6th International Conference on Big Data for Official Statistics

SESSION 3: USING BIG DATA FOR SDGS - MOBILE DATA FOR TOURISM, MIGRATION, POPULATION AND TRANSPORT

Statistics Netherlands

May Offermans

Statistics Netherlands research for use of mobile phone data in statistics

1) Introduction
2) Practical challenges and privacy
3) Examples
4) Conclusion
5) Opportunities, standardisation and collaboration

1) Introduction

- CBS has years of experience in research of mobile phone data to look for application in official statistics.
- Need in society for fast statistics because of new developments in society (platform economy and tourism). Gekeken is naar de toepassing in de statistiek
- Results for application are impressive
- Same development in countries in Europe (Germany, Italy, Belgium, France etc.)

2) Practical challenges and privacy (1)

a) Data access

-> There are three operators in the Netherlands

-> Privacy authority and public relations/opinion

b) System for privacy protection

a) No connections on a personal level or tracking (this is not acceptable).
b) We make automated counts based anonymous analyses
c) Connect data on a meso/macro level on space and time
d) Based on 100x100 metergrids

c) Technical Challenges

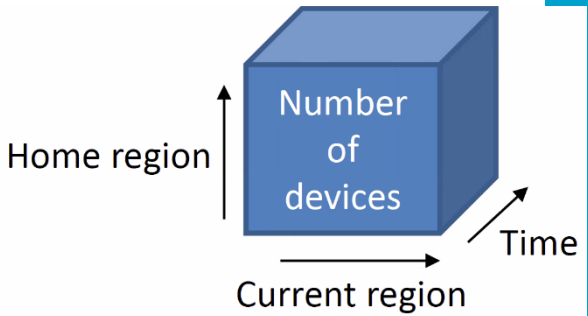
a) Determine geolocation
b) Determine resident
c) Very large quanties of signaling/cdr data (>5 miljard records a day)
d) Processing time in hours (using SPARK)
e) Implemting and run systems at the site of the operator

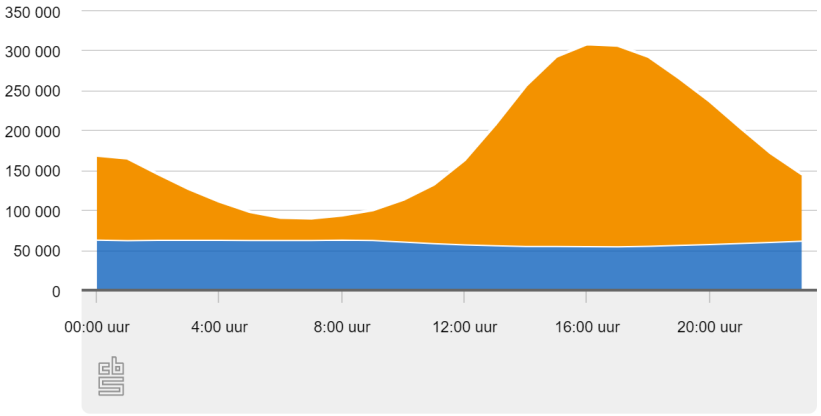
3) Examples

3) Examples (1) Kingsday in Amsterdam

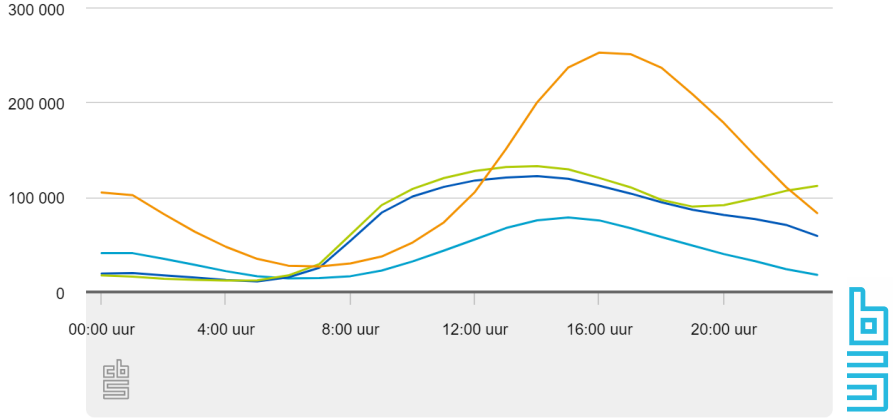
Daytime population, how many people are there and where are they from?

