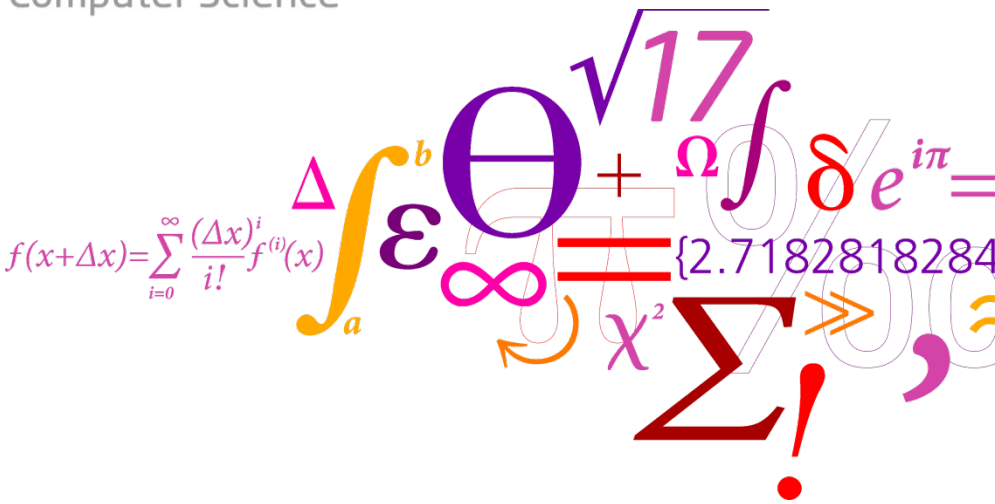
LiRA Status Workshop: WP3
Production and Validation

Ekkart Kindler, Shahrzad M. Pour and
Markus Berthold

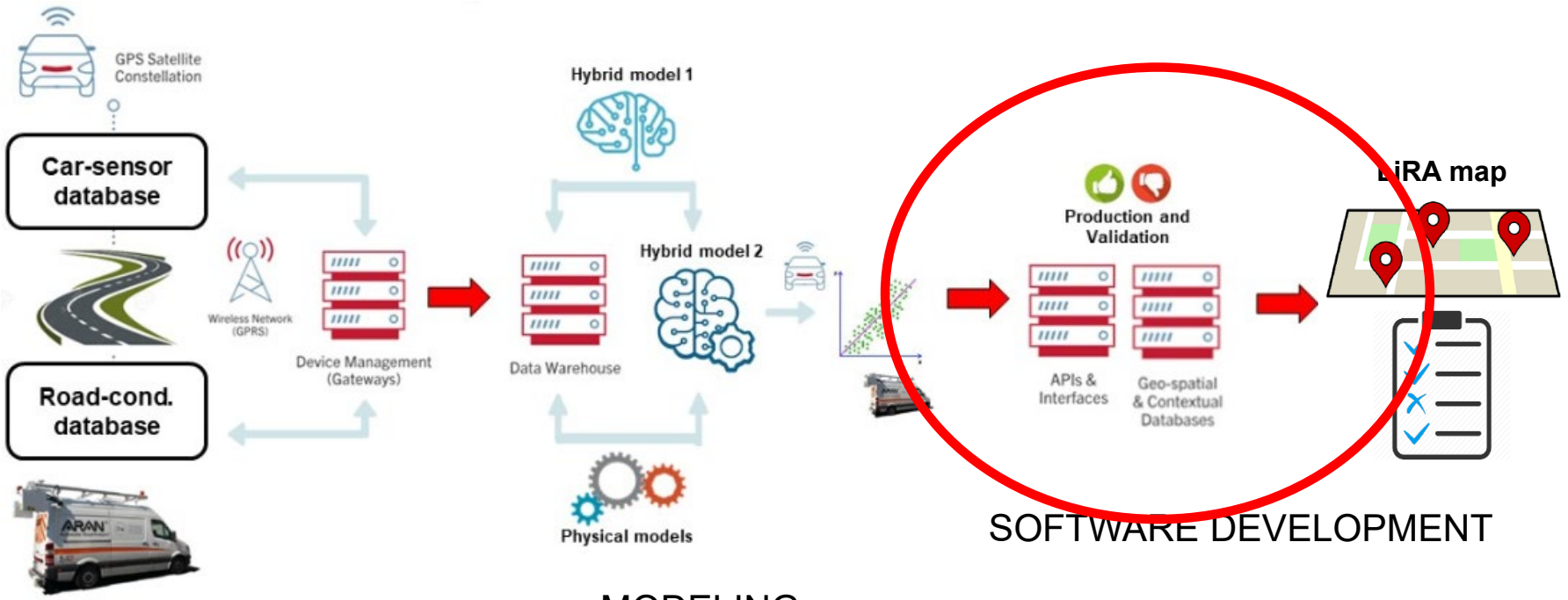
DTU Compute
Department of Applied Mathematics and Computer Science



Overview

DTU Compute
Department of Applied Mathematics and Computer Science
Ekkart Kindler

First slide from Kick-off Meeting in February 2019



Changes of Plan

DTU Compute
Department of Applied Mathematics and Computer Science
Ekkart Kindler

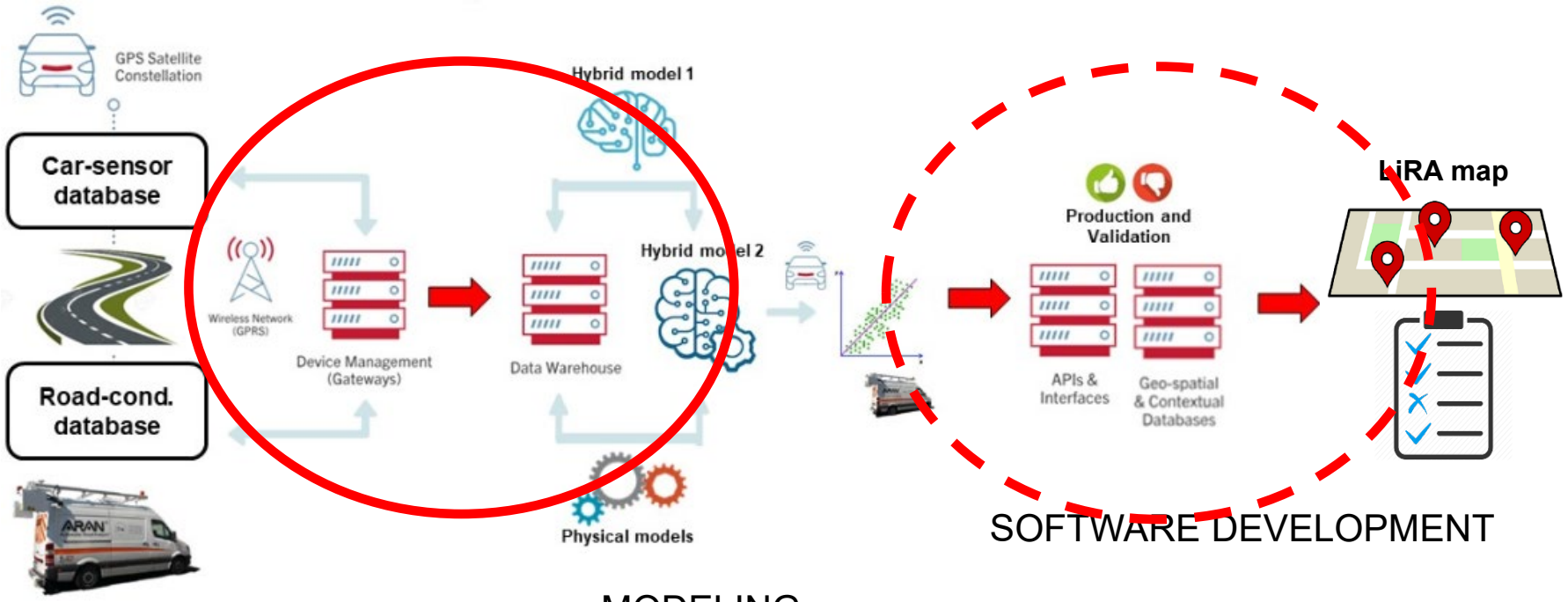
- Post Doc position for WP3 filled only from Oct. 14, 2019 and for 2 years full-time instead (3 years 2/3 part-time)

