

**Fakultät Maschinenbau  
Fachgebiet IT in Produktion und Logistik**

# **Masterarbeit**

## **Ein Vergleich agiler und klassischer Vorgehensmodelle des IT-Projektmanagements bei der Einführung von Big-Data-Management-Systemen**

im Studiengang Wirtschaftsingenieurwesen  
an der Technischen Universität Dortmund  
ausgegeben am: 04.11.2014  
eingereicht am: 17.04.2015

vorgelegt von:	Jan-Niklas Grafe
Matrikelnummer:	165400
Betreuer:	Univ.-Prof. Dr.-Ing. Markus Rabe Dipl.-Inf. Anne Antonia Scheidler

## Inhaltsverzeichnis

<b>1</b>	<b>Einleitung</b> .....	<b>1</b>
<b>2</b>	<b>Big Data</b> .....	<b>6</b>
2.1	Hintergrund und Herkunft .....	6
2.2	Definition und Eigenschaften von Big Data.....	7
2.2.1	Deutungsversuche in der Literatur.....	7
2.2.2	Beschaffenheit von Big Data .....	9
2.3	Big-Data-Management-Systeme .....	12
2.3.1	Technologien für Big Data .....	12
2.3.2	Implementierung des Big-Data-Stapels .....	16
2.3.3	Datenarchitekturen für Big-Data-Management-Systeme .....	18
2.4	Ziele und Strategien von Big Data .....	20
2.4.1	Kosteneinsparung durch Big Data .....	20
2.4.2	Zeiteinsparung mit Hilfe von Big Data .....	22
2.4.3	Entscheidungsunterstützung und Prognosen durch Big Data .....	23
2.4.4	Entwicklung neuer Produkte auf Basis von Big Data .....	24
2.5	Herausforderungen im Umgang mit Big Data .....	25
2.5.1	Data Governance und Datensicherheit.....	25
2.5.2	Personelle Ressourcen und organisatorische Einbindung .....	28
2.5.3	Erfolgsmessung bei Big-Data-Projekten .....	30
<b>3</b>	<b>Vorgehensmodelle im IT-Projektmanagement</b> .....	<b>32</b>
3.1	Definition und Herkunft von Vorgehensmodellen .....	32
3.2	Herausforderungen und Erfolgsfaktoren im IT-Projektmanagement .....	34
3.3	Klassische Vorgehensmodelle .....	36
3.3.1	Charakteristika .....	36
3.3.2	Beispiel Wasserfallmodell .....	37
3.4	Agile Vorgehensmodelle .....	39
3.4.1	Charakteristika .....	39

---

3.4.2	Beispiel Scrum .....	41
3.5	Vergleich von Vorgehensmodellen .....	44
3.5.1	Typische Vergleichskriterien.....	44
3.5.2	Nutzwertanalyse.....	47
<b>4</b>	<b>Einführung von Big-Data-Management-Systemen.....</b>	<b>51</b>
4.1	Einführungsprozesse verschiedener Big-Data-Strategien.....	51
4.2	Vergleich agiler und klassischer Vorgehensmodelle .....	56
4.2.1	Kriterienkatalog und Ausprägungen .....	56
4.2.2	Kriterien-Erfüllungsgrade agiler und klassischer Vorgehensmodelle .....	64
4.2.3	Kriterien-Bedeutsamkeiten für einzelne Big-Data-Strategien .....	70
4.2.4	Auswertung im Rahmen von Nutzwertanalysen .....	76
4.3	Handlungsempfehlungen.....	84
<b>5</b>	<b>Beispielhafte Anwendung im Rahmen einer Fallstudie .....</b>	<b>89</b>
5.1	Rahmenbedingungen und Ausgangssituation .....	89
5.2	Projektdefinition.....	91
5.2.1	Projektumfang und Vorgehensweise .....	91
5.2.2	Produktspezifikationen .....	92
5.3	Projektdurchführung .....	94
5.3.1	Projektverlauf mit klassischen Vorgehensmodellen.....	94
5.3.2	Projektverlauf mit agilen Vorgehensmodellen.....	96
5.4	Projekterfolg und Ergebnisbewertung.....	97
<b>6</b>	<b>Schlussbetrachtung.....</b>	<b>100</b>
6.1	Zusammenfassung.....	100
6.2	Fazit .....	105
6.3	Ausblick.....	106

## Abkürzungsverzeichnis

Abb.	Abbildung
AG	Aktiengesellschaft
BCG	Boston Consulting Group
BITKOM	Bundesverband Informationswirtschaft, Telekommunikation und neue Medien
bzgl.	bezüglich
bzw.	beziehungsweise
ca.	circa
CRM	Customer Relationship Management
d.h.	das heißt
EDW	Enterprise Data Warehouse
ELT	Extrahieren, Laden, Transformieren
ERP	Enterprise Resource Planning
ETL	Extrahieren, Transformieren, Laden
et al.	et alii (lat. und andere)
Gl.	Gleichung
GmbH	Gesellschaft mit beschränkter Haftung
HANA	High-Performance Analytic Appliance (Produkt der SAP SE)
HD Video	High Definition Video
HP	Hewlett Packard
IBM	International Business Machines Corporation
IDC	International Data Corporation
IAIS	Fraunhofer-Institut für Intelligente Analyse- und Informationssysteme
IT	Informationstechnik
MGI	McKinsey Global Institute
MPP	Massively Parallel Processing
NASA	National Aeronautics and Space Administration
NESSI	Networked European Software and Service Initiative

noSQL	Not Only Structured Query Language
p.	page (engl. Seite)
pp.	pages (engl. Seiten)
RDBMS	Relational Database Management System (Relationale Datenbank)
ROI	Return on Investment
SCM	Supply Chain Management
s.u.	siehe unten
Tab.	Tabelle
vgl.	vergleiche
XML	Extensible Markup Language
z.B.	zum Beispiel

## Abbildungsverzeichnis

Abbildung 1: Emerging Technologies Hype Cycle 2013 .....	1
Abbildung 2: Vorgehensweise der vorliegenden Ausarbeitung .....	4
Abbildung 3: Big-Data-Evolution .....	12
Abbildung 4: Der Big-Data-Stapel.....	16
Abbildung 5: Big-Data-Einführungsprozess .....	18
Abbildung 6: Zusammenspiel von Hadoop mit bestehenden Lösungen.....	19
Abbildung 7: Modelle zur organisatorischen Einbindung eines Analytikbereichs.....	29
Abbildung 8: Bekanntheitsgrad und Anwendung von Vorgehensmodellen .....	33
Abbildung 9: Zentrale Phasen der Softwareentwicklung .....	36
Abbildung 10: Wasserfallmodell nach Boehm.....	38
Abbildung 11: Scrum im Überblick.....	42
Abbildung 12: Magisches Projektdreieck .....	44
Abbildung 13: Parameter klassischer und agiler Vorgehensmodelle.....	46
Abbildung 14: Erfüllungsgrad der Teilnutzwerte - Kosteneinsparung .....	78
Abbildung 15: Erfüllungsgrad der Teilnutzwerte - Zeiteinsparung .....	80
Abbildung 16: Erfüllungsgrad der Teilnutzwerte - Entscheidungsunterstützung .....	82
Abbildung 17: Erfüllungsgrad der Teilnutzwerte - Neue Produkte.....	83
Abbildung 18: Datenarchitektur der Bank zu Projektbeginn .....	90
Abbildung 19: Vorgehensweise zur Projektumsetzung .....	92
Abbildung 20: Ziel-Datenarchitektur nach Projektabschluss .....	93
Abbildung 21: Projektdurchführung mit dem Wasserfallmodell .....	95
Abbildung 22: Scrum Burndown Chart für Kosteneinsparungsprojekt.....	97

## Tabellenverzeichnis

Tabelle 1: Big-Data-Technologien im Überblick .....	13
Tabelle 2: Typische Probleme und Herausforderungen bei IT-Projekten .....	35
Tabelle 3: Vier Leitsätze des agilen Manifests .....	40
Tabelle 4: Einflusskriterien erfolgreicher Projektarbeit nach Holzbaur .....	45
Tabelle 5: Fünf kritische Erfolgsfaktoren für agile und klassische Vorgehensmodelle.....	47
Tabelle 6: Paarvergleich zur Ermittlung der Kriteriengewichtung .....	49
Tabelle 7: Alternativenbewertung zur Ermittlung des Nutzwerts .....	49
Tabelle 8: Kriterienkatalog mit Kriterienausprägungen.....	57
Tabelle 9: Kriterien-Erfüllungsgrade agiler und klassischer Vorgehensmodelle .....	65
Tabelle 10: Kriterien-Bedeutsamkeiten für einzelne Big-Data-Strategien.....	71
Tabelle 11: Nutzwertanalyse - Kosteneinsparung.....	77
Tabelle 12: Nutzwertanalyse - Zeiteinsparung.....	79
Tabelle 13: Nutzwertanalyse - Entscheidungsunterstützung.....	81
Tabelle 14: Nutzwertanalyse - Neue Produkte.....	83
Tabelle 15: Zusammenfassung der Nutzwertanalysen .....	84

**Vorwort**

An dieser Stelle möchte ich mich bei den Personen bedanken, die mich während der Erstellung meiner Masterarbeit auf unterschiedlichste Weise unterstützt haben.

Besonderer Dank gilt dabei Herrn Prof. Dr.-Ing. Markus Rabe und Frau Dipl.-Inf. Anne Antonia Scheidler seitens der Technischen Universität Dortmund. Darüber hinaus bedanke ich mich auf Seiten der KPMG AG Wirtschaftsprüfungsgesellschaft beim gesamten Team von Herrn Marc Ennemann und dort insbesondere bei Herrn Michael Brenner. Weiterhin danke ich allen beteiligten Kolleginnen und Kollegen sowie meiner Familie und Freunden.



## 1 Einleitung

Wer seit Beginn der 2010er Jahre einen Blick in einschlägige Bücher, Magazine oder Internet-Portale der IT-Fachwelt wirft, wird unweigerlich mit dem Begriff „Big Data“ in Berührung kommen (vgl. [Mcb12]). Es handelt sich dabei um einen Sammelbegriff für umfangreiche Datenbestände, die zumeist durch Verfügbarmachung, Zusammenführung sowie Auswertung aggregiert, analysiert und visualisiert werden. Eine exakte Definition, welche eine klare Abgrenzung des Begriffs erlaubt, gibt es aktuell nicht. (vgl. [Wei13] pp.131-148). Big Data ist derzeit ein dominierendes und viel diskutiertes Thema. Das IT-Marktforschungsunternehmen Gartner belegt diese Behauptung mit ihrem jährlich erscheinenden „Hype Cycle for Emerging Technologies“ aus 2013 (Abb. 1). Darin befindet sich das Thema Big Data in einer der fünf Reifegrad-Phasen aufkommender Technologien, welche erste Erfolge, aber auch überhöhte Erwartungen widerspiegelt. Es ist typisch für Technologien, die sich in dieser Phase befinden, dass sehr viel und heterogen über sie kommuniziert wird. Betrachtet man den Hype Cycle für 2014, so hat Big Data den Hochpunkt der Kurve überschritten und ist auf dem Weg zu einer produktiv nutzbaren Technologie (vgl. [Gar13]).

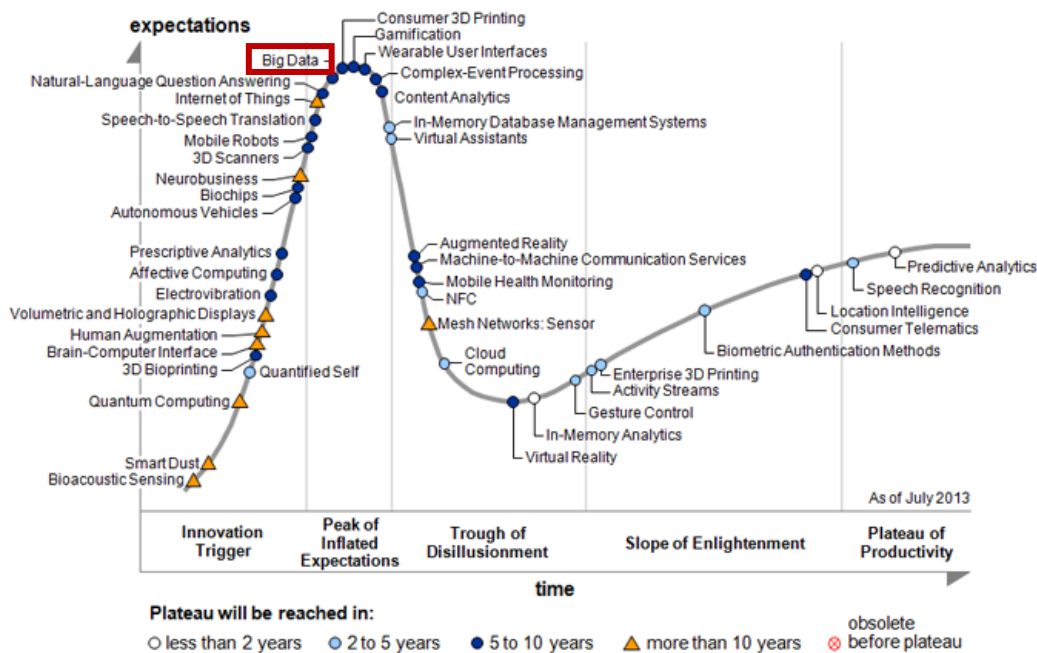


Abbildung 1: Emerging Technologies Hype Cycle 2013 (vgl. [Gar13])

Diesen Trend bestätigt auch der Bundesverband Informationswirtschaft, Telekommunikation und neue Medien (BITKOM) im Rahmen einer repräsentativen Befragung von Unternehmen in Deutschland zu den Potentialen und dem Einsatz von Big Data. Darin geben 31 Prozent der Unternehmen an, konkrete Pläne für den Einsatz von Big-Data-Lösungen zu haben und weitere 28 Prozent diskutieren ihre Nutzung (vgl. [Bit14a]). Eine derartige Befragung sowie der Emerging Technologies Hype Cycle erklären jedoch nur die Relevanz des Themas. Der

Grund, weshalb die Fachwelt sich mit Big Data beschäftigt, liegt vielmehr im rasanten Wachstum des weltweiten Datenaufkommens und der gleichzeitig exponentiell steigenden Rechenleistung von Computern. Letztere wird durch das sogenannte Moore'sche Gesetz beschrieben, welches besagt, dass sich die Komplexität integrierter Schaltkreise etwa alle zwei Jahre verdoppelt (vgl. [Mor65]). Die International Data Corporation (IDC) geht 2014 davon aus, dass sich die weltweite Datenmenge bis 2020 alle zwei Jahre auf 44 Zettabytes, was 44 Trillionen Gigabytes entspricht, verdoppelt. Laut IDC ist dieses exponentielle Wachstum vor allem durch das „Internet der Dinge“, welches die Entwicklung hin zu eingebetteten Systemen und Sensoren in physischen Gegenständen und Infrastruktur beschreibt, sowie Cloud Computing, Mobile Apps und Social Media begründet. Darüber hinaus spielen Meta- und Transaktionsdaten eine herausragende Rolle für diesen rapiden Anstieg (vgl. [Idc14]). Parallel entwickeln sich Datenspeichersysteme, Schaltkreise und Prozessoren seit Jahrzehnten mit sinkenden Kosten und steigender Leistung weiter, sodass auch große Datenmengen günstig und effizient verarbeitet und gespeichert werden können (vgl. [Far06]).

Vor diesem Hintergrund stellt sich für Organisationen die Frage, inwieweit diese Daten sinnvoll und gewinnbringend für Geschäftszwecke genutzt werden können. Das McKinsey Global Institute, welches das forschende Organ der Unternehmensberatung McKinsey repräsentiert, behauptete schon 2011, dass Big Data die „nächste Grenze für Innovation, Wettbewerb und Produktivität“ darstelle (vgl. [MCB+11]). Daraus lässt sich schließen, dass sich Organisationen heute und zukünftig unweigerlich mit Big Data beschäftigen müssen, um weiterhin innovativ, wettbewerbsfähig und produktiv zu sein.

Der Global Technology Adoption Index von Dell ist eine branchenübergreifende Studie zum Thema Cloud Computing, IT-Sicherheit und Big Data, für die über 2000 IT-Entscheidern aus mittelgroßen Organisationen befragt werden. Sie zeigt auf, dass eine effektive Nutzung von Big Data derzeit vor allem an Umsetzungsproblemen und zu hohen Kosten scheitert. So sagen 61 Prozent der Befragten, dass große Datenmengen bereits vorliegen, jedoch verstehen nur 39 Prozent, wie man daraus nachhaltigen Nutzen zieht (vgl. [Del14]). Diese Studie veranschaulicht, dass effiziente Lösungen zur Einführung von Big-Data-Management-Systemen notwendig sind, um die Hürde zur operativen Nutzung von Big Data zu überschreiten. Zu Beginn kommt es dabei vor allem auf ein effektives und effizientes Projektmanagement an.

Die Einführung eines Big-Data-Management-Systems ist in der Regel mit einem bedeutenden organisatorischen Aufwand verbunden. Dieser Aufwand resultiert zum einen aus der Novität der Aufgabe Big Data sinnvoll zu analysieren. Zum anderen müssen die notwendigen Systeme und Technologien an der richtigen Stelle verknüpft werden. Um diesen Herausforderungen gerecht zu werden, wird der Einführungsprozess eines solchen Systems

zumeist mit Hilfe von agilen oder klassischen IT-Projektmanagement-Vorgehensmodellen strukturiert und im Rahmen eines einmaligen, zeitlich begrenzten Projekts durchgeführt.

Das Gesamtziel dieser Arbeit liegt darin, eine Aussage darüber zu treffen, inwiefern die Auswahl eines Vorgehensmodell-Typen, abhängig von der in der Organisation verfolgten Big-Data-Strategie, Einfluss auf den Projekterfolg hat. Die Dimensionen des Erfolgs reichen dabei von der Kosteneffizienz über die Termintreue bis hin zur Kundenzufriedenheit. Ergebnis sind fundierte Handlungsempfehlungen für die jeweiligen Organisationen bzw. die dahinter stehenden Strategien, die mit Big Data verfolgt werden.

Um diese Handlungsempfehlungen zu untermauern und eine Basis für die Vergleichsanalyse zu schaffen, ist es ein erstes Teilziel dieser Ausarbeitung, den derzeitigen Stand der Technik sowie die damit verbundenen Herausforderungen ausführlich darzustellen. Zunächst wird der Themenkomplex Big Data mit Hilfe von aktuellen Forschungsergebnissen und bewährten Vorgehensweisen aus der Praxis aufgearbeitet. Anschließend soll mit Blick auf Big-Data-Projekte eine umfassende Charakterisierung klassischer und agiler Vorgehensmodelle im IT-Projektmanagement gegeben werden, sodass essentielles Grundlagenwissen für die anschließenden Abschnitte vorliegt.

Das zweite Teilziel der vorliegenden Arbeit ist der quantitative Vergleich von IT-Projektmanagement-Vorgehensmodellen in Bezug auf Big-Data-Strategien. Dazu sollen klassische und agile Vorgehensmodelle mit Hilfe eines Kriterienkatalogs auf ihre Effektivität und Effizienz bei der Einführung von Big-Data-Management-Systemen hin analysiert werden. Die Auswertung dieser Analyse soll in Form von strategiebezogenen Handlungsempfehlungen erfolgen. Um einen ersten Nachweis über die praktische Anwendbarkeit zu führen, sollen die erarbeiteten Ergebnisse im Rahmen einer Fallstudie validiert werden. Darin ist es das Ziel, dem Leser ein für die jeweilige Big-Data-Strategie hinreichendes Vorgehensmodell und den damit verbundenen Projektverlauf sowie den zu erwartenden Projekterfolg aufzuzeigen.

Die methodische Vorgehensweise gliedert sich in zwei Teile: Im ersten Teil geht es darum, aktuelle Informationen zu den Themenbereichen Big Data Management und IT-Projektmanagement zusammenzutragen. Darauf aufbauend wird im zweiten Teil eine Nutzwertanalyse verschiedener Big-Data-Strategien und Vorgehensmodell-Typen durchgeführt. Die Ergebnisse der Nutzwertanalyse werden anschließend im Rahmen einer Fallstudie beispielhaft angewandt. Die prozessuale Darstellung in Abbildung 2 veranschaulicht die Vorgehensweise visuell.

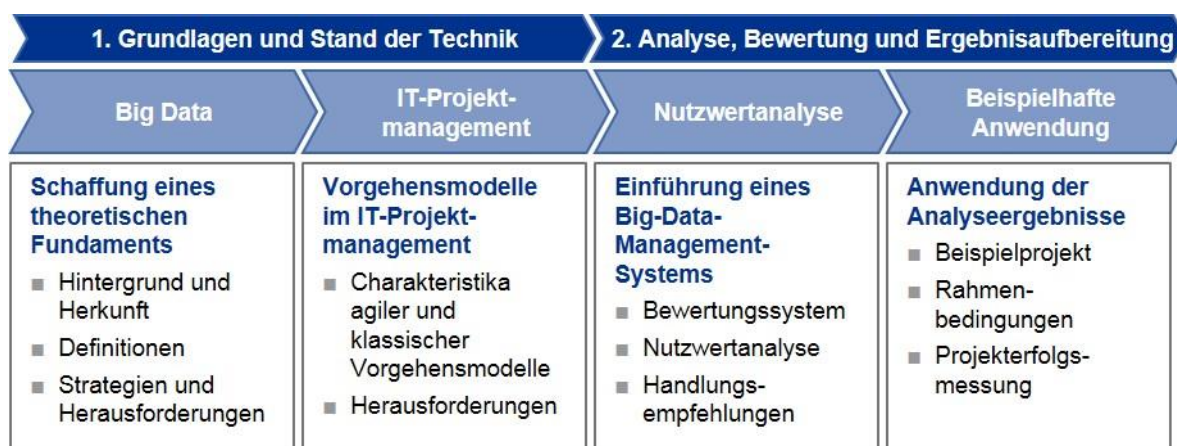


Abbildung 2: Vorgehensweise der vorliegenden Ausarbeitung

Damit ein grundlegendes Verständnis für die Motivation dieser Arbeit hergestellt wird, bedarf es im Themenbereich Big Data zunächst einer Erläuterung des Hintergrunds sowie der Herkunft der Begrifflichkeit. Im daran anschließenden Abschnitt wird Big Data mit Eigenschaften belegt, die eine Klassifizierung verschiedener Big-Data-Strategien erlauben. Da solch große und unterschiedlich strukturierte Datenmengen nutzbar gemacht werden sollen, werden sogenannte Big-Data-Management-Systeme benötigt. Der fachliche und technische Aufbau dieser Systeme sowie typische Herausforderungen beim Umgang mit ihnen sind Aspekte, die eine fundierte Basis zur weiteren Bearbeitung liefern.

Um IT-Projektmanagement-Vorgehensmodelle in Bezug auf Big-Data-Management-Systeme miteinander vergleichen zu können, müssen diese vorab umfänglich charakterisiert und voneinander abgegrenzt werden. Analog zum Themengebiet Big Data ist es auch hier erforderlich, den Hintergrund und die Herkunft von IT-Projektmanagement-Vorgehensmodellen zu erläutern. Im Anschluss werden beide Vorgehensmodell-Typen anhand von Beispielen definiert und vorgestellt, sodass im zweiten Teil grundsätzliche Unterschiede bekannt sind und zielgerichtet in den Big-Data-Kontext gesetzt werden können.

Der zweite Teil der vorliegenden Arbeit beginnt mit der Definition von Einführungsprozessen für Big-Data-Management-Systeme. Diese Prozesse stellen die Basis für die darauf folgenden Nutzwertanalysen von Vorgehensmodellen und Big Data-Strategien dar. Die ersten Schritte der Nutzwertanalyse bilden die Erarbeitung eines Kriterienkatalogs sowie die Plausibilisierung des dazugehörigen Bewertungssystems. Hierbei geht es insbesondere darum, klare und nachvollziehbare Maßstäbe mit Blick auf die Analyse und Auswertung zu definieren. Sowohl die Kriterien als auch die Bewertungsskalen gehen im weiteren Verlauf begründet und verifiziert in die Nutzwertanalyse ein. Die Ergebnisse dieser Quantifizierung sind gleichzeitig Ansatzpunkt für die darauf folgenden Handlungsempfehlungen.

Um die erarbeiteten Ergebnisse und Handlungsempfehlungen zu validieren, wird daran anschließend eine beispielhafte Anwendung im Rahmen einer Fallstudie durchgeführt. Unter Berücksichtigung verschiedener Rahmenbedingungen und Produkthanforderungen wird ein Big-Data-Einführungsprojekt abhängig vom angewendeten Vorgehensmodell-Typen auf seinen Erfolg untersucht.

## 2 Big Data

In diesem Kapitel geht es darum, ein theoretisches Fundament im Themenfeld Big Data zu schaffen, um somit eine belastbare Ausgangsbasis für spätere Analysen herzustellen. Zunächst wird dabei Big Data im übergeordneten Sinne betrachtet. In diesem Zuge werden Hintergrund und Herkunft skizziert sowie eine detaillierte Beschreibung der Eigenschaften von Big Data gegeben. Daran anschließend werden Systeme zum Umgang mit Big Data, sogenannte Big-Data-Management-Systeme, und die damit einhergehenden Herausforderungen erläutert. Zur Differenzierung der unterschiedlichen Anwendungszwecke von Big Data werden darüber hinaus vier verschiedene Big-Data-Strategien vorgestellt.

### 2.1 Hintergrund und Herkunft

Schon in der eingangs zitierten Studie des McKinsey Global Institute (MGI) aus dem Jahr 2011 wird das exponentielle Datenwachstum sowie die Analysierbarkeit umfangreicher und vielfältiger Datenmengen als Haupttreiber des Interesses an Big Data identifiziert. Das Interesse besteht insofern, als dass Big Data einen signifikanten Nutzen für die Wirtschaft, Konsumenten und Regierungen darstellen kann. Dieser Nutzen äußert sich in Form von Produktivitätssteigerungen, die durch gezielte Sammlung, Analyse und Auswertung von großen heterogenen Datenmengen ermöglicht werden (vgl. [MCB<sup>+</sup>11] p. 15). Kurz darauf wurden die Technologien rund um Big Data im „Hype Cycle for Emerging Technologies 2011“ des IT-Marktforschungsunternehmens Gartner zum ersten Mal mit aufgenommen (vgl. [Gar11a]).

Das Ausmaß des prognostizierten Datenwachstums wird in der jährlich erscheinenden Studie „Digital Universe“ des Forschungsunternehmens International Data Corporation (IDC) im Auftrag von EMC verdeutlicht. In besagter Studie aus dem Jahr 2014 geht IDC davon aus, dass das sogenannte digitale Universum von 2013 bis 2020 um den Faktor 10 von 4,4 Trillionen auf 44 Trillionen Gigabytes anwachsen und sich somit alle zwei Jahre verdoppeln wird. Zudem stellt IDC fest, dass von der vorhandenen Datenmenge derzeit nur 22 Prozent analysierbar sind. Dieser Anteil steigt bis 2020 auf 35 Prozent, was hauptsächlich durch die zunehmende strukturierte Datenproduktion eingebetteter Systeme zu begründen ist (vgl. [Idc14]).

Die Ursache für dieses rasante Wachstum erkennt die BITKOM in ihrem Praxisleitfaden für Big Data vor allem in Technologien wie RFID, Machine-to-Machine-Lösungen, Smartphones, Cloud Computing und der intensiven Nutzung von Social Media (vgl. [Bit12]). Die große Datenmenge allein kann das technologische Phänomen Big Data jedoch nicht hinreichend beschreiben. Innovative Konzepte sowie die notwendige Rechenleistung zur Datenverarbeitung und -analyse, welche solch umfangreiche Datenfluten nutzbar machen können, stellen eine zweite wichtige Säule von Big Data dar (vgl. [HL13] p.14).

In diesem Zusammenhang ist es erstaunlich, dass der Technologie-Journalist Chris Anderson im Wired Magazin bereits im Jahr 2008 ein wegweisendes und vorausschauendes Plädoyer für die aufkommende Ära von Big Data hält. Er spricht dort vom „Ende der Theorie“ aufgrund immer besser werdender Datenanalysemöglichkeiten:

„This is a world where massive amounts of data and applied mathematics replace every other tool that might be brought to bear. Out with every theory of human behavior, from linguistics to sociology. Forget taxonomy, ontology, and psychology. Who knows why people do what they do? The point is they do it, and we can track and measure it with unprecedented fidelity. With enough data, the numbers speak for themselves [And08].“

Die Verknüpfung großer Datenmengen mit ausgefeilten Analysen auf Basis gezielter Messungen lässt hochgenaue Rückschlüsse auf das menschliche Verhalten zu, was Anderson zufolge wiederum Klassifizierung und Ontologie obsolet erscheinen lässt.

Zusammenfassend lässt sich demnach feststellen, dass Teilaspekte von Big Data bereits 2008 und auch schon früher diskutiert wurden – jedoch nur vereinzelt unter dieser Bezeichnung. Symptomatisch für die Etablierung des Begriffs ist der im Jahr 2010 veröffentlichte erste eigene Wikipedia-Artikel zum Thema Big Data, welcher diverse vertrauenswürdige Quellenangaben aufweist. Im selben Jahr beginnt der Hype und erste Anwendungsbeispiele werden vorgestellt (vgl. [Mcb12]). Der Inhalt der Begrifflichkeit bewegt sich nach wie vor zumeist im Bereich des Handling und der Analyse großer Datenbestände und findet seinen Ursprung in dem rapide steigenden Datenaufkommen sowie den sich kontinuierlich verbessernden Technologien zur Verwertung.

## **2.2 Definition und Eigenschaften von Big Data**

Wie in den vorangegangenen Abschnitten bereits angeklungen ist, gibt es keine einheitliche Festlegung, welche fachlichen und technischen Themenfelder der Begriff Big Data repräsentiert. In diesem Kapitel werden zum einen Deutungsversuche in der einschlägigen Literatur vorgestellt und zum anderen typische Eigenschaften von Daten, die unter den Begriff Big Data fallen, erläutert.

### **2.2.1 Deutungsversuche in der Literatur**

Die europäische Technologieplattform Networked European Software and Service Initiative (NESSI) hat in ihrem Whitepaper zum Big-Data-Konzept einen treffenden Definitionsvorschlag geliefert:

„Big Data“ is a term encompassing the use of techniques to capture, process, analyse and visualize potentially large datasets in a reasonable timeframe not accessible to standard IT technologies. By extension, the platform, tools and software used for this purpose are collectively called “Big Data technologies” ([Nes12], p.6).”

Diese Umschreibung des Themengebiets umfasst sowohl die Erfassung, die Weiterverarbeitung, die Analyse als auch die Visualisierung potentiell großer Datenmengen in einem vergleichbar kurzen Zeitraum. Diese Verarbeitung ist darüber hinaus nur mit speziellen IT-Technologien möglich. Die NESSI merkt jedoch im gleichen Zuge an, dass es sich bei Big Data um kein revolutionär neues Konzept handle, sondern dass es als ein sich weiterentwickelndes Technologiefeld gesehen werden kann (vgl. [Nes12], p.6).

Einen deutlich unpräziseren, dafür aber umso innovativeren und globaleren Definitionsversuch liefert der Oxford-Professor und Big-Data-Experte Viktor Mayer-Schönberger. Er beschreibt in seinem Standardwerk zu Big Data den Begriff folgendermaßen:

„[...] big data refers to things one can do at a large scale that cannot be done at a smaller one, to extract new insights or create new forms of value, in ways that change markets, organizations, the relationship between citizens and governments, and more ([MC13], p.6).“

Nach Mayer-Schönberger definiert sich Big Data demnach aus der Fähigkeit heraus, neue Erkenntnisse zu gewinnen bzw. Werte zu schaffen, die in dieser Form nur über die Eigenschaft der großen Masse an Daten erzeugt werden können. Kleinere Datenmengen seien dazu nicht in der Lage. In diesem Zusammenhang spricht er auch davon, dass sich die Gesellschaft davon lösen müsse, ständig ausschließlich nach Kausalitäten zu suchen. Dies sei eine Obsession aus vergangenen Zeiten, in denen Datenknappheit herrschte. Big Data mache Korrelationen aussagekräftiger als statistische Stichprobenanalysen und somit zähle zukünftig vielmehr das Was als das Warum. Um diese Behauptung zu stützen zieht Mayer-Schönberger unter anderem eine Analogie aus der Nanotechnologie heran. Dort werden Dinge kleiner und nicht größer, um so auf Molekularebene die sich ändernden physikalischen Materialeigenschaften auszunutzen. Mit dem Wissen über diese neuen Eigenschaften lassen sich Materialien mit völlig neuen Eigenschaften herstellen, wie beispielsweise flexiblere Metalle oder biegsame Keramik. Bezogen auf den Big-Data-Kontext bedeutet dies, dass mit der Nutzbarmachung extrem großer Datenmengen neue, zuvor unmögliche Dinge realisierbar sind, die mit kleinen Datenmengen nicht durchführbar wären (vgl. [MC13], pp.6-12).

Davenport hingegen hält den Begriff Big Data für irreführend und sagt ihm deshalb eine relativ kurze Lebensdauer voraus. Er stellt fest, dass das Adjektiv big, also groß, nicht das treffende Unterscheidungskriterium der neuen Daten sei. Dabei bezieht er sich auf eine



Studie von NewVantage Partners aus dem Jahr 2012, in der man zu dem Schluss kommt, dass das Problem der fehlenden Struktur der Daten weitaus bedeutender ist als deren reine Menge. In genannter Umfrage gaben 30 Prozent der Befragten an, dass sie sich mit Big Data beschäftigen, um Daten aus unterschiedlichen Quellen analysieren zu können. Weiterhin behauptet er, dass die Begrifflichkeit von vielen Verkäufern und Beratern fälschlicherweise für schon lange bekannte Technologien, wie Reporting oder Business Intelligence verwendet wird, um so die Produktattraktivität zu steigern (vgl. [Dav14] pp.6-7).

Es zeigt sich, dass Big Data in der Literatur unterschiedlich gedeutet wird und die inhaltlichen Schwerpunkte sich abhängig vom Standpunkt des Autors verschieben. Festzuhalten ist jedoch ein weitgehender Konsens darüber, dass es sich bei Big Data künftig um ein essentiell wichtiges Thema für die Gesellschaft, die Wirtschaft und Regierungen handelt. Die Potentiale, die sich aus diesen innovativen Technologien ergeben, müssen im Sinne des menschlichen Fortschritts erarbeitet und ausgenutzt werden.

### **2.2.2 Beschaffenheit von Big Data**

Bei Big Data geht es keinesfalls nur um die reine Datenmenge. Das stellt das Marktforschungsunternehmen Gartner in einem Artikel aus dem Jahr 2011 fest. Sie konstatieren weiterhin, dass neben der Menge auch andere Dimensionen berücksichtigt werden müssen. Um das Phänomen Big Data ganzheitlich beschreiben zu können und die daraus resultierenden Muster nutzenbringend zu erkennen und somit für bessere Geschäftsentscheidungen einzusetzen, bedarf es einer klar formulierten Systematisierung (vgl. [Gar11b]). Die von Gartner aufgeführten Dimensionen Volume (Menge), Velocity (Geschwindigkeit) und Variety (Vielfalt) gehen auf einen Forschungsbericht des Analysten Doug Laney zurück und sind aktuell auch als 3-V-Modell bekannt. In seinem Bericht aus dem Jahr 2001 spricht Laney von 3D Data Management, welches einen innovativen formalisierten Ansatz zum Umgang mit Daten in den Dimensionen Volume, Velocity und Variety repräsentiert (vgl. [Lan01], p.1). Im Folgenden sollen nun die Charakteristika der einzelnen Dimensionen detailliert erläutert werden:

#### **Volume**

Wie der Name schon sagt, handelt es sich bei Big Data oft um sehr große Datenmengen. Die Gründe für diese Ausmaße wurden bereits in Kapitel 2.1 erörtert. Darüber hinaus bedarf es zum Verständnis dieser Dimension einer Vorstellung davon, wie und wo diese großen Massen entstehen. Das Business-Intelligence-Unternehmen Domo hat 2012 in seinem Blog eine Infografik veröffentlicht, die darstellt wie viele Daten in jeder Minute generiert werden. So erhält beispielsweise Google in jeder Minute über 2 Millionen Suchanfragen, Facebook-Nutzer teilen weit mehr als 650.000 Inhalte und Youtube-Nutzer laden durchschnittlich 48 Stunden neues Videomaterial hoch. Daneben werden unzählige E-Mails übermittelt und etwa

100.000 Twitter-Nachrichten versendet (vgl. [Jam12]). Traditionelle Datenbanksysteme sind diesen enormen Datenansammlungen im Petabyte-Bereich kaum noch gewachsen. Es gibt zwar Datenbanksysteme, die diese Mengen abbilden können, jedoch stehen die Kosten oft nicht im Verhältnis zum Wert der Daten (vgl. [KTH13]). Deshalb müssen neue Lösungen für Architekturen zur Speicherung und Verarbeitung herangezogen werden. Diese Lösungen bestehen zum einen in der Speicherung über Cloud-Systeme und zum anderen in der Verarbeitung über MapReduce- und noSQL-Algorithmen (vgl. [Nes12], p.7).

### **Velocity**

Das Attribut Velocity bedeutet im Wortsinn Geschwindigkeit und bezeichnet inhaltlich den Wechsel hin zu Datenströmen sowie die ständige Verfügbarkeit von Zugriff und Übermittlung. Die Geschwindigkeit von Daten ist laut Gartner in diesem Kontext aus zwei Richtungen zu betrachten. Zum einen beschreibt sie das Tempo, mit dem Daten generiert werden. Beispiele hierfür sind bereits in der Dimension Volume genannt worden. Zum anderen geht es darum, wie schnell die Daten verarbeitet werden müssen, um den jeweiligen Anforderungen gerecht zu werden (vgl. [Gar11b]). Diese Anforderungen bestehen in der ständigen Zugänglichkeit der Daten, der Anwendbarkeit und Relevanz der Daten sowie der notwendigen Bereitstellungs-geschwindigkeit für Auswertungen (vgl. [PCC<sup>+</sup>12], p. 11).

### **Variety**

Die Vielfältigkeit der Daten ist neben der reinen Menge die bedeutendste Eigenschaft von Big Data. Mit der Vervielfältigung der Datenquellen geht eine massive Ausdehnung verschiedener Datenformate einher. Diese reicht von strukturierten Informationen, wie Log-Dateien, bis hin zu freiem Text und Media-Dateien (vgl. [Nes12], p. 8). Die Vervielfältigung und die Entstehung neuer Datenquellen innerhalb und außerhalb von Organisationen setzt die IT in Bezug auf Integration, Architektur und Governance unter Druck (vgl. [PCC<sup>+</sup>12], p. 11). Traditionelle relationale Datenbanksysteme sind den zum Teil massiven Strukturunterschieden von Daten im Kontext von Big Data kaum noch gewachsen. Zudem gibt jeder Typ von Big Data die ihm zugrunde liegenden Analysealgorithmen vor (vgl. [KTH13]). In der Literatur unterscheidet man drei Abstufungen der Struktur von Daten: strukturierte Daten, halbstrukturierte Daten und unstrukturierte Daten.

Ausschließlich strukturierte Daten lassen sich mit klassischen Datenbanksystemen gut verarbeiten. Ihnen wurde eine Struktur im Rahmen der Eingliederung in die Datenbank auferlegt. Dementsprechend sind der Speicherort, die Bedeutung und die Relation zu anderen Daten bekannt. Ein Beispiel für strukturierte Daten sind Kundenstammdaten mit Text- und Zahlenanteilen, wie Name, Geburtsdatum oder Adresse. Die Unterscheidung von halbstrukturierten und unstrukturierten Daten ergibt sich als problematisch und schwammig,

da oft zumindest ein geringer Teil an Struktur vorliegt. Wenn Daten ein organisatorisches Schema oder Tags wie beispielsweise bei XML-Webdokumenten (Extensible Markup Language) vorweisen, fallen sie in die Kategorie halbstrukturierter Daten. In diesem Fall ist es wesentlich einfacher, die Daten zu organisieren und Analysen auf ihnen durchzuführen. Typische halbstrukturierte Daten, die häufig zunächst als unstrukturierte Daten angesehen werden, sind Text aus XML-Dateien, E-Mails, Webserver-Logdateien sowie Sensordaten. Haben Daten hingegen gänzlich keine organisatorischen Merkmale, so spricht man von unstrukturierten Daten. Dies beinhaltet beispielsweise freien Text, Bilder, Videos, Audio-dateien und zum Teil auch Social-Media-Inhalte. Derlei unstrukturierte Daten haben kein Schema, keine Ontologie und keine Metadaten, die Informationen über die eigentlichen Inhalte liefern können (vgl. [Vor13], [Nas12]).

Die Gesellschaft für Informatik schlägt in einem weiteren Schritt vor, halbstrukturierte und unstrukturierte Daten in drei Kategorien aufzuteilen. Die erste Kategorie beinhaltet Daten, die aus der Kommunikation von Mensch zu Mensch entstehen. Dies sind beispielsweise Web-Protokolldateien, virtuelle Communities oder Social-Media-Netzwerke. In der zweiten Kategorie geht es um den Austausch von Menschen und Maschinen. So fallen E-Commerce-Transaktionsdaten, Daten aus medizinischen Geräten oder auch Daten, die bei der Nutzung von Maschinen wie z.B. Geldautomaten entstehen, in diese Klasse. Der dritten Kategorie werden Daten zugeordnet, die aus der Kommunikation zweier Maschinen bzw. Diensten heraus entstehen, wie zum Beispiel Sensordaten, GPS-Positionsinformationen oder Überwachungsbilder (vgl. [KTH13]).

Neben den genannten drei Dimensionen Volume, Velocity und Variety existieren in der Literatur zum Teil noch weitere Dimensionen, welche die Eigenschaften von Big Data spezifizieren. So beschreibt die NESSI über die Dimension Value (Wert) den subjektiven wirtschaftlichen Wert der Daten, die bereits aus z.B. regulatorischen Gründen erfasst und archiviert werden, aber noch nicht effektiv und wertschöpfend analysiert und verwertet werden (vgl. [Nes12], p. 8). Des Weiteren fügen beispielsweise Page et al. die Dimension Veracity (Richtigkeit) hinzu, um damit die Problematik der Qualität und Quelle der rezipierten Daten aufzuzeigen. Die Qualität wird unter anderem von Unvollständigkeit und Inkonsistenz beeinflusst. Da mit Hilfe von Big Data Geschäftsentscheidungen auf Basis von Daten getroffen werden, ist deren Nachvollziehbarkeit und Begründbarkeit unverzichtbar (vgl. [PCC<sup>+</sup>12], p. 11).

Beispiele für Daten, die unter den Begriff Big Data fallen, sowie die Evolution hin zu Big Data unter Berücksichtigung der einzelnen Entwicklungsstufen zeigt Abbildung 3 von King aus dem Jahr 2013. Sie fasst die gesammelten Erkenntnisse über die Beschaffenheit von Big Data zusammen. So definieren sich die Entwicklungsstufen in dieser Darstellung über die Datenmengen-Größenordnungen, in denen die aufgeführten Systeme und Technologien

üblicherweise arbeiten. Parallel zum wachsenden Datenvolumen steigen die Datenvielfalt und deren Komplexität.

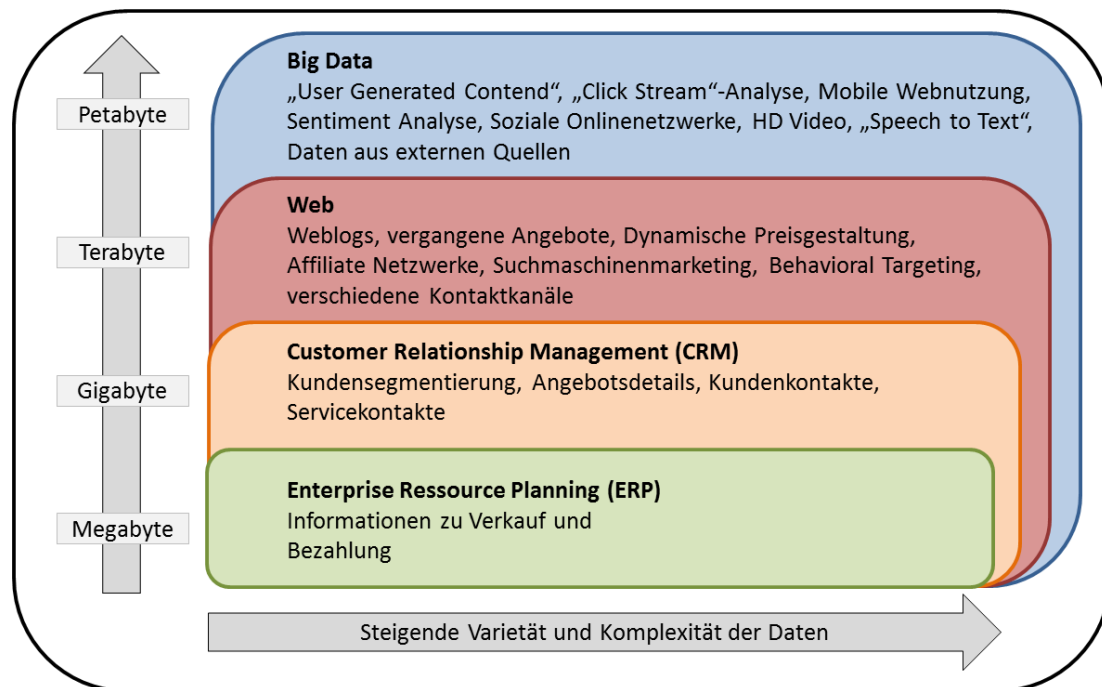


Abbildung 3: Big-Data-Evolution ([Kin14], p.36 in Anlehnung an Teradata, 2012)

Zusammenfassend ist festzustellen, dass die Beschaffenheit von Big Data am ehesten über die Menge und die strukturelle Vielfalt beschrieben werden kann. Dabei ist die größte Herausforderung, diese neuartigen, verschieden organisierten und oft unstrukturierten Daten sinnvoll zu verknüpfen und auszuwerten (vgl. [Dav14], p.8). Wie eine solche Verwertung im Rahmen von Big-Data-Management-Systemen vonstattengehen kann, wird im weiteren Verlauf dieser Ausarbeitung noch thematisiert.

## 2.3 Big-Data-Management-Systeme

Dieses Kapitel befasst sich mit dem Aufbau von Big-Data-Management-Systemen. Dabei geht es zum einen um die notwendigen Technologien, welche die Nutzbarmachung von Big Data erst ermöglichen, und zum anderen um die Technologie-Schichten und mögliche Datenarchitekturen von Big-Data-Management-Systemen.

### 2.3.1 Technologien für Big Data

Um große Mengen verschieden strukturierter Daten handhaben zu können, bedarf es spezieller Big-Data-Technologien. In diesem Abschnitt werden die in der Literatur am häufigsten genannten grundlegenden Technologien vorgestellt. Hierbei ist zu beachten, dass die fachliche Tiefe der Darstellung dem Ziel dieser Ausarbeitung entsprechend hinreichend ist. Das bedeutet, dass tiefgehende Informationen zur Funktionsweise der einzelnen Technologien hier nicht gegeben werden. Vielmehr geht es darum, aufzuzeigen welche

Technologien im Kontext von Big Data neben den traditionellen Data-Warehouse-Lösungen existieren. Hierzu wird in Tabelle 1 zunächst ein Überblick über die Technologien gegeben, die in diesem Abschnitt vorgestellt werden.

Technologie	Kurzdefinition
Hadoop	Freies, skalierbares und verteilt arbeitendes Framework zur parallelen Verarbeitung von Big Data auf mehreren Servern
MapReduce	Programmiermodell zur parallelen Berechnung großer Datenmengen auf Computerclustern
NoSQL	NoSQL (Not only SQL) bezeichnet Datenbanken, die nicht ausschließlich relational arbeiten
In-Memory-Technologie	Datenbanksysteme, die den Arbeitsspeicher eines Computers als Datenspeicher nutzen, um höhere Zugriffsgeschwindigkeiten zu realisieren
Machine Learning	Wissensgewinnung künstlicher Systeme auf Basis von Erfahrungswerten zahlreicher Iterationen
ETL / ELT	<ul style="list-style-type: none"> <li>▪ ETL ( Extraktion, Transformation, Laden): Daten extrahieren, Umwandeln und in Zieldatenbank laden</li> <li>▪ ELT (Extraktion, Laden, Transformation): Neues Prozess-Paradigma im Big-Data-Zeitalter</li> </ul>

Tabelle 1: Big-Data-Technologien im Überblick (Quellen s.u.)

## Hadoop

Wie in der Kurzdefinition bereits angedeutet, handelt es sich bei Hadoop nicht nur um eine großes Computercluster sondern vielmehr um ein Framework, welches sich gut eignet, um große Datenmengen verschiedener Struktur und Quelle zu speichern und weiterzuverarbeiten. Hadoop ist eine Open-Source-Plattform, die von der Apache Software Foundation entwickelt wurde und sich in ständiger Weiterentwicklung befindet. Das Besondere an diesem Framework ist, dass es ein vergleichsweise einfaches Programmiermodell liefert, das die verteilte Bearbeitung umfangreicher Datenbestände auf beliebig großen Gruppen von Computern, sogenannten Clustern, ermöglicht. Das Intel IT Center hat diese Technologie in einem Whitepaper 2013 vorgestellt und nennt dabei diverse Gründe, weshalb sie große Vorteile im Umgang mit Big Data bereitstellt. Zunächst erlaube Hadoop die Speicherung von Daten in ihrem nativen Format. Es sei keine Umwandlung oder Bereinigung notwendig, bei der Informationen verloren gehen könnten. Weiterhin sei Hadoop bereits im großen Stil in der Praxis durch Unternehmen wie Facebook und Yahoo! erprobt und im Einsatz. Ein weiterer Nutzen sei die Tatsache, dass keine Lizenzgebühren anfallen und das Cluster inkrementell zu vergleichsweise sehr geringen Kosten erweitert und darüber hinaus auf Standardservern betrieben werden kann. Der wohl größte Vorteil ist laut Intel IT

Center jedoch die weite Verbreitung von Hadoop in Bezug auf Big Data. Sehr viele Big-Data-Experten entwickeln Anwendungsfälle auf diesem Framework, was dazu führt, dass es fortlaufend verbessert wird und das Hadoop-Wissen sämtliche Branchen erreicht (vgl. [Int13], p. 7).

