BIG GNSS DATA

Philipp Mitterschiffthaler
Big Data

DATA IS EXPLODING
Today, every 2 minutes we are generating the same amount of data that was created from the beginnings of time until the year 2000.

Over 90% of all the data in the world was created in the past 18 months.

WHAT IS BIG DATA?

DATA IS EXPLODING
Today, every 2 minutes we are generating the same amount of data that was created from the beginnings of time until the year 2000.

Over 90% of all the data in the world was created in the past 18 months.

EVERY MINUTE we send over 200 million emails, generate 1 million Facebook likes and over 250 thousand Tweets, and upload from 200,000 photos to Facebook.

GOOGLE processes over 40 thousand queries per second; or over 3.5 billion per single day.

If you burned all the data created in just one day onto DVDs, you could stack them on top of each other and reach the moon - twice.

100 HOURS OF VIDEO are uploaded to YouTube every minute and it would take you around 15 years to watch every video uploaded by users in one day.

The number of Bits of information stored in the digital universe is thought to have exceeded the number of stars in the physical universe in 2007.

The boom of the Internet of Things will mean that the amount of devices connected to the Internet will rise from about 13 billion today to 50 billion by 2020.

Today, companies use less than 10% of the data they own, and less than 1% of the data they have access to.

Bernard Marr, 16.2.15

Philipp Mitterschifftaler
Big Data

Volume

Velocity

Variety
relational data models

established in business

very well for isolated systems

changes in data model are very time-consuming

Joins will get very complex and confusingly
graph based data models

Simply structured!
flexible for data
flexible regarding to the data scheme
interoperability
loosely coupled system
RDF – Ressource Description Framework

Is designed to be read by computers

Is written in xml

Can be queried with SPARQL

So called Triplestores

<table>
<thead>
<tr>
<th>subject</th>
<th>predicate</th>
<th>object</th>
</tr>
</thead>
<tbody>
<tr>
<td>San Sebastian</td>
<td>Is part of</td>
<td>Spain</td>
</tr>
<tr>
<td>San Sebastian</td>
<td>Is</td>
<td>cultural capital</td>
</tr>
<tr>
<td>San Sebastian</td>
<td>has population of</td>
<td>180.000</td>
</tr>
</tbody>
</table>
RDF – Ressource Description Framework
www – Internet of websites and documents
LOD – Internet of data

Linked Open data

http://lod-cloud.net
Acknowledgement

- django
- SciPy
- OpenLayers 3.0
- G-Nut/Anubis
- jQuery
- Data-Driven Documents
Thank you!

Philipp Mitterschiffthaler
Federal Office of Metrology and Surveying
Austria

philipp.mitterschiffthaler@bev.gv.at