

Corporate Performance Management

Big Data

Eine Revolution für die Business Intelligence?

Tilman Hagen, Senior BI Consultant

Marcel Weiß, BI Consultant

Braincourt GmbH

Inhaltsverzeichnis

1	Einleitung	3
2	Big Data	4
3	Big Data – eine Revolution für die Business Intelligence?	6
4	Synergien durch eine integrierte BI- und Big Data-Anwendungslandschaft	8
5	Zusammenfassung und Fazit.....	9
6	Ihre Ansprechpartner	10
	Literaturverzeichnis.....	10

1 Einleitung

Die Digitalisierung durchdringt immer mehr unsere Lebens- und Arbeitsbereiche. Sie stellt die Grundlage dar, um Informationen zugänglich zu machen und Menschen mit Menschen, Menschen mit Maschinen sowie Maschinen mit Maschinen zu verbinden. Hieraus resultiert ein explosionsartiges Ansteigen des weltweiten Datenaufkommens. Dass aus diesen Datenmengen wertvolle neue Erkenntnisse gewonnen und auch ein wirtschaftlicher Nutzen generiert werden kann, zeigen immer mehr Unternehmen und Projekte.

Spätestens seit der CeBIT 2012 sorgt ein Phänomen nicht nur in der Fachwelt für Aufsehen, sondern beschäftigt mehr und mehr die öffentliche Wahrnehmung: *Big Data*. Neben den in der Öffentlichkeit zu diesem Thema stattfindenden Diskussionen, z. B. bei Fragen des Datenschutzes, rückt *Big Data* zunehmend in das Bewusstsein von Entscheidungsträgern in Wirtschaftsunternehmen. Eine zunehmende Bedeutung von *Big Data* zeigen auch diverse Studien, wie beispielsweise der Experton *Big Data* Vendor Benchmark, der für den deutschen *Big Data*-Markt ein jährliches Wachstum von etwa 23 Prozent vorhersagt.¹

Big Data rückt in den Fokus von Unternehmen und Gesellschaft.

In den meisten Unternehmen werden bereits heute analytische Anwendungen genutzt, um eine Speicherung und Auswertung von Datenbeständen aus verschiedenen Systemen zu ermöglichen und daraus Erkenntnisse für die Unternehmenssteuerung zu gewinnen. Die damit einhergehenden Verfahren und Prozesse werden unter dem Begriff *Business Intelligence (BI)* zusammengefasst.

Der in den Unternehmen etablierten *BI* stehen nun die Konzepte von *Big Data* gegenüber. Sie wird dabei von Teilen der Fachpresse gar als „verstaubt und altmodisch“ dargestellt.

Handelt es sich bei *Big Data* um eine Revolution in der Verarbeitung und Auswertung von Daten und der damit verbundenen Wissensgewinnung oder stellt *Big Data* nicht vielmehr eine durch die Zeit hervorgebrachte Weiterentwicklung der etablierten *BI* dar?

Dieser Artikel soll zunächst das Phänomen *Big Data* erklären. Aufbauend darauf wird eine mögliche Antwort auf die genannte Fragestellung skizziert und Möglichkeiten beschrieben, wie sich die klassische *Business Intelligence* und *Big Data* kombinieren lassen.

¹ Vgl. Experton Group [Exper14].

2 Big Data

Das Schlagwort *Big Data* ist heute ein Synonym für die Handhabung exponentiell wachsender Datenmengen, die immer häufiger aktualisiert werden und aus unterschiedlichen Quellen stammen sowie unterschiedlicher Art sein können. Dadurch, dass der *Big Data*-Begriff von verschiedenen (Interessen-) Gruppen mit unterschiedlichen Vorstellungen, Erwartungen und insbesondere Versprechen versehen wurde, existiert bis heute kein einheitliches Begriffsverständnis.

