

e-COMMA

Datenspezialist

TM 02: Technische Grundlagen



Datenspezialist | BM 02: Technische Grundlagen

Lehrgangsinhalt

1. Datenspezialist: Einführung:
 - 1.1 Rollenbeschreibung
 - 1.2 Hauptaufgaben
 - 1.3 AARRR
2. Daten im Internethandel
 - 2.1 Überblick
 - 2.2 Big Data im Internethandel
3. Datenbanken
 - 3.1 Überblick
 - 3.2 Instrumente



Datenspezialist | BM 02: Technische Grundlagen

Quellen und Literaturempfehlungen:

S. Fan et al. 2015. Demystifying Big Data Analytics for Business Intelligence Through the Lens of Marketing Mix.

James Fahl, 2017, Data Analytics

Michael Devellano, 2017, Automate and Grow.

David A. Schweidel, 2017, Profiting from the Data Economy

Philip Kotler, 2017, Marketing 4.0: Moving from Traditional to Digital

<http://www.ibmbigdatahub.com/infographic/extracting-business-value-4-vs-big-data>

<http://www.business2community.com/big-data/big-data-and-analytics-value-chain-cross-section-0589031>

<https://www.linkedin.com/pulse/20140814193544-49814607-dynamic-pricing-the-future-of-ecommerce-in-india>



1 Datenspezialist: Einführung



1. Datenspezialist: Einführung

1.1 Rollenbeschreibung

Der Datenspezialist arbeitet mit Datensätzen, die die Organisationsentwicklung unterstützen. Die Daten werden von ihm gesammelt, aufbereitet, verglichen und geordnet.

Diese Fachkraft kann Datenbanken, die für programmierte Maßnahmen relevant sind, identifizieren und durchsuchen. Sie/Er verwaltet und pflegt die Unternehmensdatenbanken und selektiert, reduziert, interpretiert Daten und transformiert sie in relevante Informationen.



1. Datenspezialist: Einführung

1.2 Hauptaufgaben

1. Identifikation, Auswahl und Organisation offener Datenbanken für Entwicklungsprozesse im Marketing und Vertrieb

Datenspezialisten beginnen i.d.R. jedes Datenkonvertierungsverfahren mit einer gründlichen Analyse der Kundendaten. Sie recherchieren Datenquellen und korrigieren vorhandene Datensätze nach Bedarf, um die Genauigkeit der erfassten Informationen zu gewährleisten.



1. Datenspezialist: Einführung

1.2 Hauptaufgaben

2. Erstellung und Implementierung von Marketingroutinen anhand verfügbarer Daten

Datenspezialisten analysieren bestehende Systeme und wählen ein für ihr Unternehmen geeignetes Programm aus. Ggf. programmieren Datenspezialisten eine Datenbank oder ein Programm zur Datenkonvertierung selbst.



1. Datenspezialist: Einführung

1.2 Hauptaufgaben

- 3. Beschreibung der Kapitalrendite (ROI) bei Marketing- und Vertriebsmaßnahmen**
- 4. Definition und Verwaltung von Analytics-Verfahren**
- 5. Dokumentation und Berichterstattung**

Datenspezialisten müssen Kunden konsistent über den Fortschritt eines Konvertierungsprogramms informieren. Sie müssen Berichte über Workflows und Workflow-Störungen, Ausnahmen, Kosten und Analyseergebnisse vorlegen.



1. Datenspezialist: Einführung

1.2 Hauptaufgaben

6. Beitrag zu strategischen Analysen der Organisation (E-Marketing-Strategie, Community-Management-Strategie etc.)

7. Technische Unterstützung bieten

Daten-Spezialisten pflegen zudem Datenbanken und beantworten alle Fragen, die Anwender zum System haben.

Schlüsselwörter für Datenspezialisten lauten: AARRR – Acquisition, Activation, Retention, Reference, Revenues



1. Datenspezialist: Einführung

1.3 AARRR

Acquisition (Akquise): Wie viele Nutzer besuchen Ihre Seite? Woher kommen sie? Wer sind sie? Wie viele Seitenansichten? wie viele Klicks? Bounce-Rate?

Activation (Aktivierung): Welchen ersten Eindruck hat der Nutzer? Was betrachtet er? Was macht er?

Retention (Bindung): Kommt der Nutzer zurück?

