

The need for research on
data science techniques and new data sources
in official statistics

Michael Reusens - Data Science Coordinator – Statistics Flanders

Statistics Flanders data science hub

➔ Create new and better statistics for Flanders using new data sources and data science techniques

➔ Since 2019

- 1. Projects:** create new and better official statistics using data science techniques and new data sources
- 2. Data science education** for Flemish government
- 3. Network** and collaborate

Traditional official statistics production

➔ Using surveys: designed data

⊕ Labor intensive

⊕ High cost

⊕ Infrequent

⊕ Response burden

⊕ dropping response rates

➔ More and more administrative data

Data science hub project focus

- ② **Do more with currently used data** (survey and administrative)
 - ⊕ Machine learning, now- and forecasting
 - ⊕ Exploiting unstructured data fields (e.g. Natural Language Processing)

- ② **Exploit new data sources: found data**
 - ⊕ Earth observation, social media, web data, sensor data, ... (found data)

Designed data vs found data

- ➔ **Designed data:** data resulting from a data collection processed tailored for statistics production
 - ⊕ Many existing methods and quality frameworks

- ➔ **Found data:** data resulting from any process or application, without statistics production in mind
 - ⊕ Unknown optimal methods and lack of quality frameworks

Descriptive to predictive

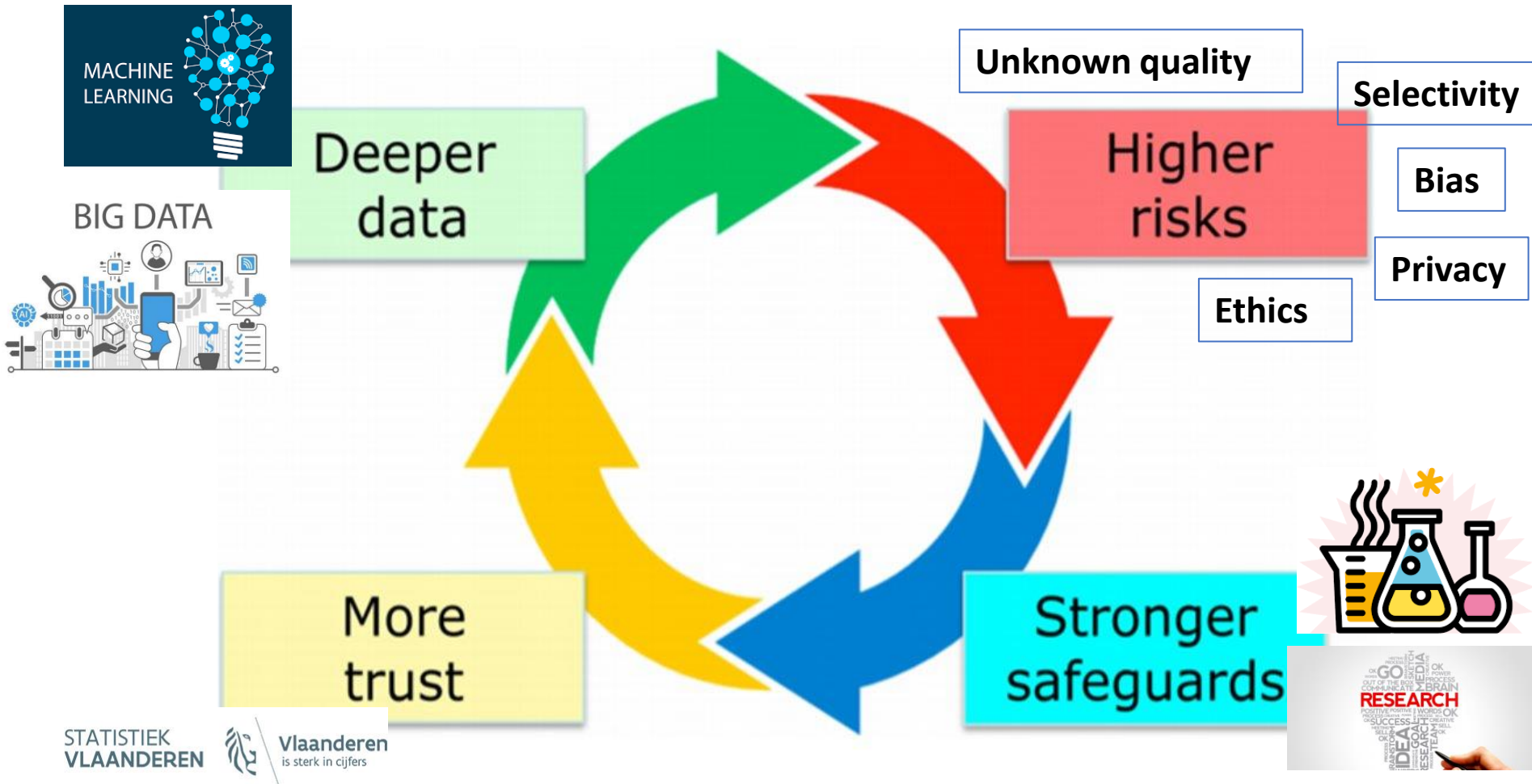
➔ Descriptive analysis

- ⊕ Well studied in official statistics community

➔ Predictive methods

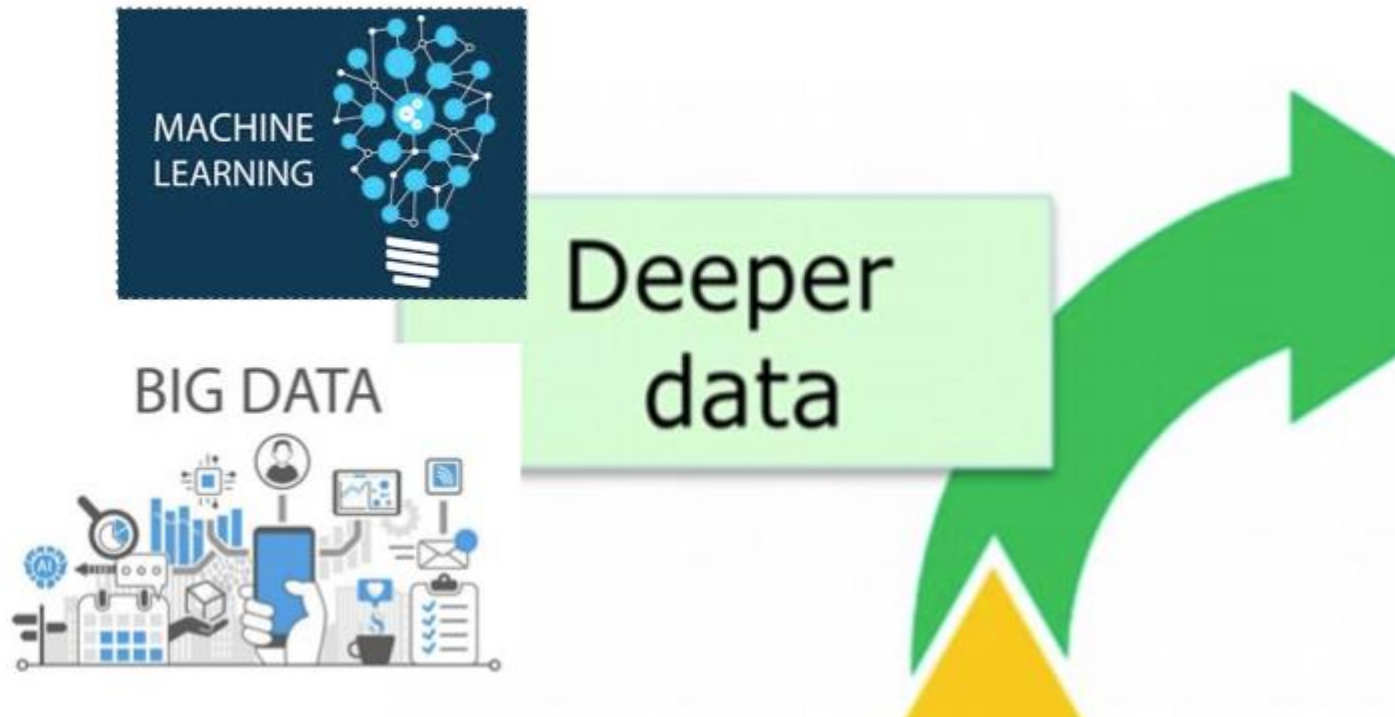
- ⊕ Lesser known in official statistics community
- ⊕ Need more know-how (which methods work best?)
- ⊕ Need quality frameworks (how well do the methods work?)
- ⊕ --> More uncertainty

