

# MOVEMENT DATA IN GIS

Geobeer Switzerland 2021-03-18

Anita Graser



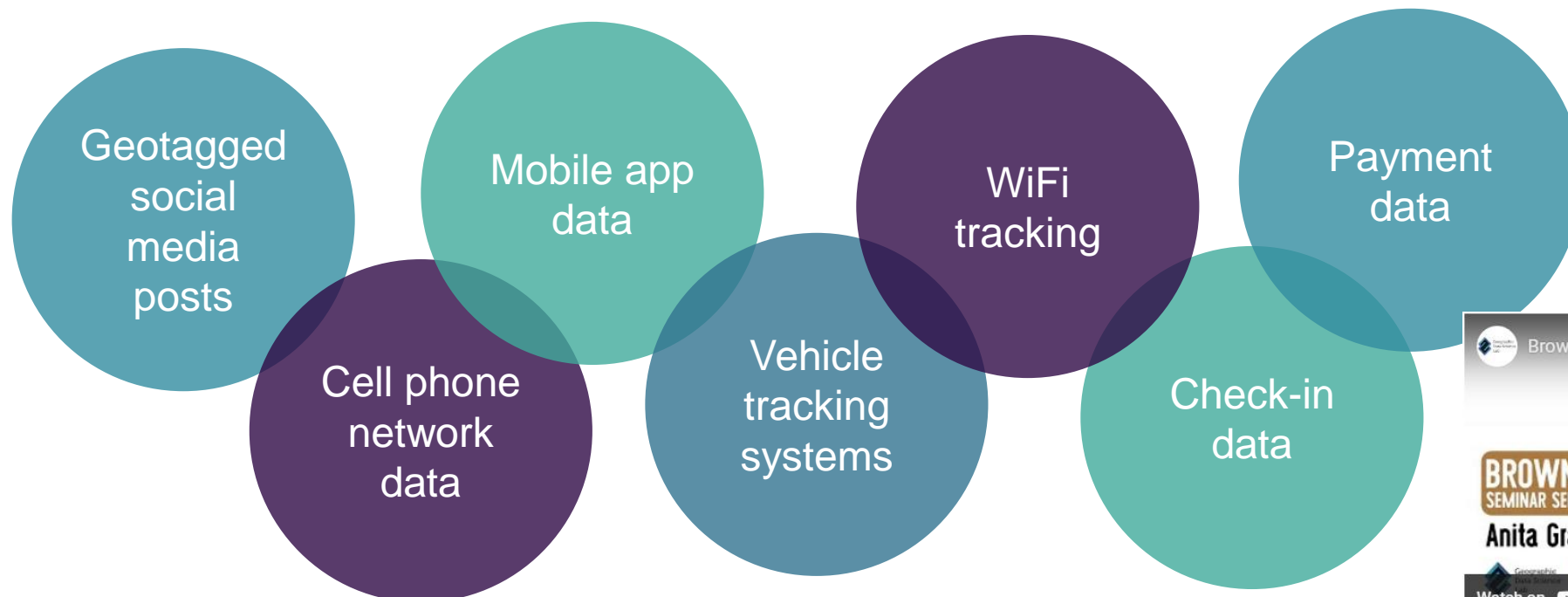
# MOVEMENT DATA SCIENCE

✓ **Opportunistic reuse** of data

→ Black box / undocumented data collection