### **MapReduce**

Als MapReduce bezeichnet man das Programmier-Rahmenwerk in einer Hadoop-Umgebung. Es vereinfacht die Verarbeitung von Big Data und stellt dem Programmierer eine Methode zur Verfügung, mit der er komplexe Verarbeitungsaufgaben auf mehrere Computer in einem Cluster verteilen kann. Das Intel IT Center skizziert den Ablauf einer MapReduce-Operation in zwei wesentlichen Schritten. Im ersten Schritt, dem „Mapping“, wird der zu verarbeitende Datenbestand in mehrere unabhängige Datenblöcke aufgeteilt, um eine parallele Bearbeitung zu ermöglichen und einzuleiten. Die daraus resultierenden Zwischenergebnisse werden im Anschluss sortiert, was als „Shuffling“ bezeichnet wird, und zum zweiten Schritt, dem „Reducing“, weitergeleitet. Die Reduce-Funktion ist genau wie die Map-Funktion vom Programmierer festgelegt. Sie bestimmt, wie die Zwischenergebnisse clusterübergreifend aggregiert werden und schließlich als Verarbeitungsergebnis ausgegeben werden. Üblicherweise findet die Verarbeitung jedes einzelnen Datenblocks aus Effizienzgründen auf genau dem Computer im Hadoop-Cluster statt, auf dem der Datenblock initial gespeichert wurde (vgl. [Int13], p. 8). Diese Operationen können mit Hilfe von innovativen Scriptsprachen, wie Apache-Pig und -Hive, die auf dem Hadoop-Framework aufsetzen, durchgeführt werden. Dabei stellt Pig Funktionalitäten zum Lesen, Schreiben, Umwandeln und Zusammenführen von Daten bereit. Hive hingegen ermöglicht die Umwandlung von Daten in ein relationales Format, auf welches in einem zweiten Schritt durch klassische SQL-Abfragen zugegriffen werden kann (vgl. [Dav14], p. 121).

### **NoSQL**

NoSQL (not only Structured Query Language) stellt einen Überbegriff für eine Reihe von Datenbank-Management-Systemen dar, die auch über relationale Datenbanken hinweg Anwendung finden und deshalb im Kontext von Big Data genannt werden. NoSQL-Datenbanken spielen also immer eine Rolle, wenn die zu verarbeitenden Daten nicht einem relationalen Datenmodell entsprechen und eine höhere Skalierbarkeit und Flexibilität benötigt wird. Beispiele für Anwendungsszenarien sind Social Media, Prognosemodelle, Web-Applikationen, umfangreiche Geschäftsanalysen und Echtzeit-Applikationen. Die BITKOM empfiehlt die Nutzung von NoSQL-Datenbanken immer dann, wenn die Anwendungsanforderungen nicht mit relationalen Datenbanken wirksam abgebildet werden können (vgl. [Bit14b], p. 44).

## **In-Memory-Technologie**

Kernaspekt der In-Memory-Technologie ist die Verlagerung von Datenspeicherung, Datenverarbeitung und Datenanalyse in den Arbeitsspeicher eines Computers. Dadurch kann die Zugriffsgeschwindigkeit auf Daten um mehrere Zehnerpotenzen erhöht werden. Dies ist der Grund, weshalb In-Memory-Technologien insbesondere bei Big-Data-Prognosemodellen und Echtzeit-Analysen zum Einsatz kommen. Die Technologie zur Speicherung und Verarbeitung von Daten im Arbeitsspeicher ist nicht neu. Durch kontinuierlich fallende Kosten für Speichermedien und Rechenperformance sowie den zunehmenden Komfort bei Applikationen greifen nicht wenige Unternehmen auf diese Big-Data-Technologie zurück. Die Ziele einer In-Memory-Implementierung sind divergent und variieren von Zeitersparnis über verbesserte Unternehmensentscheidungen in Echtzeit bis hin zur gezielteren Kundenanalyse und -ansprache. In-Memory-Technologien werden von zahlreichen Herstellern angeboten. Ein typisches Beispiel ist HANA (High-Performance Analytic Appliance) von SAP (vgl. [Bit14b], p. 46).

## **Machine Learning**

Beim maschinellen Lernen geht es um Methoden und Anwendungen, in denen Programme selbstständig mit Hilfe statistischer und logischer Analysen auf Basis zahlreicher Iterationen neues Wissen erwerben. Dies widerspricht dem klassischen hypothesenbasierten Ansatz eines Analytikers, der die Eignung seiner Hypothese mit einem gegebenen Datensatz testet und validiert. Big Data ist für diesen Ansatz zu voluminös, zu variabel und zu volatil (vgl. [Dav14], p. 116). Machine Learning ist demnach ein Verfahren, welches vorteilhafte Lösungsansätze für Probleme, die mit klassischen Methoden nur unter hohem Kostenaufwand durchführbar wären, automatisch ermittelt und weiterentwickelt. Ein Anwendungsfall von Machine Learning ist die sogenannte Vorhersagen-Modellierung. Hierbei wird ein Modell aus bekannten Daten eingespielt und diverse Qualitätskriterien für die Prognose definiert. Die beobachtete Genauigkeit bietet eine Rückmeldung für das Lernverfahren und das Programm passt sich dementsprechend an (vgl. [Bit14b], p. 67).

## **ETL / ELT**

Big-Data-Umgebungen erfordern ein Überdenken der bestehenden Softwareprozesse bei der Datenintegration. Die Begrifflichkeit Extraktion-Transformation-Laden (ETL) stellt dabei einen erprobten Ansatz für klassische Data-Warehouse-Lösungen dar. Es geht darum, Daten aus vielen verschiedenen Quellen zu integrieren und dabei die Qualität der gewonnenen Informationen mit Hilfe von Datenbereinigungen sicherzustellen. Die Bereinigung und die Umformung der Daten zur Herstellung der Analysierbarkeit stellt in diesem Fall das Transformieren dar. Im Kontext von Big Data wird der ETL-Prozess vor

große Herausforderungen gestellt, da große Datenvolumina mit hoher Geschwindigkeit generiert und extrahiert werden und deshalb nicht mehr adäquat transformiert werden können. Zudem stellen Big-Data-Analysen nicht den Anspruch angepasste, dimensionale Datenformate zu verarbeiten. Sie bewältigen unvollständige und unstrukturierte Datenschemata, sodass es Anlass für ein Umdenken bei jenen Softwareprozessen gibt. Der Aufwand für die Datenextraktion ist im Umfeld von Big Data häufig vergleichbar gering. Direkt darauf folgt das Laden der Rohdaten in einen kostengünstigen Speicher, wie beispielsweise ein Hadoop-Cluster. Die BITKOM nennt dieses Paradigma „Speichere jetzt – Verarbeite später“. Durch die Skalierbarkeit der Rechenleistung innerhalb eines Hadoop-Rechnernetzes kann der Transformationsschritt selbst bei großen Datenmengen noch hinreichend schnell ausgeführt werden (vgl. [Bit14b], pp. 93-96).

### 2.3.2 Implementierung des Big-Data-Stapels

In einer Studie des International Institute for Analytics haben Davenport und Dyché den sogenannten Big-Data-Stapel entworfen (Abb. 4). Dieser beschreibt die verschiedenen Technologie-Schichten eines Big-Data-Management-Systems. Sämtliche Komponenten des Stapels sind dabei auf die Beschaffenheit von Big Data abgestimmt, sodass eine bewegliche und ganzheitliche Lösung zur Speicherung und Verarbeitung von großen unstrukturierten Datenmengen entsteht.



Abbildung 4: Der Big-Data-Stapel (in Anlehnung an [DD13], p. 10)

Die unterste Ebene des Big-Data-Stapels stellt laut Davenport und Dyché der Speicher dar. Dieser werde aufgrund von Technologien, wie z.B. Hadoop, immer günstiger, flexibler und skalierbarer. Die tatsächliche Speicherung von Big Data erfolgt zumeist auf Standardservern mit konventionellen Festplatten, die über virtuelle Cluster koordiniert werden. Die Big-Data-Plattform-Infrastruktur als zweite Schicht umfasst typischerweise eine Reihe von Funktionen, welche die Hochleistungsdatenverarbeitung ermöglichen. Auch diese Schicht wird in der



Praxis häufig durch eine Hadoop-Umgebung bereitgestellt. Nichtsdestotrotz verweisen Davenport und Dyché darauf, dass Hadoop nicht zwangsläufig die einzige Lösung sein müsse. Vielmehr bestehe eine Plattform-Infrastruktur, insbesondere bei großen Unternehmen, zumeist aus einer Mischung von traditionellen und neuartigen Systemen. Die Daten selbst erhalten aufgrund ihrer fundamentalen Bedeutung für den Nutzen von Big Data eine eigene Schicht. Davenport und Dyché fordern in ihrem Modell vor diesem Hintergrund eine separate Data Governance. Da es zur Data Governance im Kontext von Big Data durchaus kontroverse Auffassungen gibt, wird im weiteren Verlauf der Arbeit noch deutlich (vgl. [DD13], pp. 9-11; [Man14]).

Die Ebene des Anwendungscode variiert stark in Abhängigkeit von der zu bearbeitenden Aufgabe des Big-Data-Management-Systems. In einem Hadoop-Umfeld käme hier beispielsweise das Programmier-Rahmenwerk MapReduce zum Einsatz, um sämtliche Operationen zu instruieren. Die vorletzte Schicht des Big-Data-Stapels ist ebenfalls stark abhängig vom Anwendungsfall und stellt die geschäftsbezogene Sicht auf die Daten dar. Davenport und Dyché beziehen sich hierbei auf eine intermediäre Datenstruktur wie z.B. ein statistisches Modell, ein strukturierter Text oder eine relationale Tabelle. Diese Struktur ließe sich daraufhin mit Hilfe hinlänglich bekannter Werkzeuge wie z.B. SQL tiefgehend analysieren. Die Unternehmenssicht stellt somit die Ausrichtung der Daten auf die Analyseaufgabe dar. Eine herausragende Bedeutung kommt der Anwendungsschicht zuteil. Sie erlaubt entgegen bestehender Datenvisualisierungen aus Business-Intelligence-Systemen eine intuitive grafische Datenanalyse, die auch durch einen Mitarbeiter ohne ausgeprägte Big-Data-Kenntnisse durchgeführt werden kann. Hierbei ist jedoch darauf zu achten, dass die Visualisierung nicht mehr als zwei Dimensionen umfasst, da das menschliche Gehirn diese nur mit Mühe verstehen kann. Dies kann beispielsweise bei multivariaten Vorhersagemodellen zu einer großen Herausforderung werden (vgl. [DD13], pp. 11-13).

Die Implementierung von Elementen des Big-Data-Stapels ist eine umfassende organisatorische Aufgabe und erfordert ein hohes Maß an Planung. Das IBM Institute for Business Value hat 2012 in Zusammenarbeit mit der Saïd Business School der Universität von Oxford im Rahmen einer Studie eine Befragung von Unternehmen zu ihrem Big-Data-Entwicklungsstand durchgeführt. Sie haben die Studienteilnehmer darüber befragt, welche Big-Data-Aktivitäten in ihrem Unternehmen gerade verfolgt werden. Aus den Ergebnissen lassen sich vier Hauptphasen der Einführung ableiten (Abb. 5).



Abbildung 5: Big-Data-Einführungsprozess ([Ibm12], p. 13)

Etwa ein Viertel der Befragten befinden sich in der Informationsphase. In dieser ersten Phase der Einführung geht es darum eine fundierte Wissensbasis aufzubauen, Führungskräfte zu sensibilisieren und Geschäftschancen durch Big Data zu identifizieren. Beinahe die Hälfte der Studienteilnehmer befindet sich in der Planungsphase. Diese Phase umfasst die Erstellung von Business Cases und Roadmaps für mögliche Big-Data-Strategien und die damit verbundenen individuellen Big-Data-Stapel. Die dritte Phase der Einführung wird unter dem Begriff Prüfen zusammengefasst. Knapp ein Viertel der teilnehmenden Organisationen befinden sich in dieser Phase, welche die Potentialbewertung sowie den Start erster kleinerer Pilotprojekte beinhaltet. Innerhalb eines begrenzten Spielraums prüfen Unternehmen dieser Gruppe den Geschäftsnutzen und die Konformität der Big-Data-Strategie mit der Unternehmensstrategie. Die vierte und letzte Phase, die das IBM Institute for Business Value identifiziert, impliziert die operative und technologische Umsetzung auf breiter Basis. In dieser Phase befinden sich der Studie zufolge jedoch lediglich sechs Prozent der befragten Unternehmen (vgl. [Ibm12], pp. 13-14).

### 2.3.3 Datenarchitekturen für Big-Data-Management-Systeme

Die Datenarchitektur einer typischen bestehenden Analyseumgebung in einem großen Unternehmen setzt sich stark vereinfacht aus drei Ebenen zusammen: operative Systeme, die als Datenquellen dienen, einem Data-Warehouse mit mehreren Data-Marts sowie einer Sammlung von Analysewerkzeugen zur Entscheidungsunterstützung (vgl. [Dav14], p. 126). Die Rohdaten gelangen dabei in strukturierter Form aus den operativen Systemen, wie z.B. einem Enterprise-Resource-Planning-System, über den in Kapitel 2.3.1 beschriebenen ETL-Prozess in relationale Datenbanken im Data-Warehouse. Die ursprünglichen Rohdaten werden nach der Transformation üblicherweise wieder gelöscht, sodass nur strukturierte relationale Daten im Data-Warehouse vorliegen. Diese Daten können in einem dritten Schritt mit geeigneten Anwendungen für Analysen und Visualisierungen aufbereitet werden (vgl. [Bit14b], p. 100).

Die BITKOM hat 2014 in einem Praxisleitfaden für Big-Data-Anwendungen Lösungsszenarien für eine Implementierung von Big-Data-Funktionen in eine bestehende Datenarchitektur aufgestellt. Innerhalb dieses Leitfadens kristallisieren sich dabei in einer Befragung von Unternehmen, die auf Hadoop basierende Big-Data-Projekte durchführen, drei Modelle für das Zusammenwirken von Hadoop und herkömmlichen Lösungen heraus: die Nutzung von Hadoop-Clustern als preisgünstiger Langzeitspeicher von Daten aller Art, die Exploration und Erforschung von Daten in einer Hadoop-Umgebung sowie Hadoop als eine unternehmensweite Plattform und Big-Data als Shared-Service. Die Komplexität und Ausgereiftheit der Lösung ist hierbei der Reihenfolge nach aufsteigend. Ergebnis der Befragung ist, dass die Komplementarität von herkömmlichen Systemen und Big-Data-Management-Systemen gegeben ist und insbesondere bei großen Unternehmen eine Koexistenz-Strategie sinnvoll erscheint (vgl. [Bit14b], pp. 100-102).

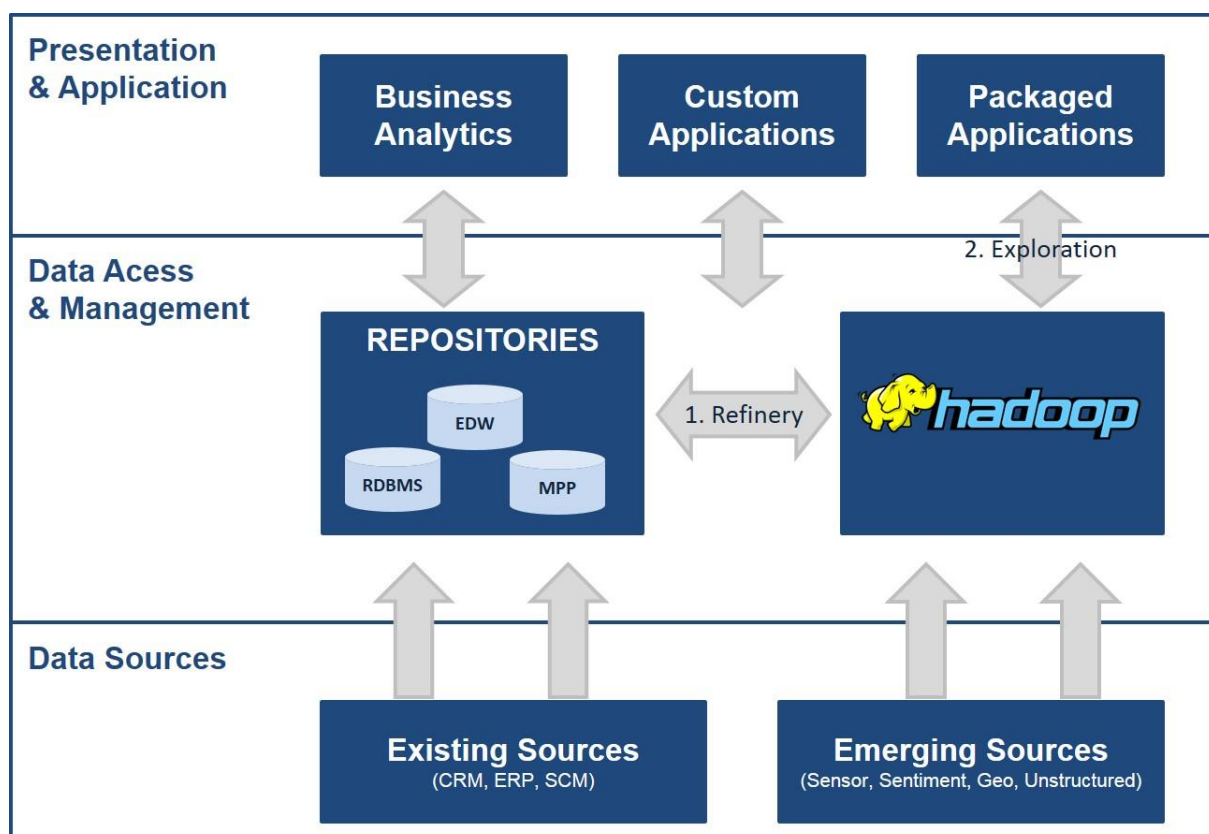


Abbildung 6: Zusammenspiel von Hadoop mit bestehenden Lösungen ([Bit14b], p. 102)

Wie das Zusammenspiel von bestehenden Lösungen mit einer Hadoop-Umgebung schematisch aussehen kann, stellt die BITKOM in Abbildung 6 dar. Der prinzipielle Aufbau gestaltet sich wie oben angedeutet in drei Schichten. Die Datenquellen als operative Systeme, das Datenmanagement innerhalb von Data-Warehouses sowie die Anwendungs- und Präsentationsebene mit entsprechenden Analysewerkzeugen. Wie in den vorangegangenen Kapiteln bereits erläutert, kommen auf der Ebene der operativen Systeme im Zuge von Big Data zahlreiche neue Datenquellen, wie z.B. Sensor- und Geodaten hinzu. Ein

klassisches Data-Warehouse käme bei der Verknüpfung bestehender und neuer Datenquellen schnell an seine Grenzen. Damit diese Kombination gelingt, wird auf der Datenmanagement-Ebene eine Big-Data-Umgebung in Form eines Hadoop-Clusters implementiert. Durch geeignete Anwendungen auf der obersten Schicht können Daten innerhalb der Hadoop-Plattform erforscht und somit neue Kenntnisse gewonnen werden. Diese Datenexploration stellt einen Vorprozess zur Speicherung im Data-Warehouse dar (vgl. [Bit14b], pp. 100-102).

Davenport konstatiert zusammenfassend in diesem Kontext, dass es unwahrscheinlich sei, dass Big-Data-Technologien wie Hadoop Data-Warehouses gänzlich ersetzen werden. Nichtsdestotrotz stellen sie eine neue Möglichkeit dar, Entscheidungen im Unternehmen zu unterstützen. Diese Prognose werde insbesondere dadurch bestärkt, dass große Unternehmen enorme Geldsummen in den Aufbau von Data-Warehouse-Plattformen investiert haben und diese darüber hinaus für ihren Zweck entsprechend gut funktionieren. Es gilt also, eine Kombination aus etablierten Data-Warehouses für Standardaufgaben und innovativen Big-Data-Umgebungen für Datenexploration zu gestalten, die sich an der angestrebten Big-Data-Strategie orientiert (vgl. [Dav14], p. 124-131).

## **2.4 Ziele und Strategien von Big Data**

Sämtliche unternehmerischen Aktivitäten dienen im Idealfall der Erreichung primärer Unternehmensziele und der Operationalisierung einer langfristigen Unternehmensstrategie. Dies können beispielsweise die Ergebnisverbesserung und die Umsatzsteigerung sein. Begibt man sich auf eine tiefere Abstraktionsebene, so finden sich auch für Big-Data-Aktivitäten unterschiedliche Zielrichtungen (vgl. [BKG14], p. 45). Das Anliegen, das mit einer Big-Data-Initiative in einer Organisation verfolgt wird, hat großen Einfluss auf den Einführungsprozess der notwendigen Big-Data-Management-Systeme. In diesem Abschnitt werden anhand von in der Literatur angeführten Anwendungsbeispielen mögliche Ziele und Strategien von Big Data vorgestellt. Dadurch soll ein Eindruck darüber vermittelt werden, in welche Richtungen eine Big-Data-Anwendung Nutzen und Mehrwerte schaffen kann. Davenport identifiziert hierbei vier verschiedene Zielrichtungen: Die Einsparung von Kosten durch Big-Data-Technologien, die Erhöhung der Verarbeitungsgeschwindigkeit, die Realisation innovativer Produkte mit Hilfe von Big Data sowie die Nutzung von Big Data zur Prognose und Entscheidungsunterstützung (vgl. [Dav14], pp. 58). Jede dieser Zielrichtungen wird zur Veranschaulichung mit Anwendungsbeispielen untermauert.

### **2.4.1 Kosteneinsparung durch Big Data**

Ende des Jahres 2013 hat das Marktforschungsunternehmen IDC im Rahmen einer Studie Entscheider von 282 Unternehmen in Deutschland befragt, welche Ausrichtungen und Absichten sie mit Big Data verfolgen werden. Darin geben 52 Prozent der Befragten an, dass

es eines der Hauptziele sei, größere Datenmengen speichern zu können (vgl. [Idc13a]). Mit der Fähigkeit große Datenmengen zu speichern, gehen oft auch hohe Kosten für die notwendigen Speichersysteme einher. Ein Hadoop-Cluster als eine Big-Data-Technologie kann hier zu enormen Kosteneinsparungen führen. Die BITKOM führt diesen Kostenvorteil darauf zurück, dass Hadoop zum einen ein freies Framework ist, für das keine Lizenzgebühren anfallen. Zum anderen läuft es zumeist auf kostengünstiger Hardware ohne Herstellerbindung. Das Ausmaß der Kostendifferenz lässt sich gut durch einen Vergleich der Ausgaben für die Speicherung von einem Terabyte Daten mit unterschiedlichen Speichertechnologien veranschaulichen. Die BITKOM hat dazu die Anschaffungskosten einer Lösung, folglich die Hardware und Software, die Installation sowie die Wartung für das erste Betriebsjahr veranschlagt und verschiedene Technologien gegenübergestellt. Typische Kosten für ein Hadoop-Cluster liegen demnach bei 250 bis 1.000 US-Dollar für ein Terabyte Daten. Die tatsächlichen Kosten hängen letztlich davon ab, mit welcher Intensität die gespeicherten Daten verarbeitet werden sollen. Für ein klassisches Network Attached Storage hingegen werden komplexitätsabhängig bis zu 20.000 US-Dollar pro Terabyte, für ein individuell angepasstes Enterprise Data Warehouse sogar bis zu 80.000 US-Dollar pro Terabyte angesetzt. Für datenintensive Unternehmen ergeben sich hierdurch substantielle Kostenpotentiale (vgl. [Bit14b], p.36). Dieser Vergleich berücksichtigt jedoch nicht die Tatsache, dass eine traditionelle Technologie zuverlässiger und sicherer zu verwalten ist und darüber hinaus zur Implementierung und zum Betrieb eines Hadoop-Clusters womöglich zusätzliches Personal benötigt wird. Es ist demzufolge sinnvoll die Kostenpotentiale nach dem Total-Cost-of-Ownership-Ansatz abzuwägen (vgl. [Dav14], p. 58).

Kosteneinsparungen auf Basis von Ausgabenreduzierung für Datenspeicherung als Primärziel einer Big-Data-Initiative sind in der Praxis tendenziell selten. Sie gehen aber häufig mit der Ausnutzung anderer Möglichkeiten durch Big Data, wie z.B. der Verbesserung der Analyse- und Prognosefähigkeit einher. Die BITKOM hat neben vielen weiteren Anwendungsfällen ein Beispiel für den Einsatz von Big Data in der Produktion und Instandhaltung geliefert, welches diesen Sachverhalt treffend illustriert. Bei dem Anwender handelt es sich um ein Unternehmen, das sich primär mit Energietechnik als Lieferant und deren Instandhaltung als Servicedienstleister befasst. Die Sensorik moderner Produktionsanlagen liefert kontinuierlich Informationen über den Zustand der Maschine. Diese Informationen werden als Daten mit Hilfe von Big-Data-Technologien auf relevante Ereignisse hin analysiert. Die Diagnose erfolgt wissensbasiert, d.h. auf Basis der Instandsetzungshistorie werden Trends erkannt und es können präventiv Maßnahmen eingeleitet werden. Dieser Abgleich führt dazu, dass Fehlalarme von relevanten Fehlermeldungen unterschieden werden können und sich der Supportmitarbeiter zielgerichteter und schneller um die Bedürfnisse des Kunden kümmern kann. Somit hat sich durch die Identifikation und Vorabselektion

von Fehlalarmen und relevanten Alarmen die Servicequalität bei geringeren Kosten erhöht. (vgl. [Bit12], p. 71).

### **2.4.2 Zeiteinsparung mit Hilfe von Big Data**

Ein weiteres Ziel einer Big-Data-Initiative ist nach Davenport zufolge die Einsparung von Zeit durch gezielte Nutzung von Big-Data-Werkzeugen. Dabei geht es in erster Linie darum, die benötigte Zeit für einen bekannten Prozess zu reduzieren. Zeitaufwändige, oft auch teils manuelle Analyseprozesse können mit performanceorientierten Big-Data-Management-Systemen deutlich verkürzt werden, wodurch aktuellere und bessere Untersuchungsergebnisse entstehen. Dies wird insbesondere dann wichtig, wenn sich die Rahmenbedingungen für ein Unternehmen häufig ändern oder Entscheidungen schnell getroffen werden müssen (vgl. [Dav14], p. 61). Ein in der Literatur häufig zitiertes Beispiel liefert das amerikanische Handelsunternehmen Macy's:

Macy's ist in den Vereinigten Staaten überregional mit etwa 800 Filialen vertreten. Das mehrere Zehntausend Artikel umfassende Sortiment soll zu bestmöglichen Preisen angeboten werden und wird deshalb standortspezifisch an den Absatz und die jeweilige Wettbewerbssituation angepasst. Das führt dazu, dass es für fast jeden Artikel im Sortiment regionsabhängig verschiedene Preise im System gibt. Die dadurch entstehenden etwa 270 Millionen Preispunkte müssen auf Basis des Absatzes der Vergangenheit möglichst schnell berechnet werden. Bevor Macy's In-Memory-Datenbanken eingeführt hat, benötigte diese Analyse ca. 30 Stunden Rechenzeit. Durch die Umstellung auf ein Big-Data-Management-System kann diese Zeit auf unter zwei Stunden reduziert werden. So ist es nun möglich, mitunter mehrfach am Tag Preise an den Absatz und den Wettbewerb anzupassen. Gerade im Handel spielt die Preisgestaltung eine enorm wichtige Rolle. Dies führt dazu, dass Macy's gegenüber seinen Wettbewerbern diesbezüglich einen signifikanten Wettbewerbsvorteil generiert (vgl. [Bit12], p.60).

Ein weiterer Anwendungsfall von Big-Data-Technologie, welcher ebenfalls das Ziel der Zeiteinsparung verfolgt, ist die sogenannte Fraud Detection bzw. Betrugserkennung. Dieser tritt insbesondere im Zusammenhang mit Kreditkartentransaktionen auf. Es geht hierbei aber weniger um Performanceerhöhung als vielmehr um die verbesserte Analysefähigkeit bei großen Datenbeständen, um Muster schneller zu erkennen und automatisch entsprechende Betrugserkennungsregeln zu erstellen. Ein solches System wurde beispielsweise vom Fraunhofer-Institut für Intelligente Analyse- und Informationssysteme (Fraunhofer IAIS) zusammen mit der Paymint AG entwickelt, um Betrugsmuster und Anomalien schnell zu identifizieren und somit Kartenmissbrauch zu verhindern. Die Zeitersparnis liegt dabei in erster Linie bei der Erstellung von Regelsätzen für neue Betrugsmuster sowie bei dem manuellen Aufwand bei der Abwicklung bisher unbekannter Fraud-Fälle (vgl. [Fra12]).

### 2.4.3 Entscheidungsunterstützung und Prognosen durch Big Data

Die Unterstützung von Unternehmensentscheidungen ist eine traditionelle Aufgabe der Datenanalyse und nicht erst seit dem Big-Data-Zeitalter bekannt. Klassische Business-Intelligence-Systeme verarbeiten schon seit Jahrzehnten vorhandene strukturierte Unternehmensdaten aus der Vergangenheit, um z.B. Absatzprognosen oder Preisanpassungen zu berechnen. Derartige Analysen mit großen, teils unstrukturierten Datenmengen, deren Quellen auch außerhalb der Organisation liegen können, anzureichern, kann ebenso ein Ziel einer Big-Data-Initiative sein. In dieser Zieldimension gibt es viele verschiedene Stoßrichtungen, die zum einen davon abhängen, welche Information man generieren möchte und zum anderen davon, welche Arten von Daten man in die Verarbeitung mit einbezieht. Typische Fragestellungen beziehen sich häufig auf Kundenbedürfnisse, Preispolitik, Nachfrageprognosen oder auch Bestandsmanagement. Davenport stellt vor diesem Hintergrund fest, dass die Fragestellungen oft dieselben wie früher sind, die genutzten Tools bzw. die nutzbaren Daten sich heute jedoch stark unterscheiden. Somit kann durch wertvolleres detaillierteres Wissen sogar ein Wettbewerbsvorteil gegenüber der Konkurrenz erzielt werden (vgl. [Dav14], pp. 64-67).

Die deutsche Drogeriemarkt-Kette dm hat durch den Einsatz von Big-Data-Technologien die Qualität und Präzision der Mitarbeiterereinsatzplanung entscheidend verbessert. Die haus-eigene Planungssoftware von dm gerät bei Sondersituationen schnell an ihre Grenzen, wodurch es häufig zu Personalengpässen bzw. -überschüssen kommt. Es handelt sich hierbei um ein Prognoseproblem, welches sich sehr gut in dieser Strategierichtung wiederfindet. Mit Hilfe des Big-Data-Dienstleisters Blue Yonder gelingt dm eine Lösung, die sowohl interne als auch externe Daten verschiedener Struktur berücksichtigt. So werden nicht nur die Tagesumsätze der Vergangenheit in die Analyse mit einbezogen, sondern auch individuelle Parameter wie Paletten-Anlieferzeiten, Öffnungszeiten, Wettervorhersagen, Markttag oder auch Ferien im Nachbarland. Diese Fülle an Informationen wird verdichtet und führt somit zu einer besser prognostizierbaren Mitarbeiterereinsatzplanung. Dies wiederum ermöglicht eine effizientere Arbeitsorganisation und sorgt für zufriedener Mitarbeiter und Kunden (vgl. [Bit12], p. 58).

Ein weiterer Anwendungsfall besteht darin, Informationen aus mehrkanaligen Kundenbeziehungen zu analysieren, um beispielsweise mit individuellen Angeboten zu werben und so den Umsatz zu steigern. Dies sind zumeist Daten unterschiedlichster Struktur, wie z.B. Transaktionsdaten, Klicks auf Websites, Sprachaufzeichnungen aus Callcentern oder auch Notizen von Vertriebsmitarbeitern. Um seine Kunden besser zu verstehen und gezielter auf Beschwerden und Abwanderungstendenzen einzugehen, nutzt die amerikanische Versicherungsgesellschaft United Healthcare in diesem Kontext eine Spracherkennungssoftware und wandelt Sprachaufnahmen aus Telefonaten in Textdaten um. Diese Daten sind

besonders wertvoll, da gerade über diesen Kanal Kundenemotionen viel genauer identifiziert werden können. Die Gesprächsdaten werden dann auf Begriffe hin untersucht, die auf eine besondere Unzufriedenheit hinweisen. Somit kann das Versicherungsunternehmen präventiv auf diese Kunden einwirken und dadurch einem potentiellen Verlust vorbeugen. Es zeigt sich, dass durch die richtige Formulierung von Fragestellungen und die gezielte Nutzung von neuen Datenquellen im Rahmen von Big-Data-Technologien früher eher intuitive und ungenaue Entscheidungsvorlagen heute deutlich ausgefeilter und granularer bereitgestellt werden können (vgl. [Dav14], pp. 64-67).

#### **2.4.4 Entwicklung neuer Produkte auf Basis von Big Data**

Eine besonders herausfordernde Strategie, die man mit einem Big-Data-Projekt verfolgen kann, ist die Entwicklung innovativer Produkte und Dienstleistungen auf Basis von Big Data. Hierbei ist vor allem ein hohes Maß an Flexibilität sowie die Bereitschaft, Ideen auch wieder zu verwerfen gefordert. Es geht darum Marktbedürfnisse zu erkennen und Big-Data-Lösungen zu entwickeln. Dies wiederum erfordert intensive Zusammenarbeit zwischen Marketingabteilung und Produktentwicklung, wobei von allen Projektteilnehmern ein hohes Maß an Big-Data-Know-how gefordert wird (vgl. [Dav14], pp. 62-63). Um als Organisation nachhaltig datengetriebener zu operieren, bedarf es auch auf Managementebene an ausgeprägtem Wissen über Big-Data-Technologien und -Potentiale für die betreffende Branche. Langfristig kann gerade diese Strategie zu schwer kopierbaren Wettbewerbsvorteilen führen (vgl. [SKM<sup>+</sup>12], p. 62).

Das Beratungsunternehmen Boston Consulting Group (BCG) hat 2014 untersucht, wie traditionelle Geschäftsmodelle mit Hilfe von Big Data verändert, angereichert oder neu entwickelt werden können. Laut BCG werden sich Unternehmen langfristig branchenübergreifend damit beschäftigen müssen Big Data als ein Geschäft anzusehen, welches einen Mehrwert für Kunden und das Unternehmen selbst liefert. Sie identifizieren in ihrer Studie verschiedene Geschäftsmodelle, die auf Big Data basieren. Dabei können einige nur von Organisationen genutzt werden, bei denen Datenanalyse die Kernkompetenz darstellt. Andere wiederum eignen sich auch für Organisationen aus tendenziell datenfernen Branchen.

Ein Beispiel für ein solches Geschäftsmodell liefert die National Bank of Australia. Laut BCG gründete diese schon 2008 ein Joint Venture mit dem auf Datenanalyse spezialisierten Unternehmen Quantum, um viele Millionen Transaktionsdaten auszuwerten, aufzubereiten und so die Erkenntnisse anonymisiert an Dritte weiterzuverkaufen. In diesem Beispiel liefert die National Bank of Australia große Mengen an wertvollen bestehenden Daten und Quantum die Analysefähigkeit. Beide Unternehmen konzentrieren sich auf ihre Kernkompetenzen und haben trotzdem partnerschaftlich einen Mehrwert durch Big Data.



BCG illustriert in einem weiteren Beispiel, wie ein Unternehmen mit einer eigenen Geschäftseinheit mit Hilfe von Big Data Produkte entwickelt. Das Handelsunternehmen Tesco analysiert mit Hilfe seines Big-Data-Tochterunternehmens Dunnhumby Millionen von Kundentransaktionsdaten, um so Erkenntnisse über das Kaufverhalten auf Postleitzahlen-ebene zu gewinnen. Diese Einblicke verkauft Dunnhumby an Hersteller wie Unilever, Nestlé oder Heinz weiter (vgl. [PSC<sup>+</sup>14]).

## 2.5 Herausforderungen im Umgang mit Big Data

Mit der Nutzung von Big Data in einer Organisation gehen unweigerlich auch verschiedenartige Herausforderungen einher. Einige dieser Herausforderungen können auch Einfluss auf die Einführung eines Big-Data-Management-Systems haben. Genau diese Aufgabenstellungen werden in diesem Abschnitt thematisiert. Eine besondere Aufgabe stellt die Gestaltung der Data Governance und die Berücksichtigung von Sicherheitsaspekten dar. Hier geht es darum, Prozesse und Verantwortlichkeit festzulegen und Compliance-Richtlinien zu definieren. Eine weitere Herausforderung betrifft die personelle Komponente einer Big-Data-Initiative. Sogenannte Datenwissenschaftler (Data Scientists) stellen eine neu aufkommende Berufsgruppe dar und sind oft essentieller Bestandteil von Big-Data-Projekten. Abschließend geht es um die Problemstellung, inwieweit der Erfolg einer Big-Data-Initiative quantitativ bewertet werden kann.

### 2.5.1 Data Governance und Datensicherheit

Beim Aufbau von Big-Data-Management-Systemen in Organisationen treten neuartige, unbekannte Herausforderungen in Bezug auf Prozesse, Verantwortlichkeiten und Compliance auf. Deshalb ist es wichtig, einen strukturierten Ansatz bei der Implementierung einer Data Governance zu verfolgen (vgl. [Bit12], p. 16). Dazu ist es notwendig ein Verständnis über die Begrifflichkeit zu gewinnen. Das Data Governance Institute (DGI) definiert Data Governance wie folgt:

“Data Governance is a system of decision rights and accountabilities for information-related processes, executed according to agreed-upon models which describe who can take what actions with what information, and when, under what circumstances, using what methods [Tho14].”

Diese detaillierte Definition beleuchtet die innere Komplexität einer vollständigen Data Governance. Bezieht man die Definition auf das Big-Data-Zeitalter, so wird klar, dass zur konsequenten Sicherung von Datenqualität und Datensicherheit ein umfassendes System von Regeln und Verantwortlichkeiten benötigt wird. In einem Artikel der Computerwoche in Zusammenarbeit mit IBM wird aufgezeigt, dass Big Data nicht mit einem klassischen Daten-Management auskommt. Vielmehr müsse dieses erweitert werden, um eine angemessene Data Governance mit großen und oft unstrukturierten Datenmengen zu gewährleisten. So

zitiert die Computerwoche dort ein aktuelles Beispiel eines Unternehmens, welches sein Data-Warehouse mit Hilfe von Hadoop-Clustern entlastet hat. Da innerhalb dieser Cluster kein Daten-Lebenszyklus integriert war, konnte dem Unternehmen noch ein viele Jahre in der Vergangenheit liegendes Vergehen nachgewiesen werden, was zu hohen Strafzahlungen führte. Eine Data-Warehouse-Archivierung hätte diese E-Mail-Konversationen im Rahmen des Lebenszyklus-Managements längst gelöscht. Es ist diesem Fall dementsprechend essentiell, die für Data-Warehouses geltenden Archivierungs-, Sicherungs- und Zugriffsregeln auf das Hadoop-Cluster auszuweiten. Neben den für die Speicherung und Archivierung geltenden Regeln bedarf es auch Standards für die Analyse und Verarbeitung von Big Data. Beispielsweise nennt der Autor des Artikels in diesem Kontext die Anonymisierung personenbezogener Daten mit Hilfe von Aggregationsalgorithmen. Eine gänzlich neue Herausforderung sei die Data Governance für Datenströme in Echtzeit, wie z.B. Live-Videoaufnahmen. Hier spiele insbesondere die Zugriffsberechtigung eine exponierte Rolle. Zusammenfassend müsse bei der Implementation von Big-Data-Management-Systemen in einer Organisation darauf geachtet werden, dass sämtliche genannte Aspekte nicht isoliert betrachtet werden. Bestehende Data Governance-Systeme müssen um diese Problematiken ergänzt werden, sodass sie auch für Big Data greifen (vgl. [Man14]).

Die wohl prominenteste Herausforderung für die Etablierung von Big Data scheint die Datensicherheit sowohl für Organisationen als auch für Individuen zu sein. Sie ist in der Regel unmittelbar mit einer funktionierenden Data Governance verbunden und wird deshalb in diesem Zuge mit aufgeführt. Das Fraunhofer IAIS hat in seiner Studie zum Innovationspotential von Big Data neben zahlreichen anderen Aspekten auch Barrieren für die Durchsetzung untersucht. Dabei stechen Datenschutz und Datensicherheit als wesentliche Hindernisse deutlich heraus (vgl. [SKM<sup>+</sup>12], p. 48).

Das Handelsblatt Research Institute hat in einer Studie den Anspruch der Deutschen an den Datenschutz mit dem konkreten Umgang eben dieser mit persönlicher Daten untersucht. Der Anspruch der Deutschen an den Datenschutz ist traditionell aber auch per Gesetz sehr hoch. So wünschen sich in einer Umfrage 62 Prozent der befragten Personen einen besseren Datenschutz im Big-Data-Zeitalter, obwohl Deutschland im internationalen Vergleich bereits deutlich strengere Datenschutzregelungen besitzt. Das Misstrauen gegenüber Unternehmen ist mit 87 Prozent außerordentlich hoch. Parallel zur Umfrage hat das Handelsblatt Research Institute ein Experiment zur Untersuchung der Bereitschaft zur Herausgabe persönlicher Daten durchgeführt. Das Ergebnis steht dem Anspruch nach hohem Datenschutz in starkem Kontrast. Es werden bereits für den Erhalt verschwindend kleiner Vorteile tiefgreifende persönliche Daten preisgegeben. Derartiges Verhalten bezeichnet man in der Literatur als

Privacy Paradox (vgl. [HL13], p. 58). Sehr treffend hat der Internet-Unternehmer Austin Hill im Jahr 2000 das Privacy Paradox in einem Artikel im Wall Street Journal beschrieben:

“Ask 100 people if they care about privacy and 85 will say yes. Ask those same 100 people if they'll give you a DNA sample just to get a free Big Mac, and 85 will say yes [Web00].”

Die Möglichkeiten, die Big Data bietet, stellen aufgrund ihrer rasanten Entwicklungsgeschwindigkeit eine enorme Herausforderung für die Gesetzgebung dar. Schon innerhalb kürzester Zeit kann auf Basis persönlicher Daten ein völlig neues Geschäftsmodell entstehen (vgl. [HL13], p. 60). Mayer-Schönberger und Cukier fordern deshalb einen Paradigmenwechsel in Sachen Datensicherheit. Der bestehende Ansatz, dem Individuum die Kontrolle zu überlassen, ob, wie und von wem seine persönlichen Informationen verarbeitet werden, sei im Internetzeitalter zu einer „notice and consent“-Politik geraten, die der Nutzer kaum überblicken könne. Durch Big-Data-Technologien bestehe der Wert dieser Daten oft durch die mehrfache und vielfältige Verwendung, oft auch zu Zwecken, die zum Zeitpunkt der Datensammlung noch nicht bekannt waren. Mayer-Schönberger und Cukier schlagen deshalb vor, dass der Verwender der Daten für die Nutzung verantwortlich gemacht wird. Die Nutzer von Big-Data-Technologien dürfen die persönlichen Daten abhängig von den Auswirkungen auf das Individuum auch mehrfach verwenden, sofern sie den zukünftigen Regeln der Datenschutzbehörden entsprechen. Würde eine solche Regel von der Organisation gebrochen, so müsse sie mit Sanktionen rechnen. Insgesamt fördere dieser Paradigmenwechsel zum einen die Kreativität durch die Möglichkeit der Mehrfachnutzung von Daten und zum anderen entlaste er das Individuum von der Verantwortung für die Daten (vgl. [MC13], p. 173). Das Handelsblatt Research Institute hält diesen Ansatz für eine Lösung, die das Privacy Paradox auflösen könnte. Dort stellt man fest, dass „wenn Menschen wissen, dass sie von Unternehmen [...] auch noch Jahre nach einer einmal gegebenen Einverständniserklärung die Löschung ihrer Daten verlangen und sie bei Missbrauch zur Rechenschaft ziehen können, dann dürfte das helfen, die Schere zwischen den geäußerten Einstellungen und dem eigenen Verhalten zu schließen [...] ([HL13], p. 61).“

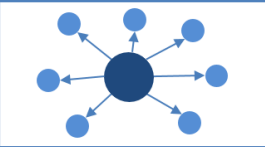
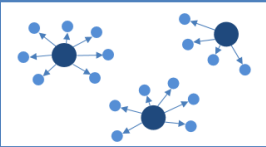
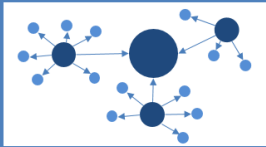
Für Unternehmen ergibt sich hieraus eine unübersichtliche Situation. Viele Anwendungsfälle belegen die enormen Potentiale, die sich durch Big Data ergeben. Nichtsdestotrotz belegen Studien die scheinbar unumstößliche öffentliche Meinung, dass die Nutzung persönlicher Informationen zur Verbesserung von Produkten, Dienstleistungen oder zur Prozesseffizienzsteigerung tendenziell kritisch gesehen wird. Dem gegenüber steht wiederum das tatsächliche Verhalten der Individuen, die schon für sehr geringe Anreize bereit sind Daten preiszugeben. Schlussendlich muss die öffentliche Debatte entscheiden, wie eine Neuordnung des Datenschutzes gesetzlich geregelt werden kann und so Unternehmen eine sichere Handlungsbasis für Tätigkeiten im Themenfeld Big Data gibt. Unternehmen können

zu dieser Debatte beitragen, indem sie glaubhaft und detailliert Transparenz bei der Nutzung persönlicher Daten zeigen und darstellen, welche öffentlichen Vorteile durch die vielfache und vielfältige Nutzung der Daten entstehen (vgl. [HL13], pp. 62-63).

### **2.5.2 Personelle Ressourcen und organisatorische Einbindung**

Viele bestehende Denkweisen verändern sich durch den Eintritt in das Big-Data-Zeitalter. Organisationen, die einen Nutzen aus den neuen Technologien ziehen wollen, müssen hierbei insbesondere die Ressource Mensch berücksichtigen. Neue hochqualifizierte Mitarbeiter, sogenannte Datenwissenschaftler, die erweiterte Analysefähigkeiten in Bezug auf Big Data besitzen, stellen eine unverzichtbare Komponente von Big-Data-Projekten dar. Diese Ressource entscheidet oft über die erfolgreiche Implementierung eines Big-Data-Management-Systems (vgl. [Dav14], p. 84). Das IT-Unternehmen IBM sieht in einem Datenwissenschaftler die evolutionäre Weiterentwicklung des klassischen Datenanalysten. Neben dem formell ähnlichen Fähigkeitsprofil in den Bereichen Informatik, Modellierung, Statistik und Mathematik zeichnet den Datenwissenschaftler ein besonders ausgeprägter Scharfsinn in Bezug auf geschäftsrelevante Datenzusammenhänge und Geschäftsmodelle aus. Zudem versuchen sie verschiedene Datenquellen nutzenbringend zu verknüpfen, um so wertvolle Erkenntnisse im Sinne des Geschäftsmodells zu gewinnen. Dies gelingt ihnen, indem sie bestehende Strukturen in Frage stellen und im Rahmen von was-wäre-wenn-Analysen Datenlandschaften erkunden. Der besondere Wert besteht weiterhin darin, dass ein Datenwissenschaftler in der Lage ist die ermittelten Erkenntnisse in Schlussfolgerungen und Entscheidungsvorlagen für das Management aufzubereiten und die Organisation somit datengetriebener agieren kann (vgl. [Ibm14]). Die individuellen Fähigkeiten in Bezug auf Big-Data-Technologien variieren laut Davenport jedoch von Person zu Person. Dies führe dazu, dass die optimale Lösung immer ein ausgeglichenes Team von Datenwissenschaftlern sei. Diese Mitarbeiter müssen nicht immer neu eingestellt werden. Durch Training und Förderung bestehender Datenanalysten könne ein Team zusammengestellt werden, welches die Anforderungen der jeweiligen Big-Data-Strategie vollends erfülle (vgl. [Dav14], p. 97).

Interdisziplinarität ist ein Schlüsselfaktor, um die Potentiale von Big Data vollständig auszunutzen. Minelli et al. stellen fest, dass es insbesondere in großen Organisationen oft Vorbehalte bezüglich übergreifender Kollaboration gibt. Zudem verlasse man sich zu sehr auf firmeneigene Informationen. Um diese Widerstände zu überwinden, schlagen Minelli et al. drei verschiedene Organisationsstrukturen für die Anordnung eines sogenannten Analytikbereichs vor (siehe Abb. 7) (vgl. [MCD13], p. 146).