Reference (Empfehlung): Teilt der Nutzer seine Erlebnisse?

Revenues (Umsatz): Wie viel Umsatz wird erwirtschaftet?



2 Daten im Internethandel



2. Daten im Internethandel

2.1 Überblick: Relevante Indexe

Die meisten Internethändler nutzen heutzutage die folgenden Indexe:

- **NC (new customer / Neue Kunden);**
- **RC (retained customer / gebundene Kunden);**
- **ROI (return on investment / Kapitalrendite);**
- **CLV (Customer-Lifetime-Value);**
- **ROI CLV;**
- **RR (return rate / Wiederkehrrate);**
- **CR (conversion rate / Konversionsrate);**
- **CPO (cost per order / Kosten pro Bestellung);**
- **CPNC (cost per new customer / Kosten pro Neukunde);**
- **CPRC (cost per retained customer / Kosten pro Bestandskunde);**



2. Daten im Internethandel

2.1 Überblick: Datenquellen

Internethandel:

Bestellungen

Produkte

Warenkörbe

Besuche

Nutzer

Marketing-Kampagnen

verweisende Links

Suchbegriffe

durchsuchbare Produktkataloge

Soziale Daten:

Facebook

Twitter

Google

Daten zum Traffic:

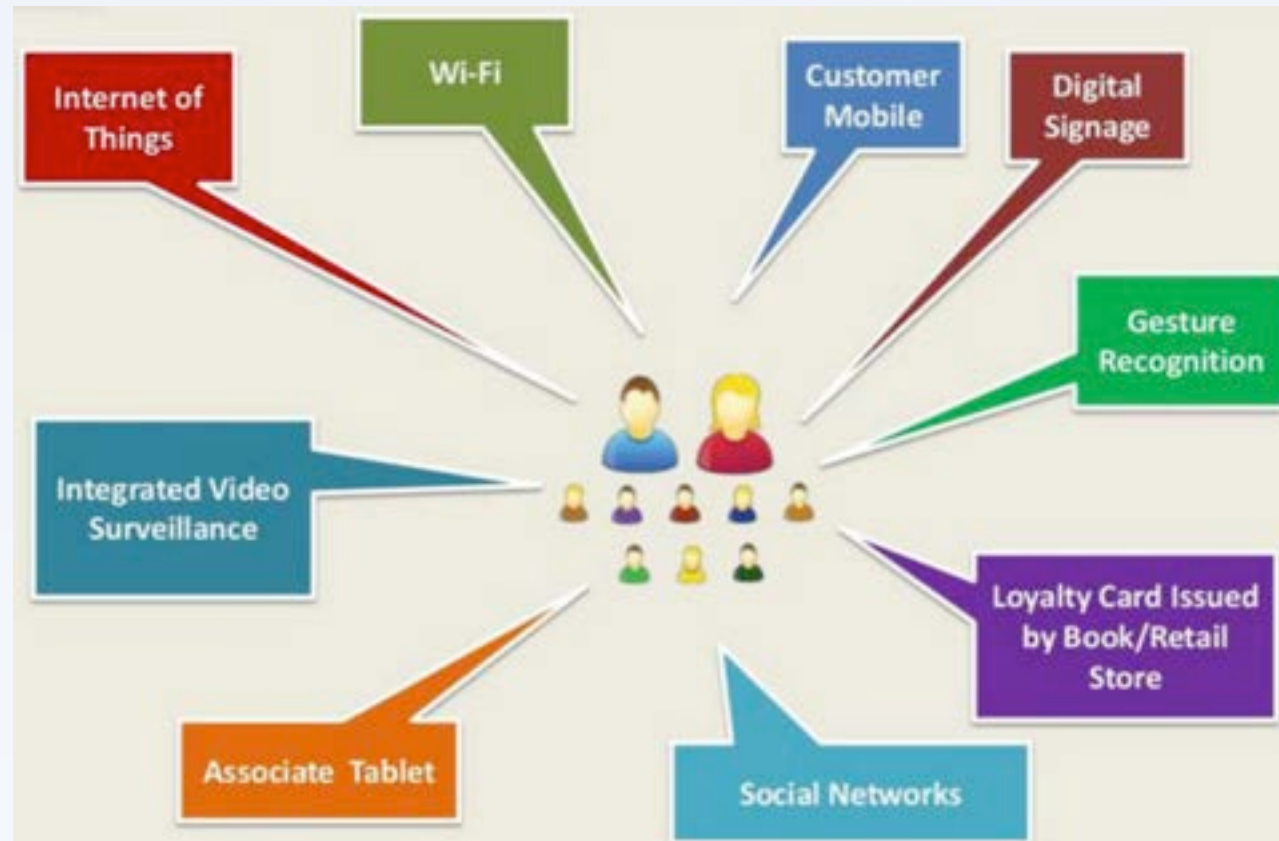
Cookies

Google Analytics



2. Daten im Internethandel

2.1 Überblick: Datenquellen



2. Daten im Internethandel

2.1 Überblick: Analytische Fragen

- Welche Bestseller gibt es pro Kategorie?
- Ist das am häufigsten aufgerufene Produkt auch das Bestverkaufte?
- Welche Produkte verkaufen sich am besten unter den Benutzern, die bereits einen Artikel in der jeweiligen Produktkategorie gekauft haben?
- Wie oft kehrt eine bestimmte Benutzergruppe (z.B. neue Benutzer) in Ihren Shop zurück?
- ...

Das Problem ist leider, dass die Beantwortung dieser Fragen nicht direkt mehr Profit bedeutet. Unternehmen lassen sich oft entmutigen, da sich die Antworten nur schwer in die Praxis übertragen lassen.



2. Daten im Internethandel

2.1 Überblick: Erforderliche Daten

- kollaboratives Filtern