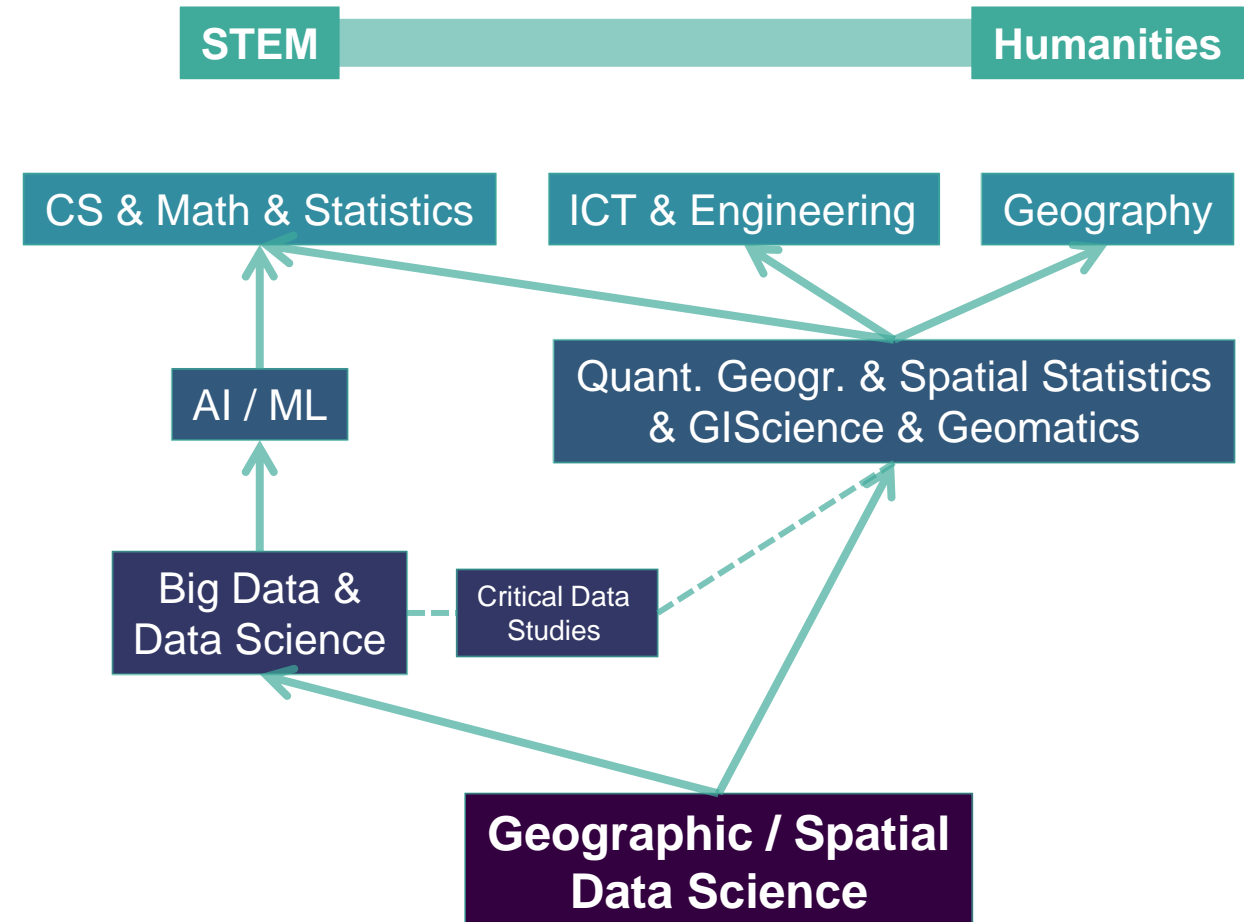
→ Usually biased & messy data

*“All metadata records are incomplete as it is impossible to foresee future uses”* Janowicz et al. (2020) GeoAI