	Zentralisiert	Dezentralisiert	Föderal
Inhalt	<ul style="list-style-type: none"> <li>➤ Zentrale Shared-Services-Organisation, die mehrere Bereiche beliefert</li> </ul> 	<ul style="list-style-type: none"> <li>➤ Jeder Bereich mit eigener Analytik-Abteilung</li> </ul> 	<ul style="list-style-type: none"> <li>➤ Zentrale Koordination mit lokaler Ausführung</li> </ul> 
Eigentümer	<ul style="list-style-type: none"> <li>➤ C-Level Entscheider</li> </ul>	<ul style="list-style-type: none"> <li>➤ Bereichsleiter / Funktionsleiter</li> </ul>	<ul style="list-style-type: none"> <li>➤ C-Level Entscheider mit Koordinatoren in jeweiligen Bereichen</li> </ul>
Pro	<ul style="list-style-type: none"> <li>➤ Skaleneffekte in Bezug auf Infrastruktur und Prozesse</li> <li>➤ Wenig Widerstände bei der Kommunikation der Big-Data-Vision</li> <li>➤ Cross-funktionale Kollaboration</li> </ul>	<ul style="list-style-type: none"> <li>➤ Bereichsunabhängigkeit</li> <li>➤ Risiko der zu schwachen Verbreitung in Bereichen, die stärker fokussiert werden müssten</li> </ul>	<ul style="list-style-type: none"> <li>➤ Agilität und Flexibilität bleibt erhalten</li> <li>➤ Zugang zu den notwendigen Fachspezialisten gegeben</li> <li>➤ Autonomie der funktionalen Bereiche</li> </ul>
Contra	<ul style="list-style-type: none"> <li>➤ Cost-Center-Sicht</li> <li>➤ Abstimmung mit Bereichen nicht gegeben</li> <li>➤ Langsam und unflexibel</li> </ul>	<ul style="list-style-type: none"> <li>➤ Interdisziplinarität wird nicht gefördert</li> <li>➤ Kurzfristige Ergebnisse im Fokus</li> <li>➤ Verminderte Kontrollierbarkeit und Ausrichtbarkeit an Unternehmensvision</li> </ul>	<ul style="list-style-type: none"> <li>➤ Koordination und Planung oft aufwändig</li> </ul>

**Abbildung 7: Modelle zur organisatorischen Einbindung eines Analytikbereichs (vgl. [MCD13], p. 148)**

Im Falle eines zentral angeordneten Analytikbereichs beliefert dieser alle anderen Bereiche mit Analyseservices. Dies hat den Vorteil, dass Skaleneffekte genutzt werden können und von einer integrierten Dateninfrastruktur profitiert wird. Gleichzeitig kann aber nicht die Flexibilität gewährleistet werden, um Big Data in den einzelnen Bereichen zu einem integralen Bestandteil zu machen. Bei dezentral organisierten Analytikbereichen in jedem Geschäftsbereich besteht laut Minelli et al. die Gefahr, dass redundante oder konfliktäre Ziele mit Big Data verfolgt werden. Nichtsdestotrotz ist dies ein Ansatz, der die Belange aller Bereiche berücksichtigt und das gesamte Unternehmen schnell datengetriebener agieren lässt. Das föderale Modell versucht die Vorteile der beiden vorangegangenen Organisationsstrukturen zu vereinen, indem es den Bereichen Eigenverantwortung bezüglich Big Data überlässt, jedoch über einen zentralen Ausschuss alle Tätigkeiten koordiniert (vgl. [MCD13], p. 146).

Davenport stellt zusammenfassend fest, dass Big-Data-Fertigkeiten in großen Organisationen in den meisten Fällen in bestehende Analytikbereiche integriert werden. Er zieht vor diesem Hintergrund den Vergleich zu Technologie- und Datenarchitekturen in Organisationen. Auch diese wachsen und verschmelzen mit der bestehenden Organisationsstruktur. Trotzdem ist im Rahmen eines Big-Data-Projekts abhängig vom Ziel

und der Strategie oft eine Ausrichtung an eine funktionale Abteilung, wie beispielsweise das Marketing, notwendig (vgl. [Dav14], p. 177).

### 2.5.3 Erfolgsmessung bei Big-Data-Projekten

Den Return on Investment (ROI) von Big-Data-Initiativen in einer Organisation zu bestimmen, ist eine viel diskutierte Herausforderung. Die Möglichkeiten zur Ermittlung eines ROI hängen laut Davenport maßgeblich davon ab, welche Big-Data-Strategie die Organisation verfolgt und wie Big Data dort zur Wertschöpfung beiträgt. Auf der einen Seite nutzen Organisationen Big Data, um neue Geschäftsmöglichkeiten zu generieren. Auf der anderen Seite werden mit Hilfe von Big-Data-Technologien oft exakt die Analysen durchgeführt, die in der Vergangenheit auch mit einem klassischen Data-Warehouse-Ansatz bearbeitet wurden (vgl. [Dav14], p. 183). Im Falle einer derartigen Aufrüstung sind die Analysen jedoch kostengünstiger, schneller oder genauer, was die Berechenbarkeit des ROI deutlich vereinfacht (vgl. [For14]).

In einer viel beachteten, auf Wikibon veröffentlichten Fallstudie, hat David Floyer verschiedene Analyseumgebungen auf ihre finanzielle Performance hin untersucht. Eine Data-Warehouse-Umgebung wird dort einer auf Hadoop basierenden Big-Data-Lösung gegenübergestellt, um zu ermitteln, welche Investition sich schneller amortisiert und mehr Cash-Flow generiert. In genannter Fallstudie wird die Big-Data-Umgebung als Investition favorisiert, da der Breakeven-Punkt gegenüber dem Data-Warehouse-Ansatz mit 24 Monaten bereits nach drei Monaten überschritten ist. Dieser massive Vorsprung ist jedoch nicht dadurch begründet, dass traditionelle Data-Warehouses obsolet sind. Vielmehr hat die Big-Data-Lösung besser zu den Anforderungen des Projekts gepasst. Als Fazit stellt Floyer fest, dass beide Lösungen in der analytischen Datenarchitektur einer Organisation eine spezielle Rolle spielen und sich nicht vollständig ersetzen können (vgl. [Flo13]).

Betrachtet man Big Data als Technologie, um neue Geschäftsmöglichkeiten zu generieren, so wird die Berechnung der Rentabilität deutlich komplexer und die Ergebnisse sind oft vage. Der Marketing-Manager Samer Forzley stellt fest, dass die Methoden zur Berechnung des ROI von IT-Projekten nicht geeignet sind, um den ROI einer Big-Data-Initiative zu bestimmen. Er nennt als Beispiel den Umzug der IT von einer klassischen Server-Lösung in eine Cloud. Mit der Kapitalwertmethode könne hier über die Projektlaufzeit die Einsparung sehr einfach berechnet und eine Projektinitiierung entsprechend davon abhängig gemacht werden. Die Einsparungen bzw. der Wert, den Big Data generiert, entstammt hingegen neuen Erkenntnissen aus Daten, die zu Umsatz- und Effizienzsteigerungen führen können. Das Ausmaß dieser Verbesserungen ist vor Projektbeginn jedoch nicht abschätzbar, sodass auch kein Kapitalwert berechnet werden kann. Forzley stellt diesbezüglich eine Analogie zu Branchen, wie der Pharmaindustrie oder der Bergbauindustrie her, in denen vor Projekt-

beginn ebenfalls keine Abschätzung darüber gemacht werden kann, welchen Wert das Projekt letztendlich liefert. Diese Branchen müssen sich auf die Realloptionsanalyse verlassen. Diese Methode versucht zu berücksichtigen, dass sich das Risiko bzw. die Ergebnisprognose während eines Projekts verändern kann. Im Falle von Big Data müsse also regelmäßig überprüft werden, ob die Erkenntnisse den geplanten Projekterfolg liefern können, oder ob die Ausrichtung der Big-Data-Initiative geändert werden muss (vgl. [For14]).

In einer Studie von IDG Research und Kapow Software konstatiert man darüber hinaus, dass bisherige Big-Data-Projekte nur in den seltensten Fällen den avisierten ROI erreicht haben. Dies läge in erster Linie daran, dass die Projekte oft länger dauern als erwartet und ein hohes Maß an externer Beratung in Anspruch nehmen. Deshalb solle man als Erfolgskennzahl neben dem ROI auch den Zeitaufwand für die Erarbeitung nutzbarer Erkenntnisse in Betracht ziehen. Vor diesem Hintergrund mache es Sinn, zunächst nur einen geringeren Projektumfang zu wählen und Big-Data als Basistechnologie zu betrachten. Ist der Nachweis über die Machbarkeit erbracht und das Projekt profitabel, so könne man die Initiative ausweiten (vgl. [Idc13b]).

Einen weiteren Ansatz liefert die BITKOM in ihrem Leitfaden zum Management von Big-Data-Projekten aus dem Jahr 2013. Sie beziehen sich auf ein Whitepaper der Firmen HP und Vertica, welche zur Erfolgsbewertung von Big-Data-Initiativen den Return on Information definiert haben. Darin behaupten sie, dass Organisationen, die in der Lage sind einen hohen Return on Information zu realisieren, einen maßgeblichen Wettbewerbsvorteil gegenüber der Konkurrenz herstellen können. Die Kennzahl berechnet sich anders als der klassische ROI, sodass hier erstmals messbare Werte aus Big-Data-Projekten monetarisiert werden können. In die Berechnung werden das Volumen der analysierbaren Daten, die Anzahl der zugriffsberechtigten Benutzer, die Analysetiefe sowie die Zeit zwischen Datenerzeugung und Analyse mit einbezogen. Diese Werte werden mit den Kosten für die Anschaffung und den Betrieb der Lösung in Relation gesetzt. Im Ergebnis ist damit eine verbesserte Analysierbarkeit des Erfolgs von Big-Data-Projekten möglich. (vgl. [Bit13], pp. 37-38).

Die zahlreichen, heterogenen Ansätze zur Erfolgsmessung bzw. Profitabilitätsbestimmung von Big-Data-Projekten, die in diesem Abschnitt vorgestellt wurden, zeigen auf, dass es sich dabei in der Tat um eine viel diskutierte Herausforderung von Big Data handelt und es bislang keine einheitliche Antwort dafür gibt.

### 3 Vorgehensmodelle im IT-Projektmanagement

In diesem Kapitel werden Grundlagen zum IT-Projektmanagement im übergeordneten Sinne sowie klassische und agile Vorgehensmodelle als strukturierte Ansätze des IT-Projektmanagements vorgestellt. Dabei wird zum einen das Wasserfallmodell als ein den klassischen Vorgehensmodellen zugehöriger Ansatz näher betrachtet und zum anderen Scrum als ein typisches agiles Vorgehensmodell untersucht. Beide Betrachtungen werden vor dem Hintergrund der Einführung eines Big-Data-Management-Systems durchgeführt. Weiterhin werden charakteristische Vergleichskriterien für das IT-Projektmanagement aufgestellt, welche eine quantitative Untersuchung mittels der Nutzwertanalyse erlauben.

#### 3.1 Definition und Herkunft von Vorgehensmodellen

Um sich dem Begriff des Vorgehensmodells fundiert zu nähern, bedarf es zunächst einer Definition des Projektbegriffs im Kontext von IT und Softwareentwicklung. Trepper definiert den Begriff in Anlehnung an Wieczorrek und Mertens wie folgt:

„Ein Projekt ist gekennzeichnet durch eine exakte Definition der zu erledigenden Aufgaben. Es ist eindeutig abgrenzbar von operativen Unternehmensaufgaben, hat einen eindeutigen Start- und Endtermin, ist in seiner Form einzigartig und birgt ein hohes Risiko. Des Weiteren herrscht Konkurrenz um knappe Ressourcen und es ist ein kritischer Faktor für den Erfolg des Unternehmens ([Tre12], p. 25 in Anlehnung an [WM07], pp. 7-9).“

Wieczorrek und Mertens stellen die Verknüpfung zu IT-Projekten insofern her, als dass dort im Gegensatz zu nicht IT-bezogenen Projekten Mitarbeiterressourcen ab einem bestimmten Projektstatus nicht mehr ausgetauscht werden sollten. Zudem sei die Anwendung von Vorgehensmodellen sinnvoll, da derartige Projekte sich häufig in ähnliche, immer wieder eintretende Abschnitte aufteilen lassen (vgl. [WM07], p. 9). Ein Vorgehensmodell ist somit der Versuch die Entwicklungsstufen bei IT-Projekten zu strukturieren und zu systematisieren. Ruf und Fittkau liefern dazu die folgende Definition:

„Eine planvolle, systematische Beschreibung der einzelnen Prozessschritte wird als Vorgehensmodell bezeichnet. Es handelt sich um eine Folge von aufeinander aufbauenden und genau definierten Prozessschritten. In Vorgehensmodellen wird der organisatorische Ablauf für die Entwicklung eines IT-Systems beschrieben. Im Vordergrund steht die Frage: „Was ist wann zu tun?“ ([RF08], p. 25)“

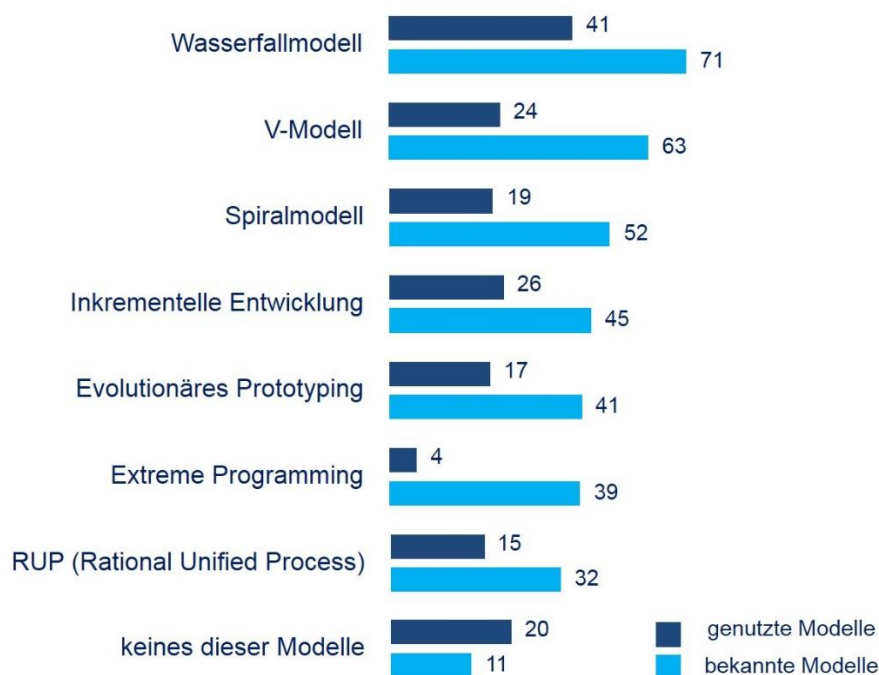
Zusammenfassend sind Vorgehensmodelle demnach eine konkrete Abfolge von Schritten im Prozess der Softwareentwicklung und helfen bei der Projektplanung und -abwicklung. Die strukturierte Aufteilung des Projekts in Prozessschritte dient als Orientierungshilfe für die Projektbeteiligten und liefert genaue Regeln, mit welchem Fertigstellungsgrad und welcher



Qualität die entsprechenden Produkte zu welchem Zeitpunkt vorliegen sollen (vgl. [SDB+10], p. 46).

Des Weiteren ist es laut Ruf und Fittkau aufgrund variierender IT-Produkte, unterschiedlicher Qualitätsanforderungen und differierender Projektgrößen sehr unwahrscheinlich, dass sich ein einheitliches Vorgehensmodell für alle IT-Projekte herauskristallisiert. Sie verweisen jedoch darauf, dass die IT-Projekte innerhalb eines Unternehmens häufig mit ähnlichen Anforderungen und Methoden realisiert werden und somit ein unternehmenseinheitliches Vorgehensmodell entwickelt werden kann. Dieses werde dann projektspezifisch modifiziert, um den Anforderungen gerecht zu werden (vgl. [RF08], p. 28).

Die Anwendung und die Bekanntheit von Vorgehensmodellen hat die Computer Zeitung im Jahr 2005 im Rahmen einer Umfrage in deutschen Unternehmen untersucht (Abb. 8). Aus den Ergebnissen der Studie schließt Kleuker, dass es kein ultimatives, überlegenes Vorgehensmodell gibt. Vielmehr zeigt sich, entgegen der These von Ruf und Fittkau, dass es nicht untypisch ist, verschiedene Vorgehensmodelle innerhalb eines Unternehmens anzuwenden. Dies wird dadurch belegt, dass die Summation der Prozentangaben mehr als 100 Prozent ergibt. (vgl. [Kle13], pp. 25-26).



**Abbildung 8: Bekanntheitsgrad und Anwendung von Vorgehensmodellen**  
 ([Kle13], p. 26 zitiert aus Computer Zeitung, Nr. 46, 14.11.2005)

Die Ergebnisse dieser Umfrage sind im Jahr 2015 nur noch bedingt repräsentativ, da sich der Anteil der Nutzer von agilen Vorgehensmodellen deutlich erhöht hat. Zu beachten ist beispielsweise, dass Scrum, als ein sehr populäres agiles Vorgehensmodell, in obiger Auswertung nicht erscheint. Dort wird lediglich Extreme Programming als einziger Vertreter agiler Methoden aufgeführt. Eine Softwaretest-Umfrage aus dem Jahr 2011 belegt diese

Unterrepräsentanz von agilen Methoden. Nach wie vor entwickelt die Mehrheit der Teilnehmer dieser aktuelleren Umfrage mit klassischen Vorgehensmodellen und dort insbesondere mit dem V-Modell. Dessen ungeachtet nutzen bereits ein Viertel aller Befragten agile Vorgehensmodelle. Scrum stellt in dieser Teilmenge mit 57 Prozent den klaren Favoriten dar (vgl. [HSV<sup>+</sup>12], pp. 7-9).

In dieser Ausarbeitung wird zum einen das Wasserfallmodell, welches bereits in den 1970er Jahren von Royce veröffentlicht sowie später von Boehm verfeinert wurde, als Vertreter der klassischen Vorgehensmodelle vorgestellt. Zum anderen wird stellvertretend für die flexibleren sogenannten agilen Vorgehensmodelle Scrum detailliert erläutert. Agile Vorgehensmodelle versuchen die Nachteile klassischer Vorgehensmodelle, wie z.B. mangelnde Flexibilität und Kundennähe, zu beheben und gehen auf das Agile Manifest aus dem Jahr 2001 zurück (vgl. [SDB<sup>+</sup>10], p. 62).

### **3.2 Herausforderungen und Erfolgsfaktoren im IT-Projektmanagement**

Aus zahlreichen Studien geht hervor, dass IT-Projekte häufig nicht zu einem erfolgreichen Ende geführt werden. Die Standish Group veröffentlicht zu diesem Thema seit 1994 den sogenannten Chaos Report über den Erfolg von IT-Projekten. Darin kann ein Projekt in drei verschiedene Ergebnisklassen kategorisiert werden: Projekt erfolgreich abgeschlossen, eingeschränkter Projekterfolg sowie Projekt nicht erfolgreich und abgebrochen. Im Jahr 1994 wurden der Studie zufolge 16,2 Prozent der Projekte erfolgreich, 52,7 Prozent mit Einschränkung, also z.B. mit Budget- oder Zeitüberschreitungen, sowie 31,1 Prozent nicht erfolgreich abgeschlossen. Die Schlüsselfaktoren für den Projekterfolg sieht die Standish Group insbesondere bei der Einbindung des Kunden, der Unterstützung durch die Geschäftsführung sowie in einer eindeutigen Anforderungsbeschreibung. Im Umkehrschluss führen laut Standish Group mangelnde Kundenintegration, unvollständige Anforderungen und unzureichende Ressourcen häufig zu Projektabbrüchen (vgl. [Sta94], pp. 4-10). Die jährlich erscheinenden Untersuchungen der Standish Group werden zwar vielfach in der Fachliteratur zitiert, jedoch müssen sie sich auch wissenschaftlicher Kritik aussetzen. Hauptkritikpunkte sind dabei die nicht hinreichend granulare Ergebniskategorisierung sowie die wissenschaftlich kaum replizierbaren und ausgesprochen pessimistischen Ergebnisse (vgl. [EV10]).

Es ist ein Faktum, dass IT-Projekte im Zeitalter von Big Data weiter an Bedeutung gewinnen werden. Ruf und Fittkau prognostizieren darüber hinaus, dass der IT-Projekterfolg kontinuierlich wachsenden Einfluss auf die langfristige Sicherung des Unternehmenserfolgs haben wird. Sie stellen weiterhin fest, dass sich bei IT-Projekten wiederkehrende Probleme und Herausforderungen identifizieren lassen (Tab. 2), die in den spezifischen Risiken von IT-Projekten begründet sind. Um diesen Problemen souverän zu begegnen, macht es Ruf und

Fittkau zufolge Sinn, sich im Rahmen eines Risikomanagements einen Überblick über die Risiken und deren Auswirkungen sowie Eintrittswahrscheinlichkeiten zu verschaffen (vgl. [RF08], pp. 3-7).

Problem / Herausforderung	Inhalt
Kosten	Die für das IT-Projekt geplanten Kosten werden häufig überschritten
Termin	Die beabsichtigte Projektlaufzeit wird ungeplant deutlich überschritten
Komplexität	Wesentliche Faktoren der Komplexität sind: <ul style="list-style-type: none"> <li>▪ Integration des Projekts in die Unternehmensstrategie</li> <li>▪ Einbeziehung von betrieblichen Anforderungen</li> <li>▪ Menschliches Potential und Kompetenz</li> <li>▪ Beherrschung der technischen Herausforderung</li> </ul>
Qualität	Bei IT-Projekten gelingt es häufig nicht, die geplante Qualität einzuhalten (z.B. Performance, Bedienbarkeit, Änderbarkeit, Robustheit, Verständlichkeit, etc.)
Zielerreichung	Die entwickelten IT-Systeme sind teilweise nicht in der Lage die Ausgangsproblematik zu beheben
Interne Konkurrenz	Ständiger Kampf eines Projekts um die knappen Ressourcen des Unternehmens
Wettbewerb	Auch die Konkurrenz führt IT-Projekte durch und verschafft sich somit häufig Wettbewerbsvorteile, indem er die Aufgaben besser löst

**Tabelle 2: Typische Probleme und Herausforderungen bei IT-Projekten (vgl. [RF08], pp. 4-5)**

Die Risiken, die zu den oben genannten Problemen und Herausforderungen führen, lassen sich laut Tiemeyer in fünf Kategorien gruppieren. Als erste Risikogruppe führt er die unzureichende Projektdefinition auf, welche sich in ungenauen Zielsetzungen oder zu oberflächlichen Anforderungsspezifikationen äußert. Weiterhin stellt er den falschen oder fehlenden Einsatz von Planungs- und Steuerungsinstrumenten, was unter anderem auch die geeignete Nutzung von Vorgehensmodellen umfasst, als zweite Risikogruppe dar. Die dritte Kategorie betrifft die am Projekt beteiligten Personen und deren Eignung in Bezug auf die Aufgabenstellung. Die operative Projektplanung und die Durchführung repräsentieren die vierte und fünfte Risikogruppe. Zusammenfassend weist Tiemeyer darauf hin, dass mit wachsender Größe und Komplexität des IT-Projekts sowie zunehmender Anzahl von Schnittstellen zu anderen Projekten das Risiko eines Projektmissersfolgs überproportional zunimmt. Analog zu Ruf und Fittkau verweist er auf ein vorbeugendes Risikomanagement sowie die transparente Kommunikation von Projektleitung, Projektteam und Stakeholdern, damit die Misserfolgswahrscheinlichkeit gesenkt werden kann (vgl. [Tie14], pp. 4-7).

### 3.3 Klassische Vorgehensmodelle

In diesem Abschnitt werden die Eigenschaften klassischer Vorgehensmodelle im IT-Projektmanagement vorgestellt. Als ein typischer Vertreter dieser Gruppe wird das Wasserfallmodell kritisch untersucht.

#### 3.3.1 Charakteristika

In der Neu- oder Weiterentwicklung von Softwareprodukten sind nach Kleuker häufig wiederkehrende Phasen anzutreffen (Abb. 9). Am Anfang eines Projektes steht dabei immer eine Anforderungsanalyse, in welcher der Wunsch des Kunden in eine Produktspezifikation übersetzt wird. Im Anschluss werden im Rahmen des Grobdesigns die Produktspezifikationen in ein Modell überführt sowie die zugrunde liegende Software-Architektur definiert. Das Feindesign bereitet die Implementierung auf granularer Ebene vor und definiert die innere Struktur der Software. Bei der Implementierung beginnt die Programmierung der zuvor entwickelten Modelle und Architekturen. Das Ziel der Implementierung ist ein lauffähiges Produkt, das getestet und zusammengeführt werden kann. Test und Integration stellen abschließend die Erfüllung der Kundenspezifikation sicher. Parallel zu sämtlichen Phasen sorgt eine durchgängige Qualitätssicherung dafür, dass die Qualitätskriterien für jedes Teilprodukt erfüllt sind. Klassische Vorgehensmodelle im IT-Projektmanagement durchlaufen genannte Phasen in der Regel von oben nach unten (vgl. [Kle13], pp. 25-26).

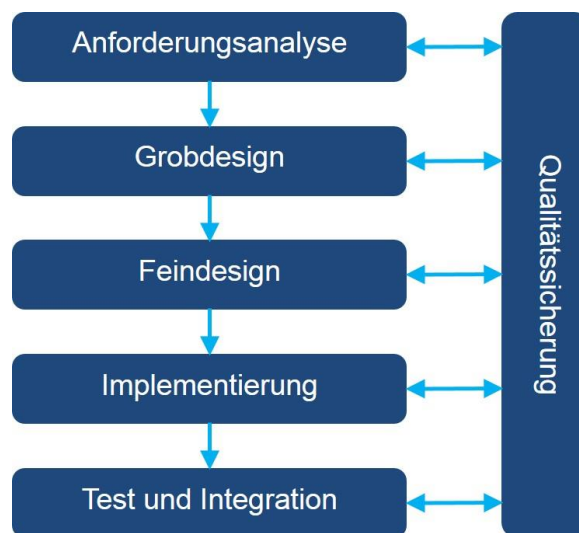


Abbildung 9: Zentrale Phasen der Softwareentwicklung ([Kle13], p. 26)

Ruf und Fittkau bezeichnen klassische Vorgehensmodelle auch als sequentiell, da sie sich durch eine Abfolge festgelegter Phasen auszeichnen. Die Anzahl und Abfolge der Phasen unterscheiden sich jedoch je nach Autor und Vorgehensmodell, was sich unter anderem am Wasserfallmodell zeige. Dort werde beispielsweise, abhängig von Autor und Anwendungsfall, der Betrieb und die Instandhaltung als letzte Phase nach dem Test und der Integration

bis hin zur Stilllegung der IT-Lösung durchlaufen. Kennzeichnend für klassische bzw. sequentielle Vorgehensmodelle ist laut Ruf und Fittkau das am Ende jeder Phase stehende Phasenprodukt, welches gleichzeitig einen Genehmigungspunkt darstellt und im Projekt als Meilenstein fungiert. Zudem darf die anschließende Phase nur gestartet werden, wenn die Qualitätsbedingungen für die Vorgängerphase erfüllt sind (vgl. [RF08], pp. 26-31).

Innerhalb der Gruppe der klassischen Vorgehensmodelle gibt es aufbauend auf dem grundlegenden Prozess des Wasserfallmodells von Royce aus dem Jahr 1970 noch weitere, den sich ändernden Anforderungen angepasste Vorgehensmodelle. Zu diesen Modellen zählt beispielsweise das V-Modell, welches von der NASA zu Beginn der 1990er Jahre vorgestellt wurde. Es ist mehrfach weiterentwickelt worden und unter dem Namen V-Modell XT seit 2004 der Entwicklungsstandard für die Planung und Durchführung von IT-Projekten des Bundes (vgl. [LXS<sup>+</sup>13]).

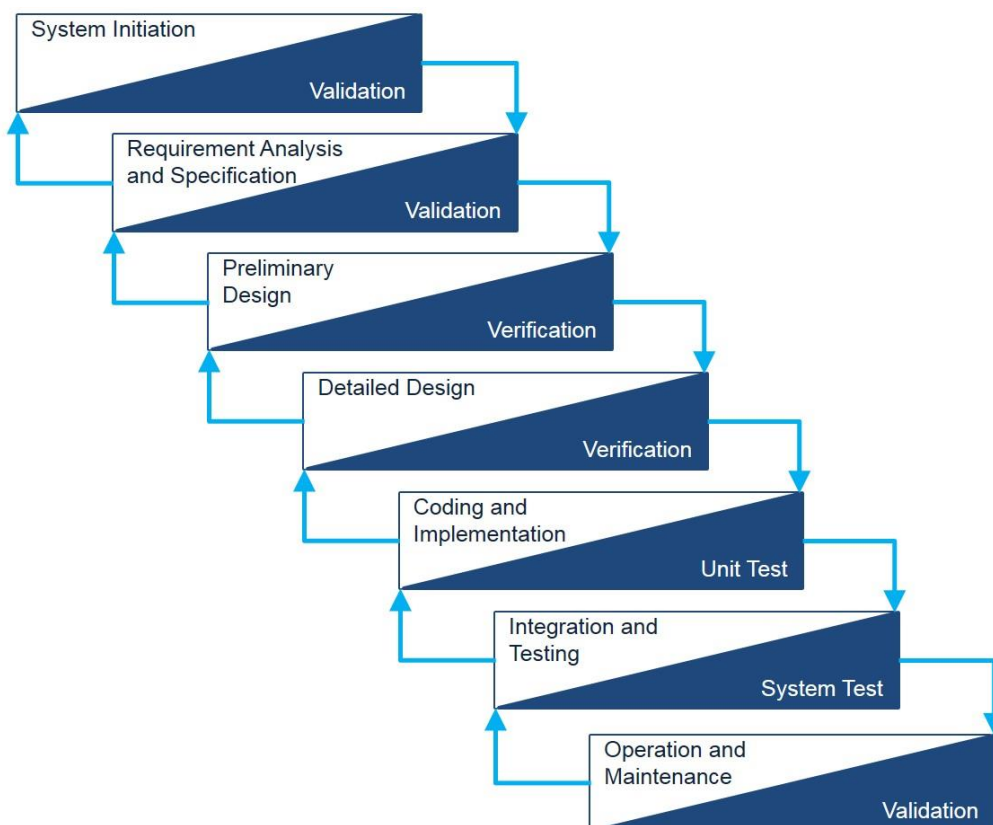
In der Literatur werden als Zwischenform von klassischen und agilen Vorgehensmodellen die sogenannten inkrementellen und iterativen Vorgehensmodelle genannt. Diese Modelle haben den Anspruch die Nachteile von klassischen Methoden zu beheben und bekannten Herausforderungen besser zu begegnen. Durch die Aufteilung des IT-Produkts in selbstständig herstellbare Teilprodukte, die parallel oder aufeinander aufbauend entwickelt werden, sind bei iterativen Verfahren lauffähige Systeme schneller verfügbar und Risiken können frühzeitiger erkannt werden. Inkrementelle Modelle erweitern diese Vorgehensweise durch die Betrachtungsweise der Systementwicklung als evolutionären Prozess, welcher mit einem Prototypen beginnt und über mehrere Iterationen zum Endprodukt gelangt (vgl. [RF08], pp. 32-34). Ein Beispiel für ein solches Vorgehensmodell ist das Spiralmodell nach Boehm aus dem Jahr 1988, welches dem Risikomanagement eine herausragende Bedeutung zukommen lässt (vgl. [SDB<sup>+</sup>10], pp. 57-58). Diese Zwischenformen sind jedoch kein Analysegegenstand der vorliegenden Ausarbeitung.

### **3.3.2 Beispiel Wasserfallmodell**

Das Wasserfallmodell repräsentiert eine der einfachsten Formen von klassischen Vorgehensmodellen. Dies ist gleichzeitig der Grund, weshalb es auch in der Praxis, trotz bekannter Mängel gegenüber anderen Modellen, noch weit verbreitet ist. Es stellt die Grundform eines sequentiellen Vorgehensmodells dar und besteht in seiner Ursprungsform aus den fünf in Kapitel 3.3.1 aufgestellten Lebenszyklusphasen: Anforderungsanalyse, Grobdesign, Feindesign, Implementierung sowie Test und Integration. Die Inhalte einer Phase bauen jeweils auf den Ergebnissen der vorherigen Phase auf. Diese Eigenschaft verleitet laut Trepper leicht dazu die Spezifikationen und Anforderungen des Produkts ausschließlich am Anfang des IT-Projekts zu definieren, sodass spätere Änderungen oder Ergänzungen kaum berücksichtigt werden können. Darüber hinaus kritisiert er, dass die Softwaretests sehr

spät im Entwicklungsprozess verankert sind und bei spät erkannten, tiefgreifenden Mängeln hohe Kosten das Projektbudget belasten. Einen weiteren Kritikpunkt sieht er in der mangelnden Kundeneinbindung während des Projekts. Das Wasserfallmodell hat die Eigenschaft, dass erste lauffähige Versionen, auf die der Kunde eine Rückmeldung geben könnte, vergleichsweise spät vorliegen (vgl. [Tre12], pp. 30-31). Kleuker fügt den Kritikpunkt hinzu, dass bei Projekten, die das Vorgehensmodell Wasserfallmodell anwenden, häufig über lange Zeit der Eindruck erweckt wird, dass die Ergebnisse erfolgreich seien. Dies liege darin begründet, dass bis zur Implementierung ausschließlich mit Modellen und Texten gearbeitet werde. Dies wiederum führe in den letzten Phasen häufig zu ernstzunehmenden Zeitverzögerungen durch verschleppte Probleme (vgl. [Kle13], pp. 26-27).

Das erste Mal publiziert Royce im Jahr 1970 eine Projektmethodologie für die Softwareentwicklung, die im strukturellen Aufbau einem Wasserfall ähnelt. Nachteil des Begriffs Wasserfall ist, dass Rückschritte von Stufen nicht in der Natur eines Wasserfalls liegen. Dies ist nach wie vor ein großer Kritikpunkt an diesem ersten Wasserfallmodell, obwohl Royce sogenannte Feedback-Schleifen integriert hat, welche einen Rückschritt bei gravierenden Qualitätsmängeln erlauben würden (vgl. [Roy70]).



**Abbildung 10: Wasserfallmodell nach Boehm (vgl. [Boe81])**

Die von Boehm im Jahr 1981 veröffentlichte ergänzte Version des ursprünglichen Wasserfallmodells von Royce wird gemeinhin als Stand der Technik anerkannt (Abb. 10). Dabei steht am Ende einer jeden Phase entweder eine Validierung, eine Verifizierung oder ein

Test. Eine Validierung stellt in diesem Fall eine Überprüfung der Aufgabenerfüllung dar. Die Verifizierung hingegen hinterfragt, ob das Ergebnis der Phase den Projektvorgaben entspricht. Zu Beginn eines IT-Projekts steht Boehm zufolge eine Vorstudie und Systemplanung, in der das zu lösende Problem näher beschrieben wird. Darauf folgt die Anforderungsanalyse, welche das System spezifiziert. Im Grobdesign und Feindesign werden die aufgestellten Spezifikationen modelliert, um dann im Rahmen der Implementation in Form von Code realisiert zu werden. Die vorletzte Phase umfasst die Integration, das Testen und die Abnahme des Produkts. Abschließend wird die Software eingeführt und über ihre Lebensdauer gewartet und gepflegt. Am Ende einer jeden Phase liegt ein Meilenstein, der nur bei Erfüllung aller Qualitätsanforderungen überschritten werden darf. Ein Rückschritt in die vorangegangene Phase darf nur bei erheblichen Qualitätsabweichungen erfolgen (vgl. [Boe81]).

Ruf und Fittkau sowie Schatten et. al. empfehlen die Anwendung des Wasserfallmodells bei IT-Projekten, bei denen das Endprodukt schon in einem frühen Stadium detailliert spezifizierbar ist und bei deren Verlauf nur wenige Änderungsanforderungen auftreten. Zudem sei es gut geeignet bei der Erweiterung bereits vorhandener Programme. Des Weiteren weisen sie darauf hin, dass das Wasserfallmodell Grundlage für viele spätere klassische Vorgehensmodelle sei und es das zentrale Paradigma der Ablauffolge im sequentiellen IT-Projektmanagement geprägt habe. Nichtsdestotrotz sei das Vorgehensmodell durch die oben beschriebenen Nachteile in der Praxis von modernen IT-Projekten nur noch vergleichsweise selten anzutreffen und häufig durch flexiblere Modelle ersetzt. (vgl. [RF08], p. 32; [SDB<sup>+</sup>10], p. 49). Dieser These widerspricht Trepper, der behauptet, dass das Wasserfallmodell aufgrund seiner Simplizität, trotz wohlbekannter Unzulänglichkeiten, in der Praxis noch immer häufig in Gebrauch sei (vgl. [Tre12], p. 30).

### **3.4 Agile Vorgehensmodelle**

In diesem Abschnitt geht es um die Vorstellung von Charakteristika agiler Vorgehensmodelle im IT-Projektmanagement. Als in der Praxis relevantester Vertreter dieser Gruppe wird die Scrum-Methodologie kritisch vorgestellt.

#### **3.4.1 Charakteristika**

Agile Vorgehensmodelle fußen, ihrer Bezeichnung folgend, auf ihrer Agilität und Flexibilität. Um sich dem Kern derartiger Modelle nähern zu können, bedarf es zunächst einer Aufarbeitung des Begriffs der Agilität. Trepper stellt 2013 als Zusammenfassung verschiedener Definitionen Agilität wie folgt dar:

„Agilität ist die Fertigkeit schnell, flexibel und situationsbezogen in chaotischen und dynamischen Situationen zu agieren, indem eine Balance zwischen Strukturierung und Flexibilität geschaffen wird, um daraus einen Nutzen für den Kunden und sich selbst zu generieren (vgl. [Tre13], p. 67).“

Kritikpunkte klassischer Vorgehensmodelle, wie die eingeschränkte Flexibilität, die mangelnde Kundeneinbeziehung oder ein scheinbar überhöhter Dokumentationsaufwand werden dieser Definition zufolge durch Agilität behoben (vgl. [SDB<sup>+</sup>10], p. 62). Den Kern agiler Vorgehensmodelle bilden zahlreiche Rückkopplungsprozesse und ein wiederkehrendes iteratives Vorgehen. Das Ziel agiler Methoden in der Softwareentwicklung ist es, in einer möglichst kurzen Zeitspanne lauffähige Entwicklungsergebnisse zu erhalten, um somit die Zeit bis zur Markteinführung (Time-to-Market) zu verkürzen. Darüber hinaus soll durch die hohe Anzahl an Rückkopplungen besser auf neue oder veränderte Produktanforderungen reagiert werden können (vgl. [RF08], pp. 34-37).

Grundlage sämtlicher agiler Projektmanagementmethoden und Vorgehensmodelle sind die vier Leitsätze des sogenannten Agilen Manifests aus dem Jahr 2001 (Tab. 3). Das Agile Manifest wurde von 17 Autoren und Vertretern der bis dahin bekannten agilen Methoden verfasst. Neben den hier aufgeführten vier Leitsätzen umfasst das Manifest zusätzlich noch zwölf Prinzipien, welche die Vorstellungen der Verfasser detaillieren (vgl. [BBB<sup>+</sup>01]).

<b>Individuen und Interaktionen...</b>	...mehr als Prozesse und Werkzeuge
<b>Funktionierende Software...</b>	...mehr als umfassende Dokumentation
<b>Zusammenarbeit mit dem Kunden...</b>	...mehr als Vertragsverhandlung
<b>Reagieren auf Veränderung...</b>	...mehr als das Befolgen eines Plans

**Tabelle 3: Vier Leitsätze des agilen Manifests (vgl. [BBB<sup>+</sup>01])**

Inhaltlich stellen die vier Leitsätze einen qualitativen Vergleich von Bestandteilen agiler Vorgehensmodelle dar. Sämtliche Bestandteile sind wichtig, jedoch haben die zuerst genannten Inhalte dem Agilen Manifest zufolge bei agilen Methoden einen höheren Stellenwert als die zuletzt genannten. Der erste Leitsatz stellt die Projektteilnehmer und deren Umgang miteinander in den Vordergrund. Demnach ist ein lückenhaft geplantes Projekt mit guter Interaktion einem detailliert geplanten Projekt mit schwacher Interaktion vorzuziehen. Der zweite Leitsatz lässt erkennen, dass bei agilen Methoden die Teamleistung besser über funktionierende Software als über eine umfassende Dokumentation festgestellt werden kann. Der dritte Leitsatz stellt die Kundenzusammenarbeit in den Vordergrund. Nur in einer gemeinschaftlichen Beziehung kann demzufolge gute Leistung erbracht werden. Der vierte und letzte Leitsatz priorisiert die Flexibilität und Fähigkeit auf Änderungen zu reagieren gegenüber der starren Befolgung eines zuvor definierten Plans (vgl. [Tre13], pp. 67-69).



Das wohl am verbreitetsten genutzte agile Vorgehensmodell im IT-Projektmanagement ist Scrum. Neben Scrum existieren noch zahlreiche weitere Modelle, wie z.B. Extreme Programming oder Kanban. Sämtliche agile Methoden folgen den Werten des Agilen Manifests und unterscheiden sich demnach oft nur im Detail (vgl. [LXS\*13]).

Neben den hier dargestellten Vorteilen agiler Vorgehensmodelle, welche primär danach streben die Mängel klassischer Methoden zu beheben, zeigen Padberg und Tichy auf, dass es auch Schwächen gibt. So sei die eingeschränkte Dokumentation während des Projekts zwar förderlich in Bezug auf die Entwicklungsgeschwindigkeit, jedoch können hierdurch zu einem späteren Zeitpunkt, z.B. in der Softwarewartung, Probleme durch nicht schriftlich festgehaltenes Wissen auftauchen. Darüber hinaus führe das iterative Vorgehen zu häufigen Anforderungsänderungen und somit zu zahlreichen Umstrukturierungen im Programmcode. Die These, dass die Anforderungsermittlung bei agilen Methoden gegenüber klassischen Vorgehensmodellen effizienter sei, bleibe deshalb fraglich. Weiterhin weisen sie darauf hin, dass agile Vorgehensmodelle hohe Anforderungen an das Entwicklungsteam stellen. Die eingebundenen Mitarbeiter sollten eine ausgeprägte fachliche Expertise, Eigenverantwortung, Disziplin sowie einen sicheren Umgang mit Kunden vorweisen können. Padberg und Tichy halten darüber hinaus den Einsatz agiler Vorgehensmodelle bei großen Entwicklungsprojekten mit über 20 beteiligten Entwicklern für ungeeignet, da ein hohes Maß an Kommunikation und Interaktion gefordert wird. Dies kann bei derartigen Größenordnungen nur begrenzt gewährleistet werden. (vgl. [PT07], pp. 1-9).

### 3.4.2 Beispiel Scrum

Scrum ist ein agiles Vorgehensmodell des IT-Projektmanagements, welches aus einer geringen Anzahl an sehr präzisen Regeln besteht. Es fußt, analog zu sämtlichen anderen agilen Vorgehensmodellen, auf dem Agilen Manifest. In Scrum gibt es drei Rollen: die des Product Owners, die des Scrum Masters sowie die des Entwicklungsteams. Jeder Rolle kommen klar festgelegte Aufgaben und Verantwortlichkeiten zu. Innerhalb kurzer Entwicklungszyklen, welche in Scrum Sprints genannt werden, ist die Erstellung von Produktinkrementen, die über ein priorisiertes Lastenheft, das sogenannte Product Backlog, gesteuert werden (vgl. [Pic08], p.1). Eine detaillierte Darstellung der Scrum-Methodologie wird im Folgenden gegeben (Abb. 11).

In Scrum ist der Product Owner für den ökonomischen Erfolg des Produktes verantwortlich. Er beginnt die Produktentwicklung mit einer klaren Produktvision. Passend zur Produktvision erstellt und priorisiert er die Anforderungen im sogenannten Product Backlog. Sämtliche relevanten Stakeholder werden frühzeitig in den Definitionsprozess von Produktvision und Product Backlog mit eingebunden. Nichtsdestotrotz bleibt der Product Owner verantwortlich für Ausgestaltung der Produkthanforderungen. Zu Beginn jedes Sprints findet die Sprint-

planung statt. Inhalt der Sprintplanung ist eine Vereinbarung zwischen Product Owner und Entwicklungsteam darüber, welche Anforderungen aus dem Product Backlog im nächsten Sprint zu erledigen sind.

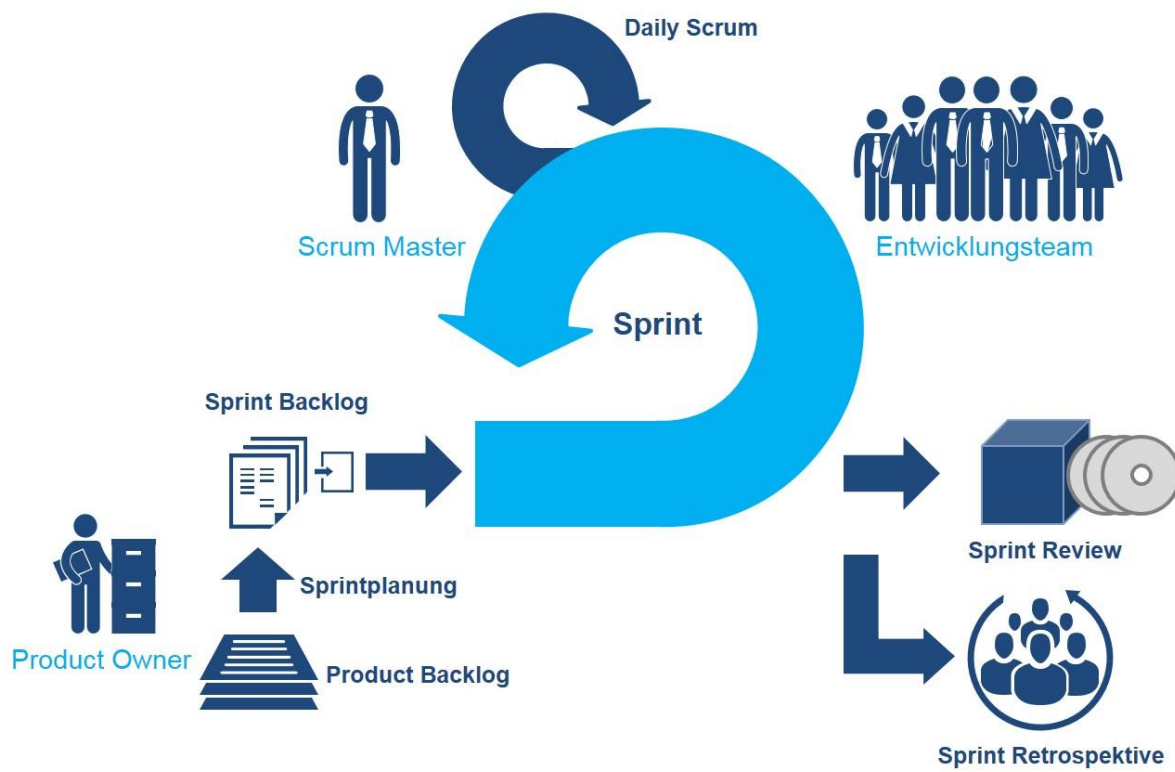


Abbildung 11: Scrum im Überblick (in Anlehnung an [Ita15])

Diese Anforderungen werden unter Zustimmung von Product Owner und Entwicklungsteam in das Sprint Backlog übertragen. Somit tätigt das Entwicklungsteam selbstständig eine Abschätzung darüber, wieviel Produkthanforderungen es im nächsten Sprint abarbeiten kann. Ein Sprint hat eine für das jeweilige Projekt konstant gleiche Länge, die in der Regel zwischen einer und vier Wochen liegt. Durch die gleichbleibende Sprintlänge ergibt sich eine gute Vergleichbarkeit von Entwicklungszyklen, was in einer guten Prognostizierbarkeit für zukünftige Zyklen resultiert. Die Entwicklung erfolgt in direkt aufeinander folgenden Sprints. Während der Sprints arbeitet das Entwicklungsteam selbstorganisiert und störungsfrei an der Abarbeitung des selbst festgelegten Sprint Backlogs. Während des Sprints können somit keine neuen oder geänderten Anforderungen an das Team gestellt werden. Diese können erst bei der nächsten Sprintplanung berücksichtigt werden. Während eines Sprints gibt es täglich sogenannte Daily Scrums. Die etwa 15 Minuten andauernden, meist im Stehen abgehaltenen und jeden Werktag um die gleiche Uhrzeit stattfindenden Besprechungen beinhalten die aktuelle Einsatzplanung sowie einen Statusabgleich aller Beteiligten Entwickler. Jedes Teammitglied muss sich im Daily Scrum drei Fragen stellen:

1. Was wurde seit dem letzten Daily Scrum erledigt?
2. Sind während der Entwicklung Probleme aufgetreten und wenn ja – welche waren das?
3. Was ist das Ziel bis zum nächsten Daily Scrum?

Während eines Scrum-Projekts sorgt der sogenannte Scrum Master dafür, dass alle mit dem Scrum-Vorgehensmodell einhergehende Regeln und Prozesse eingehalten werden und die Teams entsprechend arbeiten. Zudem schützt er das Team gegen Störungen von außen und beseitigt Hindernisse in Bezug auf die Methodologie. Er löst jedoch keine direkten Probleme des Teams, sondern hilft ihm die vorliegenden Herausforderungen selbst zu meistern. Der Scrum Master hat keine Weisungsbefugnis gegenüber dem Entwicklungsteam und folgt dem Grundsatz „Führen durch Dienen“. Am Ende jedes Sprints findet das Sprint Review statt. Zu diesem Termin präsentiert das Team dem Product Owner basierend auf der vorliegenden Sprintplanung die Ergebnisse des Sprints. Es werden darin zudem häufig neue Anforderungen an das Produkt definiert, die daraufhin in das Product Backlog eingestellt werden. Neben dem Sprint Review, in dem es um die Ergebnisse geht, gibt es darüber hinaus eine Sprint Retrospektive, in welcher das Entwicklungsteam sich durch Selbstreflexion in Bezug auf die Einhaltung von Scrum-Regeln selbst optimiert (vgl. [Pic08], pp. 7-24).

Innerhalb eines Sprints wächst das Produkt inkrementell. Mit jeder Sprint-Iteration entwickelt sich die Software somit über alle Technologieschichten hinweg evolutionär weiter. Dabei knüpft die Entwicklung im jeweils nächsten Sprint an die bisherigen Ergebnisse an und baut auf dem Stand des vorangegangenen Sprints auf. Parallel zum eigentlichen Produkt wachsen auch Tests und Dokumentation schrittweise. Zur Visualisierung und Kontrolle der Projektaktivitäten kommt bei Scrum das sogenannte Burndown Chart zum Einsatz. Es handelt sich dabei um ein zweidimensionales Diagramm, welches die noch zu erledigenden Aufgaben des Projekts in Relation zur noch verfügbaren Arbeitszeit setzt. Auf der Abzisse wird dabei der Zeitverlauf in Sprints abgetragen, auf der Ordinate die noch zu leistenden Aufgaben in Zeit oder Anzahl. Das Burndown Chart wird kontinuierlich vom Product Owner aktualisiert (vgl. [Pic08], pp. 83-85).

