twitch

Tweets
Internet 2.0

Broadband
Blogs
Games
Smartphones

YouTube

Merely 1% of available data is processed. [8]

190.000 jobs vacant for Big Data specialists in the US. [7]

"… 5 exabytes of information created between the dawn of civilization through 2003 now created every 2 days…" – Eric Schmid (CEO, Google)

NATIONAL SECURITY AGENCY · UNITED STATES OF AMERICA ·

The Digital Universe [1]

4,4 zB 2014

44 Zettabytes 2020

⚠ !Warning! ⚠
Data volume reaching critical mass!

„Big Data is an abstract term for any type and amount of data which can't be processed with conventional ways of data storaging and thus requires new and better technologies" [5]
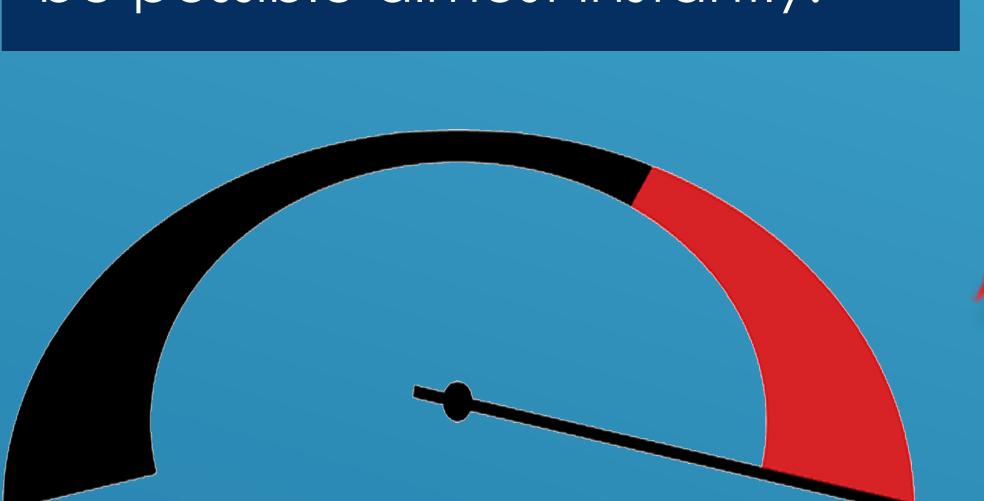
hadoop
Hadoop is a framework written in Java, capable of controlling distributed processes on huge clusters.
It can be used to process Big Data. [6]

BIG DATA BANG

Velocity [2]
Data needs to be processed almost instantly. Just like the access to stored data should be possible almost instantly.
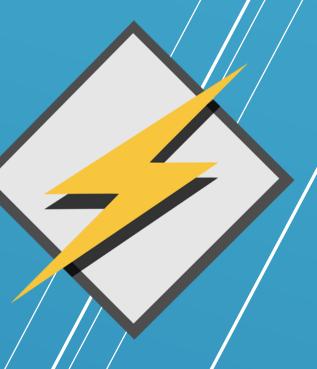
Variety [2]
Vastly different types of unstructured data like audio, text, visuals need to be processed by a Big Data System.

Volume [2]
Prior to the Big Data Bang unthinkable amounts of data need storage and processing.

Database Database Database Database

Big Data Players [4]

Google
15.000 pB stored.

facebook
300 pB stored.

ebaY
100 pB processed/day.

Big Data Pipeline [3]

Aquisition, Recording ▶ Extraction, Cleaning, Annotation ▶ Integration, Aggregation, Annotation ▶ Analysis, Modeling ▶ Interpretation

Big Data Chances
Scientific and medical progress.
Economical growth.
Homeland defense.

Big Data Dangers
Total surveillance.
Automatic weaponsystems.
Glass human being.

Sources
[1] www.emc.com/leadership/digital-universe/2014iview/executive-summary.htm
[2] www.gartner.com/newsroom/id/1731916
[3] Divyakant Agrawal et al.: Challenges and Opportunities with Big Data, 2012
[4] followthedata.wordpress.com/2014/06/24/data-size-estimates/
[5] www.gi.de/service/informatiklexikon/detailansicht/article/big-data.html
[6] hadoop.apache.org/
[7] www.nytimes.com/2012/02/12/sunday-review/big-datas-impact-in-the-world.html
[8] John Gantz, David Reinzel: THE DIGITAL UNIVERSE IN 2020: Big Data, Bigger Digital Shadows, and Biggest Growth in the Far East, IDC Iview, 2012