

BIG DATA

Definition, Merkmale und Technologien

Als Big Data werden Daten bezeichnet, die sich in den Merkmalen Volumen, Geschwindigkeit, Datentypen, Qualität, Wert oder Variabilität von klassischen, strukturierten Datensätzen unterscheiden. Diese

Daten benötigen sowohl in Aufnahme, Speicherung als auch Analyse neue, spezialisierte Systeme und Methoden, ermöglichen aber sogleich genauere und umfangreichere Analysen und Vorhersagen als zuvor.

MERKMALE (6Vs)

VOLUME



Sehr große Anzahl an Einträgen und / oder Attributen.

VELOCITY



Sehr schnell generierte oder sich ändernde Daten.

VARIETY



Unstrukturierte Daten wie Bilder und Text.

VERACITY



Qualität des Inhalts und der Herkunft der Daten.

VALUE



Einsatzzweck und Wert der vorhandenen Daten.

VARIABILITY



Fluktuationen und hohe Varianz wie Saisonalität.

www.datadrivencompany.de

TECHNOLOGIEN (Beispiele)

Akquise & Extraktion



Talend



Kafka



Azure Data Factory

Speicherung & Bereitstellung



MongoDB



Hadoop



AWS S3

Verarbeitung & Analyse



python



KNIME



Google AutoML

Deployment & Visualisierung



Docker



Kibana



Amazon API Gateway

(und viele, viele Weitere)