Für die Befragung wurde ein zweistufiges Verfahren gewählt. Im ersten Schritt wurde als Erhebungsmethode eine Onlinebefragung umgesetzt, die im Anschluss durch persönliche Interviews mit ausgewählten Führungskräften vertieft wurde. Im Rahmen der Onlinebefragung wurden Topentscheider aus diversen Branchen befragt. Die Onlinebefragung ist dabei auf den deutschen Sprachraum begrenzt und wurde im Zeitraum von Mitte Mai bis Mitte August durchgeführt.

Struktur der Befragung

Für die Studie wurde eine Onlinebefragung auf Basis von SurveyMonkey entwickelt. Damit lässt sich ein breites Spektrum an Executives aus unterschiedlichen Branchen erreichen. Darüber hinaus entsteht ein klares Meinungsbild zum Thema Big Data ohne Beeinflussung durch den Interviewer. Eine anonyme Erhebung erhöht zudem die Wahrscheinlichkeit einer subjektiv zutreffenden Beantwortung. Die Struktur des Fragebogens unterliegt einer Gliederung in fünf Teilbereiche. Für die Strukturanalyse der Stichprobe wurden zu Beginn Fragen zu Person und Unternehmen der Befragten gestellt.

Dies ermöglicht im Nachgang eine tiefere Analyse der Studienergebnisse, z.B. strukturiert nach Unternehmensgröße oder Branche. Die fünf inhaltlichen Kernbereiche der Befragung beziehen sich auf folgende Teilaspekte:

- Begriff und Bedeutung von Big Data
- Ziele und Strategien für Big Data
- Daten und Technologien
- Herausforderungen für die IT
- Erforderliche Ressourcen und Kompetenzen

Der Schwerpunkt der Befragung bezieht sich auf geschlossenen Fragen und vorstrukturierte Antworten mit einer 5-stufigen Likertskala. Der Fragebogen wurde durch einen internen Pretest getestet und anschließend per eMail direkt an potentielle Teilnehmer der Studie zugestellt.

Als Grundlage dienten die Ergebnisse der Onlinebefragung und es wurden mit ausgewählten Teilnehmern vertiefende Interviews geführt. Ausschlaggebend für die Auswahl der Interviewpartner war dabei eine möglichst ausgeprägte Erfahrung in der Umsetzung von Big Data Projekten. Daher wurde der Fragebogen für die Telefoninterviews direkt aus der Onlinebefragung abgeleitet. Die durchgeführten Interviews wurden digital aufgezeichnet, transkribiert und durch eine qualitative Datenanalyse ausgewertet. Die Ergebnisse der Interviews fließen in die Darstellung der standardisierten Befragungsergebnisse ein.

Struktur der Stichprobe

An der Umfrage beteiligten sich insgesamt 111 Teilnehmer aus unterschiedlichen Branchen. Die genaue Verteilung der Stichprobe auf unterschiedliche Stichproben kann Abbildung 6 entnommen werden. Insgesamt sind in der Stichprobe ausreichend viele Branchen für ein fundiertes Gesamtbild enthalten. Zum Teil lassen sich durch qualitative Zusatzinterviews auch Detailreports zu unterschiedlichen Industriesegmenten ableiten.

Der Großteil der befragten Teilnehmer (52,08%) stammt aus Unternehmen mit weniger als 1.000 Mitarbeitern. Ein Fünftel der Befragten (20,83%) sind in Unternehmen mit einer Größe von 1.000 bis 5.000 Mitarbeitern beschäftigt. Bei mehr als 5.000 Mitarbeiter großen Unternehmen sind 27,10% angestellt. Dies erlaubt bei unterschiedlichen Teilfragen eine Differenzierung des Antwortverhaltens nach Großkonzernen und klein- und mittelständischen Unternehmen.

Branche	Antworten
	Prozent
Automobilwirtschaft	2,08%
Baugewerbe/Bau	10,42%
Chemie	0,00%
Energieversorgung	8,33%
Erbringung von Finanz- und Versicherungsdienstleistungen	6,25%
Erbringung von sonstigen Dienstleistungen	8,33%
Erziehung und Unterricht	4,17%
Gastgewerbe/Beherbergung und Gastronomie	0,00%
Gesundheits- und Sozialwesen	4,17%
Handel	4,17%
Kunst, Unterhaltung und Erholung	0,00%
Land- und Forstwirtschaft, Fischerei	2,08%
Logistik	6,25%
Maschinenbau	12,50%
Nahrungsmittel	2,08%
Öffentliche Verwaltung, Verteidigung, Sozialversicherung	6,25%
Pharma	2,08%
Telekommunikationsindustrie	6,25%
Verarbeitendes Gewerbe/Herstellung von Waren	4,17%
Verkehrswirtschaft	2,08%
Sonstiges	8,33%
	100,00%

Abbildung 6: Stichprobe nach Branchen

3.1. Begriff und strategische Bedeutung von Big Data

Der erste Teilbereich der Befragung bezieht sich auf das Begriffsverständnis und die strategische Bedeutung von Big Data in den befragten Unternehmen. Die Ergebnisse sind in Abbildung 7 dargestellt.

Aus Sicht von 83,33% der Befragten ist Big Data vor allem durch den Begriff Datenbanken gekennzeichnet. Vier von zehn Unternehmen assoziieren mit dem Schlagwort Big Data die Themen Realtime (43,75%) und Cloud Computing (41,67%). Die ersten Aussagen weisen bereits darauf hin, dass es bei Big Data aus Sicht von IT-Executives vor allem um die Speicherung und das Handling von großen Datenmengen geht. Businessaspekte spielen dabei nur eine untergeordnete Rolle.

Mit vergleichbarer Gewichtung werden die Schlagworte Internet der Dinge (35,42%), Sensornetzwerke (33,33%), Services (35,42%) sowie Social Networks (37,50%) mit Big Data in Verbindung gebracht. Bei genauerer Betrachtung zeigt sich, dass die genannten Bereiche stark branchenabhängig sind und das Internet der Dinge sowie Sensornetzwerke besonders im Bereich Maschinenbau, Energieversorgung und Logistik eine große Rolle spielen.

Dies ist vor allem auf den Trend zurückzuführen, physische Produkte verstärkt an das Internet anzubinden. Darüber hinaus verbinden 22,92% der befragten Führungskräfte Big Data mit Mobile Apps & Location Based Services.

Die geringe Nennung kann unter anderem darauf zurückgeführt werden, dass für die befragten Teilnehmer der Einsatz von Apps unklar oder gegebenenfalls noch kein Geschäftsmodell zur gewinnbringenden Nutzung bekannt ist. Zu den sonstigen Antworten zählen zu jeweils 2,08% die folgenden Nennungen: Smart Meter, Analytics und Zeitreihenverwaltung.

Verantwortung für Big Data liegt überwiegend in der IT

Bei der Frage nach dem zentralen Ansprechpartner für Big Data zeigt sich, dass dieser überwiegend in der IT-Abteilung zu finden ist. Die Ergebnisse zu dieser Frage sind in Abbildung 8 visualisiert. Für das Thema ist in 43,75% der Fälle der CIO verantwortlich. Ein weiterer wichtiger Ansprechpartner ist der Chief Executive Officer mit 22,92% der Antworten. 20,83% der Befragten geben an, dass im Unternehmen keine Führungskraft oder kein Ansprechpartner für das Thema vorhanden ist.

In nur 4,16% der Unternehmen ist der Chief Marketing Officer oder Chief Strategy Officer der wesentliche Ansprechpartner für das Thema Big Data. Zu den sonstigen Antworten zählen z.B. Chief Technical Officer, Chief Operation Officer, Chief Process and Innovation Officer sowie IT Manager mit jeweils unter 2% der Antworten.

Der Begriff "Big Data" ist aus meiner Sicht durch die folgenden Teilbegriffe gekennzeichnet.	Antworten		Prozent aller Fälle
	Anzahl	Prozent	
Datenbanken	40	24,54%	83,33%
Internet der Dinge	17	10,43%	35,42%
Realtime	21	12,88%	43,75%
Cloud Computing	20	12,27%	41,67%
Sensorennetzwerke	16	9,82%	33,33%
Services	17	10,43%	35,42%
Social Networks	18	11,04%	37,50%
Mobile Apps & Location-Based Services	11	6,75%	22,92%
Sonstiges	3	1,84%	6,25%
	163	100,00%	339,58%

Abbildung 7: Begriffsverständnis von Big Data

In größeren Unternehmen besteht vielfach eine klar geregelte Zuständigkeit für Big Data. Eine genauere Betrachtung der Ergebnisse zeigt, dass in einem Drittel der Kleinunternehmen mit bis zu 500 Mitarbeitern derzeit kein Ansprechpartner für Big Data etabliert ist. Dies kann ursächlich an der schwachen Verbreitung standardisierter Lösungen für Big Data liegen. Da sich noch wenige etablierte Referenzmodelle durchgesetzt haben, sind bislang v.a. Großunternehmen an entsprechenden Lösungen beteiligt.

Big Data ist in der Unternehmenspraxis weitgehend ein junges Thema

Der Big Data Report untersucht auch die Frage, wie lange sich deutsche Unternehmen bereits mit der Thematik Big Data befassen. Die Ergebnisse dieser Teilfrage sind in Abbildung 9 dargestellt. Dabei findet in 29,17% der Unternehmen bisher keine Auseinandersetzung mit Big Data statt. Dazu zählen allerdings größtenteils Kleinunternehmen mit bis zu 500 Beschäftigten. Hier gaben die meisten Unternehmen an, sich nicht aktiv mit Big Data zu beschäftigen.

Bei größeren Unternehmen ist Big Data dagegen in vielen Fällen erst seit weniger als drei Jahren ein Thema. 22,92% der Teilnehmer geben an, dass ihr Unternehmen sich seit weniger als einem Jahr mit Big Data befasst. In 35,42% der befragten Unternehmen findet seit ein bis drei Jahren eine zielgerichtete Analyse des Themenfeldes Big Data statt.

Lediglich 12,51% der Befragten beschäftigen sich mit dem Big Data seit mehr als drei Jahren. Eine Detailanalyse zeigt, dass dies v.a. auf Konzerne zutrifft, die sich bereits länger gezielt Big Data zuwenden. Das kann darauf zurückzuführen sein, dass entscheidungsunterstützende Supportsysteme und Management von Datenbanken hier bereits sehr lange eine übergeordnete Rolle spielen.

Big Data verfügt bereits heute über eine hohe Relevanz in der IT-Strategie

Immer mehr Unternehmen weisen dem Thema Big Data eine hohe Bedeutung zu. Etwa sieben von zehn Unternehmen schätzen das Thema Big Data als wichtig ein. Die genaue Verteilung ist in Abbildung 10 skizziert.

Für 16,67% der Unternehmen trifft die Aussage zur Wichtigkeit von Big Data in der IT-Strategie voll und ganz zu. Bei 18,75% der Studienteilnehmer ist dies überwiegend der Fall. Ein Drittel der Befragten ist der Meinung, dass die Aussage teils/teils zu trifft (33,33%).

Wer ist innerhalb Ihres Unternehmens der wesentliche Ansprechpartner für das Thema Big Data?	Antworten	
	Anzahl	Prozent
Chief Information Officer (CIO)	21	43,75%
Chief Executive Officer (CEO)	11	22,92%
Kein wesentlicher Ansprechpartner	6	12,50%
Keine Führungskraft	4	8,33%
Sonstiger Ansprechpartner	4	8,33%
Chief Strategy Officer (CSO)	1	2,08%
Chief Marketing Officer (CMO)	1	2,08%
	48	100,00%

Abbildung 8: Ansprechpartner für Big Data

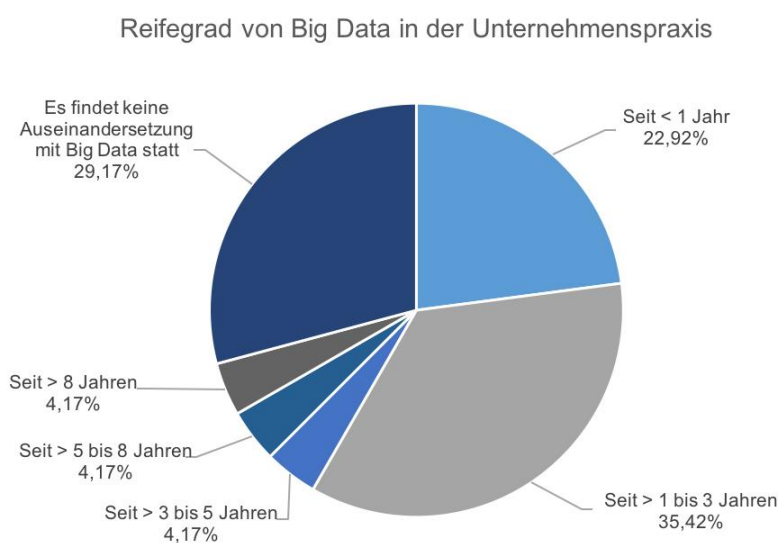


Abbildung 9: Reifegrad von Big Data in der Unternehmenspraxis

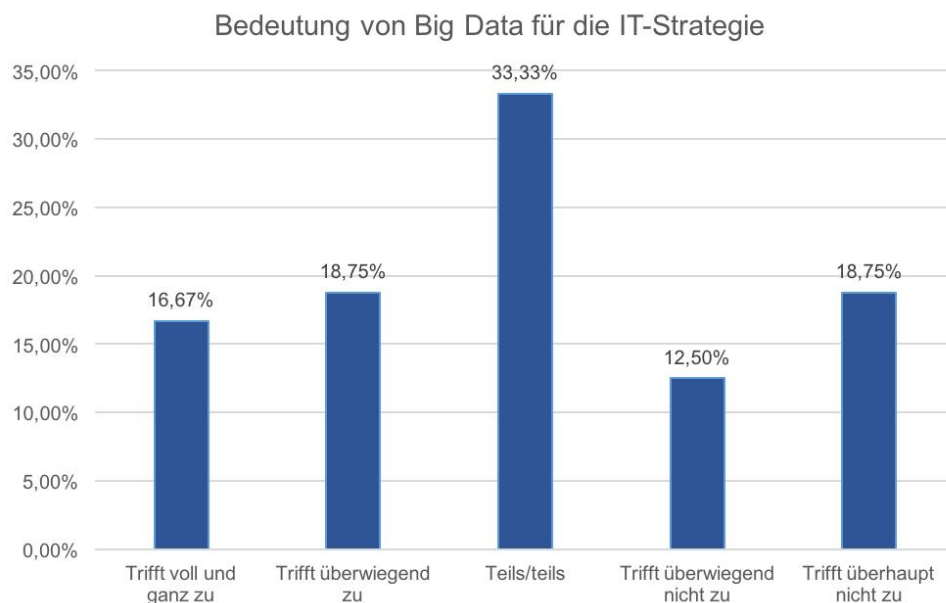


Abbildung 10: Bedeutung von Big Data für die IT-Strategie

Big Data für KMU noch wenig relevant

Das letzte Drittel der Teilnehmer stimmen der Aussage überwiegend nicht (12,50%) oder überhaupt nicht (18,75%) zu. Je nach Unternehmensgröße zeigen sich hier deutliche Unterschiede. So ist, vor allem für kleine Unternehmen bis 500 Mitarbeiter, der Einsatz von Big Data derzeit noch kaum relevant.

Interessant sind darüber hinaus die Aussagen der befragten Executives in Bezug auf die weitere Entwicklung. Diese sind in Abbildung 11 dargestellt.

In acht von zehn Unternehmen gewinnt das Thema Big Data in Zukunft bei der IT-Strategie wesentlich an Bedeutung. Der Aussage, ob Big Data in Zukunft an Bedeutung zunimmt, stimmten ein Viertel der Teilnehmer voll und ganz zu. Für 35,42% der Befragten trifft dies überwiegend zu. Gemischter Ansicht sind 18,75% der befragten Unternehmen. Die Bedeutung von Big Data in der IT Strategie steigert sich bei einem Fünftel der Unternehmen nicht (überwiegend nicht und überhaupt nicht jeweils 10,42%).

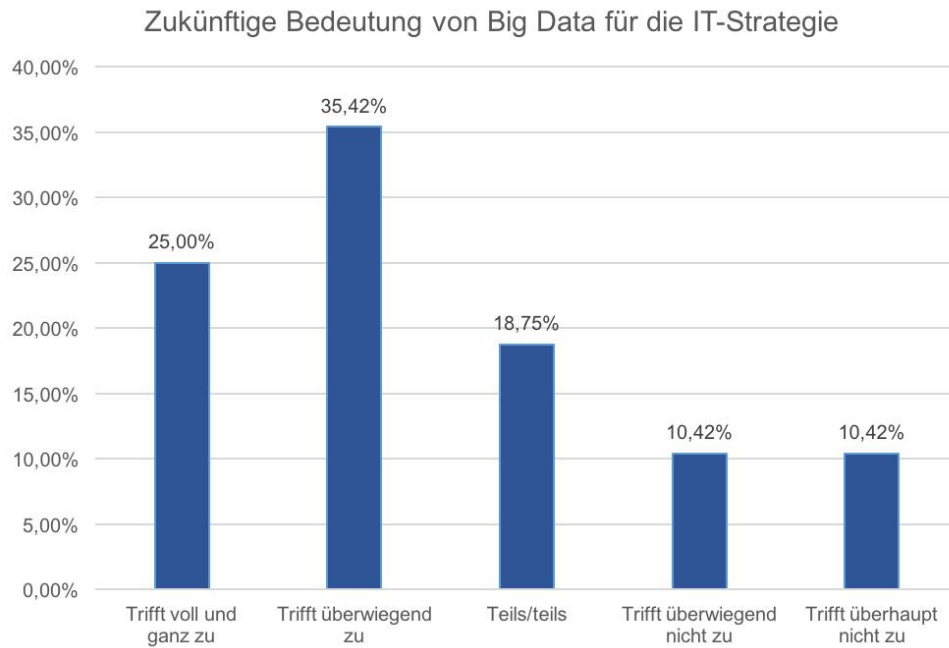


Abbildung 11: Zukünftige Bedeutung von Big Data für die IT-Strategie

Die genauere Auswertung der Ergebnisse zeigt, dass für 9 von 10 Kleinunternehmen die Wichtigkeit von Big Data in Zukunft nicht an Bedeutung gewinnt. Zurückzuführen ist das grundsätzlich darauf, dass kleine Unternehmen nicht in dem gleichen Maß wie große Unternehmen datengetrieben sind und die strategischen Vorteile nicht bekannt sind oder erkannt werden.

Big Data Projekte auf dem Vormarsch

Dass die strategische Bedeutung von Big Data zunimmt, ist am Status quo der aktuell in der Praxis laufenden Big Data Projekte zu beobachten. Big Data steht vor dem Durchbruch.

Entsprechende Projekte wurden in fast jedem zehnten Unternehmen bereits abgeschlossen (8,33%) und in jedem vierten Unternehmen bereits gestartet (25,00%). 20,83% haben bereits eine konkrete Planung angestoßen, künftig Big Data Projekte umzusetzen und bei weiteren 25,00% sind Projekte unter Beobachtung.

Lediglich bei 20,83% finden Big Data Projekte keine Beobachtung. Auch in diesem Fall ist dies wieder auf die Unternehmensgröße zurückzuführen. In sieben von zehn Kleinunternehmen finden Big Data Projekte keine Anwendung.

