

# Bicycle networks: From data quality assessment to computational analysis

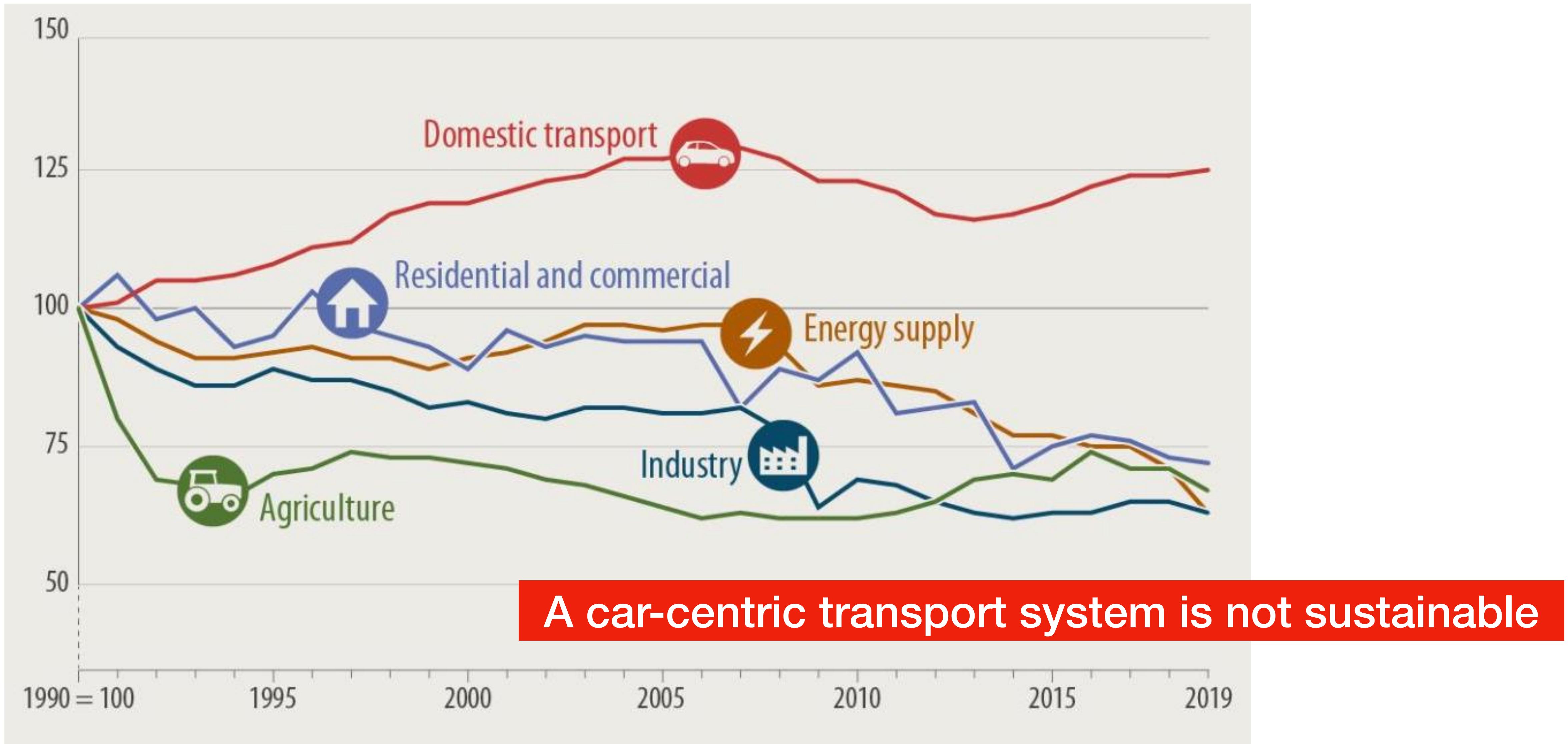
Michael Szell  
Ane Rahbek Vierø

NERDS (NEtwoRks, Data, and Society)  
Computer Science Department

IT UNIVERSITY OF COPENHAGEN



# Transport plays a key role in the climate crisis



# You can't beat geometry: Cars will always be inefficient

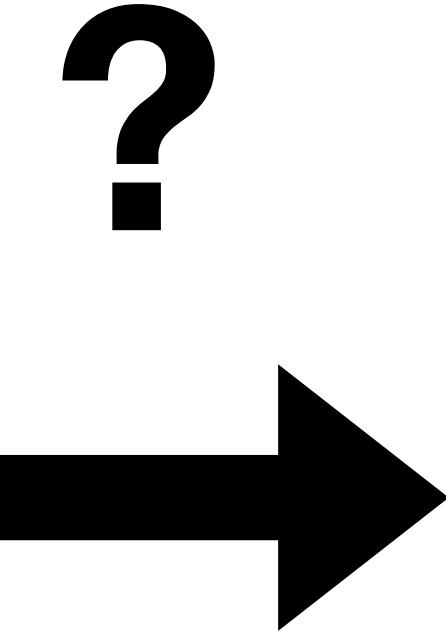


Harms and Kansen, Netherlands Institute for Transport Policy Analysis (2017)  
Szell, Urb Plan 3, 1-20 (2018)  
Euclid (300 BC)