Eine Begriffsdefinition bietet das Gartner IT Glossary:

„Big Data is high-volume, high-velocity and high-variety information assets that demand cost-effective, innovative forms of information processing for enhanced insight and decision making.“

Vor allem geprägt wurde der Begriff *Big Data* durch die Entstehung neuer Arten von Massen- und Echtzeitdaten. Der Großteil dieser Daten wird heute von Maschinen und IT-Systemen „produziert“. Hinzu kommen neue Anwendungen und Technologien wie Cloud Computing und RFID, sowie das „Internet der Dinge“. Einen wesentlichen Teil dieses Datenwachstums tragen auch das Internet, die zunehmende Verbreitung mobiler Endgeräte und soziale Medien wie Facebook, Twitter und Co. bei.² Wir Menschen tragen z. B. mit jedem Klick, jedem Beitrag auf Online-Portalen oder jedem Kauf in einem Online-Shop zu *Big Data* bei.

Die Entstehung von *Big Data* hat damit eine Vielzahl von Einflussfaktoren und Ursachen. Dementsprechend vielfältig sind auch die in den relevanten Systemen vorkommenden Datenarten. So beschreibt *Big Data* nicht nur alleine die riesigen Datenmengen, sondern auch die damit verbundene Datenvielfalt. Zu den bekannten strukturierten Daten, wie sie etwa in transaktionsverarbeitenden Systemen wie ERP-, CRM- oder SCM-Lösungen zu finden sind, kommen unstrukturierte Daten in Form von Bildern, Videos, Sprachaufnahmen, Dokumente oder E-Mails. Daher wird im *Big Data*-Umfeld auch von sogenannten polystrukturierten Daten gesprochen.

Verursacht durch die ersten Einsatzgebiete, ist *Big Data* heute stark durch Technologiebegriffe wie Hadoop oder In Memory-Computing geprägt. Diese ermöglichen es, die riesigen Datenmengen zu beherrschen und erlauben Nutzern solcher neuartigen *Big Data*-Systeme, Analysen in Echtzeit durchzuführen und die daraus gewonnenen Erkenntnisse direkt oder indirekt für die Entscheidungsunterstützung zu nutzen.

[Vielfältige Datenquellen](#)

[Big Data Daten sind polystrukturierte Daten.](#)

[Big Data und neue Technologien](#)

² Vgl. Bitkom [Bitk12].

Wie bereits die Gartner-Definition andeutet, wird Big Data vor allem durch drei Merkmale charakterisiert, die auch in der Fachwelt akzeptiert sind. Zusammengefasst werden diese Merkmale durch das sogenannte 3V-Modell.³ Dieses 3V-Modell beschreibt die Herausforderungen, die es im Rahmen von *Big Data* zu bewältigen gilt:

3V-Modell zur Charakterisierung von Big Data

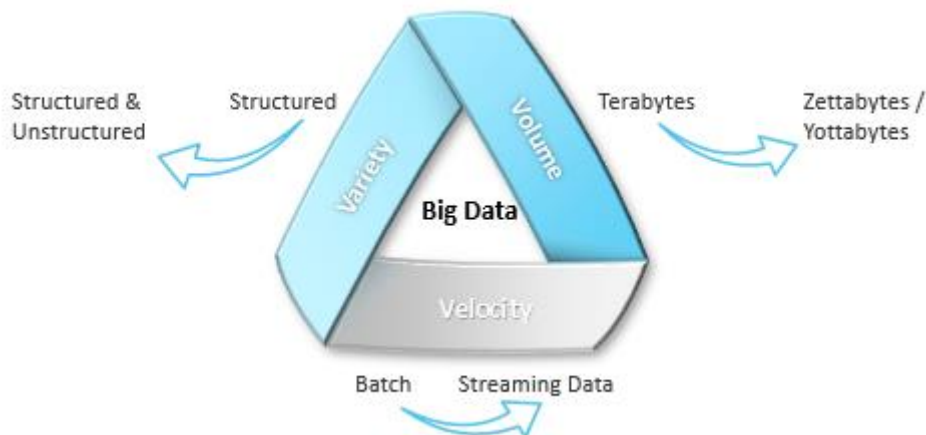


Abbildung 1: Charakteristika von *Big Data*⁴

- **Volume** bezieht sich auf die Herausforderung der Integration, Verarbeitung und Speicherung von riesigen Datenmengen (Datensätze, Dateien und Messdaten). Diese reichen von einigen Terabytes bis hin zu Petabytes.
- **Variety** beschreibt die Vielfalt der heute vorkommenden Datenquellen und Datenarten, mit denen man konfrontiert wird. So unterscheidet man etwa unstrukturierte, semistrukturierte und strukturierte Daten, welche als polystrukturiert zusammengefasst werden.
- **Velocity** bezieht sich zum einen auf die enorme Frequenz mit der Daten heute in den verschiedenen Anwendungsdomänen erzeugt werden. Zum anderen sind damit auch die Sicherstellung einer schnellen Reaktion sowie eine zeitnahe Weiterverarbeitung der Datenmassen, die im Idealfall in Echtzeit erfolgt, verbunden.⁵

Doch wie können konkrete und nützliche Szenarien für *Big Data* aussehen und welche Unternehmensbereiche eignen sich besonders für den Einsatz?