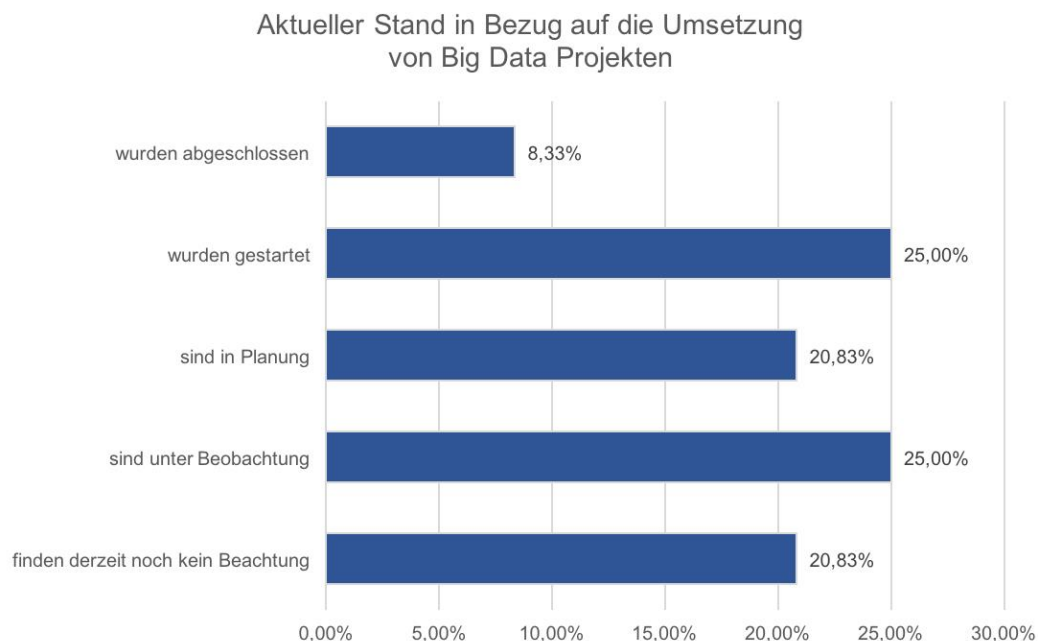


Abbildung 12: Aktueller Stand in Bezug auf die Umsetzung von Big Data Projekten

Die skizzierten Ergebnisse zeigen, dass das Thema Big Data inzwischen auf der strategischen Agenda der meisten befragten Unternehmen angekommen ist. Abstriche zeigen sich nur bei klein- und mittelständischen Unternehmen. Hier weist Big Data häufig noch eine zu hohe Komplexität auf. Daher befinden sich mögliche Projekte bei diesen Unternehmen in den meisten Fällen nur unter Beobachtung.

Trotz der hohen Bedeutung von Big Data ist der Erfahrungshorizont noch gering. Nur 8,33% der befragten Unternehmen haben ein Big Data Projekt bereits tatsächlich abgeschlossen.

Daher ist in den kommenden Jahren mit einer dynamischen Marktentwicklung im Bereich Big Data zu rechnen. Mit der steigenden Anzahl umgesetzter Projekte werden sich auch der Erfahrungsgrad und die Standardisierung im Bereich Big Data erweitern. Dann ist auf Basis von erfolgreich abgeschlossenen Projekten mit einer Verfeinerung der bereits heute vorliegenden Referenzmodelle zu rechnen.

Bis zu dieser Entwicklung sind v.a. frühe Erfolgsbeispiele mit Erfahrungsberichten aus der Praxis wesentlich, um die Lernkurve für Neuprojekte im Big Data Umfeld günstig zu gestalten.

3.2. Ziele und Strategien von Big Data

Ein weiterer Schwerpunkt der empirischen Befragung konzentriert sich auf die Ziele und Strategien von Big Data Projekten. Die Ergebnisse dieser Teilfrage sind in Abbildung 13 dargestellt.

In der CIO Studie sehen die Hälfte aller Beantwortungen die Analyse von Kundenbedürfnissen (52,08%) und die Optimierung von Prozessen (52,08%) als Kernziele von Big Data an. Weitere wichtige Ziele stellen die Unterstützung von Unternehmensentscheidungen (47,92%), Erzeugung von neuen Serviceleistungen (43,75%), Erhöhung der Kundenzufriedenheit (43,75%), Gestaltung individueller Kundenlösungen (39,58%), Kostenreduzierung (39,58%) und die Verbesserung der Kundenbindung (37,50%) dar.

Wertschöpfung aus Kundenbeziehungen

Der starke Kundenfokus lässt sich v.a. aus der Wertschöpfung durch Kundenbeziehungen ableiten. Die Neukundengewinnung ist im Vergleich zur Kundenbindung um ein vielfaches teurer. Durch die Analyse der genauen Kundenbedürfnisse können neue Serviceleistungen erzeugt und individuelle Kundenlösungen generiert werden. Weiterhin können die Daten genutzt werden um die Retention und Kundenzufriedenheit zu erhöhen.

Damit lässt sich auch festhalten, dass die Fachabteilungen Marketing und Vertrieb aktuell wichtigste Partner der IT bei Umsetzungsprojekten sind.

Optimierung von Prozessen und Produktivität

Die Optimierung von Prozessen basiert auf den Grundprinzipien einer schlanken Organisation und der Verbesserung von Entscheidungsprozessen. Dies führt zwangsläufig zur Reduktion von Kosten und der Erhöhung der Produktivität.

Die Identifikation von Langzeittrends und Verkaufsprognosen werden von 29,17% bzw. 20,83% der Befragten als Ziele von Big Data priorisiert. Im Rahmen von Business Intelligence Projekten sind diese Ziele ebenfalls vorhanden. Im Kontext von Big Data können jedoch auch große Datenmengen unstrukturierter Qualität in die Analyse einfließen und das Ergebnis in Echtzeit analysiert werden.

Weitaus weniger oft als die Verbesserung der Kundenbindung und der Entwurf neuer Produkte und Leistungen, ist der Fokus auf die Neukundengewinnung (22,92%) oder auf die Förderung von Kaufentscheidungen (18,75%) gerichtet. Aus Sicht der CIO Studie sind das Kampagnenmanagement und die Erkennung von Compliance Problemen untergeordnete Ziele. Diese sind ausschließlich für 14,58% bzw. 8,33% der Unternehmen relevant. Zu den sonstigen Antworten zählen Qualitätsdatenmanagement (Messdaten) sowie die Abdeckung gesetzlicher Vorgaben mit jeweils 2,08%. Zwei Studienteilnehmer gaben unter sonstigen Zielen an, dass sie keinerlei Ziele mit Big Data verfolgen (4,16%).

Welche Ziele sind für Ihr Unternehmen beim Thema Big Data besonders wichtig?	Antworten		Prozent aller Fälle
	Anzahl	Prozent	
Analyse von Kundenbedürfnissen	25	10,87%	52,08%
Optimierung von Prozessen	25	10,87%	52,08%
Unterstützung von Unternehmensentscheidungen	23	10,00%	47,92%
Erzeugung von neuen Serviceleistungen	21	9,13%	43,75%
Erhöhung der Kundenzufriedenheit	21	9,13%	43,75%
Gestaltung individueller Kundenlösungen	19	8,26%	39,58%
Kostenreduzierung	19	8,26%	39,58%
Verbesserung der Kundenbindung	18	7,83%	37,50%
Identifikation von Langzeittrends	14	6,09%	29,17%
Gewinnung von Neukunden	11	4,78%	22,92%
Verkaufsprognosen	10	4,35%	20,83%
Förderung von Kaufentscheidungen	9	3,91%	18,75%
Kampagnenmanagement	7	3,04%	14,58%
Erkennung von Compliance Problemen	4	1,74%	8,33%
Sonstiges	4	1,74%	8,33%
	230	100,00%	479,17%

Abbildung 13: Ziele für Big Data Projekte

Neben der Ausrichtung von Big Data Projekten auf dezidierte Ziele ist der Reifegrad bzw. die explizite Abbildung einer strategischen Vorgehensweise wesentlich. Die in Abbildung 14 dargestellten Ergebnisse weisen jedoch darauf hin, dass eine entsprechende Strategie in den meisten befragten Unternehmen noch nicht dokumentiert ist. Lediglich jedes sechste Unternehmen gibt an, dass für Big Data eine explizite Strategie entwickelt und dokumentiert wurde. Für nur 2,08% der befragten Unternehmen trifft diese Aussage voll und ganz zu, 12,50% der Unternehmen wählten die Antwort „überwiegend“. In drei von sieben Unternehmen (29,17%) liegt teilweise eine dezidierte und explizit dokumentierte Strategie vor. Bei über der Hälfte der befragten Teilnehmer steht keine dokumentierte Strategie zur Verfügung (überwiegend nicht 27,08% und überhaupt nicht 29,17%).

Nachholbedarf bei der Strategieentwicklung

Damit lässt sich als Status quo dokumentieren, dass etliche Big Data Projekte heute noch ohne klare strategische Ausrichtung betrieben werden.

In 33,33% der Unternehmen wurden bereits Projekte zum Thema abgeschlossen oder gestartet, aber lediglich in 14,58% der Fälle liegt eine explizit dokumentierte Strategie vor. Bei einem Großteil der befragten Unternehmen herrscht demnach Nachholbedarf bei der Strategieentwicklung für Big Data.

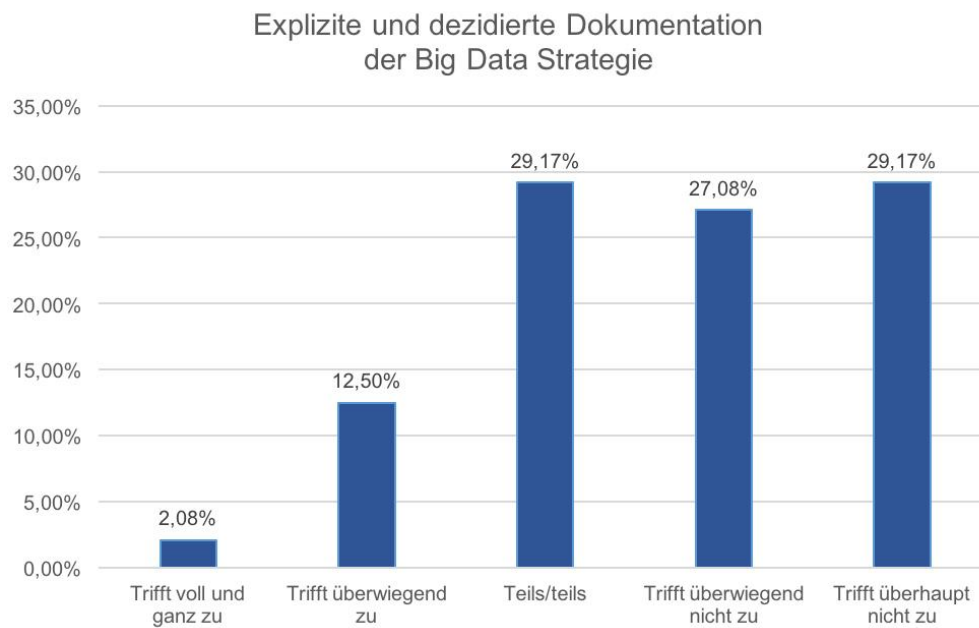


Abbildung 14: Reifegrad der Big Data Strategie

Die Bedeutung von Big Data wird weiter steigen bzw. weist bereits heute in vielen Unternehmen ein hohes Niveau auf. Bei der Frage nach der Einbettung von Big Data in die Gesamtstrategie zeigt sich jedoch ein gemischtes Bild (siehe Abbildung 15). Überwiegend wird die Frage nach der Einbettung von Big Data in die IT-Strategie gemischt bewertet (37,50%).

Bei einem Drittel der Befragten ist Big Data überwiegend nicht (12,50%) oder überhaupt nicht (16,67%) in die IT-Strategie integriert. Darunter fallen mit 87,50% zu einem Großteil Kleinunternehmen mit bis zu 500 Mitarbeiter/innen.

Die Ergebnisse reflektieren die Tatsache, dass Big Data einen bereichsübergreifenden Ansatz erfordert. Dies zeigt sich auch bei der Einbettung in Bereichsstrategien und dem Anspruch an eine ganzheitliche Zusammenarbeit innerhalb des gesamten Unternehmens. Daher sind häufig bereichsübergreifende Strategien für eine volle Erschließung der mit Big Data verbundenen Potentiale erforderlich.

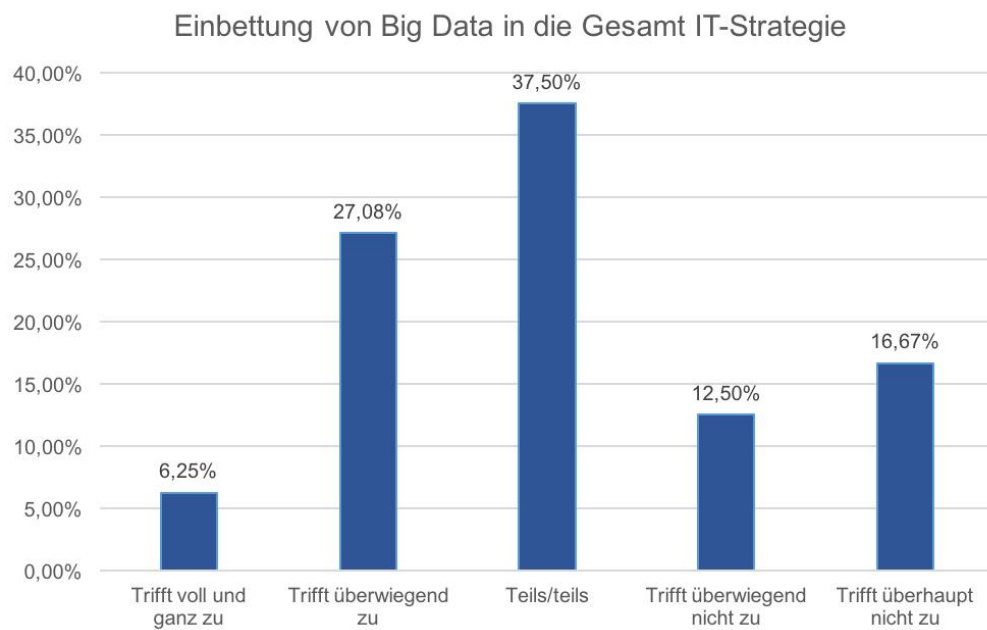


Abbildung 15: Einbettung von Big Data in die IT-Strategie

3.3. Daten und Technologien

Ein weiterer Schwerpunkt der Untersuchung bezieht sich auf die aktuell in Unternehmen genutzten Datenquellen, die eingesetzten Technologien und die Aktualität der Datenanalyse. Die Ergebnisse der Befragung sind in Abbildung 16 dargestellt. Bei der Analyse von Datenquellen sind Mehrfachantworten erlaubt. Davon ausgenommen sind die Fragen zur Aktualität der Datenauswertung und zur Reaktion auf Datenupdates.

Interne Daten im Fokus

Der Status quo in Bezug auf aktuell relevante Datenquellen zeigt, dass bei der Hälfte aller Unternehmen interne Prozessdaten (51,17%), maschinenproduzierte Daten (45,83%) und Produktionsdaten (43,75%) relevant sind. Zu maschinenproduzierten Daten zählen u.a. Sensornetzwerke sowie Daten aus dem Internet der Dinge. Zu einem Großteil stammen diese Ergebnisse aus den Branchen Maschinenbau (100%), Energieversorgung (75%) und Telekommunikation (66,66%). Diese nutzen häufig sogar eine Kombination aus den drei skizzierten Datenquellen.

Die Ergebnisse zeigen auf, dass der Schwerpunkt der Datenanalyse vorwiegend an intern verfügbaren Daten ansetzt. Diese lassen sich im Verhältnis zu externen Daten einfacher erfassen und verarbeiten. Mit Hinblick auf die eingangs skizzierten Zielsetzungen für Big Data zielt dies v.a. auf die Faktoren Produktivität und Effizienz.

Eine stärkere Ausrichtung von Big Data Projekten auf den Kunden muss sich auch mit der Analyse externer Daten befassen. Für jedes dritte Unternehmen sind produktbezogene Lebenszyklusdaten (33,33%), Geodaten (31,25%) und Daten aus Social Media Kanälen (29,17%) relevant. Damit ist bereits ein Fokus auf externe Datenquellen vorhanden, dieser ist jedoch deutlich weniger stark ausgeprägt. Darüber hinaus stellt sich die Frage, wie unternehmensexterne Daten in geeignete Servicemodelle überführt werden können. Dies lässt sich häufig nur auf Basis von Einzelbeispielen und spezifischen Industrieanwendungen beantworten. Damit ist auch die Definition von Standardmodellen für die Verbindung von Daten und Services schwierig.

Nutzung von Maschinendaten im Produktlebenszyklus

Bei der Analyse fällt weiterhin auf, dass die Analyse von Lebenszyklusdaten mit 83,33% im Maschinenbau von besonders hoher Bedeutung ist. Dies weist darauf hin, dass der Maschinenbau inzwischen auf eine genaue Erfassung von Maschinendaten im Produktlebenszyklus setzt und daraus neue Servicemodelle entwickelt. Speziell für diese Branche bietet die digitale Transformation erhebliche Ansätze für neue Produkt- und Servicemodelle. Dies reflektiert sich bei der Frage nach der Nutzung unterschiedlicher Datenquellen.

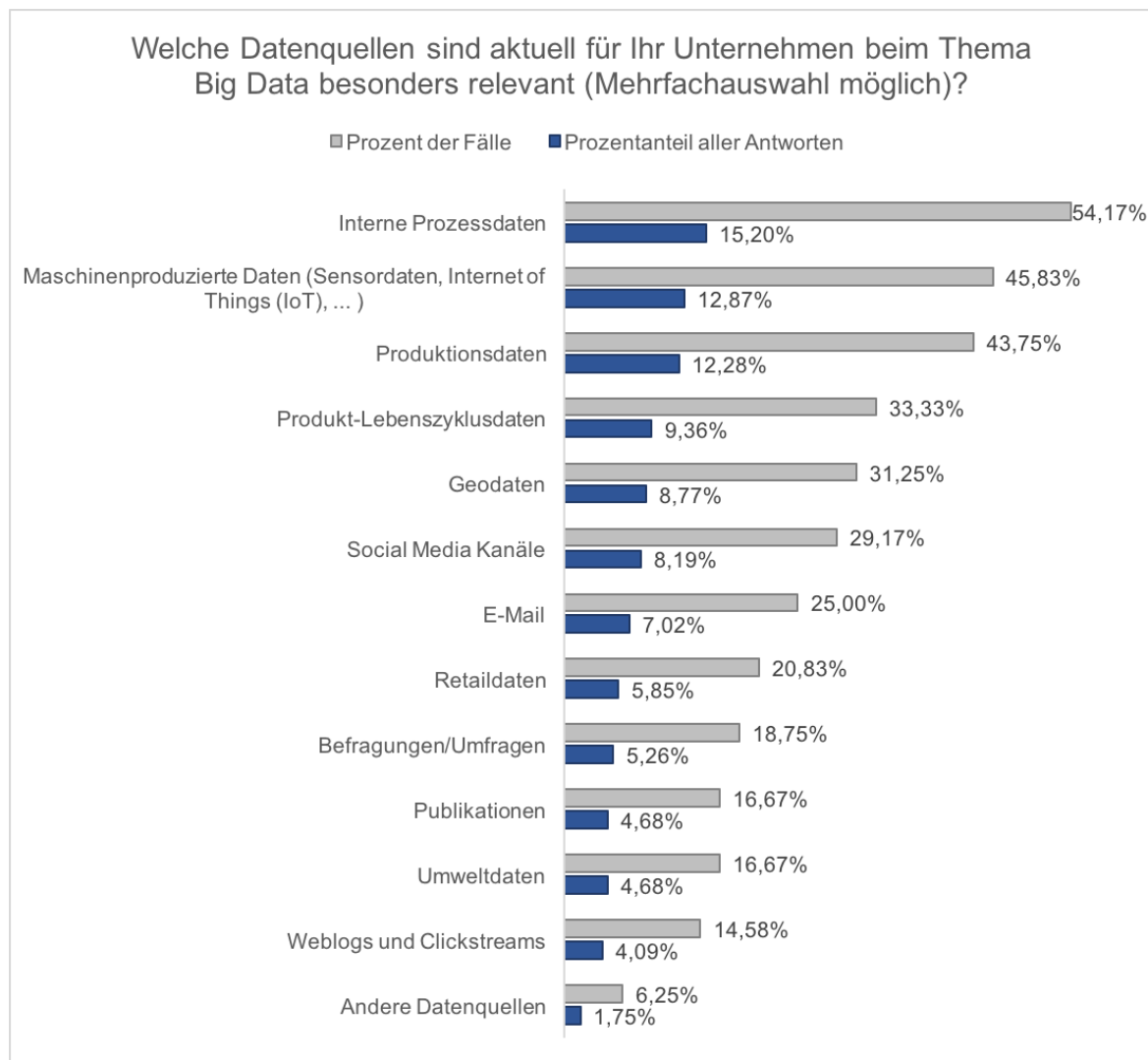


Abbildung 16: Relevanz von Datenquellen für Big Data Projekte

Im Kontext von Big Data steht darüber hinaus bei einem Viertel aller Teilnehmer der vorliegenden Befragung die E-Mail im Fokus (25,00%). 20,83% sehen Retaildaten als wesentliche Datenquelle. Befragungen (18,75%) und Publikationen (16,67%) finden zum Großteil in der Gesundheitsbranche Anwendung. Darüber hinaus setzen 16,67% der Unternehmen Umweltdaten für Big Data Projekte ein.

