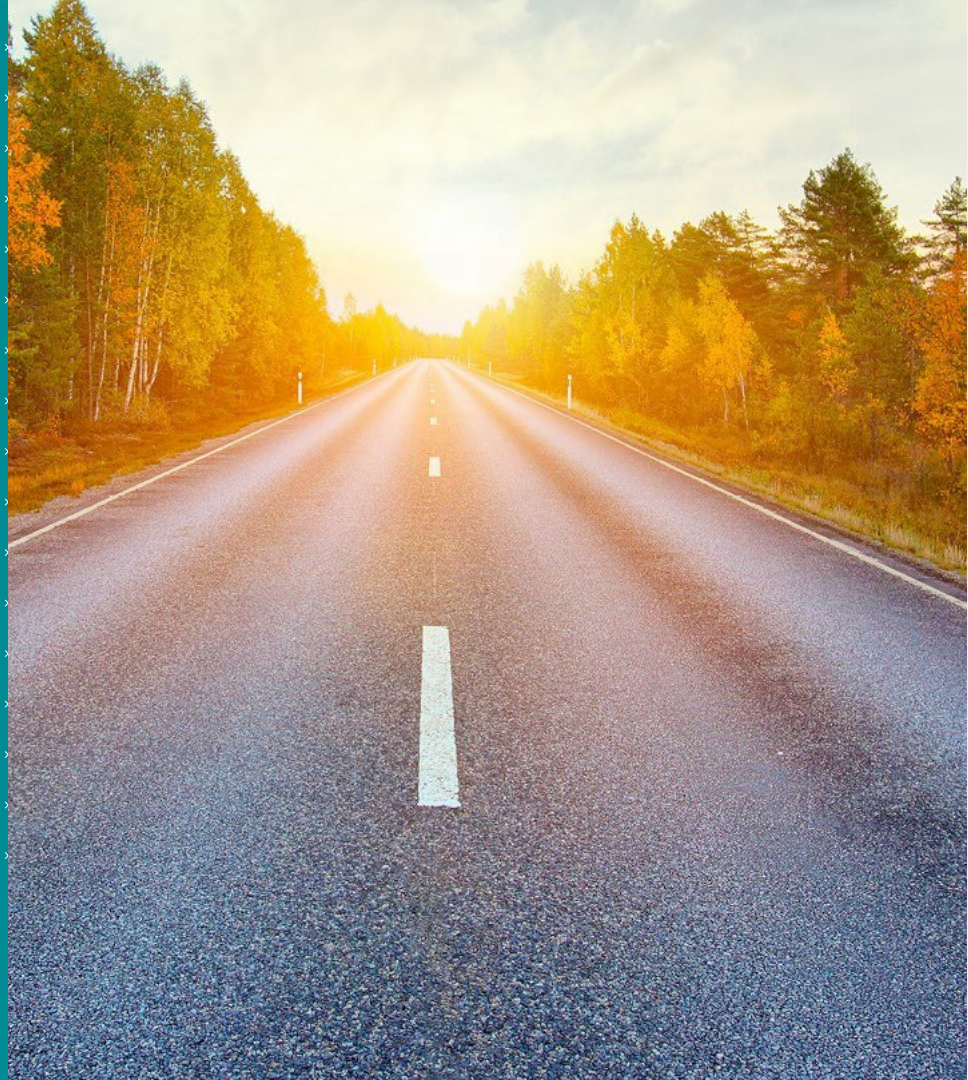
Big data!

- Nytta & Utmaningar

Stockholm, Nov 25, 2019

Cristofer Englund

RISE Research Institutes of Sweden AB



Data finns och skapas ständigt runt oss

Big Data (three V's)
- Volume
- Velocity
- Variety

Who is Cristofer?