- After LiRA SC meeting in September 2019: shift of focus: LiRA Data Warehouse

But work on Production and Validation Tasks still actively pursued

Overview

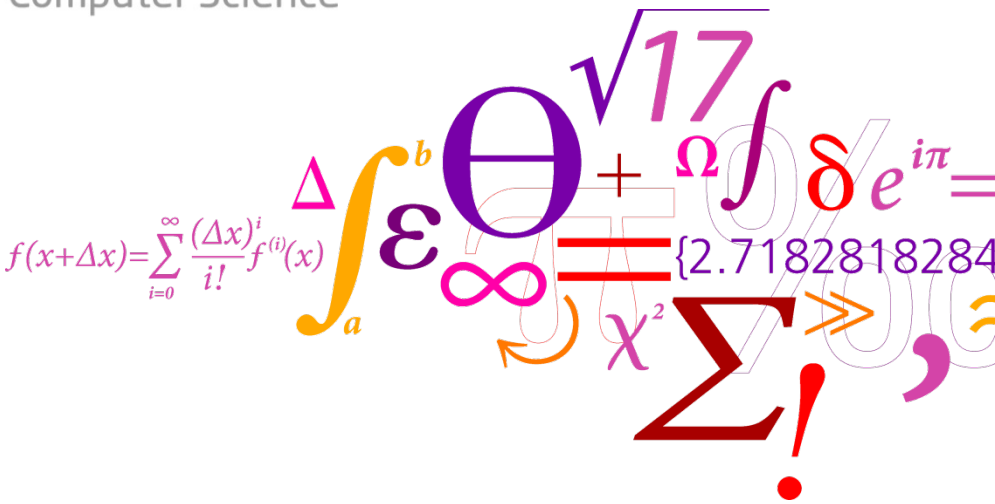
DTU Compute
Department of Applied Mathematics and Computer Science
Ekkart Kindler



LiRA Data Warehouse:
Concepts, data model, architecture and data collection and processing pipelines

Ekkart Kindler, Shahrzad M. Pour and
Markus Berthold

DTU Compute
Department of Applied Mathematics and Computer Science



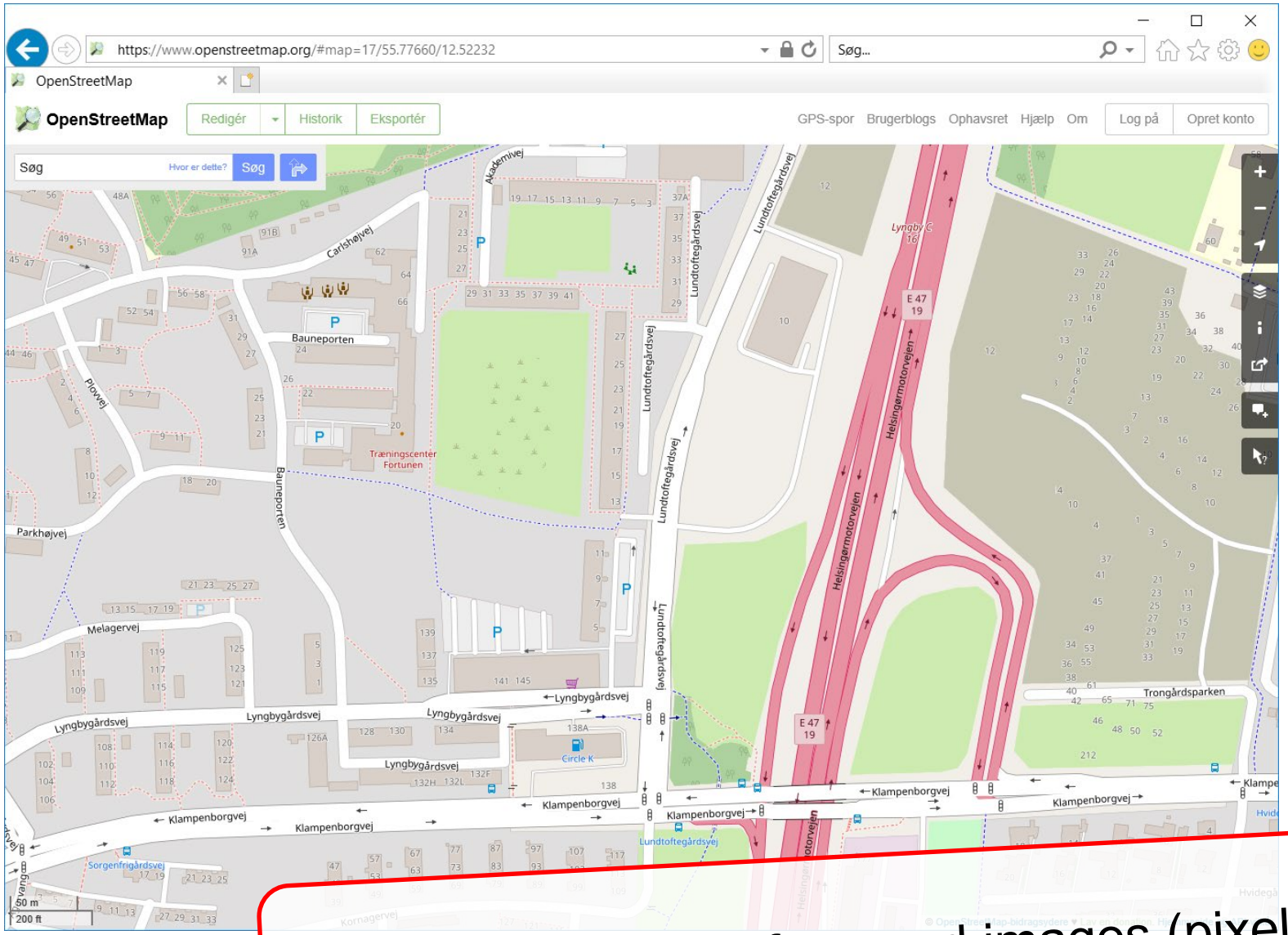
Outline

DTU Compute
Department of Applied Mathematics and Computer Science
Ekkart Kindler

- Introduction and overview
- Concepts (domain analysis)
- Design and architecture
- Detailed models
- Data processing pipelines
- Demo

Open Street Map (OSM)

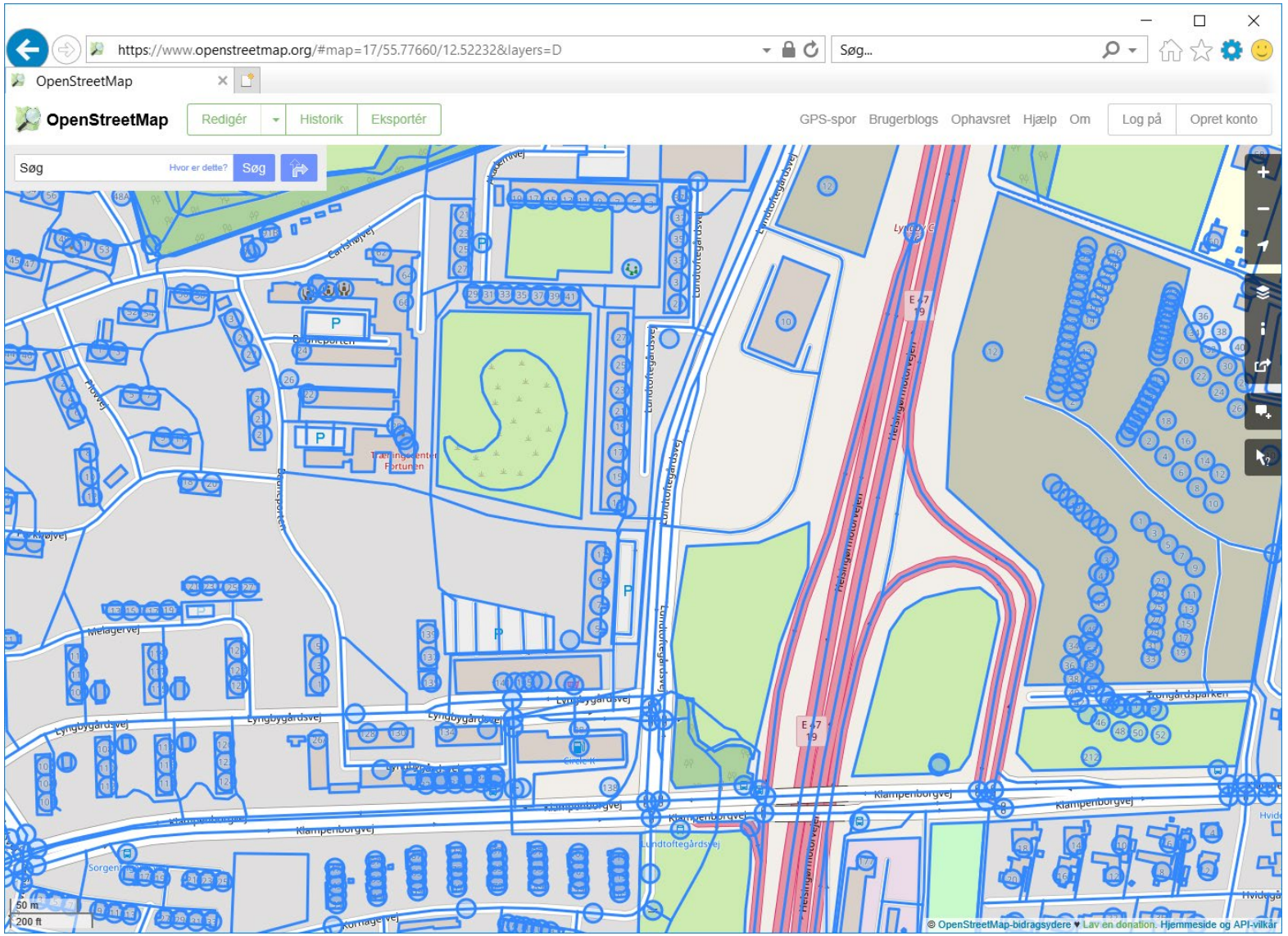
DTU Compute
Department of Applied Mathematics and Computer Science
Ekkart Kindler



Basically, just geo-referenced images (pixels)!