Diese lassen sich in geeigneten Anwendungen mit anderen Datenquellen kombinieren. Weblogs und Clickstreams werden von 14,58% der befragten Unternehmen genutzt. Zu den weiteren Datenquellen zählen mit 2,08% Grid Daten für Smart Grids und kundenspezifische Maschinendaten. Damit zeigt sich eine sehr hohe Bandbreite bei den genutzten Datenquellen für Big Data Projekte.

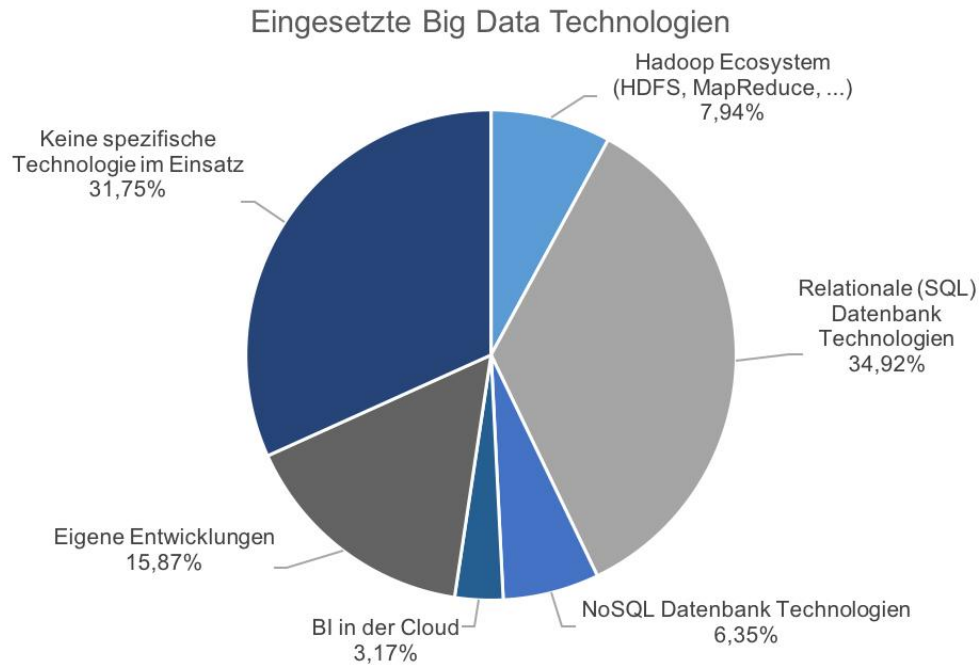


Abbildung 17: Eingesetzte Big Data Technologien

Spezifische Big Data Technologien derzeit kaum im Einsatz

Die Frage nach eingesetzten Big Data Technologien in Unternehmen zeigt deutlich, dass Big Data noch ein sehr junges Thema in deutschen Unternehmen ist.

In 31,75% der Fälle wird derzeit noch keine spezifische Technologie für Big Data eingesetzt. 15,87% der Unternehmen setzen auf eigene Entwicklungen. Neuere Technologien wie Hadoop und MapReduce (7,92%), NoSQL Datenbanken (6,35%) oder BI in der Cloud (3,17%) finden derzeit kaum Anwendung. Dies ist im Wesentlichen auf die schwache Verbreitung und die geringe Anzahl abgeschlossener Big Data Projekte (8,33%) zurückzuführen.

Darüber hinaus sind neue Technologien wie Hadoop mit einer umfassenden Investition und Neugestaltung der eigenen IT-Infrastruktur verbunden.

Diese Aufwendungen werden heute häufig noch gescheut. Daher setzen Unternehmen bei der Umsetzung von Pilotprojekten eher auf klassische Technologien. Beispielsweise werden relationale Datenbanken im Kontext von Big Data bei 34,92% der befragten Unternehmen eingesetzt. Spezifische Technologien für Big Data haben sich noch nicht durchgesetzt. In der Praxis wird allerdings häufig der Ansatz einer Two-Speed Architektur zwischen etablierten Kernsystemen und neuen Big Data Anwendungen verfolgt.

Wo werden die Daten aktuell in Ihrem Unternehmen ausgewertet?	Antworten		Prozent aller Fälle
	Anzahl	Prozent	
Internes Rechenzentrum	27	51,92%	56,25%
Virtual Private Cloud	6	11,54%	12,50%
Hybrid Cloud	1	1,92%	2,08%
Public Cloud	2	3,85%	4,17%
Aktuell erfolgt keine systematische Datenauswertung	16	30,77%	33,33%
Sonstiges	0	0,00%	0,00%
	52	100,00%	108,33%

Abbildung 18: Datenhaltung

Zurückhaltung bei Cloud Strategien

Dies reflektiert sich bei der Frage nach den eingesetzten Verfahren zur Datenhaltung. Die entsprechenden Ergebnisse sind in Abbildung 18 visualisiert.

Nach wie vor sind deutsche Unternehmen beim Thema Cloud sehr zurückhaltend. In Summe beschäftigen sich erst 17,31% aller befragten Unternehmen mit Datenanalysen in unterschiedlichen Cloudvarianten (z.B. Virtual Private Cloud, Hybrid Cloud, Public Cloud). Jedes zweite Unternehmen (51,92%) analysiert die Daten im internen Rechenzentrum. Häufig ist auch noch keine klare Strategie für die Haltung unterschiedlicher Datentypen erkennbar. Es dominieren unterschiedliche Datensilos, die sich bislang nur im Bedarfsfall und mit Zusatzaufwand für eine integrierte Datenauswertung nutzen lassen. Die damit verbundenen Vorteile von Cloudlösungen werden grundsätzlich gesehen, jedoch bestehen hier nach wie vor erhebliche Vorbehalte in Bezug auf die Datensicherheit.

Fokus auf klassische Ansätze der Datenauswertung

Der konservative Ansatz bei der Datenhaltung reflektiert sich auch bei der Datenauswertung. Vorwiegend werden klassische Reportingstrategien umgesetzt.

Die Vorgehensweise in Bezug auf die Datenauswertung in teilnehmenden Unternehmen ist von erheblichem Interesse. Schließlich lassen sich die Vorteile von Big Data v.a. auf Basis der Datenauswertung und der Verwertung der Analysen für neue Services erzielen. Allerdings geben 35,42% der Befragten an dieser Stelle an, dass in ihren Unternehmen keine Methode zur systematischen Datenauswertung im Einsatz ist (siehe Abbildung 19). Dazu zählen 82,35% der befragten Unternehmen mit weniger als 5.000 Mitarbeitern/innen. Die verbleibenden 17,65% dieser Teilstichprobe basiert auf Großunternehmen mit mehr als 25.000 Mitarbeitern/innen.

Wie werden die Daten aktuell in Ihrem Unternehmen ausgewertet?	Antworten		Prozent aller Fälle
	Anzahl	Prozent	
Reporting / Standardberichtswesen	27	30,34%	56,25%
Dashboard	13	14,61%	27,08%
OLAP	11	12,36%	22,92%
Data Mining	10	11,24%	20,83%
Predictive Analytics (Prognoseverfahren)	10	11,24%	20,83%
Text Mining	1	1,12%	2,08%
Aktuell erfolgt keine systematische Datenauswertung	17	19,10%	35,42%
Sonstiges	0	0,00%	0,00%
	89	100,00%	185,42%

Abbildung 19: Datenauswertung

Datenbestände ohne Auswertungslogik

Daher stellt sich die Frage, auf welcher Grundlage die Unternehmen dieser Teilstichprobe Entscheidungen treffen bzw. wie genau mit den gesammelten Daten verfahren wird. Offensichtlich werden in etlichen Unternehmen Daten erfasst, die dann aber nicht in strategische Entscheidungsprozesse einfließen.

Die Ergebnisse der Untersuchung zeigen weiterhin, dass bei den Verfahren zur Datenauswertung klassische Methoden dominieren. Selbst in den Unternehmen, die bereits aktiv Daten für Analysezwecke nutzen, fließen diese überwiegend nur in das Standardreporting ein. Bei 56,25% der Fälle ist dies im Sinne eines Berichtswesens der Fall. Dashboards kommen in 27,08% der befragten Unternehmen zum Einsatz. Bei fortgeschrittenen Analysemethoden ist die Erfahrungsbasis noch gering ausgeprägt. Auf Basis eines Data Warehouse werden nur in 22,92% der Unternehmen die Daten mittels OLAP analysiert.

Datenauswertung in den meisten Unternehmen nach Bedarf

Die Datenauswertung findet bei den meisten Unternehmen heute noch nach Bedarf statt. Algorithmenbasierte Verfahren zur Analyse der Daten werden nur teilweise eingesetzt. Dazu zählen Data Mining mit 20,83%, Predictive Analytics mit 20,83% und Text Mining mit 2,08% (siehe Abbildung 19).

Die Datenauswertung findet in diesem Kontext bei 14,58% der befragten IT-Executives laufend und in Echtzeit statt. Aus Abbildung 20 lässt sich entnehmen, dass die Aktualität der Datenauswertung für viele Unternehmen heute noch nicht die höchste Priorität hat. 27,08% der Teilnehmer gaben an, dass in ihrem Unternehmen die Daten täglich ausgewertet werden. Der Großteil der Unternehmen (33,33%) wertet seine Daten bei Bedarf aus.

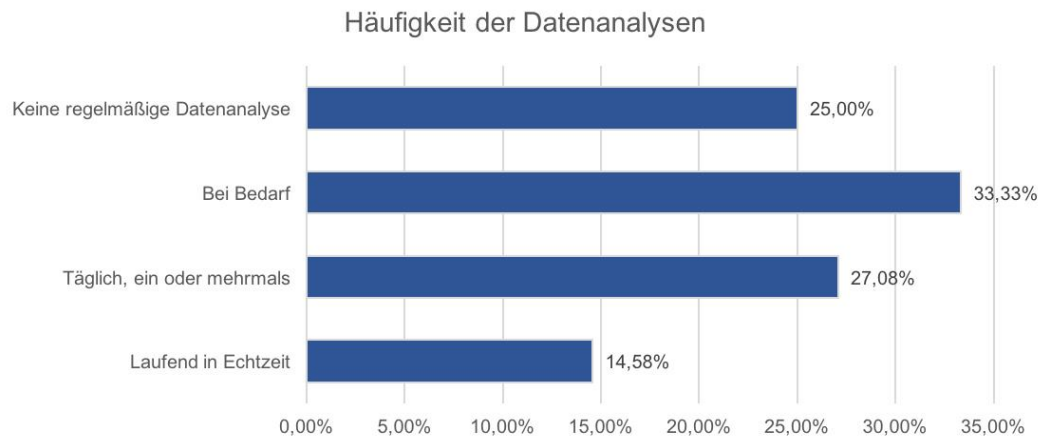


Abbildung 20: Häufigkeit der Datenanalysen

Weitere Ergebnisse zur Reaktion auf dynamische Datenbestände sind in Abbildung 21 dargestellt. Der Umgang mit derartigen Datenbeständen ist stark abhängig vom Anwendungsfall. Bei jedem fünften Studienteilnehmer erfolgt bereits eine automatisierte Reaktion auf entsprechende Ereignisse (20,83%). Eine manuelle Reaktion ist in 50,00% der Fälle zu beobachten.

29,17% der befragten Unternehmen reagieren überhaupt nicht auf eine Veränderung im Datenbestand. Das Ergebnis verdeutlicht, dass der Großteil der Unternehmen einer automatisierten Auswertung noch kritisch gegenübersteht und auf eine manuelle Reaktion vertraut. Damit lassen sich Kontrollmöglichkeiten mit Hinblick auf Reaktionen bei Datenänderungen deutlich erweitern. Automatische Reaktionsmuster werden dagegen kritisch bewertet, da hiermit häufig unkontrollierte Aktionen assoziiert werden, die sich ggf. keiner klaren Verantwortung zuordnen lassen.

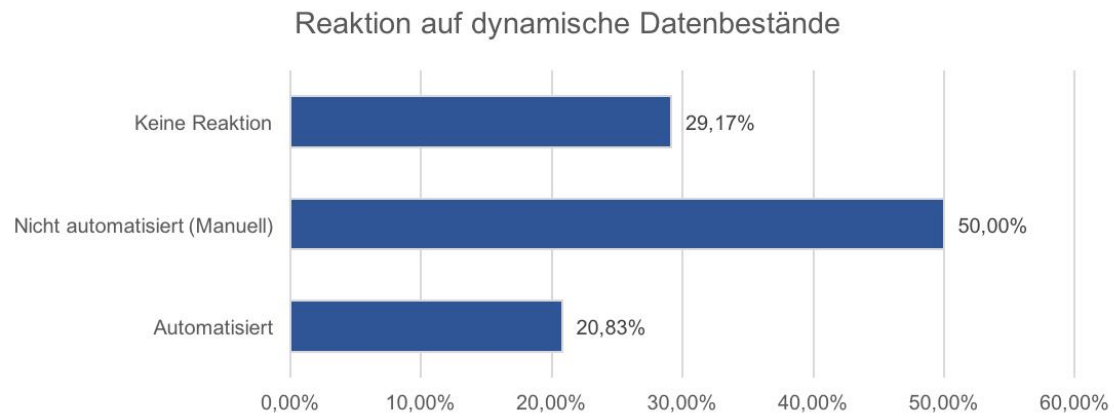


Abbildung 21: Reaktion auf dynamische Datenbestände

3.4. Herausforderungen für die IT

Der Big Data Report untersucht darüber hinaus die mit Big Data verbundenen Anforderungen an die IT. Fast 80% der Unternehmen geben an, dass Big Data die IT vor erhebliche Anforderungen stellt (siehe Abbildung 22). 35,42% der befragten Teilnehmer sind der Meinung, dass die gesetzte Aussage zur Steigerung der Anforderungen an die IT durch Big Data voll und ganz zu trifft. Vier von zehn Teilnehmern stimmen dem überwiegend zu (41,67%).

Dies bringt zum Ausdruck, dass die befragten IT-Executives mit erheblichen Herausforderungen bei der Umsetzung von Big Data Projekten rechnen. Big Data ist daher kein Selbstläufer. Eine exakte Analyse der relevanten Anforderungsbereiche für die Unternehmenspraxis ist angezeigt.

Datenmanagement als kurzfristige Priorität

Dabei betrachten die befragten IT-Executives v.a. das Datenmanagement als erhebliche Herausforderungen (79,17%). Dies bezieht sich auf die Konsolidierung unterschiedlicher Datenquellen und die Verfügbarkeit der Daten. Daten werden heute nach wie vor in unterschiedlichen Datenbeständen gehalten und liegen nur in wenigen Fällen integriert vor. Häufig müssen für Big Data Projekte völlig neue Architekturen der Datenhaltung aufgebaut werden. Daher wird sich ein großer Teil der kurzfristigen Projekte auf die Konsolidierung der Datenbestände konzentrieren.