- mit Informationen zu Nutzeraktivitäten werden automatisch die Zusammenhänge zwischen Elementen auf der Internetseite, einem Suchbegriff und dem gewählten Link hergestellt.
- Empfehlungen
Produkte
Angebote
- Klassifizierung
Nutzer, die weitereinkaufen

- Regression
Hinweise auf oder das Fehlen von Trends
Vorhersage von Beständen
die zukünftige Beliebtheit eines Produktes antizipieren
den zukünftigen Anklang von Werbemaßnahmen antizipieren
- Beurteilung der Auswirkungen von Marketingaktivitäten auf den Umsatz oder die Anzahl Nutzer
- Kategorisierung und Segmentierung
Kunden
Produkte



2. Daten im Internethandel

2.1 Überblick: Erforderliche Daten - Beispiele

Können wir dank *Big Data* die Korrelation zwischen den sozialen Medien und unseren Systemdaten finden, berücksichtigen wir dabei, dass:

- 40% der Nutzer ein Produkt erworben haben, nachdem Sie es in sozialen Medien gelikt oder geteilt haben;
- 71% der Nutzer in sozialen Medien primär auf Empfehlung kaufen;
- wir für bestimmte Kunden Einkaufsempfehlungen, basierend auf deren Social-Media-Verhalten, ausarbeiten können.



2. Daten im Internethandel

2.1 Überblick: Erforderliche Daten - Beispiele

Beispiel: T-Mobile

- Abrechnungen, Social-Media-Daten;
- Auswahl von Kunden für die Migration auf Premium-Modelle;
- Identifikation von Kunden mit hohem Lifetime-Customer-Value.



2. Daten im Internethandel

2.1 Überblick: Erforderliche Daten - Beispiele

Beispiel: STARBUCKS

- Daten zu Kundenbestellungen sammeln;
- Personalisierte Werbeanzeigen;
- Personalisierte Gutscheine;
- Kunden auswählen, die das Interesse am Angebot verlieren;
- ehemalige Kunden erneut anwerben.



2. Daten im Internethandel

2.1 Überblick: Erforderliche Daten - Beispiele

Beispiel: EasySize

Analyse von Bestellungen und Retouren - anhand der Ergebnisse wird entschieden, welche Marken in welchen Größen zu einer bestimmten Person passen.