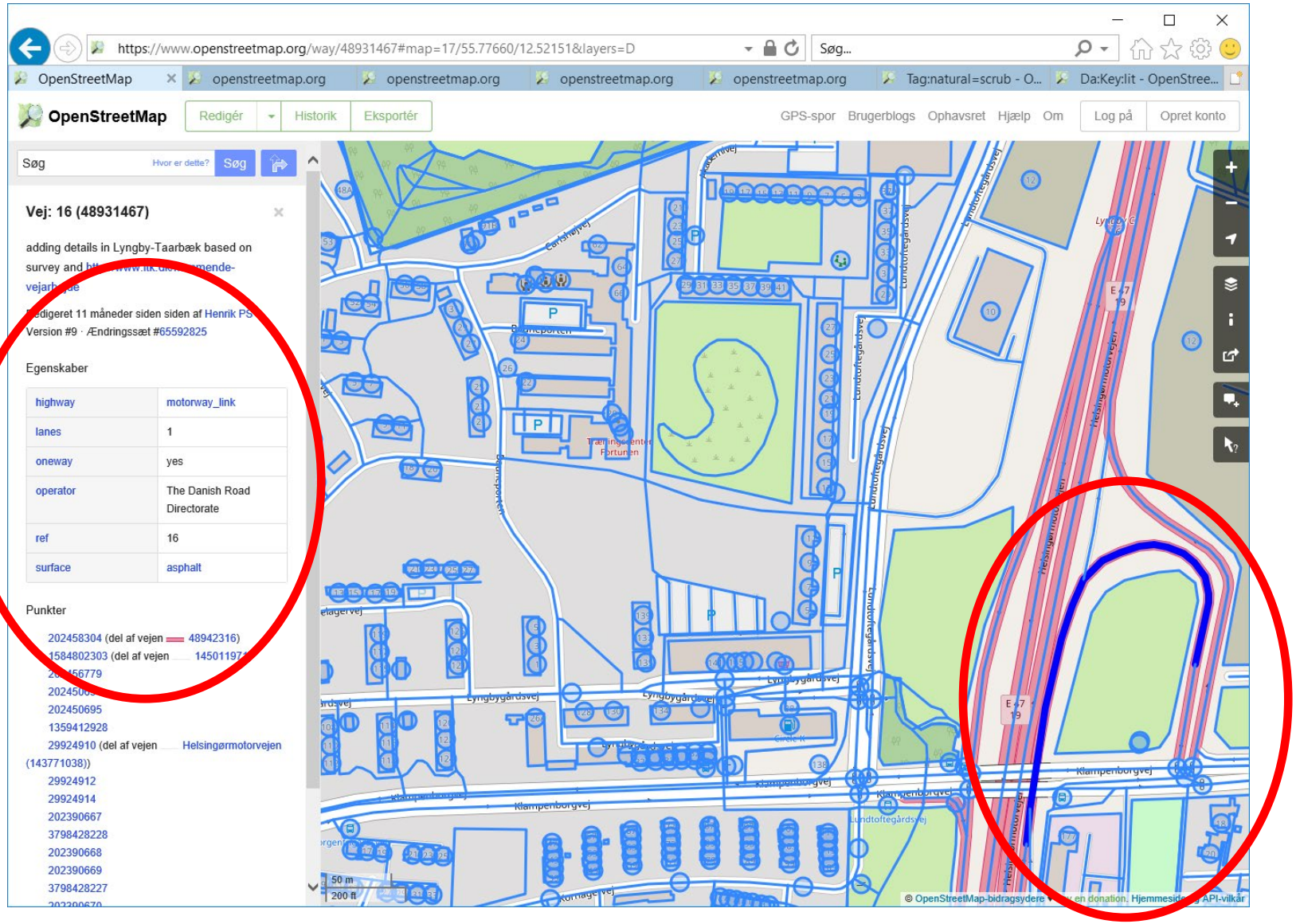
OSM: Map Data

DTU Compute
Department of Applied Mathematics and Computer Science
Ekkart Kindler



OSM: Way with Tags

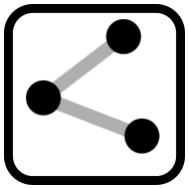
DTU Compute
Department of Applied Mathematics and Computer Science
Ekkart Kindler



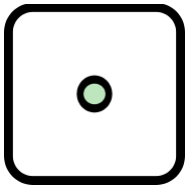
Core Concepts

DTU Compute
Department of Applied Mathematics and Computer Science
Ekkart Kindler