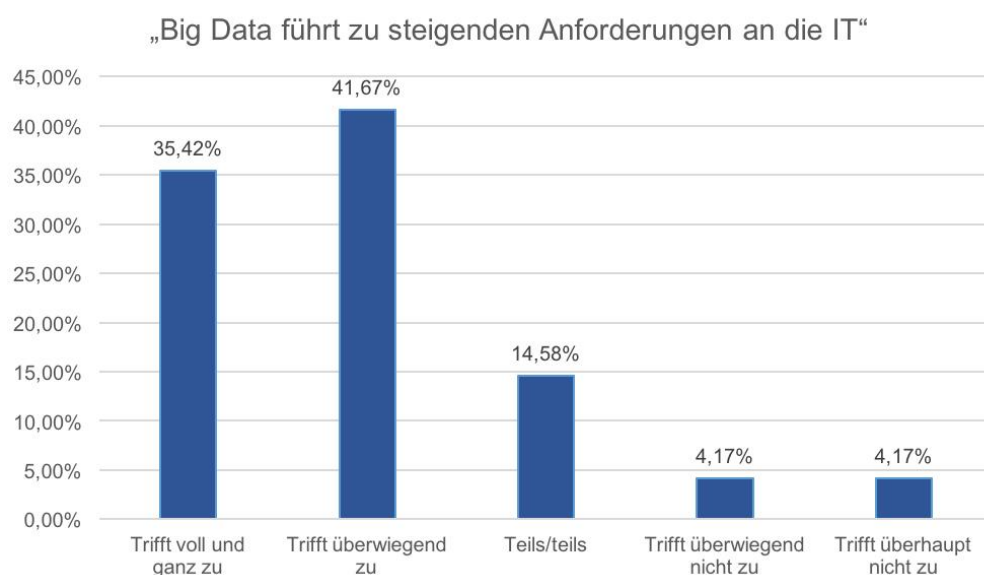


Abbildung 22: Steigende Anforderungen durch Big Data

Analyse der wesentlichen Anforderungsbereiche

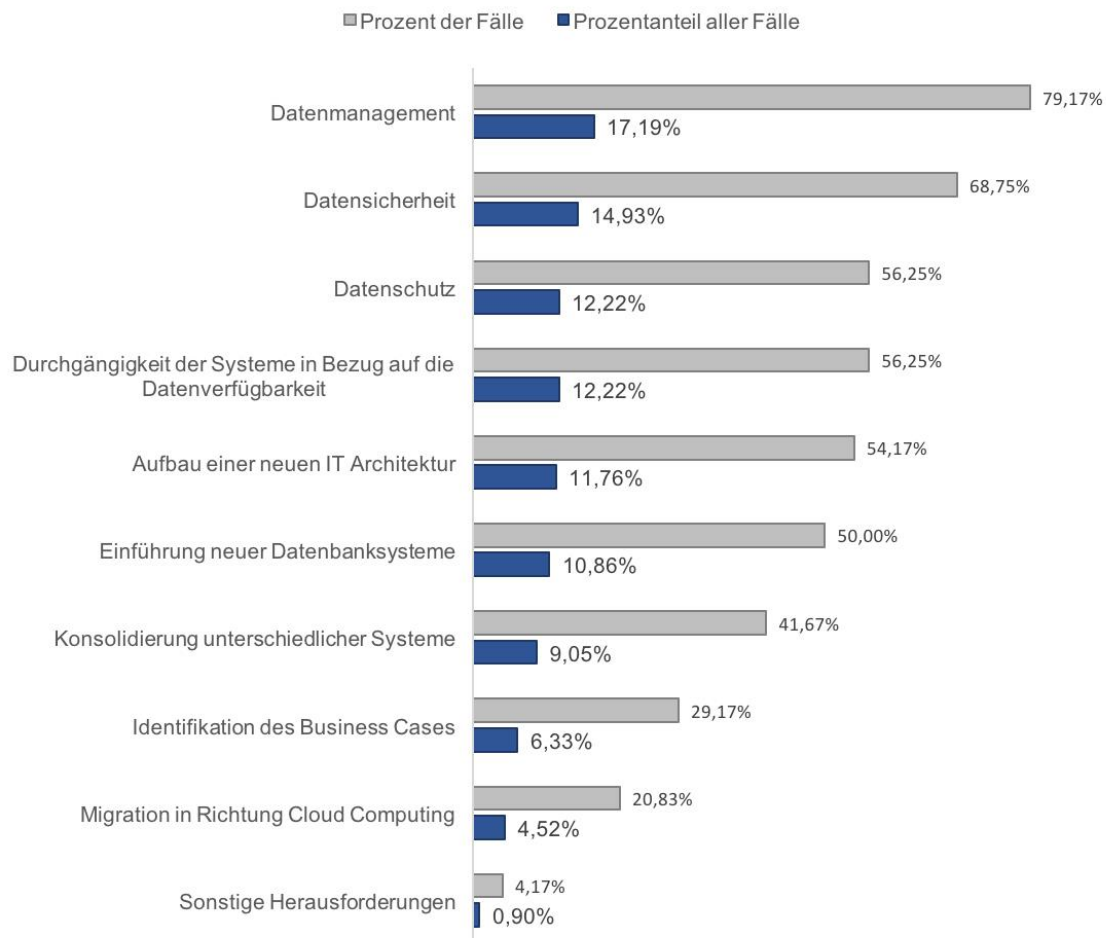


Abbildung 23: Analyse der wesentlichen Anforderungsbereiche

Fokus auf Datensicherheit und Datenschutz

Die Datensicherheit steht mit 68,75% der Nennungen bei den skizzierten Anforderungen an die IT an vorderer Stelle – direkt gefolgt vom Datenschutz (56,25%). Dies reflektiert die Sorge der befragten Unternehmen in Bezug auf den Verlust oder den Missbrauch von Daten. Aus Sicht der Befragten liegen heute kaum fundierte Strategien für wirkungsvolle Sicherheitssysteme vor. Dies gilt v.a. bei einer Migration der Datenhaltung bei Cloud Computing.

Die oben dargestellten Ausführungen zum Datenmanagement reflektieren sich bei der Durchgängigkeit der Systeme (56,25%), dem Aufbau einer neuen IT Architektur (54,17%) sowie der Einführung neuer Datenbanksysteme (50,00%) und der Konsolidierung unterschiedlicher Systeme (41,67%). Erneut zeigt sich ein wesentliches Handlungsfeld beim Aufbau der für Big Data wesentlichen Grundstruktur der Datenhaltung.

Eine Interpretation der Ergebnisse führt zu dem Ergebnis, dass Big Data für die meisten befragten IT-Executives v.a. ein Datenbankthema ist. Der Fokus liegt auf dem Datenschutz sowie der Datensicherheit, um einen Missbrauch von Daten zu verhindern.

Weniger relevant sind dagegen Themen wie die Identifikation der Business Cases und die damit verbundene Gestaltung neuer Services (29,17%) oder die Migration in die Cloud (20,83%).

Die Ergebnisse zeigen, dass sich die meisten Unternehmen noch ganz am Anfang befinden und zunächst die gewaltigen Datenmengen in den Griff bekommen wollen.

Zusätzliche Geschäftsmöglichkeiten durch Big Data stehen zumindest bei der IT kaum im Fokus. Gegebenenfalls werden diese Fragestellungen in den Verantwortungsbereich der Fachfunktionen übergeben.

Auch die Zurückhaltung bei der Migration in die Cloud reflektiert v.a. das Sicherheitsbewusstsein deutscher Unternehmen bei unternehmens- und kundenkritischen Daten. Daher sind die Ergebnisse der Befragung in dieser Hinsicht vollkommen konsistent.

3.5. Erforderliche Ressourcen und Kompetenzen

Die Lösung der mit Big Data verbundenen Herausforderungen erfordert aus Sicht einer Mehrzahl der befragten IT-Executives einen erheblichen Ressourcenaufwand in der IT (60,42%). Die exakte Verteilung ist in Abbildung 24 dargestellt.

Interessant ist in diesem Zusammenhang, dass Unternehmen besonders dann einen hohen Ressourcenaufwand in der IT ausmachen, wenn sie sich intensiver mit dem Thema Big Data befassen haben. Soweit erste Erfahrungen mit Big Data gemacht werden, lässt sich auch der Ressourcenaufwand besser abschätzen. Damit bleibt festzuhalten, dass der erforderliche Ressourcenaufwand für eine Erschließung der mit Big Data verbundenen Mehrwerte durchaus als hoch eingestuft wird.

Konzentration auf die Erweiterung der eigenen Fähigkeiten

Die befragten IT-Executives identifizieren darüber hinaus die wesentlichen Investitionsbereiche für Big Data. Diese sind in Abbildung 25 skizziert. Den Investitionsaufwand in die Entwicklung neuer Fähigkeiten durch Trainings schätzen 72,92% der Unternehmen als größten Handlungsbereich ein. Damit wird bei Big Data v.a. in die Weiterentwicklung der eigenen Mitarbeiter/innen investiert. Dies macht deutlich, dass innerhalb der meisten Unternehmen bislang wenig Wissen über das Themenfeld Big Data vorliegt und das verfügbare Know-How durch Trainings in der IT sowie in den Fachbereichen ausgebaut werden muss.

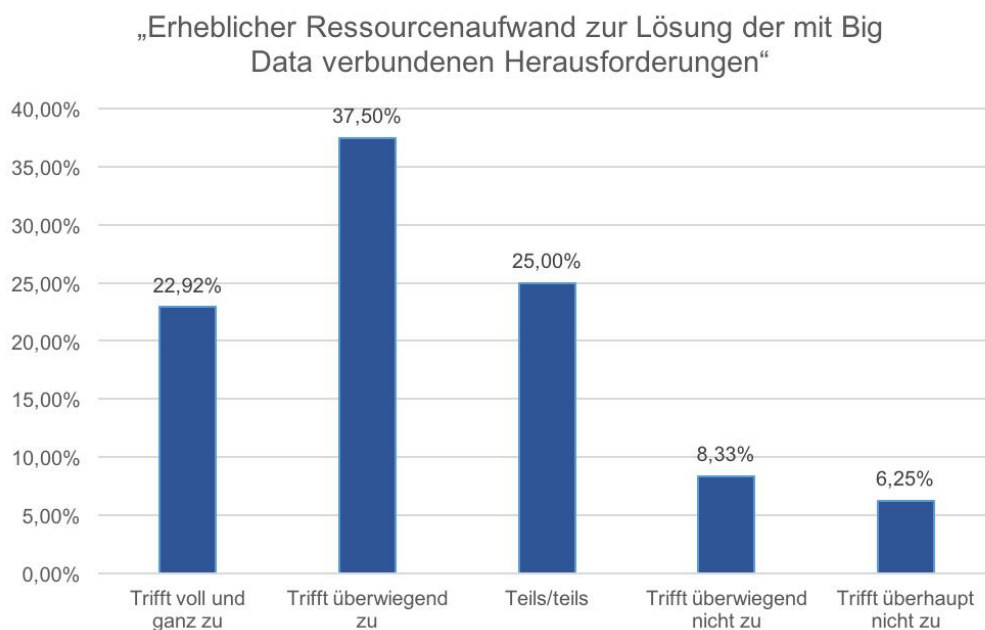


Abbildung 24: Ressourcenaufwand für Big Data

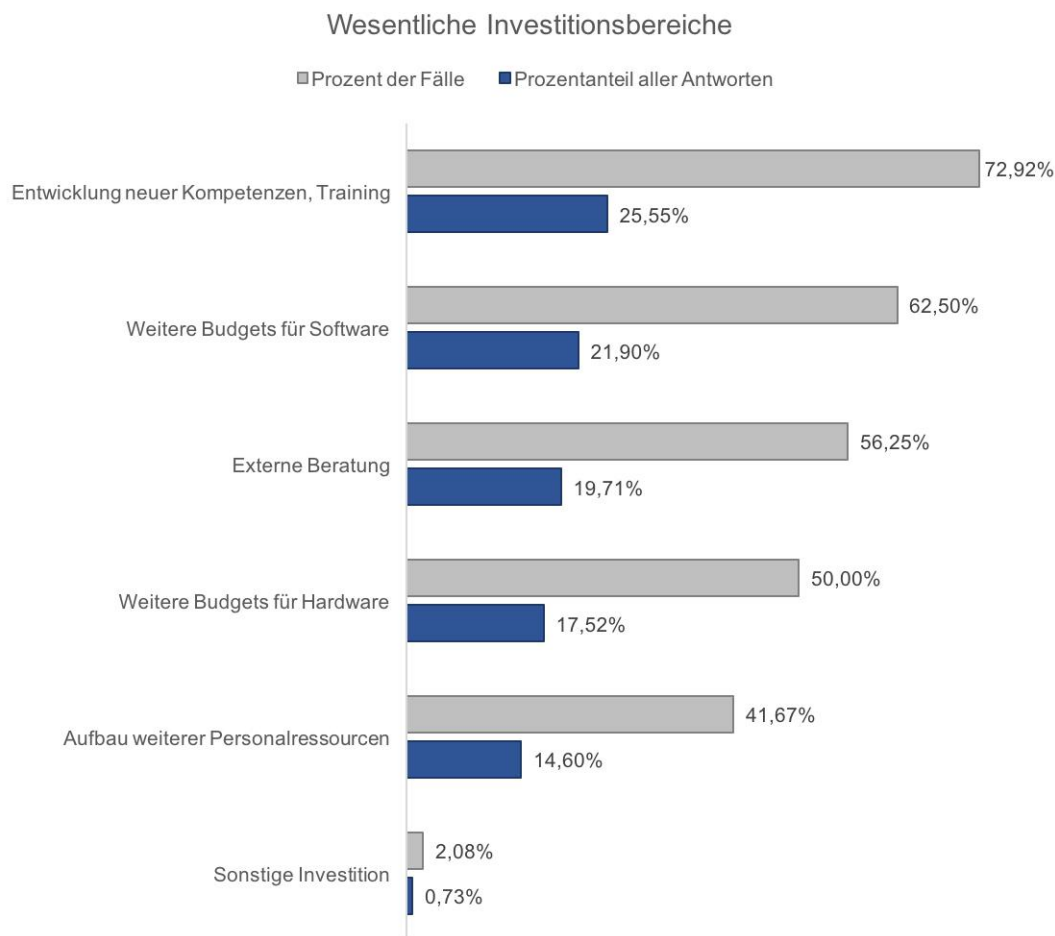


Abbildung 25: Wesentliche Investitionsbereiche

Das Fehlen der erforderlichen Kompetenzen zeigt sich auch darin, dass die Hälfte der befragten Unternehmen auf ein breites Beratungswissen von extern zugreifen muss (56,25%).

Ein weiterer Großteil an Investitionen wird von 62,50% der Befragten im Bereich Software gesehen. Dabei dürfte jedoch ein Großteil auf die Arbeitsleistung zur Implementierung und für externe Dienstleistungen anfallen, da beispielsweise das Hadoop Ökosystem als Open Source Plattform kostenlos zur Verfügung steht.

Big Data tangiert die gesamte Business- und IT-Architektur

Weitere Investitionsbereiche liegen in den Bereichen Hardware (50,00%) und in der Rekrutierung neuer Personalressourcen (41,67%). Die relative Gleichverteilung der einzelnen Handlungsfelder macht deutlich, dass es kaum Aspekte gibt, die nicht durch Big Data erfasst werden. Dies macht erneut deutlich, dass es sich bei Big Data um ein ganzheitliches Thema handelt, welches die Business- und IT-Architektur in Summe tangiert.

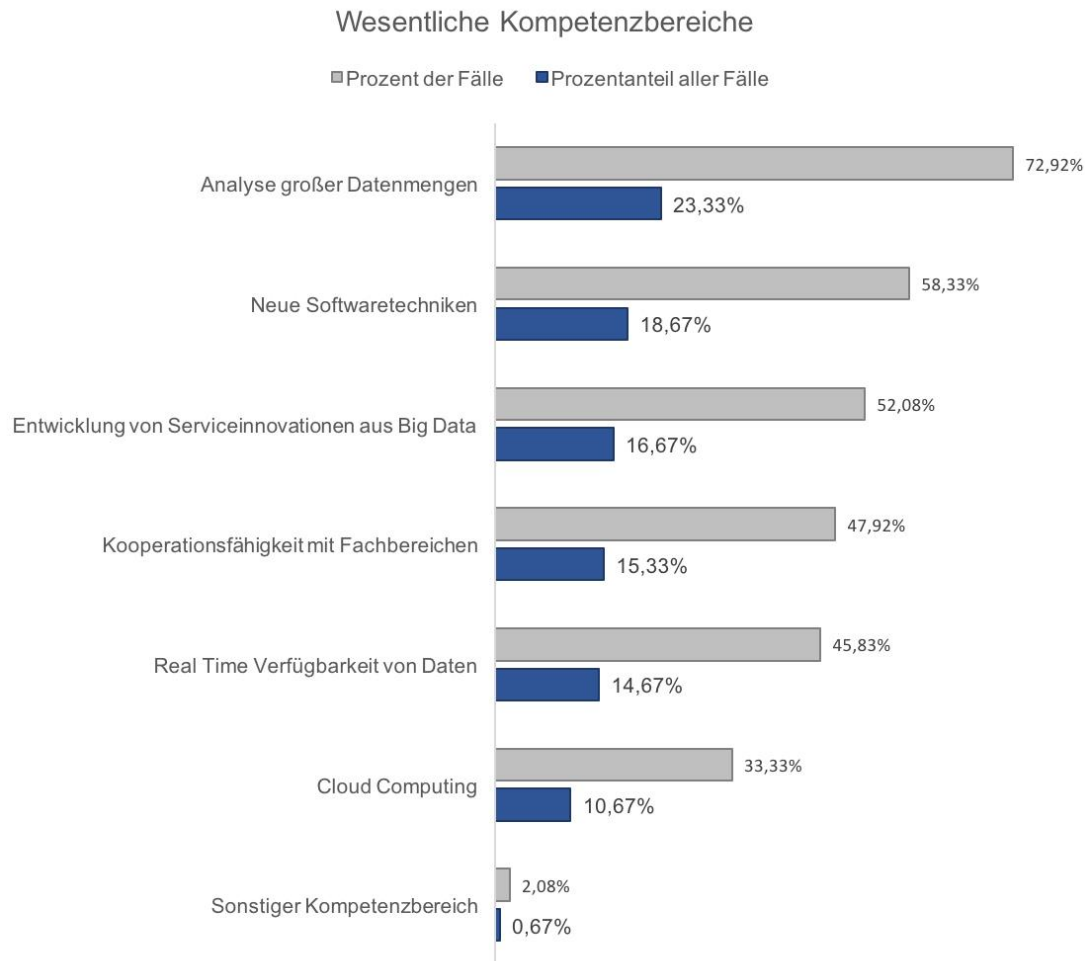


Abbildung 26: Wesentliche Kompetenzbereiche

Neue Kompetenzen für die Datenanalyse gefragt

Eine wesentliche Aufgabenstellung bei Big Data liegt in der Entwicklung neuer Kompetenzen (siehe Abbildung 25). Daher wurden abschließend die für Big Data erforderlichen Kompetenzbereiche genauer erfasst. Die Ergebnisse dieser Teilfrage sind in Abbildung 26 visualisiert.

Die meisten Unternehmen sind der Meinung, dass die IT ihre Kompetenzen im Bereich der Analyse von großen Datenmengen ausbauen muss (72,92%). Etwa zwei Drittel der Unternehmen wollen ihr Know-How im Bereich neuer Softwaretechniken weiterentwickeln und die Mitarbeiter/innen entsprechend trainieren (58,33%).

Dieses Ergebnis bestätigt erneut, dass neuere Technologien wie Hadoop, MapReduce oder NoSQL Datenbanken bislang noch nicht flächendeckend etabliert sind.

Die Entwicklung von Serviceinnovationen stellt für 52,08% der Befragten ein wesentliches Feld der Kompetenzerweiterung dar. Für jedes zweite Unternehmen ist es demnach schwierig, neue Produkte und Geschäftsfelder aus Big Data abzuleiten. Dies adressiert auch die Zusammenarbeit zwischen Business und IT bzw. den Kreativprozess bei der Entwicklung von Ideen für die Nutzung von Big Data. Hier sind weitere Vorgehensmodelle, z.B. aus dem Bereich Design Thinking sowie Methoden mit Hinblick auf die Entwicklung neuer Geschäftsmodelle erforderlich. Folgerichtig muss in 47,92% der Fälle die Kooperationsfähigkeiten mit den entsprechenden Fachbereichen ausgebaut werden.

Bei der Verfügbarkeit der Daten in Echtzeit wollen 45,83% der befragten Unternehmen ihre Fähigkeiten ausbauen. Daran ist ersichtlich, dass eine Verfügbarkeit von Daten in Realtime für viele im ersten Moment eine untergeordnete Rolle spielt und Unternehmen zunächst die Masse an Daten bewältigen wollen.

Den Ausbau von Kompetenzen zum Thema Cloud Computing sehen ein Drittel der Studienteilnehmer als wichtig an (33,33%). Die geringe Anzahl an Nennungen ist dabei auf die Zurückhaltung im Kontext der Migration in die Cloud zurückzuführen.

3.6. Zusammenfassung der Ergebnisse

Die folgende Zusammenfassung der Ergebnisse bezieht sich auf die wichtigsten Resultate der empirischen Untersuchung aus Sicht der befragten IT-Executives. Dabei erfolgt auch eine Bezugnahme auf die in Kapitel 2 dargestellten theoretischen Grundlagen.

Big Data ist aus Sicht der IT vorwiegend Datenmanagement

Der Begriff Big Data ist aus Sicht der Befragten v.a. durch die Perspektive auf Datenbanken geprägt. Für acht von zehn Unternehmen geht es hauptsächlich um die Speicherung und Bearbeitung von großen Datenmengen. Themen wie die Datenverfügbarkeit in Echtzeit oder Cloud Computing sind für ungefähr 40% der Teilnehmergruppe relevant. Im Kontext von Big Data sind diese zum jetzigen Zeitpunkt jedoch weniger relevant, da das Hauptaugenmerk auf den Herausforderungen in den Bereichen Datenmanagement und Datensicherheit liegt. Das Internet der Dinge ist für ein Drittel der Teilnehmer von Bedeutung. Dazu zählen zum Großteil Unternehmen in den Bereichen Maschinenbau, Energieversorgung und Logistik. Dies ist auf die zunehmende Vernetzung physischer Produkte mit dem Internet zurückzuführen. In der IT sind für 37,50% der befragten Unternehmen Social Networks von Bedeutung. Social Media als Datenquelle wird hauptsächlich von Unternehmen eingesetzt, die sich auf Analyse von Kundenbedürfnissen und individuelle Kundenlösungen konzentrieren.