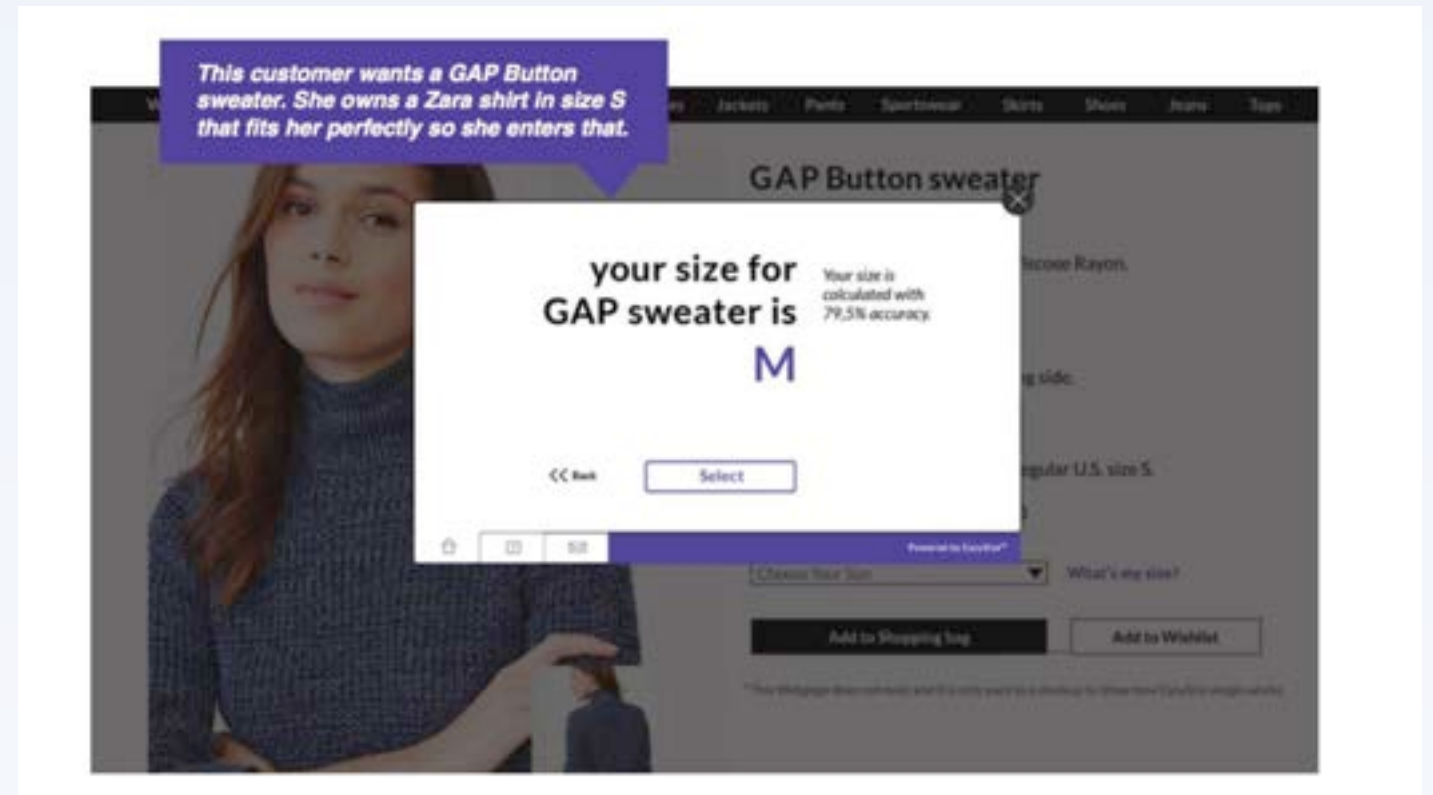
The screenshot shows a product page for a "GAP Button sweater". A purple callout box at the top left states: "This customer wants a GAP Button sweater. She owns a Zara shirt in size S that fits her perfectly so she enters that." A white modal window titled "Let's find your size" is overlaid on the page. It contains three dropdown menus: "I have a" with "shirt" selected, "From" with "Zara" selected, and "In size" with "S" selected. Below the dropdowns is a "Convert" button and the text "Approve selection.". At the bottom of the modal, it says "Powered by EasySize™". The background shows the product image of a woman wearing a blue sweater, a navigation bar with categories like "Jackets", "Pants", "Sportswear", "Skirts", "Shoes", "Jeans", and "Tops", and buttons for "Add to Shopping bag" and "Add to Wishlist". A small disclaimer at the bottom of the page reads: "This Webpage does not exist and it is only used as a mockup to show how EasySize plugin works".