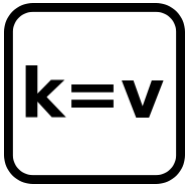
- Way / Section



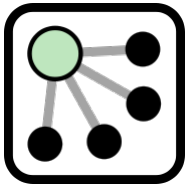
- Node / Point



- Tags / Attributes

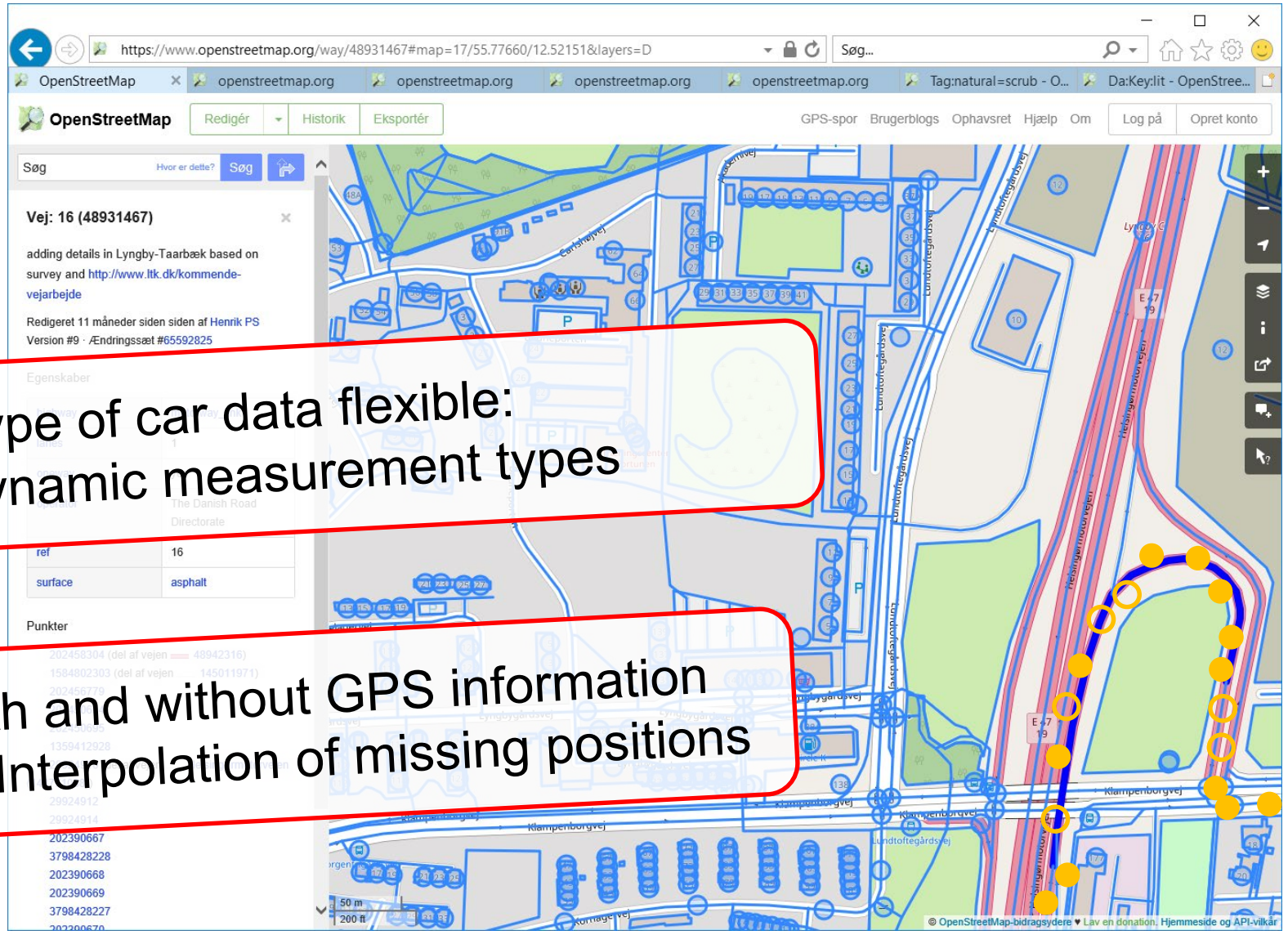


- Relations



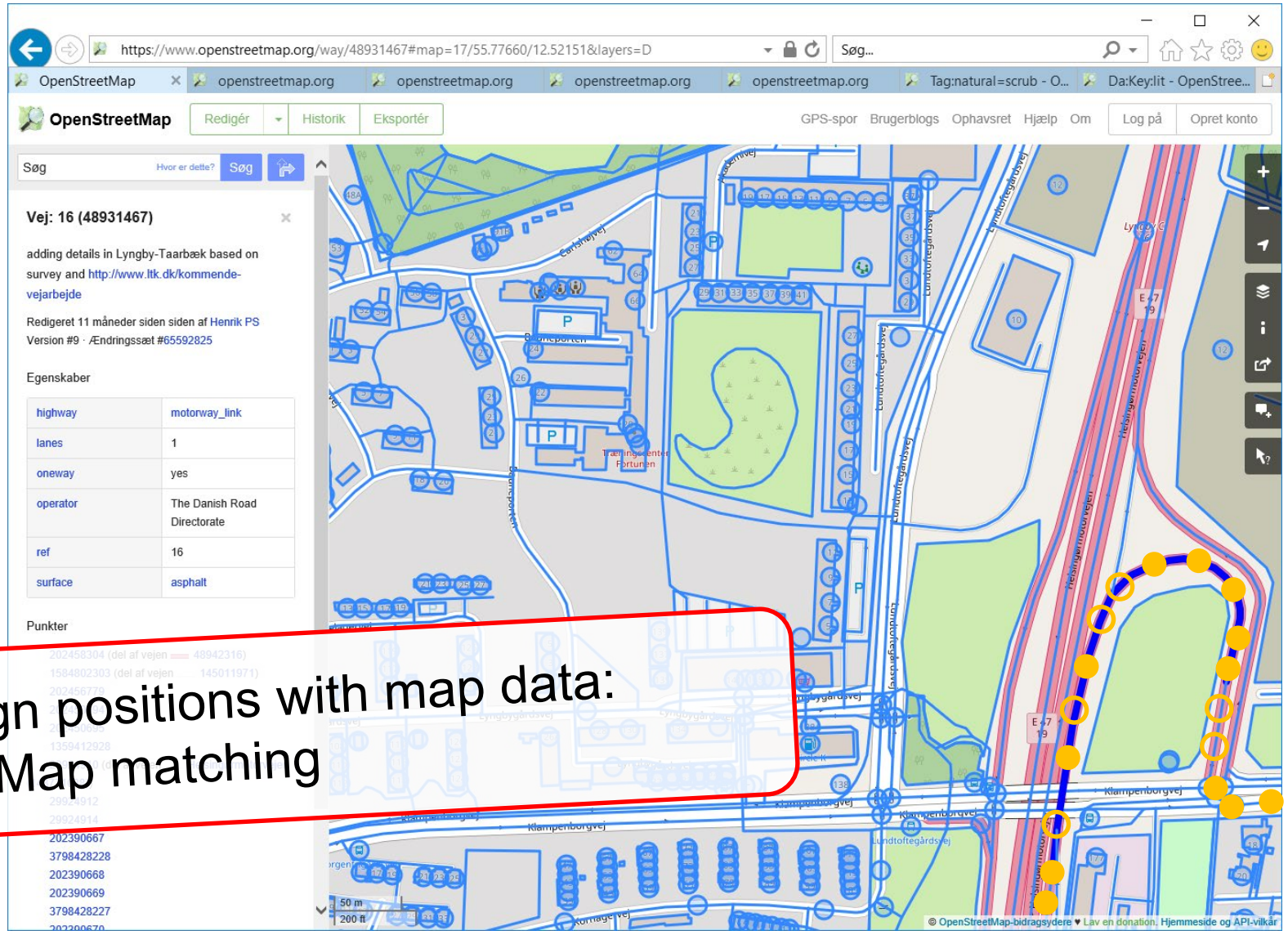
Car data

DTU Compute
Department of Applied Mathematics and Computer Science
Ekkart Kindler



Car data

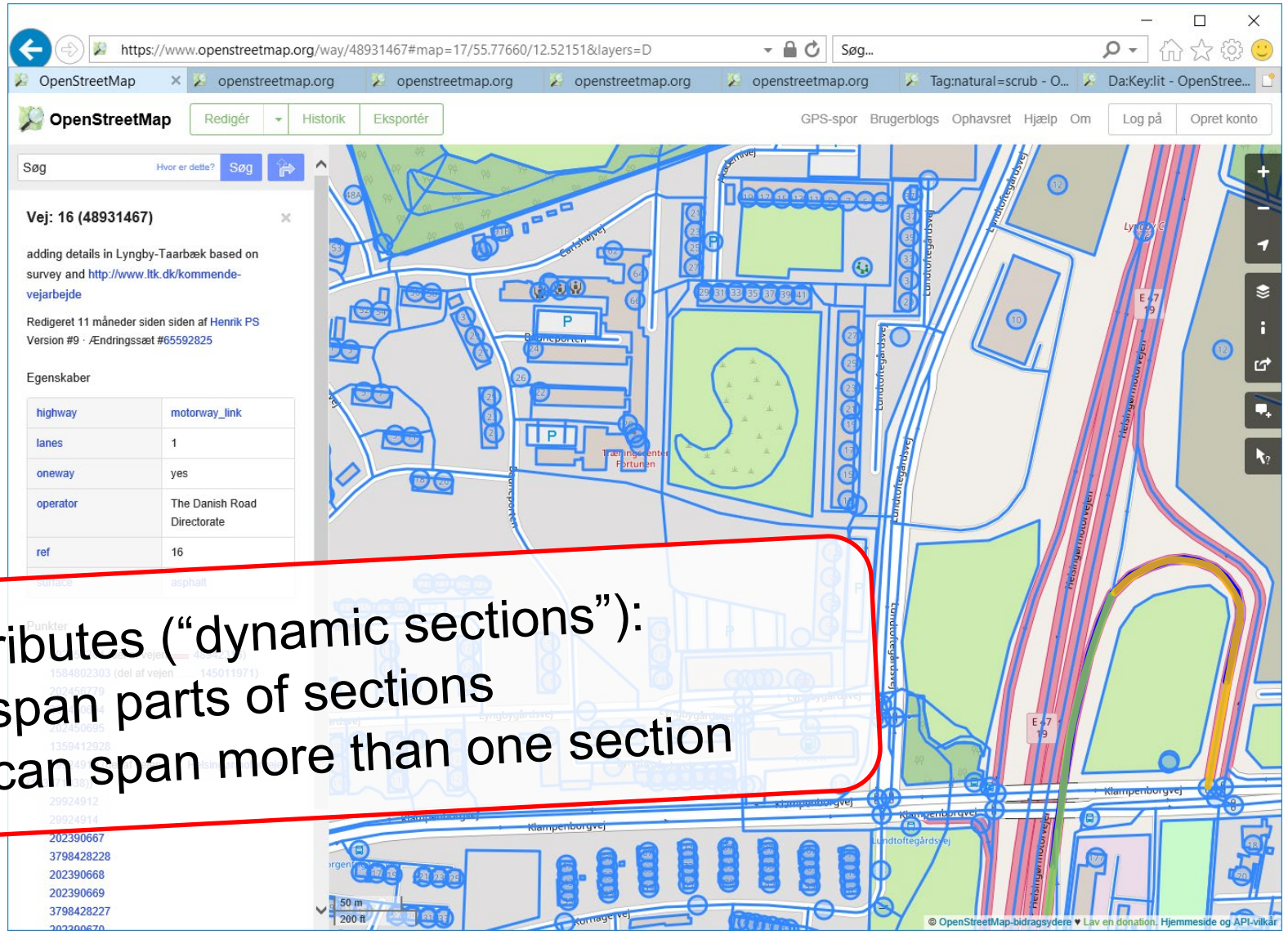
DTU Compute
Department of Applied Mathematics and Computer Science
Ekkart Kindler



Align positions with map data:
→ Map matching

Road state data

DTU Compute
Department of Applied Mathematics and Computer Science
Ekkart Kindler



Attributes (“dynamic sections”):