IT als zentraler Ansprechpartner

Die IT sieht sich mit Hinblick auf diese thematische Perspektive als wesentlicher Treiber. Der wichtigste Ansprechpartner ist in sieben von zehn Fällen der CIO oder direkt der CEO. In etwa 20% der Unternehmen ist derzeit keine Führungskraft oder kein spezieller Ansprechpartner für das Thema Big Data verantwortlich. Damit nimmt die IT in der Diskussion rund um Big Data eine wichtige Rolle ein. Jedoch muss mit Hinblick auf die weitere Entwicklung geprüft werden, ob der skizzierte Themenausschnitt (Datenmanagement) das Thema Big Data in ausreichende Breite und Wertschöpfung abdeckt.

Big Data ist ein junges Thema mit hoher Entwicklungsdynamik

Ein aktiver Dialog zu Big Data erfolgt in über 70% der befragten Unternehmen. Die Mehrheit der Stichprobe (58,34%) beschäftigt sich jedoch seit weniger als drei Jahren mit dem Themengebiet. Zu den weiteren 30 Prozent zählen primär Kleinunternehmen mit bis zu 500 Beschäftigten. In Summe sind Kleinunternehmen beim Einsatz von Big Data sehr zurückhaltend. Das gibt auch der aktuelle Status quo wieder. In jedem fünften Unternehmen sind keine Bestrebungen vorhanden, Big Data Projekte durchzuführen. Dazu zählen vorwiegend Kleinunternehmen. Offensichtlich sind die Hürden beim Thema Big Data v.a. bei kleineren Unternehmen noch sehr hoch.

In vier von zehn Unternehmen wird Big Data derzeit als wichtig eingestuft. Für sechs von zehn Unternehmen nimmt die Bedeutung in Zukunft stark zu. Der gegenwertige Umsetzungsstand von Big Data Projekten verdeutlicht ebenfalls die strategische Bedeutung von Big Data. Jedes zehnte Unternehmen hat bereits ein Big Data Projekt abgeschlossen. In jeder vierten Organisation wurden Projekte gestartet und in 45 Prozent der Unternehmen sind Projekte in Planung oder unter Beobachtung.

Big Data verfügt über ein breites Zielspektrum

Die mit Big Data verbundenen Unternehmensziele verfügen über eine hohe Bandbreite. Aus Sicht der befragten IT-Executives geht es in erster Linie um die Analyse von Kundenbedürfnissen, die Erzeugung neuer Serviceleistungen und individuelle Kundenlösungen. Wesentlich sind auch Ziele wie die Optimierung von Prozessen und die Unterstützung bei Unternehmensentscheidungen. Bei den gegenwärtig analysierten Datenquellen zeigt sich jedoch ein deutlicher Fokus auf unternehmensinterne Daten. Daher ist davon auszugehen, dass die Umsetzung von Big Data Projekten mit internen Daten für IT-Executives einfacher zu gestalten ist. Aus dieser Perspektive dominieren aktuell entsprechende Projekte zur Optimierung interner Prozesse.

Eine dezidierte und explizit dokumentierte Strategie für Big Data liegt allerdings in den meisten Unternehmen noch nicht vor. Lediglich jedes sechste Unternehmen gab an, dass für Big Data eine entsprechende Dokumentation vorhanden ist.

Fokus auf interne Daten

Die Analyse des Status quo aktuell relevanter Datenquellen gibt Aufschluss über zukünftige Potentiale von Big Data. Als wesentliche Datenquelle stufen 50% der befragten Unternehmen derzeit interne Prozessdaten, maschinenproduzierte Daten und Produktionsdaten ein. Zu maschinenproduzierten Daten zählen u.a. Sensornetzwerke sowie das Internet der Dinge. Ein Drittel der Studienteilnehmer bezieht wesentliche Daten aus Social Media. Dazu zählen insbesondere Unternehmen, welche die Analyse der Kundenbedürfnisse und die Gestaltung individueller Kundenlösungen zum Ziel haben. Die Analyse von produktbezogenen Lebenszyklusdaten ist besonders im Maschinenbau von Bedeutung, da sich diese vermehrt mit Vernetzung physischer Produkten mit dem Internet beschäftigen.

Keine spezifischen Big Data Technologien im Einsatz

Zur Analyse von Daten werden aktuell kaum spezifische Big Data Technologien genutzt. In weniger als jedem zehnten Unternehmen sind Technologien wie Hadoop, MapReduce, NoSQL Datenbanken oder BI in der Cloud im Einsatz. In 35% der Fälle finden relationale Datenbanken Anwendung. Dies deckt sich mit den Ergebnissen zum Umsetzungsstand von Big Data Projekten in der Praxis und weist auf einen schwachen Reifegrad des Themas in der Breite hin.

Cloud Computing in der Erprobung

In jedem zweiten Unternehmen werden die relevanten Daten für Big Data Projekte vorwiegend in internen Rechenzentren gehalten. Beim Thema Cloud sind alle befragten Organisationen sehr zurückhaltend. Ansätze zum Thema Cloud werden vorwiegend in Pilotprojekten und noch nicht in der Breite umgesetzt. Die Analyse der Herausforderungen zeigt zum einen, dass Unternehmen die Themen Datenschutz und Datensicherheit sehr ernst nehmen. Zum anderen stellen die mit Cloud Computing verbundenen Anpassungen der IT-Architektur Unternehmen vor erhebliche Herausforderungen.

Fokus von Big Data in der IT liegt auf dem Datenmanagement

Die IT konzentriert sich in der aktuellen Entwicklungsphase von Big Data auf das Datenmanagement. Die Datenauswertung spielt nur eine untergeordnete Rolle. In einem Drittel der befragten Unternehmen erfolgt derzeit keine systematische Datenauswertung. Bei der Hälfte der Stichprobe erfolgt eine Auswertung mittels Reporting und Standardberichtswesen. Nur in jeder fünften Organisation kommen komplexere Verfahren wie Data Mining, Predictive Analytics oder Text Mining zum Einsatz. Dies gilt besonders für den Mittelstand. 35% der befragten mittelständischen Unternehmen führen derzeit keine systematische Datenauswertung durch. Eine Reaktion in Echtzeit auf dynamische Datenbestände und die Aktualität der Datenanalyse hat daher für viele Unternehmen eine untergeordnete Priorität. Der Fokus liegt auf Strukturen für die Datenverwaltung und das Datenmanagement.

Neue Herausforderungen in der IT

Etwa acht von zehn Unternehmen sagen aus, dass Big Data die IT vor erhebliche Herausforderungen stellt. Die wesentlichen Herausforderungen liegen dabei in den Bereichen Datenmanagement, Datensicherheit, Datenschutz und in der Durchgängigkeit der Systeme bei der Datenverfügbarkeit. Das Ergebnis spiegelt wieder, dass der Fokus aktuell darauf liegt, große Datenmengen zu bewältigen und den Missbrauch von Daten zu verhindern. Weitere wesentliche Herausforderungen beziehen sich auf die Weiterentwicklung der IT-Architektur sowie auf die Einführung neuer Datenbanksysteme.

Investitionen in Wissensaufbau und externe Beratung

Für die Lösung der mit Big Data verbundenen Herausforderungen ist aus Sicht der meisten Befragten ein erheblicher Ressourcenaufwand erforderlich. Dieser bezieht sich vorwiegend auf die Entwicklung neuer Kompetenzfelder bei den eigenen Mitarbeiter/innen. Dies macht erneut deutlich, dass innerhalb der befragten Unternehmen aktuell noch zu wenig Expertise zu Big Data vorliegt. Mehr als die Hälfte der befragten Unternehmen will daher auch auf externe Beratung im Umfeld von Big Data zugreifen.

Eine tiefere Analyse der erforderlichen Kompetenzfelder zeigt auf, dass der größte Handlungsbedarf im Umgang mit großen Datenmengen gesehen wird. Sieben von zehn Unternehmen sind der Meinung, dass die IT ihre Kompetenz im Bereich der Analyse von großen Datenmengen ausbauen muss.

Etwa zwei Drittel der befragten Unternehmen wollen ihr Know-how im Bereich neuer Softwaretechniken ausbauen. Dies bestätigt, dass neue Technologien wie Hadoop, MapReduce oder NoSQL Datenbanken bisher noch nicht flächendeckend bekannt sind. Weitere Kompetenzfelder im Bereich der Echtzeitverfügbarkeit der Daten sowie in Bezug auf Cloud Computing werden in der Befragung weniger stark gewichtet. Dies unterstreicht nochmals, dass derartige Fragen im Moment noch eine untergeordnete Rolle spielen.

Die CIO Studie macht in erster Linie deutlich, dass Big Data in Deutschland noch ein sehr junges Thema ist. Eine aktive Auseinandersetzung mit der Thematik findet in vielen der befragten Unternehmen seit weniger als drei Jahren statt. Die zukünftige Bedeutung von Big Data für die IT-Strategie wird allerdings als hoch eingestuft. Handlungsbedarf besteht vor allem bei der Definition von Zuständigkeiten für Big Data, der Weiterentwicklung der Zusammenarbeit zwischen Business und IT, einer expliziten und dezidierten Dokumentation für Big Data sowie dem entsprechenden Kompetenzaufbau innerhalb der IT. Ohne das notwendige Rüstzeug und das Vertrauen in neuere Big Data Technologien lassen sich die mit Big Data verbundenen Mehrwerte kaum realisieren.

4. Fallstudien:

Einblicke in Big Data Lösungen der Unternehmenspraxis

Die Ergebnisse der empirischen Befragung zum Status quo und den Perspektiven von Big Data weisen auf einen deutlichen Entwicklungsbedarf hin. Dieser lässt sich u.a. durch eine Orientierung an Fallbeispielen für die Umsetzung von Big Data Projekten decken. Daher bildet die Auswertung von Fallstudien neben der Analyse theoretischer Grundlagen und empirischer Perspektiven einen wesentlichen Kern des Big Data Reports.

Fallstudien erweitern die Sicht auf die Umsetzung von Big Data Strategien in der Unternehmenspraxis. Damit lassen sich die skizzierten theoretischen Frameworks für Big Data besser nachvollziehen. Zum anderen werden die Ergebnisse der empirischen Analyse deutlich vertieft.

Um Fallstudien in ihrem realen Umfeld aufgreifen, bewerten und vergleichen zu können, wird in der vorliegenden Studie die Fallstudienmethodik nach Yin (2013) umgesetzt. Diese basiert auf einem fünfstufigen Forschungsprozess in den Teilphasen (1) Design, (2) Prepare, (3) Collecting, (4) Analyzing und (5) Reporting. Aus Darstellungsgründen werden an dieser Stelle nicht alle methodischen Teilschritte dargestellt. Weitere Informationen zur Vorgehensweise sind über das Research Lab for Digital Business verfügbar.

Im Grundsatz basieren die Fallstudien auf einer Voranalyse relevanter Anwendungsfelder für Big Data Projekte. Dabei soll das Fallstudienspektrum die Anwendung von Big Data in der Unternehmenspraxis möglichst gut abdecken. Die Anzahl der zu untersuchenden Fallstudien definiert sich einerseits über die zur Verfügung stehenden Beispiele und andererseits über den zur Verfügung stehenden Untersuchungszeitraum. In Summe beinhaltet der Big Data Report bislang sieben Fallstudien für Big Data in der Unternehmenspraxis.

Durch die Fallstudienanalyse werden Gemeinsamkeiten und Unterschiede zwischen einzelnen Anwendungsfällen untersucht. Darüber hinaus lassen sich allgemeine Prinzipien für den Umgang mit Big Data aus einer Cross Case Analyse ableiten.

Zur Datenerhebung wurden zunächst verfügbare Sekundärdaten für jeden Einzelfall evaluiert. Dabei lassen sich Informationen von Forschungsinstituten, wissenschaftliche Publikationen, Marktforschungsberichte sowie die veröffentlichten Unternehmensinformationen selbst heranziehen. Für die Bewertung der Sekundärdaten und die Erhebung neuer Daten wurden halbstrukturierte Leitfadeninterviews umgesetzt. Die Interviews wurden telefonisch durchgeführt, digital aufgezeichnet und anschließend transkribiert. Dies sichert eine systematische und vergleichbare Datenbasis.

Für die Ergebnisauswertung der Interviews ist ein Evaluationsschema erforderlich. Dies orientiert sich an der Struktur der Interviews bzw. an den Konzepten und Modellen der Literaturanalyse. Da die Fallstudien aus unterschiedlichen Branchen stammen, ist das Evaluationsschema generisch gehalten.

Das Evaluationsschema umfasst die Teilergebnisbereiche (1) Beschreibung, (2) Unternehmen und Organisation, (3) Angebot, Services, Geschäftsmodell, (4) Kunden und Zielgruppen, (5) Kanäle, (6) Devices, (7) IT-Infrastruktur, Frontend, Backend, (8) Partner, (9) Lessons Learned.

Auf Grund des Umfangs der erhobenen Informationen lassen sich die Fallstudien an dieser Stelle nicht in allen Details beschreiben. Die skizzierten Praxisfälle geben jedoch einen guten Einblick in die Umsetzung von Big Data Projekten in der Unternehmenspraxis.

Übersicht Fallstudien

- (1) SBB Serviceoptimierung
- (2) Planung und Risikosteuerung von Ölplattformen
- (3) Smart Energy Stadt Berlin
- (4) Smart City Monaco
- (5) Predictive Maintenance in der Automobilindustrie
- (6) Industrie 4.0 für Chemie und Pharma
- (7) Connected Car eCall Service

4.1. SBB Serviceoptimierung

Die Schweizer Bundesbahn (SBB) verfügt über eine klare Zielsetzung. Diese liegt in einer Erhöhung der Passagierzahlen ohne zusätzlichen Streckenbau. Um dies erreichen zu können, muss der Betrieb des Streckennetzes optimiert werden.

Im Auftrag der SBB hat Tibco daher ein System für das Management der Streckenaktivitäten entwickelt. Dies umfasst Sensoren zur Aufnahme der Verkehrs- und Wartungsdaten, die Übertragung der Informationen, die Verarbeitung im Backend sowie die Ausgabe der automatisch entwickelten Handlungsempfehlungen über das Frontend.

In der Schweiz gibt es eine der höchst frequentierten Bahnstrecke der Welt. Das Netz soll ohne Ausbau, d.h. durch die Erhöhung der Auslastung eine steigende Passagierzahl transportieren. Die eingesetzte Lösung hat bereits 1,7 Millionen Verspätungsminuten eingespart. Hinzu kommen weitere Kosteneffekte durch eine Einsparung der Anfahrts- und Abbremsenergie.

Die SBB hat die Plattform unter der Bezeichnung *Rail Control System* entwickelt. Der Ansatz wurde durch eine Projektierungsphase initiiert. Das Hosting wird über die Tibco Cloud und die Lösung als Software-as-a-Service umgesetzt. In diesem Kontext sind neue Services durch Big Data Anwendungen entstanden.

Durch die Sammlung und Verarbeitung großer Datenmengen ist es nun möglich, diese in Echtzeit auszuwerten und somit eine automatische Steuerung des Schienensystems der SBB ausführen zu können. Dies lässt sich über die Analytics-Komponente der Plattform umsetzen, die mit der Hilfe eines vorgegebenen Regelwerkes eigenständig Handlungsempfehlungen ausarbeitet und diese umsetzen kann.

Das von einem Partner aufgebaute Sensorennetzwerk an den Zügen und im Gleisumfeld ist über Router bzw. Hotspots oder mittels Simkarten über das Mobilfunknetzwerk zur Datenaufnahme vernetzt.



Das System wird zentral über eine Leitstelle mittels Webportal gesteuert. Dies wird durch Tablets ergänzt, welche den Zugführern zur Verfügung gestellt werden. Sie dienen dazu, das Bahnpersonal zu informieren und Informationen zurückzuspielen. So bekommen die Zugführer Signalinformationen direkt auf das Tablet gespielt und können reaktiv die Geschwindigkeit des Zuges verringern. Bei Signaländerung kann dann ohne große Energiezufuhr weitergefahren werden.

Die Fast Data Plattform ist das Backbone der Lösung. Sie verarbeitet Informationen in Echtzeit und leitet Aktionen ab bzw. trifft Voraussagen für zukünftige Ereignisse.

Das Backend besteht aus drei Layern:

1. **Integration**

Die Integrationsebene ist für den Datenaustausch und Schnittstellen mit und von Drittsystemen zuständig. Aufbauend auf einer SOA-Infrastruktur werden Services standardisiert und via „API Exchange“ bereitgestellt.

2. **Event Processing**

Sind diese Daten angebunden, werden sie von dem Big Data Event Processing nahezu in Echtzeit verarbeitet, analysiert und ausgewertet. Wurden Pattern gefunden, werden diese an die dritte Ebene weitergeleitet. Schließlich sind in diesem Layer auch Business Services, wie Payment oder ein Customer Relationship Management (CRM) System hinterlegt.

3. **Visualisierung**

Das als Tibco Spotfire angebotene Tool bildet das User Interface der Data Scientists bzw. der Data Analysts. Es zeigt die aggregierten Daten in visuellen Dashboards (Live Data Marts) aufbereitet an und erzeugt somit neues Wissen. Auch Analysen zur Vorhersage von zukünftigen Ereignissen werden hier dargestellt.

Die Umsetzung von Big Data Lösungen hat auch Auswirkungen auf die Kooperation zwischen den beteiligten Partnern. Das Geschäftsmodell von Tibco liegt ursprünglich in der Integration und Vernetzung von IT-Infrastrukturen. Vor allem serviceorientierte Architekturen und die lose Kopplung von Business und IT gehören früh zu den Kernkompetenzen des Unternehmens. Mit In-Memory zur Verarbeitung großer Datenmengen ging Tibco den ersten Schritt in Richtung Big Data/IoT. Getrieben von diesem Trend baute das Unternehmen sein Portfolio hin zu Echtzeitverarbeitungs- und Event-Processing-Produkten aus.

Dies skizziert einen Veränderungsprozess auf Anbieterseite, weg von reinen Integrationsprojekten und hin zum Produkt- und Lösungsgeschäft. Aus Sicht des Geschäftsmodells wurde die Entwicklung der Plattform über einen Lizenzansatz und die Projektimplementierung abgerechnet. Darüber hinaus werden aktuell weitere Geschäftsmodelle getestet.

Neben klassischen Modellen liegen zunehmend neuartige Ansätze wie Software- bzw. Plattform-as-a-Service inklusive Cloud-Hosting vor. Zukünftig ist es denkbar, dass Softwarehersteller in Kooperation mit anderen Unternehmen Lösungen entwickeln und diese partnerschaftlich vertreiben.

Das Partnernetzwerk der vorliegenden Fallstudie umfasst ein Ökosystem aus Sensorik- und Hardwarepartner sowie die SBB als Generalunternehmer.

Die skizzierte Anwendung lässt sich v.a. auf Basis neuerer technologischer Entwicklungen umsetzen. Noch vor einigen Jahren wäre die Umsetzung auf Grund hoher Technologiekosten (z.B. für RAM) und zu geringer Leistung, Konnektivität und Verfügbarkeit des Netzes gescheitert.