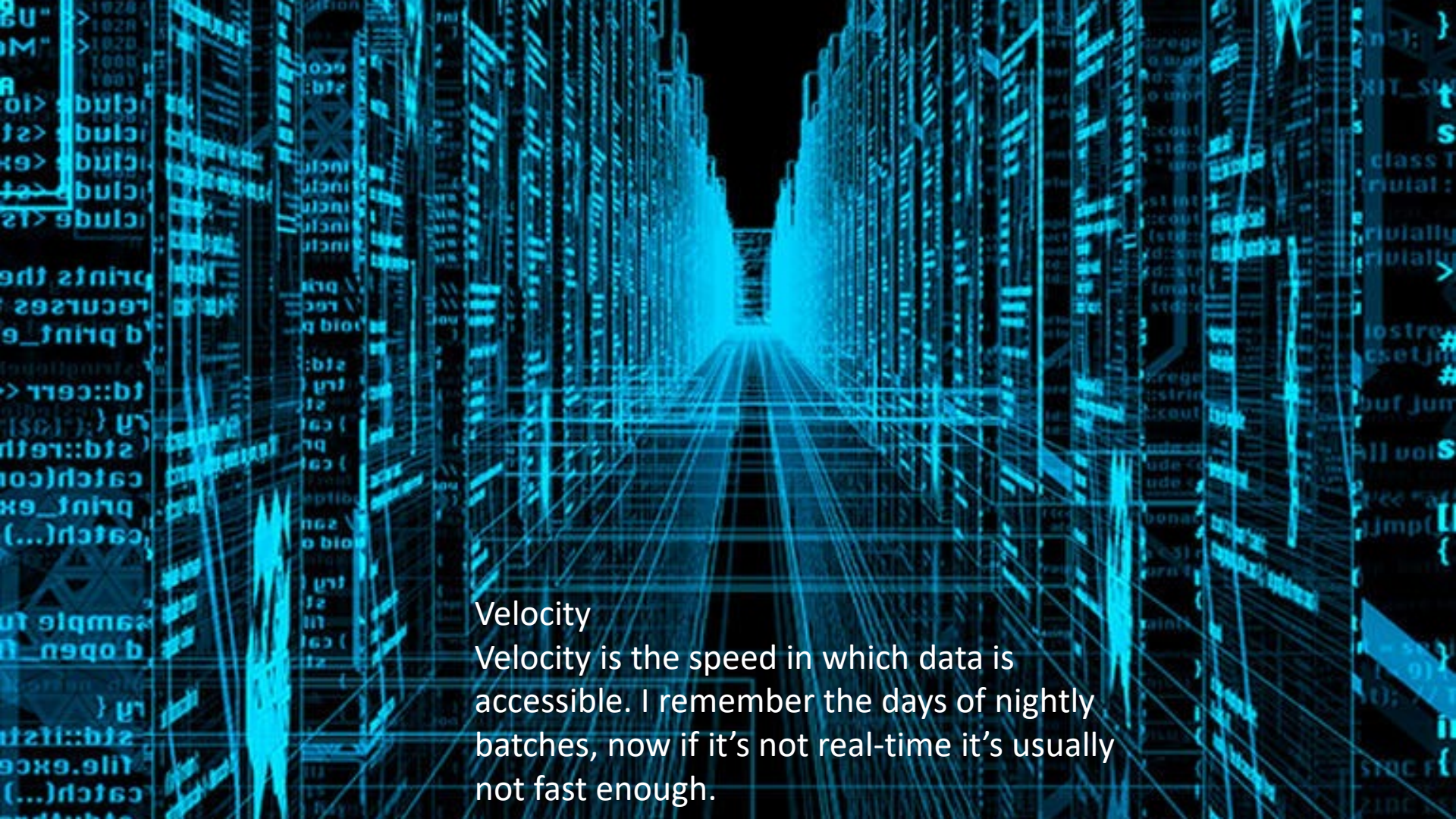
Research manager, Göteborg

Associate Professor, Halmstad University

Research interest: AI, Machine Learning, cooperative and automated vehicles, behavior modelling, traffic systems

Volume

Volume is how much data we have – what used to be measured in Gigabytes is now measured in Zettabytes (ZB) or even Yottabytes (YB). The IoT (Internet of Things) is creating exponential growth in data.



Velocity

Velocity is the speed in which data is accessible. I remember the days of nightly batches, now if it’s not real-time it’s usually not fast enough.

Variety describes one of the biggest challenges of big data. It can be unstructured and it can include so many different types of data from XML to video to SMS. Organizing the data in a meaningful way is no simple task, especially when the data itself changes rapidly.