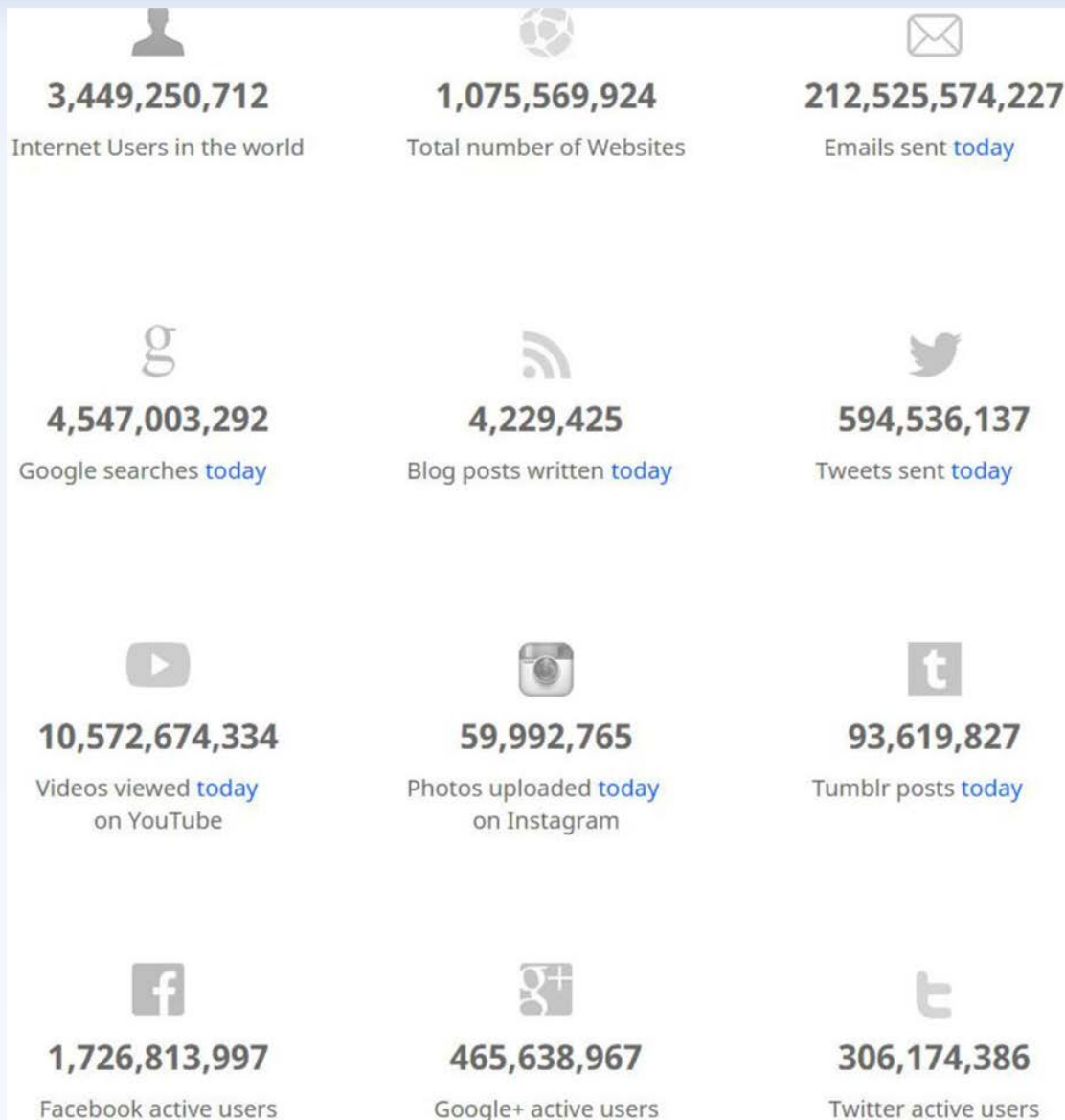
2. Daten im Internethandel

2.1 Überblick: Erforderliche Daten - Beispiele

Beispiel: EasySize

Ergebnisse: 35-40% weniger
Retouren





2. Daten im Internethandel

2.2 Big Data im Internethandel

Big Data

Beschreibt, wie viele Daten Teil unseres Lebens sind, ausgelöst durch beschleunigte technologische Fortschritte.



2. Daten im Internethandel

2.2 Big Data im Internethandel

OPTIMIZED

Vorteile von Big Data im Internethandel

- Produktportfolio;
- Preisgestaltung;
- Kundenerfahrung online und im Geschäft;
- Werbung/Marketing;
- Budget für Kundenservice;
- Bestand.