4.2. Planung und Risikosteuerung von Ölplattformen

Für Ölplattformen ist die Umsetzung eines fundierten Betriebs- und Servicekonzepts wesentlich. Dieses basiert u.a. auch auf der Früherkennung möglicher Schwachstellen. Daher hat das Beratungshaus CSC gemeinsam mit der Ölförderindustrie eine Plattform zur Produktions- und Risikosteuerung konzeptioniert und umgesetzt. Das wesentliche Ziel liegt darin, Wartungszyklen zu optimieren und das Angebot optimal an die Nachfrage anzupassen.

Hierzu werden relevante interne und externe Daten, wie z. B. Wind, Wetter, Druck oder Materialflüsse über Sensoren und Datenquellen Dritter im Kontext einer Big Data Plattform vernetzt. Das Betriebs- und Risikomodell kann damit automatisch gesteuert werden. Somit lässt sich beispielsweise bei sinkender Nachfrage an den Ölbörsen die Produktionsmenge herunterfahren oder bei defekten Ölbohrern das hinterlegte Notfallkonzept umsetzen.

Aus der Zusammenarbeit unterschiedlicher Unternehmen wurde mittlerweile ein standardisierter Lösungsansatz unter dem Label *Digital Oil Field* entwickelt. Dieser ist bereits bei verschiedenen Ölplattformen im Einsatz.

Die Kernidee *Digital Oil Field* folgt einem mehrstufigen Prozess. So wird im ersten Schritt die Ölförderung in einem digitalen Modell in Echtzeit abgebildet. Daraus lassen sich Potentiale zur Prozess- und Wartungsoptimierung ableiten. Dies wiederum senkt die Kosten. Durch die Integration des Systems in die Ölbörsen lässt sich zeitgleich die Profitabilität steigern. Sinkt der Gaspreis unter einen bestimmten Wert, können beispielsweise kostenintensive Produktionsprozesse (Fracking) gedrosselt werden.

Bei der Umsetzung des *Digital Oil Field* ist der Betreiber der Ölplattform der Business Owner. Die verwendeten Sensoren sind für Brücken, Temperatur, Druck oder Vibration im Einsatz und werden von weiteren Partnern zugeliefert. Dabei sind die Sensoren u.a. an Bohrköpfen angebracht. Die Akteure steuern hauptsächlich die Schieberegler in den Pipelines. Da in Fördergebieten mit extremer Kälte zu rechnen ist, wird bei der Datenübertragung weitgehend auf Kabel verzichtet. Es sind neuartige Geräte im Einsatz, die gemeinsam mit Elektronikherstellern entwickelt wurden.



Zur Steuerung der Infrastruktur gibt es einen zentralen Leitstand, der das Monitoring aller Maschinen übernimmt. Dort sind Dashboards über Portalzugänge ersichtlich. Zur Wartung werden mobile Endgeräte genutzt, die u.a. Augmented Reality Technologien beinhalten, um die Fehlersuche und -behebung zu vereinfachen. Die Komponenten im Backend sind aufgrund ihrer zahlreichen Aufgabengebiete vielschichtig. So setzt das *Digital Oil Field* eine Integrationsebene ein, auf der die Sensordaten zusammengefasst werden. Hierzu gehören auch Partnerlösungen zur Bilddaten- und Videodatenanalyse. Schließlich liegt eine prozessuale und regelbasierte Schicht vor, die die Simulations- und Steuerungsfunktion übernimmt. Die Risikomodelle werden über die Data Mining Komponente zusammengestellt. Ebenso wird die Absatzanalyse über die gleiche Funktion in Echtzeit in das Entscheidungsmodell integriert. Die Streamingkomponente bindet die Daten, wie Wetterprognosen oder Aktienkurse von den Datenbanken externer Service Provider an. Das Hosting findet auf der CSC Cloud, einer hybriden Cloudumgebung, statt.

Für die Umsetzung entsprechender Lösungen ist ein Ökosystem unterschiedlicher Kompetenzträger erforderlich. Im vorliegenden Fall übernimmt CSC eine zentrale Koordinatorenfunktion. Für die Hardware steht General Electric als Partner zur Verfügung. Weiterhin gibt es Technologiepartner wie Oracle oder Hadoop und angebundene Servicedienstleister, u. a. für die Videoanalyse. Die Angebote und Lösungen der einzelnen Teilpartner müssen sinnvoll integriert werden. Dabei entstehen kaum Aufwände für Neuentwicklungen.

Die vorliegende Fallstudie macht deutlich, dass viele Big Data Anwendungen heute bereits einen hohen technologischen Reifegrad erreicht haben. Daher liegen die Einschränkungen mit Hinblick auf eine breite Umsetzung in vielen Fällen weniger im Technologiebereich. Im Kern erfordert Big Data jedoch neue Formen der Zusammenarbeit zwischen unterschiedlichen Wertschöpfungspartnern und ein breites Kompetenzspektrum.

Damit ist auch eine neue Rolle der IT-Abteilung verbunden. Da die Lösungen durch die Fachbereiche vorangetrieben werden und bereits vorliegende Lösungselemente im Sinne einer Serviceorchestrierung zusammenzuführen sind, wandeln sich die Anforderungen an die IT. Dies erhält zusätzliche Aufgaben als Partner der Businessunits bei der Service- und Datenorchestrierung und der Umsetzung von Service Level Agreements. Dagegen werden Teilaufgaben beim Aufbau und Betrieb eigener Infrastrukturen abgebaut.

Die skizzierten Änderungen haben jedoch auch Einfluss auf das Geschäftsmodell der Anbieter. In klassischen Beratungsprojekten werden Dienstleistungen nach Tagessatz berechnet, IT Projekte folgen häufig einer Kostenlogik nach Zeit und Material. Für das *Digital Oil Field* allerdings gibt es beim Geschäftsmodell fundamentale Veränderungen. So erfolgt die Abrechnung per Anzahl der verwendeten Geräte, dem Datenvolumen sowie auf Basis unterschiedlicher Services.

4.3. Smart Energy Stadt Berlin

Die *Stadt Berlin* hat gemeinsam mit *Bosch Software Innovations* und *eMessage* (ein Betreiber von Pagerfunksystemen in Deutschland) auf Basis der gesetzlichen Anforderungen ein neues System zur Steuerung von rund 50.000 Anlagen für Energieerzeugung, -speicherung und -verbrauch entwickelt.

Das Management der Berliner Energieerzeugung und -abnahme erfolgt im Rahmen des Aufbaus einer intelligenten Energiesteuerung. Dazu zählt ein virtuelles Kraftwerk, welches sowohl kommerziell als auch netzstützend Angebot, Nachfrage und Zwischenspeicherung im Energiemarkt in Einklang bringt. Dabei werden relevante Systeme über Sensoren zur Datenaufnahme vernetzt, um stets das aktuelle Angebot und die momentane Nachfrage im Energiebereich transparent zu machen.

Basierend darauf steuert die Prozess-, Regel- und Analyseplattform im Hintergrund die Zu- und Abschaltung der Stromerzeuger (Kraftwerke) bzw. Stromverbraucher (z.B. Straßenlaternen). Aus kommerzieller Sicht werden die Preise ebenfalls auf Basis von Angebot und Nachfrage optimiert.

Für die Umsetzung der Lösung wurde ein gemeinsames Nachrichtenformat definiert. Zusätzlich war es erforderlich, die bisherigen Nachrichtensender und -empfänger durch neue, wesentlich kleinere Geräte abzulösen. Somit spart der Betreiber nicht nur Flächen-, sondern auch hohe Wartungskosten. Durch ein neues Nachrichtenformat steigt die Erreichbarkeit der Geräte.

Zusätzlich kann durch signierte Nachrichten die Sicherheit der Gesamtlösung erhöht bzw. die Möglichkeit für Manipulationen gesenkt werden. Zu den verwendeten Geräten gehören Steuereinheiten, die in den Blockheizkraftwerken, in den Straßenbeleuchtungen, Nachtspeicheröfen, PV-Anlagen oder Schleusen verbaut sind. Diese messen die aktuellen Produktions- oder Verbrauchsmengen und können bei Bedarf die Anlage regeln.

Alle Einheiten sind dauerhaft über das Pagerfunknetz von *eMessage* miteinander und mit dem zentralen Backend vernetzt. Perspektivisch ist die Lösung dafür konzipiert, auch Laufwasserkraftwerke, Ladesäulen, Power2Heat- bzw. Windanlagen anzubinden.



Auf das Frontend kann via Webportal zugegriffen werden. Jeder Kunde kann sich die Ansicht modular zusammenstellen. Zu sehen ist das Monitoring (Zeitreihen), das Journal (Logfiles), der Verfügbarkeitskalender sowie die Statusüberwachung und Lage der Anlagen. Die Besonderheit innerhalb des Projektes liegt in der grafischen Anpassung des neuen Systems an das Altsystem. Das Ziel der Stadt Berlin lag darin, die Eingewöhnung im Sinne von Lernkurven möglichst gering zu halten, da die Mitarbeiter auch in kritischen Situationen schnell und sicher mit der Software arbeiten müssen. Daher wurde die Visualisierungskomponente nahezu vollständig an das Altsystem angeglichen.

Stromnetz Berlin hat die zentrale Koordinationsfunktion des Projektes inne. Zu den weiteren Partnern gehören *Keben* als # Gateway- und Empfangseinheitenhersteller sowie *eMessage* für das Netzwerk und das Sendeformat.

Mögliche Anwender für entsprechende intelligente Systeme der Energiesteuerung finden sich global von Energiehändlern über Unternehmen mit Erzeugungs-, Speicherungs- und Verbrauchsanlagen sowie Verteilnetzbetreibern. Die Perspektiven entsprechender Systeme liegen nicht nur in der Einsparung von Prozess- und Betriebskosten. Mit entsprechenden Lösungen ist auch die Möglichkeit verbunden, neue und innovative Geschäftsmodelle am Markt zu etablieren.

Erneut zeigen sich auch Veränderungen am Geschäftsmodell auf Anwenderseite. *Bosch Software Innovations* bietet die Lösung als Software-as-a-Service auf Basis der Bosch IoT Suite an. Auf Kundenwunsch kann die Anwendung on premise auf Kundensystemen installiert werden.

Die Lösung ist entsprechend sowohl für SaaS als auch für vor Ort Installationen konzipiert. Im Anwendungsfall wurde bei Stromnetz Berlin ein Signaturserver, ein Prozessserver sowie ein Messageserver beim Kunden installiert.

Der Preis der Lösung setzt sich dynamisch aus unterschiedlichen Parametern zusammen. Dies umfasst beispielsweise die Leistung der Anlagen, verwendete Module sowie Anlagentypen. Das Ziel der Lösung liegt darin, über das Preismodell an den Einsparungen des Betreibers zu partizipieren.

Zusammenfassend weist die Fallstudie auf einen erheblichen, quantifizierbaren Wert hin, der durch vernetzte IT-Lösungen auf Basis von Big Data geschaffen wurde. Die Herausforderung liegt in der kontinuierlichen Abstimmung im partnerschaftlichen Verbund. Auch die Anbindung der überholten Bestandsinfrastruktur ist eine Herausforderung. Hier sind häufig unterschiedliche Systeme und Geschwindigkeiten bei der Weiterentwicklung der IT-Architektur zu berücksichtigen.

4.4. Smart City Monaco

Der *Stadtstaat Monaco* setzt im Sinne der Smart City Idee auf moderne IT-Infrastrukturen, die die Lebensbedingungen in der Stadt verbessern sollen. Hierzu wurde gemeinsam mit *Bosch Software Innovations* ein Proof-of-Concept entwickelt und getestet.

Dieser basiert auf einer zentralen Datenplattform und der Vernetzung einzelner Domänen. Die Domänen beziehen sich auf unterschiedliche Funktionen.

Zum Funktionsumfang gehört beispielsweise ein Crowd Management für Fahrstühle. Dieses ist an die entsprechende Videoüberwachung angeschlossen und löst einen automatischen Alarm an der Leitstelle aus, sobald wegen der Überfüllung von Fahrstühlen interveniert werden muss.

Darüber hinaus soll die Plattform den Verkehr in der Stadt optimieren. So können die Einwohner und Besucher Informationen über Baustellen oder verfügbare Parkplätze abrufen. Außerdem wurden Müllfahrzeuge mit einem GPS-Gerät ausgestattet, so dass überfüllte Straßen bei der Routenplanung zu vermeiden sind.

Dies bildet die Grundlage für einen weiteren Anwendungsfall im Bereich Waste Management. Bürger werden nicht nur über den Tag der Abholung informiert, sondern ebenfalls über die aktuelle Geoposition des Fahrzeugs. Über die Mobile App ist es möglich, den Standort einer Verschmutzung zu dokumentieren (Foto) bzw. der Stadtverwaltung zu melden.

Der Kern der Lösung liegt in der Bosch IoT-Plattform. Sie fungiert u.a. als Service-Brokering Plattform, vernetzt bestehende Services wie Baustellen- oder Müllwageninformationen und bietet Perspektiven für neue Dienste. Weiterhin steuert die Plattform das Prozessmanagement. Letztlich liegt die Stärke der Plattform in der Integration bestehender IT-Infrastrukturen.

Die Dienste der neu geschaffenen Lösung wurden bereits eingangs skizziert. Hierzu zählen das Crowd Management, das Traffic Management sowie das Waste Management. Alle Dienste basieren auf Big Data Modellen, welche eine Vielzahl von Daten aufnehmen, nach bestimmten Regeln (Standort oder Nutzerprofil) auswerten und dem Nutzer visualisieren.



Der Nutzer, Bürger oder Besucher der Stadt Monaco kann sich im Appstore die Smart City App herunterladen. Diese bietet die Möglichkeit, ein Nutzerprofil für Fußgänger oder Autofahrer einzufügen und basierend darauf, nur die relevanten Dienste angezeigt zu bekommen.

Weiterhin nutzt die Stadtverwaltung einen Portalzugang, über welchen Beschwerden über verunreinigte Plätze mit Foto und Beschreibung eingehen und abgearbeitet werden können.

Die Smart City Plattform bildet das zentrale System im Backend. Sie beinhaltet ein Prozess- und Regelmanagement, eine Integrationsebene sowie einen Profileserver, auf dem Nutzer ihr Profil ablegen können. Darüber hinaus sind im Backend die angebundenen Bestandssysteme abgebildet.

Das Partnerportfolio für die vorliegende Fallstudie ist umfangreich. Neben einer intensiven Zusammenarbeit mit örtlichen Behörden zur Definition der Anwendungsfälle, sind viele Service Provider, wie für die Müllabfuhr, die Fahrstühle oder dem Parkraummanagement, direkt involviert.

Das Geschäftsmodell mit Smart City Lösungen ist eine große Herausforderung. So amortisieren sich die Investitionen innerhalb der Wahlperioden meist nicht. Weiterhin ist die Zahlungsfähigkeit bei Städten oftmals sehr angespannt.

Weiterhin agiert das Unternehmen als Technologielieferant für Bauprojektträger, innerhalb von neuen Projekten oder Investitionen, z. B. in Verkehrsleitsystemen oder Anschaffung neuer Müllfahrzeuge. Das Geschäftsmodell für Städte ist daher noch nicht final definiert.

Insgesamt liegt global ein hoher Bedarf an Smart City Lösungen vor. Diese sind einfacher in Neubauprojekten zu realisieren, als auf Basis der bestehenden Infrastruktur. Neben dem noch unklaren Geschäftsmodell ist vor allem die Zusammenarbeit mit verschiedenen Behörden unterschiedlicher Parteien eine große Herausforderung.

4.5. Predictive Maintenance in der Automobilindustrie

Predictive Maintenance basiert auf der Auswertung großer Datenmengen in der Produktion und der Vorhersage produktionsrelevanter Faktoren. *IBM* hat gemeinsam mit der deutschen Automobilindustrie ein Vorhersagemodell auf der Basis einer Data Mining Lösung entwickelt.

Dabei geht es um die Einflussfaktoren auf die Ausschussrate bei der Produktion von Zylinderköpfen. Durch ein tieferes Verständnis produktionsrelevanter Faktoren kann der Produktionsprozess besser analysiert, nachvollzogen und gesteuert werden. Dazu werden alle relevanten, internen und externen Datenquellen zusammengefasst und ein Modell zur Vorhersage der Ausschussrate entwickelt. Schließlich lassen sich relevante Faktoren aktiv während der Produktion steuern.

Der Lösungsansatz beinhaltet u.a. die Entwicklung eines Datenmodelles. Dazu werden zunächst interne Daten wie Druck, Strom, Zeit und Wärme aggregiert. Das Modell wird darüber hinaus durch externe Daten angereichert (z.B. Wetterdaten). Diese Informationen bilden die Inputparameter des Vorhersagemodells.

Auf der Grundlage empirischer Zeitreihen werden nun der Einfluss der Inputdaten auf die Ausschussrate sowie der kausale Effekt bei Variation einzelner Daten geschätzt. Die Lösung umfasst entsprechend ein selbstlernendes System zur Vorhersage von Wartungseinsätzen und Qualitätsabweichungen. Über den Zeitverlauf lässt sich bei steigender Prognosegenauigkeit durch die Variation der Inputdaten ein erheblicher Einfluss auf relevante Zielgrößen erzielen.

Im Kern führt das selbstlernende Vorhersagemodell zu neuen Möglichkeiten, die in der Automobilindustrie bereits genutzt werden. Durch den Ansatz lässt sich der Ausschuss an Zylinderköpfen verringern. Damit sinken die Fehlerkosten. Gleichzeitig hat das Modell einen positiven Einfluss auf die Fehlerrate und die Produktionsqualität. Als Datenquelle kommt neben den Logfiles der Maschinen sowie den bereits hinterlegten Stammdaten auch ein Netzwerk aus Sensoren zum Einsatz. Dieses wird durch einen Hardwarepartner bereitgestellt. Die dauerhafte Netzwerk-anbindung in den Produktionshallen erfolgt über WLAN oder über zentrale Knotenrechner.



Die Industrieanwendung hat unterschiedliche Möglichkeiten, um auf die Informationen zuzugreifen. Es handelt sich dabei um einen zentralen Portalzugang, der über SPSS visualisierte Reports zum Produktionsprozess bietet. Darüber hinaus wurden zur Visualisierung Applikationen für Tablet auf der Grundlage von HTML5 programmiert.

Das Backend der Lösung ist durch ein Big Data Verarbeitungsprotokoll sowie die IBM Big Data Referenzarchitektur geprägt. Diese setzt sich aus einer DB2 Datenbank bzw. In Memory Technologie (IBM Informix) zur Speicherung der Daten zusammen.

Weitere relevante Komponenten beziehen sich auf ein auf Hadoop basierendes Infosphere zur verteilten Verarbeitung der Daten, IBM SPSS für Predictive Analytics Anwendungen, Cognos als Data Warehouse bzw. Business Intelligence Tool sowie einen standardisierten Big Data Governance Prozess.

Die Fallstudie verdeutlicht erneut die Notwendigkeit für eine Zusammenarbeit in Lösungsnetzwerken. Neben den Unternehmen der Automobilindustrie sind in den Lösungsansatz unterschiedliche Unternehmen eingebunden. Dies umfasst neben der IBM auch Partner für die Entwicklung und Implementierung der Sensortechnologie.

Insofern ist für die Umsetzung entsprechender Big Data Projekte eine hohe Handlungsmotivation auf Anwenderseite erforderlich. Durch den hohen Skalierungseffekt lassen sich umfangreiche IT-Projekte zu akzeptablen Kosten realisieren. Ein wesentlicher Problembereich liegt dabei nach wie vor im Fehlen definierter Marktstandards. Dies erschwert eine langfristige Planung auf Anwenderseite. Gleichzeitig müssen sehr häufig Individuallösungen umgesetzt werden, die von der etablierten IT Infrastruktur gelöst werden. Dies erschwert die Berechnung individueller Wirtschaftlichkeitsanalysen.