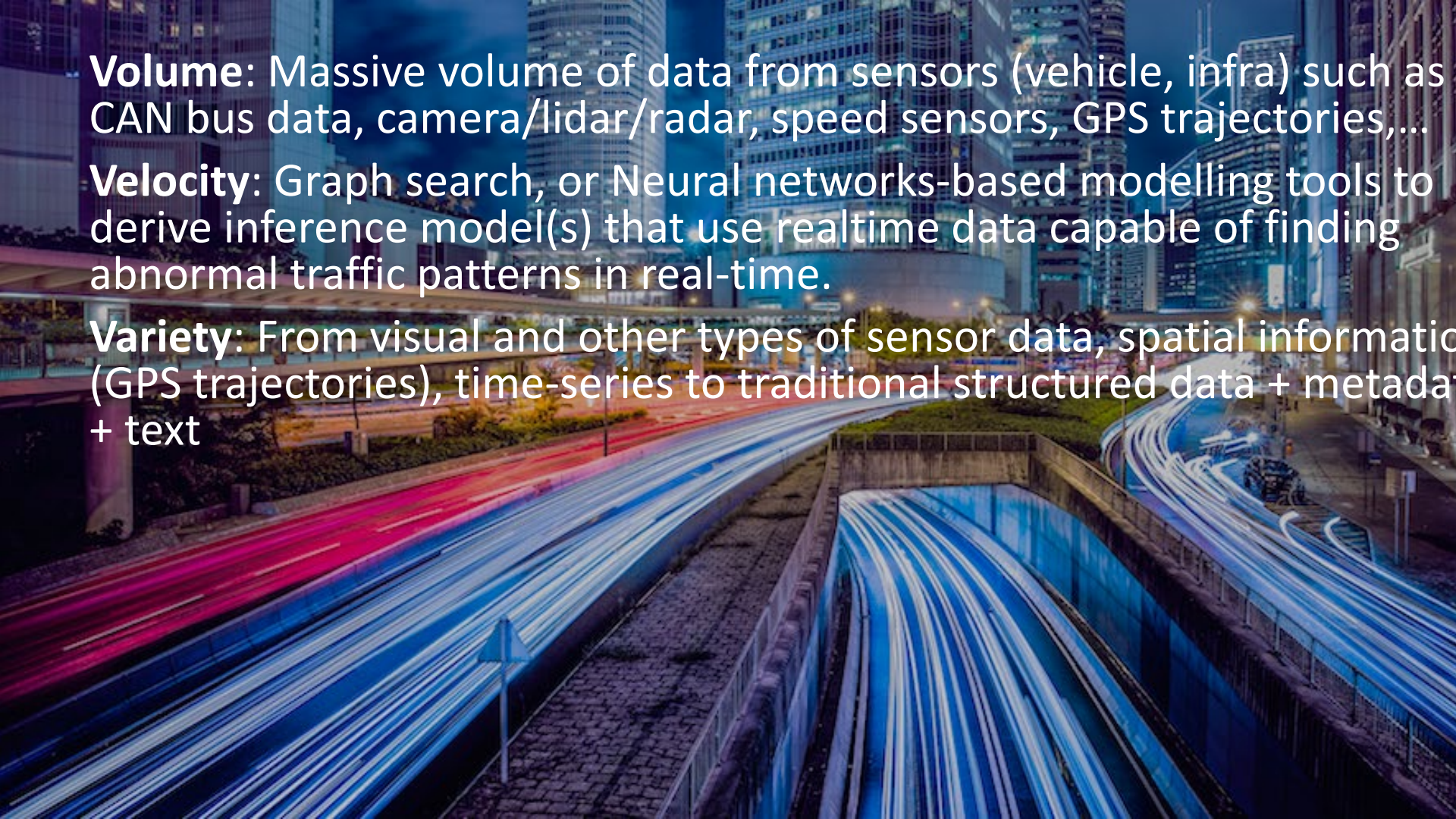
CHALLENGES
AHEAD

”Data is cheap but it’s dumb”

”Information is the oil of the 21st century, and analytics is the combustion engine”