Origen Destination Cube



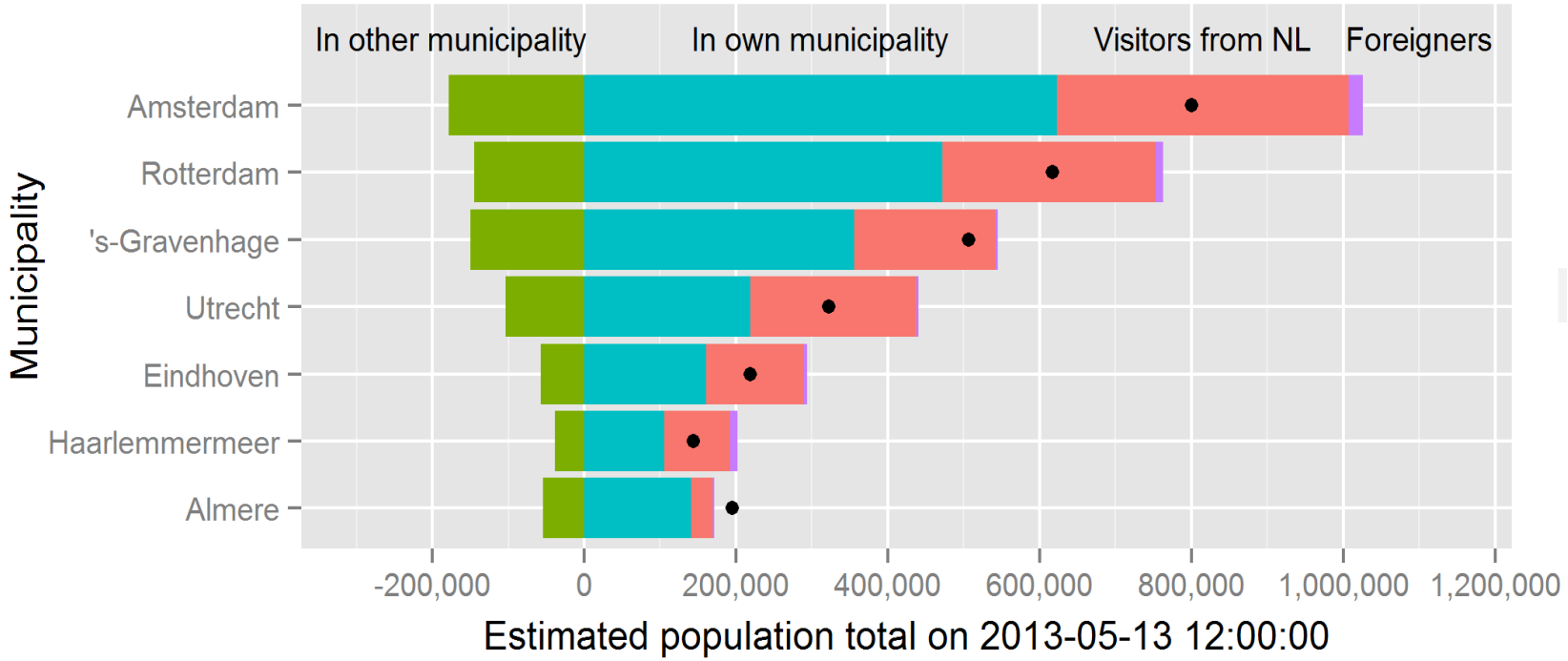


Bron: CBS, Mezuro



