

L5_2 Big Data – Information

Was ist Big Data?

Für den Begriff Big Data fehlt bisher eine allgemeingültige und eindeutige Definition. Die Festlegung, ab wann von Big Data gesprochen werden kann, stützt sich meist auf das sogenannte 3V-Modell:

- ➔ Volume
- ➔ Variety
- ➔ Velocity

Volume beschreibt die enorme *Datenmenge*, die täglich produziert wird. 2018 wurden Daten in der Größenordnung von mehreren Trillionen Bytes (10^{18}) erzeugt.

Würden alle diese Daten auf DVDs gespeichert und nebeneinandergelegt, ergibt sich eine Strecke, die nahezu der Länge des Äquators entspricht.

Variety bezieht sich auf die *Vielfalt* der Datenquellen und Datentypen. Neben den Daten aus dem Internet zählen auch Daten, die im öffentlichen Raum gesammelt werden (z.B. Navigationsgeräte, Videoüberwachung, Satelliten etc.).

Velocity bezeichnet die *Geschwindigkeit*, mit der Daten generiert, ausgewertet und weiterverarbeitet werden können. Heutzutage erfolgt dies meist im Bruchteil von Sekunden bzw. in Echtzeit.

Big Data stellt Instrumentarien und Methoden zur Verfügung, mit denen große, schnell anwachsende Datenmengen mit komplexer Datenstruktur erfasst, gespeichert, durchsucht, analysiert und visualisiert werden können.

Die Grundlage für Big Data sind die Möglichkeiten, immer mehr Daten kostengünstiger auf immer kleineren Speichermedien zu sichern. Wie das Beispiel der personalisierten Werbung auf Google und Facebook zeigt, können diese Daten zu verschiedenen Zwecken effektiv genutzt werden. So dienen diese Informationen auch zur gezielten Einflussnahme bei politischen Wahlen, indem politische Botschaften auf kleine Zielgruppen passgenau zugeschnitten werden.

Nicht nur die großen Internetgiganten verfügen über große Datenmengen und die Möglichkeiten diese auszuwerten. Im Gesundheitswesen wird die gigantische Informationsflut von Arztbriefen, elektronischen Patientenakten, Operationsberichten, Pathologiebefunden etc. digital erfasst und analysiert, um daraus medizinische Diagnosen und Therapien für Patienten abzuleiten.

Auch in den Betrieben wird Big Data ein hoher Stellenwert zugesprochen. So lassen sich mit der Sammlung und Auswertung von Kundendaten, Kaufverhalten, Mitarbeiterdaten oder Daten entlang von Produktionsketten bei der Fertigung von Produkten Geschäftsprozesse optimieren, strategische Entscheidungen ableiten und beschleunigen oder Kosten reduzieren.