”Big data is not about the data”

Peter Sondergaard, Senior Vice President, Gartner

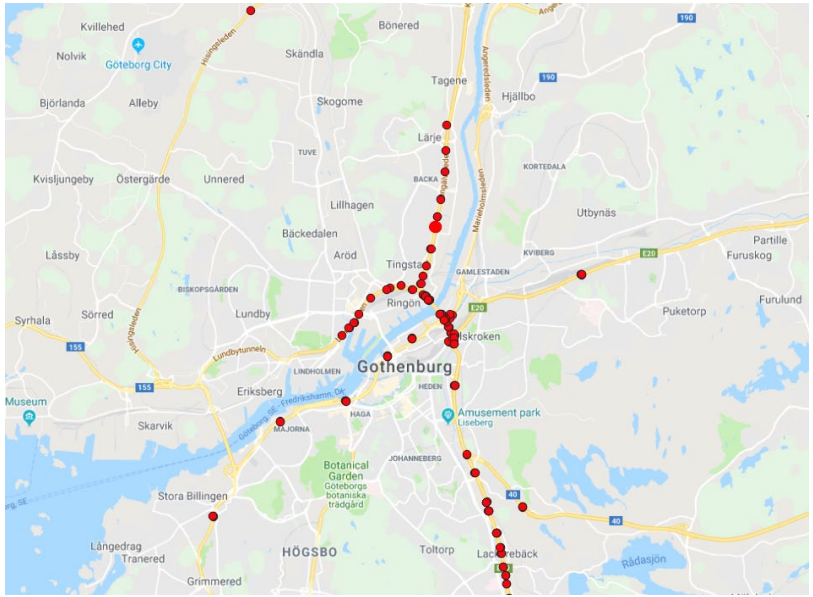


Volume: Massive volume of data from sensors (vehicle, infra) such as CAN bus data, camera/lidar/radar, speed sensors, GPS trajectories,…

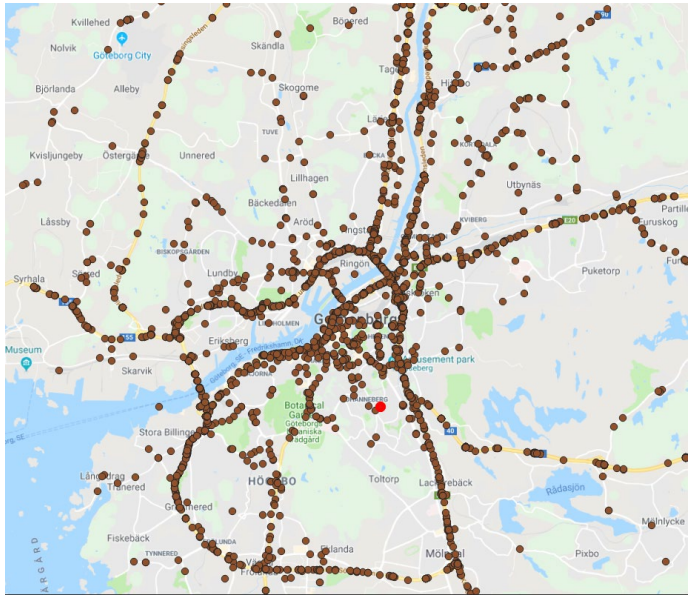
Velocity: Graph search, or Neural networks-based modelling tools to derive inference model(s) that use realtime data capable of finding abnormal traffic patterns in real-time.

Variety: From visual and other types of sensor data, spatial informatio (GPS trajectories), time-series to traditional structured data + metada + text