- span parts of sections
- can span more than one section

Data Collection: Steps

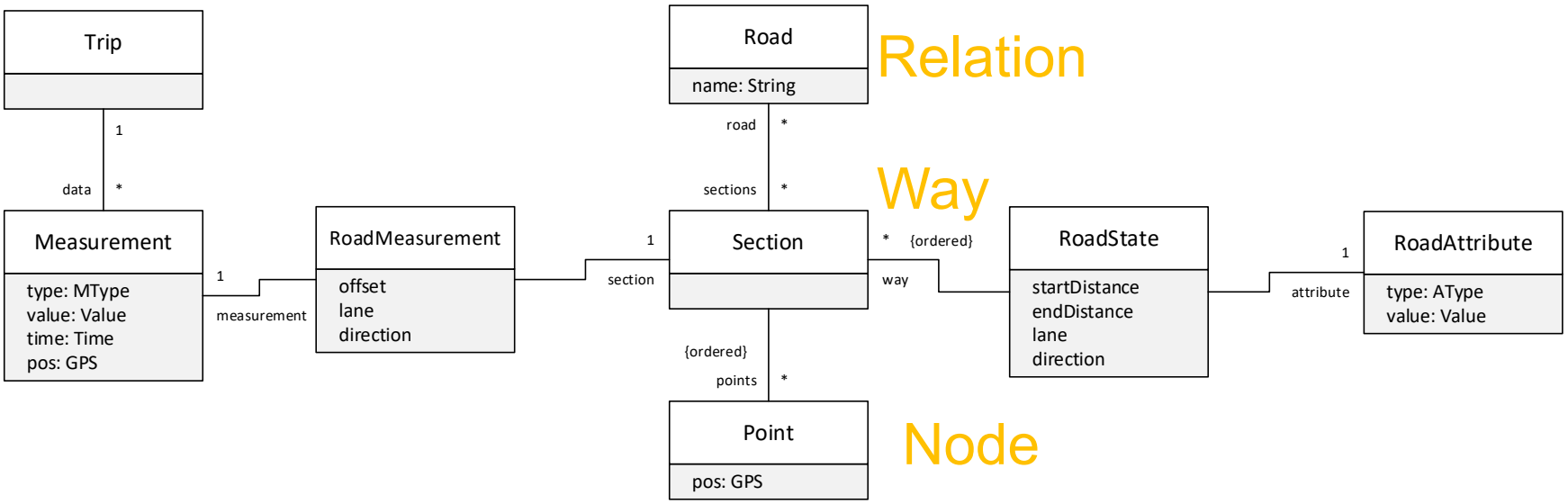
DTU Compute
Department of Applied Mathematics and Computer Science
Ekkart Kindler

- Collect data raw data from cars (per trip)
- Clean data
- Calculate position for data without positions (Interpolation)
- Map positions to sections

Simplified Data Model

DTU Compute
Department of Applied Mathematics and Computer Science
Ekkart Kindler

Car Data (from GM) Static road data (OSM, Sweco, …) Dynamic road data …



Outline

DTU Compute
Department of Applied Mathematics and Computer Science
Ekkart Kindler

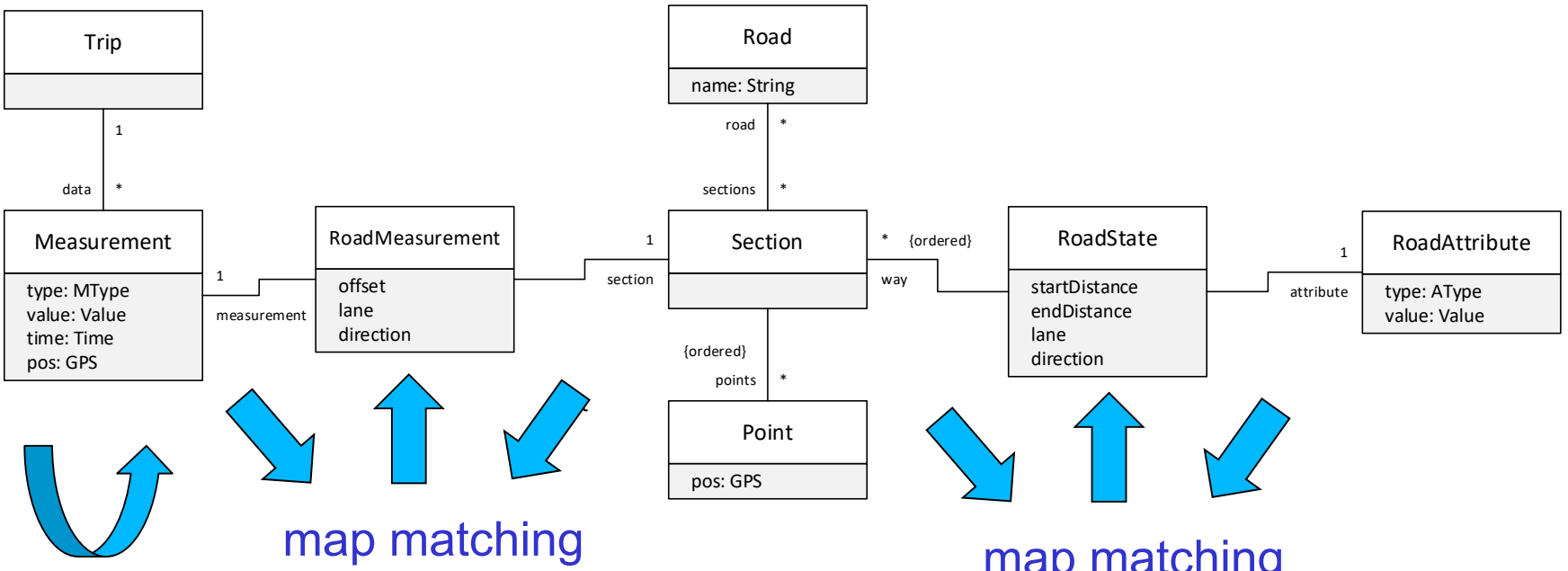
- Introduction and overview
- Concepts (domain analysis)
- Design and architecture
- Detailed models
- Data processing pipelines
- Demo

Design

DTU Compute
Department of Applied Mathematics and Computer Science
Ekkart Kindler

data collection (GM)

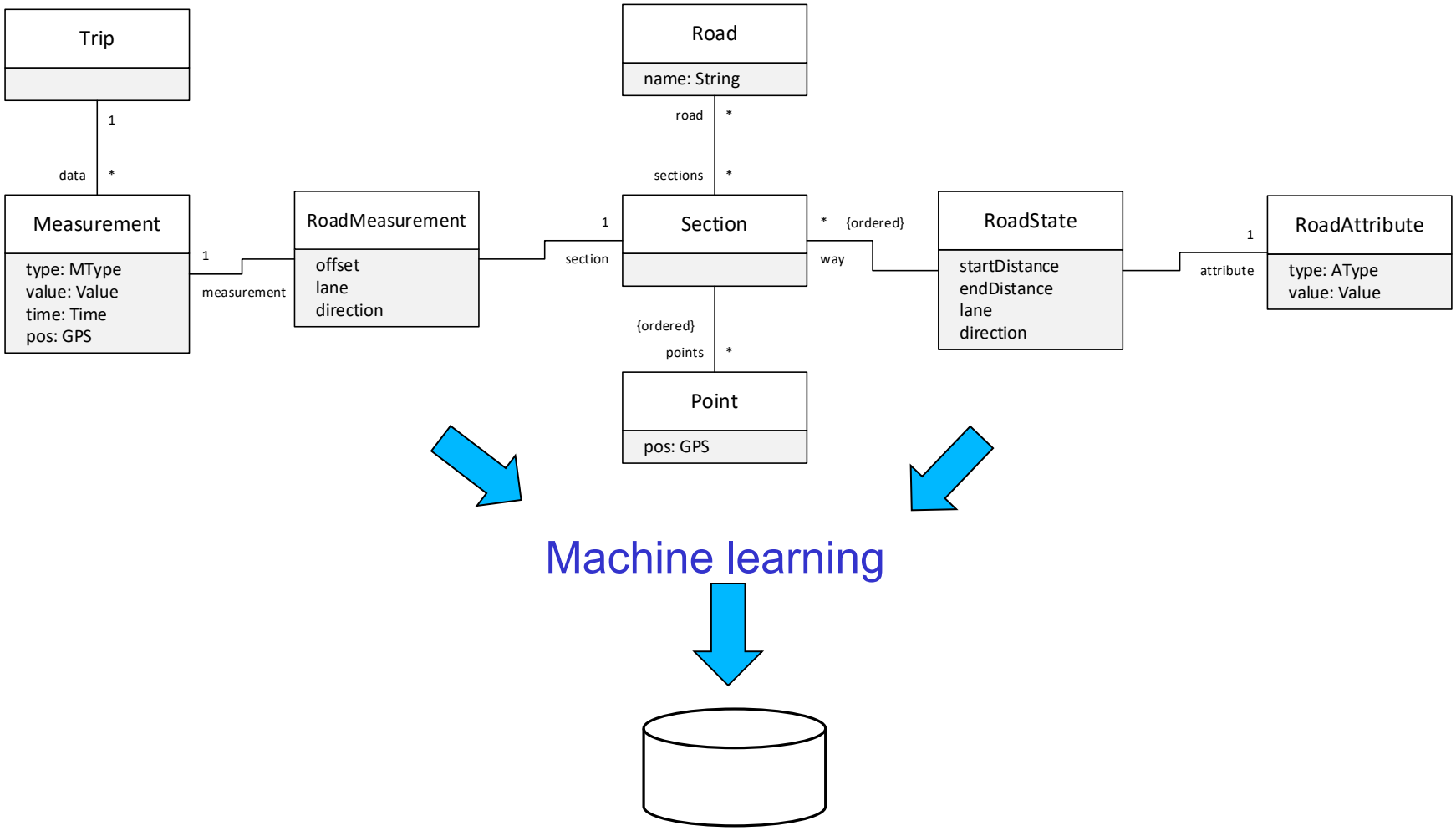
Initial data (VD)