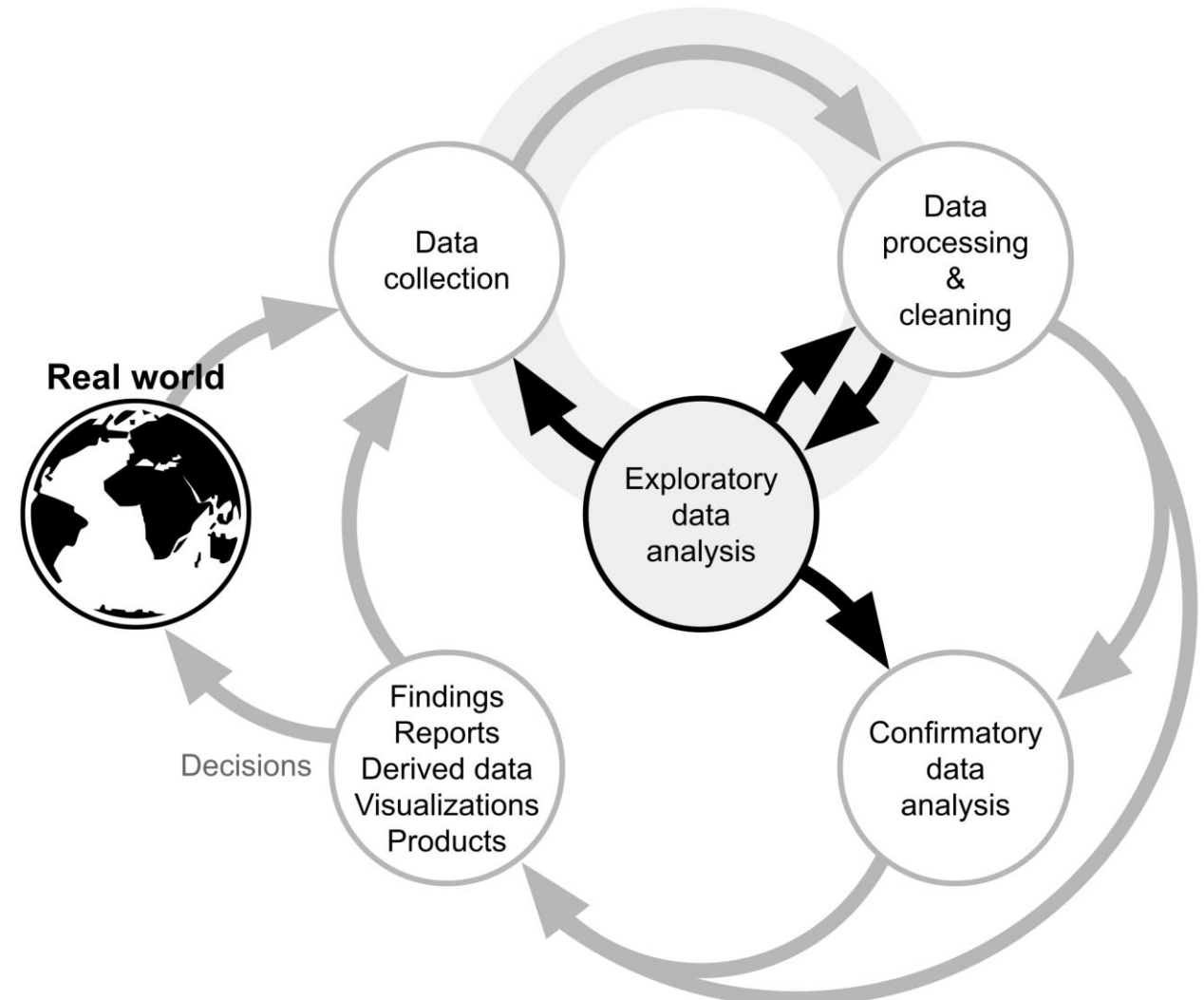
# GEOGRAPHIC DATA SCIENCE

- ✓ **Gathering** data, **massaging** it into a tractable form, making it **tell its story**, and **presenting** that story to others
- ✓ Dealing with data that incorporates **spatial and often temporal elements**
- ✓ Turning **Big Spatiotemporal Data** into **insight and understanding**



# CHALLENGES

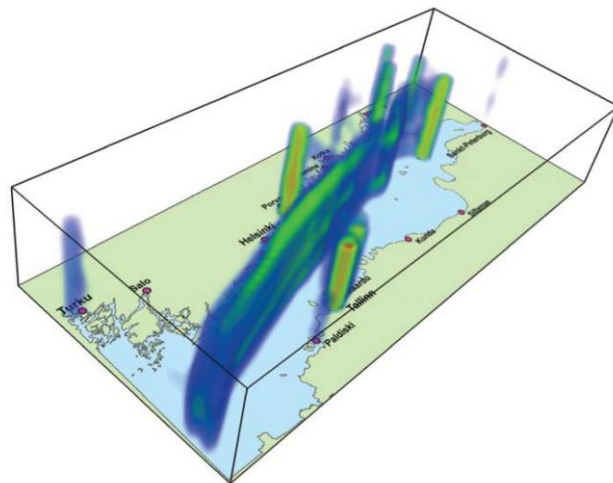
- ✓ Accept **messiness** in data
- ✓ Need to understand
  - ✓ **Causes** of bias & messiness
  - ✓ **Consequences** of using such data in analyses
- ➔ Data visualization & exploratory approaches



