



[Übungsblätter](#)

Definition:

Unter dem Begriff "**Big Data**" (Massendaten) bezeichnet man Datenmengen, welche zu komplex, quantitativ zu groß oder zu schwach strukturiert sind, um sie mit herkömmlichen Methoden (manuell, Software) zu analysieren.

Ermöglicht wird diese flächenmäßige Sammlung von Daten durch die Digitalisierung und das Internet, welche in ihrer Kombination riesige Datenmengen generieren.

Bei Personen betrifft das z.B. Einkaufsverhalten im Online-Shop.

Deutlich mehr Daten fallen aber im Bereich der Steuerung von Maschinen an (Sensordaten).

Um diese Daten sinnvoll zu verwerten, verwendet man Algorithmen.

Merkmale:

- volume (Volumen der Daten)
- velocity (Geschwindigkeit der Generierung und des Transfers von Daten)
- variety (Bandbreite der Daten)
- value (Wert der Datenmengen)
- validity (Qualität der Datenmengen)

Entstehung der Daten:

Durch die steigende Digitalisierung in allen Lebensbereichen nimmt die Menge an Daten ständig zu, und verdoppelt sich alle zwei Jahre.

Die gesamte Datenmenge der Menschheit, die sich bis zum Jahr 2000 angesammelt hat, wird jetzt schon in wenigen Minuten erreicht.

In folgenden Bereichen fallen z.B. riesige Datenmengen an:

- Cloud Computing (externe Datenspeicherung)
- Media Streaming (Netflix)
- Überwachungssysteme (öffentliche Plätze)
- mobiles Internet (Smartphones)
- Fitness- und Gesundheit (Armbänder, Uhren)
- Navigationssysteme (GPS)
- Smart Homes (vernetzte Technik in Häusern)
- Autonomes Fahren (Roboterautos)

Anwendung:

In einer Welt, die immer mehr vernetzt und komplexer hinsichtlich ihrer Datenstruktur wird, ergibt sich für Big Data ein breit gefächertes Anwendungsgebiet:

- Bonitätsprüfung (Kreditscoring)
- Früherkennung von Krankheiten (medizinische Diagnostik)
- Optimierung der Ernte (Smart Farming)
- Überprüfung von Finanztransaktionen (Fraud-Detection)

- verbesserte Wettervorhersage (optimierte Berechnungsmodelle mit großen Datenvolumen)
- Wissensabfragen (riesige Datenmengen aus allen Fachbereichen)
- optimierte Produktionsabläufe (Überwachung von Maschinen) etc.

Gefahren:

a) Datenschutz:

Der fehlende Datenschutz für Privatpersonen ist sicherlich die größte Gefahr bei einer großflächigen Anwendung von Big-Data Modellen.

Sind die Daten erst einmal gesammelt, so droht früher oder später die wirtschaftliche Nutzung.

Und hier sind die Grenzen fließend, Anwendung von Daten, um eine Wahl zu beeinflussen sind nur ein aktuelles Beispiel (Facebook 2018).

b) Entsolidarisierung:

Ein weit unterschätztes Gefahrenpotential der Big Data Anwendung stellt die Entsolidarisierung innerhalb einer Gemeinschaft dar - ein wichtiger Eckpfeiler der sozialen Marktwirtschaft.

Wenn ein Versicherungskonzern in der Lage ist ein maßgeschneidertes Angebot für jeden seiner Kunden zu machen, heißt das im Umkehrschluss auch, dass bestimmten Bevölkerungsteilen (z.B. Personen mit Erbkrankheiten) dies zukünftig aufgrund der höheren Risiken auch verweigert werden kann.

c) Fehlende Normen:

Sowohl die Datenerfassung als auch die Datenverwendung erfolgt aufgrund rein technischer Parameter, unabhängig von der komplexen Lebenssituation einer untersuchten Person.

Sie weist zudem der Vergangenheit einen höheren Stellenwert zu als der Zukunft.

Begangenes Fehlverhalten in der Vergangenheit wird dazu herangezogen zukünftiges Verhalten zu prognostizieren.

Der jeweilige Staat als oberster Normenhüter schützt derzeit sein Bürger nicht vor einem Missbrauch ihrer Daten, sondern beteiligt sich selbst als eifriger Datensammler.

d) Datenbesitz:

Wenige große weltweit operierende Konzerne wie Google, Amazon, und vor allem Facebook kontrollieren derzeit die Big Data Welt.

Wie sie mit diesen Daten umgehen, unterliegt keiner Kontrolle.

Ob sie mit den jeweiligen Regierungen sogar zusammenarbeiten, die sie eigentlich kontrollieren sollten, entzieht sich der öffentlichen Wahrnehmung.

PDF-Übungsblätter

[Big Data Gefahren Übungsblatt](#)

[Big Data Merkblatt](#)

[Big Data Übungsblatt 1](#)

[Big Data Übungsblatt 2](#)