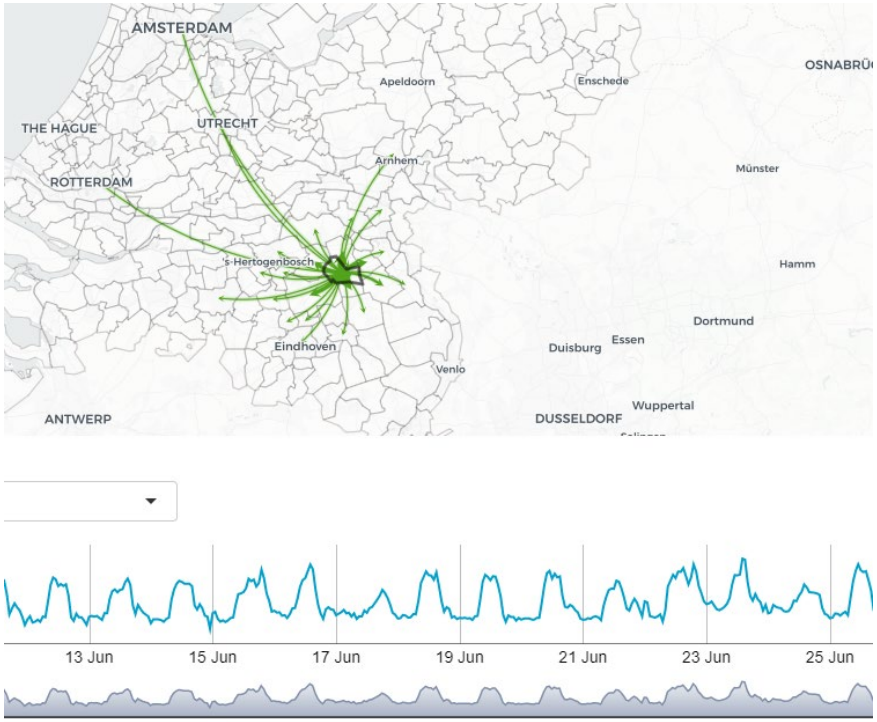
Bron: CBS, Mezuro

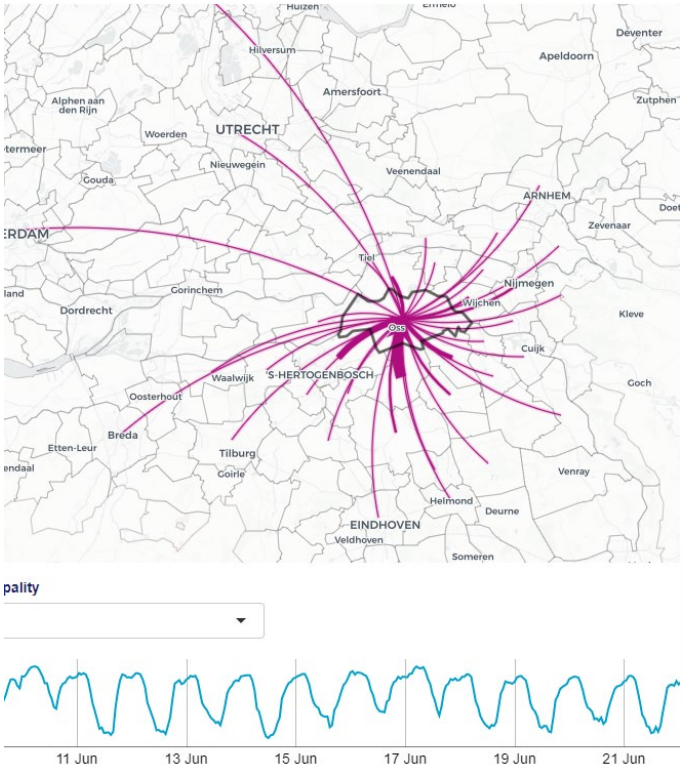
3) Example Amsterdam (2)