Durch den immer mehr an Bedeutung gewinnenden „Produktionsfaktor Information“ kommen grundsätzlich alle Unternehmensbereiche in Frage, in denen

„Produktionsfaktor Information“ in vielen Unternehmensbereichen relevant

³ Vgl. Gartner[Gart12]

⁴ Eigene Darstellung in Anlehnung an [ZiEaDe12], S. 5.

⁵ Vgl. [ZiEaDe12], S. 5. ff.

große Datenmengen anfallen oder beherrscht werden sollen. In einem Grundsatzpapier sieht der BITKOM in einer Vielzahl von Bereichen Potenziale für den Einsatz:

- Marketing und Vertrieb
Z. B. zur Kostenreduktion, zur Erhöhung des Umsatzes bei Kaufvorgängen, für die Prävention von Kundenabwanderungen oder für ein besseres Zielgruppenmarketing.
- Forschung- und Entwicklung
Z. B. zur Echtzeitauswertung wissenschaftlicher Experimente oder für die Produktentwicklung und -verbesserung auf Basis von Social Media-Trendanalysen.
- Produktion, Service und Support
Z. B. für die Produktionsoptimierung mit Maschinen- und Sensordaten oder die Früherkennung von Produktionsproblemen.
- Distribution und Logistik
Z. B. zur Optimierung der Lieferkette oder einer komplexen Verkehrsleumatik.
- Finanz- und Risikocontrolling
Z. B. für die Echtzeitreaktion auf Geschäftsinformationen, Simulationen und Vorhersagen, die Betrugserkennung oder das Risiko-Controlling.

3 Big Data – eine Revolution für die Business Intelligence?

Ein Versuch, die Frage zu beantworten, ob es sich bei *Big Data* tatsächlich um eine Revolution für die *Business Intelligence* handelt, macht ein allgemeingültiges Verständnis der *BI* erforderlich.

Auch hier bietet das Gartner IT Glossary einen Ansatzpunkt:

“Business Intelligence (BI) is an umbrella term that includes the applications, infrastructure and tools, and best practices that enable access to and analysis of information to improve and optimize decisions and performance.”

BI wird in der Regel als ein integrierter auf das Unternehmen bezogener IT-basierter Gesamtansatz verstanden, dessen Ziel es ist, Informationen für die Entscheidungsunterstützung und die Unternehmenssteuerung bereitzustellen. Um diese Aufgabe erfüllen zu können, benötigen BI-Anwendungen einen zentralen Datenspeicher konsistenter und strukturierter Daten, das sogenannte Data Warehouse. Ein Großteil dieser Daten stammt aus operativen Systemen wie ERP-, CRM- oder SCM-Anwendungen. Auf diese zentral gespeicherten und harmonisierten Daten können über definierte Schnittstellen direkt oder indirekt via Data

BI als ein IT-basierter
Gesamtansatz für die
Unternehmenssteuerung

Marts entsprechende Analyse- und Reporting-Anwendungen zugreifen. In vielen Fällen können die quantitativen Daten innerhalb des Data Warehouse um qualitative Kommentierungen ergänzt werden.

Aus den Eigenschaften von *BI* und von *Big Data* resultieren somit eine Reihe von Unterschieden. So sind *Big Data*-Lösungen beispielsweise darauf ausgelegt, auch mit unstrukturierten und möglicherweise nicht konsistenten Daten umgehen zu können. Zudem haben diese Lösungen oftmals eine heterogene, dezentrale Daten- und Datenverarbeitungsorganisation. Diese wird durch Technologien, wie das bereits genannte Hadoop ermöglicht. Zusätzlich schafft das In Memory-Computing eine Grundlage um in Echtzeit Daten zu analysieren und auszuwerten.