# Data-driven planning can support a sustainability shift



# We need data to plan more sustainable cities

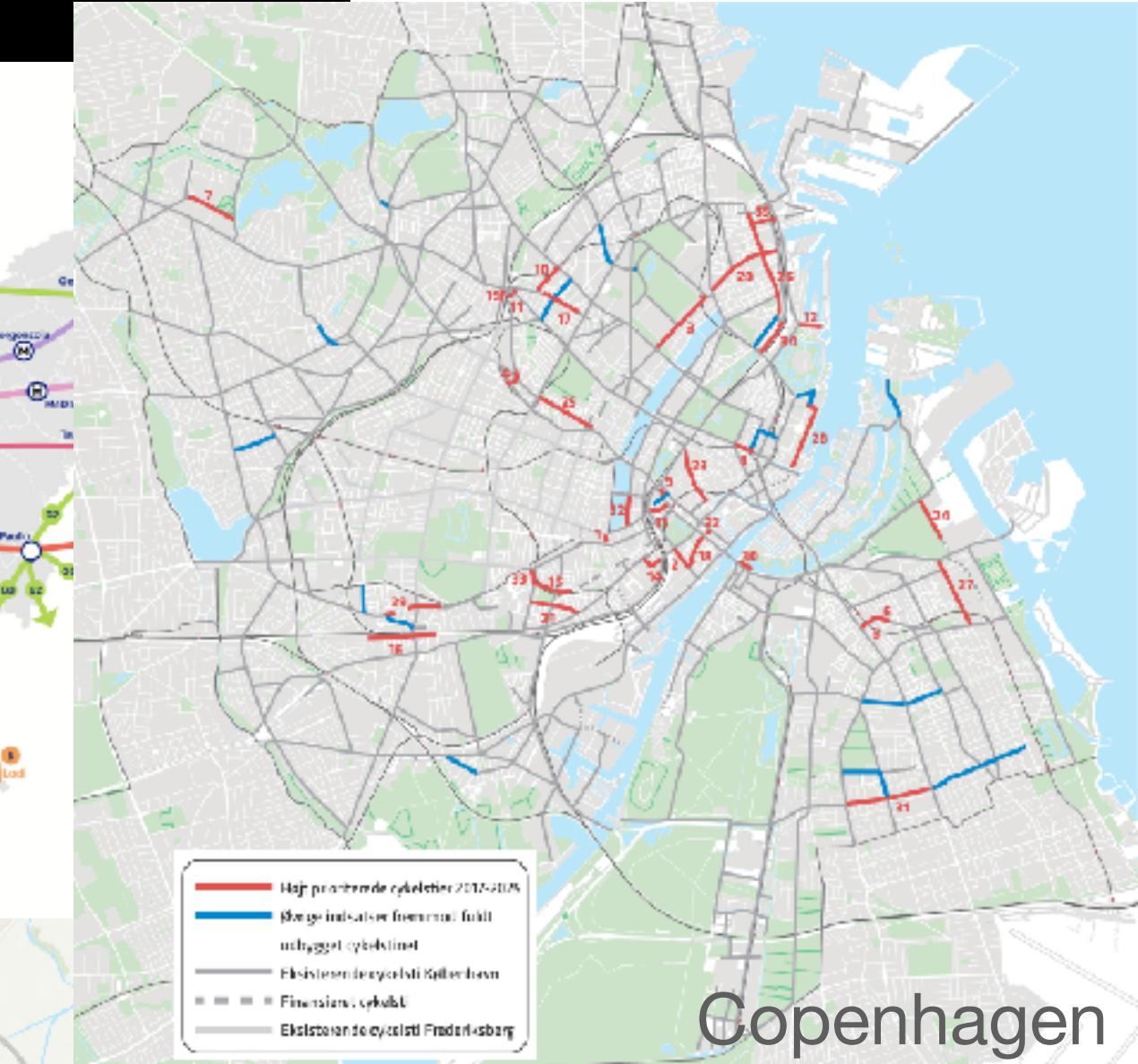