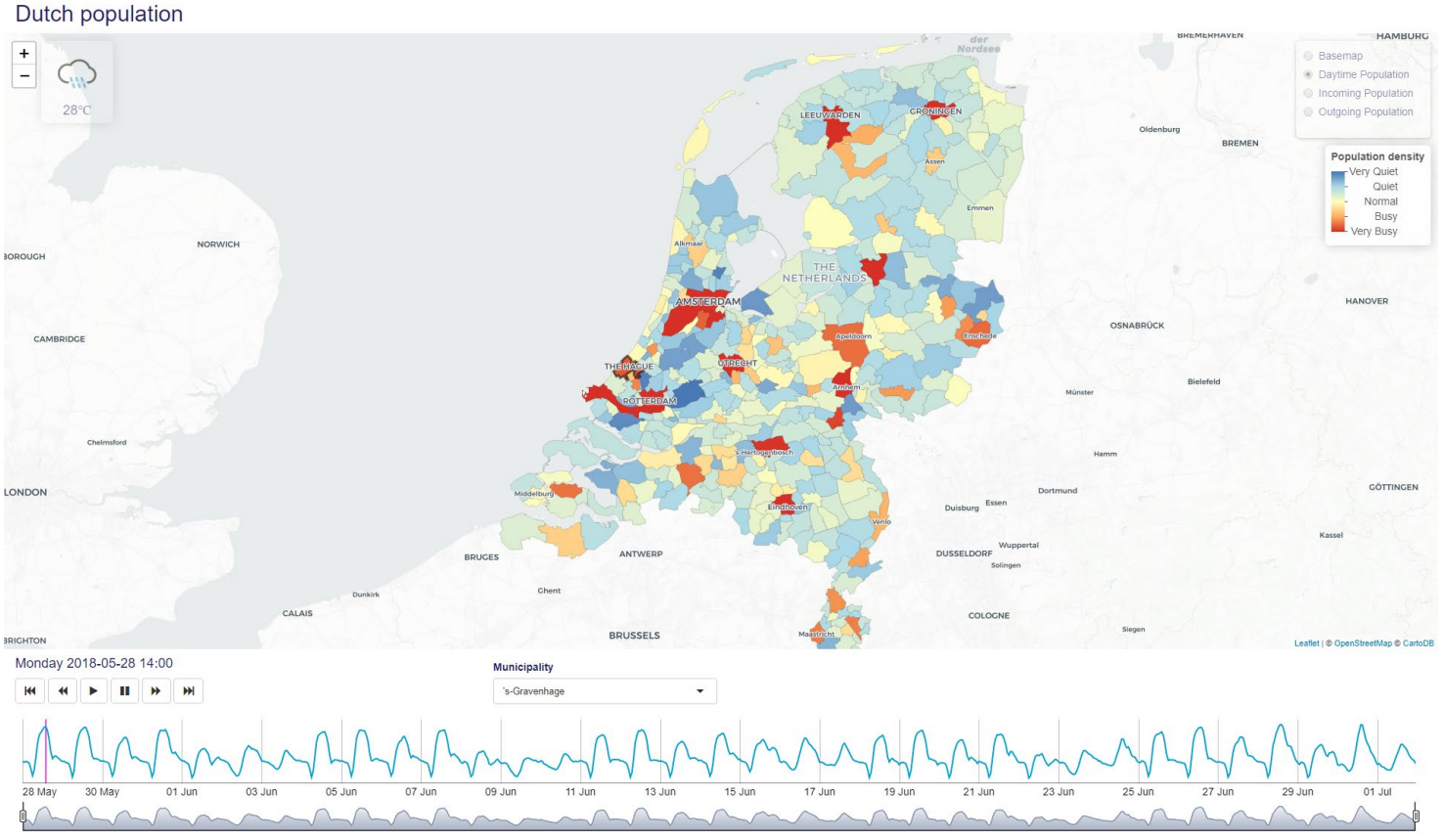
3) Example Daytime population the Netherlands (3)

https://dashboards.cbs.nl/v1/dtp/





3) Example Daytime population the Netherlands (4)

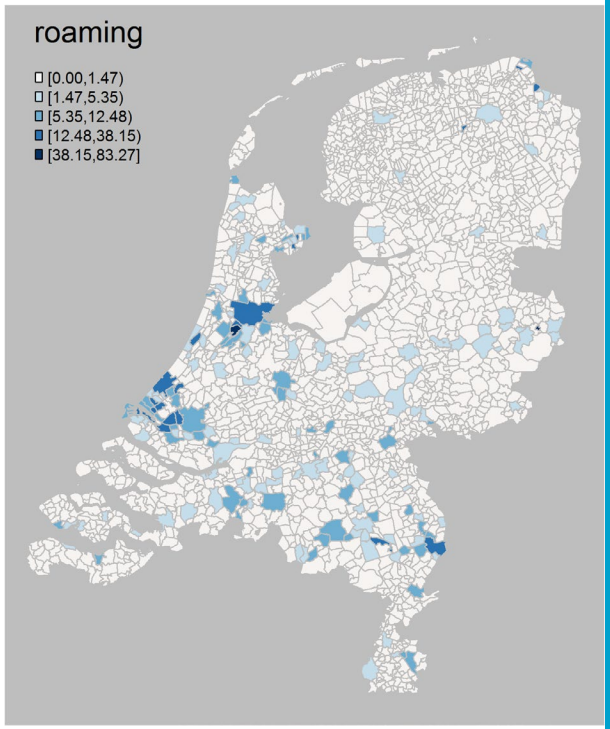


3) Tourism and commuters from other countries (5)

