

A graphic with the words "BIG DATA" in large, white, bold, sans-serif capital letters. The text is centered against a dark blue background. Surrounding the text is a network of glowing blue nodes connected by thin, light blue lines, resembling a molecular structure or a data network. The nodes vary in size and brightness, with some appearing as bright white dots and others as smaller, dimmer blue spheres.

**BIG
DATA**

Big Data projects at statistical agencies abroad

On overview

Piet J.H. Daas
Senior methodologist & project leader Big Data research



**Statistics
Netherlands**

19 Nov., NAS BEA

Big Data

- There are many 'Big Data' projects out there
 - Fortunately an Inventory was done by UN Big Data project
- But:
 - What is a Big Data project?
 - With or without admin. data
 - Need to discern between plans (ideas) and practical application
 - Many talk about it, only a few actually 'do' it
 - Few countries currently actually produce Big Data based statistics !

GWG survey data for Europe

- Global Working Group on Big Data for Official Statistics (UN) has send out a survey:
“Global assessment of the use of Big Data for Official Statistics”
- Result:
114 projects from 43 countries/organizations (10 Sept. 2015)

Exploration vs ‘intended for’ production (about 1:1)

Results (wordcloud)



Project groups

1. Web data/web scraping (21)
collect prices, job vacancies, enterprise information ...
2. Scanner data (18)
for CPI
3. Mobile phone data (15)
Tourism, border crossings, 'day time population'
4. Social media/Google trends (8)
Fast indicators & now-casts: sentiment, unemployment ...
5. Satellite/aerial imagery data (6)
Land use, crops, 'poverty' ...
6. Other (46)
Smart meters, transport (land, water), health, credit-card, patents Incl. admin data

Concluding remarks

- Big Data is an interesting data source for official statistics with great *potential*
- Only a few statistical offices are currently actually using it!
 - Current uses are:
 - Scanner data for CPI (at least 4 countries; NL, AUS, BEL, SWI)
 - Data from the web for CPI or alike (NL, CAN, KOR)
 - Road sensor data for traffic index (NL) or commuting (Fin)?

The Future



The
future
of
statistics
looks

BIG



The data deluge