Trecker et. al. bewerten Scrum als ein ausgesprochen kundenorientiertes agiles Vorgehensmodell, welches durch häufige Iterationen eine Lösung zur schnellen Reaktion auf neue oder geänderte Kundenanforderungen bietet. Durch die direkte Einbindung des Kunden in den Entwicklungsprozess gelangen jedoch auch Anforderungen ins Projekt, die bei anderen Vorgehensmodellen auf Managementebene herausgefiltert würden. Positiv wirke sich die Rolle des Scrum Masters aus, da er die Ziel- und Produktorientierung im Projekt fördere. Weiterhin führe die weitgehend selbstständige Arbeit des Entwicklungsteams zu einem hohen Maß an intrinsischer Motivation. Das größte Hindernis zur

Anwendung von Scrum, so Trecker et. al., sei die vergleichsweise schlechte Prognostizierbarkeit der Projektkosten und des Projektzeitrahmens zu Projektbeginn (vgl. [TAN<sup>+</sup>13]).

### 3.5 Vergleich von Vorgehensmodellen

In diesem Abschnitt werden typische, in der Literatur bereits existierende Vergleichskriterien von klassischen und agilen Vorgehensmodellen vorgestellt, die sich für eine quantitative Analyse eignen. Weiterhin werden Inhalt und Methodik der Nutzwertanalyse geschildert, um darauf aufbauend im zweiten Teil der vorliegenden Ausarbeitung die Gegenüberstellung von klassischen und agilen Vorgehensmodellen durchführen zu können.

#### 3.5.1 Typische Vergleichskriterien

Um Methoden des Projektmanagements adäquat vergleichen zu können, müssen ausgehend von Zielen und Erfolgsfaktoren Kriterien gefunden werden, die für verschiedene Vorgehensmodelle gleichermaßen anwendbar sind. Die Determinanten eines Projekts werden in der Literatur häufig mit Hilfe des sogenannten magischen Projektdreiecks beschrieben. Holzbaur liefert 2007 eine Version des Dreiecks, dessen Ausprägungen er mit Ergebnis, Zeit und Ressourcen bezeichnet (Abb. 12).

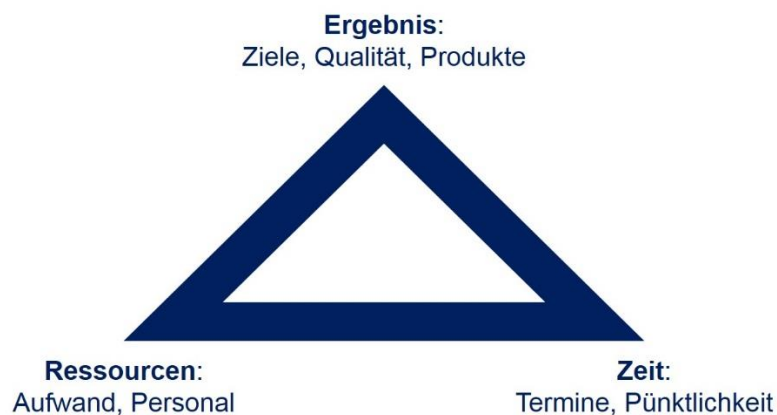


Abbildung 12: Magisches Projektdreieck (vgl. [Hol07], p. 93)

Die Dimension Ergebnis umfasst dabei sowohl Qualitätsaspekte als auch Wertschöpfung und Zielerfüllung. Ressourcen können Holzbaur zufolge in Geld, Arbeitszeit, Infrastruktur oder auch Personal gemessen werden. Inwieweit das Projekt den gesetzten Termin einhalten kann, beschreibt die Dimension der Zeit. Zu beachten ist, dass eine Änderung einer Dimension des Dreiecks in jedem Fall auch die anderen beiden Dimensionen beeinflusst (vgl. [Hol07], pp. 93-94).

Aufbauend auf dem magischen Projektdreieck definiert Holzbaur Einflusskriterien für erfolgreiche Projektarbeit. Dabei hebt er hervor, dass es neben messbaren, eher technischen Größen, auch auf weiche, qualitative Faktoren ankommt. Holzbaur bezeichnet diese Kriterien

als harte und weiche Faktoren. In Tabelle 4 ist eine Auswahl der Faktoren aufgeführt und näher beschrieben (vgl. [Hol07], pp 95-96).

Bezeichnung	Detailierung
Produkteigenschaften (harte Faktoren)	Dieses Kriterium umfasst die Art und den Umfang des Produkts in einem definierten Zielsystem
Projekteigenschaften (harte Faktoren)	<ul style="list-style-type: none"> <li>▪ Bezüglich der Aufgabe: Dazu gehören Innovationsgrad, Auftraggeber und Größe</li> <li>▪ Bezüglich der Durchführung: In dieser Ausprägung geht es um Management, Umfeld, Werkzeuge und Organisationsformen</li> </ul>
Erfahrung und Qualifikation (harte Faktoren)	Die Eignung des Projektteams in Bezug auf die Aufgabe, angewandte Methoden und Werkzeuge sowie Teamarbeit
Rahmenbedingungen bzgl. Ressourcen (harte Faktoren)	Hierbei geht es darum, wieviel Zeit, Geld, Personal und Werkzeuge zur Verfügung stehen. Weiterhin spielen Regeln, Struktur und Flexibilität bzgl. Terminen und Kosten eine Rolle
Rahmenbedingungen bzgl. Organisation (harte Faktoren)	Hierzu zählen Qualitätsmanagement und Grad der Dokumentation und Richtlinien. Weiterhin geht es um den Einsatz von Managementtechniken und Vorgehensmodellen
Menschliche Faktoren (weiche Faktoren)	Die Motivation der Projektteilnehmer ist genau wie die intensive Kommunikation im Projekt von herausragender Bedeutung

**Tabelle 4: Einflusskriterien erfolgreicher Projektarbeit nach Holzbaur (vgl. [Hol07], pp. 95-96)**

Holzbaur bezieht sich bei seinen Einflusskriterien auf das Projektmanagement im übergeordneten Sinne. In der Literatur gibt es darüber hinaus jedoch auch Autoren, die Erfolgsfaktoren und Einflussfaktoren im Kontext klassischer und agiler Vorgehensmodelle im IT-Projektmanagement analysiert haben. So haben Boehm und Turner sowie Leffingwell untersucht, wie sich Vorgehensmodell-Typen in Bezug auf die primär verfolgten Ziele unterscheiden. Boehm und Turner stellen fest, dass Vorhersagbarkeit, Stabilität und hohe Sicherheit die maßgeblichen Ziele klassischer Vorgehensmodelle seien. Agile Vorgehensmodelle hingegen haben dem Agilen Manifest zufolge Flexibilität in Form von Reaktionsfähigkeit auf veränderte Anforderungen, frühzeitig funktionierende Software sowie Kundenzufriedenheit als primäre Ziele (vgl. [BT04], p. 27; vgl. [BBB<sup>+</sup>01]).

Eine treffende Zusammenführung dieser Zieldimensionen liefert Leffingwell im Jahr 2009, indem er aufzeigt, dass agile Vorgehensmodelle das magische Projektdreieck mit seinen Parametern auf den Kopf stellt. Abbildung 13 zeigt, dass eine Änderung der Produktanforderungen bei klassischen Vorgehensmodellen vermieden werden soll, da die Anforderung eine aus dem Konzept heraus feststehende Größe ist. Aus dem Plan einer klassischen Methode resultieren die Schätzungen für notwendige Ressourcen und die Zeit. Dies sind nach Leffingwell jedoch keine primär einschränkenden Größen. Auf der anderen

Seite sind bei agilen Vorgehensmodellen Anforderungsänderungen durchaus möglich. Der Zeitplan und die Ressourcen stehen jedoch zumeist fest und stellen somit Restriktionen dar.



Abbildung 13: Parameter klassischer und agiler Vorgehensmodelle (in Anlehnung an [Lef09])

Einen konkreten Vorschlag für Kriterien sowie deren Ausprägungen in erfolgreichen Projekten liefern Boehm und Turner im Jahr 2004 mit der Untersuchung von sogenannten Optimalbedingungen für klassische und agile Vorgehensmodelle. Sie behaupten, dass je mehr die tatsächlich vorliegenden Bedingungen im Projekt von den Optimalbedingungen abweichen, desto unwahrscheinlicher sei der Erfolg des Projekts und umso sinnvoller sei es, eine Mischform von klassischen und agilen Methoden anzuwenden. Sie fassen ihre Beobachtungen in fünf kritischen Erfolgsfaktoren zusammen, welche die relative Eignung des jeweiligen Vorgehensmodell-Typs determinieren. In Tabelle 5 sind die Erfolgsfaktoren von Boehm und Turner in den Dimensionen Projektgröße, Kritikalität, Dynamik, Personal und Kultur aufgeführt. Die Projektgröße untersucht, wie gut die unterschiedlichen Vorgehensmodell-Typen geeignet sind sowohl große als auch kleine Projekte abzubilden. Der Erfolgsfaktor Kritikalität beschreibt Boehm und Turner zufolge die Fähigkeit von Vorgehensmodellen der Entwicklung sicherheitskritischer Produkte einen Rahmen zu geben. Sie unterscheiden dabei zwischen Produkten, die im Betrieb Verantwortung für zahlreiche Menschenleben haben werden und Produkten, die bei Fehlern lediglich mit geringen Kosten zu Buche schlagen. Die Dynamik, als dritter Erfolgsfaktor, beschreibt die Umgebung, in der die Vorgehensmodelle vornehmlich erfolgreich angewendet werden können. In der Dimension Personal untersuchen sie, welches Qualifikationsniveau im Projektteam vorliegen muss, damit das Vorgehensmodell effektiv betrieben werden kann. Abschließend tätigen sie mit dem Erfolgsfaktor Kultur eine Aussage über die Art und Weise der Kommunikation und Interaktion im Projekt. Sie geben darüber hinaus zu beachten, dass ein Projekt, welches nur vier der fünf Dimensionen hinreichend erfüllt einem Risikomanagement unterzogen werden sollte, um abzuwägen, ob eine Mischform zweckmäßiger sei (vgl. [BT04], pp. 51-56).

Erfolgsfaktor	Agile Vorgehensmodelle	Klassische Vorgehensmodelle
Projektgröße	Gut geeignet für kleinere Produkte und Teams. Die Abhängigkeit von implizitem Wissen erschwert die Skalierbarkeit	Methoden, die entstanden, um große Produkte und Teams zu steuern. Schwer auf kleine Projekte zu übertragen
Kritikalität	Nicht auf sicherheitskritischen Produkten getestet. Mögliche Schwierigkeiten durch einfaches Design und mangelnde Dokumentation	Methoden, die entstanden, um hochkritische Produkte zu entwickeln. Schwer auf unkritische Projekte zu übertragen
Dynamik	Einfaches Design und kontinuierliche Überarbeitung eignen sich hervorragend für dynamische Umgebungen. Dies sind gleichzeitig potentielle Quellen für teure Nacharbeit in hochstabilen Umgebungen	Detaillierte Planung und Vorabdesign sind hervorragend geeignet für hochstabile Umgebungen. Dies sind gleichzeitig Quellen für teure Nacharbeit in hochdynamischen Umgebungen
Personal	Es bedarf einer dauerhaften Präsenz von hochqualifiziertem Personal mit Erfahrung in agilen Vorgehensmodellen	Insbesondere in der Projektdefinitionsphase bedarf es einer kritischen Menge an hochqualifiziertem Personal. Diese Menge kann im Projektverlauf reduziert werden, es sei denn das Projekt findet in einer hochdynamischen Umgebung statt
Kultur	Erfolgreich in einer Kultur, in der Mitarbeiter durch eine Vielzahl an Freiheitsgraden motiviert werden und sich wohl fühlen (Erfolgreich durch Chaos)	Erfolgreich in einer Kultur, in der Mitarbeiter durch klare Rollen, Verantwortlichkeiten und Prozesse motiviert werden und sich wohl fühlen (Erfolgreich durch Ordnung)

**Tabelle 5: Fünf kritische Erfolgsfaktoren für agile und klassische Vorgehensmodelle ([BT04], p. 55)**

### 3.5.2 Nutzwertanalyse

Die Nutzwertanalyse ist eine ursprünglich aus den Vereinigten Staaten stammende und von Zangemeister in den 1970er Jahren verfeinerte Planungsmethode, um die Entscheidungsfindung in multidimensionalen Systemen zu vereinfachen. Sie wird primär zum quantitativen Vergleich von Handlungsalternativen angewandt. Zangemeister definiert die Nutzwertanalyse wie folgt:

„Die Nutzwertanalyse ist eine Planungsmethodik zur systematischen Entscheidungsvorbereitung bei der Auswahl komplexer Projektoalternativen. Ihr besonderes Kennzeichen ist darin zu sehen, dass der Bewertung von Alternativen unter vergleichsweise geringen praktischen Schwierigkeiten auch eine Vielfalt von Zielkriterien zugrunde gelegt werden kann. Die Nutzwertanalyse gewinnt daher in allen sozio-ökonomisch-technischen Bereichen, wo die herkömmlichen, eindimensionalen Methoden der ökonomischen Investitionsrechnung alleine zur Entscheidungsfindung nicht ausreichen, zunehmend an Bedeutung ([Zan70], p. 1).“

Die Nutzwertanalyse findet demnach immer dann Anwendung, wenn eine Entscheidungssituation vorliegt und die rein monetäre Bewertung nicht möglich oder nicht hinreichend ist, eine Vielzahl von Alternativen zu vergleichen sind oder ein multidimensionales Zielsystem vorliegt. Die Eigenschaften jeder Alternative werden dabei in Nutzengrößen umgewandelt und miteinander verglichen. Zu beachten ist jedoch, dass die Nutzwertanalyse keine Objektivität gewährleistet. Sämtliche Gewichtungen und Bewertungen basieren auf den individuellen Erfahrungen und Präferenzen der durchführenden Personen. Vielmehr stellt die Nutzwertanalyse die subjektiven Ziele und Werte des Entscheidungsträgers in strukturierter und systematisierter Form dar (vgl. [Zan70], pp. 55-60).

Die Nutzwertanalyse kommt in der Praxis in zahlreichen Varianten zum Einsatz. In dieser Ausarbeitung wird der Ansatz des REFA – Verband für Arbeitsgestaltung, Betriebsorganisation und Unternehmensentwicklung verwendet.

In einem ersten Schritt werden dabei  $m$  relevante Bewertungskriterien definiert, die für beide Alternativen gleichermaßen geeignet sind. Im zweiten Schritt wird die Gewichtung  $G_k$  der einzelnen Kriterien mit Hilfe eines Paarvergleichs ermittelt. In Tabelle 6 ist diese Vorgehensweise illustrativ skizziert. Sämtliche Kriterien werden dabei entsprechend ihrer Wichtigkeit für den jeweiligen Untersuchungszweck miteinander verglichen. Sie sind entweder wichtiger, gleich wichtig oder weniger wichtig als das gegenüber gestellte Kriterium. Die Gewichtung eines Kriteriums ergibt sich aus der kumulierten Punktzahl  $P_k$  geteilt durch die Summe der kumulierten Punktzahlen aller Kriterien (Gl. (1)).

Gl. (1)

$$G_k = \frac{P_k}{\sum_{k=1}^m P_k} * 100\%$$



als wichtiger	Kriterium 1	Kriterium 2	Kriterium 3	Kriterium 4	Kriterium 5	Kriterium 6	Kriterium 7	Kriterium 8	Kriterium 9	Kriterium 10	Summe $P_k$	Gewichtung $G_k$
Kriterium 1	1	0,5	0,5	1	0,5	1	1	1	1	1	7,5	16,67%
Kriterium 2	0	1	0,5	0,5	0,5	0,5	0	0	1	0,5	3,5	7,78%
Kriterium 3	0,5	0,5	1	0,5	0,5	0,5	0	0,5	1	0,5	4,5	10,00%
Kriterium 4	0,5	0,5	0,5	1	0,5	0,5	0,5	0	0,5	0,5	4	8,89%
Kriterium 5	0	0,5	0,5	0,5	1	0	0	0,5	0,5	1	3,5	7,78%
Kriterium 6	0,5	0,5	0,5	0,5	1	1	0,5	1	1	1	6,5	14,44%
Kriterium 7	0	1	1	0,5	1	0,5	1	1	0,5	0,5	6	13,33%
Kriterium 8	0	1	0,5	1	0,5	0	0	1	1	0,5	4,5	10,00%
Kriterium 9	0	0	0	0,5	0,5	0	0,5	0	1	0,5	2	4,44%
Kriterium 10	0	0,5	0,5	0,5	0	0	0,5	0,5	0,5	1	3	6,67%
Summe											45	100,00%

**Bewertungsrichtlinie**  
 0 = weniger wichtig als das andere Kriterium  
 0,5 = gleich wichtig wie das andere Kriterium  
 1 = wichtiger als das andere Kriterium

Tabelle 6: Paarvergleich zur Ermittlung der Kriteriengewichtung (in Anlehnung an [Ref91])

Der dritte Schritt umfasst dem REFA-Modell zufolge die eigentliche Bewertung der Alternativen und die daraus resultierende Ermittlung der Nutzwerte. In Tabelle 7 ist dieser Prozess schematisch mit fiktiven Werten dargestellt.

Kriterium	Gewichtung $G_k$	Alternative 1		Alternative 2	
		Wert $\omega_{i,k}$ (max. 10)	Teilnutzwert $n_{i,k}$	Wert $\omega_{i,k}$ (max. 10)	Teilnutzwert $n_{i,k}$
Kriterium 1	16,67%	6	10,00%	1	1,67%
Kriterium 2	7,78%	6	4,67%	2	1,56%
Kriterium 3	10,00%	5	5,00%	4	4,00%
Kriterium 4	8,89%	8	7,11%	8	7,11%
Kriterium 5	7,78%	8	6,22%	7	5,44%
Kriterium 6	14,44%	8	11,56%	8	11,56%
Kriterium 7	13,33%	4	5,33%	5	6,67%
Kriterium 8	10,00%	8	8,00%	7	7,00%
Kriterium 9	4,44%	3	1,33%	7	3,11%
Kriterium 10	6,67%	2	1,33%	5	3,33%

Nutzwert  $N_i$  in Bezug auf Idealprofil (=100%)

$N_1 = 60,56\%$

$N_2 = 51,44\%$

Tabelle 7: Alternativenbewertung zur Ermittlung des Nutzwerts (in Anlehnung an [Ref91])

Jede Entscheidungsalternative wird innerhalb eines Kriteriums mit einem individuellen Wert  $\omega_{i,k}$  belegt. Im illustrierten Beispiel kann der Wert  $\omega_{i,k}$  Beträge zwischen 0 und 10 einnehmen. Dabei stellt der Maximalwert  $\omega_{max}$  das Optimum dar. Der Teilnutzwert  $n_{i,k}$  für das jeweilige Kriterium ergibt sich aus dem Quotienten des Werts  $\omega_{i,k}$  und dem Maximalwert  $\omega_{max}$  multipliziert mit der Gewichtung  $G_k$  des Kriteriums (Gl. (2)).

Gl. (2)

$$n_{i,k} = \left( \frac{\omega_{i,k}}{\omega_{max}} \right) * G_k$$

Der Gesamtnutzwert  $N_i$  einer Alternative ergibt sich aus der Summe aller Teilnutzwerte  $n_{i,k}$  und ist in Bezug auf das Idealprofil von 100 Prozent Erfüllungsgrad zu bewerten (Gl. (3)).

Gl. (3)

$$N_i = \sum_{k=1}^m n_i(k)$$

Die Alternative mit dem höchsten Nutzwert stellt die zu bevorzugende Lösung dar (vgl. [Ref91]).

## 4 Einführung von Big-Data-Management-Systemen

Aufbauend auf den gewonnenen Kenntnissen aus der Literatur, geht es nun darum, zielgerichtet einen Vergleich von klassischen und agilen Vorgehensmodellen des IT-Projektmanagements bei der Einführung von Big-Data-Management-Systemen durchzuführen. Im Folgenden werden die vier in Kapitel 2.4 erläuterten Big-Data-Strategien mit ihren individuellen Anforderungen und Einführungsprozessen untersucht. Dazu wird zunächst ein Überblick über die unterschiedlichen Einführungsprozesse der Strategien gegeben. Im Anschluss findet der quantitative Vergleich über die Eignung der Vorgehensmodell-Typen im Rahmen einer Nutzwertanalyse statt. Ergebnis sind Handlungsempfehlungen über das am besten geeignete Modell für die jeweilige Big-Data-Strategie.

### 4.1 Einführungsprozesse verschiedener Big-Data-Strategien

Betrachtungsgegenstand dieses Abschnitts sind die Einführungsprozesse der in Kapitel 2.4 erläuterten Strategien und Ziele von Big-Data-Initiativen. Im Detail handelt es sich dabei um die vier verschiedenen Strategien von Big Data in Organisationen nach Davenport: die Kostenreduktion mit Hilfe von Big Data, die Zeiteinsparung durch Big Data, die Entscheidungsunterstützung durch Big Data sowie die Entwicklung neuer Produkte auf Basis von Big Data. Der Grund, weshalb zur Analyse die Strategien nach Davenport gewählt wurden, liegt in erster Linie in ihrer guten Abgrenzbarkeit und Heterogenität. Es ist genauso vorstellbar verschiedene Arten von Big Data oder Branchen als Ordnungskriterium anzuwenden. Bei der Untersuchung dieser Alternativen hat sich jedoch herausgestellt, dass die für die Analyse wichtigen Eigenschaften, wie z.B. differenzierbare Einführungsprozesse und verschiedenartige Zielrichtungen, am besten durch die Kategorisierung in Strategien repräsentiert werden.

Zu beachten ist, dass die vier Strategien keinen Anspruch auf Vollständigkeit stellen. Demnach gibt es womöglich darüber hinaus noch weitere Stoßrichtungen oder auch Mischformen der hier aufgeführten Verfahrensweisen. Um jedoch die Auswertbarkeit im Rahmen einer quantitativen Analyse zu gewährleisten, beschränkt sich diese Ausarbeitung auf die genannten vier klar abgrenzbaren Strategien.

Übergreifend ist festzustellen, dass in der Literatur keine einheitlichen Anhaltspunkte zur operativen Einführung eines Big-Data-Management-Systems vorliegen. Grundsätzlich befinden sich Organisationen, sofern sie ein geschäftliches Interesse an Big Data haben, in einer der vier in Kapitel 2.3.2 vorgestellten Einführungsphasen: Informieren, Planen, Prüfen und Umsetzen. Diese hat das IBM Institute for Business Value im Rahmen einer Befragung vom Entwicklungsstand von Big Data in Unternehmen im Jahr 2012 identifiziert. Das Ziel dieser Ausarbeitung ist es, eine Handlungsempfehlung darüber zu geben, welches Vorgehensmodell des IT-Projektmanagements am geeignetsten für die jeweilige Big-Data-

Strategie ist. Da sich die vorgestellten Vorgehensmodelle primär auf die Planung und Umsetzung von Softwareentwicklungsprojekten beziehen, wird die Phase des Informierens deshalb hier vernachlässigt. Weiterhin muss festgestellt werden, dass sich die Einführungsprozesse der verschiedenen Big-Data-Strategien auf technischer Ebene unter Umständen nur marginal unterscheiden. Es kann vorkommen, dass mehreren Strategien die gleiche Datenarchitektur, wie sie z.B. in Kapitel 2.3.3 dargestellt ist, zugrunde liegt. Viele Big-Data-Anwendungen haben diesbezüglich ähnliche Anforderungen, wie z.B. ein vor das Data-Warehouse angekoppeltes Hadoop-Cluster zur flexiblen Exploration von Daten. Ungeachtet dessen weisen die vier Strategien in Bezug auf die Prioritätensetzung, die Primärziele sowie die Leitparadigmen im Projektmanagement deutliche Unterschiede auf. Dies ermöglicht eine Differenzierung hinsichtlich der vorzugsweise anzuwendenden Vorgehensmodelle. Die erwähnten Prioritäten und Primärziele der einzelnen Strategien im Hinblick auf das Projektmanagement werden, basierend auf den in der Literatur genannten individuellen Eigenschaften, Zielen und Herausforderungen, im Folgenden dargestellt.

### **Kosteneinsparung durch Big Data**

Zunächst muss in diesem Kontext festgelegt werden, was die Kosteneinsparungsstrategie auf Basis von Big-Data-Technologie im Detail bedeutet. In den meisten Fällen geht mit der Einführung von Big-Data-Management-Systemen eine Kostenreduktion in diversen Bereichen einher. Oft ist es dabei jedoch so, dass die Verringerung der Kosten nicht das Primärziel der Big-Data-Initiative ist, sondern indirekt über andere Verbesserungspotentiale generiert wird. Ein Beispiel dafür ist die in Kapitel 2.4.1 erwähnte Vorabselektion bei der Identifikation von Fehlalarmen in der Instandhaltung, wodurch in erster Linie die Servicequalität optimiert werden soll, mittelbar jedoch auch die Kosten reduziert werden. In der vorliegenden Ausarbeitung soll die Big-Data-Strategie Kosteneinsparung hingegen ausschließlich die direkte Ausgabenreduzierung für die Datenspeicherung und -verarbeitung umfassen.

Grundsätzlich handelt es sich somit um ein überschaubares Konstrukt an notwendigen Technologien, was wiederum einen großen Einfluss auf den Projektverlauf hat. Durch die schon zu Projektbeginn zumeist sehr klaren Produktanforderungen können Arbeitspakete zielorientiert definiert werden. Die Implementierung einer Big-Data-Struktur birgt zahlreiche Möglichkeiten, wie z.B. die Analyse externer unstrukturierter Daten, die über den reinen Kosteneinsparungseffekt hinausgehen. Es ist bei der Einführung unter der Prämisse Kosten ausschließlich über die Technologie einzusparen, deshalb umso wichtiger einen klaren Fokus zu bewahren und zielgerichtetes Projektmanagement anzuwenden. Organisatorisch ist ein solches Projekt häufig in einer zentralisierten IT-Abteilung anzusiedeln. Für die Anwender in den einzelnen Fachabteilungen ändert sich operativ kaum etwas. Die notwendigen Big-Data-Kenntnisse beschränken sich aus diesem Grund auf eine vergleichbar

kleine Gruppe von Mitarbeitern in der IT-Abteilung. Da es sich bei dieser Strategie um eine rein technologische Umstellung von traditionellen Systemen auf Big-Data-Umgebungen handelt, lässt sich der Erfolg des Projekts bzw. der Grad der Kosteneinsparung effektiv am Return on Investment gegenüber den Altsystemen aufzeigen. Zu beachten ist jedoch der Aspekt, dass traditionelle Systeme in der Regel erprobt sind und ausgesprochen zuverlässig und sicher arbeiten. Dies kann bei der Umstellung auf neuartige Big-Data-Umgebungen zu außerplanmäßigen Zusatzkosten führen. Die Projektkomplexität sowie der daraus resultierende Projektaufwand sind maßgeblich abhängig von der Größe der umzustellenden Datenarchitektur. Die Umstellung erfolgt dabei in jedem Fall schrittweise und klar voneinander abgegrenzt.

### **Zeiteinsparung mit Hilfe von Big Data**

Die zweite Strategie verfolgt das Ziel mit Hilfe von Big Data Zeit einzusparen. Auch hier ist eine Abgrenzung notwendig, um Überschneidungen mit anderen Strategien für die Analyse auszuschließen. Zeiteinsparung durch Big Data kann, wie in Kapitel 2.4.2 anhand von Anwendungsfällen illustriert, in zwei markanten Ausprägungen erfolgen. Auf der einen Seite können bekannte Analyseprozesse durch performanceorientierte Big-Data-Anwendungen, wie z.B. In-Memory-Technologie, beschleunigt werden. Auf der anderen Seite besteht die Möglichkeit durch verbesserte Analysefähigkeit und die Gewinnung neuer Erkenntnisse indirekt Zeit einzusparen. Zweitgenannte Ausprägung hat in der hier angewandten Kategorisierung von Big-Data-Strategien Überschneidungspunkte mit der Strategie der Entscheidungsunterstützung und Prognoseverbesserung durch Big Data. Aus diesem Grund wird im Rahmen der Nutzwertanalyse davon ausgegangen, dass die Strategie der Zeiteinsparung ausschließlich performanceorientiert verfolgt wird.

Anders als bei der Strategie der Kosteneinsparung, welche Big-Data-Technologien als Fundament der Datenarchitektur nutzt, geht es bei der Zeiteinsparung mehr darum wenige oder zumindest einzelne bekannte Analyseprozesse schneller abzuarbeiten und, wenn möglich, mit bisher ungenutzten Daten anzureichern. Dies erfordert bei Projektbeginn detaillierte fachliche und technische Informationen sowie hohe Transparenz über die zu beschleunigenden Prozesse. Die spätere operative Nutzung des Big-Data-Management-Systems erfolgt zumeist durch die Fachabteilungen. Die Herstellung der genannten Prozesstransparenz sowie die reibungslose Eingliederung in das Tagesgeschäft gelingen nur, wenn die betroffenen Fachabteilungen interdisziplinär eng mit in das Einführungsprojekt eingebunden werden. Die konkrete Ausgestaltung des Zielprozesses sowie spätere Nutzungsmöglichkeiten der eingesparten Zeit können im Projektverlauf variieren. Die Komplexität dieser Strategie besteht zu großen Teilen in der hardware- und softwareseitigen Harmonisierung von klassischen Systemen und innovativen Big-Data-Hochleistungskomponenten. Dies erfordert Expertenwissen aus verschiedenen Fach-

gebieten. Die Interdisziplinarität der Projektteilnehmer, die hohe Komplexität der Aufgabe sowie die teils erst im Projektverlauf auftauchenden Produkthanforderungen erfordern ein hohes Maß an Koordination. Dies ist ein weiterer Grund dafür, dass Big-Data-Initiativen unter der Prämisse der Zeiteinsparung häufig in einem vergleichbar kleinen Rahmen projektiert werden und nur mit erhöhtem Aufwand skalierbar sind.

### **Entscheidungsunterstützung und Prognosen durch Big Data**

Die klassische unternehmerische Datenanalyse mit Fragestellungen aus unterschiedlichsten Funktionsbereichen kann mit Hilfe von Big-Data-Technologien substantiell angereichert werden. Der größte Mehrwert liegt dabei, wie in Kapitel 2.4.3 dargestellt, in der Fähigkeit externe und teils unstrukturierte Daten in die Entscheidungsfindung und Prognosen miteinzubeziehen. Diese Eigenschaft ist gleichzeitig das ausschlaggebende Unterscheidungskriterium zur klassischen Business Intelligence. Der Einführungsprozess von Big-Data-Technologien im Kontext dieser Strategie ist entscheidend abhängig von der Ansiedlung der Datenanalysten in der Organisationsstruktur. Die bisher verfügbaren Informationen werden in der Regel zentral bereitgestellt. Aufgrund des stark diversifizierten und abteilungsabhängigen Informationsbedarfs werden genannte Informationen nutzerfreundlich innerhalb eines Business-Intelligence-Systems aufbereitet und den dezentralen Einheiten passgenau zur Verfügung gestellt. Bei der Einführung von Big-Data-Management-Systemen werden bestehende Datenarchitekturen auf allen Technologieschichten aktualisiert. Die größte Herausforderung dieser Strategie ist die Fähigkeit relevante interne und externe Daten zu erkennen und die richtigen Fragen in Bezug auf die Datenanalyse zu formulieren. Dazu benötigt man ein ausgewogenes Team von einerseits Big-Data-Technologieexperten und andererseits Fachexperten der Geschäftsbereiche. Darüber hinaus helfen Datenwissenschaftler die Lücke zwischen technischen und fachlichen Aspekten zu schließen. Die Anforderungen an ein Big-Data-Management-System, welches auf Basis dieser Strategie entwickelt wird, ändern sich im Projektverlauf mitunter mehrfach. Dies ist vornehmlich durch die strukturell unterschiedlichen Quelldaten begründet. Zusammenfassend kann festgestellt werden, dass der Einführungsprozess dem Ablauf einer traditionellen Business-Intelligence-Implementierung ähnelt, den Schnittstellen jedoch eine signifikantere Bedeutung zuteilwird und die Formulierung der richtigen Fragen eine größere Herausforderung darstellt. Nach initialen Pilotierungen ist ein schrittweiser Aufbau von Analysen, Entscheidungshilfen und Prognosen typisch.

### **Entwicklung neuer Produkte auf Basis von Big Data**

Big-Data-Initiativen, die das Ziel verfolgen neue Produkte auf Basis von Big-Data-Technologien zu entwickeln, stellen bezüglich des Projektmanagements spezielle Anforderungen, die nicht von jedem Vorgehensmodell bedient werden können. Grundsätzlich ist festzustellen, dass es sich hierbei nicht um eine Optimierungsstrategie handelt, sondern mit Hilfe innovativer Produkte zusätzliche Umsatzpotentiale erschlossen werden sollen. Deshalb kann der Einführungsprozess auch als ein Produktentwicklungsprozess angesehen werden, welcher wiederum andere Kriterien in den Vordergrund stellt. Damit neuartige Big-Data-Produkte entwickelt werden können, muss zunächst eine geeignete Big-Data-Infrastruktur zur Datenexploration bereitgestellt werden. Darauf aufbauend werden in einem iterativen Prozess Produktentwürfe ausgearbeitet, welche unter Umständen auch wieder verworfen werden können. Um Fehlentwicklungen zu vermeiden, ist eine interdisziplinäre Zusammenarbeit mit der Marketingabteilung und dem Produktmanagement zur zielgerichteten Ermittlung von Kundenbedürfnissen anzustreben. Diese Faktoren führen dazu, dass sich die Produkthanforderungen im Verlauf der Big-Data-Einführung mehrfach ändern können und das Projektende zu Beginn kaum abgeschätzt werden kann. Dies wiederum stellt einen hohen Anspruch an die Flexibilität des angewandten Vorgehensmodells. Organisatorisch sind Einführungsprojekte dieser Art eher in einem vergleichbar kleinen Rahmen innerhalb eines abgeschlossenen Geschäftsbereichs oder des Produktentwicklungsbereichs anzusiedeln. Die mitunter ausgeprägte Produktkomplexität erfordert ein hohes Qualifikationsniveau im Projektteam. Insbesondere die ausreichende Anzahl an Datenwissenschaftlern spielt hier eine wichtige Rolle. Eine Strategie dieser Art ist nur bedingt auf andere Bereiche übertragbar, was den Anspruch an die Skalierbarkeit einschränkt. Nichtsdestotrotz kann sich eine Organisation langfristig immer weiter auf Big-Data-Produkte spezialisieren. Damit ein Einführungsprojekt wie im vorliegenden Betrachtungsfall gelingt, ist eine zielorientierte Projektleitung durch erfahrene Führungskräfte, die sowohl technisches als auch betriebswirtschaftliches Verständnis vorweisen können, hilfreich.

## 4.2 Vergleich agiler und klassischer Vorgehensmodelle

Dieses Kapitel beinhaltet die Entwicklung eines Kriterienkatalogs, welcher die quantitative Analyse von Vorgehensmodell-Typen des IT-Projektmanagements bei verschiedenen Big-Data-Strategien erlaubt. Weiterhin wird untersucht, inwieweit agile und klassische Vorgehensmodelle die definierten Kriterien erfüllen. Zusätzlich wird die Bedeutsamkeit einzelner Kriterien für die jeweiligen Strategien analysiert, um zusammenfassend im Rahmen einer Nutzwertanalyse zu berechnen, welches Vorgehensmodell für welche Big-Data-Strategie am vorteilhaftesten ist.

### 4.2.1 Kriterienkatalog und Ausprägungen

Die Auswahl der Kriterien und deren Ausprägungsmöglichkeiten haben maßgeblichen Einfluss auf das Ergebnis einer Nutzwertanalyse. Deshalb ist es notwendig einen ganzheitlichen Blick auf das Projektmanagement bei der Einführung von Big-Data-Management-Systemen zu werfen. Die aus der Literatur gewonnenen Kenntnisse zu typischen Vergleichskriterien im Projektmanagement und Vorgehensmodellen aus Kapitel 3.5.1 werden hierzu kombiniert und den Analyseanforderungen entsprechend angepasst.

Durch die Kombination der Erkenntnisse von Holzbaur, Leffingwell sowie Boehm und Turner ergeben sich für den Kriterienkatalog insgesamt 14 Vergleichskriterien, die sich in vier Kriteriengruppen klassifizieren lassen. Grundsätzlich sind sämtliche Kriterien im Kontext von IT-Projekten zur Einführung von Big-Data-Management-Systemen zu sehen. Die Ausprägung eines Kriteriums kann, abhängig vom Erfüllungsgrad des Vorgehensmodells, von sehr gut bis sehr schlecht variieren. Insgesamt sind bei jedem Kriterium fünf ordinal skalierte Abstufungen möglich. Die Ordinalskala lässt wiederum keine Aussage über die Abstände der Abstufungen zueinander zu, sondern liefert lediglich ein Urteil darüber, ob die jeweilige Ausprägung besser oder schlechter als eine andere Ausprägung ist. Erfüllt beispielsweise ein klassisches Vorgehensmodell ein Kriterium ausgesprochen gut, nimmt es für das entsprechende Kriterium die Ausprägung „sehr gut“ ein und erhält dafür die Maximalpunktzahl vier. Ist der Erfüllungsgrad für das Kriterium hingegen auffallend schwach, erhält das Vorgehensmodell die Ausprägung „sehr schlecht“ und somit die Minimalpunktzahl null. In Tabelle 8 wird ein zusammenfassender Überblick über die im Folgenden dargestellten Kriterien und deren Ausprägungsmöglichkeiten gegeben.



Gruppe	#	Bezeichnung	Ausprägungen					
			Wertung	sehr gut	gut	mittel	schlecht	sehr schlecht
			Punkte	4	3	2	1	0
Ergebnis	1	<b>Termintreue</b>	Terminein-haltung ist realistisch	leichte Termin-überschreitung realistisch (= 5% d. Projektdauer)	mittlere Termin-überschreitung realistisch (= 5%-30% d. Projektdauer)	schwere Termin-überschreitung realistisch (>30% d. Projektdauer)	unkalkulierbare Termin-überschreitung realistisch	
	2	<b>Kundenzufriedenheit</b>	sehr zufrieden, Produkt übertrifft Erwartungen	zufrieden, Produkt entspricht Erwartungen	geht so, leichte Produkt-abweichungen	unzufrieden, substantielle Produkt-abweichungen	sehr unzufrieden, Projektabbruch	
	3	<b>Kosteneinhaltung</b>	Budget-einhaltung ist realistisch	leichte Budget-überschreitung realistisch (=5% d. Budgets)	mittlere Budget-überschreitung realistisch (=5%-30% d. Budgets)	schwere Budget-überschreitung realistisch (>30% d. Budgets)	unkalkulierbare Budget-überschreitung realistisch	
Management	4	<b>Flexibilität</b>	Anforderungs-änderungen erzeugen minimalen Zusatzaufwand	Anforderungs-änderungen erzeugen geringen Zusatzaufwand	Anforderungs-änderungen erzeugen mittleren Zusatzaufwand	Anforderungs-änderungen erzeugen massiven Zusatzaufwand	Anforderungs-änderungen nicht möglich	
	5	<b>Planung und Kontrolle</b>	Sehr gute Vorhersagbar-keit der Projekt-aktivitäten	gute Vorhersagbar-keit der Projekt-aktivitäten	zufrieden-stellende Vorhersagbarkeit der Projekt-aktivitäten	kaum Überdeckung von Planung und Durchführung	Planung und Kontrolle nicht möglich	
	6	<b>Risikomanagement</b>	Risiken als zentrales Steuerungs-instrument	Risiken haben Einfluss auf Planung und Steuerung	Risiken sind bekannt aber kaum berücksichtigt	Risiken kaum bekannt	Risiken nicht berücksichtigt	
	7	<b>Data Governance und Datensicherheit</b>	Data Governance / Datensicherheit als ein Primärziel	Data Governance / Datensicherheit haben Einfluss auf Planung	Data Governance / Datensicherheit werden berücksichtigt	Data Governance / Datensicherheit werden kaum berücksichtigt	keine Berücksichti-gung von Data Governance / Datensicherheit	
Harte Faktoren	8	<b>Produktkomplexität</b>	hohe Komplexität über alle Technologie-schichten realisierbar	hohe Komplexität in Teilbereichen realisierbar	mittlere Komplexität über alle Technologie-schichten realisierbar	mittlere Komplexität in Teilbereichen realisierbar	nur einfache, kleinere Produkte realisierbar	
	9	<b>Skalierbarkeit</b>	uneingeschränkte Skalierbarkeit der Projektgröße gegeben	Skalierbarkeit der Projektgröße in weiten Teilen gegeben	eingeschränkte Skalierbarkeit der Projektgröße gegeben	kaum Skalierbarkeit der Projektgröße gegeben	keine Skalierbarkeit der Projektgröße gegeben	
	10	<b>Qualifikationsniveau</b>	sehr ausgewogenes Qualifikations-niveau im Projektteam	ausgewogenes Qualifikations-niveau im Projektteam	tendenziell hoch-bzw. niedrig-qualifizierte Projekt-teilnehmer	sehr hoch- bzw. niedrig-qualifizierte Projekt-teilnehmer	ausschließlich hoch- bzw. niedrig-qualifizierte Projekt-teilnehmer	
	11	<b>Kritikalität</b>	Produkt bei Fehler verantwortlich für viele Menschen-leben	Produkt bei Fehler verantwortlich für ein Menschen-leben	Produkt bei Fehler verantwortlich für hohen Geldverlust	Produkt bei Fehler verantwortlich für geringen Geldverlust	Keine besondere Verantwortung bei Fehler	
	12	<b>Transparenz</b>	sehr gute Nachvollziehbar-keit der Abläufe	gute Nachvollziehbar-keit der Abläufe	mäßige Nachvollziehbar-keit der Abläufe	kaum Nachvollziehbar-keit der Abläufe	undurchschau-bares Chaos	
Weiche Faktoren	13	<b>Kreativität</b>	Kreativität erwünscht und gefördert	Kreativität erwünscht aber kaum gefördert	Kreativität nur begrenzt gewünscht	Kreativität unerwünscht	Kreativität wird sanktioniert	
	14	<b>Kommunikation</b>	offene Kommunikation, motivierendes Umfeld, Selbst-verantwortung	offene Kommunikation, motivierendes Umfeld	offene Kommunikation	gelenkte Kommunikation	strikte Kommunikations-richtlinien, keine Freiheiten	

Tabelle 8: Kriterienkatalog mit Kriterienausprägungen

Die erste Kriteriengruppe beinhaltet Faktoren, die sich von Holzbaurs sogenanntem magischen Projektdreieck ableiten lassen und direkt das Projektergebnis bzw. den Projekterfolg beeinflussen.

### **Termintreue**

Die Erfolgsmessung bei Projekten ist, wie in Kapitel 2.4.3 am Beispiel von Big-Data-Projekten illustriert, eine erhebliche Herausforderung im Projektmanagement. Wie aus der Definition des Projektbegriffs in Kapitel 3.1 hervorgeht, besitzt jedes Projekt einen eindeutigen Start- und Endtermin. Eine Beurteilung des Projekterfolgs auf Basis der Termintreue liegt somit nahe. Das hier angeführte Kriterium Termintreue stellt das Pendant zur Zeitdimension in Holzbaurs magischem Projektdreieck dar. Es gestattet eine Aussage darüber, inwieweit das jeweilige Vorgehensmodell dazu in der Lage ist, den geplanten Projekttermin einzuhalten. Um diese Fähigkeit adäquat bewerten zu können, repräsentieren die Abstufungen der Ausprägungen die realistische Eintrittswahrscheinlichkeit der Einhaltung oder Überschreitung des geplanten Endtermins. Das Optimum dieses Kriteriums ist entgegen der Erwartung nicht die realistische Chance auf Unterschreitung des geplanten Endtermins. Dies wäre auf eine mangelhafte Projektplanung zurückzuführen. Vielmehr bedeutet eine besonders gute Ausprägung und hohe Punktzahl bei diesem Kriterium, dass die Einhaltung des geplanten Endtermins realistisch ist. Eine markant schwache Ausprägung hingegen bedeutet, dass der geplante Endtermin mit hoher Wahrscheinlichkeit unkalkulierbar überschritten wird. Dies wiederum hat eine geringe Punktzahl zur Folge. Unkalkulierbarkeit bedeutet in diesem Kontext eine Überschreitung des Endtermins um einen nicht absehbaren Zeitraum, der keine Planung mehr erlaubt und das Projekt mit hoher Wahrscheinlichkeit scheitern lässt.

### **Kundenzufriedenheit**

Die Kundenzufriedenheit kann grundsätzlich in verschiedenen Dimensionen sowohl objektiv als auch subjektiv gemessen und analysiert werden. Sie spiegelt implizit die Qualität des gelieferten Produkts wider und ist somit analog zur Termintreue ein maßgeblicher Einflussfaktor des Projekterfolgs. Aus diesem Grund findet sich die Kundenzufriedenheit in der Kriteriengruppe Ergebnis wieder. Indirekt kann die Kundenzufriedenheit bekanntermaßen über Beschwerdeanalysen gemessen werden. Im vorliegenden Kriterienkatalog ist die Kundenzufriedenheit jedoch, in Anlehnung an Holzbaurs Dimension Ergebnis im Projektdreieck, definiert als Grad der Erfüllung der Erwartungen des Kunden an das zu entwickelnde Produkt. Im positivsten Fall ist der Kunde sehr zufrieden und das Produkt übertrifft die Erwartungen. Im Kontext von Big Data sind dies beispielsweise unerwartet gute und präzise Analyseergebnisse. Dem gegenüber ist die negativste Ausprägung ein Projektabbruch in Verbindung mit einem sehr unzufriedenen Kunden.

### **Kosteneinhaltung**

Eine dritte grundlegende Dimension zur Beurteilung des Projekterfolgs ist die Ressourcennutzung. Sie repräsentiert die Verhältnismäßigkeit des geleisteten Projektaufwands in Relation zum entstandenen Ergebnis. Erfolgreiche Projekte charakterisiert, neben der hohen Termintreue und der überdurchschnittlichen Kundenzufriedenheit, die zielgerichtete Budgetplanung und -einhaltung. Eine Budgetunterschreitung wiederum wäre ein Indikator für fehlerhafte Planung und nicht ausgenutzte Potentiale. Die im Projektmanagement gängigste Form von Aufwand sind Kosten. Das dritte Kriterium der Gruppe Ergebnis bewertet somit die monetären Ressourcen eines Big-Data-Einführungsprojekts. In der Regel werden die Kosten von Projekten im Rahmen von Budgets abgebildet. Untersuchungsgegenstand ist im vorliegenden Fall die Fragestellung, inwiefern das Vorgehensmodell die Einhaltung des geplanten Budgets gewährleistet. Die Kriterienausprägungen können zwischen sehr gut, was einer realistischen Chance auf Budgeteinhaltung entspricht, und sehr schlecht, was wiederum unkalkulierbaren Budgetüberschreitungen gleichkommt, differieren. Auch hier bedeutet die Unkalkulierbarkeit eine nicht mehr planbare Überschreitung des Budgets, verbunden mit einer hohen Projektabbruchwahrscheinlichkeit. Ebenso wie die beiden vorangegangenen Kriterien ist auch die Dimension der Ressourcen von Holzbaurs magischem Projektdreieck abgeleitet.

Die zweite Gruppe beinhaltet Kriterien, welche das operative Management des Projekts betreffen. Sie stellen in erster Linie ein Maß für die Steuerbarkeit von Vorgehensmodellen dar.

### **Flexibilität**

Die Flexibilität, als erstes Kriterium der Gruppe Management, spielt im Kontext von Softwareentwicklungsprojekten eine herausragende Rolle. Nahezu kein IT-Projekt kommt ohne Anforderungsänderungen im Projektverlauf aus. Dem Kunden ist in vielen Fällen zu Projektbeginn noch nicht vollständig bewusst, welche Eigenschaften und Fähigkeiten das fertige Produkt haben soll. Durch die, wie in Kapitel 2.1 beschriebene, fortlaufende Erschließung neuer Datenquellen und das ständig wachsende Datenaufkommen tritt diese Problematik insbesondere im Kontext von Big-Data-Projekten auf. Im vorliegenden Kriterienkatalog ist die Flexibilität definiert als die Fähigkeit, auf Anforderungsänderungen zu reagieren. Die Reaktionsfähigkeit äußert sich dabei primär in dem zusätzlich notwendigen Aufwand zur Implementierung von Änderungswünschen. Der Aufwand indes kann sowohl in den genannten Erfolgsdimensionen Zeit, Qualität und Ressourcen als auch in erhöhtem Koordinationsaufwand anfallen. Eine Abstufung des notwendigen Aufwands stellt somit eine hinreichende Skalierung der Ausprägungen dar. Damit ein Vorgehensmodell die maximale Punktzahl erreicht, muss eine Anforderungsänderung mit minimalem Zusatzaufwand

bearbeitet werden können. Die schlechteste Bewertung wird vergeben, sobald das Modell generell nicht in der Lage ist Änderungen der Produktspezifikation zu berücksichtigen.

### **Planung und Kontrolle**

Um ein Projekt erfolgreich zum Ziel zu führen, ist eine hinreichende Disposition der Planung und Kontrolle von Projektaktivitäten fundamental wichtig. Nicht jedes Vorgehensmodell stellt diese Eigenschaft gleichermaßen gut zur Verfügung. Eine gute Planung und Kontrolle zeichnet sich durch eine detaillierte Vorhersagbarkeit von Aktivitäten im Projektverlauf aus. Es geht im Kern darum, inwieweit ein Vorgehensmodell Aufschluss darüber gibt, zu welchem Zeitpunkt und in welchem Ausmaß welche Projektaktivitäten durchgeführt werden. Zudem wird untersucht, wie stark die tatsächliche Projektdurchführung von der initialen Planung abweicht. Im besten Fall befähigt das Vorgehensmodell die Projektleitung dazu, schon sehr früh sehr genau vorherzusagen zu können wann welche Projektaktivitäten anfallen. Im schlechtesten Fall hingegen ist die Planung und Kontrolle gänzlich nicht möglich, da es keine Überdeckung von Planung und Durchführung gibt.