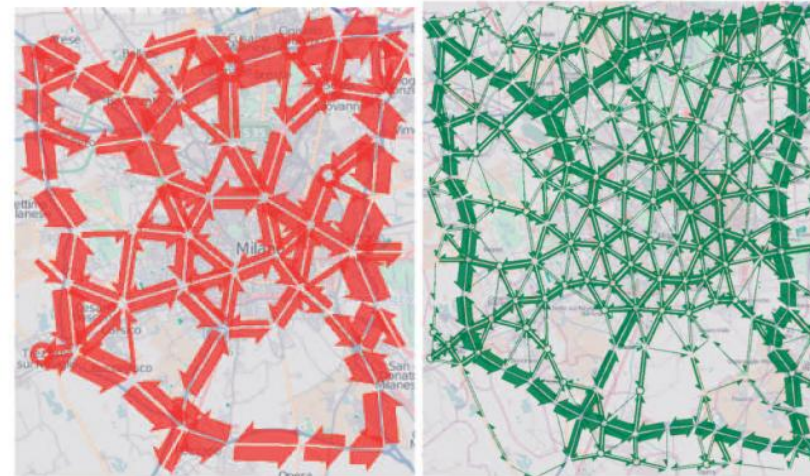
# EXPLORING MOVEMENT DATA

- ✓ Complex spatiotemporal phenomena
- ✓ Context & scale dependent
- ✓ Spatial, temporal & attribute uncertainty

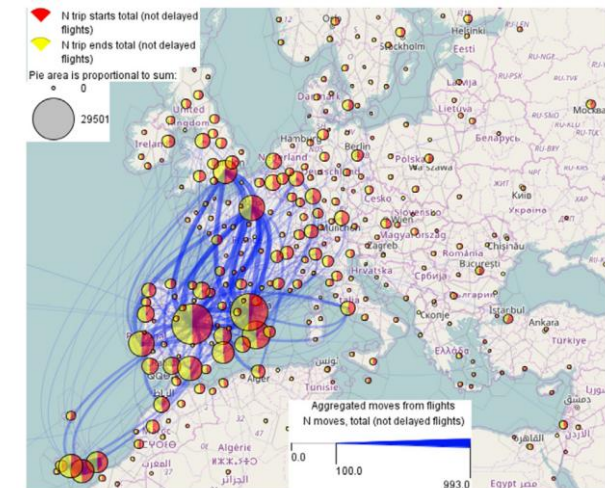
*+ Lack of established tools & practices*