- data cleaning
- GPS interpolation

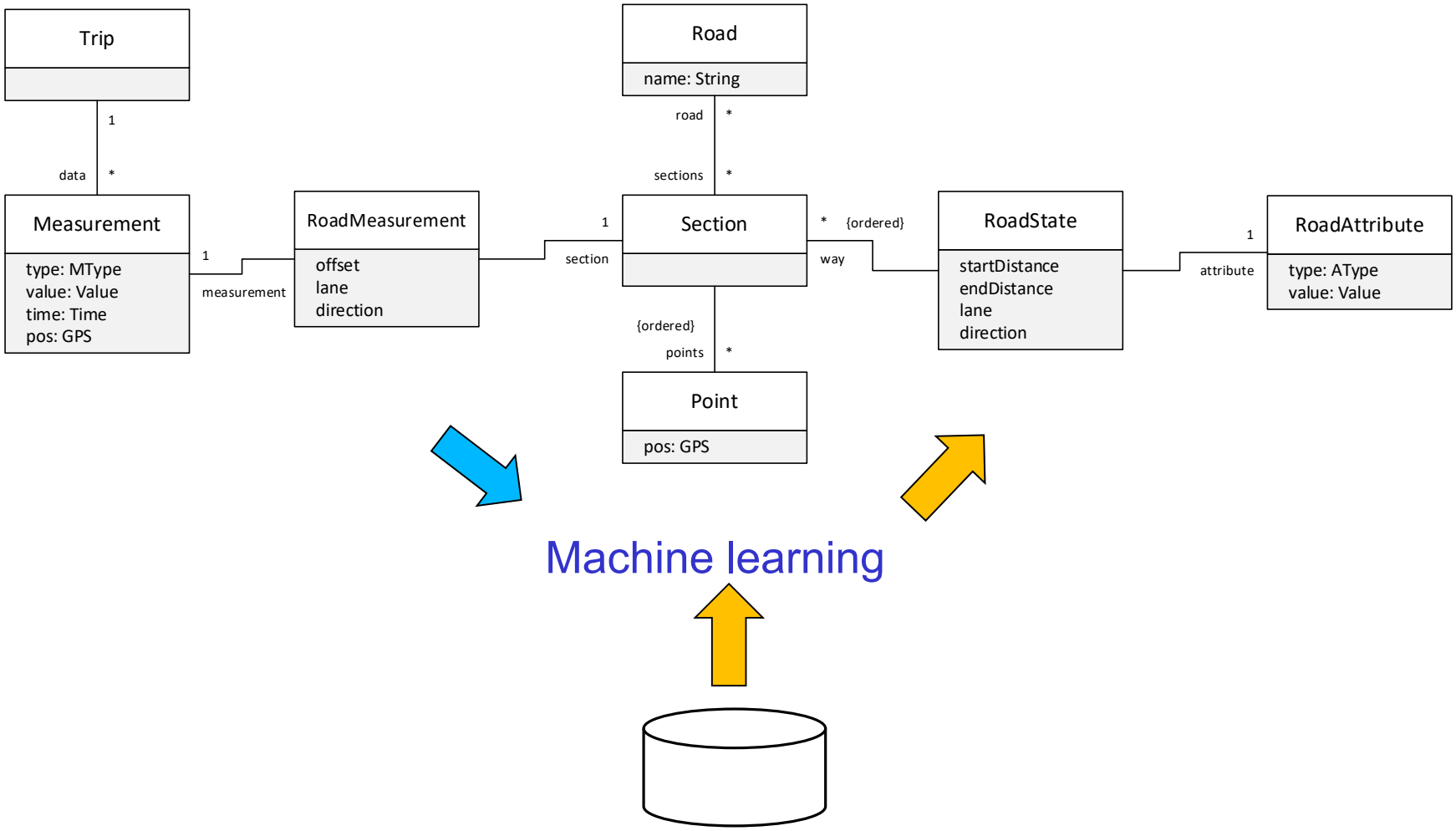
Design

DTU Compute
Department of Applied Mathematics and Computer Science
Ekkart Kindler



Design

DTU Compute
Department of Applied Mathematics and Computer Science
Ekkart Kindler



Status and Plan

DTU Compute
Department of Applied Mathematics and Computer Science
Ekkart Kindler

- Dec. 2019:
  Initial version on Working paper on
  LiRA Concepts, Design and Architecture of Data Warehouse

- January 2020:
  First internal implementation of minimal version of Data Warehouse (data collection and preprocessing)
- Requirements (Validation and Production): First technology experiments and report (MSc Project January 2020)

IoT Design methodology [1]

DTU Compute
Department of Applied Mathematics and Computer Science
Ekkart Kindler

Device:

AutoPi

- Device
  - token Id
- Trip
  - Start and end position
  - Start and end time
  - Device Id
  - List of events
  - Tag (Personal/Business)

Physical entity:

Vehicle

- Event types
  - track position (GPS)
  - accelerometer
  - device events
  - obd
  - rpi
- Event
  - Object Id
  - Timestamp
  - Type
  - Attributes (flat/structured)

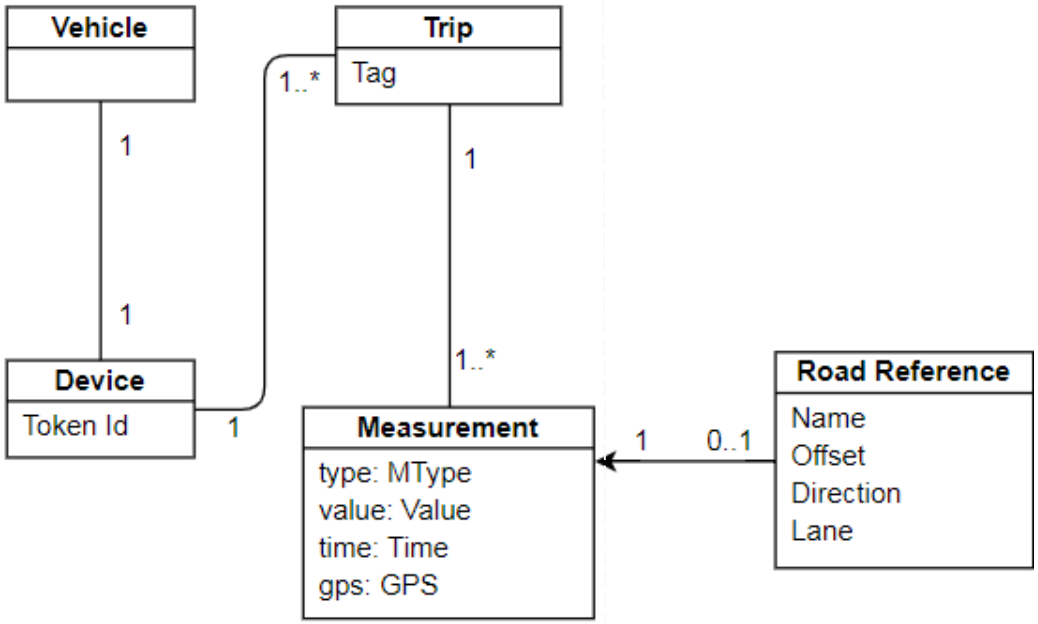
Virtual Entity:

Trip

1. (Introduced by Bahga & Madisetti, 2015):