4.6. Industrie 4.0 für Chemie und Pharma

Industrie 4.0 basiert auf der Digitalisierung von Produktionsprozessen und der darauf aufbauenden Nutzung relevanter Produktionsdaten für neue Servicemodelle. Aktuelle Schätzungen gehen davon aus, dass bisher nur ein kleiner Anteil der relevanten Produktionsdaten verfügbar ist und aktiv genutzt wird. Daher hat *BridgingIT* gemeinsam mit der Chemie- und Pharmaindustrie ein Konzept zur Nutzung von Daten in relevanten Produktionsprozessen entwickelt.

Das Konzept ermöglicht eine Steigerung der Flexibilität in der Produktion durch eine Optimierung der Planung, das schnelle Wechseln von Werkzeugen und geringere Stillstandszeiten. Durch die Optimierung der Prozesse wird eine Steigerung der Prozess- und Erzeugnisqualität realisiert. Dafür wurde ein umfangreiches Datenmodell entwickelt. Dies umfasst die Datenerzeugung, -integration, -speicherung und -auswertung.

Der Lösungsansatz umfasst zwei Phasen. In der Vorstudie werden alle Rahmenparameter definiert, ein Datenmodell mit internen und externen Informationen erstellt, die Einflussfaktoren untersucht, ein internationaler Technologievergleich durchgeführt sowie ein Konzept und eine Testarchitektur definiert.

Das Zielbild liegt in der Entwicklung hin zu einer kleinteiligen, personalisierten und flexiblen Produktion. Dazu soll die Fertigung ereignisabhängig und eigenständig gesteuert werden. Die Maschine stellt ihre Priorisierung selbstständig zusammen.

Perspektivisch sollen auf dieser Basis kundenrelevante Informationen in den Produktionsprozess integriert werden. Dabei geht es darum, die Vertriebs- und Kundeninteraktionen zu digitalisieren, um ein Vorhersagemodell für den folgenden Absatz zu entwickeln. Außerdem ist eine weltweite Umsetzung bei allen Produktionsanlagen geplant.

Die Digitalisierung von Produktionsprozessen bietet eine Vielfalt neuer Servicemodelle. Zu den Anwendungsfällen gehört die Vorhersage über die Wahrscheinlichkeit eines Defekts der Maschine, so dass Servicetechniker gezielt eingesetzt werden und Stillstandszeiten sowie Servicekosten optimiert werden können.



Industrie 4.0

Durch die digitale Abbildung der Maschine ist es außerdem möglich, den Werkzeugtausch und dessen Justierung zu verbessern. Das schafft Flexibilität in der Größe und dem Wechsel von Fertigungschargen. Weiterhin lassen sich Qualitätsabweichungen automatisch feststellen und prognostizieren.

Diese proaktive Vorgehensweise führt die Erzeugnisqualität dauerhaft auf ein sehr hohes Niveau. Sobald diese trendbasierten Funktionen vollumfänglich implementiert sind, das Regelwerk fehlerfrei funktioniert und Aktoren digital angesteuert werden können, kann eine ereignisbasierte Architektur den Produktionsprozess eigenständig und automatisiert steuern.

Grundlage des Datenmodells sind die aus den Maschinen gewonnenen Sensorinformationen wie Druck, Zeit und Temperatur. Diese werden über einen Bus vernetzt. Das Frontend fokussiert sich innerhalb der Fallstudie auf ein zentrales Portal zur Auswertung der Daten und Steuerung der Maschinen.

Im Backend sind drei Architekturebenen abgedeckt. Der Bus sammelt auf der Integrationslayer die Daten der Sensoren. Anschließend werden diese in der zweiten Layer prozessiert, verarbeitet, analysiert und in der BridgingIT Cloud gespeichert. Bei der Verarbeitung erfolgt die Anbindung von Drittsystemen, wie ein ERP oder ein Produktionssteuerungssystem. Im letzten Schritt werden die aggregierten Daten aufbereitet und angezeigt. Das Ergebnis ist eine Handlungsempfehlung basierend auf einem vorherigen Ereignis.

Erneut zeigen sich signifikante Veränderungen auf Geschäfts- und Betriebsmodellenebene. Diese sind für Anbieter und Anwender relevant. Das Anbietermodell entwickelt sich zunehmend in Richtung von Software- und Plattform-as-a-Service Leistungen. So werden weniger klassische Projekttag- und Festpreise angeboten. Die Kooperation basiert eher auf der gemeinschaftlichen Entwicklung von Lösungen und der tiefen Integration in Anwenderprozesse. Die neuen Produkte haben den Charakter von Serviceleistungen. So wird keine Softwarelizenz für on-premise Lösungen verkauft, sondern über das Hosting in der eigenen Cloud ein Software-as-a-Service Angebot umgesetzt.

Erhebliche Herausforderungen bei der Umsetzung von Big Data Lösungen liegen entsprechend in der Erarbeitung einer gemeinsamen Sprache zwischen den Stakeholdern und Projektbeteiligten. Die Umsetzung entsprechender Referenzarchitekturen verfügt über einen hohen Innovationsgrad. Daher ist eine überproportional hohe Innovationsbereitschaft auf Anwender- und Anbieterseite bei der Umsetzung entsprechender Lösungen erforderlich.

4.7. Connected Car eCall Service

Mercedes Benz führte in 2012 gemeinsam mit *Bosch* ein System zur Ausführung eines automatischen Notrufes ein – den Mercedes Benz eCall. Er identifiziert über die Auslösung des Steuergeräteairbags einen Aufprall, kann diesen klassifizieren und im Notfall die Bosch Notrufleitstelle alarmieren. Diese erhält den aktuellen Standort des Fahrzeuges, die Anzahl der Insassen sowie weitere Informationen zum Fahrzeug. Basierend darauf verständigen und koordinieren die Agenten die nächstgelegene, passende Rettungseinheit. Der Mercedes Benz Notruf ist heute in mehr als 22 Märkten in jedem Neuwagen als Serienausstattung vorhanden und kann 24 Stunden lang, 7 Tage die Woche, innerhalb weniger Sekunden Kontakt zum Fahrer aufnehmen und die passende Rettung koordinieren.

Als wesentliche Motivation sah die *Daimler AG* eine erhöhte Nachfrage nach Sicherheitslösungen seitens der Kundschaft. Um sich zeitlich vor der gesetzlichen Einführung eines eCall-Systems zu positionieren, setzte der Automobilkonzern die flächendeckende Einführung innerhalb bestimmter Mercedes Benz Modelle um.

Mittlerweile hat sich aus der einstigen Projektorganisation eine neue Markengesellschaft, die MB Connect, gegründet. Sie hat den Auftrag, vernetzte Services für das Automobil von morgen zu entwickeln und arbeitet als eigenständiges Ressort parallel zur internen IT-Abteilung.

Bereits vor einigen Jahren versuchten sich einige Automobilhersteller im Verkauf von vernetzten Services. Damals zeigte sich, dass das beratungsintensive Produkt mit geringen Margen schwierig über eine Subskription zu vertreiben ist. Daher entschloss sich die *Daimler AG*, den MB Notruf in die Serienausstattung zu integrieren. Dort ist er unter dem Punkt 350 bzw. neuerdings 351 als Sicherheitspaket zu finden.

Neben positiven Imageeffekten und neuem Wissen in der Unfallforschung ist das System ein wichtiger Enabler für neue vernetzte Services, wie den Concierge Service, innerhalb des Connect Me Portfolios.



Für den Endkunden entsteht ein Sicherheitsfeature, welches ihm im Falle eines Unfalls die passende Hilfeleistung – auch ohne sein Interagieren – direkt zum Ort des Geschehens senden kann. Die Rettungskette wird durch die verwendeten Technologien fast vollständig automatisiert.

Möglich wird dies mithilfe einer intelligenten Big Data Lösung im Hintergrund, die die Vielzahl an Events aufnimmt, die Daten integriert und überträgt, basierend auf den Präferenzeinstellungen einen Agenten in der passenden Sprache zuteilt und entsprechend des Standortes den optimalsten Rettungsdienst im Umkreis identifiziert.

Nachdem ein Subskriptionsmodell in der Vergangenheit nicht zum gewünschten Erfolg geführt hat, ist der automatische Notruf nun in der Serienausstattung zu finden. Der Service ist über den gesamten Lebenszyklus des Fahrzeuges verfügbar, selbst beim Halterwechsel. Durch die Änderung im Geschäftsmodell hat es Daimler geschafft, hohe Kostenblöcke, wie die regelmäßigen Zahlungen und die Wechselkosten bei Halterwechsel, zu eliminieren.

Das System steht jedem Neuwagen- und Gebrauchtwagenkäufer bestimmter Fahrzeugklassen ab einem gewissen Baujahr zur Verfügung. Es ist nicht personen- sondern fahrzeuggebunden.

Per Aufprallsensor wird der Unfall detektiert. Das Signal gelangt anschließend an das Airbagsteuergerät, woraufhin die Airbags ausgelöst werden. Von hier ab sendet die Telematikbox über das Mobilfunknetz ein Signal an das Backend, wo sich ein neuer Fall öffnet. Das Auto hat eine fest verbaute Simkarte, um Daten- und Sprachkanal zu übertragen. Aktuell hat der Fahrer zwei Frontendzugriffe. Vor dem Kauf stellt dies das Onlineportal bzw. die mobile App dar. Während der Service-nutzung ist die im Auto verbaute Headunit das zentrale Frontend. Hier kann der Nutzer einen manuellen Vorfall starten oder nimmt auditiv die Sprachverbindung wahr.

Im Backend wurde eine umfangreiche Plattform zur Kommunikation mit dem Fahrzeug, das sogenannte Daimler Vehicle Backend aufgebaut. Es steuert nahezu den gesamten Datenkanal von und aus dem Auto. Parallel hat der Zulieferer *Bosch* eine speziell für eCalls ausgelegte Plattform aufgebaut. Diese empfängt den gesetzlich definierten Minimum Set of Data (MSD), der zur Prozessierung des Notrufes an die Leitstellen weitergereicht wird.

Die Partnerlandschaft bei der vorliegenden Fallstudie besteht aus einem Hardwarepartner für die Telematikbox, dem Mobilfunkanbieter für die Daten- und Sprachverbindung sowie Bosch für einen Teil des Backends und die mehrsprachigen Agenten. Daimler hält die Kundenbeziehung aufrecht und ist daher Business Owner bzw. auch in der Rolle des Generalunternehmers.

5. Fallstudienvergleich und Implikationen

Die skizzierten Fallstudien weisen darauf hin, welche Mehrwerte sich bereits heute aus Big Data Projekten ableiten lassen. Darüber hinaus geben die Fallstudien Hinweise auf die Entwicklungslogik für Big Data Projekte in der Praxis. Auf der Grundlage einer vergleichenden Analyse (Cross-Case Analyse) der Fallstudien werden diese Prinzipien nachfolgend explizit dargestellt.

Darüber hinaus umfasst das folgende Kapitel einen Quervergleich der Ergebnisse aus Theorie, Befragung und Fallstudien. Dies gibt einen fundierten Überblick zum Entwicklungsstand von Big Data aus allen drei im Kontext des Big Data Reports skizzierten Perspektiven.

Abschließend werden die wesentlichen Implikationen für Big Data Lösungen der Zukunft diskutiert. Diese lassen sich auch als Leitlinien bzw. Agenda für die zukünftige Forschung im Themenfeld sowie für die Umsetzung entsprechender Projekte in der Unternehmenspraxis verstehen.

5.1. Cross-Case Analyse

Kapitel 4 des Big Data Reports bietet Einblicke in die Besonderheiten einzelner Fallstudien in Bezug auf die Umsetzung von Big Data Lösungen (Within-Case Analyse). Durch die Anwendung eines gemeinsamen Auswertungsschemas lassen sich die Fallstudien miteinander vergleichen (Cross-Case Analyse).

Dabei sollen Gemeinsamkeiten und Unterschiede identifiziert und auf höherer Abstraktionsebene zusammengefasst werden. Trotz der völlig unterschiedlichen Anwendungsfälle lassen sich auf diese Weise allgemeine Prinzipien für Big Data ableiten.

Fokus auf Mehrwerte

Aus Sicht der beteiligten Partner in Big Data Projekten ist vor allem die Schaffung neuer Mehrwerte relevant. Daher müssen sich Big Data Projekte bereits in der Projektanlage durch Wertschöpfungskonzepte auszeichnen. Durch eine dauerhafte Internetverbindung und Prozessautomation lassen sich selbststeuerbare Dienste für Unternehmen entwickeln, die auch ohne Interaktion der Anwender agieren. Die Fallstudienprojekte schaffen mit der dauerhaften Netzwerkanbindung von Sensordaten und den Möglichkeiten von Akteuren eine Optimierung im Betriebs- bzw. Serviceprozess, in der Risikominimierung oder in der eigenständigen Steuerung von Maschinen in Echtzeit. Ziel ist es in diesem Umfeld, die Kostenseite transparent darzustellen und unnötige Aufwände einzusparen. Eingesparte Betriebskosten führen unmittelbar zu einer Verbesserung des Gesamtergebnisses.

Automatisierung

Bei nahezu allen Fallstudien liegt eine zentrale Leitstelle vor, die per Ferndiagnose und -wartung die Steuerung übernimmt. Das Personal vor Ort kann somit eingespart werden. Ein Eingriff ist nur noch ereignisgesteuert, vorab geplant bzw. durch Data Mining-Algorithmen vorausgesagt. Es zeigt sich jedoch, dass es stets Bedarf an personalisierter Kommunikation und Intervention gibt, die nicht digitalisiert werden kann. Dennoch ist ein deutlicher Trend zur Automatisierung erkennbar und mit Hilfe neuerer Technologien umsetzbar.

Organisation

Die Besonderheit der fallstudien-spezifischen Projektorganisation liegt darin, dass die Teams wesentlich agiler und daher mit mehr Freiräumen arbeiten können. Durch diese Agilität kann ein Zeitvorsprung im Wettbewerb umgesetzt werden. Häufig wird daher auch eine Abspaltung von etablierten Strukturen umgesetzt.

Das Beispiel Daimler weist jedoch auf die offene Frage hin, wie die Zusammenarbeit und Abgrenzung mit etablierten Strukturen aufgesetzt werden soll. In einigen Bereichen treten Schnittstellen bei Befugnissen sowie mehrfach vorhandene Kompetenzen und Kapazitäten auf. Dies schürt Spannungen innerhalb der Gesamtorganisation. Auf Anbieterseite zeigt sich eine Bündelung aller relevanten Arbeitsfelder und Wissensgebiete, wie Mobile Applications, Analytics oder Cloud Computing in einer Organisationseinheit.

Daraus lässt sich ableiten, dass Big Data Projekte kaum von einer Fachabteilung alleine umsetzbar sind. Es bedarf der Zusammenarbeit vielfältiger Kompetenzträger aus IT und Business.

Reifegrad

Ein Großteil der befragten Unternehmen befindet selbst noch im Wandel. Bei den geplanten Maßnahmen ist ein kurzfristig keine Ende absehbar. Durch die Veränderung der Geschäftsmodelle, Serviceorientierung und die Integration von IT- und Businessbezug sind neue Strukturen und Personalprofile erforderlich. Einige Unternehmen sind noch in der Findungsphase und wollen sich über Machbarkeitsstudien mehr Wissen und Orientierung aneignen.

Dabei sind Beratungsleistungen von Anbieterseite durchaus relevant. Dies weist ebenfalls auf eine noch fehlende Kompetenzbasis zu Big Data in vielen Unternehmen hin. Externe Partner können in dieser Hinsicht einen aktiven Beitrag zur Innovationsdynamik leisten. Voraussetzung hierfür sind neue Kooperationsmodelle, die eine gemeinschaftliche Umsetzung von Big Data Lösungen erlauben.

Produkt- und Lösungsorientierung

In den überwiegenden Fällen lässt sich eine hohe Lösungs- bzw. Serviceorientierung ausmachen. Dies führt auch zu einer Neuorientierung auf Anbieterseite. Der Fokus liegt nicht mehr auf der Zulieferung reiner Komponenten, sondern eher auf ganzheitlichen Problemlösungen. Die Aufgabe der Lösungsanbieter liegt darin, bestehende Komponenten zu verbinden und daraus neue Servicesysteme zu kreieren.

Devices

Bei den meisten untersuchten Fallstudien besteht bereits eine dauerhafte Netzwerk- anbindung zur Datenübertragung. Sensoren erzeugen in allen genannten Beispielen die notwendigen Daten, wie Druck, Feuchtigkeit, Schließung/Öffnung. In vielen Fallstudien wurden darüber hinaus Aktoren zur Steuerung der verarbeiteten Informationen implementiert. Zur vollständigen Eigensteuerung der Systeme ist daher nicht nur eine IT-Logik in Form eines Regel- und Prozesswerkes notwendig, sondern auch eine Logik von Sensoren und Aktoren. Erst die perfekte Integration dieser Komponenten kann die manuelle und personalisierte Interaktion im Betrieb zumindest teilweise ersetzen.

IT Infrastruktur

Durch die zentrale Steuerung bzw. die Selbststeuerung der Systeme ist ein lokales Frontend nur zur Wartung notwendig. Alle genannten Projekte führen die Informationen an einer zentralen Stelle zusammen. Dies ermöglicht eine integrierte Datenhaltung und Datenauswertung. Durch die Leitstelle werden die dezentralen Devices angesteuert. Für die Fallstudien werden häufig neue Systeme aufgesetzt, die von der etablierten IT Architektur getrennt sind. Dies ist mit Vorteilen bei der Agilität und Nachteilen bei der Standardisierung verbunden. Auf Endkundenseite gilt das Credo „Mobile First“. Einige Fallstudien weisen eine mobile Applikation als einziges Kundenfrontend auf. Bei fast allen Fallstudien ist eine mobile Visualisierung zumindest integriert.

Die meisten Unternehmen in den skizzierten Fallstudien verwenden eine dreistufige IT-Architektur. Der Integrationslayer dient dazu, Datenquellen, wie die der Partner, die Sensoren, die Telematikeinheiten, die externe Services (z.B. Wetterinformationen), Smartphones oder die etablierte IT-Infrastruktur anzubinden. Im nächsten Schritt werden diese Daten im Verarbeitungslayer prozessiert und analysiert. Dabei lassen sich über ein Regelwerk Handlungsempfehlungen für den Betrieb ableiten. Die Ausgabe erfolgt schließlich im Anzeigenlayer. Dieser hält Dashboards, Reports oder Visualisierungen der Daten bereit.

Grundlage der Architektur ist ein Datenmodell, welches bestehende Dienste anbindet bzw. neue Anwendungen erzeugt. Bei der Integration verschiedener Datenquellen sind temporale Aspekte wesentlich. Dies betrifft v.a. die Definition für Datenabrufe in Echtzeit. Dies hat Auswirkungen auf die Taktung bei der Verarbeitung von Daten bzw. auf die Datenintegrität und Datengenauigkeit. Aus dieser Sicht ist der Begriff „Echtzeit“ für jeden Anwendungsfall gezielt zu definieren.

In Bezug auf das Hosting und die Datenhaltung zeigen sich bei den skizzierten Fallstudien deutliche Unterschiede. Dabei liegen Cluster im Bereich on premise vor, jedoch sind bei den Fallstudien auch hybride Cloudlösungen im Einsatz. Die Mehrheit der skizzierten Unternehmen setzt jedoch zumindest bei Big Data Projekten bereits auf Cloud Computing. Die Datenhaltung erfolgt dabei zum Teil auf eigenen Systemen, zum Teil jedoch auch extern bzw. bei Partnerunternehmen.