Eine weitere Differenzierung zwischen der klassischen *BI* und *Big Data* liegt in den Fragestellungen, die mithilfe der Anwendungen beantwortet werden sollen. So wurde etwa mit traditionellen BI-Lösungen hauptsächlich versucht Fragen zu beantworten, die sich auf historische Daten stützen. Also Fragen wie „Welche Produkte verkaufen sich in welchen Kundengruppen besonders gut?“ oder „Wie haben sich die Umsätze oder Absatzmengen für einen Geschäftsbereich, ein Produkt oder eine Leistung im Vergleich zum letzten Quartal entwickelt?“. Mit *Big Data* können Trends und Muster aufgedeckt werden, die bisher noch nicht erkannt wurden, wobei nicht nur Teildaten berücksichtigt und analysiert werden sondern möglichst komplette Datenbestände. *Big Data* schafft somit die Voraussetzung für einen Paradigmen-Wechsel im Controlling und in der Unternehmenssteuerung der dazu führt, dass im Rahmen von Analysen und Entscheidungsprozessen nicht so sehr die Betrachtung der Vergangenheit im Fokus steht, sondern zunehmend die Vorhersage zukünftiger Entwicklungen. Eine wesentliche Methode hierfür ist Predictive Analytics.

Zurückkommend auf die beiden Definitionen von Gartner, kann festgehalten werden, dass *BI* und *Big Data* sich nicht gegenseitig ausschließen oder gar konträr zu einander verhalten. Vielmehr werden Ansätze zur Erreichung desselben Ziels – die Informationsbereitstellung für die Entscheidungsunterstützung – aus zwei unterschiedlichen Blickwinkeln betrachtet: Während mit dem Begriff *Business Intelligence* insbesondere konkrete Anwendungen und BestPractices, wie z. B. Management Dashboards oder OLAP, beschrieben werden, beschreibt *Big Data* die besonderen Eigenschaften solcher Anwendungen.

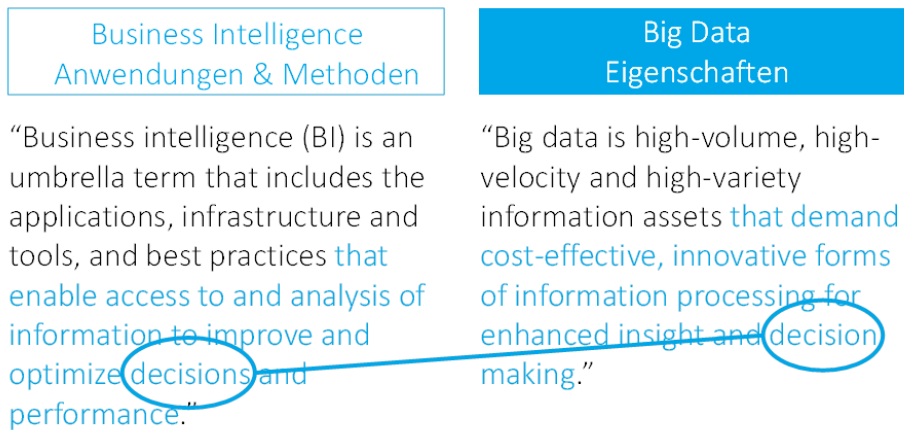


Abbildung 2: Definitionsgegenüberstellung von BI und Big Data

Auf Basis dieser Erkenntnis erscheint es so, dass es sich bei *Big Data* tatsächlich nicht um eine Revolution, sondern um eine Erweiterungsform der klassischen *BI* um die Eigenschaften der eingangs beschriebenen drei V's handelt. Voraussichtlich mündet die Entwicklung zukünftig in einer integrierten *BI*- und *Big Data*-Anwendungslandschaft.

Entscheidungsunterstützung steht im Fokus von *BI* und *Big Data*.

4 Synergien durch eine integrierte BI- und Big Data-Anwendungslandschaft

In einer integrierten *BI*- und *Big Data*-Anwendungslandschaft können *Big Data* und die zum Einsatz kommenden (neuen) Technologien verschiedene Aufgaben wahrnehmen.