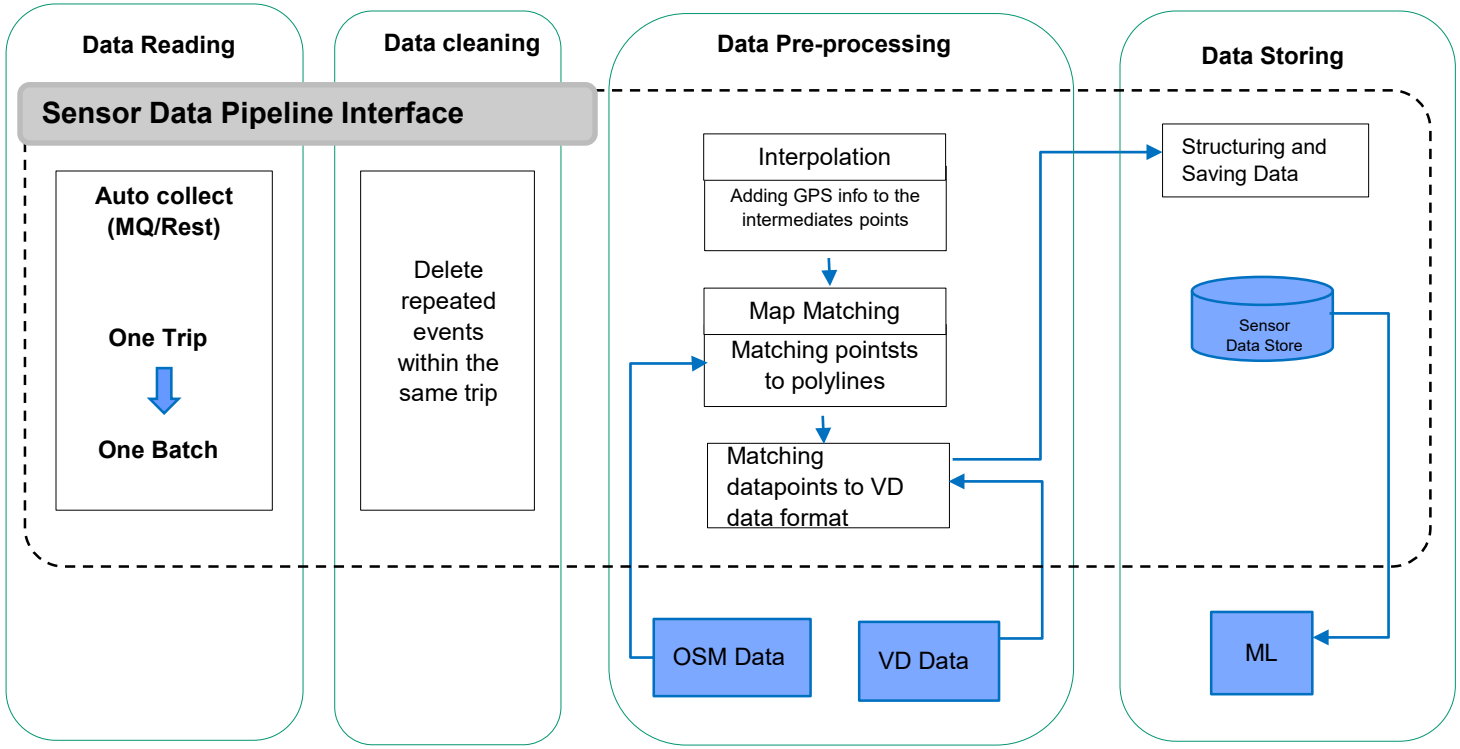
Raw data Domain Model

DTU Compute
Department of Applied Mathematics and Computer Science
Ekkart Kindler



Raw data Pipeline Interface

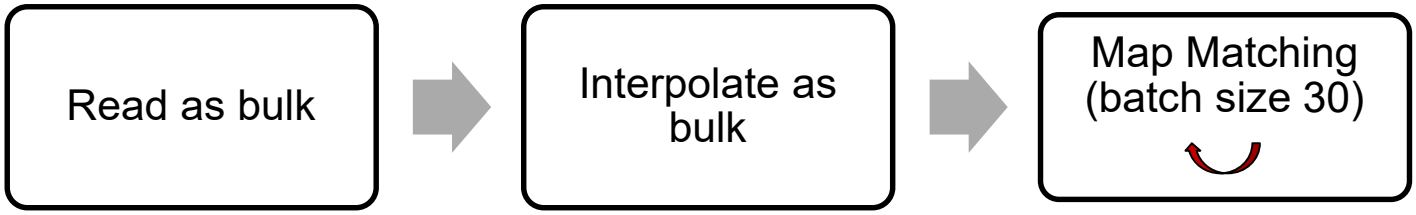
DTU Compute
Department of Applied Mathematics and Computer Science
Ekkart Kindler



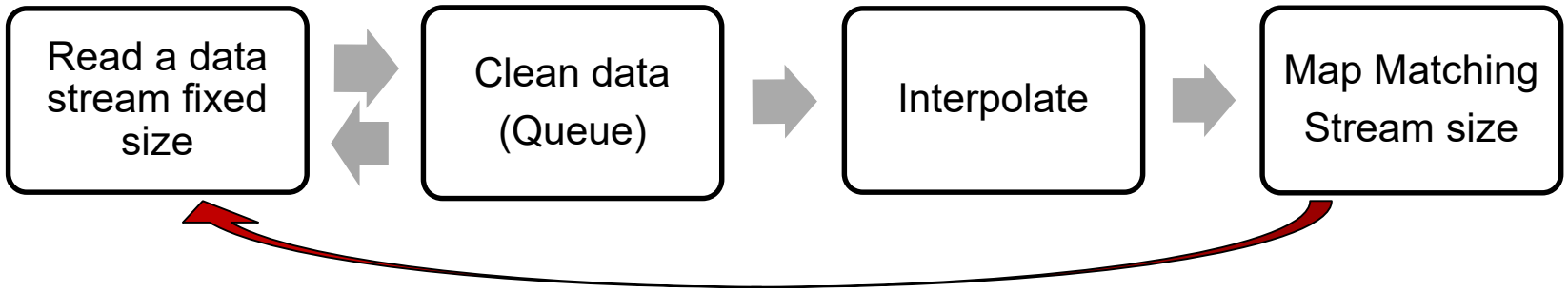
Data Reading

DTU Compute
Department of Applied Mathematics and Computer Science
Ekkart Kindler

- Batch Streaming
  - JSON input file



- Data Streaming
  - *GreenMobility* Server
  - Read by stream size of fixed number



Data Cleaning

DTU Compute
Department of Applied Mathematics and Computer Science
Ekkart Kindler

- Removing the repeated track position events
- Remove events which require extrapolation
- Can be enriched to outlier removal

Pre-Processing: Interpolation

DTU Compute
Department of Applied Mathematics and Computer Science
Ekkart Kindler

- each non "track-position" event at a given timestamp needs to be interpolated as a GPS location between other GPS locations obtained by "track-position" events

- Development of three types:
  - Linear
  - Quadratic
  - Cubic splines

$$(x_0, y_0), \ldots, (x_j, y_j), \ldots, (x_k, y_k)$$

$$L(x) := \sum_{j=0}^{k} y_j \ell_j(x)$$

$$\ell_j(x) := \prod_{\substack{0 \le m \le k \\ m \ne j}} \frac{x - x_m}{x_j - x_m} = \frac{(x - x_0)}{(x_j - x_0)} \cdots \frac{(x - x_{j-1})}{(x_j - x_{j-1})} \frac{(x - x_{j+1})}{(x_j - x_{j+1})} \cdots \frac{(x - x_k)}{(x_j - x_k)},$$

Pre-Processing: Map Matching

DTU Compute
Department of Applied Mathematics and Computer Science
Ekkart Kindler

- Map-matching is the process of aligning a sequence of observed user positions with the **road** network on a digital map.

- Hidden Markov Model Map matching

  (Paul Newson and John Krumm, 2009)

- Map Matching Service from OSRM library
  - from Open Source Routing Machine(OSRM)
  - It matches/snaps given GPS points to the road network

- Map Matching to **Segments**

haven't been fulfilled yet:

- Require the annotator function from

OSRM library

Data Storing

DTU Compute
Department of Applied Mathematics and Computer Science
Ekkart Kindler

- Postgresql database management system

Id
objectId
t
token
ts
tag
lat
lon
message
IsComputed
MapMatched_lat
MapMatched_lon
MapMatched_wayPointName

Road States

DTU Compute
Department of Applied Mathematics and Computer Science
Ekkart Kindler

- A Geographic Information System (GIS) :

  facilitating the end user's maintenance planning

  - Domain analysis of Rosy and Vejman
    - Rosy:
      - A pavement management system by Sweco
      - Various layers of information (defect, general traffic)
      - On top of OSM per section
    - Vejman:
      - A pavement management system
      - (general knowledge and road data, operational data)
      - General map (OSM or other topological map) and a lighter version of map
      - Condition Index