2. Daten im Internethandel

2.2 Big Data im Internethandel

Vorteile von Big Data im Internethandel

Gezielte Echtzeit-Werbeaktionen, die direkt per Smartphone übertragen werden, während die Kunden einkaufen, bspw. durch Analyse der Kaufhistorie, oder basierend auf dem Online-Verhalten bspw. in sozialen Netzwerken, Geo-Location, etc.



Unterstützt durch Daten aus Online-Quellen können Einzelhändler nun festlegen, welche Waren an bestimmten Orten gelagert und wo diese im Laden platziert werden sollen (z.B.: eine schwangere Frau, die Babyprodukte am Eingang eines Ladens sieht).

Durch die Anpassung der Angebote an jeden einzelnen Kunden steigt die Zahl der wiederkehrenden Kunden. Kunden suchen heutzutage nach dem einfachsten und bequemsten Weg zum Einkaufen. Big Data ermöglicht es Händlern, Kundenbedürfnisse zu verstehen, bevor sie den Laden betreten.

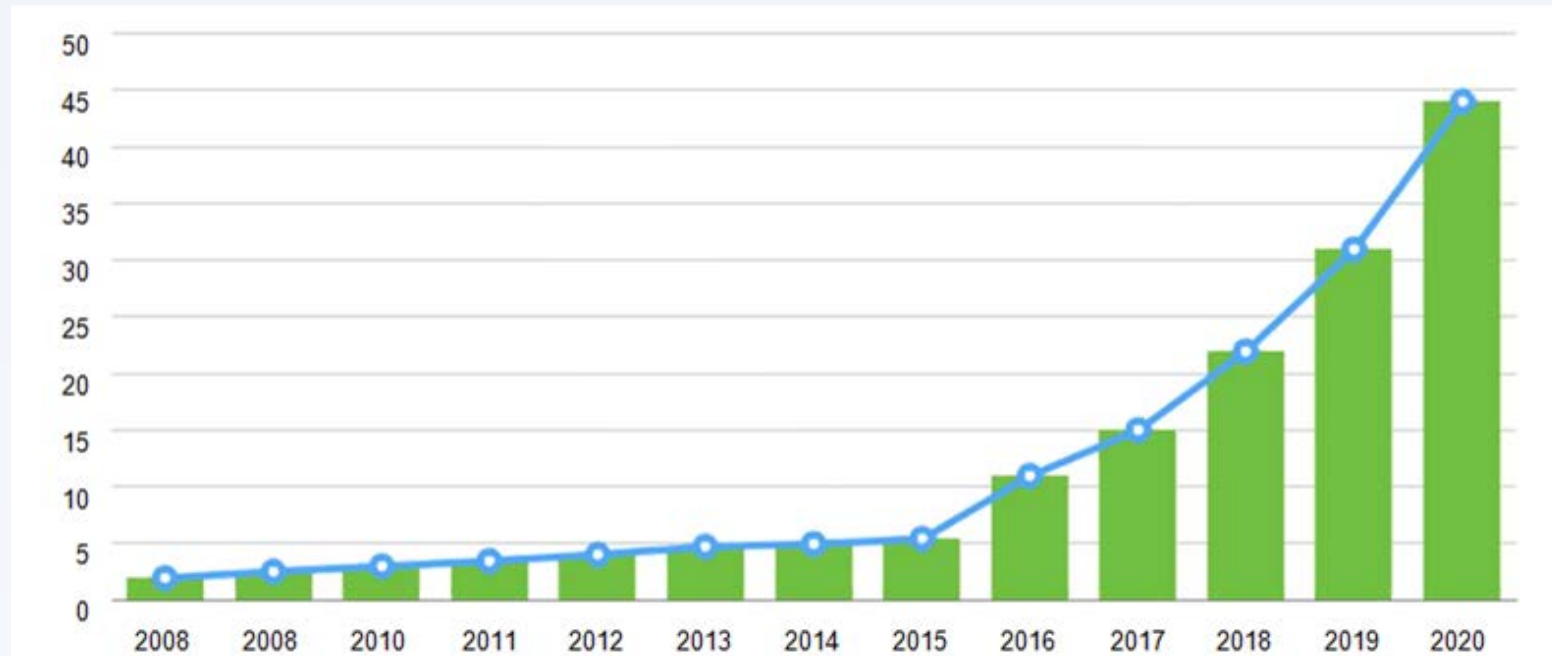


2. Daten im Internethandel

2.2 Big Data im Internethandel

BIG DATA WÄCHST UM 40% JÄHRLICH

Die Datenmengen wachsen mit einer durchschnittlichen jährlichen Rate von 40% und werden bis 2020 fast 45 ZB umfassen.