What we did
– Data driven traffic management policy



Fixed infrastructure sensors



Traffic related events

What we did
– Data driven traffic management policy

Analysis

Prediction

What we did
– Freefloating car service analysis

Google

CAR 2GO

DriveNow

Weatherstations

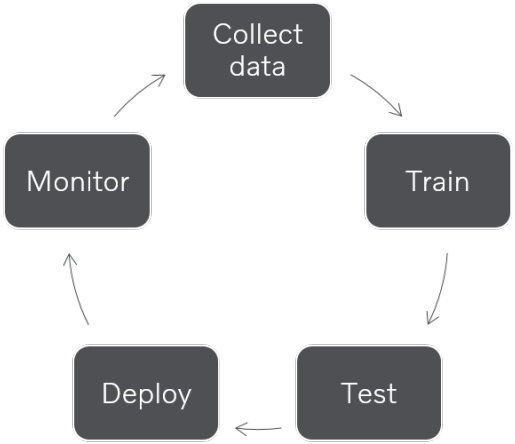
RI. SE

CHALMERS

Energimyndigheten

RI. SE

What we did
– Freefloating car service analysis



- Data collection
- Data alignment
- Data cleaning
- Data pre-processing
- Data visualization
- Data clustering

Google

CAR 2GO

DriveNow

Weatherstations

RI. SE

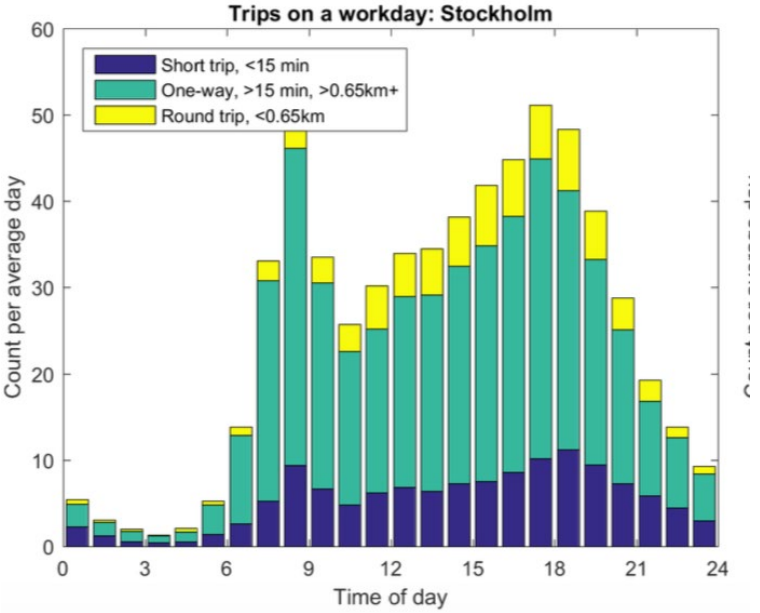
CHALMERS

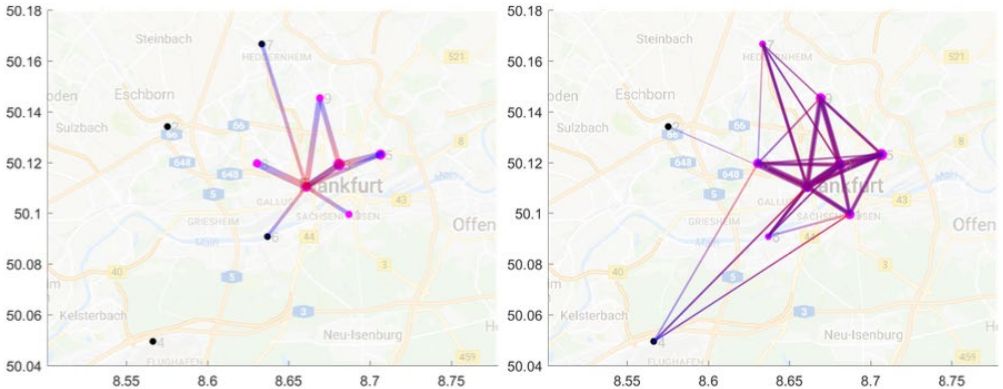
Energimyndigheten

RI. SE

Free floating car sharing behaviour analysis

- Business analytics
  - How is the service used?
  - How can the service be improved?
  - Between where is the service used?
  - Should the operator re-distribute vehicles?





(a) Frankfurt-Morning (b) Frankfurt-Evening

To organize, visualize, clear, model and draw conclusion
you often need data analytics skills on the PhD level
Education!
Tools!

Wrap-up

Volume
Velocity
Variety

Strategic and operational planning
Business operation

Education

cristofer.englund@ri.se