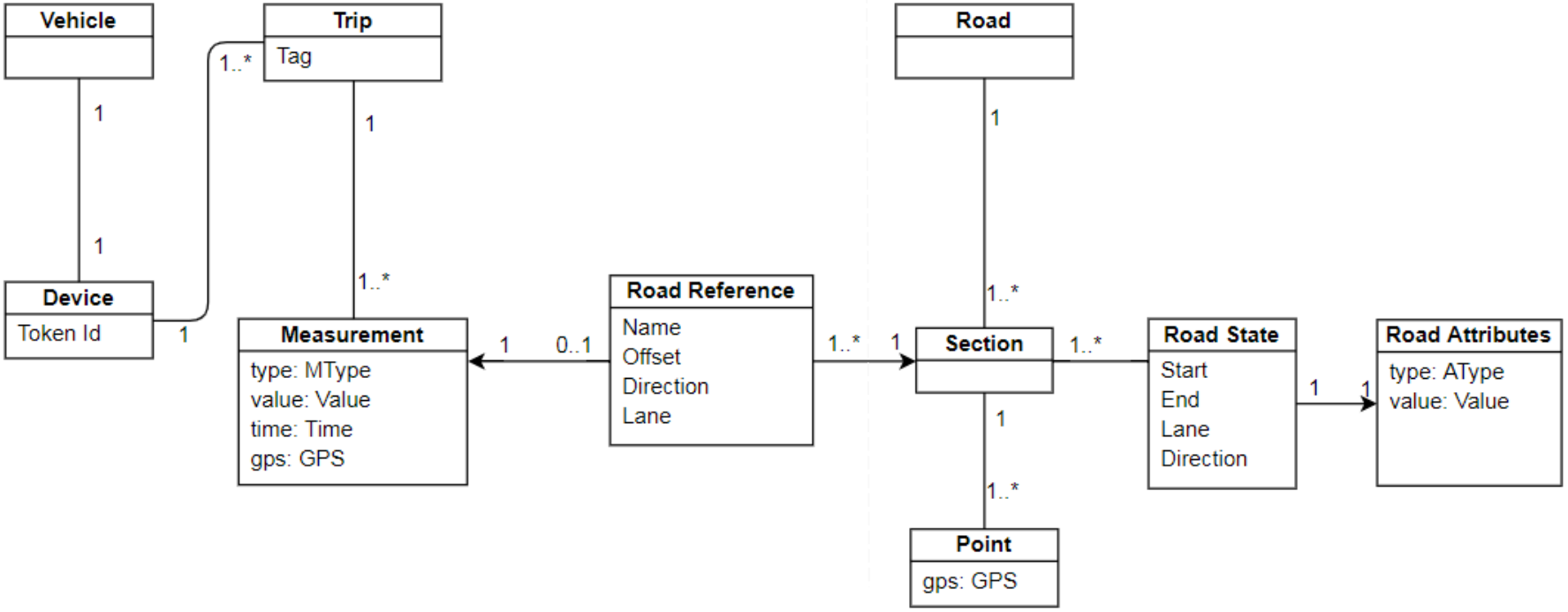
- Chosen attributes from Rosy and Vejman:
  - Sectioning
  - Open Street Map (OSM)
  - Condition Index

Domain Model

DTU Compute
Department of Applied Mathematics and Computer Science
Ekkart Kindler

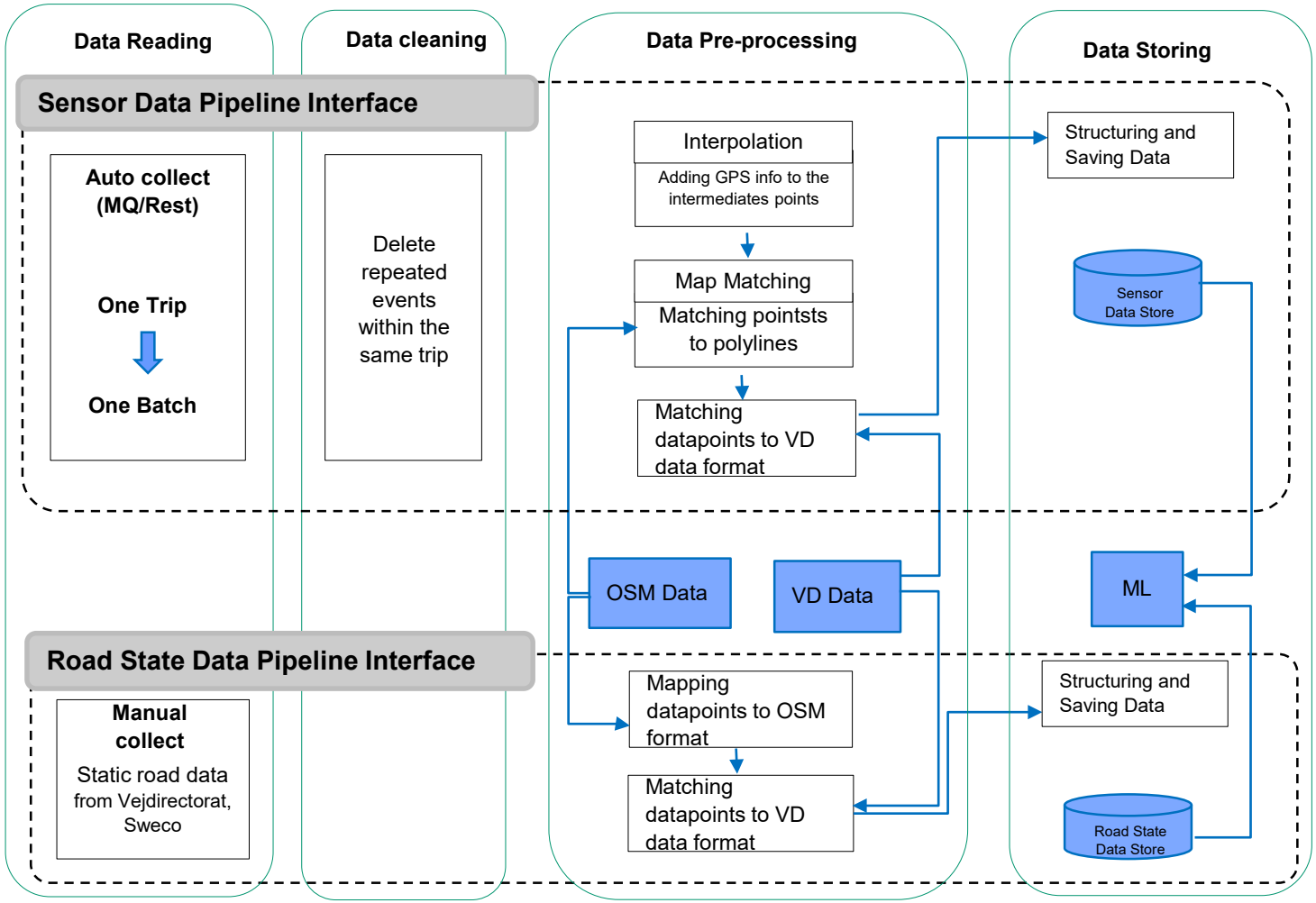
Raw data

Road state



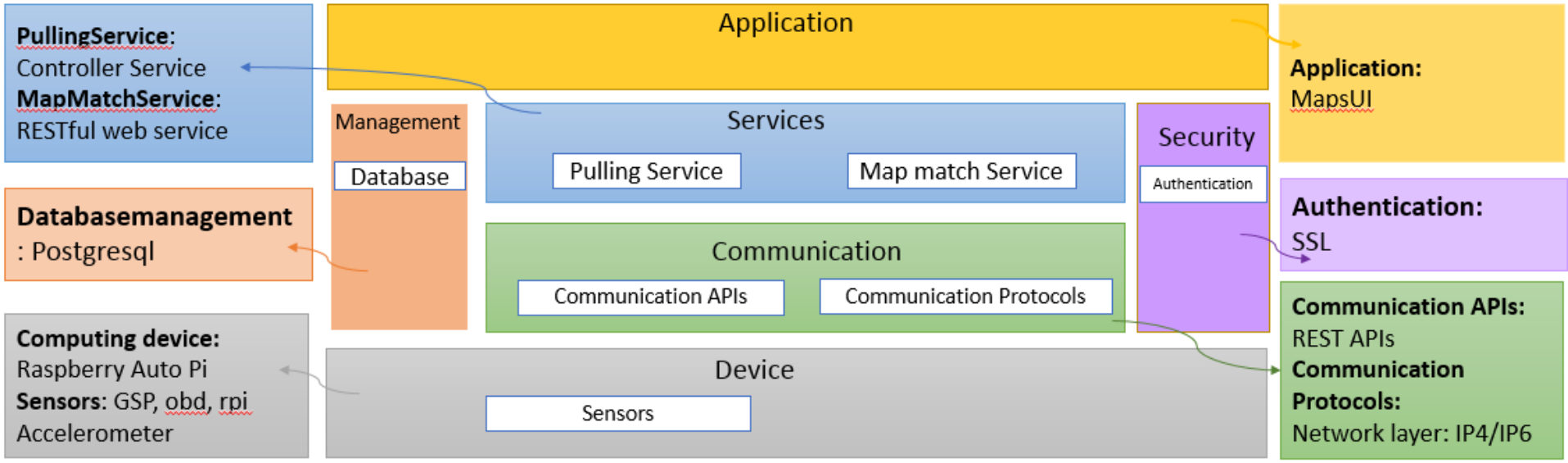
Data Pipeline Interface:

DTU Compute
Department of Applied Mathematics and Computer Science
Ekkart Kindler



Operational View Specification

DTU Compute
Department of Applied Mathematics and Computer Science
Ekkart Kindler



Operational View Specification in LiRA project inspired by the figure of Operational View Specification represented by (Bahga & Madisetti, 2015)

future works

DTU Compute
Department of Applied Mathematics and Computer Science
Ekkart Kindler