2. Daten im Internethandel

2.2 Big Data im Internethandel

DAMALS...

Vertrieb



HEUTE...

datengesteuerte Preisfindung
und Empfehlungen



2. Daten im Internethandel

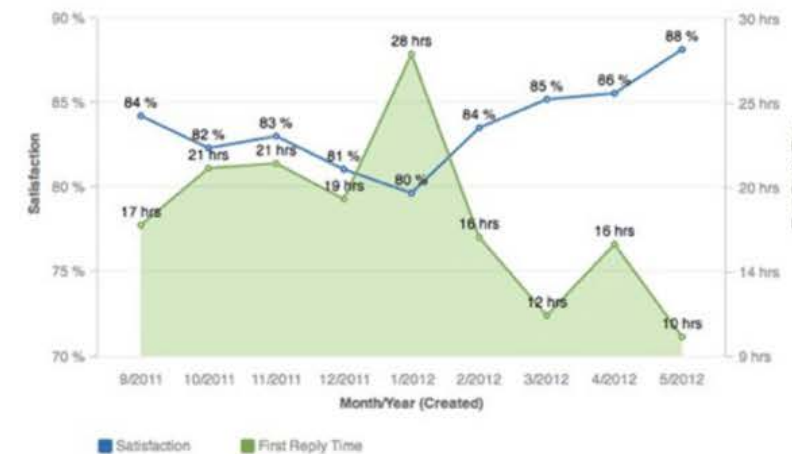
2.2 Big Data im Internethandel

VERWANDLUNG VON KUNDENDIENSTEN

DAMALS...
Unzufriedene Kunden



HEUTE...
Kundenerkenntnisse



3 Datenbanken



3. Datenbanken

3.1 Überblick

Datenbanken sind Hilfsmittel zur Verwaltung aller Datenarten, die von einem Unternehmen erzeugt werden. Im Bereich Internethandel und Online-Marketing sind die Daten von Online-Nutzern und -Kunden am wichtigsten.

Die **Datensätze** werden gesammelt und gespeichert. Die Daten beziehen sich auf die Herkunft des Nutzers, persönliche Daten, was er angesehen und angeklickt hat, Kaufdetails, Verweildauer auf der Website etc.

Die Herausforderung liegt in der Verwaltung dieser Datenmengen. Hilfreiche Tools sind **Datenbank-Management-Systeme (DBMS)**. Diese Softwareanwendungen sind Schnittstellen für Datenspezialisten. Sie helfen bei der Aufarbeitung nützlicher Informationen aus der Datenmenge.



3. Datenbanken

3.1 Überblick: DBMS

Datenbank-Management-Systeme (DBMS) helfen bei der Verwaltung aller möglichen Daten. Sie bieten die Möglichkeit, lediglich bestimmte Daten anzuzeigen und verschiedene Datensätze miteinander zu verknüpfen.

DBMS nutzen verschiedene Programmierungsstandards wie **SQL** oder **ODBC**.

Für weitere Informationen zu diesen Programmiersprachen siehe:

<https://www.youtube.com/watch?v=7Vtl2WggqOg>

<https://www.youtube.com/watch?v=nWeW3sCmD2k>

<https://www.youtube.com/watch?v=VkMXJvaWeTE>



3. Datenbanken

3.2 Instrumente

Hadoop

Die Verarbeitungssoftware für verteilte Daten (Apache Software Foundation) ist so weit verbreitet, dass die Begriffe "Hadoop" und "Big Data" manchmal synonym verwendet werden.

Hadoop ist bekannt für die Fähigkeit, extrem große Datenmengen in strukturierten und unstrukturierten Formaten zu verarbeiten und verlässlich Datenblöcke auf Knotenpunkten im Cluster zu replizieren, um die Daten auf diese Weise lokal zur Weiterverarbeitung zur Verfügung zu stellen

Die Apache Software Foundation unterstützt zudem eine Reihe verwandter Projekte, was die Möglichkeiten von Hadoop erweitert.