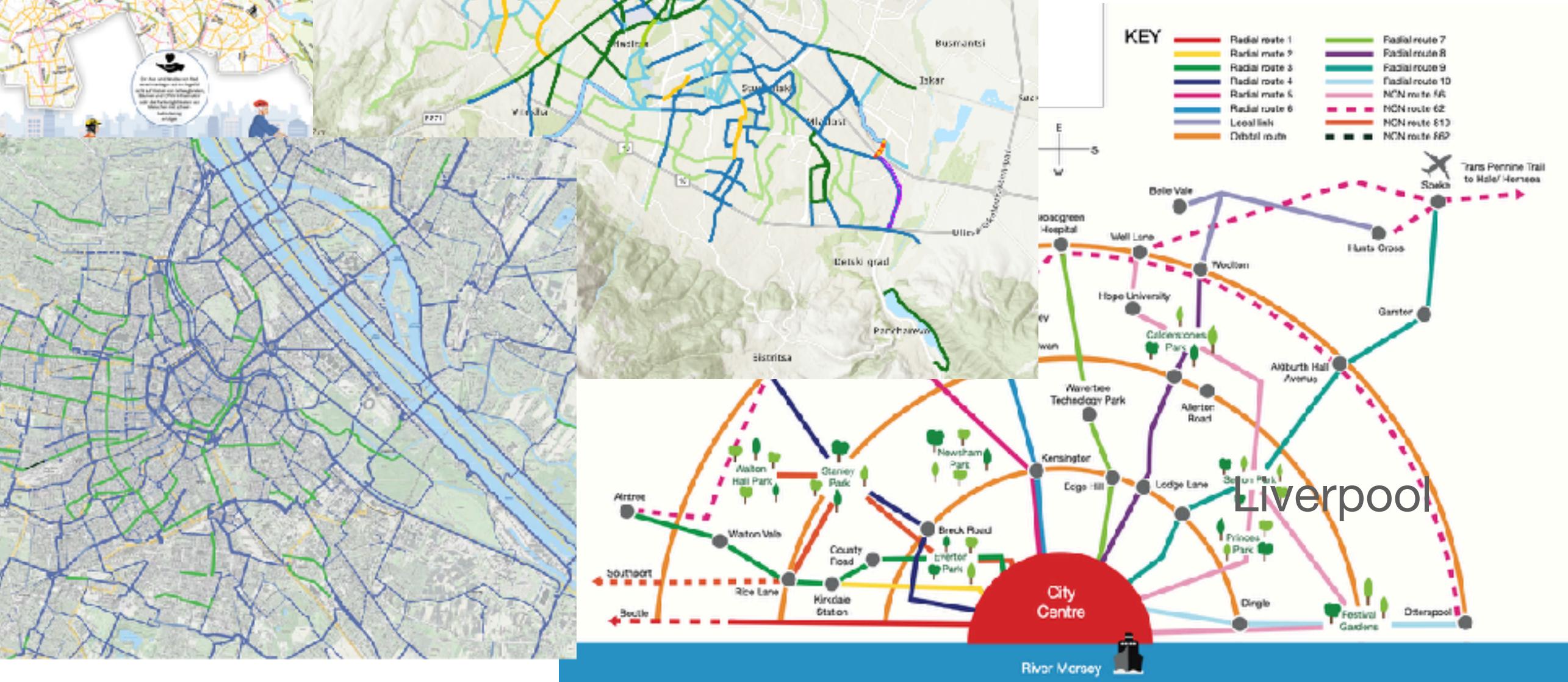
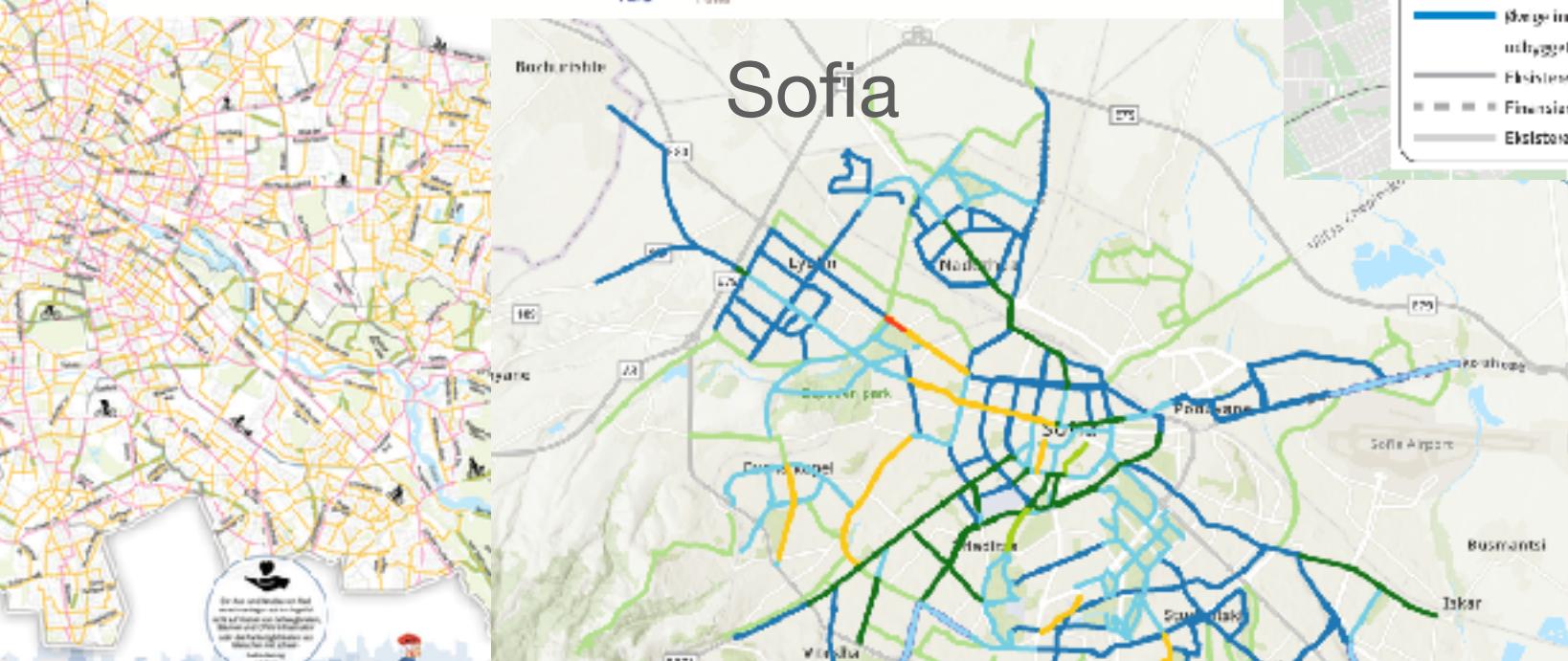