The smart and trusted cycle



Source: Ricciato e.a., 2019

Research



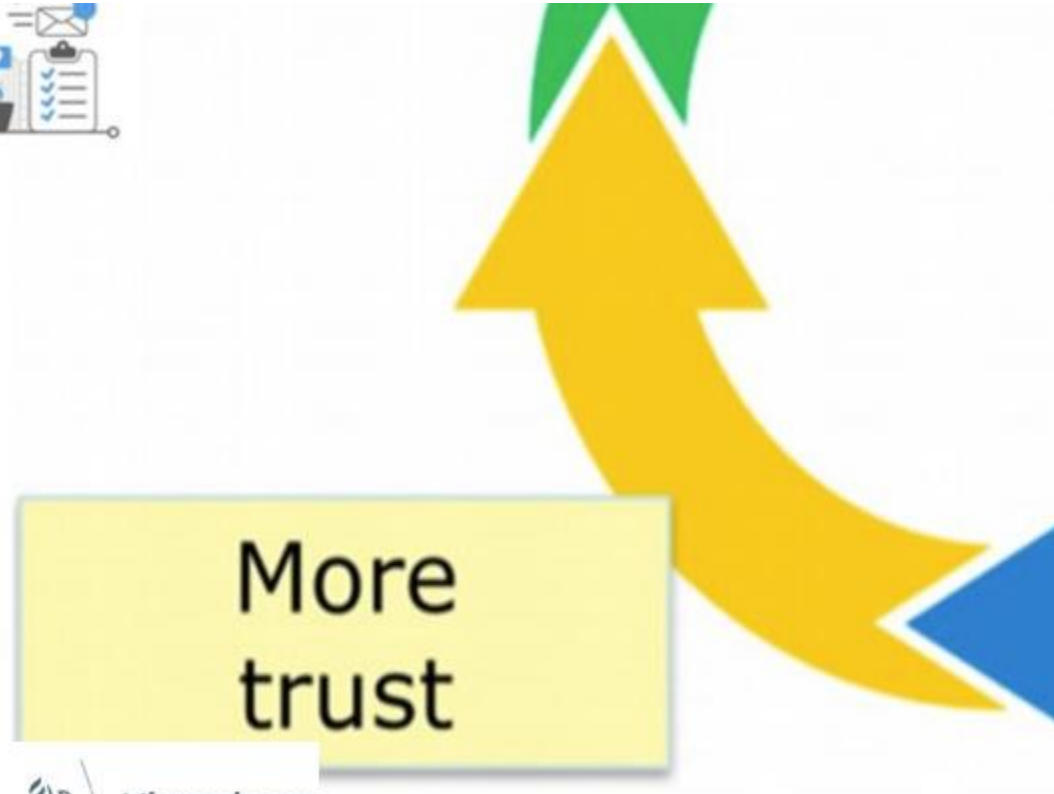
Research



Research



Data science hub



STATISTIEK
VLAANDEREN



Vlaanderen
is sterk in cijfers

Research collaborations

- ⊗ Official research collaboration with LIRIS, KU Leuven
 - ⊕ PhD Student: Manon Reusens
 - ⊕ Yearly master thesis students

- ⊗ Master theses and internships with other knowledge institutes

Thank you

➔ Let's take a look at the research of the past academic year!

➔ Contact: michael.reusens@vlaanderen.be