3. Datenbanken

3.2 Instrumente

MapReduce

Stellt man sich Hadoop als den ‚Elefantentreiber‘ der Big Data vor, dann wäre MapReduce dessen Rettungsleine.

Ein Programmiermodell und Software-Rahmenwerk für Textverarbeitungsanwendungen. MapReduce sorgt für eine schnelle Verarbeitung großer Datenmengen parallel auf großen Rechenknoten-Clustern.

Kommt oft bei Hadoop und vielen anderen Datenverarbeitungsanwendungen zum Einsatz.



3. Datenbanken

3.2 Instrumente

GridGain

GridGain ist eine Java-basierte Middleware für schnellere In-Memory-Verarbeitung von Big Data in Echtzeit.

GridGain ist mit dem Hadoop-Datenverarbeitungssystem kompatibel.

Erfordert das Windows-, Linux- oder Mac OSX-Betriebssystem.



3. Datenbanken

3.2 Instrumente

HPCC Systems

HPCC wurde von LexisNexis Risk Solutions entwickelt und steht für "High Performance Computing Cluster".

HPCC Systems liefert auf einer einzigen Plattform eine Architektur und eine Programmiersprache zur Datenverarbeitung.

Es stehen sowohl kostenlose Community-Versionen als auch kostenpflichtige Enterprise-Versionen zur Verfügung.



3. Datenbanken

3.2 Instrumente

Storm

Storm unterscheidet sich von anderen Tools durch sein verteiltes, echtzeitfähiges, fehlertolerantes Verarbeitungssystem und steht somit im Gegensatz zu den Batch-Verarbeitungssystemen von Hadoop.

Es ist schnell und besonders gut skalierbar und wird oft als „Hadoop der Echtzeit“ bezeichnet.

Fehlertolerant und kompatibel zu fast allen Programmiersprachen, wobei typischerweise Java verwendet wird.



3. Datenbanken

3.2 Instrumente

Cassandra

Cassandra ist eine hoch skalierbare, verteilte NoSQL-Datenbank für große, und über mehrere Rechenzentren und die Cloud hinweg verteilte Datenmengen.

Wird von vielen Organisationen mit großen, aktiven Datensätzen verwendet (darunter Netflix, Twitter, Urban Airship, Constant Contact, Reddit, Cisco und Digg).

Der kommerzielle Support und die Bereitstellung von Dienstleistungen erfolgen über Drittanbieter.



HBase

HBase ist der nicht-relationale Datenspeicher für Hadoop.

Als spaltenorientiertes Datenbankmanagementsystem eignet sich HBase gut für verstreute Datensätze und wurde in Java programmiert. Es unterstützt Textverarbeitungen wie Avro, REST und Thrift.

Der Funktionsumfang umfasst:

- lineare und modulare Skalierbarkeit;
- strikt-konsistentes Auslesen und Schreiben von Daten;
- automatische Failover-Unterstützung uvm.



3. Datenbanken

3.2 Instrumente

MongoDB

MongoDB wurde ursprünglich von 10gen entwickelt, um riesige Datenbanken zu unterstützen.

Es ist eine in C++ geschriebene NoSQL-Datenbank mit dokumentenorientierter Speicherung, voller Indexunterstützung, Vervielfältigungsmöglichkeiten sowie hoher Verfügbarkeit. Es skaliert horizontal ohne Kompromisse bei der Funktionalität.

Kommerzieller Support über 10gen.



3. Datenbanken

3.2 Instrumente

Neo4j

Neo4j bietet im Vergleich zu relationalen Datenbanken eine bis zu 1000-mal höhere Leistung.

Speichert Daten strukturiert in Graphen anstelle von Tabellen und ist ein festplatten-basierter, vollständig transaktionaler Java-Engine.

Unternehmen können erweiterte und Enterprise-Versionen über Neo Technology beziehen.



3. Datenbanken

3.2 Instrumente

CouchDB

CouchDB speichert Daten in JSON-Dokumenten, auf die über das Web oder über JavaScript zugegriffen werden kann.

Bietet verteilte Skalierung mit fehlertoleranter Speicherung

Schlüsselmerkmale:

- ‚On-the-fly‘-Dokumentenumwandlung;
- Echtzeit-Änderungsbenachrichtigungen;
- Nutzerfreundliche Online-Verwaltungskonsole.