Demšar & Virrantaus (2010)  
Space-time density of trajectories

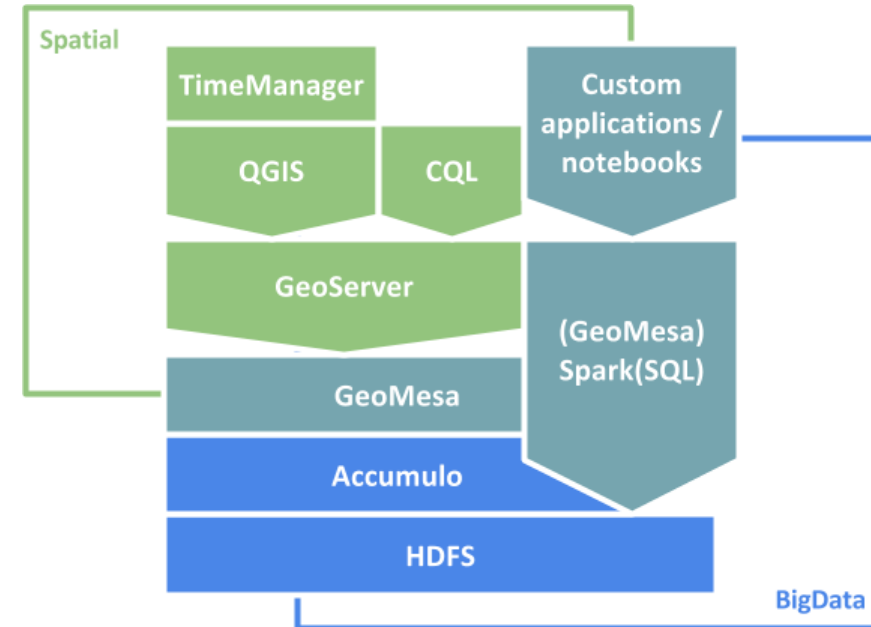
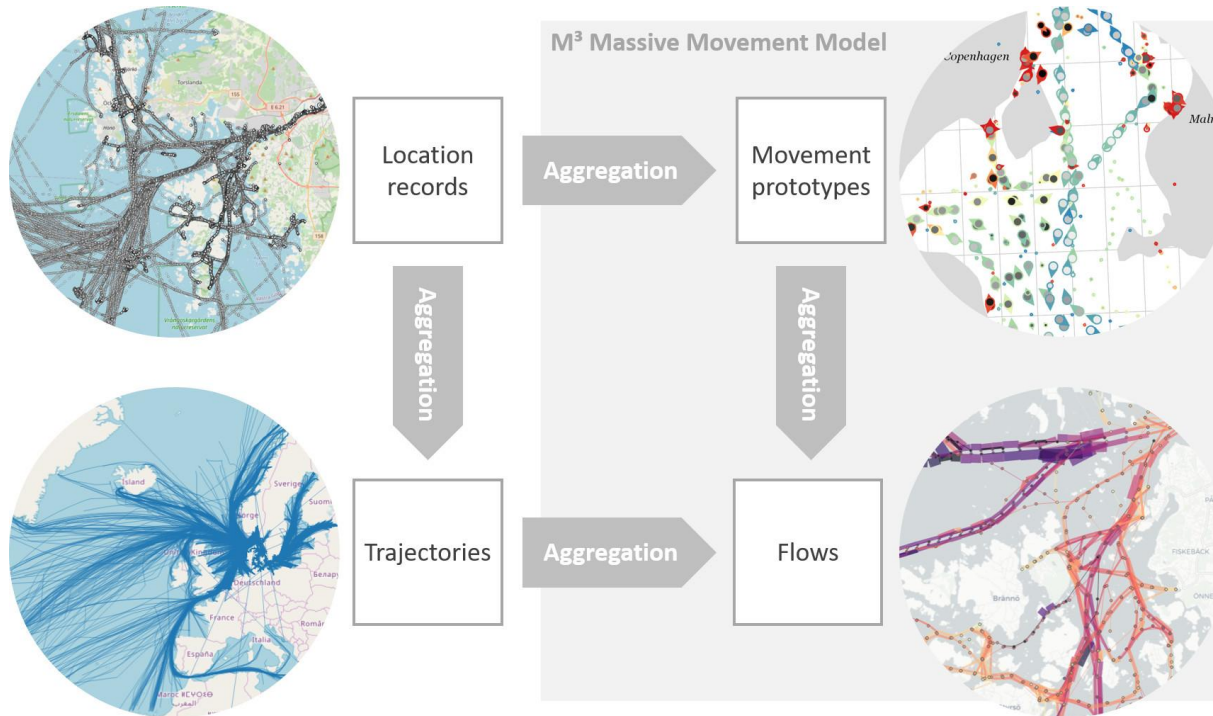


Andrienko & Andrienko (2011)  
Spatial generalization and aggregation of  
massive movement data



Andrienko et al. (2017)  
Visual exploration of movement and  
event data with interactive time masks

# AGGREGATING MOVEMENT DATA



- Graser et al. (2020) The M<sup>3</sup> massive movement model
- Graser et al. (2020) Exploratory Trajectory Analysis for Massive Historical AIS Datasets
- Graser et al. (2020) Extracting Patterns from Large Movement Datasets



<https://www.youtube.com/watch?v=dRE9Zl7jpUA>

# A protocol for identifying problems in continuous movement data

This notebook provides an open-source implementation of the protocol presented in

Graser, A. (2021) An exploratory data analysis protocol for identifying problems in continuous movement data. *Journal of Location Based Services*.  
<http://dx.doi.org/10.1080/17489725.2021.1900612>.