Die Rückmeldung der Fallstudien weist darauf hin, dass Anbieter Flexibilität bei der Auswahl der Lösungen zur Datenhaltung zeigen müssen. Ein mandantenfähiges Cloudsystem, wie SaaS oder PaaS, ist jedoch ein wesentlicher Erfolgsfaktor für ein skalierbares Geschäftsmodell.

Partnermodelle

Alle untersuchten Beispiele weisen eine intensive Zusammenarbeit mit einem oder mehreren Partnern auf. Dies reicht von der einfachen Anbindung von Wetterdaten bis hin zu der partnerschaftlichen Entwicklung spezieller Hardwarekomponenten. Kooperationen sind im Big Data Zeitalter unerlässlich.

Aus den Fallstudien lassen sich generische Rollenbilder allgemein ableiten. So liegt häufig ein zentrales Projektmanagement bzw. eine Generalunternehmung vor. In vielen Fällen ist diese bei den Unternehmen selbst verortet, zum Teil erfolgt jedoch auch ein Outsourcing an Wertschöpfungspartner. Außerdem spielen unterschiedliche Partner eine wichtige Rolle, z.B. für Datenhaltung, Analyse, Hardware, Sensorik und Aktorik. Darüber hinaus lassen sich auch Rollen für Beratungsunternehmen, Designagenturen oder Third Party Service Provider (u.a. für Daten oder andere Bestandsservices) definieren. In öffentlich geförderten Projekten ist außerdem häufig eine wissenschaftliche Einrichtung beteiligt. Kritischer Erfolgsfaktor bei der Zusammenarbeit ist die ständige Abstimmung und die Angleichung der Roadmaps zwischen den beteiligten Partnern. Big Data Projekte erfordern in dieser Hinsicht fast immer eine Multi-Partner-Struktur.

5.2. Theorie, Befragung und Fallstudien im Vergleich

Eine integrierte Analyse der einzelnen Ergebnisse der vorliegenden Studie aus Theorie, empirischer Befragung und Fallstudien gibt einen konsistenten Überblick zu Status quo und strategischen Perspektiven von Big Data.

Dabei lohnt sich zunächst ein Rückblick auf die Theorie. In der Forschung werden vorwiegend allgemeine Konzepte und Referenzmodelle zur Strukturierung des Themengebiets eingesetzt. Dabei verfügen die bis dato vorliegenden Modelle über unterschiedliche Schwerpunkte. Ein ausgereiftes Modell für die interne Perspektive auf Big Data bietet das Referenzmodell des Branchenverbandes bitkom. Dabei lassen sich Ansatzpunkte für Big Data in den Bereichen Datenhaltung, Datenzugriff, Datenverarbeitung, Visualisierung, Datenintegration, Governance und Sicherheit definieren.

Die empirische Befragung greift das Thema aus der Perspektive von IT-Executives auf. In Bezug auf das bitkom-Modell fokussiert die Praxis in der IT vorwiegend auf den Bereich der Datenhaltung und -integration. Im Grundsatz werden konservative Modelle der Datenhaltung präferiert. Die Integration unterschiedlicher Datenquellen und IT-Systeme bereitet hier besondere Probleme. Allgemein ist der Reifegrad der Big Data Strategien daher eher als schwach einzustufen. Neuere Technologien für das Management und die Verarbeitung großer Datenmengen sind kaum im Ansatz. Aus dieser Sicht bietet das Referenzmodell der bitkom etliche Perspektiven für die weitere Entwicklung von Big Data in der Praxis.

Kritisch ist jedoch der fehlende Businessbezug dieser Perspektiven zu bewerten. Dies gilt sowohl für das Referenzmodell der bitkom, als auch für die Ergebnisse der empirischen Befragung. In der Forschung liegen aktuell unterschiedliche Modelle vor, die im Kern das Servicemodell sowie den Mehrwert für den Anwender in den Mittelpunkt stellen. Eine derartige Businessperspektive zeigt sich beispielsweise beim IoT Infrastrukturmodell der Harvard Business School oder beim Ansatz hybrider Geschäftsmodelle der Universität St.Gallen.

Bei einer Analyse der skizzierten Fallstudien fällt ebenfalls auf, dass diese ausschließlich aus Sicht der Anwendung und für den Mehrwert auf Userseite aufgebaut sind. Diese Denken in Services und Mehrwerten führt zu neuen Anforderungen an die IT. Zukünftig müssen IT Services das Businesskonzept eines Unternehmens deutlich schneller adaptieren und die dafür erforderlichen IT Services bereitstellen. Die etablierte IT Enterprise Architektur wird diesen Anforderungen nach Agilität nicht immer nachkommen können.

Daher ist in vielen Fällen der Ansatz einer Two-Speed Architektur erforderlich, der agile Businessanwendungen mit etablierten Kernsystemen verbindet. Parallel ist die Integration dieser beiden Systemansätze weiter voranzutreiben. Cloud Computing wird in dieser Diskussion eine wichtige Rolle spielen, ist gegenwärtig als Lösungsansatz in der Praxis jedoch noch nicht akzeptiert.

Eine Erweiterung der Kompetenzen bei der Umsetzung von Big Data Strategien wird den Fokus stärker auf unstrukturierte und unternehmensexterne Daten lenken. Aktuell beziehen sich noch viele Big Data Initiativen auf unternehmensinterne Daten. Diese weisen häufig eine ausgeprägte Struktur auf. Darüber hinaus sind diese Datentypen beherrschbar und im Datenbestand relativ stabil. Für eine volle Ausschöpfung der mit Big Data verbundenen Mehrwerte ist jedoch eine stärkere Hinwendung zu unternehmensexternen Daten erforderlich. Dies zeigt sich auch bei der empirischen Analyse in der Fokussierung auf kundenorientierte Businessziele.

Die skizzierten Fallstudien weisen darüber hinaus darauf hin, dass sich die Qualität der Zusammenarbeit zwischen IT, Fachbereichen und externen Parteien verändert. Zukünftig stehen Serviceökosysteme mit einer deutlich stärkeren Integration der Partner im Fokus. Dies verändert auch die Rolle der Anbieter. Aus Lieferanten für einzelne Komponenten werden System- und Lösungspartner. Dabei kann sich die Zusammenarbeit in Einzelfällen sogar auf den Betrieb gemeinsamer Geschäftsmodelle beziehen.

Schließlich zeigen die Fallstudien auf, dass nicht zwingend neue Technologien für die Umsetzung von Big Data Ansätzen erforderlich sind. Die Anforderungen an die Technologie ergeben sich aus dem direkten Anwendungsfall. Dieser definiert auch den Anspruch an Echtzeit und die Komponenten der Datenauswertung. Zum Teil lassen sich daher auch kostengünstig und mit etablierten Strukturen ambitionierte Big Data Projekte realisieren.

5.3. Implikationen für Big Data Lösungen

Auf der Grundlage der skizzierten Resultate aus Theorie und empirischer Forschung lassen sich unterschiedliche Implikationen für Big Data Lösungen skizzieren. Diese bauen auf den folgenden sechs Prinzipien für die Weiterentwicklung von Big Data im Unternehmenskontext auf.

Prinzip 1:

Serviceorientierung

Big Data ist kein Selbstzweck. Das Datenmanagement muss sich an der Verbindung von IT- und Businessdiensten ausrichten. Daher ist in der IT eine höhere Serviceorientierung einzufordern, die sich bereits bei der Entwicklung entsprechender Referenzmodelle für die Datenhaltung und das Datenmanagement zeigt. In diesem Sinne sind Big Data Lösungen aus Sicht der beteiligten Stakeholder zu konfigurieren. Der aktuelle Fokus in der Unternehmenspraxis ist in dieser Hinsicht zu stark nach innen gerichtet. Erst auf Basis der Anforderungen aus dezidierten Use Cases für Big Data lassen sich echte Konsequenzen für die IT ableiten. In der Zwischenzeit muss die IT an der eigenen Kompetenzbasis arbeiten und eine möglichst agile Infrastruktur zu Verfügung stellen, die unterschiedliche Anwendungsfälle optimal reflektiert.

Prinzip 2:

Fokus auf Mehrwerte

Mit der skizzierten Serviceorientierung ist auch eine klare Ausrichtung auf den Mehrwert von Big Data verbunden. Dieser leitet sich immer aus der Realisierung relevanter Unternehmensziele ab.

Daher sind die Auswirkungen von Big Data auf das aktuelle Geschäftsmodell sowie mögliche Opportunitäten für neue Geschäftsmodelle sorgfältig zu prüfen. Dies ist in erster Linie ein Kreativprozess. In diesem Sinne sind fundierte Ansätze im Innovationsmanagement gefragt, die tragfähige und passende Ideen für Big Data stimulieren. Erst auf Basis der Ideenfindung sind konkrete Datenquellen, Geschäfts- und IT-Prozesse sowie die daran angebundenen Fragestellungen für das Datenmanagement eindeutig zu klären. Dies bezieht sich auch für die Auswahl der an der Umsetzung von Big Data Projekten beteiligten Partnern.

Prinzip 3:

Two Speed Architecture

Für die IT werden sich eigene Ansätze im Big Data Segment an einer Two Speed Architecture orientieren. Damit ist die agile Umsetzung wertschöpfender Big Data Strategien bei gleichzeitiger Weiterentwicklung etablierter Bestandssysteme gemeint. Die bislang in etlichen Unternehmen umgesetzten IT-Architekturen sind nicht für die Gestaltung von Big Data Lösungen aufgebaut. Zudem befindet sich die IT selbst in einer Zeit massiver Veränderungen. Diese lassen sich kaum im Sinne eines konsistenten Gesamtansatzes aus einem Guss lösen. Daher stellt sich die Frage, wie sich Bestandssysteme weiterentwickeln lassen – ohne an Agilität in wettbewerbsrelevanten Bereichen zu verlieren. Erste Ansätze für die Umsetzung sind in den vorgestellten Fallbeispielen vorhanden.

Prinzip 4: **Service Ökosysteme**

Unternehmen werden Lösungen für Big Data selten im Alleingang entwickeln. Erfolgreiche Fallbeispiele weisen darauf hin, dass immer mehrere Parteien an einer Gesamtlösung beteiligt sind. Häufig ist dies mit der Bildung von Service Ökosystemen verbunden. Die relevanten Teilleistungen agieren dann vernetzt, externe Anbieter werden zu Wertschöpfungspartnern und etablierte Rollen in der Umsetzung entsprechender Projekte stehen zur Diskussion. Dies trifft bereits auf die Entwicklung von Big Data Strategien zu. Unternehmen sollten sich daher damit befassen, mit welchen Partnern ein strategischer Dialog über Big Data Konzepte der Zukunft Sinn macht.

Prinzip 5: **Integration von Business und IT**

Nicht erst seit Big Data wird in der Diskussion rund um die Weiterentwicklung der IT über eine stärkere Integration mit einzelnen Fachabteilungen diskutiert. Spätestens bei Big Data ist diese Zusammenarbeit jedoch obligatorisch. In der Regel werden sogar mehrere Fachbereiche gemeinschaftlich durch die Umsetzung einer Big Data Lösung tangiert. In einigen Fällen kann dies auch das Unternehmen als Ganzes betreffen. Daher stellt sich die Frage, wie eine tiefere Integration zwischen Business und IT gestaltbar ist. Dies ist in letzter Konsequenz v.a. eine Kulturfrage. Die Kultur der Zusammenarbeit zwischen unterschiedlichen Funktionen ist für den Erfolg von Big Data Projekten ein entscheidender Erfolgsfaktor. In dieser Hinsicht müssen alle beteiligten Partner die erforderlichen Kooperationskompetenzen nachhaltig aufbauen.

Prinzip 6: **Technology follows Strategy**

Big Data wird häufig mit technologischen Fragestellungen verbunden. Hadoop, MapReduce, NoSQL Datenbanken oder BI in der Cloud? Im Endeffekt muss die IT die eigenen Kompetenzen in diesen Bereichen weiter ausbauen. Häufig lassen sich Big Data Projekte jedoch auch ohne Investitionen in neue Technologien umsetzen. Insofern muss sich die IT an der fachlichen Big Data Strategie orientieren. Etliche Referenzmodelle orientieren sich zu stark an theoretischen Referenzmodellen und verfügen über wenig Anbindung an fachliche Fragestellungen. Langfristig muss sich die IT jedoch immer an relevanten Businessmodellen orientieren.

Prinzip 7: **Sicherheit in der Cloud**

Unternehmen werden sich bereits in naher Zukunft stärker mit dem Thema Cloud befassen. Cloudstrategien bieten bereits heute fundierte Lösungen für etliche Fragestellungen der Datenhaltung. Darüber hinaus lassen sich über Cloud Computing deutlich agilere Modelle der IT-Architektur umsetzen. Schließlich werden künftig vermehrt externe Daten aus der Cloud für Big Data Projekte genutzt. Daher müssen sich Unternehmen deutlich stärker mit entsprechenden Konzepten befassen. Dies bezieht sich auf grundlegende Anwendungsmodelle, jedoch zunehmend auch auf Frage der Datensicherheit. An dieser Stelle sind auch die IT-Industrie selbst und die Forschung gefragt, um die Vorteile von Cloud Computing mit deutlich ausgereifteren Konzepten der Datensicherheit zu kombinieren.

Prinzip 8:**Investition in die eigene Kompetenz**

Schließlich verlangt Big Data eine kritische Auseinandersetzung mit der eigenen Kompetenzbasis. Aus Sicht der empirischen Befragung besteht in dieser Hinsicht der größte Handlungsbedarf. Unternehmen müssen sich eigene Kompetenzen für das Management großer Datenmengen aneignen und darüber hinaus die Wissensressourcen der externen Partner nutzen. Das Thema ist besonders kritisch, denn ohne ausreichende eigene Expertise lassen sich etliche der bereits angeführten Prinzipien kaum umsetzen. Dies bezieht sich in erster Linie auf das Thema IT und die mit Big Data verbundenen technologischen Innovationen. Darüber hinaus dehnt sich die Bandbreite der relevanten Kompetenzbereiche jedoch auch auf das Management von Wertschöpfungssystemen, die Gestaltung funktionaler Kooperationen und die Entwicklung relevanter Geschäftsmodelle aus. Die gute Nachricht ist, dass das Handlungsfeld der IT damit rapide an Bedeutung gewinnt. IT erzeugt bei Big Data einen messbaren ökonomischen Mehrwert. Diesen Status muss sich der IT-Bereich durch eigene Expertise und ein gutes Partnermanagement erhalten.

6. Fazit und Ausblick

Der Big Data Report analysiert den Entwicklungsstand in Bezug auf den Umgang von Unternehmen mit großen Datenmengen aus verschiedenen Perspektiven. Dabei ist klar, dass das Thema an Bedeutung gewinnt. Daten erzeugen Wettbewerbsvorteile und führen zu neuen Modellen der Wertschöpfung.

Gleichzeitig ist die Thematik komplex. Big Data erfordert ein Umdenken in mehreren Bereichen. Dies gilt zum einen für die IT selbst. Die etablierte IT-Architektur steht in Frage und neue Kompetenzen sind gefragt. Mehr denn je hat sich das Handlungsumfeld für CIOs dynamisiert. Dies ist wie immer mit Chancen und Risiken verbunden.

Neben der IT verändert sich jedoch auch der fachliche Rahmen. Dabei kann Big Data durchaus zu komplett neuen Geschäftsmodellen führen. Zumindest ist eine signifikante Unterstützung für etablierte Geschäftsmodelle greifbar. Damit sind durchaus neue Formen der Zusammenarbeit und Wertschöpfung verbunden.

Auf Basis der skizzierten Rahmenbedingungen kann ein Ausblick auf die weitere Entwicklung erfolgen. Dabei ist davon auszugehen, dass sich die Weiterentwicklung von Big Data in deutschen Unternehmen mit unterschiedlichen Geschwindigkeiten vollzieht. Die dargestellten Anforderungen werden von einzelnen Unternehmen unterschiedlich gut gelöst. Big Data ist daher immer auch mit möglichen Wettbewerbsvorteilen verbunden.

Auf Grund der Komplexität und Individualität von Big Data Lösungen lassen sich funktionale Lösungen nur sehr eingeschränkt kopieren. Für innovative Unternehmen ist dies eine gute Nachricht. Damit wird sich die Spannweite zwischen Unternehmen mit starker und schwacher Big Data Expertise deutlich und dynamisch erweitern.

Für die weitere Forschung bietet der Themenbereich weitere Fragestellungen. Diese reichen von der Entwicklung funktionaler Modelle einer Two Speed Architecture bis hin zur Umsetzung neuer Sicherheitskonzepte im Cloud Computing. Zum Teil werden sich derartige Konzepte in Kooperationen mit der Praxis entwickeln.

Der Big Data Report kann daher nur als Zwischenbericht einer langfristigen Entwicklung aufgefasst werden. Weitere Referenzmodelle und Fallbeispiele werden folgen und den Dialog rund um fundierte Big Data Lösungen der Zukunft weiter stimulieren.

7. Literatur

Bassi et al. (2013): Enabling Things to Talk:
Designing IoT Solutions with the IoT Architectural Reference Model. Springer

Berkeley (2014): What is Big Data? Zugriff am 22.05.2015
unter <http://www.datascience.berkeley.edu/what-is-big-data>

bitkom (2014): Big-Data-Technologien – Wissen für Entscheider. Zugriff am 01.05.2015
unter https://www.bitkom.org/Bitkom/Publikationen/Publikation_775.html

Denning, P.J. (1990): Saving all the bits.
In: The Science of Computing for American Scientist, Vol. 78, Nr.5

Fleisch, E., Weinberger, M. & Wortmann, F. (2014): Geschäftsmodelle im Internet der Dinge.
Zugriff am 08.06.2015 unter http://www.iot-lab.ch/?page_id=10543

Forbes (2014): 12 Big Data Definitions. Zugriff am 22.05.2015
unter <http://www.forbes.com/sites/gilpress/2014/09/03/12-big-data-definitions-whats-yours>

Gartner (2013): Gartner Newsroom. Zugriff am 28.04.2015
unter <http://www.gartner.com/newsroom/id/2636073>

Ke, R., Guangyu, H., Yong, L., Yongjiang, S. & Liang, G. (2015): Understanding business ecosystem using a 6C framework
in Internet-of-Things-based sectors. In: International Journal of Production Economics. Vol. 159(1), 41-55

Laney (2001), 3D Data Management: Controlling Data Volume, Velocity, and Variety. Zugriff am 15.05.2015
unter <http://blogs.gartner.com/doug-laney/files/2012/01/ad949-3D-Data-Management-Controlling-Data-Volume-Velocity-and-Variety.pdf>

Morris & Truskovski (2003), The Evolution of Storage Systems. IBM Systems Journal, Vol. 42/2, 205-217

Opentracker (2015): Definitions of Big Data. Zugriff am 22.05.2015
unter <http://www.opentracker.net/article/definitions-big-data>

Porter, M.E. & Heppelmann, J.E. (2014): Wie smarte Produkte den Wettbewerb verändern.
In: Harvard Business Manager, 12/2014, 34-60

MIT Technology Review: The Big Data Conundrum: How to Define it? Zugriff am 01.05.2015
unter <http://www.technologyreview.com/view/519851/the-big-data-conundrum-how-to-define-it>

Tekinger, F. & Keane, J.A. (2013): Big Data Framework.
In: IEEE International Conference on Systems, Man, and Cybernetics (SMC), 1494-1499

Yin, R.K. (2013): Case Study Research: Design and Methods. Sage Publications