### **Risikomanagement**

Risiken spielen bei der Durchführung von Projekten eine wichtige Rolle, da sie maßgeblich Einfluss auf das Projektergebnis haben können. Sie treten in unterschiedlichen Dimensionen, wie z.B. Technologie- oder Organisationsrisiken, auf. Unabhängig von der Risikoausprägung sollte ein Vorgehensmodell idealerweise ein angemessenes Management der Risiken gewährleisten. Vor dem Hintergrund der Einführung eines Big-Data-Management-Systems sind insbesondere die in Kapitel 2.5.1 genannten Gefahren einer vernachlässigten Data-Governance und Datensicherheit als Risiken zu berücksichtigen. Der Einfluss von Risiken auf die Planung und Durchführung von IT-Projekten dieser Art fließt über das Kriterium Risikomanagement in den Kriterienkatalog ein. Es bewertet ein Vorgehensmodell dahingehend, dass die Untersuchung und Wirkung von Risiken berücksichtigt und beurteilt werden. Ist das Kriterium sehr gut ausgeprägt, stellt das Risikomanagement das zentrale Steuerungsinstrument des jeweiligen Vorgehensmodells dar. Andersherum wird das Kriterium sehr schwach bewertet, wenn Risiken gänzlich unbekannt sind und vernachlässigt werden.

### **Data Governance und Datensicherheit**

Data Governance und Datensicherheit sind bei Softwareentwicklungsprojekten und insbesondere bei Big-Data-Management-Systemen eine bedeutende Herausforderung. Dies ist, wie in Kapitel 2.5.1 beschrieben, auf das hohe Datenvolumen sowie die erhebliche Komplexität der Daten im Big-Data-Zeitalter zurückzuführen. Das Kriterium erlaubt eine Bewertung dahingehend, wie Data Governance und Datensicherheit im Rahmen eines Vor-

gehensmodells berücksichtigt werden können. Ist das Vorgehensmodell fähig, einen sehr hohen Stellenwert von Data Governance und Datensicherheit zu gewährleisten, so erhält es eine entsprechend positive Ausprägung. Kann dies nicht sichergestellt werden und die Data Governance findet keine Beachtung, erfolgt eine schlechte Bewertung. Dieses Kriterium ist bei Big-Data-Projekten besonders dann von Bedeutung, wenn Daten aus zahlreichen internen und externen Quellen zur Analyse herangezogen werden.

Die dritte Kriteriengruppe enthält sogenannte harte Faktoren nach Holzbaur. Zudem beinhaltet sie Aspekte, die Boehm und Turner als maßgebliche Unterscheidungskriterien von klassischen und agilen Vorgehensmodellen bezeichnen. Sie beziehen sich im Wesentlichen auf das Rahmenwerk sowie die strukturellen Eigenschaften des Projekts.

### **Produktkomplexität**

Der Begriff der Komplexität besitzt, obwohl jeder Mensch eine individuelle, grundlegende Vorstellung davon hat, keine einheitliche Definition. Die Ausprägungen von Komplexität unterscheiden sich in Abhängigkeit von der betrachteten Disziplin. Im Projektmanagement entsteht Komplexität in erster Linie durch die Struktur und die Anzahl der Elemente sowie deren Relationen untereinander. Zudem ist die Unsicherheit bei der Planung für die Komplexität von Belang. Das vorliegende Kriterium Produktkomplexität bezieht sich hingegen vollständig auf einen Big-Data-Kontext. Es ist ein Indikator dafür, inwieweit ein Vorgehensmodell in der Lage ist eine hohe Komplexität über alle Technologieschichten des in Kapitel 2.3.2 skizzierten Big-Data-Stapels abzubilden. Dies wiederum ist direkt verknüpft mit der Planbarkeit von Aktivitäten und der Transparenz von Prozessen im Projekt. Im besten Fall kann das Vorgehensmodell über alle sechs Schichten eine hohe technologische Komplexität realisieren. Die Negativausprägung des Kriteriums hingegen spricht dem Vorgehensmodell lediglich die Fähigkeit zu, kleine und einfache Produkte zu realisieren, die bei Weitem nicht alle Technologieschichten umfassen.

### **Skalierbarkeit**

Die Größe eines Projekts hängt von zahlreichen Faktoren ab. Maßgeblich dafür sind z.B. die Unternehmensgröße, das Projektziel und im Kontext von Big Data auch die verfolgte Strategie. Nicht jedes Vorgehensmodell kann alle Projektgrößen gleichermaßen effizient abbilden. Das Kriterium Skalierbarkeit gibt Auskunft darüber, wie variabel ein Vorgehensmodell in Bezug auf die realisierbare Projektgröße ist. Prinzipiell kann nahezu jede Projektgröße mit jedem Vorgehensmodell umgesetzt werden. Sie unterscheiden sich jedoch in dem zusätzlich notwendigen Aufwand zur Realisation einer anderen Projektgröße. Eine uneingeschränkte Skalierbarkeit ohne zusätzlichen organisatorischen Aufwand stellt dabei das Optimum dar. In diesem Fall ist es völlig indifferent, ob ein Vorgehensmodell mit

wenigen oder mit vielen Projektteilnehmern durchgeführt wird. Das negative Extrem ist gegeben, wenn das Vorgehensmodell ausschließlich auf eine bestimmte Projektgröße beschränkt ist und diesbezüglich keine Variation erlaubt.

### **Qualifikationsniveau**

Das Qualifikationsniveau der am Projekt beteiligten Mitarbeiter ist ein weiteres Kriterium der Gruppe harter Faktoren. Üblicherweise wird ein höheres Qualifikationsniveau positiver bewertet als ein niedriges. Im Sinne des Gesamterfolgs eines Projekts, welcher maßgeblich mit einer effizienten Ressourcennutzung zusammenhängt, wird bei dem vorliegenden Hintergrund von dieser Sichtweise abgesehen. Das Kriterium Qualifikationsniveau bewertet Vorgehensmodelle entsprechend der Ausgewogenheit der zur Anwendung notwendigen Qualifikationsprofile. Das bedeutet, dass ein Vorgehensmodell, welches den Anforderungen entsprechend mit einer Mischung aus hoch- und niedrigqualifizierten Mitarbeitern betrieben werden kann, anzustreben ist. Es wird somit indirekt auch die Komplexität der Methode beurteilt. Bei Big-Data-Projekten sind personelle Ressourcen, wie in Abschnitt 2.5.2 detailliert beschrieben, eine besondere Herausforderung. In vielen Fällen werden hochqualifizierte Datenwissenschaftler benötigt, um die Aufgabenstellung angemessen zu bearbeiten. In Bezug auf das vorliegende Kriterium ist jedoch festzustellen, dass je ausgewogener das nötige Qualifikationsniveau für ein Vorgehensmodell ist, desto besser wird es bewertet. Kann ein Vorgehensmodell jedoch nur mit einem ausschließlich hoch- bzw. niedrigqualifizierten Team funktionieren, so wird dieses Kriterium einen negativen Ausschlag geben.

### **Kritikalität**

Die Kritikalität als Kriterium zur Bewertung von Vorgehensmodellen entstammt den Überlegungen von Boehm und Turner und erlaubt ein Urteil darüber, wie sicher Produkte sind, die mit dem begutachteten Vorgehensmodell entwickelt werden. Die Sicherheit eines Produkts zu messen ist im Allgemeinen auf viele Arten möglich. Boehm und Turner schlagen zum Vergleich von agilen und klassischen Modellen, wie im Abschnitt 3.5.1 beschrieben, eine Einteilung nach Höhe des Verlustes bei Produktfehlschlägen vor. Die Auswirkungen möglicher Produktfehlschläge tragen sie auf einer individuell entwickelten Skala ab. Dort ist es als sehr positiv zu werten, wenn ein Vorgehensmodell ein hohes Maß an Entwicklungssicherheit gewährleistet, sodass die Endprodukte verantwortlich für mehrere Menschenleben sein können. Dies erfordert detaillierte und gut dokumentierte Entwicklungsarbeit auf einem sehr hohen Qualitätsniveau. Dem gegenüber ist es schlecht, wenn die Sicherheit des Vorgehensmodells lediglich die Entwicklung von Produkten erlaubt, die keinerlei Verantwortung für Fehler übernehmen können, da es bei ihnen häufig zu Produktfehlern kommt. Eine hoch sicherheitskritische Anwendung wäre im Zusammenhang mit Big Data beispielsweise eine Software, die durch Massendatenanalyse Stauwarnzeichen auf digitalen Verkehrsschildern

kontrolliert. Ein Produktfehler hätte in diesem Fall womöglich den Verlust vieler Menschenleben zur Folge.

### **Transparenz**

Die Transparenz im Projektmanagement hat direkten Einfluss auf die Planbarkeit und Kontrollierbarkeit der Aktivitäten. Dies wiederum hat in vielen Fällen eine unmittelbare Wirkung auf den Projekterfolg. Das vorliegende Kriterium der Transparenz gibt Auskunft darüber, wie gut die Nachvollziehbarkeit der Projektabläufe im Rahmen eines Vorgehensmodells ist. Die Nachvollziehbarkeit kann beispielsweise durch detaillierte Dokumentation verbessert werden. Dieser Logik zufolge ist es positiv zu bewerten, wenn ein Vorgehensmodell hohe Transparenz gewährleistet, da es die notwendigen Werkzeuge und Prozesse zur Herstellung von Transparenz in sich trägt. Auf der anderen Seite erhält ein Vorgehensmodell eine schlechte Bewertung, wenn keine Transparenz vorliegt, sondern ein durch die Methodik begründetes undurchschaubares Chaos. Ein solcher Zustand liegt vor, wenn das Vorgehensmodell keinerlei Dokumentation der Projektaktivitäten vorschreibt und durch die Komplexität der Prozesse im Projekt eine retrograde Nachvollziehbarkeit nicht gewährleistet werden kann.

Die vierte und letzte Kategorie von Kriterien umfasst die sogenannten weichen Faktoren nach Holzbaur. Primär geht es hierbei darum, auch schwer messbare weiche Faktoren mit in die Analyse einzubeziehen.

### **Kreativität**

Weiche Faktoren sind nach Holzbaur nur schwer quantifizierbar und hängen in den meisten Fällen von den am Projekt beteiligten Individuen ab. Damit bei der Analyse ein ganzheitlicher Blick in alle Bereiche des Projektmanagements bereitgestellt werden kann, sollen auch solche Faktoren den Einzug in den Kriterienkatalog finden. Kreativität ist der Definition nach die Fähigkeit etwas Neues, vorher nicht Dagewesenes zu kreieren. Im Kontext von Vorgehensmodellen im IT-Projektmanagement geht es jedoch weniger um die individuelle Fähigkeit, sondern vielmehr darum, inwiefern das Vorgehensmodell Rahmenbedingungen für Kreativität schafft. Das Kriterium Kreativität ist vor diesem Hintergrund somit definiert als der Grad der erwünschten und geförderten Kreativität in Vorgehensmodellen. Es gilt dabei zu untersuchen, ob und inwiefern ein Vorgehensmodell Kreativität benötigt oder verhindert. Bei Big-Data-Projekten spielt Kreativität oft eine bedeutende Rolle, um innovative und womöglich riskante Wege beschreiten zu können. Ist Kreativität essentieller Bestandteil des Vorgehensmodells, so erhält es aus diesem Grund eine sehr gute Bewertung. Im Gegensatz dazu können Vorgehensmodelle Kreativität womöglich sogar sanktionieren. Dies führt zu einer entsprechend schlechten Ausprägung des Kriteriums.

## Kommunikation

Neben dem Freiraum für Kreativität stellt die Kommunikation in Projekten einen zweiten wichtigen weichen Faktor dar, der sich nur indirekt quantifizieren lässt. Eine offene Kommunikation in einem motivierenden Umfeld wirkt sich insbesondere bei interdisziplinären Big-Data-Projekten sehr positiv aus und reduziert die Gefahr eines Projektfehlschlags um ein Vielfaches. Boehm und Turner fügen, wie in Abschnitt 3.5.1 dargestellt, in ihrer Vergleichsdimension Kultur an, dass Projektteilnehmer auch durch klare Strukturen, Rollen und Verantwortlichkeiten motiviert werden können. Im Kontext von Big-Data-Initiativen, bei denen ein hohes Maß an Flexibilität, Kreativität und Kommunikation gefordert ist, kann dies jedoch nicht gelten. Das vorliegende Kriterium Kommunikation beurteilt somit, ob und inwiefern ein Vorgehensmodell die Voraussetzungen schafft, um Kommunikation zu ermöglichen. Im besten Fall besitzt ein Vorgehensmodell die Prämisse offener, freier Kommunikation in einem motivierenden Umfeld und einem hohen Maß an Selbstbestimmung. Dies stellt gleichzeitig die Idealausprägung dar und führt zur Maximalpunktzahl. Auf der anderen Seite kann es sein, dass ein Vorgehensmodell keine Freiheiten erlaubt, sämtliche Kommunikation in strikten Richtlinien geregelt hat und jegliche Selbstbestimmung unterdrückt. Dies wiederum ist als die negativste Ausprägung des Kriteriums anzusehen. Die in Kapitel 3.3 und 3.4 beschriebenen klassischen und agilen Vorgehensmodelle unterscheiden sich diesbezüglich im Erfüllungsgrad deutlich voneinander. Ein Kernparadigma agiler Methoden ist die Interaktion zwischen Individuen, wohingegen klassische Modelle strukturiertere Ansätze verfolgen.

### 4.2.2 Kriterien-Erfüllungsgrade agiler und klassischer Vorgehensmodelle

Die im vorangegangenen Abschnitt definierten Kriterien werden in diesem Kapitel herangezogen, um agile und klassische Vorgehensmodelle isoliert voneinander auf ihren jeweiligen Erfüllungsgrad zu bewerten. Dabei wird bei allen Kriterien untersucht, inwieweit das Vorgehensmodell das einzelne Kriterium mit welcher Ausprägung erfüllt. Die Bewertungen basieren auf den in Kapitel 3 erarbeiteten Erkenntnissen zu Vorgehensmodellen im IT-Projektmanagement. Abhängig von der zutreffenden Ausprägung wird dem Vorgehensmodell für jedes Kriterium eine Punktzahl von null bis vier zugewiesen. Die zugewiesenen Werte  $\omega_{i,k}$  gehen anschließend in die Nutzwertanalysen für die jeweiligen Big-Data-Strategien ein. Die nun folgenden Plausibilisierungen sind zusammengefasst in Tabelle 9 dargestellt.



Gruppe	#	Bezeichnung		Klassische Vorgehensmodelle	Agile Vorgehensmodelle	
		Wertung	Punkte $\omega_{i,k}$			
Ergebnis	1	<b>Termintreue</b>		Durch strukturierte Planung ist eine Termineinhaltung realistisch	Durch zahlreiche Anforderungsänderungen ist eine Verzögerung zwischen 5% und 30% realistisch	
		Wertung / Punkte		gut	3	mittel
	2	<b>Kundenzufriedenheit</b>		Da nur geringe Flexibilität gewährleistet werden kann, sind Produktabweichungen realistisch	Durch die hohe Flexibilität und Kundenorientierung ist es realistisch, dass das Produkt den Erwartungen entspricht	
Wertung / Punkte			mittel	2	gut	3
Management	3	<b>Kosteneinhaltung</b>		Leichte Budgetüberschreitungen durch Nachbesserungen und unvorhersehbare Ereignisse realistisch	Leichte Budgetüberschreitungen durch hohe Anzahl an anfänglich nicht bekannten Anforderungen realistisch	
		Wertung / Punkte		mittel	2	mittel
	4	<b>Flexibilität</b>		Da Projektplanung nur einmal zu Projektbeginn durchgeführt wird, und Änderungen nicht geplant sind, sehr unflexibel	Durch zahlreiche Iterationen und Kundenfeedbacks können Anforderungsänderungen mühelos bearbeitet werden	
Wertung / Punkte			schlecht	1	sehr gut	4
Management	5	<b>Planung und Kontrolle</b>		Durch strukturierte Planung der Arbeitspakete sind die Aktivitäten sehr gut vorhersehbar	Vorhersagbarkeit der Aktivitäten eingeschränkt gegeben, da sich Aktivitätenschwerpunkte im Projektverlauf verlagern können	
		Wertung / Punkte		sehr gut	4	mittel
	6	<b>Risikomanagement</b>		Risiken nehmen Einfluss auf den Projektplanungsprozess und werden im Projektverlauf berücksichtigt	Risiken können durch die hohe Dynamik leicht in den Hintergrund rücken	
Wertung / Punkte			gut	3	mittel	2
Management	7	<b>Data Governance und Datensicherheit</b>		Data Governance und Datensicherheit können durch gute Planbarkeit berücksichtigt werden	Durch den inkrementellen Entwicklungsprozess sind Herausforderungen der Data Governance und Datensicherheit schwer zu bewältigen	
		Wertung / Punkte		mittel	2	schlecht
	Management	8	<b>Produktkomplexität</b>		Modell ausgelegt für extensive Produkte. Hohe Komplexität ist mindestens in Teilbereichen realisierbar	Vorgehensmodell eher für technisch einfachere, kleinere Produkte geeignet. Mittlere Komplexität realisierbar
Wertung / Punkte				gut	3	mittel
Management		9	<b>Skalierbarkeit</b>		Das Vorgehensmodell lässt sich durch strukturierten Aufbau ohne großen Aufwand in seiner Größe skalieren	Da sehr hohe Dynamik und beträchtlicher Kommunikationsbedarf besteht, ist die Skalierbarkeit kaum gegeben
	Wertung / Punkte			gut	3	schlecht
	Management	10	<b>Qualifikationsniveau</b>		Mit Ausnahme der Projektplanungsphase ist während des Projekts ein ausgewogenes Qualifikationsniveau hinreichend	Durch hohe Selbstverantwortung und teils autonome Planungsaktivitäten ist ein tendenziell hohes Qualifikationsniveau nötig
Wertung / Punkte				gut	3	mittel
Management		11	<b>Kritikalität</b>		Vorgehensmodelle durch hohen Dokumentationsgrad und transparente Struktur höchst verantwortungsvoll einsetzbar	Aufgrund hoher Dynamik und mäßiger Transparenz sind Produkte nur bedingt verantwortungsvoll einsetzbar
	Wertung / Punkte			sehr gut	4	mittel
	Management	12	<b>Transparenz</b>		In der Struktur verankerte Dokumentationspflichten gewährleisten gute Nachvollziehbarkeit der Abläufe	Nachvollziehbarkeit nur bedingt gegeben durch hohen Grad an Dynamik sowie eine potentiell hohe Anzahl von Anforderungsänderungen
Wertung / Punkte				gut	3	mittel
Management		13	<b>Kreativität</b>		Durch strikte Projektstruktur kommt der Kreativität prinzipiell kein großer Stellenwert zu	Projektteam wird aktiv ermutigt kreative Ideen in den Entwicklungsprozess einfließen zu lassen
	Wertung / Punkte			schlecht	1	sehr gut
	Management	14	<b>Kommunikation</b>		Projektkommunikation wird durch Projektleitung gelenkt und strukturiert	Kernelement des Vorgehensmodells ist die offene Kommunikation in einem motivierenden Umfeld mit hoher Selbstverantwortung
Wertung / Punkte				schlecht	1	sehr gut

Tabelle 9: Kriterien-Erfüllungsgrade agiler und klassischer Vorgehensmodelle

Die Fähigkeit zur Einhaltung des geplanten Endtermins variiert bei klassischen und agilen Vorgehensmodellen stark. So ist es bei klassischen Modellen durchaus realistisch, dass die geplante Projektdauer nur leicht überschritten wird. Dies ist primär durch die klare Struktur und die gute Planbarkeit der eindeutig voneinander abgrenzbaren Phasen begründet. Es lässt sich somit insgesamt zielgerichtet auf den Endtermin hinarbeiten. Dies wiederum hat jedoch Einfluss auf andere Aspekte, wie im Folgenden noch deutlich wird. Agile Vorgehensmodelle hingegen tendieren durch ihre ausgeprägte Flexibilität und eine potentiell hohe Anzahl an Anforderungsänderungen während des Projektverlaufs dazu, den geplanten Endtermin in einem erhöhten Maß zu überschreiten. Durch die inkrementelle Vorgehensweise und den überdurchschnittlichen Kommunikationsbedarf während der Entwicklung kann es schnell zu Verzögerungen kommen. Diese wiederum führen dazu, dass Entwicklungsinhalte nicht zum geplanten Zeitpunkt fertiggestellt werden und das Projekt den Endtermin überschreitet.

Die Kundenzufriedenheit, als zweites Ergebniskriterium, ist stark davon abhängig, wie das Vorgehensmodell die Kundenwünsche in die Produktentwicklung mit einfließen lassen kann. Bei klassischen Modellen werden in der Regel zu Projektbeginn auf Basis der vom Kunden gelieferten Spezifikationen genaue Anforderungen an das Endprodukt definiert. Dieser Anforderungskatalog stellt gleichzeitig die Grundlage für die gesamte Projektplanung dar. Möchte der Kunde seine Spezifikationen im Projektverlauf aufgrund neuer Erkenntnisse noch einmal ändern, so ist dies durch die geringe Flexibilität maximal in frühen Phasen möglich. Dies führt häufig dazu, dass durch leichte Produktabweichungen die Erwartungen nicht vollends erfüllt werden. Agile Vorgehensmodelle dagegen versuchen diese Schwäche auszuräumen und sind bereits durch das Agile Manifest dazu verpflichtet die Kundenorientierung in den Vordergrund zu stellen. Durch die konsequente und regelmäßige Einbindung des Kunden können Anforderungsänderungen identifiziert werden und neue Spezifikationen in das Produkt einfließen. Diese Flexibilität stellt sicher, dass das fertige Produkt den Kundenerwartungen entspricht.

Die Wahrscheinlichkeit, dass das geplante Projektbudget eingehalten wird, ist bei beiden Vorgehensmodell-Typen gleichermaßen gering. Sowohl bei klassischen als auch bei agilen Modellen sind leichte Budgetüberschreitungen realistisch. Dies ist jedoch durch verschiedene Faktoren begründet. So führen bei klassischen Modellen Aufwände durch unvorhersehbare Ereignisse und Produktnachbesserungen dazu, dass die Kosten höher als geplant ausfallen. Es ist bei komplexen Big-Data-Projekten praktisch nicht möglich sämtliche Eventualitäten in einem zu Projektbeginn definierten Projektplan abzubilden. Dies führt, genau wie die nicht hinreichend planbaren, im Projektverlauf aufkommenden Kundenanforderungen, zu zusätzlichen Kosten. Bei agilen Vorgehensmodellen spielt die mangelnde Termintreue eine ausschlaggebende Rolle für die Budgetüberschreitung. Viele ungeplante

Anforderungsänderungen können häufig nur durch zusätzlichen Kapazitäteneinsatz entwickelt werden.

Die Flexibilität steht in Wechselwirkung mit den Erfüllungsgraden zahlreicher anderer Kriterien. Sie ist ein Maß dafür, wieviel zusätzlichen Aufwand Anforderungsänderungen im Projektverlauf des jeweiligen Vorgehensmodells erzeugen. Hierbei unterscheiden sich klassische und agile Vorgehensmodelle fundamental voneinander. Agile Methoden sind hervorragend dafür ausgelegt Anforderungsänderungen im laufenden Projektbetrieb einfließen zu lassen. Durch zahlreiche Iterationen, hohe Dynamik und die enge Kundeneinbindung in den Entwicklungsprozess sind Änderungen grundlegender Bestandteil der Logik des Vorgehensmodells und verursachen nur minimalen Zusatzaufwand. Dem gegenüber stehen klassische Vorgehensmodelle mit statischen, genau geplanten Projektplänen, die im schlechtesten Fall nur einmal zu Projektbeginn auf Basis der Kundenspezifikation erarbeitet werden. Insbesondere dadurch, dass bei klassischen Modellen erst vergleichsweise spät laufende Software zum Testen zur Verfügung steht und genau dort häufig Anforderungsänderungen und Anpassungswünsche auftauchen, entstehen hier massive Zusatzaufwände durch Neuplanungen.

Die Möglichkeit ein Projekt planvoll und zielgerichtet steuerbar zum Erfolg zu führen und eine gute Vorhersagbarkeit der Projektaktivitäten zu haben, ist dem Kriterium Planung und Kontrolle nach als positiv zu bewerten. Dies trifft ebenso übergreifend auf Projekte im Themenfeld Big Data zu. Klassische und agile Vorgehensmodelle unterscheiden sich diesbezüglich jedoch deutlich. Durch ihren strukturierten Aufbau sowie klar voneinander abgrenzbare Phasen und Arbeitspakete erlauben klassische Modelle eine ausgesprochen gute Planung und Kontrolle. Projektpläne können detailliert ausgearbeitet werden, Rollen und Verantwortlichkeiten sind klar zugeordnet und die Steuerbarkeit ist durch gerichtete Kommunikationswege gegeben. Agile Modelle hingegen setzen andere Schwerpunkte. Die Vorhersagbarkeit der Aktivitäten ist durch die hohe Dynamik, Flexibilität und Kundenorientierung nur eingeschränkt gewährleistet.

In Bezug auf das Risikomanagement sind zwischen den Vorgehensmodell-Typen nur geringe Unterschiede zu verzeichnen. Tendenziell sind klassische Modelle jedoch besser in der Lage Risiken zu antizipieren. Der Grund dafür liegt primär in der sehr frühzeitig stattfindenden ganzheitlichen Planung des Projekts, bei der eine umfassende Berücksichtigung der Risiken unabdingbar ist. Kristallisieren sich erhebliche Risiken heraus, so ist eine Anpassung der Projektplanung durchaus realistisch. Agile Vorgehensmodelle agieren zwar auch auf Basis eines Projektplans, durch die hohe Dynamik im Entwicklungsprozess können jedoch schnell vorher unbekannte Aspekte in den Vordergrund rücken. Dies würde ein iteratives, sich dem Projekt anpassendes Risikomanagement erfordern, was hohen organisatorischen Zusatzaufwand mit sich bringt und deshalb nur bedingt zur Anwendung

kommt. Aus diesem Grund können bei agilen Modellen Risikoaspekte leichter vernachlässigt werden als bei klassischen Modellen.

Eine adäquate Data Governance und Datensicherheit spielt, wie in Abschnitt 2.5.1 detailliert ausgeführt, eine wichtige Rolle bei der Einführung von Big-Data-Management-Systemen. Die verfügbaren klassischen und agilen Vorgehensmodelle sind jedoch nur mittelbar in der Lage dieser Herausforderung gerecht zu werden. Insbesondere agile Modelle erschweren eine simultane Weiterentwicklung der bestehenden Data Governance und Datensicherheit durch ihren inkrementellen von häufigen Richtungsänderungen geprägten Projektverlauf. Klassische Vorgehensmodelle weisen zwar eine bessere Planbarkeit auf, vermeintliche Randthemen wie Data Governance und Datensicherheit, üben jedoch keinen entscheidenden Einfluss auf die Projektplanung aus. Trotzdem findet Data Governance als übergeordnete Leitlinie Berücksichtigung, sodass klassische Modelle in diesem Kontext einen höheren Erfüllungsgrad aufweisen.

In der Reihe der harten Faktoren liegt der Erfüllungsgrad klassischer Vorgehensmodelle in allen Kriterien über dem der agilen Vorgehensmodelle. Die Einstufungen sind zu einem großen Teil auf die in Kapitel 3.5.1 von Boehm und Turner definierten Optimalbedingungen für klassische und agile Methoden zurückzuführen. Die beste Ausprägung eines Kriteriums hingegen richtet sich danach, was für eine Big-Data-Management-System-Einführung am besten wäre. Eine hohe Produktkomplexität, als erster harter Faktor, kann nicht gleichermaßen gut von klassischen und agilen Modellen in Big-Data-Projekten realisiert werden. Da eine hohe Produktkomplexität in vielen Fällen direkt mit einer ausgedehnten Projektgröße zusammenhängt, sind hier klassische Modelle durch ihre bessere Planbarkeit und Kontrollierbarkeit eher heranzuziehen als agile Modelle. Letztere zeichnen sich vielmehr durch ihre ausgeprägte Dynamik und eine hohe Abhängigkeit vom impliziten Wissen der Projektteilnehmer aus. Aufgrund dessen sind sie primär für kleinere Produkte in kleineren Teams gut geeignet.

Letztgenannte Aspekte lassen sich auf den Erfüllungsgrad des Kriteriums Skalierbarkeit übertragen. Der Erfolg agiler Vorgehensmodelle ist in erster Linie davon abhängig, wie offen die Kommunikation ist, wie selbstverantwortlich die Teammitglieder agieren und wieviel Freiraum der Kreativität gegeben wird. Diese Eigenschaften können nur sehr eingeschränkt auf extensive Projektgrößen skaliert werden. Dem gegenüber gewährleisten klassische Vorgehensmodelle insbesondere bei Projekten mittleren und großen Ausmaßes sehr gute Skalierbarkeit. Lediglich bei sehr kleinen Projekten können sie aufgrund ihres hohen Dokumentations- und Koordinationsaufwands nur mit eingeschränkter Effizienz angewendet werden.

Das Kriterium Qualifikationsniveau, als Maß dafür, wie ausgewogen die Qualifikationsprofile des Projektteams in einem Vorgehensmodell sein können, fällt bei klassischen und agilen Methoden ebenfalls unterschiedlich aus. Agile Vorgehensmodelle erfordern eine ständige Anwesenheit von vergleichsweise hoch qualifizierten Kräften, die in der Lage sind selbstverantwortlich Arbeitspakete zu definieren und gleichzeitig mit der agilen Methodologie vertraut sind. Bei klassischen Modellen wiederum ist lediglich in der komplexen Projektinitiationsphase ein vergleichbar hohes Qualifikationsniveau erforderlich. Im operativen Projektbetrieb kann die Menge an hochqualifiziertem Personal reduziert werden, sodass sich bezüglich der Mitarbeiterprofile ein ausgewogenes Niveau einstellt.

Die Kritikalität des durchzuführenden Projekts ist ein weiterer von Boehm und Turner aufgestellter Erfolgsfaktor, dem agile und klassische Vorgehensmodelle unterschiedlich begegnen. Im Kontext von Big Data kann die Kritikalität als die Fähigkeit eines Vorgehensmodells, sicherheitskritische Produkte in Projekten entwickeln zu können, gesehen werden. Klassische Modelle haben einen aus der Methodik begründeten hohen Anspruch an eine detaillierte Dokumentation. Dies und die Tatsache, dass Sicherheitsroutinen in die Modellstrukturen implementiert werden können, haben zur Folge, dass eine insgesamt hohe Nachvollziehbarkeit der Aktivitäten gewährleistet werden kann. Dem gegenüber sind agile Modelle häufig tendenziell schlecht dokumentiert, sodass nur eingeschränkte Nachvollziehbarkeit gegeben ist. Dies hat insbesondere bei späteren Wartungsarbeiten am fertig entwickelten Produkt schlechte Auswirkungen. Begründet sind solche Dokumentationsmängel durch die hohe Anzahl an Anforderungsänderungen und den dynamischen Arbeitsprozess. Im gleichen Zuge erübrigt sich deshalb die Erläuterung der Erfüllungsgrade des Kriteriums Transparenz. Da die Transparenz im vorliegenden Kriterienkatalog über die Nachvollziehbarkeit der Abläufe bewertet wird, ist die Einstufung sehr ähnlich zur Kritikalität. Die Nachvollziehbarkeit wird in erster Linie über eine detaillierte Dokumentation bereitgestellt, was, wie bereits erörtert, klassische Vorgehensmodelle besser gewährleisten als agile.





Abschließend spiegeln die weichen Kriterien wider, inwiefern die Vorgehensmodelle emotionale und menschliche Einflussfaktoren berücksichtigen und fördern. Weiche Faktoren zeichnen sich dadurch aus, dass sie nur schwer messbar sind. Eine grundsätzliche Kategorisierung in Erfüllungsgradabstufungen lässt sich jedoch durchführen. Grundsätzlich ist festzustellen, dass agile Vorgehensmodelle in dieser Gruppe deutlich bessere Eigenschaften besitzen als klassische. Das Kriterium Kreativität steht in diesem Zusammenhang dafür, wie stark Kreativität im Projekt erwünscht ist und gefördert wird. Im Kontext von Big-Data-Projekten ist dies ein essentiell wichtiges Kriterium, um mit den innovativen Technologien zielgerichtet kundenorientierte Produkte zu entwickeln. In agilen Modellen, wie z.B. Scrum, werden Projektteilnehmer aktiv ermutigt kreative Ideen in den Entwicklungsprozess einfließen zu lassen. Es gibt keine strikten Vorgaben daran, wie das Zielprodukt ausgestaltet

sein muss. Vielmehr wird in vielen Iterationen im Rahmen eines kreativen Prozesses ein optimal auf die Kundenwünsche angepasstes Produkt entworfen. Durch ihre sehr klare, in Phasen unterteilte Projektstruktur sowie bereits sehr früh umfassend definierten Arbeitspakete sind klassische Vorgehensmodelle diesbezüglich deutlich schwächer anzusehen. Kreativität spielt eine untergeordnete Rolle und ist bestenfalls in der initialen Ausarbeitung der Projektplanung auf Basis der gelieferten Kundenspezifikationen notwendig.

Eine ähnliche Einstufung gilt für die Kommunikation in Vorgehensmodellen. Ein Kernparadigma agiler Modelle ist die Interaktion zwischen Individuen. Entsprechend offen ist das Kommunikationsklima im Projektteam. In einem motivierenden Umfeld müssen die Projektmitarbeiter eigenverantwortlich Aufgabenpakete definieren, abarbeiten und vor der nächsten Iteration dem Kunden präsentieren. Durch häufiges Feedback, sowohl von methodischen Assistenten als auch von Kunden, wird somit ein umfassender Rahmen für die Kommunikation geschaffen. Dies ist insbesondere bei interdisziplinären Big-Data-Projekten, bei denen unterschiedliche Vorkenntnisse der Projektteilnehmer vorliegen, ein wichtiger Faktor. Klassische Vorgehensmodelle erfüllen dieses Kriterium analog zum Kriterium Kreativität nur schlecht. Die Gründe dafür sind im methodischen Aufbau des Modells zu finden. Die strikte Struktur erfordert klare Kommunikationswege, die auf den Funktionszuordnungen im Projekt basieren. Die Projektleitung nimmt dabei eine Lenkungsfunktion ein, sodass offene Kommunikation gänzlich nicht bereitgestellt werden kann.

#### **4.2.3 Kriterien-Bedeutsamkeiten für einzelne Big-Data-Strategien**

Zur Ermittlung der Gewichtung der einzelnen Kriterien ist es notwendig deren Bedeutsamkeit für jede der vier zu analysierenden Big-Data-Strategien zu bewerten. Jedes Kriterium kann dabei als wichtig, neutral oder unwichtig für die jeweilige Strategie eingestuft werden. Diese Beurteilung dient als weitere Grundlage zur Durchführung der in Kapitel 3.5.2 erläuterten Nutzwertanalyse. Auf Basis der Einstufungen werden zur Bestimmung der Kriterienbewertungen Paarvergleiche der Kriterien untereinander durchgeführt. Die Kriterien-Bedeutsamkeiten für die in Kapitel 2.4 vorgestellten Big-Data-Strategien sind in Tabelle 10 zusammenfassend dargestellt.

Gruppe	#	Bezeichnung	Big-Data-Strategien			
			Kosteneinsparung	Zeiteinsparung	Entscheidungsunterstützung	Neue Produkte
						
Ergebnis	1	Termintreue	+	+	0	-
	2	Kundenzufriedenheit	0	0	+	+
	3	Kosteneinhaltung	+	0	0	-
Management	4	Flexibilität	-	-	0	+
	5	Planung und Kontrolle	+	+	0	0
	6	Risikomanagement	0	0	+	0
	7	Data Governance und Datensicherheit	-	-	+	+
Harte Faktoren	8	Produktkomplexität	-	+	-	+
	9	Skalierbarkeit	+	0	0	-
	10	Qualifikationsniveau	0	0	-	-
	11	Kritikalität	+	0	0	0
	12	Transparenz	0	+	+	-
Weiche Faktoren	13	Kreativität	-	0	+	+
	14	Kommunikation	-	-	0	+

**Bewertungsrichtlinie**  
 + = wichtig  
 0 = neutral  
 - = unwichtig

Tabelle 10: Kriterien-Bedeutsamkeiten für einzelne Big-Data-Strategien

### Kosteneinsparung durch Big Data

Die Bedeutsamkeit der einzelnen Kriterien für die Strategie Kosteneinsparung basiert auf den in Kapitel 2.4.1 beschriebenen Charakteristika der Strategie sowie den in Abschnitt 4.1 getroffenen Annahmen zum Einführungsprozess eines Big-Data-Management-Systems unter dieser Prämisse. Zur effizienten Einführung eines Big-Data-Management-Systems unter dem Leitmotiv der Kosteneinsparung kristallisieren sich auf dieser Basis somit fünf Kriterien als besonders wichtig heraus. Da diese Strategie aus reiner Kostensicht verfolgt wird, sind insbesondere Ergebniskriterien und harte Faktoren von hoher Bedeutung. Vor dem Hintergrund, dass eine Organisation mit dem Big-Data-System ausschließlich Kosten reduzieren möchte, liegt eine strikte Kosten- und Terminpolitik schon während der Projekteinführung nahe. Aus diesem Grund muss das verwendete Vorgehensmodell hinreichend in der Lage sein, vornehmlich diese beiden Ergebniskriterien zu erfüllen. Hieraus lässt sich ableiten, dass zur Erreichung der Termin- und Kostenziele im Projekt eine strukturierte Planung und Kontrolle notwendig ist. Weiterhin sind die Produktspezifikationen des Zielsystems bereits zu

Projektbeginn weitestgehend bekannt. Dieser Informationsvorteil kann bessere Projektergebnisse induzieren, sofern eine zielgerichtete Planung und Kontrolle der Projektaktivitäten durch das Vorgehensmodell gewährleistet werden kann. Aus der Reihe der harten Faktoren sind eine gute Skalierbarkeit des Vorgehensmodells sowie eine ausgeprägte Fähigkeit zur Entwicklung sicherheitskritischer Produkte von hoher Bedeutsamkeit. Erstere ist durch die Tatsache begründet, dass die Ausmaße der im Rahmen der Kosteneinsparung durch Big-Data-Technologien zu ersetzenden Datenarchitekturen von Organisation zu Organisation stark variieren können. Da bei der Implementierung der Aufwand für das Projektmanagement mit der Größe des Systems ansteigt, ist eine gute Skalierbarkeit von übergeordneter Relevanz. Die hohe Bedeutung der Kritikalität resultiert daraus, dass das Big-Data-System das traditionelle System in einigen Fällen sogar vollständig ersetzt und somit auch die gesamte Verantwortung für die Sicherung und den Erhalt hochsensibler Unternehmensdaten übernimmt.

Die für die Strategie der Kosteneinsparung unwichtigen Kriterien entstammen verschiedenen Gruppen und sind auf unterschiedliche Ursachen zurückzuführen. Die Flexibilität, als Maß dafür, wie viel Zusatzaufwand durch eine Anforderungsänderung generiert wird, ist für eine Kosteneinsparungsstrategie wenig relevant, da die Anwendung der Big-Data-Technologien sich ausschließlich auf die Speicherung von Daten reduziert. Dies führt zu vergleichsweise wenig komplexen Datenarchitekturen, die in ihrer logischen Struktur den traditionellen Systemen ähneln. Daraus resultiert eine gute Definierbarkeit des Endprodukts, was wiederum den Anspruch an die Flexibilität in den Hintergrund rücken lässt. Aus dem gleichen Grund sind die Kriterien Data Governance und Datensicherheit sowie Produktkomplexität von untergeordneter Bedeutung. Da das zu entwickelnde Big-Data-Produkt den gleichen Zweck verfolgt wie die traditionellen Systeme, können Data-Governance-Strukturen und Richtlinien zur Datensicherheit in weiten Teilen übernommen werden. Weiterhin bleibt der Komplexitätsgrad bei dieser Strategie weit hinter den Möglichkeiten zurück, die eine Big-Data-Technologie, wie z.B. ein Hadoop-Cluster, mit sich bringen könnte. Weiche Faktoren, wie die Kreativität und die Kommunikation sind aus genannten Gründen ebenso unmaßgeblich. Vielmehr könnte ausgeprägte Kreativität und freie, unstrukturierte Kommunikation zu Projektverzögerungen und fehlgeleiteten Entwicklungen führen, die für sensible Big-Data-Systeme dieser Art bedrohlich sind. Alle weiteren Kriterien sind von neutraler Bedeutung. Der Erfüllungsgrad der untersuchten Vorgehensmodelle wirkt sich bei jenen neutralen Kriterien lediglich in einem moderaten Maße entweder positiv oder negativ auf die Eignung für diese Big-Data-Strategie aus.



### **Zeiteinsparung mit Hilfe von Big Data**

Die Strategie der Zeiteinsparung setzt bezüglich der Bedeutsamkeit einzelner Kriterien andere Schwerpunkte, hat jedoch auch einige Schnittstellen mit der Kosteneinsparungsstrategie. Auch hier gehen die Bewertungen zurück auf die in Abschnitt 2.4.2 aufgestellten Charakteristika der Strategie nach Davenport sowie auf die für den Einführungsprozess in Kapitel 4.1 getroffenen Annahmen. Aus der Definition der Strategie heraus, Zeit einsparen zu wollen, kann abgeleitet werden, dass die durchführende Organisation viel Wert auf effiziente Zeitnutzung legt. Daraus ergibt sich eine hohe Bedeutsamkeit der Termintreue. Im vorliegenden Kontext wird die Strategie, mit Hilfe von Big Data Zeit einzusparen, ausschließlich performanceorientiert untersucht. Das bedeutet, dass unter Anwendung komplexer Big-Data-Hochleistungssysteme, wie z.B. der In-Memory-Technologie, bekannte Prozesse beschleunigt werden. Projekte dieser Art haben einen aus ihrer Komplexität begründeten hohen Anspruch an Planung und Kontrolle sowie die realisierbare Produktkomplexität. Dieser Anspruch wird durch die Notwendigkeit, interdisziplinäre Teams zu koordinieren, ergänzt. Die Komplexität resultiert dabei in erster Linie daraus, Big-Data-Systeme und traditionelle Systeme sowohl auf Hardwareebene als auch auf Softwareebene zu harmonisieren. Die Interdisziplinarität rührt aus der Tatsache, dass zumeist wenige funktionsgebundene Prozesse der Fachabteilungen beschleunigt werden. Das fachliche Wissen zu derartigen Prozessen sowie die späteren Anwendungsfälle liegen in den einzelnen Abteilungen. Dies wiederum führt zu einem hohen Anspruch an die Transparenz. Eine Verbreitung der Big-Data-Expertise von der IT-Abteilung in die Fachbereiche erfordert eine sehr gute retrograde Nachvollziehbarkeit der Projektaktivitäten, die primär über detaillierte Dokumentationstätigkeiten gewährleistet werden kann.

Analog zur Strategie der Kosteneinsparung durch Big Data wird auch hier das Kriterium Flexibilität als vergleichsweise unwichtig eingestuft. Die Gründe dafür unterscheiden sich jedoch geringfügig. Bei der Kosteneinsparung durch Big Data resultiert die geringe Notwendigkeit an Flexibilität aus der guten Definierbarkeit des Endprodukts. Diese wiederum ist zurückzuführen auf eine geringe Komplexitätsstufe. Demgegenüber geht die gute Definierbarkeit des Endprodukts bei der Zeiteinsparung mit Big Data aus dem detaillierten Wissen über den zu beschleunigenden Prozess hervor und nicht aus der niedrigen Komplexität. Es bedarf somit auch hier nur geringfügig der Fähigkeit, Anforderungsänderungen möglichst aufwandsarm implementieren zu können. Die Ursachen für die geringe Bedeutsamkeit des Kriteriums Data Governance und Datensicherheit sind jedoch identisch mit denen der Kosteneinsparungsstrategie. Die Zielstrukturen ähneln in ihrem logischen Aufbau sehr stark den Ausgangsstrukturen, sodass ein Projektmanagement unter dem Leitmotiv der Data Governance nicht zwingend notwendig ist. Schließlich wird der weiche Faktor Kommunikation als verhältnismäßig unbedeutend eingestuft. Die inter-

disziplinären Projektteilnehmer mit unterschiedlichen Vorkenntnissen in Bezug auf Big-Data-Technologien haben die Aufgabe ein hochkomplexes Produkt zu entwickeln. Klare und intensive, jedoch auch gut koordinierte Kommunikation ist deshalb notwendig, um effizient zum Ziel zu kommen. Da das Kriterium Kommunikation nicht den Grad der Kommunikation, sondern die Art der Kommunikation bewertet, kommt dem Kriterium vor diesem Hintergrund eine geringe Bedeutung zu. Alle weiteren Kriterien liegen in ihrer Bedeutsamkeit zwischen den genannten wichtigen und unwichtigen Faktoren und werden aus diesem Grund als neutral eingestuft.

### **Entscheidungsunterstützung und Prognosen durch Big Data**

Die Strategie auf Basis von Big-Data-Technologien bessere Entscheidungen treffen zu können sowie leistungsfähigere Prognosen zu berechnen, wurde in Kapitel 2.4.3 bereits mit Hilfe von Anwendungsbeispielen charakterisiert. Verschiedene Annahmen über den Einführungsprozess eines Big-Data-Management-Systems dieser Art wurden in Abschnitt 4.1 getroffen. Basierend auf den erarbeiteten Erkenntnissen ergeben sich für die Kriterien des Kriterienkatalogs unterschiedliche Bedeutsamkeiten. So ist die Kundenzufriedenheit als Ergebniskriterium von hoher Bedeutung, da bei dieser Strategie die kundenspezifische Ausgestaltung des Produkts, die sich in der Wirksamkeit der Analysen äußert, im Fokus steht. Interne und externe Daten unterschiedlicher Quellen und heterogener Struktur werden zur Entwicklung detaillierter Analysen kombiniert und aggregiert. Insbesondere bei der Verwendung externer Daten können massive Risiken auftreten, welche die Organisation unter Umständen bedeutend schädigen können. Aus diesem Grund kommt einem adäquaten Risikomanagement eine große Bedeutung zu. Aus einem ähnlichen Grund, nämlich der bisher nicht getätigten Zusammenführung diverser Datenmengen aus unterschiedlichen Datenquellen, muss die Data Governance laufend überprüft werden. Im gleichen Zuge besteht das Risiko, dass durch den Zugriff auf externe Quellen die interne Datensicherheit gefährdet wird. Welche Daten auf welche Art und Weise zusammengeführt werden, muss während des Projekts zum Teil auch experimentell erörtert werden, sodass der Kreativität im Vorgehensmodell eine hohe Bedeutsamkeit zukommt. Deckungsgleich mit der klassischen Business Intelligence werden die Analysen in Zusammenarbeit mit den Fachbereichen entwickelt und angewandt. Eine detaillierte Nachvollziehbarkeit der Projektaktivitäten ist für die Rekapitulation der Auswertungen und Prognosen unerlässlich, was dazu führt, dass bei dieser Strategie viel Wert auf Transparenz gelegt wird.

Da es sich bei den zu entwickelnden Big-Data-Management-Systemen im Grunde um eine aktualisierte Version klassischer Business-Intelligence-Systeme handelt und deren Strukturen hinlänglich bekannt sind, ist die Fähigkeit eines Vorgehensmodells hohe Komplexität abzubilden nur von geringer Bedeutung. Die Systeme arbeiten zwar auf einer

gänzlich neuen Plattform, die Routinen und Abläufe ähneln sich jedoch. Darüber hinaus werden nicht gezwungenermaßen alle Technologieschichten des in Kapitel 2.3.2 skizzierten Big-Data-Stapels implementiert. Da das Kriterium Qualifikationsniveau als Ausprägung den Grad der Ausgewogenheit des Qualifikationsniveaus im Projektteam festlegt und keine Aussage über die notwendige Höhe des Niveaus zur erfolgreichen Durchführung des Projekts erlaubt, wird es in diesem Kontext als unbedeutend eingestuft. Nichtsdestotrotz können Projekte, die das Ziel der verbesserten Entscheidungsfindung und Prognose mit Hilfe von Big Data verfolgen, nur mit einer kritischen Menge an hochqualifizierten Projektmitarbeitern zum Erfolg geführt werden.

### **Entwicklung neuer Produkte auf Basis von Big Data**

Die Entwicklung neuer Produkte auf Basis von Big-Data-Technologien als vierte analysierte Strategie unterscheidet sich grundlegend von den vorangegangenen Strategien. Nachdem bisher größtenteils bekannte Prozesse und Analysen mit Hilfe von Big Data optimiert wurden, geht es bei dieser Strategie darum, marktorientierte und verkaufsfähige Big-Data-Produkte zu entwickeln. Der Einführungsprozess ähnelt, wie in Kapitel 4.1 festgestellt, somit vielmehr einem klassischen Produktentwicklungsprozess. Anwendungsbeispiele dieser Strategie wurden bereits in Abschnitt 2.4.4 illustriert und damit gleichzeitig die Eigenschaften dieser Vorgehensweise beschrieben. Den Kern der Entwicklung von Big-Data-Produkten bildet die ausgesprochen ausgeprägte Kundenorientierung. Durch Einbindung von Marketing und Produktmanagement in den Entwicklungsprozess kann maximale Marktorientierung gewährleistet werden. Eine überdurchschnittliche Kundenorientierung bringt in der Regel einen hohen Anspruch an die Flexibilität des Projektmanagements mit sich. Dies ist dadurch begründet, dass nicht jede Entwicklung erfolgreich am Markt platziert werden kann und sich Kundeninteressen während des Entwicklungsprozesses mehrfach ändern können. Häufig können innovative Produkte nur durch Verknüpfung von Daten bisher ungenutzter Quellen, wie z.B. große Mengen anonymisierter Kundendaten, hervorgebracht werden. Dies wiederum erfordert die durchgängige Berücksichtigung und Weiterentwicklung der Data Governance und Datensicherheit, welche für diese Strategie als wichtig angesehen wird. Eine weitere Möglichkeit Big-Data-Produkte zu entwickeln, ist die Fähigkeit komplexe Systeme zu vereinfachen und für Anwender nutzbar zu gestalten. Mitunter sind hochkomplexe Produkte über alle Technologieschichten des Big-Data-Stapels zu entwickeln, was die hohe Bedeutung des Kriteriums in diesem Kontext rechtfertigt. Unerlässlich für Projekte dieser Art ist ein kreatives, motivierendes Umfeld, offene Kommunikation sowie hohe Selbstverantwortung der Projektmitarbeiter. Die Projektkultur muss Fehler zugunsten der kreativen Entfaltung verzeihen, da nur so mutige und innovative Produktideen zutage treten. Dies führt dazu, dass die Kriterien Kreativität und Kommunikation von großer Bedeutung für diese Strategie sind.