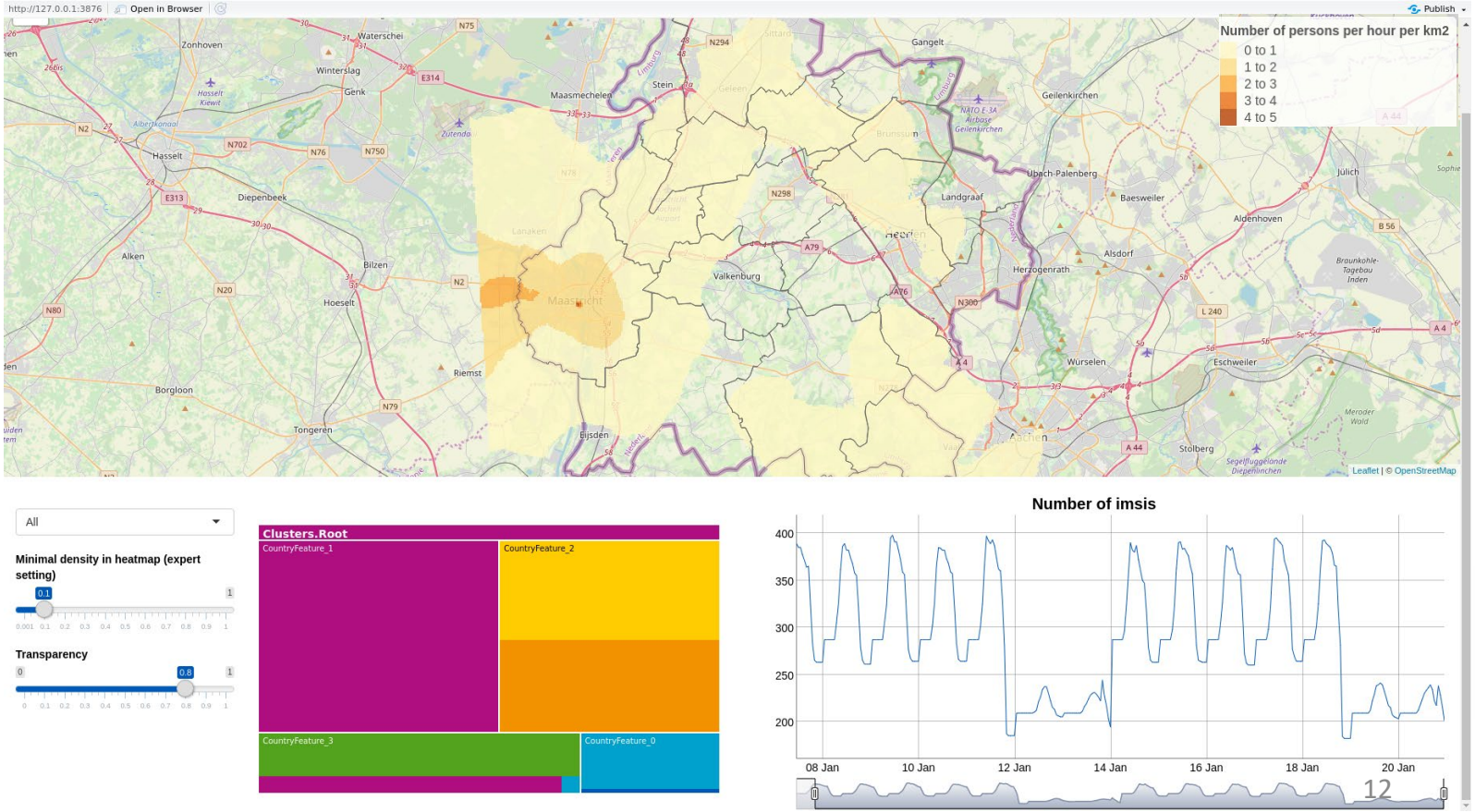
- Duration, point of entry
- Incoming and outbound tourism
- Commuter or tourist?