The individual protocol steps are demonstrated using a dataset of vessel tracking data published by the Danish Maritime Authority. The demo data covers two days (July 1st and 2nd, 2018). Since the datasets are too large for Github, they have been made available via Figshare: <https://doi.org/10.6084/m9.figshare.11577543>

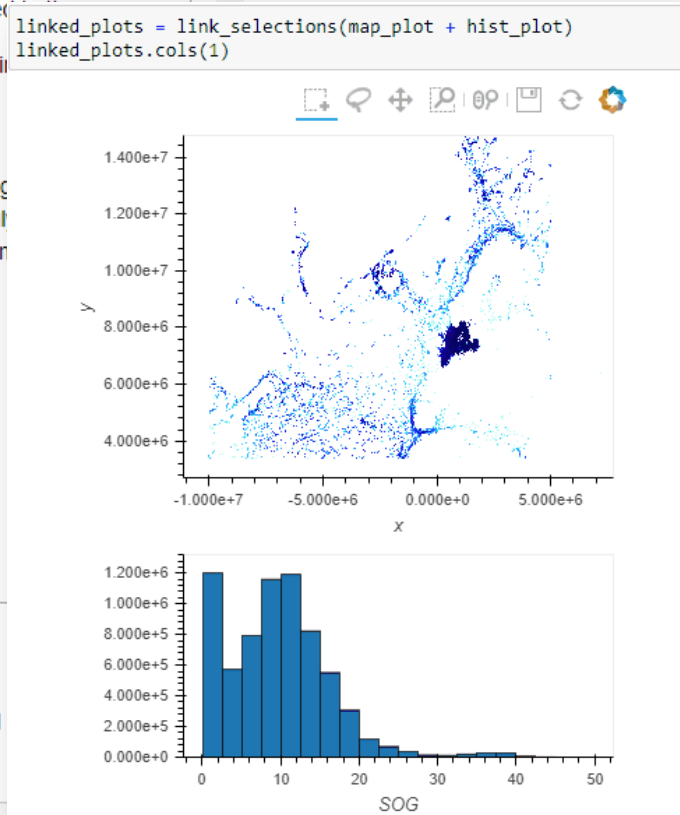
## Content

- [Setup](#)
- [A. Missing data](#)
- [B. Precision problems](#)
- [C. Consistency problems](#)
- [D. Accuracy problems](#)

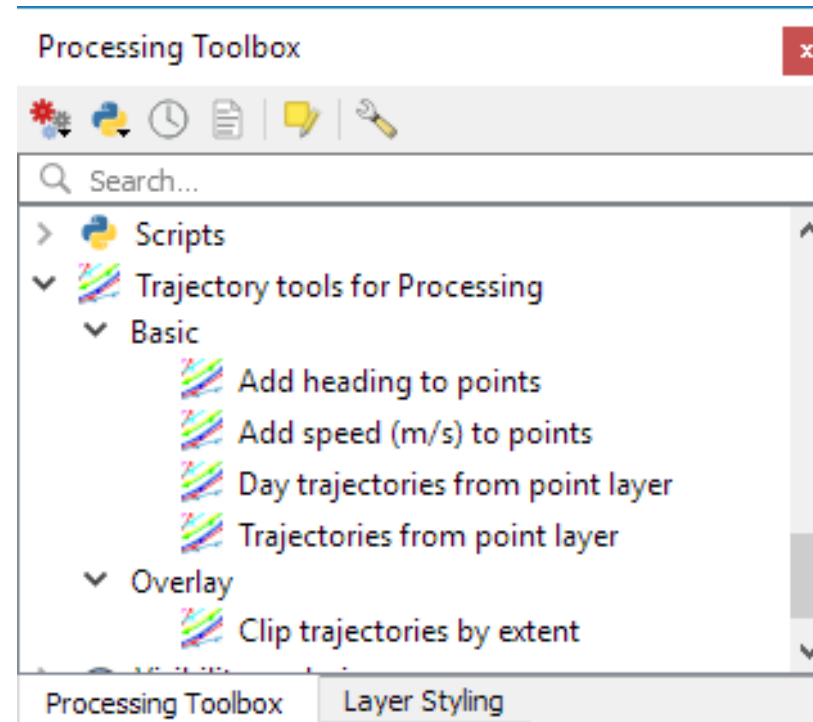
## Setup

Before running this notebook, make sure to [download](#) `dk_csv_20170701.7z` and `dk_csv_20180101.7z` and unzip the files into the data directory:

```
In [1]: input_files = [
        './data/aisdk_20170701.csv',
        './data/aisdk_20180101.csv'
    ]
```



# QGIS ♥ DATA SCIENCE





[anitagraser.com/movement-data-in-gis](http://anitagraser.com/movement-data-in-gis)

✉ [anita.graser@ait.ac.at](mailto:anita.graser@ait.ac.at)

🐦 [@underdarkGIS](https://twitter.com/underdarkGIS)