Die Kriterien der Gruppe Ergebnis sind bis auf die Kundenzufriedenheit von untergeordneter Bedeutung. Dies rührt daher, dass es bei der Entwicklung primär darum geht ein marktfähiges Produkt auf Basis von Big Data hervorzubringen. Eine Produktentwicklung kann z.B. aufgrund von falsch antizipierten Kundenwünschen durchaus länger dauern als ursprünglich geplant. Dies wiederum kann höhere Kosten zur Folge haben. Um den finanziellen Schaden möglichst gering zu halten, werden Projekte dieser Art in der Regel in kleineren Umgebungen durchgeführt. Genau hierin liegt gleichzeitig die Ursache, dass die Skalierbarkeit des Vorgehensmodells von geringer Bedeutsamkeit ist. Die Bewertung des Qualifikationsniveaus ist analog zur Strategie der Entscheidungsverbesserung durch die Definition des Kriteriums als unwichtig eingestuft. Trotzdem ist ein überdurchschnittliches Qualifikationsprofil der Projektmitarbeiter notwendig, um alle Kundenwünsche in einem Produkt darstellen zu können. Eine ausreichende Zahl an interdisziplinär ausgebildeten Datenwissenschaftlern erhöht die Chance ein erfolgreiches Produkt zu entwickeln um ein Vielfaches. Die Transparenz, als Maß für die Nachvollziehbarkeit der Projektabläufe, ist bei der Entwicklung neuer Produkte auf Basis von Big Data anfänglich nur von geringer Bedeutung. Die Begründung dafür liegt in der Tatsache, dass es sich hierbei um einen flexiblen kreativen Prozess handelt, bei dem eine ausführliche Dokumentation hinderlich wäre. Die Herstellung und der Test lauffähiger Produkte sind in diesem Fall deutlich in den Vordergrund zu stellen. Alle weiteren Kriterien liegen in Bezug auf ihre Bedeutsamkeit zwischen genannten wichtigen und unwichtigen Faktoren und werden deshalb als neutral eingestuft.

#### **4.2.4 Auswertung im Rahmen von Nutzwertanalysen**

In diesem Abschnitt geht es darum die auf Basis der im zweiten und dritten Kapitel erarbeiteten Informationen, aufgestellten Annahmen und Plausibilisierungen der Abschnitte 4.2.1 bis 4.2.3 im Rahmen von Nutzwertanalysen zu verknüpfen. Dabei werden die vier, von Davenport definierten Big-Data-Strategien isoliert betrachtet und jeweils ein Nutzwert für klassische und agile Vorgehensmodelle berechnet. Die Methodik zur Berechnung der Nutzwerte ist identisch mit der in Kapitel 3.5.2 illustrierten Vorgehensweise der Nutzwertanalyse. Zunächst werden die im vorangegangenen Abschnitt 4.2.3 aufgestellten Kriterien-Bedeutsamkeiten durch Paarvergleiche der Kriterien untereinander zu Gewichtungen verrechnet. Zur Nachvollziehbarkeit der Ermittlung der Gewichtungen sind die Paarvergleiche zu allen vier Strategien den Anhängen 1 bis 4 zu entnehmen. Die berechneten Gewichtungen gehen im Anschluss in die Nutzwertanalyse ein und zeigen auf, welche Kriterien für die untersuchte Strategie besonders bedeutsam sind. Die in Kapitel 4.2.2 definierten Kriterien-Erfüllungsgrade der einzelnen Vorgehensmodelle werden als Prozentsatz vom Maximalwert mit der Gewichtung multipliziert und ergeben gemeinsam den Teilnutzwert des betrachteten Kriteriums. Aufsummiert führen alle Teilnutzwerte eines Vorgehensmodells zum Gesamtnutzwert, welcher schließlich als Vergleichswert der Alternativen

fungiert. Die Nutzwertanalysen der untersuchten Strategien werden entsprechend der geschilderten Logik im Folgenden detailliert dargestellt. Eine inhaltliche Diskussion der Ergebnisse wird im Rahmen von Handlungsempfehlungen in Kapitel 4.3 gegeben.

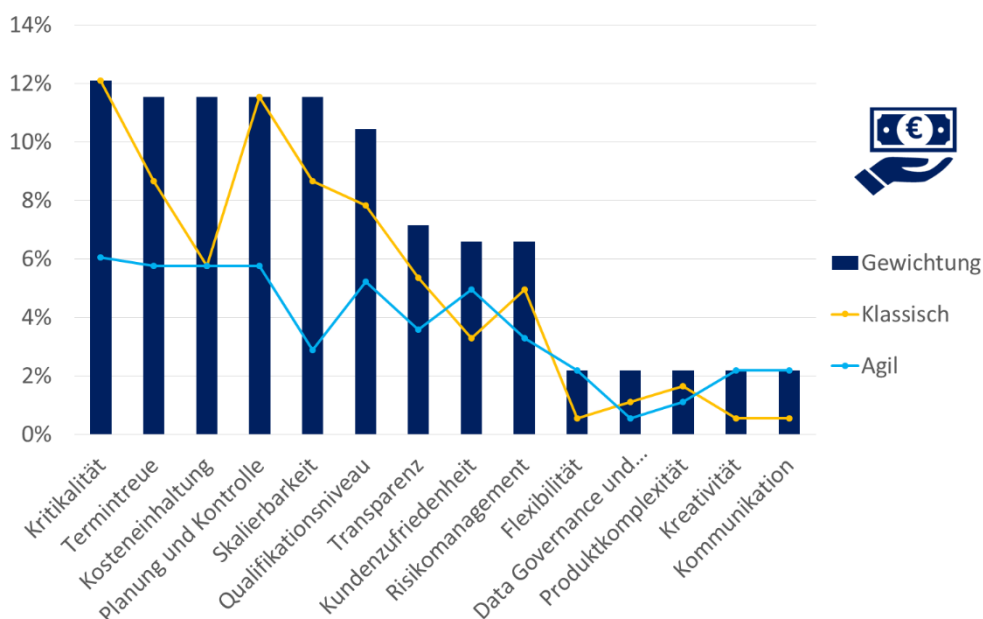
### Kosteneinsparung durch Big Data

Unter Berücksichtigung der in Kapitel 4.2.2 festgelegten und plausibilisierten Kriterien-Erfüllungsgrade  $\omega_{i,k}$  für klassische und agile Vorgehensmodelle sowie der in Kapitel 4.2.3 hergeleiteten Bedeutsamkeiten der einzelnen Kriterien ergibt sich für die Strategie der Kosteneinsparung die in Tabelle 11 dargestellte Nutzwertanalyse. Die detaillierte Ermittlung der Kriteriengewichtungen ist dem Paarvergleich für diese Strategie in Anhang 1 zu entnehmen.

	#	Kriterium	Gewichtung $G_k$	Klassische Vorgehensmodelle		Agile Vorgehensmodelle	
				Wert $\omega_{i,k}$ (max. 4)	Teilnutzwert $n_{i,k}$	Wert $\omega_{i,k}$ (max. 4)	Teilnutzwert $n_{i,k}$
Ergebnis	1	Termintreue	11,5%	3	8,7%	2	5,8%
	2	Kundenzufriedenheit	6,6%	2	3,3%	3	4,9%
	3	Kosteneinhaltung	11,5%	2	5,8%	2	5,8%
Management	4	Flexibilität	2,2%	1	0,5%	4	2,2%
	5	Planung und Kontrolle	11,5%	4	11,5%	2	5,8%
	6	Risikomanagement	6,6%	3	4,9%	2	3,3%
	7	Data Governance und Datensicherheit	2,2%	2	1,1%	1	0,5%
Harte Faktoren	8	Produktkomplexität	2,2%	3	1,6%	2	1,1%
	9	Skalierbarkeit	11,5%	3	8,7%	1	2,9%
	10	Qualifikationsniveau	10,4%	3	7,8%	2	5,2%
	11	Kritikalität	12,1%	4	12,1%	2	6,0%
	12	Transparenz	7,1%	3	5,4%	2	3,6%
Weiche Faktoren	13	Kreativität	2,2%	1	0,5%	4	2,2%
	14	Kommunikation	2,2%	1	0,5%	4	2,2%
Nutzwert $N_i$ in Bezug auf Idealprofil (=100%)				$N_1 =$	72,5%	$N_2 =$	51,5%

**Tabelle 11: Nutzwertanalyse - Kosteneinsparung**

Es zeigt sich, dass bei Anwendung der in Kapitel 3.5.2 definierten Methodik zur Nutzwertanalyse, klassische Vorgehensmodelle gegenüber agilen Modellen besser für die Einführung eines Big-Data-Management-Systems unter der Prämisse Kosteneinsparung geeignet sind. Der Nutzwert für klassische Vorgehensmodelle beträgt 72,5 Prozent in Bezug auf das Idealprofil von 100 Prozent. Im Vergleich dazu beläuft sich der Nutzwert agiler Vorgehensmodelle auf lediglich 51,5 Prozent. Relativ betrachtet sind die klassischen Modelle somit um 21 Prozentpunkte besser als die agilen.



**Abbildung 14: Erfüllunggrad der Teilnutzwerte - Kosteneinsparung**

Um diese deutliche Differenz in Bezug auf die relative Vorteilhaftigkeit auf granularer Ebene erläutern zu können, sind in Abbildung 14 die Erfüllungsgrade der Teilnutzwerte abgetragen. Entlang der Abzissenachse sind die Kriterien entsprechend ihrer aus dem Paarvergleich resultierenden Gewichtung absteigend sortiert.

Über die Gewichtungen sind die errechneten Teilnutzwerte der beiden Alternativen in gelb und blau abgetragen. Die größten Differenzen sind demnach bei den Kriterien Kritikalität, Planung und Kontrolle sowie Skalierbarkeit auszumachen. Dies ist zum einen durch die hohe Gewichtung dieser Kriterien für die Strategie der Kosteneinsparung zu begründen. Zum anderen erfüllen klassische Vorgehensmodelle diese Kriterien deutlich besser. Hinzu kommt, dass agile Modelle lediglich bei vier von vierzehn Kriterien bessere Teilnutzwerte aufweisen. Die deutliche Differenz der Gesamtnutzwerte lässt sich insbesondere hierauf zurückführen.

### **Zeiteinsparung mit Hilfe von Big Data**

Auch bei der Strategie der Zeiteinsparung kommt die bekannte Methodik aus Kapitel 3.5.2 zum Tragen. Die Bedeutsamkeit der einzelnen Kriterien unterscheidet sich jedoch von denen der anderen Strategien. Der Paarvergleich zur Umrechnung der in Abschnitt 4.2.3 bestimmten Bedeutsamkeiten in Gewichtungen ist in Anhang 2 beigefügt. Die daraus resultierende Berechnung der Nutzwerte ist Tabelle 12 zu entnehmen.

Bei der Strategie Zeit mit Hilfe von Big Data einzusparen, um so größtenteils bekannte Prozesse zu beschleunigen, kommt es zu einem ähnlichen Analyseergebnis wie bei der Kosteneinsparungsstrategie. Die klassischen Vorgehensmodelle sind hier mit einem Nutzwert von 71,8 Prozent deutlich den agilen Vorgehensmodellen mit einem errechneten Nutz-

wert von 54,1 Prozent zu bevorzugen. Die Differenz der Nutzwerte fällt in diesem Fall mit 17,7 Prozentpunkten jedoch etwas geringer aus als bei der Strategie der Kosteneinsparung. Dies ist mehrheitlich auf einen Anstieg des Nutzwerts der agilen Vorgehensmodelle um 2,6 Prozentpunkte zurückzuführen. Hinzu kommt der Rückgang des Nutzwerts der klassischen Modelle um 0,7 Prozentpunkte gegenüber den Ergebnissen der Nutzwertanalyse bei der Kosteneinsparungsstrategie.

	#	Kriterium	Gewichtung $G_k$	Klassische Vorgehensmodelle		Agile Vorgehensmodelle	
				Wert $\omega_{i,k}$ (max. 4)	Teilnutzwert $n_{i,k}$	Wert $\omega_{i,k}$ (max. 4)	Teilnutzwert $n_{i,k}$
Ergebnis	1	Termintreue	12,6%	3	9,5%	2	6,3%
	2	Kundenzufriedenheit	6,6%	2	3,3%	3	4,9%
	3	Kosteneinhaltung	6,6%	2	3,3%	2	3,3%
Management	4	Flexibilität	1,1%	1	0,3%	4	1,1%
	5	Planung und Kontrolle	12,6%	4	12,6%	2	6,3%
	6	Risikomanagement	6,6%	3	4,9%	2	3,3%
	7	Data Governance und Datensicherheit	1,1%	2	0,5%	1	0,3%
Harte Faktoren	8	Produktkomplexität	12,6%	3	9,5%	2	6,3%
	9	Skalierbarkeit	6,6%	3	4,9%	1	1,6%
	10	Qualifikationsniveau	6,6%	3	4,9%	2	3,3%
	11	Kritikalität	6,6%	4	6,6%	2	3,3%
	12	Transparenz	12,6%	3	9,5%	2	6,3%
Weiche Faktoren	13	Kreativität	6,6%	1	1,6%	4	6,6%
	14	Kommunikation	1,1%	1	0,3%	4	1,1%

Nutzwert $N_i$ in Bezug auf Idealprofil (=100%)	$N_1 =$	71,8%	$N_2 =$	54,1%
---	---------	-------	---------	-------

**Tabelle 12: Nutzwertanalyse - Zeiteinsparung**

Analog zur ersten Strategie sind in Abbildung 15 die Erfüllungsgrade der Teilnutzwerte abgetragen, um eine detailliertere Analyse des Ergebnisses zu ermöglichen. Die Kriterien sind auch hier in absteigender Reihenfolge entsprechend ihrer Gewichtung entlang der Abzisse sortiert. Als ausgesprochen wichtig gegenüber den anderen Kriterien haben sich demnach die Termintreue, die Planbarkeit und Kontrollierbarkeit, die realisierbare Produktkomplexität sowie die Nachvollziehbarkeit der Projektaktivitäten herauskristallisiert. Bei keinem dieser Kriterien erhalten agile Vorgehensmodelle einen besseren Teilnutzwert als klassische. Auch bei den moderat wichtigen Kriterien erhalten die agilen Vorgehensmodelle nur zwei Mal einen höheren Teilnutzwert. Insgesamt lässt sich die deutliche Differenz somit durch die genannten Faktoren begründen. Es ist zudem festzuhalten, dass sich die Nutzwertanalysen der Kosteneinsparungsstrategie und der Zeiteinsparungsstrategie in ihren Ausprägungen scheinbar stark ähneln. Betrachtet man die ausschlaggebenden Kriterien, so ist dies jedoch

nur bedingt der Fall. Nichtsdestotrotz ergeben die Summen der Teilnutzwerte nahezu identische Gesamtnutzwerte.

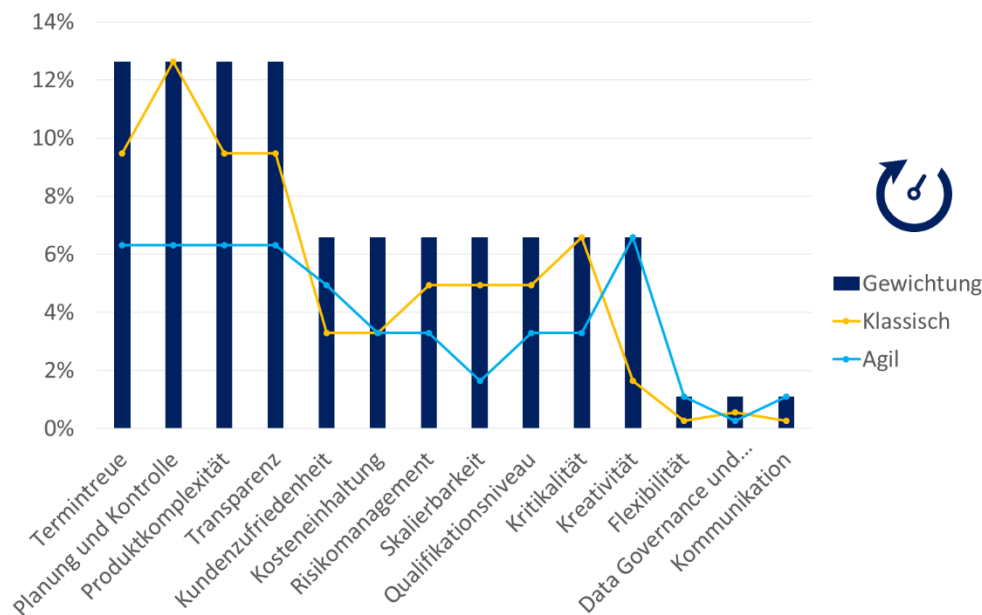


Abbildung 15: Erfüllungsgrad der Teilnutzwerte - Zeiteinsparung

### Entscheidungsunterstützung und Prognosen durch Big Data

Bei der Einführung von Big-Data-Management-Systemen mit dem Ziel bessere Entscheidungen zu treffen sowie belastbarere Prognosen aufzustellen kommt es in Bezug auf die Eignung von klassischen und agilen Vorgehensmodellen zu einem beachtenswerten Ergebnis. Unter Zuhilfenahme der gleichen Methodik wie bei den vorangegangenen Analysen und Verwendung der in Kapitel 4.2.3 für die vorliegende Strategie entwickelten Kriterien-Bedeutsamkeiten kommt es zu annähernd identischen Nutzwerten für beide Vorgehensmodell-Typen. Der zur Ermittlung der Kriteriengewichtungen durchgeführte Paarvergleich ist Anhang 3 zu entnehmen. Im Gesamtergebnis liegen agile Vorgehensmodelle mit einem Nutzwert von 59,6 Prozent nur äußerst knapp vor klassischen Vorgehensmodellen mit 59,3 Prozent. Die dazu durchgeführte Nutzwertanalyse ist im Detail in Tabelle 13 dargestellt. Die ausgesprochen geringe Differenz der Nutzwerte von 0,3 Prozentpunkten ist zwar beachtlich, jedoch ist gleichzeitig auch festzustellen, dass in Bezug auf das Idealprofil von 100 Prozent keins der beiden Vorgehensmodell-Typen hinreichend gut geeignet ist. Vielmehr werden bei der Verwendung sowohl klassischer als auch agiler Modelle in jedem Fall Kompromisse in Bezug auf diverse Kriterien gemacht werden müssen.



	#	Kriterium	Gewichtung $G_k$	Klassische Vorgehensmodelle		Agile Vorgehensmodelle	
				Wert $\omega_{i,k}$ (max. 4)	Teilnutzwert $n_{i,k}$	Wert $\omega_{i,k}$ (max. 4)	Teilnutzwert $n_{i,k}$
Ergebnis	1	Termintreue	5,5%	3	4,1%	2	2,7%
	2	Kundenzufriedenheit	12,1%	2	6,0%	3	9,1%
	3	Kosteneinhaltung	5,5%	2	2,7%	2	2,7%
Management	4	Flexibilität	5,5%	1	1,4%	4	5,5%
	5	Planung und Kontrolle	5,5%	4	5,5%	2	2,7%
	6	Risikomanagement	12,1%	3	9,1%	2	6,0%
	7	Data Governance und Datensicherheit	12,1%	2	6,0%	1	3,0%
Harte Faktoren	8	Produktkomplexität	1,1%	3	0,8%	2	0,5%
	9	Skalierbarkeit	5,5%	3	4,1%	1	1,4%
	10	Qualifikationsniveau	1,1%	3	0,8%	2	0,5%
	11	Kritikalität	5,5%	4	5,5%	2	2,7%
	12	Transparenz	12,1%	3	9,1%	2	6,0%
Weiche Faktoren	13	Kreativität	12,1%	1	3,0%	4	12,1%
	14	Kommunikation	4,4%	1	1,1%	4	4,4%
Nutzwert $N_i$ in Bezug auf Idealprofil (=100%)				$N_1 =$	59,3%	$N_2 =$	59,6%

Tabelle 13: Nutzwertanalyse - Entscheidungsunterstützung

Zur Visualisierung der Erfüllungsgrade der einzelnen Teilnutzwerte sind in Abbildung 16 die Kriteriengewichtungen und die Teilnutzwerte der Alternativen in bekannter Darstellungsform abgetragen.

Obwohl neun von vierzehn Teilnutzwerte der klassischen Vorgehensmodelle einen höheren Betrag aufweisen als die der agilen Vorgehensmodelle, sind beide Modelle für die Einführung von Big Data unter der Prämisse besserer Entscheidungsfindung etwa gleich gut geeignet. Aus Abbildung 16 wird ersichtlich, dass das Kriterium Kreativität, welches für die vorliegende Strategie von hoher Bedeutung ist, mit einem maßgeblichen Ausschlag zugunsten agiler Vorgehensmodelle zum Ergebnis der Analyse beiträgt. Bei keinem anderen Kriterium ist die Differenz der Teilnutzwerte vergleichbar groß. Die Vorgehensmodell-Typen wechseln bezüglich ihrer Dominanz bei allen weiteren Kriterien unsystematisch, sodass schließlich in der Gesamtsicht beinahe ein Gleichstand zustande kommt.

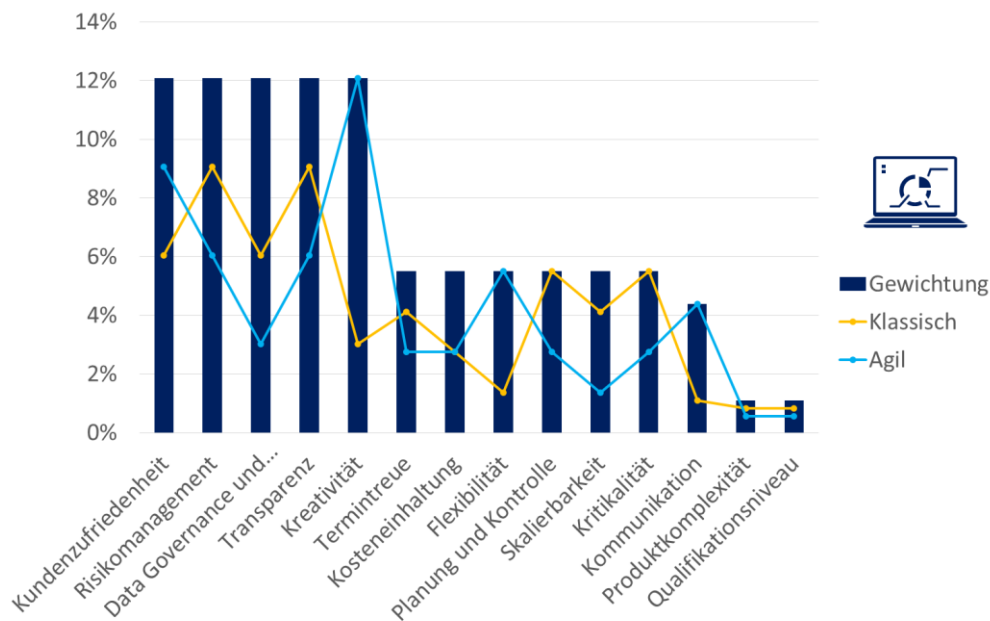


Abbildung 16: Erfüllunggrad der Teilnutzwerte - Entscheidungsunterstützung

### Entwicklung neuer Produkte auf Basis von Big Data

Ist das Ziel der Big-Data-Initiative die Entwicklung neuer Produkte und Dienstleistungen, die auf Big-Data-Technologien fußen, so sind klassische und agile Vorgehensmodelle unterschiedlich gut dafür geeignet. Auch bei dieser Strategie findet die in Abschnitt 3.5.2 skizzierte Nutzwertanalysen-Methodik ihre Anwendung. Zunächst sind dazu die für diese Strategie individuellen Kriterien-Bedeutsamkeiten im Rahmen des in Anhang 4 beigefügten Paarvergleichs zu Kriteriengewichtungen umgerechnet worden. Das Ergebnis der darauf folgenden Nutzwertanalyse ist Tabelle 14 zu entnehmen.

Es zeigt sich, dass agile Vorgehensmodelle für diese Strategie besser geeignet sind als klassische. Mit einem Gesamtnutzwert von 66,8 Prozent liegen die agilen Modelle mit 12,1 Prozentpunkten vor den klassischen Modellen mit einem Nutzwert von 54,7 Prozent. Bezogen auf das Idealprofil von 100 Prozent kann man von einer eindeutigen Vorteilhaftigkeit agiler Modelle sprechen, die jedoch nicht so klar wie bei der Kosteneinspar- und Zeiteinsparstrategie ausfällt.

	#	Kriterium	Gewichtung $G_k$	Klassische Vorgehensmodelle		Agile Vorgehensmodelle	
				Wert $\omega_{i,k}$ (max. 4)	Teilnutzwert $n_{i,k}$	Wert $\omega_{i,k}$ (max. 4)	Teilnutzwert $n_{i,k}$
Ergebnis	1	Termintreue	2,2%	3	1,6%	2	1,1%
	2	Kundenzufriedenheit	11,5%	2	5,8%	3	8,7%
	3	Kosteneinhaltung	2,2%	2	1,1%	2	1,1%
Management	4	Flexibilität	11,5%	1	2,9%	4	11,5%
	5	Planung und Kontrolle	6,6%	4	6,6%	2	3,3%
	6	Risikomanagement	6,6%	3	4,9%	2	3,3%
	7	Data Governance und Datensicherheit	11,5%	2	5,8%	1	2,9%
Harte Faktoren	8	Produktkomplexität	11,5%	3	8,7%	2	5,8%
	9	Skalierbarkeit	2,2%	3	1,6%	1	0,5%
	10	Qualifikationsniveau	2,2%	3	1,6%	2	1,1%
	11	Kritikalität	6,6%	4	6,6%	2	3,3%
	12	Transparenz	2,2%	3	1,6%	2	1,1%
Weiche Faktoren	13	Kreativität	11,5%	1	2,9%	4	11,5%
	14	Kommunikation	11,5%	1	2,9%	4	11,5%

Nutzwert $N_i$ in Bezug auf Idealprofil (=100%)	$N_1 =$	54,7%	$N_2 =$	66,8%
---	---------	-------	---------	-------

Tabelle 14: Nutzwertanalyse - Neue Produkte

Diese Aussage kann bei Betrachtung der Erfüllungsgrade der Teilnutzwerte auf detaillierterer Ebene nachvollzogen werden. In Abbildung 17 sind die Kriterien in bekannter Weise absteigend ihrer Gewichtung visualisiert.

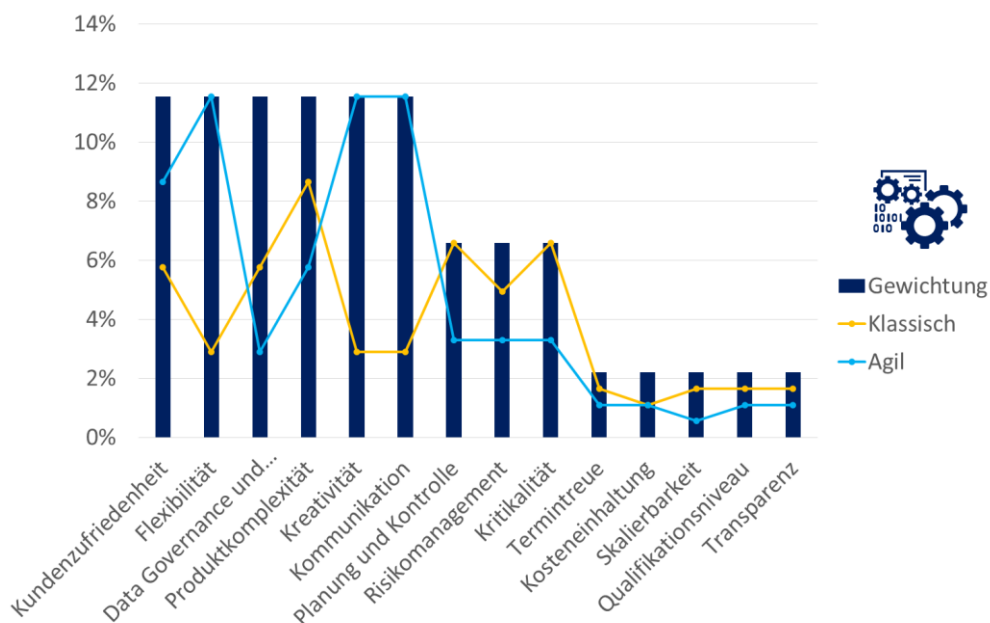


Abbildung 17: Erfüllungsgrad der Teilnutzwerte - Neue Produkte

Bei näherer Betrachtung wird deutlich, dass im vorliegenden Fall die agilen Vorgehensmodelle aufgrund von vier essentiellen Kriterien den klassischen Modellen vorzuziehen sind. Sie erfüllen die hier besonders bedeutsamen Kriterien Kundenzufriedenheit, Flexibilität, Kreativität und Kommunikation um ein deutliches Maß besser als klassische Vorgehensmodelle. Diese wiederum liegen trotz eines bedeutend schlechteren Gesamtnutzwerts bei den meisten anderen Kriterien vorn.

### 4.3 Handlungsempfehlungen

In diesem abschließenden Abschnitt des vierten Kapitels werden die im Rahmen von Nutzwertanalysen quantifizierten, relativen Vorteilhaftigkeiten von Vorgehensmodell-Typen in Bezug auf verschiedene Big-Data-Strategien argumentativ untermauert. Ziel ist es, für jede Strategie eine fundierte Handlungsempfehlung für die Einführung eines Big-Data-Management-Systems zu geben. Zur verbesserten Vergleichbarkeit sind die Ergebnisse sämtlicher Nutzwertanalysen unter Berücksichtigung von Gewichtungen und Teilnutzwerten in Tabelle 15 zusammenfassend dargestellt.

	#	Kriterium	Kosteneinsparung		Zeiteinsparung			Entscheidungsunterstützung			Neue Produkte			
			Gewichtung $G_k$	Teilnutzwert $n_{i,k}$		Gewichtung $G_k$	Teilnutzwert $n_{i,k}$		Gewichtung $G_k$	Teilnutzwert $n_{i,k}$		Gewichtung $G_k$	Teilnutzwert $n_{i,k}$	
				Klassisch	Agil		Klassisch	Agil		Klassisch	Agil		Klassisch	Agil
Ergebnis	1	Termintreue	11,5%	8,7%	5,8%	12,6%	9,5%	6,3%	5,5%	4,1%	2,7%	2,2%	1,6%	1,1%
	2	Kundenzufriedenheit	6,6%	3,3%	4,9%	6,6%	3,3%	4,9%	12,1%	6,0%	9,1%	11,5%	5,8%	8,7%
	3	Kosteneinhaltung	11,5%	5,8%	5,8%	6,6%	3,3%	3,3%	5,5%	2,7%	2,7%	2,2%	1,1%	1,1%
Management	4	Flexibilität	2,2%	0,5%	2,2%	1,1%	0,3%	1,1%	5,5%	1,4%	5,5%	11,5%	2,9%	11,5%
	5	Planung und Kontrolle	11,5%	11,5%	5,8%	12,6%	12,6%	6,3%	5,5%	5,5%	2,7%	6,6%	6,6%	3,3%
	6	Risikomanagement	6,6%	4,9%	3,3%	6,6%	4,9%	3,3%	12,1%	9,1%	6,0%	6,6%	4,9%	3,3%
	7	Data Governance und Datensicherheit	2,2%	1,1%	0,5%	1,1%	0,5%	0,3%	12,1%	6,0%	3,0%	11,5%	5,8%	2,9%
Hard Facts	8	Produktkomplexität	2,2%	1,6%	1,1%	12,6%	9,5%	6,3%	1,1%	0,8%	0,5%	11,5%	8,7%	5,8%
	9	Skalierbarkeit	11,5%	8,7%	2,9%	6,6%	4,9%	1,6%	5,5%	4,1%	1,4%	2,2%	1,6%	0,5%
	10	Qualifikationsniveau	10,4%	7,8%	5,2%	6,6%	4,9%	3,3%	1,1%	0,8%	0,5%	2,2%	1,6%	1,1%
	11	Kritikalität	12,1%	12,1%	6,0%	6,6%	6,6%	3,3%	5,5%	5,5%	2,7%	6,6%	6,6%	3,3%
	12	Transparenz	7,1%	5,4%	3,6%	12,6%	9,5%	6,3%	12,1%	9,1%	6,0%	2,2%	1,6%	1,1%
Soft Facts	13	Kreativität	2,2%	0,5%	2,2%	6,6%	1,6%	6,6%	12,1%	3,0%	12,1%	11,5%	2,9%	11,5%
	14	Kommunikation	2,2%	0,5%	2,2%	1,1%	0,3%	1,1%	4,4%	1,1%	4,4%	11,5%	2,9%	11,5%
Nutzwert $N_i$ in Bezug auf Idealprofil (=100%)				72,5%	51,5%		71,8%	54,1%		59,3%	59,6%		54,7%	66,8%

Tabelle 15: Zusammenfassung der Nutzwertanalysen

#### Kosteneinsparung durch Big Data

Das Ergebnis der Nutzwertanalyse fällt, wie im Kapitel 4.2.4 beschrieben, deutlich zugunsten klassischer Vorgehensmodelle aus. Weshalb derlei Modelle besser für die Einführung von Big-Data-Management-Systemen mit dem Ziel Kosten einzusparen geeignet sind, ist Gegenstand dieses Abschnitts. Die Kritikalität ist das für diese Strategie bedeutendste Kriterium. Dies ist dadurch begründet, dass nicht selten die gesamte Architektur bzw. hoch-sensible Teilbereiche zur Datenspeicherung auf Big-Data-Technologien umgestellt wird und

somit sämtliche Unternehmensdaten auf dem Spiel stehen. Hierzu muss das Vorgehensmodell in der Lage sein hoch sicherheitskritische Systeme zu entwickeln. Klassische Vorgehensmodelle sind durch ihre klare Struktur und den in der Methodik verankerten hohen Transparenzgrad sehr gut geeignet, um jene verantwortungsvolle Produkte zu entwickeln. In agilen Modellen hingegen ist durch die ausgeprägte Dynamik eine detaillierte Dokumentation und Transparenz in einem vergleichbaren Ausmaß nur schwerlich herzustellen.

Der starke Fokus auf die effiziente Zielerreichung mit Blick auf Zeit und Kosten sowie der unmittelbar damit verbundene hohe Anspruch an eine hinreichende Planbarkeit und Kontrollierbarkeit der Projektaktivitäten sprechen ebenso für die Verwendung klassischer Vorgehensmodelle. Mit eindeutig voneinander abgrenzbaren, sequentiell aufeinander folgenden Phasen und Arbeitspaketen erlauben sie durchgängige Nachvollziehbarkeit und Steuerbarkeit. Durch ihre iterative Herangehensweise können agile Vorgehensmodelle diese Eigenschaft nicht gleichermaßen zur Verfügung stellen. Weiterhin spricht gegen die agilen Modelle, dass sie nur bedingt auf umfangreichere Projektgrößen skaliert werden können. Kosteneinsparprojekte mit Big Data, die komplette Speicherarchitekturen ersetzen, sind häufig weit über der effizient agil realisierbaren Größenordnung. Weiterhin spielen Kriterien wie die Kreativität und offene Kommunikation im Projekt, die ein Kernelement agiler Methoden darstellen, eine untergeordnete Rolle. Der Grund dafür liegt in den bereits frühzeitig sehr klar und detailliert vorliegenden Produktspezifikationen, die sich im Projektverlauf sehr wahrscheinlich kaum ändern. Deshalb bedarf es nur wenig Kundenfeedback, was wiederum zu einem eingeschränkten Bedürfnis nach Kreativität führt.

Zusammenfassend ist das Ergebnis der Nutzwertanalyse bei der Kosteneinsparungsstrategie unter allen Strategien am deutlichsten ausgefallen. Die Differenz der Alternativennutzwerte beträgt 21 Prozentpunkte. Der Nutzwert der zu bevorzugenden Alternative erreicht auch im Vergleich mit den anderen Strategien mit 72,5 Prozent den höchsten Wert in Bezug auf das Idealprofil. Somit sind, sofern man das Ziel verfolgt mit Hilfe von Big-Data-Technologien Speicherkosten einzusparen, klassische Vorgehensmodelle den agilen eindeutig vorzuziehen. Begründet ist dies in erster Linie durch transparentere Strukturen und die bessere Planbarkeit der Aktivitäten, welche bei hochkritischen Systemen von großer Bedeutung sind.

### **Zeiteinsparung mit Hilfe von Big Data**

Möchte man auf Basis von Big-Data-Technologien Prozesse beschleunigen und somit Zeit einsparen, fällt das Ergebnis der Analyse ähnlich aus wie bei der Kosteneinsparung. Klassische Vorgehensmodelle scheinen besser geeignet zu sein als agile Modelle. Die Differenz der Nutzwerte der Alternativen ist zwar etwas geringer als bei der Strategie der Kosteneinsparung, gleichwohl erreichen klassische Vorgehensmodelle einen ähnlich hohen

Gesamtnutzwert. Big-Data-Projekte mit dem Ziel der Prozessbeschleunigung arbeiten in der Regel mit komplexer In-Memory-Technologie, die deutlich kürzere Zugriffszeiten ermöglicht. Die Harmonisierung bestehender Systeme mit Teilen der Technologien des Big-Data-Stapels aus Kapitel 2.3.2 stellt dabei eine bedeutende planerische Herausforderung dar. Die Transparenz sämtlicher Implementierungsschritte ist für eine spätere Nachvollziehbarkeit bei z.B. Wartungsarbeiten ebenso von hoher Relevanz. Da sich die durchführende Organisation damit beschäftigt Zeit einzusparen, liegt ein hoher Anspruch an die Termintreue von Projekten nahe.

All dies sind Kriterien, die klassische Vorgehensmodelle hinlänglich besser erfüllen können als agile Vorgehensmodelle. Insbesondere die Realisierung von hoher Produktkomplexität auf mehreren Ebenen kann durch vergleichbar leicht kontrollierbare und antizipierbare Abläufe im Projektgeschehen besser gewährleistet werden. Faktoren wie Flexibilität und Kommunikation spielen eine untergeordnete Rolle. Da in der Regel bekannte Prozesse der Fachabteilungen beschleunigt werden, sind gleichzeitig die Anforderungen an die Big-Data-Lösung frühzeitig bekannt. Der Kunde hat bereits in der Projektdefinitionsphase eine präzise Vorstellung vom zu entwickelnden Produkt. Dementsprechend ist mit vergleichbar wenigen Spezifikationsänderungen im Projektverlauf zu rechnen, was den Anspruch an die Flexibilität des Vorgehensmodells senkt. Wie bei der Strategie der Kosteneinsparung schon angeklungen ist, stellen klassische Vorgehensmodelle darüber hinaus durch ihre fundamentale Transparenz eine ausgesprochen gute Planbarkeit und Kontrollierbarkeit bereit. Hierdurch kann gegenüber agilen Modellen eine bessere Termintreue geboten werden.

In der Gesamtschau kann konstatiert werden, dass für die Einführung von Big-Data-Management-Systemen unter der Prämisse der Zeiteinsparung klassische Vorgehensmodelle bedeutend besser geeignet sind als agile Vorgehensmodelle. Die Begründung liegt im Kern analog zur Strategie der Kosteneinsparung in dem klar strukturierten und leicht nachvollziehbaren Aufbau der Modelle, die somit auch die planvolle und termingerechte Entwicklung sehr komplexer Systeme erlauben. Agile Vorgehensmodelle legen dem gegenüber andere Schwerpunkte, die nicht gleichermaßen gut für derartige Projekte geeignet sind.

### **Entscheidungsunterstützung und Prognosen durch Big Data**

Die Analyse agiler und klassischer Vorgehensmodelle zur Einführung von Big-Data-Technologien zur Entscheidungsunterstützung und Prognose führt zu einem besonderen Ergebnis. Beide Vorgehensmodell-Typen sind nahezu gleich gut geeignet, um in Projekten dieser Art eingesetzt zu werden. Die Erfüllungsgrade der Teilnutzwerte für die Kriterien haben jedoch unterschiedliche Schwerpunkte. Dementsprechend ist vor Projektbeginn umfassend abzuwägen, in welchen Kerngebieten das zu entwickelnde System seine Stärken

haben soll. Wie in Kapitel 4.2.4 bereits erwähnt wurde, sind beide Vorgehensmodelle mit Nutzwerten von jeweils ca. 56,5 Prozent im Verhältnis zum Idealprofil von 100 Prozent nur mäßig gut geeignet. Nichtsdestotrotz weisen beide Typen Stärken und Schwächen auf, die in Bezug auf die Strategie jeweils Vor- und Nachteile bringen.

Grundsätzlich geht es bei der Entscheidungsunterstützung mit Hilfe von Big Data darum, neue Datenquellen zu erschließen und die zumeist bereits vorhandene interne Datenanalyse in Form von Business Intelligence zu aktualisieren und anzureichern. Es sind demnach bekannte organisatorische Strukturen, was die moderate Bedeutung von Planung und Kontrolle erklärt. Vielmehr geht es darum bisher ungenutzte Daten verfügbar zu machen und im Rahmen von innovativen Analysemodellen zu verwerten. Hierzu bedarf es insbesondere an kreativem Freiraum, den vornehmlich agile Vorgehensmodelle zur Verfügung stellen. Da die Analysen, Entscheidungsvorlagen und Prognosen im Betrieb durch die Fachabteilungen genutzt und durchgeführt werden, ist von Projektbeginn an eine konsequente Einbindung des Kunden bzw. des Mitarbeiters der Fachabteilung in den Entwicklungsprozess erforderlich. Legt die individuelle Strategie ihre Schwerpunkte auf möglichst ausgefeilte und innovative Analysen durch Verknüpfung externer und interner Daten sowie auf hohe Entsprechung von Kundenerwartungen und geliefertem Produkt, so sind agile Vorgehensmodelle vorzuziehen.

Bei der Verwendung externer Daten undurchsichtiger Struktur kann es bei dieser Strategie schnell dazu kommen, dass Datenschutzrechte verletzt werden. Hierzu bedarf es im Projekt eines effektiven Risikomanagements sowie der fortlaufenden Weiterentwicklung der Data Governance. Jene Aspekte können in einem strukturierten Rahmen, wie es ein klassisches Vorgehensmodell bietet, besser verfolgt und erfüllt werden als in den vermeintlich chaotischen Strukturen agiler Modelle. Weiterhin ist die Transparenz des Entwicklungsprozesses bei dieser Strategierichtung von großer Bedeutung. Die Schritte zur Verknüpfung und Analyse müssen zur Argumentation eindeutig dokumentiert werden. Möchte die betroffene Organisation bei Verfolgung der vorliegenden Strategie ihre Schwerpunkte darauf legen, möglichst unstrukturierte, womöglich personenbezogene Daten zu analysieren und somit höhere Risiken eingehen, so ist ein klassisches Vorgehensmodell zur Einführung zu empfehlen.

Zusammenfassend kann demnach festgehalten werden, dass die Empfehlung für oder gegen einen Vorgehensmodell-Typ abhängig davon ist, welche Schwerpunkte die individuelle Strategie zur Entscheidungsunterstützung setzt. Steht die möglichst kreative Verknüpfung bisher ungenutzter und unbekannter Daten im Vordergrund, so sind agile Vorgehensmodelle zu empfehlen. Kommt es der Organisation primär darauf an sensible Daten in Entscheidungen mit einzubeziehen, so sind aus Risiko- und Data-Governance-Gründen klassische Vorgehensmodelle vorzuziehen.

### **Entwicklung neuer Produkte auf Basis von Big Data**

Da die Entwicklung neuer Produkte auf Basis von Big Data gegenüber den drei anderen Strategien eine strukturell etwas andere Stoßrichtung darstellt, müssen die Ergebnisse der Nutzwertanalyse in diesem Fall differenziert betrachtet werden. Die Einführung von Big-Data-Management-Systemen unter dieser Prämisse ähnelt vielmehr einem klassischen markt-orientierten Produktentwicklungsprozess, bei dem Kundenorientierung eine essentielle Rolle spielt. Das Ergebnis der Nutzwertanalyse in Kapitel 4.2.4 spricht die relative Vorteilhaftigkeit mit einem Gesamtnutzwert von 66,8 Prozent und einer Differenz von 12,1 Prozentpunkten eindeutig den agilen Vorgehensmodellen zu. Die Gründe dafür liegen primär in genannter Kundenorientierung und den daraus für ein Projekt resultierenden Faktoren.

Ein hohes Maß an Kundenzufriedenheit erfordert gleichzeitig auch immer ausgeprägte Flexibilität. Dies ist darauf zurückzuführen, dass Kunden zu Projektbeginn oft nur eine vage Vorstellung davon haben, wie die fertige Big-Data-Lösung später aussehen soll. Dadurch wiederum entstehen während des Projekts zahlreiche Änderungen der Produktspezifikationen, die teilweise sogar dazu führen, dass ganze Teilsysteme aufgrund mangelnder Kundenwunscherfüllung verworfen werden müssen. Die ausgeprägte Projektflexibilität und die uneingeschränkte Kundenorientierung sind Kernaspekte agiler Vorgehensmodelle, die im Agilen Manifest zum Ausdruck gebracht werden. Damit der Kunde in der Lage ist, ein Verständnis von den realisierbaren Möglichkeiten zu bekommen, muss ein kreativer Rahmen mit offener Kommunikationskultur vorliegen. Auch diese Aspekte schlagen sich in agilen Modellen deutlich ausgeprägter nieder als in klassischen.

Durch ihre starre Struktur und festgelegte Kommunikationsmodelle sind klassische Vorgehensmodelle in Bezug auf die für diese Strategie ausschlaggebenden Kriterien nur bedingt geeignet. Trotzdem können sie eine hohe Produktkomplexität, so wie sie in den meisten innovativen Big-Data-Produkten vorliegt, besser realisieren als agile Modelle. Darüber hinaus ist die kontinuierliche Weiterentwicklung der Data Governance in jenen Vorgehensmodellen besser gewährleistet als in agilen Modellen. Im Gesamtüberblick sind letztere jedoch eindeutig den klassischen Modellen vorzuziehen. Sie stellen ein motivierendes Umfeld, selbstbestimmtes Arbeiten sowie optimale Kundenorientierung sicher und liefern demnach einen geeigneten Projektrahmen für neue Produkte auf Basis von Big Data.



## 5 Beispielhafte Anwendung im Rahmen einer Fallstudie

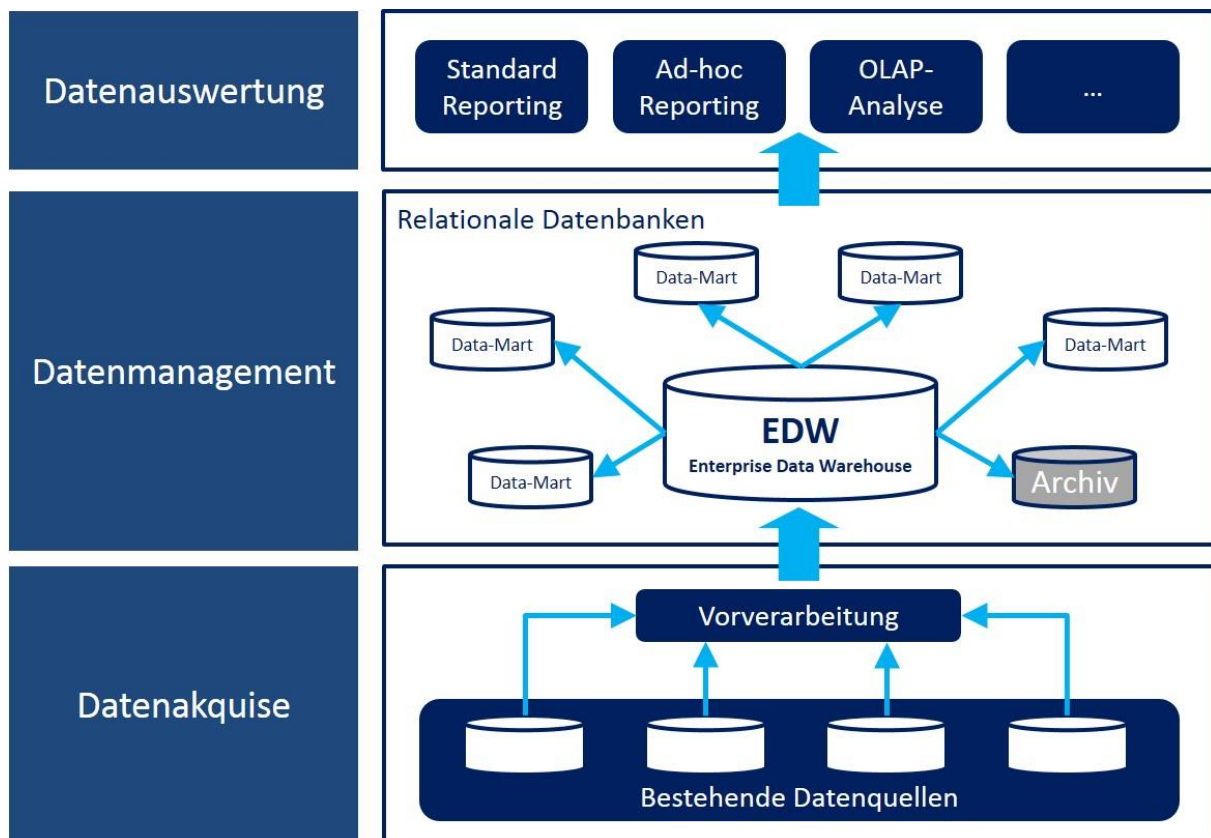
In diesem Abschnitt geht es darum, die Ergebnisse der in Kapitel 4 durchgeführten Nutzwertanalysen mit Hilfe einer beispielhaften Anwendung im Rahmen einer Fallstudie zu validieren. Da die Validierung aller vier untersuchten Strategien den Rahmen der vorliegenden Ausarbeitung sprengen würde, beschränkt sich dieser Abschnitt auf die Untersuchung der Strategie zur Kosteneinsparung durch Big-Data-Technologien. Illustriert wird diese Strategie anhand einer frei erfundenen, an die Erkenntnisse der Literaturrecherche angelehnten, Fallstudie aus dem Bankensektor. Zunächst werden dazu die Rahmenbedingungen und die Ausgangssituation skizziert. Im Anschluss werden die Projektziele und Produktspezifikationen erläutert, um darauf aufbauend illustrativ den realistischen Projektverlauf für klassische und agile Vorgehensmodelle nachzuzeichnen. Abschließend wird der individuelle Projekterfolg überprüft und die Aussagekraft der Nutzwertanalysen aus Kapitel 4 kritisch betrachtet.