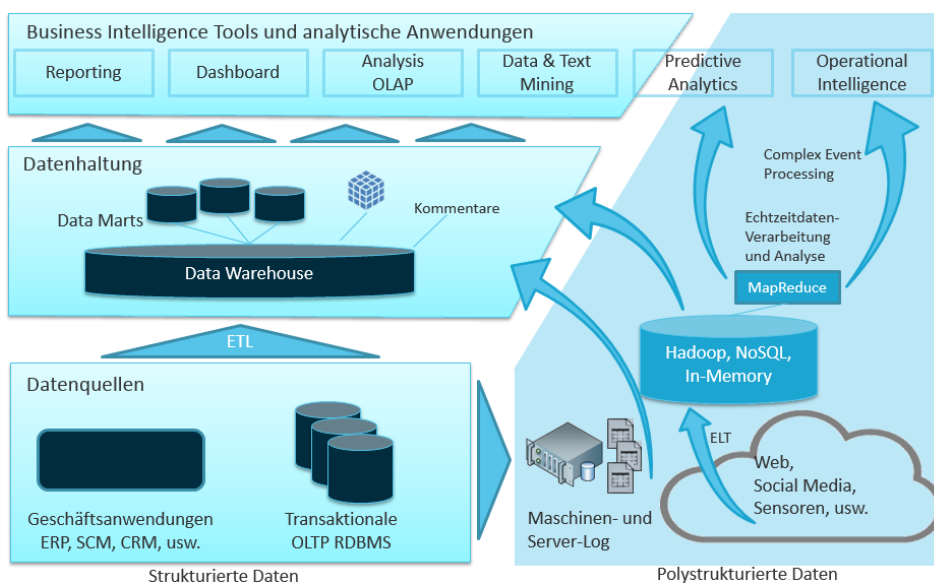


Abbildung 3: Integrierte BI- und *Big Data*-Anwendungslandschaft⁶

⁶ Eigene Darstellung in Anlehnung an BITKOM [Bitk12], S. 28.

Mögliche Synergien können dabei in unterschiedlichen Anwendungsszenarien entstehen. So können beispielsweise historische oder archivierbare Daten des Data Warehouse in ein kostengünstigeres Hadoop-Cluster (Beispiel) ausgelagert werden. Diese Daten können für Planungsprozesse oder für die Erfüllung rechtlicher Anforderungen zukünftig weiterhin genutzt werden. In einem alternativen Szenario können innerhalb eines solchen Clusters umfangreiche Analysen und Berechnungen, auf Basis großer polystrukturierter Datenmengen und mithilfe neuartiger Methoden (z. B. Text-Mining), durchgeführt und die (strukturierten) Analyseergebnisse in das Data Warehouse fortgeschrieben werden. Zudem können in einer derart integrierten Anwendungslandschaft alle Daten miteinander in Verbindung gebracht und so neue wertvolle Erkenntnisse gewonnen werden.

[Synergien in einer integrierten Anwendungslandschaft](#)

Durch neue Analysemöglichkeiten, wie Predictive Analytics oder Operational Intelligence, bietet *Big Data* darüber hinaus eine sinnvolle Erweiterung der Möglichkeiten der klassischen BI. Teilweise sind dies Konzepte, welche bereits in der Vergangenheit als Bestandteil der *BI* verstanden wurden (z. B. Operational BI), die jedoch in die Dimensionen von *Big Data* skaliert werden. So kann beispielsweise mithilfe diverser Analysewerkzeuge direkt auf die gesammelten und vielfältigen Daten zugegriffen und darin Muster und Zusammenhänge erkannt werden (Exploration).⁷ Die Muster oder sonstige aus Predictive Analytics gewonnenen Erkenntnisse können wiederum in eine typische *BI*-Anwendung z. B. zur Unterstützung einer Unternehmensplanung einfließen.

[Neue Potentiale für die BI durch Predictive Analytics und Operational Intelligence](#)

5 Zusammenfassung und Fazit

Das durch verschiedene Ursachen weltweit rasant steigende Datenaufkommen weckt das Bedürfnis, aus diesen Daten neuartige Erkenntnisse zu gewinnen und den Prozess der Erkenntnisgewinnung zu beschleunigen. Diese Entwicklung wird unter dem Begriff *Big Data* zusammengefasst. Die Beherrschung der riesigen Datenmengen macht den Einsatz neuartiger Technologien und Methoden erforderlich. Diesem Trend gegenüber steht die in den Unternehmen bereits seit vielen Jahren etablierte und bewährte *Business Intelligence*, welche ebenfalls (unternehmerische) Entscheidungen unterstützen soll.