- Machine learning technology

Roaming, uit Oost Europa

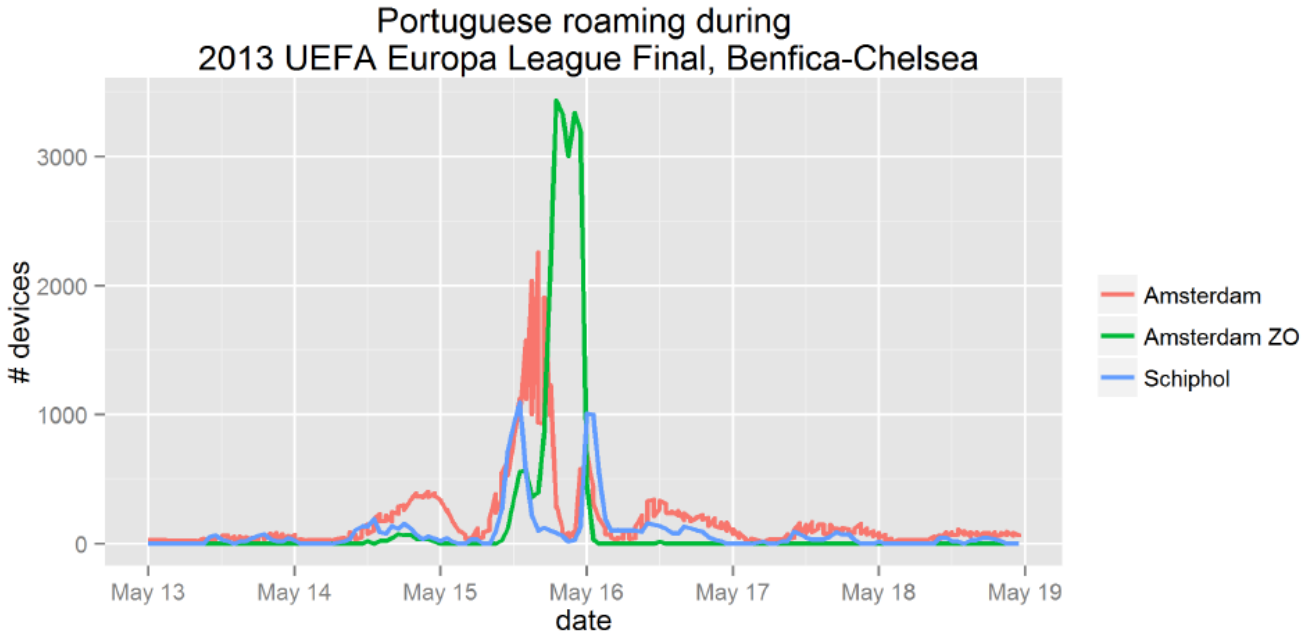


3) Testing machine learning for tourism (6)



3) Voorbeelden van toepassingen (7)

Point of entry



4) Conclusion

Mobile phone data are extremey privacy sensitve.

Using state of the art methods, they can be used in a responsable way for official statistics.

- Until now nothing has been implemented in official statistics
- Improvements and validation
- New quality frameworks
- International collaboration and standardisation (Positium, Flowminder, Eurostat)
- Discuss and explain and be transparent about privacy

Problems & opportunities

Keep momentum

Statistics Netherlands

positium

eurostat

FLOWMINDER.ORG

Problems & Opportunities –

Why collaboration and standaardisation is needed

- Solving privacy, technical and quality of data challenges use on our own takes a lot of energy and expertise

- Loss of momentum limits progress on:

  - Further development
  - Usage of data for public good, why?
    - No energy left for documentation or publication on methods
    - Perceived risks limits data access opportunities
    - Causing projects to fail
    - Requires high investments on tools or expertise

General Public: Reliability and trust

Explain the complex method (be transparant) and inform about privacy