### 5.1 Rahmenbedingungen und Ausgangssituation

Wie eingangs bereits angedeutet wurde, bewegt sich die vorliegende Fallstudie im Bankensektor. Im Fokus steht dabei eine Großbank, die innovativen Technologien zwar sehr offen gegenüber tritt, bei Betrachtung der Marktsituation und des Wettbewerbs jedoch vorsichtig agieren möchte. In der gesamten Branche herrscht eine angespannte Stimmung, da viele neue Wettbewerber mit Nischenprodukten auf den Markt drängen. Die Unternehmensleitung ruft aus diesem Grund eine Kosteneinsparungsstrategie aus. Die Bank agiert entsprechend konservativ und sucht funktionsübergreifend nach Potentialen für Kostensenkungen. Davon betroffen ist auch die IT-Abteilung, welche verantwortlich für die Administration, Instandhaltung und Beschaffung von IT-Systemen ist. Mehrere Mitarbeiter der IT-Abteilung haben Erfahrungen auf dem Gebiet von Big Data und können demnach als Datenwissenschaftler bezeichnet werden. Ungeachtet dessen arbeiten sie im Tagesgeschäft in ihrer Linienstruktur und sind nur selten mit Projektarbeit konfrontiert. Um ihren Beitrag zur Strategie des Unternehmens zu leisten, hat die Leitung der IT-Abteilung beschlossen mit Hilfe von Big-Data-Technologien Kosten einzusparen.

Das Datenaufkommen der Bank beschränkt sich zum großen Teil auf einheitlich strukturierte Transaktionsdaten. Die Big Data Dimensionen Variety und Velocity spielen demnach im vorliegenden Fallbeispiel eine untergeordnete Rolle. Nichtsdestotrotz wächst die Menge der Daten und somit die Bedeutung der Dimension Volume durch massiv steigende Transaktionszahlen rasant an. Dies hat zur Folge, dass auch die Speicherkapazitäten der Systeme und Archive der Bank mitwachsen müssen. Die Datenarchitektur der Bank weist den klassischen dreistufigen Aufbau aus operativen Systemen zur Datenakquise, einem Data-

Warehouse mit diversen Data-Marts zum Datenmanagement sowie einer Anwendungsebene zur Datenauswertung auf (Abb. 18).



**Abbildung 18: Datenarchitektur der Bank zu Projektbeginn**

Das Archiv ist im Ausgangszustand in der mittleren Ebene direkt an das Data-Warehouse gekoppelt. Mit Hilfe diverser Analysen hat die IT-Abteilung festgestellt, dass die Kosten für die pflichtgemäße Archivierung der Daten überproportional mit dem zu speichernden Volumen ansteigen. Zum Zeitpunkt der Analysen werden die Daten in klassischen Datenbanksystemen archiviert. Wie jedes andere Handelsunternehmen ist auch die Großbank von Gesetzes wegen verpflichtet Daten im Archiv aufzubewahren. Andernfalls drohen empfindliche Strafen bis hin zu Regressforderungen durch verschiedene Stakeholder. Von entsprechend hoher Bedeutung ist die Datensicherheit und Zuverlässigkeit des Archivs.

Anhand einfacher Kostenvergleiche wurde festgestellt, dass Big-Data-Technologien, wie z.B. ein Hadoop-Cluster, um ein Vielfaches kostengünstiger in Beschaffung und Betrieb sind. Dies ist unter anderem durch den Sachverhalt begründet, dass Hadoop eine Open-Source-Plattform ist und somit keine Lizenzgebühren anfallen (vgl. Kapitel 2.4.1). Darüber hinaus belegen Studien wie die von David Floyer, dass derlei Technologien sich deutlich schneller amortisieren und Cash-Flow generieren (vgl. Kapitel 2.5.3). Da das Know-how zur Implementierung innerhalb der IT-Abteilung liegt, hat die Bank beschlossen, das auf relationalen Datenbanken basierende, bestehende Datenarchiv auf ein Hadoop-Cluster umzustellen. Dadurch lassen sich den Kostenvergleichen zufolge die laufenden Kosten für die

Archivierung von einem Terabyte Daten um den Faktor 10 bis 20 reduzieren. Mit Blick auf genannte Einsparpotentiale hat das geplante Big-Data-Projekt die Zustimmung der Unternehmensleitung. Diese hat für die Umstellung ein klar begrenztes Budget und einen festen Projektendtermin festgesetzt und fordert ein kontinuierliches, transparentes Reporting.

Unabhängig vom Zweck der Aufbewahrung und Archivierung können mit Hilfe des aktiven Hadoop-Clusters weitere Mehrwerte für die Bank geschaffen werden. Beispielsweise ist es durch die aktive Archivierung möglich, Analysen über den gesamten Datenbestand seit Beginn der Datenaufzeichnung durchzuführen und die Rohdaten somit interaktiv zu durchsuchen. Diese Mehrwerte stehen jedoch nicht im Fokus der Tätigkeiten. Analog zur klar abgegrenzten Strategie der Kosteneinsparung durch Big Data, geht es im vorliegenden Fall primär darum direkt Kosten durch die Anschaffung und den Betrieb von Big-Data-Lösungen zu reduzieren (vgl. Kapitel 4.1). Dessen ungeachtet ist die Hadoop-Lösung, sofern sie zuverlässig und performant arbeitet, auch auf andere Teilbereiche der aktuellen Datenarchitektur der Bank erweiterbar. Das Projekt stellt unter dieser Prämisse gleichzeitig die initiale Anwendung von Big-Data-Technologien bei der Großbank dar.

## **5.2 Projektdefinition**

In diesem Abschnitt geht es darum die bereits umrissenen Rahmenbedingungen zu konkretisieren und zu verfeinern. Dies geschieht zum einen, indem der Umfang des Projekts dezidiert aufgeführt wird und ein Überblick über die notwendigen Tätigkeiten zur Implementierung gegeben wird. Weiterhin wird das Produkt zum besseren Verständnis in seiner Struktur näher spezifiziert.

### **5.2.1 Projektumfang und Vorgehensweise**

Ausgehend von der bereits skizzierten Kosteneinsparstrategie des gesamten Unternehmens hat jeder Funktionsbereich die Aufgabe, Einsparpotentiale zu identifizieren und nach Rücksprache mit der Unternehmensleitung umzusetzen. Der Funktionsbereich IT hat hierzu die Umstellung des auf relationalen Datenbanken basierenden Datenarchivs auf eine Hadoop-Umgebung als Projekt initiiert. Das Ziel ist es, die Speicherkapazität und Rechenleistung zu einem günstigen Preis auszudehnen. Es geht der Bank jedoch weniger um die Nutzung von großen Datenmengen, sondern vielmehr um die kostengünstige Nutzarmachung dieser vielen Daten in einem preiswerten aktiven Archiv. Der Umfang des Projekts beginnt mit der Machbarkeitsprüfung unter Zuhilfenahme von Testdaten und endet mit den finalen Systemtests sowie verschiedenen Performance-Benchmarkings.

Die Vorgehensweise zur Projektumsetzung orientiert sich an dem in Kapitel 2.3.2 von IBM und der Saïd Business School aufgestellten vierphasigen Prozess zur Implementierung des Big-Data-Stapels in Organisationen. Die initiale Informationsphase ist im vorliegenden Fallbeispiel schon lange überschritten, da die Führungskräfte hinsichtlich der Chancen durch Big

Data sensibilisiert sind und ein Projektbeschluss bereits vorliegt. Genauso hat die Bank die zweite Phase, in der individuelle Big-Data-Strategien für die jeweilige Organisation erarbeitet werden, bereits überschritten. Die darauf folgende dritte Phase der Prüfung stellt den Projektbeginn für das Big-Data-Projekt der Großbank dar. Wie in Abbildung 19 dargestellt, beginnt das Projekt mit der Prüfung der Machbarkeit mit Hilfe eines anonymisierten Datenauszugs. Gleichzeitig endet mit der Machbarkeitsprüfung die übergeordnete Phase Prüfen und ermöglicht damit den Auftakt der Umsetzungsphase.

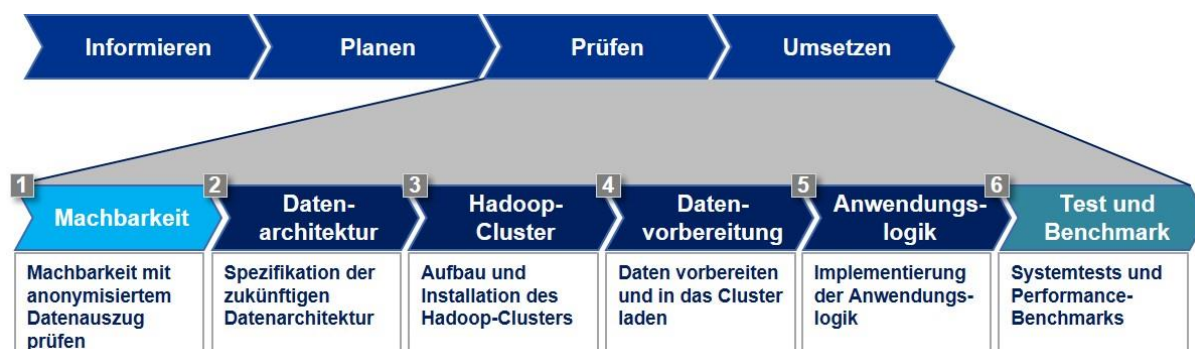


Abbildung 19: Vorgehensweise zur Projektumsetzung

Das Initial der Umsetzungsphase bildet die Ausarbeitung und Spezifikation der zukünftigen Datenarchitektur. Darin wird modelliert, inwieweit das Hadoop-Cluster mit den bestehenden Systemen interagiert. Darüber hinaus wird in dieser Unterphase die Datenmigration von der alten in die neue Struktur geplant und dokumentiert. Im Anschluss daran wird das Hadoop-Cluster vollumfänglich installiert und aufgebaut. Damit die Archivdaten in einem Hadoop-Cluster effizient gespeichert werden können, müssen sie mit Hilfe eines Transformationsprozesses vorbereitet werden. Erst danach können sie in das Cluster geladen werden. Die Implementierung der Logiken zur Zugriffssteuerung und Datenbearbeitung ist Bestandteil der darauf folgenden Unterphase. Eine besondere Herausforderung besteht hierbei darin, die für Hadoop spezifizierten Scriptsprachen Apache Pig und Hive anzuwenden (vgl. Kapitel 2.3.1). Diese Phase stellt in der Regel die größte Herausforderung für das Gesamtprojekt dar. Den Projektabschluss bilden die Tests der entwickelten Anwendungslogiken im Hadoop-Cluster sowie ein Benchmarking der Performance mit vergleichbaren Systemen. Zudem wird die Kosteneffizienz in Relation zur alten Lösung validiert und in die Zukunft interpoliert, um zu prüfen, ob die Einspareffekte in erwartetem Maße eintreten werden. Der Betrieb der Big-Data-Lösung und deren Instandhaltung sind demnach nicht mehr Projektbestandteil. Die Dauer des Projekts ist aufgrund des Aufwands zur Datenknotenverknüpfung proportional zur Größe des implementierten Hadoop-Clusters.

### 5.2.2 Produktspezifikationen

Große Teile der Produktspezifikation sind abhängig von den vorherrschenden Rahmenbedingungen. Wie bereits erwähnt, verfügt die Großbank über ein Data-Warehouse und

mehrere daran angeschlossene Data-Marts (Abb. 18). Diese relational arbeitenden Systeme werden zentral von der IT-Abteilung administriert und instandgehalten. Das Hadoop-Cluster soll den Platz des Archivs auf der Datenmanagement-Ebene einnehmen und bei Bedarf für Analysen bereitstehen (Abb. 20).

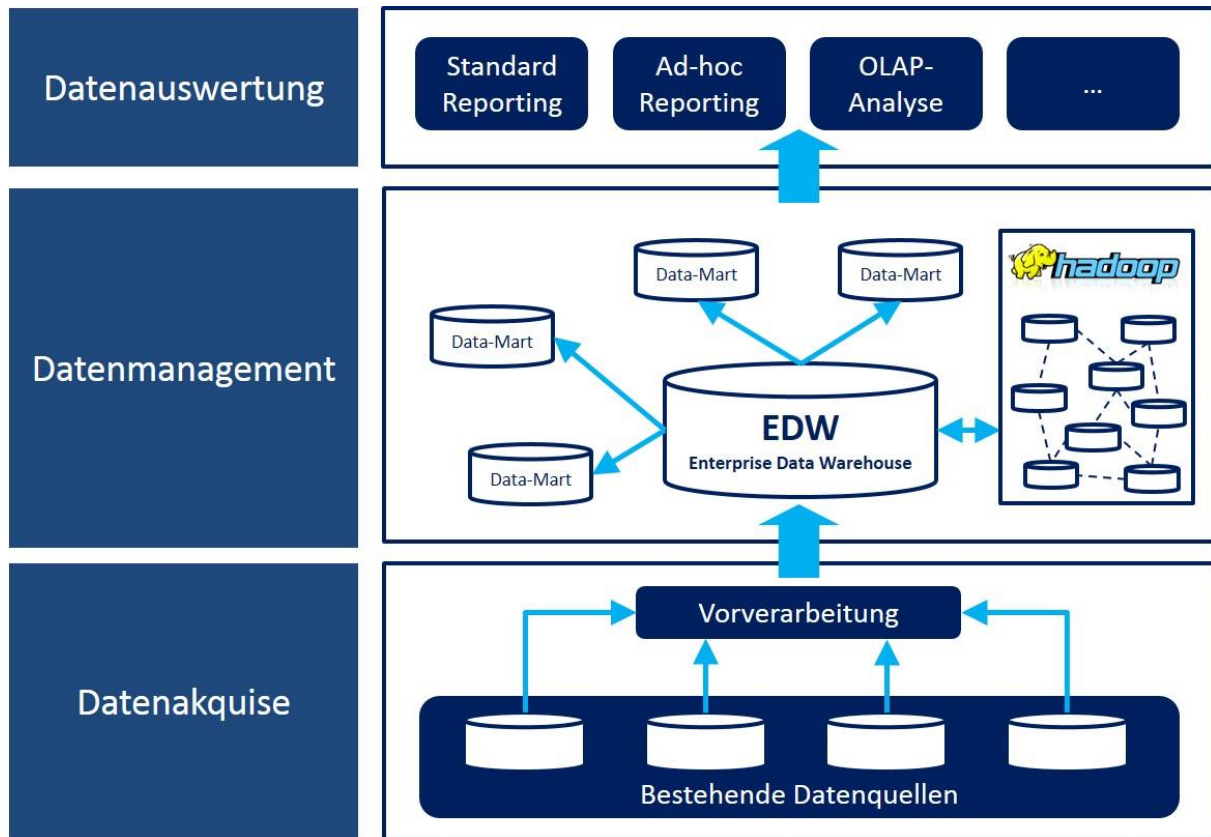


Abbildung 20: Ziel-Datenarchitektur nach Projektabschluss

Im Detail hat man beschlossen dazu ein Hadoop-Cluster mit 50 Knoten und 800 Kernen zur langfristigen Verfügbarmachung eines aktiven Archivs für Rohdaten zu beschaffen. Diese Ausmaße sind vor allem auf die Menge der zu archivierenden Daten zurückzuführen. Das Cluster soll langfristig Daten im Petabyte-Bereich verarbeiten und archivieren können sowie reibungslos an die bestehende Datenarchitektur des Data-Warehouses gekoppelt werden können. Dabei werden vornehmlich Transaktionsdaten der operativen Systeme akquiriert, anschließend vorverarbeitet und letztlich durch das Data-Warehouse in das Hadoop-Cluster geladen. Der Aufbau des Hadoop-Clusters soll so gestaltet werden, dass eine spätere Erweiterbarkeit auf andere IT-Funktionen, wie beispielsweise den Ersatz diverser Data-Marts, möglich ist. Dies ist abhängig davon, wie zuverlässig das System arbeitet. In der ersten Nutzungsphase übernimmt das Hadoop-Cluster vollständig die Aufgaben des bestehenden relational arbeitenden Archivs. Der Fokus liegt demnach ausschließlich auf den für die Archivierung aufzuwendenden Kosten. In späteren Nutzungsphasen sollen die aktiv vorgehaltenen Rohdaten zur langfristigen Datenanalyse genutzt sowie zur Datenexploration

freigegeben werden. Eine Kommunikationsschnittstelle mit der Anwendungsebene muss demnach ebenso implementiert werden.

Insgesamt ist die Produktkomplexität im Kontext der Möglichkeiten von Big-Data-Technologien als moderat anzusehen. Aus diesem Grund und angesichts der Tatsache, dass die Großbank eine sehr genaue Vorstellung vom Einsatzzweck des Zielsystems hat, liegt der Anforderungskatalog an das zu entwickelnde System bereits zu Projektbeginn sehr granular spezifiziert vor.

### **5.3 Projektdurchführung**

In diesem Abschnitt wird jeweils ein möglicher Projektverlauf für klassische und agile Vorgehensmodelle auf Basis der genannten Rahmenbedingungen für das Projektmanagement und das zu entwickelnde Produkt dargestellt. Stellvertretend für die klassischen Vorgehensmodelle wird das in Kapitel 3.3.2 erläuterte Wasserfallmodell von Boehm angewandt. Dem gegenüber wird repräsentativ für die agilen Vorgehensmodelle ein möglicher Projektverlauf unter Verwendung der in Abschnitt 3.4.2 geschilderten Scrum-Methodologie vorgestellt.

#### **5.3.1 Projektverlauf mit klassischen Vorgehensmodellen**

Verwendet man zur Durchführung des vorliegenden Projekts das Wasserfallmodell nach Boehm, so kann man die Phasen der in Abschnitt 5.2.1 aufgestellten Vorgehensweise direkt in das Modell einfließen lassen (Abb. 21). Durch die klare Abtrennung der einzelnen Phasen, was charakteristisch für klassische Vorgehensmodelle ist, können aufeinander aufbauende Tätigkeiten gut strukturiert und nachverfolgt werden.

Nachdem ein detaillierter Projektplan entwickelt wurde, beginnt das Wasserfallmodell mit der Phase der Machbarkeitsprüfung. Erst bei erfolgreicher Validierung der Durchführbarkeit einer Hadoop-Cluster-Implementierung als aktives Datenarchiv kann die darauf folgende Phase der Anforderungsanalyse und Produktspezifikation beginnen. Diese Phase umfasst die Sammlung sämtlicher bekannter Produkthanforderungen und hat als Ergebnis eine detaillierte Spezifikation des zu entwickelnden Produkts unter Berücksichtigung der bestehenden Datenarchitektur. Zur Validierung werden zahlreiche Modelle erstellt und Simulationen durchgeführt. Führen diese Simulationen dazu, dass die Machbarkeit erneut in Frage gestellt wird, so ist ein Rückschritt in die erste Phase möglich. In diesem Fall würde dort eine erneute Machbarkeitsanalyse unter geänderten Rahmenbedingungen erfolgen.

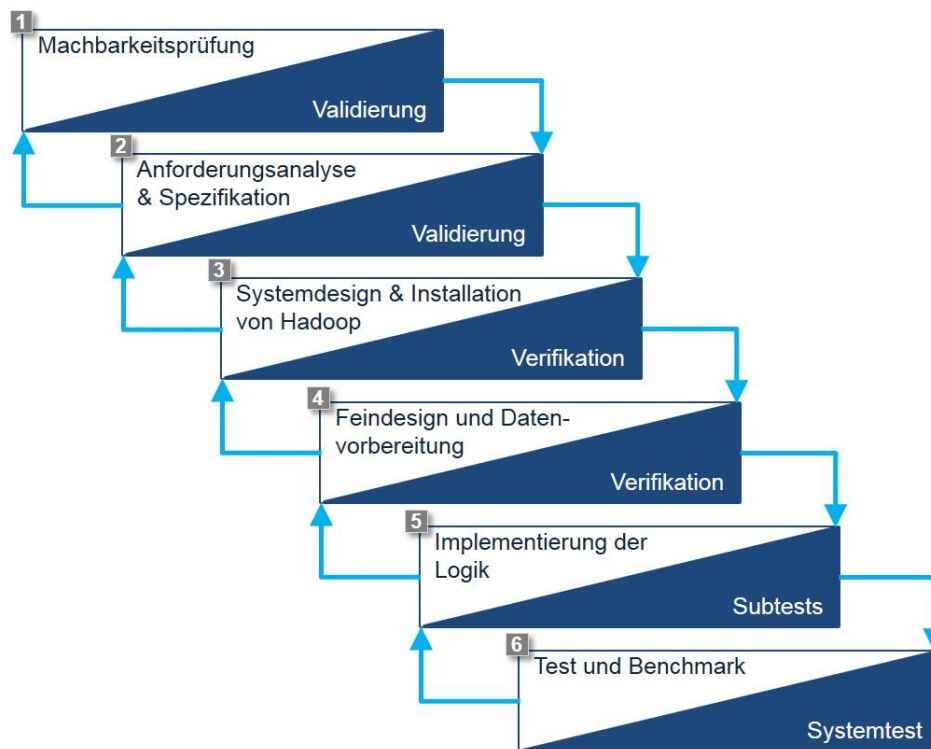


Abbildung 21: Projektdurchführung mit dem Wasserfallmodell

In der dritten Phase beginnt das operative Systemdesign zur Erfüllung des zuvor definierten Anforderungskatalogs. In einem ersten Schritt wird dazu das Hadoop-Cluster mit 50 Knoten aufgespannt und installiert. Darüber hinaus werden erste Tests zur Schnittstellenintegration durchgeführt. Diese sind notwendig, um die bisherigen Aktivitäten zu verifizieren. Auch die darauf folgende Phase beschäftigt sich mit dem Systemdesign. Dort wird auf einer feingranularen Ebene geprüft, ob die Archivfunktion vollends erfüllt werden kann. Dazu werden Testdaten in das Archiv geladen und die Gesamtkonzeption somit erneut verifiziert. Darüber hinaus ist es Bestandteil dieser Phase, die bestehenden Archivdaten dahingehend vorzubereiten, dass sie reibungslos in ein Hadoop-Cluster geladen werden können.

Die Implementierung der Logik in der fünften Phase ist für die Termintreue und Kosteneinhaltung des Projekts ausschlaggebend, da sie am meisten Aufwand verursacht. Hierzu wird im Rahmen des Wasserfallmodells bereits im Vorfeld viel planerischer Aufwand betrieben, um derartig kritische Phasen kontrolliert zu durchlaufen. Durch Subtests einzelner Funktionen werden die Phasenergebnisse gesichert. Die das Projekt abschließende sechste Phase umfasst den Systemtest und einen Benchmark mit vergleichbaren Lösungen. Entsprechen die Leistungen des entwickelten Produkts den Produkthanforderungen, kann das Projekt erfolgreich beendet werden.

Während des gesamten Projektverlaufs werden die Aktivitäten von der Projektleitung gesteuert. Kommt es zu Bearbeitungsabweichungen der bei der initialen Projektplanung definierten Arbeitspakete, so kann sie direkt reagieren und Planänderungen durchführen. Die

Planänderungen sind dabei jedoch selten von inhaltlicher Natur. Vielmehr geht es dabei darum, die Aktivitäten zeitlich neu zu verteilen, um den geplanten Endtermin einzuhalten und die budgetierten Kosten nicht zu überschreiten.

### 5.3.2 Projektverlauf mit agilen Vorgehensmodellen

Ein möglicher Projektverlauf mit agilen Vorgehensmodellen stellt für das vorliegende Projekt eine Herausforderung dar. Zur Illustration wird im Folgenden davon ausgegangen, dass das Projekt anhand der Scrum-Methodologie durchgeführt und strukturiert wird. Anders als beim Wasserfallmodell lassen sich die in Kapitel 5.2.1 definierten Phasen zur Projektumsetzung nicht eins zu eins in das Vorgehensmodell einflechten. Die inkrementelle Vorgehensweise, die Scrum zugrunde liegt, zielt darauf ab, möglichst frühzeitig ein funktionierendes, testfähiges System bereitzustellen. Dieses System wird daraufhin iterativ von Sprint zu Sprint entsprechend der in Kapitel 3.4.2 dargestellten Verfahrensweise optimiert und anhand der Kundenanforderungen weiterentwickelt. Trotzdem bauen die vorgestellten Phasen inhaltlich aufeinander auf, was dazu führt, dass die Aktivitäten auch im Rahmen von Scrum in ähnlicher Abfolge durchgeführt werden.

Zu Beginn des Projekts steht nach wie vor die Machbarkeitsanalyse. Sie überprüft vorgelagert die Durchführbarkeit des Projekts und legt den Grundstein für die weiteren Aktivitäten. Scrum hat in diesem Fall keinen Einfluss auf deren Ausführung. Im Anschluss beginnt die IT-Abteilung, welche den Product Owner darstellt, damit, die Produktanforderungen im Product Backlog zu sammeln und zu priorisieren. In die Erstellung des Product Backlogs sind alle relevanten Stakeholder eingebunden. Nachdem das Product Backlog mit einer klaren Vision vom fertigen Produkt definiert ist, beginnt die operative Sprintplanung. Darin vereinbaren der Product Owner und das Entwicklungsteam, welche Produktanforderungen im nächsten Sprint erarbeitet werden.

Die in den Phasen der Vorgehensweise aus Kapitel 5.2.1 aufgeführten Tätigkeiten werden mit Blick auf die verfügbaren Kapazitäten von Sprint zu Sprint individuell geplant und abgearbeitet. Dabei wird mit erhöhter Priorität dahin entwickelt, möglichst frühzeitig eine laufende Software zur Verfügung zu stellen. Da im vorliegenden Projekt aufgrund der hochkritischen Produktverwendung zunächst eine detaillierte Modellierung der Datenarchitektur im Rahmen des Feindesigns notwendig ist, gestaltet sich dieser Ansatz als schwierig. Demzufolge erfolgt die Abarbeitung der Arbeitspakete in identischer Reihenfolge wie beim Wasserfallmodell im vorangegangenen Abschnitt.

Um einen Überblick über die Projektaktivitäten zu bekommen und somit die Projektaktivitäten zu kontrollieren, wird typischerweise das Scrum Burndown Chart eingesetzt. Es visualisiert die Differenz zwischen optimalem und tatsächlichem Kapazitäteneinsatz im Projekt. Auf der Abzisse sind dabei die Sprints abgetragen, auf der Ordinate die Bearbeitungseinheiten.



Illustrativ ist in Abbildung 22 ein solches Burndown Chart mit fiktiven Größen und Skalierungen für das vorliegende Projekt skizziert. Der optimale Verlauf des noch zu erledigenden Arbeitsaufwands erstreckt sich linear über sechs Sprints. Der tatsächliche Verlauf hingegen bewegt sich durchgängig darüber, was bedeutet, dass das Projekt nicht zum geplanten Endtermin abgeschlossen werden kann. Die Gründe dafür liegen primär darin, dass die zu bearbeitenden Aufgaben des Product Backlogs sprintweise geplant und ausgeführt werden. Dadurch können komplexe Produktbestandteile, wie beispielsweise die Programmierung und Implementierung der Anwendungslogik für das Hadoop-Archiv nicht hinreichend antizipiert werden. Dies wiederum macht zusätzliche Sprints notwendig, um den vollen Umfang des Product Backlogs zu entwickeln und das Projekt schließlich zum Ende zu führen.

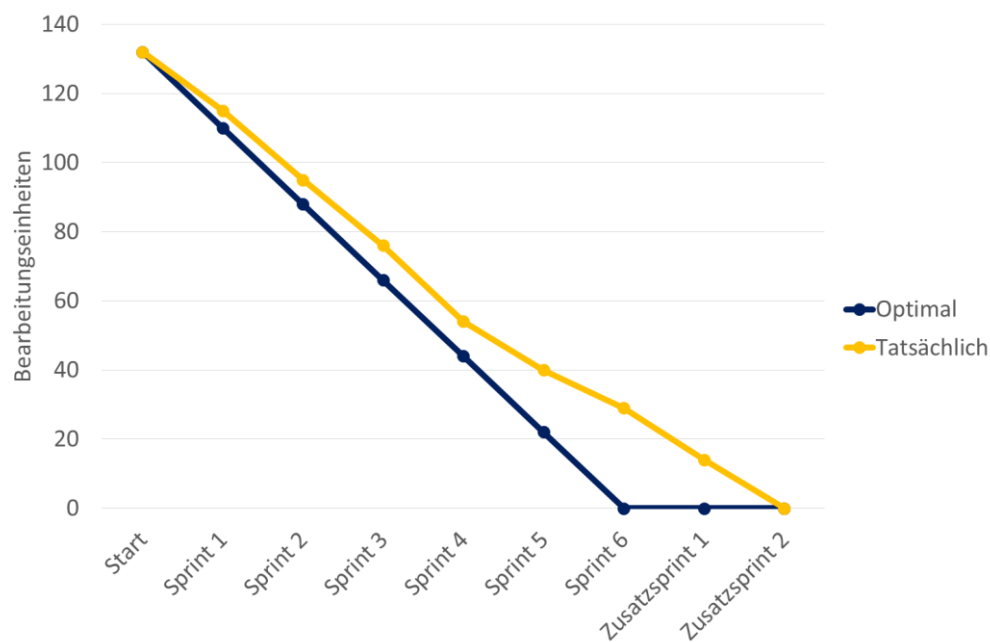


Abbildung 22: Scrum Burndown Chart für Kosteneinsparungsprojekt

## 5.4 Projekterfolg und Ergebnisbewertung

Die in Kapitel 5.3.1 und 5.3.2 vorgestellten möglichen Projektverläufe für klassische und agile Vorgehensmodelle werden in diesem Abschnitt auf ihr Erfolgspotential hin untersucht. Die Erfolgsbewertung stellt dabei die Validierung der Ergebnisse der Nutzwertanalyse aus Kapitel 4.2.4 sowie der darauf basierenden Handlungsempfehlungen aus Abschnitt 4.3 dar. Im Detail werden die möglichen Projektverläufe in Relation zu den typischen Charakteristika des entsprechenden Vorgehensmodell-Typs gesetzt und deren Wirkungsweise untersucht. Vergleicht man die vorgestellten möglichen Projektverläufe miteinander, so wird deutlich, dass die Zielrichtung des Projekts, Kosten mit Hilfe eines Hadoop-Clusters zu sparen, dafür spricht klassische Vorgehensmodelle anzuwenden. Welche Eigenschaften klassischer Vorgehensmodelle zu diesem Ergebnis führen, wird im Folgenden dargelegt.

Zunächst ist es möglich die notwendigen Tätigkeiten zur Einführung eines solchen Systems präzise und chronologisch voneinander in einzelne Phasen abzugrenzen. Dies erlaubt eine vergleichbar simple Einflechtung in das Wasserfallmodell nach Boehm. Charakteristisch für den gesamten, sequentiell strukturierten Projektverlauf sind klar definierte Termine und Meilensteine, die regelmäßig auf ihre Validität überprüft werden. Können Arbeitspakete nicht zeitgerecht fertiggestellt werden, so erlaubt das Vorgehensmodell frühzeitiges planerisches Gegensteuern. Darüber hinaus werden sämtliche Projektaktivitäten detailliert protokolliert und müssen kontinuierlich an die Projektleitung berichtet werden. Das führt zu einer sehr strukturierten und zielorientierten Arbeitsatmosphäre. Diese trägt insbesondere in kritischen Phasen, wie bei der Implementierung der Anwendungslogik, dazu bei, dass stets eine hohe Transparenz über die Aktivitäten vorliegt und frühzeitig auf Abweichungen reagiert werden kann. Gleichzeitig führt es dazu, dass das Endprodukt durch präventive Risiko-berücksichtigung mit Fokus auf hoher Zuverlässigkeit und Sicherheit entwickelt werden kann. Da das durchführende Projektteam nur vereinzelt mit der Projektkoordination konfrontiert ist, können sie demnach tendenziell produktiver arbeiten, wenn sie genaue Vorgaben über die Arbeitsinhalte erhalten.

Auf der anderen Seite zeigt sich, dass agile Methoden, wie im vorliegenden Fall Scrum, mit hoher Wahrscheinlichkeit nicht die Ergebnisqualität eines klassischen Vorgehensmodells erreicht. In erster Linie äußert sich dies im inkrementell iterativen Aufbau des Modells. Da sich die Entwicklungsarbeit übergeordnet zwar an der Grobstruktur des Product Backlogs orientiert, im Detail jedoch nur von Sprint zu Sprint geplant wird, kann es trotz umfassender Kenntnisse über die Produkthanforderungen leicht zu Verzögerungen kommen. Dementsprechend kann der durch die Unternehmensleitung strikt vorgegebene Projektendtermin unter der Prämisse einer möglichst effizienten und kostenschonenden Systemeinführung nur unwahrscheinlich eingehalten werden. Des Weiteren ist die frühzeitige Bereitstellung lauffähiger Systeme ein Kernelement agiler Modelle. Da das zu entwickelnde Produkt diese Anforderung nur begrenzt erfüllt, sind den agilen Vorgehensmodellen auch hier Grenzen gesetzt. Eine weitere Problemstellung ist die Unerfahrenheit des Entwicklungsteams in Bezug auf Projektarbeit. Bei Scrum besteht ein überproportional hoher Anspruch an die Eigeninitiative und das Verantwortungsbewusstsein der Projektmitglieder. Dies kann ebenso zu schlechteren Projektergebnissen führen.

Setzt man die vorangegangene Erfolgsbetrachtung in Bezug zu den Ergebnissen der Nutzwertanalyse aus Kapitel 4, so bestätigen sich die Resultate. Mit einem Nutzwert von 72,5 Prozent liegen klassische Vorgehensmodelle bei der Einführung von Big-Data-Management-Systemen mit der Prämisse Kosten einzusparen deutlich vor agilen Modellen mit 51,5 Prozent. Beide Werte müssen in Relation zum Idealprofil von 100 Prozent gesehen werden, was bedeutet, dass auch klassische Vorgehensmodelle keine Ideallösung darstellen. Es wird

jedoch die Handlungsempfehlung validiert, dass bei relativer Betrachtung klassische Modelle, trotz hinlänglich bekannter Schwächen in Bezug auf Flexibilität, vorzuziehen sind. Zusammenfassend attestiert das vorliegende Fallbeispiel einer Großbank demnach die Sinnhaftigkeit der durchgeführten Nutzwertanalysen. Dies wiederum bestätigt, dass der aufgestellte Kriterienkatalog zur Beurteilung der Vorgehensmodelle es erlaubt eine adäquate Aussage darüber zu treffen, mit welchem Modell Big-Data-Management-Systeme für die jeweilige Strategie effizienter eingeführt werden können. Dabei greift er hinreichend die typischen Herausforderungen und Erfolgsfaktoren im Projektmanagement von Softwareprojekten auf.

## 6 Schlussbetrachtung

Der abschließende Abschnitt der vorliegenden Ausarbeitung besteht aus drei Teilen. Zunächst werden die Erkenntnisse der Literaturrecherche zum Status Quo der Themenfelder Big Data (Kapitel 2) und IT-Vorgehensmodelle (Kapitel 3) zusammengetragen. Im Anschluss werden die Ergebnisse der Nutzwertanalysen (Kapitel 4) sowie deren Validierung im Rahmen einer Fallstudie (Kapitel 5) zusammengefasst dargestellt. Den zweiten Teil bildet ein persönliches Fazit, welches die Resultate der Arbeit kritisch betrachtet. Im dritten und letzten Teil wird ein Ausblick über die potentielle Entwicklung der Thematik sowie mögliche Ansatzpunkte zur wissenschaftlichen Weiterbearbeitung gegeben.

### 6.1 Zusammenfassung

Die Ausarbeitung lässt sich, wie in Abbildung 2 in der Einleitung visualisiert, in zwei Teile gliedern. Im ersten Teil werden Grundlagen zum Stand der Technik erarbeitet, welche die Basis für die im zweiten Teil durchgeführten Analysen und Bewertungen darstellen. Den Startpunkt bildet dabei die Aufarbeitung der Literatur zum Themenkomplex Big Data in Kapitel 2.

Zunächst hat sich in Abschnitt 2.1 herausgestellt, dass die Herkunft von Big Data in erster Linie in dem kontinuierlich wachsenden weltweiten Datenaufkommen zu suchen ist. Nichtsdestotrotz spielt nicht nur die Menge der Daten eine Rolle, sondern vordergründig auch deren Auswertbarkeit im Rahmen von Analysen. Um sich dem Begriff Big Data weiter zu nähern, beschäftigt sich Kapitel 2.2 mit der Definition und den Eigenschaften von Big Data. Bei Betrachtung unterschiedlicher Deutungsversuche der Literatur wird deutlich, dass es keine einheitliche Definition gibt. Vielmehr charakterisiert sich Big Data über die Eigenschaften der Daten. Die Dimensionen Volume, Velocity und Variety spielen bei der Beschreibung eine übergeordnete Rolle. Insbesondere letztere Dimension erlaubt eine Aussage über die Struktur der Daten, was einer besonderen Herausforderung in diesem Kontext gleichkommt. Weiterhin tauchen in der Literatur die Dimensionen Value und Veracity auf, um zusätzliche Merkmale beschreiben zu können.

Eine Aussage darüber, welche Technologien zum Handling von Big Data bestehen, wird in Kapitel 2.3 geliefert. Es beschäftigt sich mit dem Aufbau von Big-Data-Management-Systemen, welche letztlich Betrachtungsgegenstand der Ausarbeitung sind. Zunächst werden Technologien für Big Data vorgestellt, die den Umgang mit großen Mengen verschieden strukturierter Daten ermöglichen. Die Darstellung von Technologien, wie z.B. Hadoop, MapReduce oder NoSQL ermöglicht ein besseres Verständnis des Aufbaus von Big-Data-Management-Systemen. Strukturell kann ein solches System adäquat mit dem sogenannten Big-Data-Stapel beschrieben werden. Er kategorisiert die notwendigen Technologien in verschiedene Ebenen, beginnend mit dem Speicher und endend mit der tat-

sächlichen Anwendung. In Kapitel 2.3.2 wird dazu hervorgehoben, dass nicht zwangsweise alle Schichten des Stapels implementiert werden müssen. Zudem wird basierend auf einer Studie von IBM ein mögliches Phasenmodell zur Einführung von Elementen des Big-Data-Stapels geliefert. Um einen ganzheitlichen Blick auf die Datenarchitektur bei der Verwendung von Big-Data-Management-Systemen zu werfen, sind in Abschnitt 2.3.3 mögliche Anbindungen zu bestehenden Systemen illustriert. Es stellt sich dabei heraus, dass traditionelle Systeme durch Big Data nicht verdrängt, sondern vielmehr ergänzt und erweitert werden.

Maßgeblich für die Struktur der Nutzwertanalysen sind die in Abschnitt 2.4 dargestellten Ziele und Strategien von Big Data. Basierend auf den Erkenntnissen von Davenport kristallisieren sich für Unternehmen vier verschiedene Ansätze im Umgang mit Big Data heraus: die Kosteneinsparstrategie, die Zeiteinsparstrategie, die Strategie mit Big Data die Entscheidungsfindung zu optimieren sowie die Strategie zur Entwicklung innovativer Produkte auf Basis von Big Data. Für den Untersuchungszweck im Rahmen der Nutzwertanalyse werden alle Stoßrichtungen isoliert betrachtet. In der Realität ist es jedoch nicht unwahrscheinlich, dass ein Unternehmen sich mit Mischformen oder mehreren Strategien gleichzeitig beschäftigt. Um diese Ansätze zu veranschaulichen, sind jeweils diverse Anwendungsfälle herausgestellt. Den Abschluss der Aufbereitung des Big-Data-Themenkomplexes bildet die Darstellung verschiedener Herausforderungen im Umgang mit Big Data. Ein erstes Problemfeld bildet dabei die Data Governance und Datensicherheit. Es zeigt sich, dass die bestehenden Modelle kontinuierlich weiterentwickelt werden müssen, sodass die Datenverwendung auch im Big-Data-Zeitalter kontrolliert und gesichert werden kann. Des Weiteren kristallisiert sich mit dem Qualifikationsprofil des Datenwissenschaftlers eine neue Schnittstellenfunktion im Unternehmen heraus. Er verbindet typische Eigenschaften eines Datenanalysten mit denen eines Managers. Eine kritische Menge dieser Mitarbeiter vorzuhalten ist, genau wie die organisatorische Einbindung in die Aufbauorganisation, eine weitere Herausforderung. Schließlich stellt die Erfolgsmessung von Big-Data-Projekten Unternehmen vor eine schwierige Aufgabe, da die Wirkung von Big Data, abhängig von der verfolgten Strategie, unter Umständen nicht hinreichend monetär belegt werden kann. Gerade die Berechnung des Return on Investment gestaltet sich häufig als mühevoll. Aus diesem Grund hat die BITKOM eine ähnliche Kennzahl, den Return on Information, entwickelt, um derartige Bewertungen zu vereinfachen.

Die Aufarbeitung des Themengebiets Vorgehensmodelle des IT-Projektmanagements findet in Kapitel 3 statt. Auch hier wird dazu zunächst ein übergreifender Abriss über die Herkunft und Definition von Vorgehensmodellen im Allgemeinen gegeben. Dabei zeigt sich, dass Vorgehensmodelle vor allem daher rühren, dass IT-Projekte sich häufig in ähnliche, immer wieder auftretende Aufgabenstellungen gliedern lassen. Diese Eigenschaft nutzen verschie-

dene Vorgehensmodelle-Typen aus, um Projekte zu strukturieren und zu systematisieren. Es gibt klassische und agile Modelle, die Umfragen zufolge in der Praxis verschieden populär sind. Es zeigt sich, dass nur etwa ein Viertel der befragten Unternehmen IT-Systeme mit Hilfe agiler Vorgehensmodelle entwickelt. Bevor beide Perspektiven anhand repräsentativer Modelle charakterisiert werden, geht aus Abschnitt 3.2 hervor, dass es im IT-Projektmanagement diverse Herausforderungen und Erfolgsfaktoren zu berücksichtigen gilt. Prinzipiell stellt sich heraus, dass IT-Projekte nur verhältnismäßig selten im vollen Umfang erfolgreich abgeschlossen werden. Dies ist zurückzuführen auf eine große Anzahl an Risiken, wie z.B. die ungeplante Überschreitung der Projektlaufzeit oder die unerwartet hohe Produktkomplexität. Ein fundiertes Risikomanagement hilft demnach, die Erfolgswahrscheinlichkeit zu erhöhen.

In Kapitel 3.3 werden klassische Vorgehensmodelle im Detail erläutert und ihre Vor- und Nachteile anhand des Wasserfallmodells nach Boehm veranschaulicht. Klassische Vorgehensmodelle greifen auf die Tatsache zurück, dass bei Neu- oder Weiterentwicklungen von IT-Systemen häufig wiederkehrende Phasen anzutreffen sind. Es zeigt sich, dass diese Phasen klar voneinander abgrenzbar und somit verhältnismäßig leicht strukturierbar sind. Das Wasserfallmodell repräsentiert die einfachste Form von klassischen Vorgehensmodellen und ist in der Praxis noch immer häufig in Verwendung. Es wird durch den sequentiellen Ablauf der verschiedenen Entwicklungsphasen charakterisiert. Dabei darf nur in die nächste Phase vorgerückt werden, wenn tatsächlich alle Arbeitsinhalte der bisherigen Phase vollendet und validiert bzw. verifiziert sind. Problematisch wird das Wasserfallmodell, wenn unerwartete Anforderungsänderungen auftreten. Folglich sind klassische Vorgehensmodelle hinderlich, wenn das IT-Projekt einen hohen Anspruch an die Flexibilität hat.

Kapitel 3.4 beschäftigt sich mit den Eigenschaften agiler Vorgehensmodelle. Basierend auf dem Agilen Manifest versuchen sie die Schwächen klassischer Vorgehensmodelle zu beheben. Demnach stellen sie eine Reaktion auf immer wiederkehrende Probleme bei IT-Projekten dar. Sie stellen das Individuum, die Kundenorientierung und die möglichst frühzeitige Erstellung funktionierender Software in den Vordergrund. Darüber hinaus sind sie weitaus flexibler in Bezug auf Planänderungen und neue, bisher unbekannte Produktanforderungen. Beispielhaft wird das in der Praxis verbreitetste agile Modell Scrum vorgestellt. Durch die iterativen Entwicklungszyklen wird bei Scrum das sogenannte Product Backlog inkrementell entwickelt. Diese Vorgehensweise erfordert hohe Eigenverantwortung des Projektteams und die Fähigkeit sich häufig ändernden Anforderungen anzupassen. Folglich ist die Planbarkeit und Prognostizierbarkeit der Projektaktivitäten bei agilen Modellen übergreifend schwächer ausgeprägt als bei klassischen Modellen.

Um beide Vorgehensmodell-Typen angemessen miteinander vergleichen zu können, werden in Abschnitt 3.5 verschiedene Ansätze aus der Literatur sowie Kriterien aus der Praxis zum

Vergleich jener Modelle eingeführt. Sie basieren auf den typischen Erfolgsfaktoren im Projektmanagement, die über das sogenannte magische Projektdreieck beschrieben werden können. Zudem sind viele Kriterien auf die Überlegungen von Boehm und Turner zurückzuführen, die fünf Faktoren, wie z.B. die Dynamik innerhalb eines Projekts, definiert haben, anhand derer sich klassische und agile Vorgehensmodelle abgrenzen lassen. Als Vorbereitung auf die Analysen und Auswertungen in Kapitel 4, wird in Abschnitt 3.5.2 die Methodik der Nutzwertanalyse anhand des Ansatzes des REFA-Verbands erläutert.

Damit in Abschnitt 4.2 klassische und agile Vorgehensmodelle im Rahmen von Nutzwertanalysen für jede der in Kapitel 2.4 vorgestellten Big-Data-Strategien miteinander verglichen werden können, bedarf es zunächst einer Untersuchung der unterschiedlichen Einführungsprozesse der jeweiligen Strategien. Dabei ist festzustellen, dass es sich bei der Kosteneinsparungsstrategie zumeist um ein technologisch vergleichbar einfaches Projekt handelt. Es orientiert sich ausschließlich an den Einsparungsmöglichkeiten durch Big-Data-Management-Systeme und ist strukturell überschaubar. Die Systemanforderungen liegen zumeist schon bei Projektbeginn vor. Die Strategie der Zeiteinsparung stellt auf technologischer Ebene eine größere Herausforderung dar, da durch Hochleistungskomponenten, wie z.B. In-Memory-Technologie, bekannte Prozesse beschleunigt werden. Insbesondere die Anbindung an bestehende Systeme und die damit verbundene Harmonisierung sind Schlüsselfaktoren bei der Systemeinführung. Bei der Strategie der Entscheidungsunterstützung besteht die Problematik darin, bisher unbekannte interne und externe Datenquellen zu identifizieren und sinnvoll zu verknüpfen. In diesem Zuge müssen jene Daten auf innovative Art und Weise aufbereitet werden. Eine solche Aggregation und Verarbeitung kann nur mit Big-Data-Technologien, wie z.B. Hadoop und MapReduce, erfolgen. Die Strategie auf Basis von Big Data neue Produkte zu entwickeln, kann eher als Produktentwicklungsprozess angesehen werden, da es das Ziel ist marktfähige, kundenorientierte Produkte zu entwerfen, um so neue Umsatzpotentiale zu erschließen. Somit geht es dabei im Gegensatz zu den drei anderen Strategien nicht um eine Optimierungsaufgabe. Bei der Einführung eines Big-Data-Management-Systems unter dieser Prämisse stehen Freiraum für Kreativität und eine ausgewogene Projektkultur im Vordergrund.

Die für die Nutzwertanalyse notwendigen Kriterien werden in Kapitel 4.2.1 in einem Kriterienkatalog zusammengefasst. Die Auswahl der Kriterien erfolgt auf Basis der in den Kapiteln 2 und 3 erarbeiteten Erkenntnisse und insbesondere mit Blick auf die in Abschnitt 3.5.1 aufgestellten typischen Vergleichskriterien. Die Kriterien lassen sich in vier Gruppen einteilen: Ergebniskriterien, Managementkriterien, harte Faktoren und weiche Faktoren. Ergebniskriterien spiegeln die Fähigkeit von Vorgehensmodellen, die Ergebnisdimensionen des magischen Projektdreiecks zu erfüllen, wider. Bei den Managementkriterien handelt es sich um Faktoren, die das Projektmanagement operativ beeinflussen. Harte Faktoren stellen die

messbaren Größen in Projekten, wie z.B. die Produktkomplexität, dar. Schließlich erlauben die weichen Faktoren ein Urteil über die nur schwer messbaren Einflussgrößen wie Motivation, Kommunikation und Kreativität. Jedes Vorgehensmodell erfüllt die definierten Kriterien unterschiedlich gut. Deshalb werden im Anschluss die Kriterien-Erfüllungsgrade agiler und klassischer Vorgehensmodelle plausibilisiert. Dabei stellt sich heraus, dass die in Kapitel 3 beschriebenen Eigenschaften massiven Einfluss auf die Kriterienausprägung haben. Beispielsweise bestätigt sich, dass agile Modelle deutlich mehr Flexibilität zugestehen, als klassische Modelle. Die Erfüllungsgrade der Vorgehensmodelle werden in Punkten gemessen, die im Anschluss in die Nutzwertanalysen einfließen. Weiterhin ist es notwendig die Kriterien-Bedeutsamkeiten für die einzelnen Strategien zu ermitteln. Dabei stellt sich heraus, dass für die Kosten- und Zeiteinsparstrategien vordergründig gute Planbarkeit und Kontrollierbarkeit sowie ausgeprägte Ergebniserreichung wichtig sind. Dem gegenüber sind für die Strategien Entscheidungsunterstützung und Produktinnovation die Kundenzufriedenheit und weiche Faktoren von großer Bedeutung.