Vor dem Hintergrund der gleichen Zielsetzung kann *Big Data* als eine Erweiterung der *Business Intelligence* um die durch die drei *V*'s beschriebenen Eigenschaften Volume, Variety und Velocity beschrieben werden. Somit kann die eingangs formulierte Fragestellung, ob es sich bei *Big Data* um eine Revolution für die *Business Intelligence* handelt, mit einem eindeutigen „Nein“ beantwortet werden.

⁷ Vgl. BITKOM [Bitk14], S. 101.

Beide Ansätze können in einer Big Data based BI-Architecture integriert und kombiniert werden. Durch die vorhandenen Unterschiede in den Wesensmerkmalen (Strukturiertheit vs. Polystrukturiertheit, Homogenität vs. Heterogenität, Aggregation vs. Detaillierung) lassen sich außergewöhnliche Synergien und voraussichtlich vollkommen neue Erkenntnisse gewinnen.

Für solche integrierten Ansätze existieren heute erst wenige konkrete Fallbeispiele und auch der tatsächliche, praktische Nutzen ist noch nicht sehr oft belegt. Doch, auch wenn die Anzahl an konkreten Fall- und Nutzenbeispielen noch gering ist: Business Intelligence und Big Data gehören zusammen und es ist mit einer hohen Wahrscheinlichkeit zu erwarten, dass sich in den Unternehmen in den kommenden Jahren eine Integration von BI und Big Data technisch, methodisch und organisatorisch vollziehen wird.

6 Ihre Ansprechpartner

Tilman Hagen

Senior BI Consultant
Braincourt GmbH
Fasanenweg 11
70771 Leinfelden-Echterdingen
Tilman.Hagen@braincourt.com
Telefon: + 49 711 75 85 80-54

Marcel Weiß

BI Consultant
Braincourt GmbH
Fasanenweg 11
70771 Leinfelden-Echterdingen
Marcel.Weiss@braincourt.com
Telefon: + 49 711 75 85 80-44

Literaturverzeichnis

Bezeichnung	Autor, Buchtitel, Verlagsangaben/URL
[Baja14]	Bange Carsten, Janoschek Nikolai (2014): Big Data Analytics – Auf dem Weg zur datengetriebenen Wirtschaft http://www.exasol.com/fileadmin/content-de/pdf/Studien/BARC_Big_Data_Analytics_2014.pdf
[Bitk12]	BITKOM (2012): Big Data im Praxiseinsatz – Szenarien, Beispiele, Effekte, http://www.bitkom.org/files/documents/BITKOM_LF_big_data_2012_online(1).pdf
[Bitk14]	BITKOM (2014): Big Data-Technologien – Wissen für Entscheider

	http://www.bitkom.org/files/documents/BITKOM Leitfaden Big-Data-Technologien-Wissen fuer Entscheider Febr 2014.pdf
[Exper14]	Van Laak, Kirsten (2014): Experton Group veröffentlicht <i>Big Data</i> Vendor Benchmark 2015 für Deutschland, http://www.experton-group.de/press/releases/pressrelease/article/experton-group-veroeffentlicht-big-data-vendor-benchmark-2015-fuer-deutschland.html
[Gart12]	Gartner (2012) Website von Gartner, IT-Glossary – <i>Big Data</i> https://www.gartner.com/it-glossary/big-data/
[MiChDh13]	Minelli, Michael, Chambers, Michele, Dhiraj, Abiga: <i>Big Data, Big Analytics – Emerging Business Intelligence and analytic trends for today's businesses,</i> Hoboken, New Jersey, John Wiley & Sons Inc., 2013
[ZiEaDe12]	Paul C. Zikopoulos, Chris, Eaton, Dirk Deroos, Tom Deutsch, George Lapis: <i>Understanding Big Data – Analytics for Enterprise Class Hadoop and Streaming Data,</i> New York: McGraw-Hill Companies, 2012, http://public.dhe.ibm.com/com-mon/ssi/ecm/en/iml14296usen/IML14296USEN.PDF