Bei Kombination der bisherigen Erkenntnisse ergeben sich im Rahmen der Nutzwertanalysen interessante Ergebnisse, welche sich bereits vorher abgezeichnet haben. Für die Strategie der Kosteneinsparung sind demnach klassische Vorgehensmodelle besser geeignet als agile. Das gleiche Ergebnis gilt für die Zeiteinsparungsstrategie, wobei die Differenz der Nutzwerte dort etwas geringer ausfällt. Möchte man mit Hilfe von Big Data Entscheidungen unterstützen und bessere Prognosen ermöglichen, so sind der Nutzwertanalyse zufolge beide Vorgehensmodell-Typen gleich gut geeignet. Ihre Nutzwerte unterscheidet lediglich eine Differenz von 0,3 Prozentpunkten. Für die Strategie neue Produkte mit Big Data entwickeln zu wollen sind agile Vorgehensmodelle vorzuziehen. Die Ergebnisse werden in Kapitel 4.3 in Handlungsempfehlungen ausformuliert und dabei detailliert Bezug auf die Eigenschaften der Vorgehensmodelle und Big-Data-Strategien genommen.

Zur Validierung der Ergebnisse der Nutzwertanalysen wird im fünften Kapitel anhand einer Big-Data-Strategie exemplarisch im Rahmen einer Fallstudie die Projektdurchführung mit beiden Vorgehensmodellen simuliert und auf ihren Erfolg untersucht. Betrachtungsgegenstand ist dabei die Kosteneinsparungsstrategie, für welche der Nutzwertanalyse zufolge klassische Vorgehensmodelle besser geeignet sind. Zunächst werden dazu fiktive Rahmenbedingungen aufgestellt und eine plausible Ausgangssituation geschaffen. Im vorliegenden Beispiel handelt es sich dabei um eine Großbank, die ihr Datenarchiv aufgrund stark ansteigender Datenvolumina auf ein kostengünstiges Hadoop-Cluster umstellen möchte. Es zeigt sich, dass eine gut strukturierte Vorgehensweise für diesen Betrachtungsfall besser geeignet ist als ein hochdynamisches Projektmanagement mit der Fähigkeit schnell auf Anforderungsänderungen reagieren zu können. Demzufolge bestätigt die Fallstudie die Nutzwertanalysenergebnisse zumindest für die Kosteneinsparungsstrategie und



validiert damit gleichzeitig die Auswahl der Bewertungskriterien. Folglich ist der Kriterienkatalog hinreichend gut geeignet, um typische Herausforderungen und Probleme im IT-Projektmanagement beschreiben zu können.

## 6.2 Fazit

Betrachtet man das für die Ausarbeitung gesetzte Gesamtziel, so kann konstatiert werden, dass mit Hilfe eines fundierten Kriterienkatalogs und einer strukturierten Nutzwertanalyse eine Aussage darüber getroffen werden kann, welches Vorgehensmodell bei der Verfolgung unterschiedlicher Big-Data-Strategien vorzuziehen ist, damit ein Big-Data-Management-System möglichst effizient eingeführt werden kann. Dabei ist jedoch zu berücksichtigen, dass die untersuchten Strategien nach Davenport isoliert betrachtet werden. In der Unternehmenspraxis ist es nicht ungewöhnlich, dass Mischformen oder gar mehrere Strategien parallel verfolgt werden. Entsprechend muss die Aussagekraft der Analysen und Bewertungen für derartige Strategieformen noch überprüft werden. Darüber hinaus ist vorstellbar, dass bei Veränderung des Ordnungskriteriums, das bedeutet bei Analyse von z.B. Branchen, Big-Data-Arten oder Unternehmensgrößen anstelle von Big-Data-Strategien, veränderte Analyseergebnisse zum Vorschein kommen können. Aus diesem Grund sind im vorliegenden Fall strikte Rahmenbedingungen für die Anwendbarkeit der Ergebnisse gesetzt, sodass kein Anspruch auf Allgemeingültigkeit vorliegt. Nichtsdestotrotz erlaubt die angewandte Struktur eine Aussage über die relative Vorteilhaftigkeit eines Vorgehensmodell-Typen zu treffen.

Genannte Vorgehensmodell-Typen werden stark abstrahiert in die zwei Gruppen agil und klassisch klassifiziert. Dies kommt einer starken Aggregation von Informationen gleich, da innerhalb der Gruppen verschiedene Modelle existieren, die wiederum unterschiedliche Charakteristika aufweisen. Ungeachtet dessen lassen sich für sämtliche Modelle eines Typs immer wieder auftretende Eigenschaften auffinden. Bei klassischen Vorgehensmodellen ist dies beispielsweise der sequentielle Ablauf in klar abgrenzbaren Phasen, bei agilen Modellen die durchgängige Kundenorientierung und hohe Flexibilität. Auf dieser Aggregationsebene sind alle Analysen erfolgt, sodass eine hinreichend gute Aussagekraft erhalten bleibt.

Um die fundierten Handlungsempfehlungen auf Basis der Nutzwertanalysen zu untermauern, war es ein Teilziel den aktuellen Stand der Technik in den Themenkomplexen Big Data und Vorgehensmodelle im IT-Projektmanagement ausführlich darzustellen. Im Bereich Big Data ist dabei besonders auffällig, dass nach wie vor keine einheitliche Begriffsdefinition vorliegt. Genauso ist sich die einschlägige Literatur uneinig darüber, welche Zielrichtungen Big Data verfolgen wird. Es hat sich jedoch als praktikabel erwiesen Big-Data-Anwendungen in den Dimensionen Volume, Velocity und Variety zu charakterisieren. Falls nötig werden darüber

hinaus auch weitergehende Dimensionen, wie z.B. Veracity oder Value, herangezogen. Der eigentliche Mehrwert von Big Data liegt dabei jedoch häufig in der Kombination von Daten verschiedener interner und externer Quellen zur Gewinnung neuer Erkenntnisse. Bisher stellt dabei die Identifikation jener Quellen sowie die Verarbeitung von insbesondere unstrukturierten Daten die größte konzeptionelle Herausforderung für Unternehmen dar. Betrachtet man das Management von Big-Data-Initiativen, so treten die in Abschnitt 2.5 vorgestellten Herausforderungen in den Vordergrund.

Wie effektiv ein Big-Data-Management-System ist und welchen Mehrwert es beiträgt, hängt zu einem großen Teil auch von der bestehenden Datenarchitektur ab. Lassen sich Big-Data-Technologien gut darin integrieren, sind erstaunliche Ergebnisse möglich. Genauso kann eine unflexible Datenarchitektur jedoch auch dazu führen, dass jene Technologien ausgebremst werden. Übergreifend stellen Autoren dazu fest, dass Big Data die bestehenden Systeme nicht ablösen wird. Vielmehr geht es darum, produktive Koexistenzen von Systemen zu entwickeln und zu etablieren.

Zusammenfassend ist demnach festzuhalten, dass im Rahmen dieser Ausarbeitung ein umfassender Überblick über den aktuellen Stand der Technik insbesondere im Themengebiet Big Data gegeben wurde. Zudem erlauben der aufgestellte Kriterienkatalog sowie die Systematik zur Bewertung von Big-Data-Strategien eine adäquate Beurteilung der Vorteilhaftigkeit von IT-Projektmanagement Vorgehensmodellen bei der Einführung von Big-Data-Management-Systemen.

### **6.3 Ausblick**

Die erarbeiteten Ergebnisse liefern diverse Ansätze, die für weitere tiefergehende Forschungszwecke genutzt werden können. Wie bereits im Fazit angeklungen ist, verfolgen Unternehmen nicht selten Mischformen der hier genannten vier Big-Data-Strategien. In einem nächsten Schritt macht es Sinn diese noch nicht berücksichtigten Strategien zu identifizieren und aufsattelnd auf dem bestehenden Kriterienkatalog aus Kapitel 4.2.1 zu analysieren. Es kann darüber hinaus auch dazu kommen, dass Big-Data-Initiativen aus mehr als einem Big-Data-Projekt bestehen. Um sämtliche Projekte im Rahmen eines Big-Data-Programmmanagements zu organisieren, bedarf es detaillierter Erkenntnisse über die Wechselwirkungen von klassischen und agilen Vorgehensmodellen. Geht man davon aus, dass verschiedene Vorgehensmodell-Typen zur Anwendung kommen, stellt sich die Frage, inwieweit diese heterogenen Modelle innerhalb eines Programms harmonisiert und kontrolliert werden können.

Ein weiterer Ansatz für die Fortführung der Forschungstätigkeiten ist die granulare Untersuchung der Eignung verschiedener klassischer und agiler Vorgehensmodelle. In der vorliegenden Ausarbeitung wurden sämtliche Modelle aggregiert zu klassischen und agilen

Vorgehensmodellen zusammengefasst. Analog zu den genannten Strategie-Mischformen lassen sich, aufbauend auf dem entwickelten Kriterienkatalog, auch einzelne Vorgehensmodelle, wie z.B. das klassische V-Modell XT untersuchen. Die Erweiterbarkeit der Untersuchung ist durch den flexibel einsetzbaren Kriterienkatalog gegeben.

Betrachtet man die Ergebnisse aus Sicht der Praxis, so liefert die Arbeit fundierte Handlungsempfehlungen für spezifische Big-Data-Strategien. Die Fähigkeit zur Implementierung von Big-Data-Management-Systemen wird zum Erfolg von Unternehmen beitragen und stellt einen substantiellen Wettbewerbsvorteil dar. Zudem liefert die Aufarbeitung der vorhandenen Literatur zum Themenkomplex Big Data ein stabiles Gerüst innerhalb der exponentiell wachsenden Zahl von Publikationen zu diesem Thema. Da sich Organisationen laufend damit beschäftigen, möglichst effizient innovative Technologien einzuführen, können die erarbeiteten Resultate eine wichtige Orientierungshilfe darstellen.

Zurückgreifend auf den in der Einleitung zitierten Emerging Technologies Hype Cycle 2013 des Marktforschungsunternehmens Gartner, kann eine Prognose über die weitere Entwicklung von Big Data und damit verbundenen Technologien gegeben werden. Überproportional viele IT-Experten beschäftigen sich derzeit mit diesem Themenkomplex und liefern zahlreiche Anwendungsbeispiele, was typisch für die derzeitige Position an der Spitze der aufgeblähten Erwartungen ist. In Kürze werden viele dieser Ideen der Desillusionierung der Märkte zum Opfer fallen. Trotzdem wird sich Big Data Schritt für Schritt hin zu einer produktiv nutzbaren Technologie entwickeln und für viele Analysen unabdingbar werden. Dabei ist jedoch zu berücksichtigen, dass im ersten Schritt nicht alle Sektoren und Branchen gleichermaßen intensiv betroffen sind. Langfristig hingegen sehen sich sämtliche Sektoren mit den Herausforderungen von Big Data konfrontiert. Dabei muss jede Organisation, die in Zukunft wettbewerbsfähig bleiben möchte, ihre individuellen Big-Data-Potentiale erkennen und für sich nutzen. Blickt man auf die vorhandenen Rahmenbedingungen zum Aufbau dieser Expertise in Organisationen, so muss festgestellt werden, dass Deutschland gegenüber anderen Ländern außerhalb der EU datenschutzrechtlich noch Nachteile hat. Zur Eliminierung dieser Barriere bedarf es einer umfassenden Umstrukturierung der Gesetzgebung, wozu der Diskurs möglichst zeitnah angestoßen werden sollte. Dieser Diskurs sollte darüber hinaus die Frage beinhalten, wie sich deutsche Unternehmen im ausgesprochen volatilen Big-Data-Markt positionieren wollen.

## Quellenverzeichnis

### Literaturquellen

- [BKG14]** Bachmann, R., Kemper, G., Gerzer, T.: *Big Data – Fluch oder Segen?: Unternehmen im Spiegel gesellschaftlichen Wandels*, Auflage: 2014, Heidelberg: mitp, 2014
- [Boe81]** Boehm, B.: *Software Engineering Economics*, London: Prentice-Hall Inc., 1981
- [BT04]** Boehm, B., Turner, R.: *Balancing Agility and Discipline: A Guide for the Perplexed*, Boston: Pearson Education, 2004
- [Dav14]** Davenport, T. H.: *big data @ work – Chancen erkennen, Risiken verstehen*, 1. Auflage. München: Vahlen, 2014
- [Hol07]** Holzbaur, U.: *Entwicklungsmanagement – Mit hervorragenden Produkten zum Markterfolg*, 1. Auflage. Berlin: Springer, 2007
- [Kin14]** King, S.: *Big Data - Potential und Barrieren der Nutzung im Unternehmenskontext*, 1. Auflage. Wiesbaden: Springer Fachmedien, 2014
- [Kle13]** Kleuker, S.: *Grundkurs Software-Engineering mit UML: Der pragmatische Weg zu erfolgreichen Softwareprojekten*, 3. Auflage. Wiesbaden: Springer Vieweg, 2013
- [MC13]** Mayer-Schönberger, V., Cukier, K.: *Big Data – A Revolution That Will Transform How We Live, Work and Think*, 1. Auflage. London: John Murray, 2013
- [MCD13]** Minelli, M., Chambers, M., Dhiraj, A.: *Big Data, Big Analytics: Emerging Business Intelligence and Analytic Trends for Today's Businesses*, Hoboken (New Jersey): John Wiley & Sons, Inc., 2013
- [Pic08]** Pichler, R.: *Scrum – Agiles Projektmanagement erfolgreich einsetzen*, 1. Auflage. Heidelberg: dpunkt, 2008
- [Ref91]** REFA – Verband für Arbeitsgestaltung, Betriebsorganisation und Unternehmensentwicklung: *Methodenlehre der Betriebsorganisation, Planung und Steuerung Teil 5*, München: Hanser, 1991
- [RF08]** Ruf, W., Fittkau, T.: *Ganzheitliches IT-Projektmanagement – Wissen, Praxis, Anwendungen*, München: Oldenbourg, 2008
- [Roy70]** Royce, W. W. in IEEE Computer Society Press: *Managing the Development of Large Software Systems: Concepts and Techniques*, Los Alamitos (Kalifornien): Proc WESCON, 1970
- [SDB\*10]** Schatten, A., Demolsky, M., Winkler, D., Biffel, S., Gostischa-Franta, E., Östreicher, T.: *Best Practice Software-Engineering – Eine praxiserprobte Zusammenstellung von komponentenorientierten Konzepten, Methoden und Werkzeugen*, Heidelberg: Spektrum Akademischer Verlag, 2010
- [Tie14]** Tiemeyer, E. (Hrsg.): *Handbuch IT-Projektmanagement - Vorgehensmodelle, Managementinstrumente, Good Practices*, 2. Auflage, München: Hanser, 2014

- 
- [Tre12]** Trepper, T.: *Agil-systemisches Softwareprojektmanagement*, Wiesbaden: Springer Gabler, 2012
- [Wei13]** Weichert, T.: *Big Data – eine Herausforderung für den Datenschutz*. In: Geiselberger, H.; Moorstedt, T. (eds.): *Big Data – Das neue Versprechen der Allwissenheit*. Berlin: Suhrkamp, 2013
- [WM07]** Wieczorrek, H. W., Mertens, P.: *Management von IT-Projekten: Von der Planung zur Realisierung*, 2. Auflage. Wiesbaden: Springer, 2007
- [Zan70]** Zangemeister, C.: *Nutzwertanalyse in der Systemtechnik – Eine Methodik zur multidimensionalen Bewertung und Auswahl von Projektalternativen*. 2. Auflage. München: Wittemann, 1971

**Internetquellen**

- [And08]** Anderson, C., Wired Magazine: 16.07: *The End of Theory: The Data Deluge Makes the Scientific Method Obsolete*, 2008.  
[http://archive.wired.com/science/discoveries/magazine/16-07/pb\\_theory](http://archive.wired.com/science/discoveries/magazine/16-07/pb_theory)  
zuletzt abgerufen am 20.11.2014
- [BBB\*01]** Beck, K., Beedle, M., Bennekum, A. van, Cockburn, A., Cunningham, W., Fowler, M., Grenning, J., Highsmith, J., Hunt, A., Jeffries, R., Kern, J., Marick, B., Martin, R. C., Mellor, S., Schwaber, K., Sutherland, J., Thomas, D.: *Manifest für Agile Softwareentwicklung*, 2001  
<http://agilemanifesto.org/iso/de/>  
zuletzt abgerufen am 16.02.2015
- [Bit12]** Bundesverband Informationswirtschaft, Telekommunikation und neue Medien (BITKOM), *Big Data im Praxiseinsatz – Szenarien, Beispiele, Effekte*, 2012.  
[http://www.bitkom.org/de/publikationen/38337\\_73446.aspx](http://www.bitkom.org/de/publikationen/38337_73446.aspx)  
zuletzt abgerufen am 20.11.2014
- [Bit13]** Bundesverband Informationswirtschaft, Telekommunikation und neue Medien (BITKOM), *Management von Big-Data-Projekten*, 2013  
[http://www.bitkom.org/files/documents/LF\\_big\\_data2013\\_web.pdf](http://www.bitkom.org/files/documents/LF_big_data2013_web.pdf)  
zuletzt abgerufen am 31.01.2015
- [Bit14a]** Bundesverband Informationswirtschaft, Telekommunikation und neue Medien (BITKOM), *Potenziäle und Einsatz von Big Data – Ergebnisse einer repräsentativen Befragung von Unternehmen in Deutschland*, 2014.  
[http://www.bitkom.org/de/markt\\_statistik/64026\\_79283.aspx](http://www.bitkom.org/de/markt_statistik/64026_79283.aspx)  
zuletzt abgerufen am 10.11.2014
- [Bit14b]** Bundesverband Informationswirtschaft, Telekommunikation und neue Medien (BITKOM), *Big-Data-Technologien – Wissen für Entscheider*, 2014  
[http://www.bitkom.org/de/publikationen/38337\\_78776.aspx](http://www.bitkom.org/de/publikationen/38337_78776.aspx)  
zuletzt abgerufen am 15.01.2014
- [Del14]** Dell, TNS: *Dell Global Technology Adoption Index – Revealing decision points around technology adoption, use and benefits in midsize organizations*, 2014.  
[https://kapos-files-prod.s3.amazonaws.com/uploads/direct/1415199563-23-1043/Executive\\_Summary\\_Global\\_Technology\\_Adoption\\_Index.PDF](https://kapos-files-prod.s3.amazonaws.com/uploads/direct/1415199563-23-1043/Executive_Summary_Global_Technology_Adoption_Index.PDF)  
zuletzt abgerufen am 07.11.2014
- [DD13]** Davenport, T. H., Dyché, J., SAS Institute Inc.: *Big Data in Big Companies*, 2013  
<http://www.sas.com/resources/asset/Big-Data-in-Big-Companies.pdf>  
zuletzt abgerufen am 30.01.2015
- [EV10]** Eveleens, J.L., Verhoef, C., Vrije Universiteit Amsterdam in IEEE Software – Focus Project Management: *The Rise and Fall of the Chaos Report Figures*, 2010  
<http://www.cs.vu.nl/~x/chaos/chaos.pdf>  
zuletzt abgerufen am 12.02.2015

- [Far06]** Farrance, R.: *Timeline: 50 Years of Hard Drives*, 2006.  
<http://www.pcworld.com/article/127105/article.html>  
zuletzt abgerufen am 10.12.2014
- [Flo13]** Floyer, D.: *Financial Comparison of Big Data MPP Solution and Data Warehouse Appliance*, 2013  
[http://wikibon.org/wiki/v/Financial Comparison of Big Data MPP Solution and Data Warehouse Appliance](http://wikibon.org/wiki/v/Financial_Comparison_of_Big_Data_MPP_Solution_and_Data_Warehouse_Appliance)  
zuletzt abgerufen am 22.01.2015
- [For14]** Forzley, S. in *DataInformed – Big Data and Analytics in the Enterprise: Big Data ROI: Does the Payoff Match the Potential?*, 2014  
<http://data-informed.com/big-data-roi-payoff-match-potential/>  
zuletzt abgerufen am 22.01.2015
- [Fra12]** Fraunhofer-Institut für Intelligente Analyse- und Informationssysteme IAIS: *Fraud Detection in Kreditkartentransaktionen*, 2012.  
[http://www.iais.fraunhofer.de/fraud\\_kreditkarten.html](http://www.iais.fraunhofer.de/fraud_kreditkarten.html)  
zuletzt abgerufen am 15.12.2014
- [Gar11a]** Gartner, Inc.: *Gartner's 2011 Hype Cycle Special Report Evaluates the Maturity of 1.900 Technologies*, 2011.  
<http://www.gartner.com/newsroom/id/1763814>  
zuletzt abgerufen am 20.11.2014
- [Gar11b]** Gartner, Inc.: *Gartner Says Solving 'Big Data' Challenge Involves More Than Just Managing Volumes of Data*, 2011  
<http://www.gartner.com/newsroom/id/1731916>  
zuletzt abgerufen am 17.12.2014
- [Gar13]** Gartner, Inc.: *Gartner's 2013 Hype Cycle for Emerging Technologies Maps Out Evolving Relationship Between Humans and Machines*, 2013.  
<http://www.gartner.com/newsroom/id/2575515>  
zuletzt abgerufen am 07.11.2014
- [HL13]** Heilmann, D., Liegl, T., Handelsblatt Research Institute: *Big Data und Datenschutz – Der Umgang der Deutschen mit persönlichen Daten und die Konsequenzen für den Einsatz von Big-Data-Analysen*, 2013.  
<http://research.handelsblatt.com/datenschutz-und-big-data-was-die-deutschen-denken/>  
zuletzt abgerufen am 20.11.2014
- [HSV\*12]** Haberl, P., Spillner, A., Vosseberg, K., Winter, M.: *Umfrage 2011: Softwaretest in der Praxis*, 2012  
[http://www.softwaretestumfrage.de/Broschuere Softwaretest in der Praxis.pdf](http://www.softwaretestumfrage.de/Broschuere_Softwaretest_in_der_Praxis.pdf)  
zuletzt abgerufen am 11.02.2015

- [Ibm12]** International Business Machines Corporation (IBM), IBM Institute for Business Value in Zusammenarbeit mit Saïd Business School der Universität Oxford: *Analytics: Big Data in der Praxis – Wie innovative Unternehmen ihre Datenbestände effektiv nutzen*, 2012  
<http://www-935.ibm.com/services/de/gbs/thoughtleadership/studie-bigdata.html>  
zuletzt abgerufen am 11.02.2015
- [Ibm14]** International Business Machines Corporation (IBM): *What is a data scientist*, 2014  
<http://www-01.ibm.com/software/data/infosphere/data-scientist/>  
zuletzt abgerufen am 20.01.2015
- [Idc13a]** International Data Corporation (IDC): *Big Data – Business Value in deutschen Unternehmen auf dem Prüfstand*, 2013  
<http://idc.de/de/ueber-idc/press-center/57064-idc-studie-big-data-business-value-in-deutschen-unternehmen-auf-dem-prufstand>  
zuletzt abgerufen am 09.01.2015
- [Idc13b]** International Data Corporation (IDC), Kapow Software: *Making Big Data Insights Easily Consumable and Actionable*, 2013  
<http://info.kapowsoftware.com/IDGSurveyConsumableBigData#.VMFK5Eccwz>  
w  
zuletzt abgerufen am 22.01.2015
- [Idc14]** International Data Corporation (IDC), EMC Corporation: *EMC Digital Universe with Research & Analysis by IDC – The Digital Universe of Opportunities: Rich Data and the Increasing Value of the Internet of Things*, 2014.  
<http://www.emc.com/leadership/digital-universe/2014iview/executive-summary.htm>  
zuletzt abgerufen am 07.11.2014
- [Int13]** Intel IT Center: *Planning Guide: Getting Started with Big Data*, 2013  
<http://www.intel.com/content/www/us/en/big-data/getting-started-with-big-data-planning-guide.html>  
zuletzt abgerufen am 23.01.2015
- [Ita15]** it-agile GmbH: *Methodenwissen – Scrum*, 2015  
<http://www.it-agile.de/wissen/methoden/scrum/>  
zuletzt abgerufen am 17.02.2015
- [Jam12]** James, J.: *How Much Data is Created Every Minute?* in Domo business intelligence (BI) blog, 2012.  
<http://www.domo.com/blog/2012/06/how-much-data-is-created-every-minute/>  
zuletzt abgerufen am 17.12.2014
- [KTH13]** Klein, D., Tran-Gia, P., Hartmann, M.: *Big Data* in Informatiklexikon der Gesellschaft für Informatik, 2013.  
<http://www.gi.de/service/informatiklexikon/detailansicht/article/big-data.html>  
zuletzt abgerufen am 17.12.2014



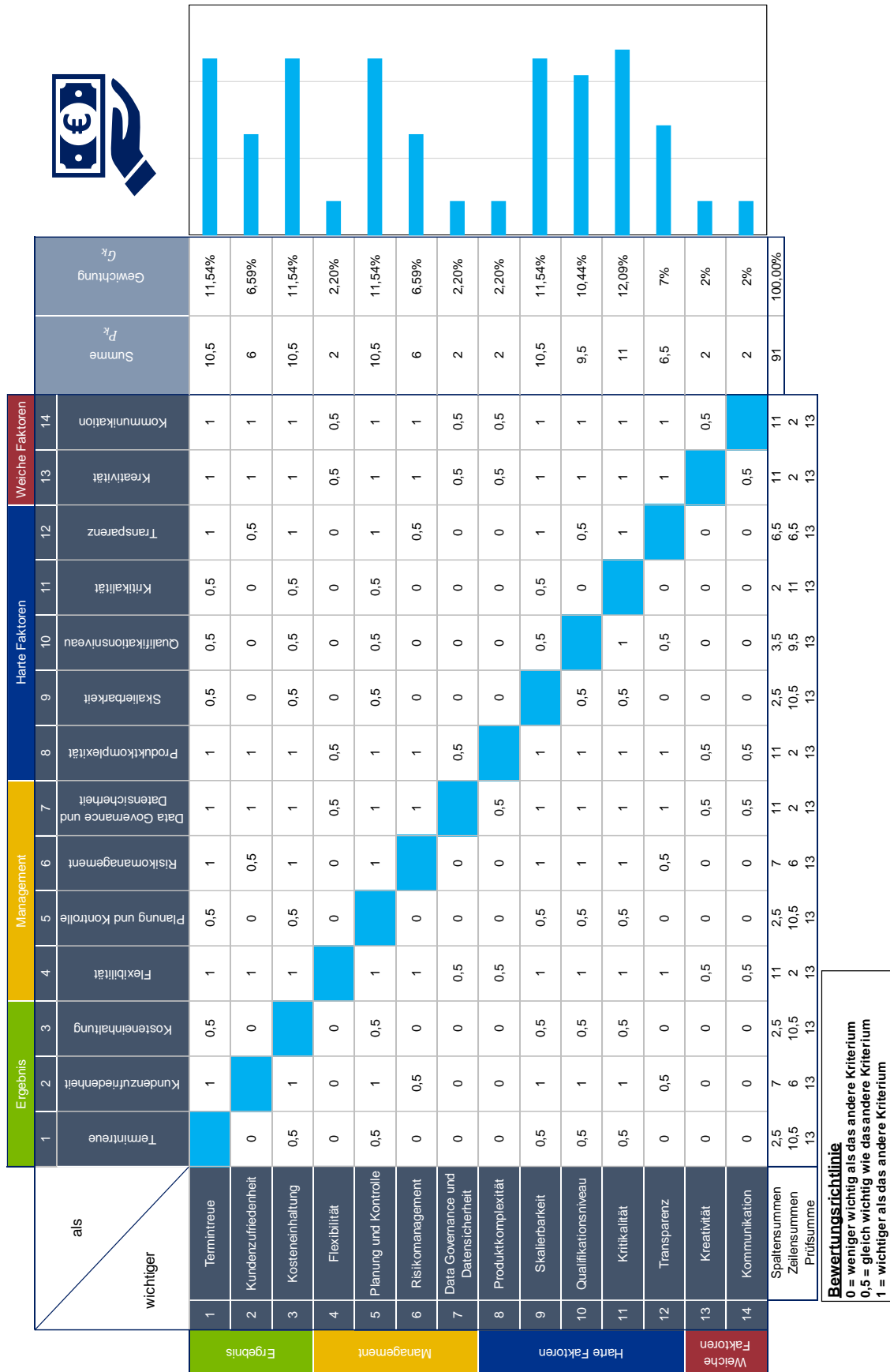
- [Lan01]** Laney, D.: *3D Data management: Controlling Data Volume, Velocity and Variety* in Application Delivery Strategies, META Group Inc., 2001.  
<http://blogs.gartner.com/doug-laney/files/2012/01/ad949-3D-Data-Management-Controlling-Data-Volume-Velocity-and-Variety.pdf>  
zuletzt abgerufen am 17.12.2014
- [Lef09]** Leffingwell, D.: *Scaling Software Agility – Best practices for Large Enterprises*, 2009  
[http://agile2009.agilealliance.org/files/session\\_pdfs/Scaling%20Software%20Agility%20Overview%20Agile%202009.pdf](http://agile2009.agilealliance.org/files/session_pdfs/Scaling%20Software%20Agility%20Overview%20Agile%202009.pdf)  
zuletzt abgerufen am 19.02.2015
- [LXS+13]** Lehmann, N., Xuan, R. P., Schatto, C., Weidner, C. in Usability und UX Wiki, TU Berlin: *Klassische Vorgehensmodelle*, 2013  
[http://www.uselab.tuberlin.de/wiki/index.php/%E2%80%9EKlassische%E2%80%9C\\_Vorgehensmodelle#Anwendungsbereiche](http://www.uselab.tuberlin.de/wiki/index.php/%E2%80%9EKlassische%E2%80%9C_Vorgehensmodelle#Anwendungsbereiche)  
zuletzt abgerufen am 16.02.2015
- [Man14]** Manhart, Klaus, Computerwoche in Zusammenarbeit mit IBM: *Data Governance im Big-Data-Zeitalter*, 2014  
<http://www.computerwoche.de/a/data-governance-im-big-data-zeitalter,3062474>  
zuletzt abgerufen am 21.01.2015
- [MCB+11]** Manyika, J., Chui, M., Brown, B., Bughin, J., Dobbs, R., Roxburgh, C., Hung Byers, A., McKinsey Global Institute: *Big Data: The next frontier for innovation, competition and productivity*, 2011.  
[http://www.mckinsey.com/insights/business\\_technology/big\\_data\\_the\\_next\\_frontier\\_for\\_innovation](http://www.mckinsey.com/insights/business_technology/big_data_the_next_frontier_for_innovation)  
zuletzt abgerufen am 07.11.2014
- [Mcb12]** McBurney, V., Toolbox for IT: *The Origin and Growth of Big Data Buzz*, 2012.  
<http://it.toolbox.com/blogs/infosphere/the-origin-and-growth-of-big-data-buzz-51509>  
zuletzt abgerufen am 21.11.2014
- [Mor65]** Moore, G. E., *Cramming more components onto integrated circuits*, 1965. In: Electronics Magazine p. 4.,  
<http://www.cs.utexas.edu/~fussell/courses/cs352h/papers/moore.pdf>  
zuletzt abgerufen am 10.12.2014
- [Nas12]** Nash, T.: *Breaking the Barrier of Big Data Analytics in BI in information management*, 2012.  
<http://www.information-management.com/newsletters/big-data-bi-value-roi-mcdonalds-david-wallerstein-10022871-1.html>  
zuletzt abgerufen am 18.12.2014
- [Nes12]** Networked European Software and Service Initiative (NESSI): *Big Data – A New World of Opportunities*, 2012  
[http://www.nessi-europe.com/Files/Private/NESSI\\_WhitePaper\\_BigData.pdf](http://www.nessi-europe.com/Files/Private/NESSI_WhitePaper_BigData.pdf)  
zuletzt abgerufen am 17.12.2014

- [PCC+12]** Page, C., Campbell, R., Coggeshall, S., Gillespie, E., Johnson, R., Olson, M. & Perkins, P.: *Demystifying Big Data: A Practical Guide To Transforming The Business of Government* Microsoft. Washington, DC, 2012.  
<http://www-304.ibm.com/industries/publicsector/fileserve?contentid=239170>  
zuletzt abgerufen am 17.12.2014
- [PSC+14]** Platt, J., Souza, R., Checa, E., Chabaladas, R., bcp.perspectives by The Boston Consulting Group: *Seven Ways to Profit from Big Data as a Business*, 2014  
[https://www.bcgperspectives.com/content/articles/information\\_technology\\_strategy\\_digital\\_economy\\_seven\\_ways\\_profit\\_big\\_data\\_business/](https://www.bcgperspectives.com/content/articles/information_technology_strategy_digital_economy_seven_ways_profit_big_data_business/)  
zuletzt abgerufen am 16.01.2015
- [PT07]** Padberg, F., Tichy, W. in *Wirtschaftsinformatik Heft 3/2007: Schlanke Produktionsweisen in der modernen Softwareentwicklung*, 2007  
[https://www.st.cs.uni-saarland.de/~padberg/artikel/padberg\\_wi07.pdf](https://www.st.cs.uni-saarland.de/~padberg/artikel/padberg_wi07.pdf)  
zuletzt abgerufen am 17.02.2015
- [SKM+12]** Schäfer, A., Knapp, M., May, M., Voß, A., Fraunhofer-Institut für Intelligente Analyse- und Informationssysteme IAIS: *Big Data – Vorsprung durch Wissen – Innovationspotenzialanalyse*, St. Augustin, 2012  
[http://www.iais.fraunhofer.de/fileadmin/user\\_upload/Abteilungen/KD/uploads\\_BDA/Innovationspotenzialanalyse\\_Big-Data\\_FraunhoferIAIS\\_2012.pdf](http://www.iais.fraunhofer.de/fileadmin/user_upload/Abteilungen/KD/uploads_BDA/Innovationspotenzialanalyse_Big-Data_FraunhoferIAIS_2012.pdf)  
zuletzt abgerufen am 16.01.2015
- [Sta94]** The Standish Group in Project Smart: *Chaos Report*, 1994  
<http://www.projectsmart.co.uk/docs/chaos-report.pdf>  
zuletzt abgerufen am 12.02.2015
- [TAN+13]** Trecker, A., Arslan, A., Nagel, C., Seigies, K. in *Usability und UX Wiki*, TU Berlin: *Agile Entwicklungsmethoden*, 2013  
[http://www.uselab.tuberlin.de/wiki/index.php/Agile\\_Entwicklungsmethoden#Agile\\_Bewertung](http://www.uselab.tuberlin.de/wiki/index.php/Agile_Entwicklungsmethoden#Agile_Bewertung)  
zuletzt abgerufen am 17.02.2015
- [Tho14]** Thomas, G, The Data Governance Institute: *Defining Data Governance*, 2014  
<http://www.datagovernance.com/defining-data-governance/>  
zuletzt abgerufen am 20.01.2015
- [Vor13]** Vorhies, B.: *The Big Deal About Big Data: What's Inside – Structured, Unstructured, and Semi-Structured Data* in *Data Magnum Blog*, 2013.  
<http://data-magnum.com/the-big-deal-about-big-data-whats-inside-structured-unstructured-and-semi-structured-data/>  
zuletzt abgerufen am 18.12.2014
- [Web00]** Weber, T., Wall Street Journal: *Protecting Consumers' Privacy May Start to Pay Off on Internet*, 12.06.2000  
<http://www.wsj.com/articles/SB96075991873321400>  
zuletzt abgerufen am 21.01.2015

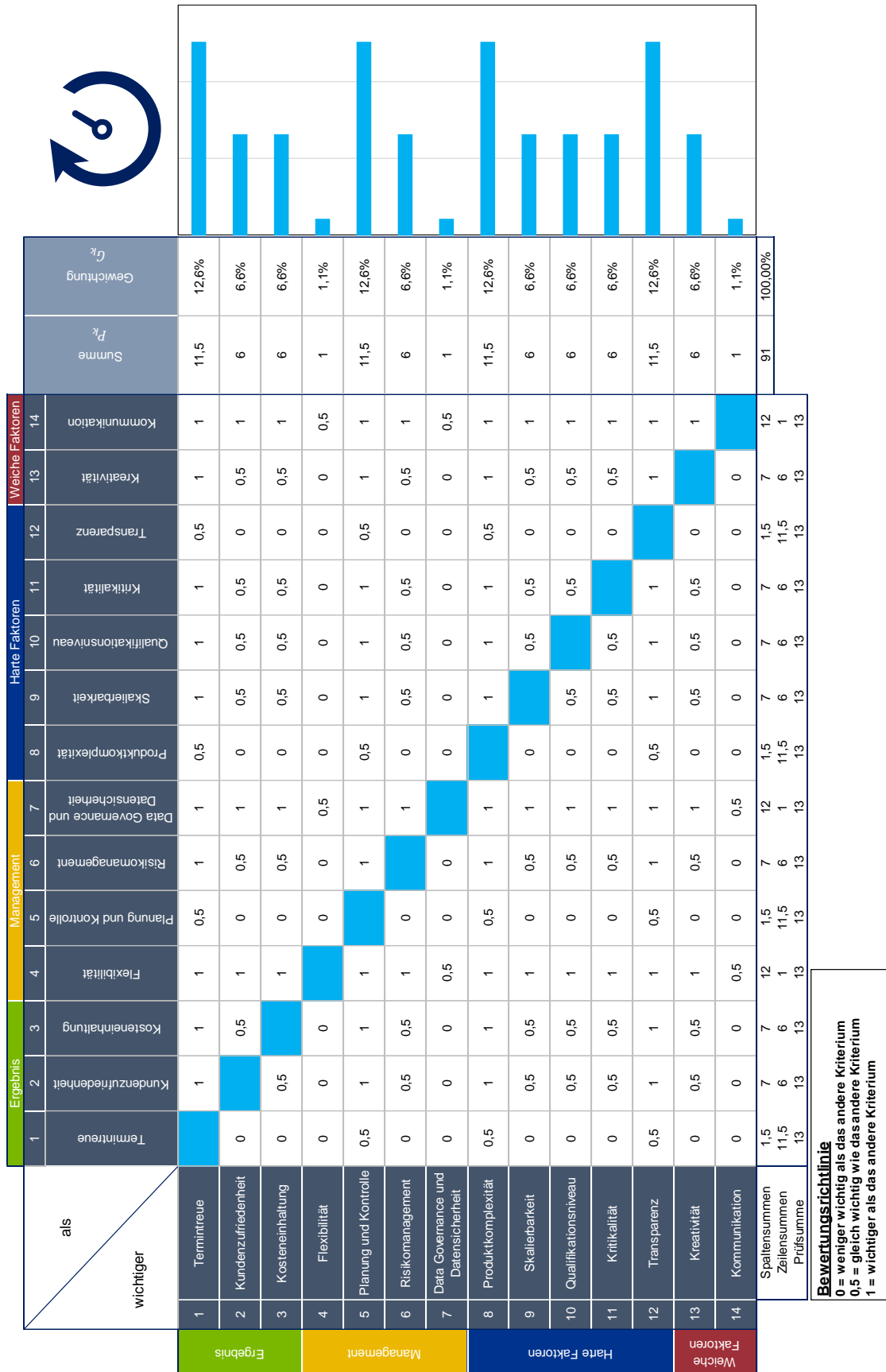
## **Anhangsverzeichnis**

- Anhang 1:** Paarvergleich zur Gewichtungsermittlung – Kosteneinsparung durch Big Data
- Anhang 2:** Paarvergleich zur Gewichtungsermittlung – Zeiteinsparung mit Hilfe von Big Data
- Anhang 3:** Paarvergleich zur Gewichtungsermittlung – Entscheidungsunterstützung und Prognosen durch Big Data
- Anhang 4:** Paarvergleich zur Gewichtungsermittlung – neue Produkte auf Basis von Big Data
- Anhang 5:** Verzeichnis elektronischer Anhänge
- Anhang 6:** Eidesstattliche Versicherung

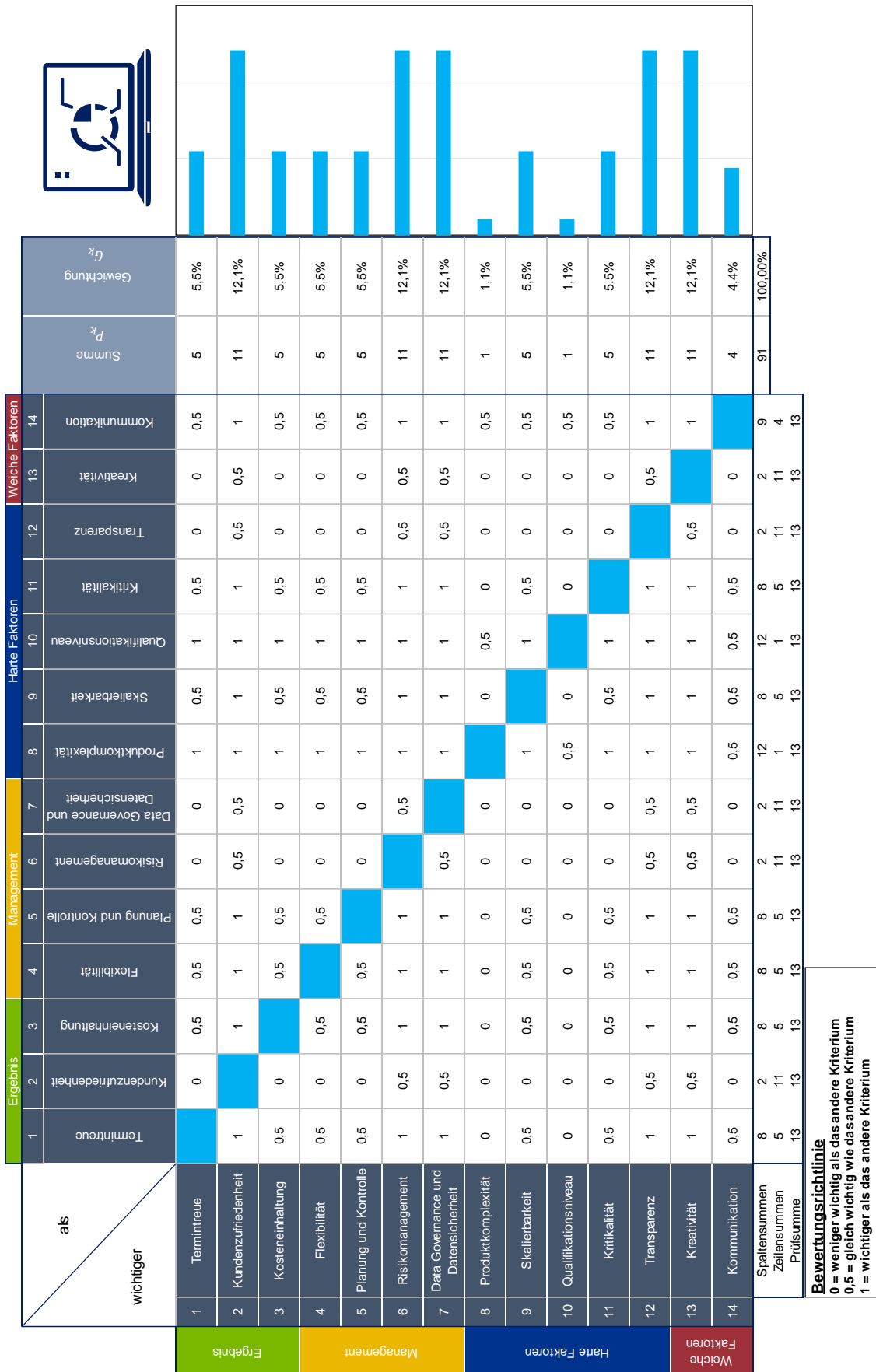
**Anhang 1: Paarvergleich zur Gewichtungsermittlung – Kosteneinsparung durch Big Data**



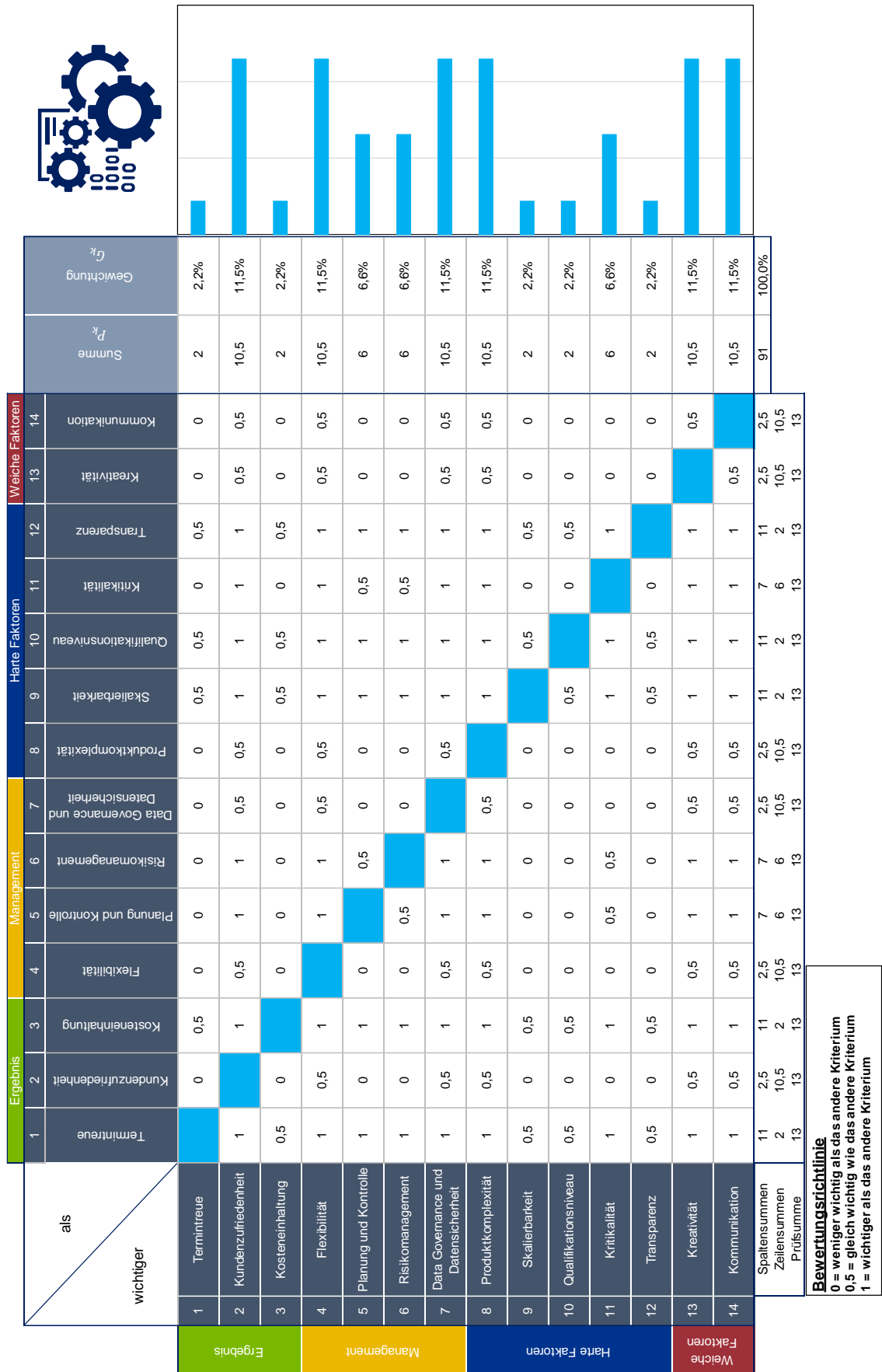
Anhang 2: Paarvergleich zur Gewichtungsermittlung – Zeiteinsparung mit Hilfe von Big Data



**Anhang 3: Paarvergleich zur Gewichtungsermittlung – Entscheidungsunterstützung und Prognosen durch Big Data**



**Anhang 4: Paarvergleich zur Gewichtungsermittlung – neue Produkte auf Basis von Big Data**



**Anhang 5:** Verzeichnis elektronischer Anhänge**Verzeichnis elektronischer Anhänge**

Anhang 5.A	Masterarbeit_Grafe_Jan-Niklas.pdf
Anhang 5.B	Masterarbeit_Grafe_Jan-Niklas.docx
Anhang 5.C	Kriterienkatalog.xlsx
Anhang 5.D	Kriterien-Erfüllungsgrad.xlsx
Anhang 5.E	Kriterien-Bedeutsamkeit.xlsx
Anhang 5.F	Nutzwertanalyse_Kosteneinsparung.xlsx
Anhang 5.G	Nutzwertanalyse_Zeiteinsparung.xlsx
Anhang 5.H	Nutzwertanalyse_Entscheidungsunterstützung.xlsx
Anhang 5.I	Nutzwertanalyse_neue_Produnkte.xlsx
Anhang 5.J	Nutzwertanalyse_Zusammenfassung.xlsx
Anhang 5.K	Burndown_Chart_Scrum_Kosteneinsparung.xlsx



**Anhang 6:** Eidesstattliche Versicherung**Eidesstattliche Versicherung**

Grafe, Jan-Niklas

165400

Name, Vorname

Matr.-Nr.

Ich versichere hiermit an Eides statt, dass ich die vorliegende Masterarbeit mit dem Titel

**„Ein Vergleich agiler und klassischer Vorgehensmodelle des IT-Projektmanagements bei der Einführung von Big-Data-Management-Systemen“**

selbstständig und ohne unzulässige fremde Hilfe erbracht habe. Ich habe keine anderen als die angegebenen Quellen und Hilfsmittel benutzt sowie wörtliche und sinngemäße Zitate kenntlich gemacht. Die Arbeit hat in gleicher oder ähnlicher Form noch keiner Prüfungsbehörde vorgelegen.

Dortmund, 17.04.2015

Ort, Datum

Unterschrift

**Belehrung:**

Wer vorsätzlich gegen eine die Täuschung über Prüfungsleistungen betreffende Regelung einer Hochschulprüfungsordnung verstößt, handelt ordnungswidrig. Die Ordnungswidrigkeit kann mit einer Geldbuße von bis zu 50.000,00 € geahndet werden. Zuständige Verwaltungsbehörde für die Verfolgung und Ahndung von Ordnungswidrigkeiten ist der Kanzler/die Kanzlerin der Technischen Universität Dortmund. Im Falle eines mehrfachen oder sonstigen schwerwiegenden Täuschungsversuches kann der Prüfling zudem exmatrikuliert werden. (§ 63 Abs. 5 Hochschulgesetz - HG - )

Die Abgabe einer falschen Versicherung an Eides statt wird mit Freiheitsstrafe bis zu 3 Jahren oder mit Geldstrafe bestraft.

Die Technische Universität Dortmund wird gfls. elektronische Vergleichswerkzeuge (wie z.B. die Software „turnitin“) zur Überprüfung von Ordnungswidrigkeiten in Prüfungsverfahren nutzen.

Die oben stehende Belehrung habe ich zur Kenntnis genommen:

Dortmund, 17.04.2015

Ort, Datum

Unterschrift