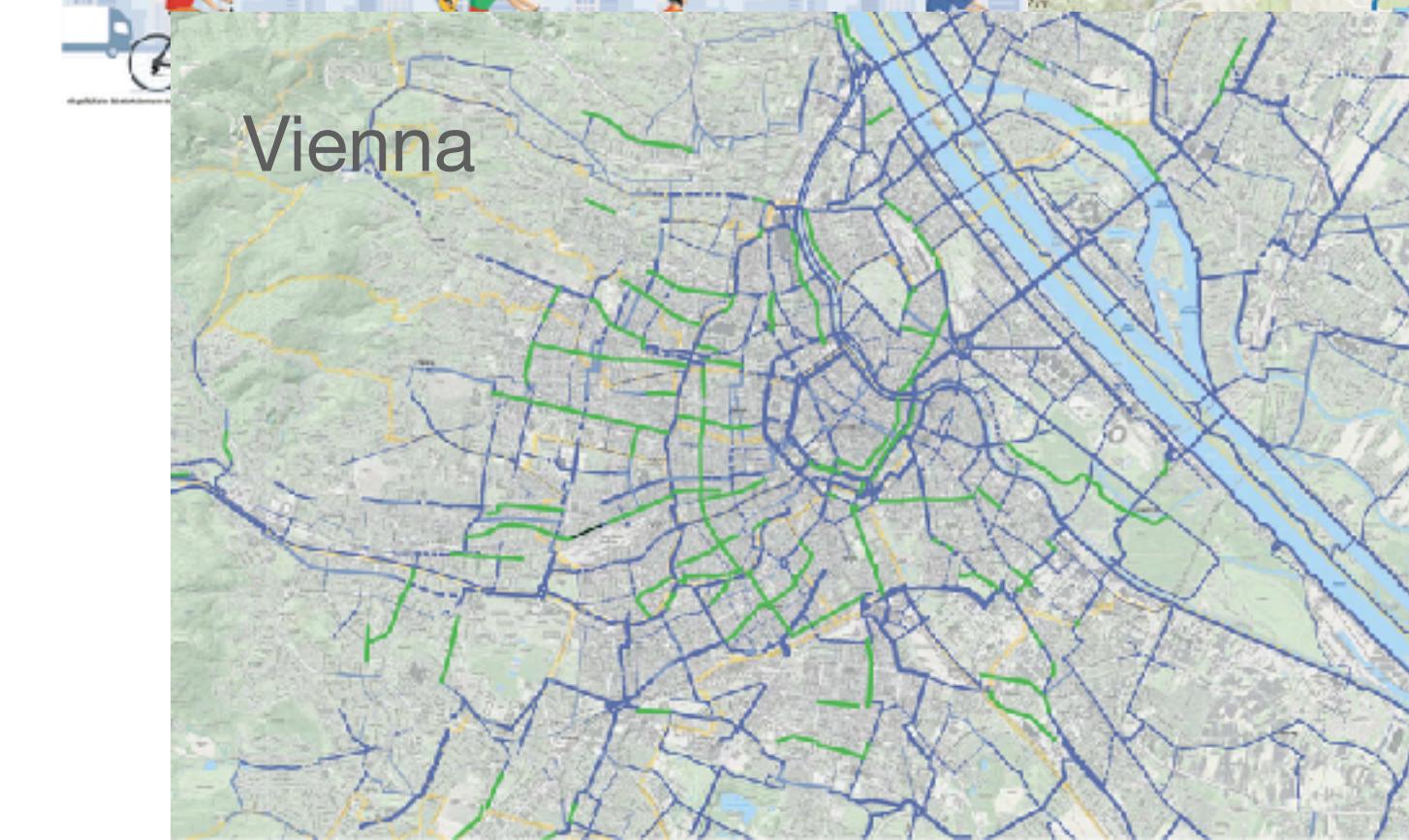
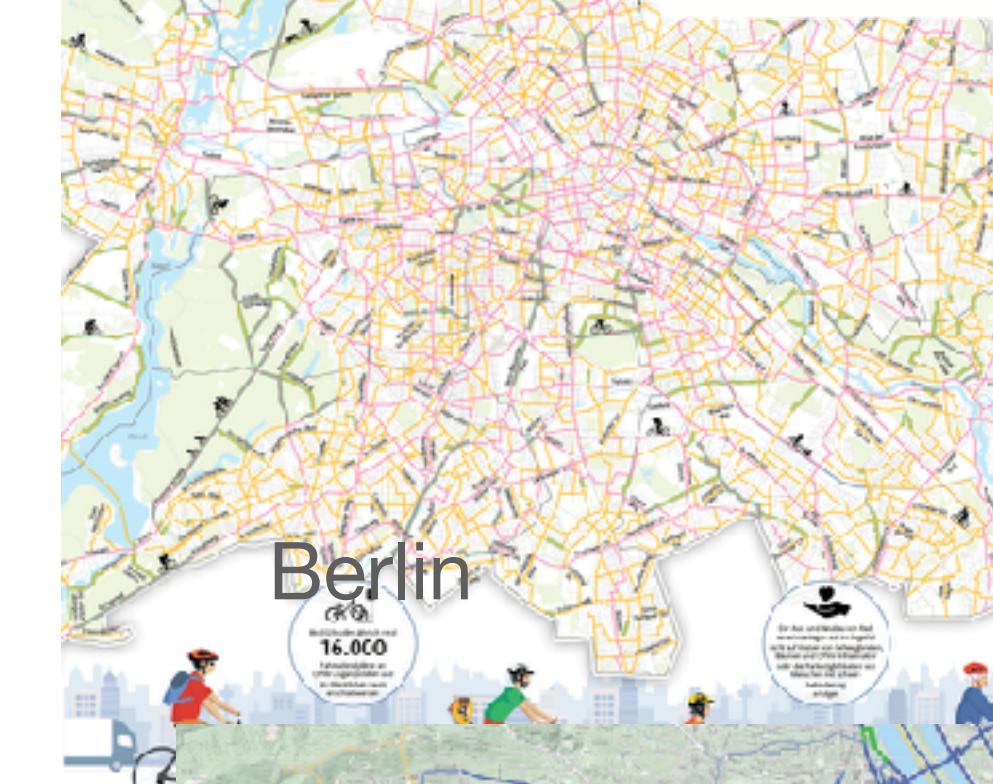


## Radverkehrsnetz

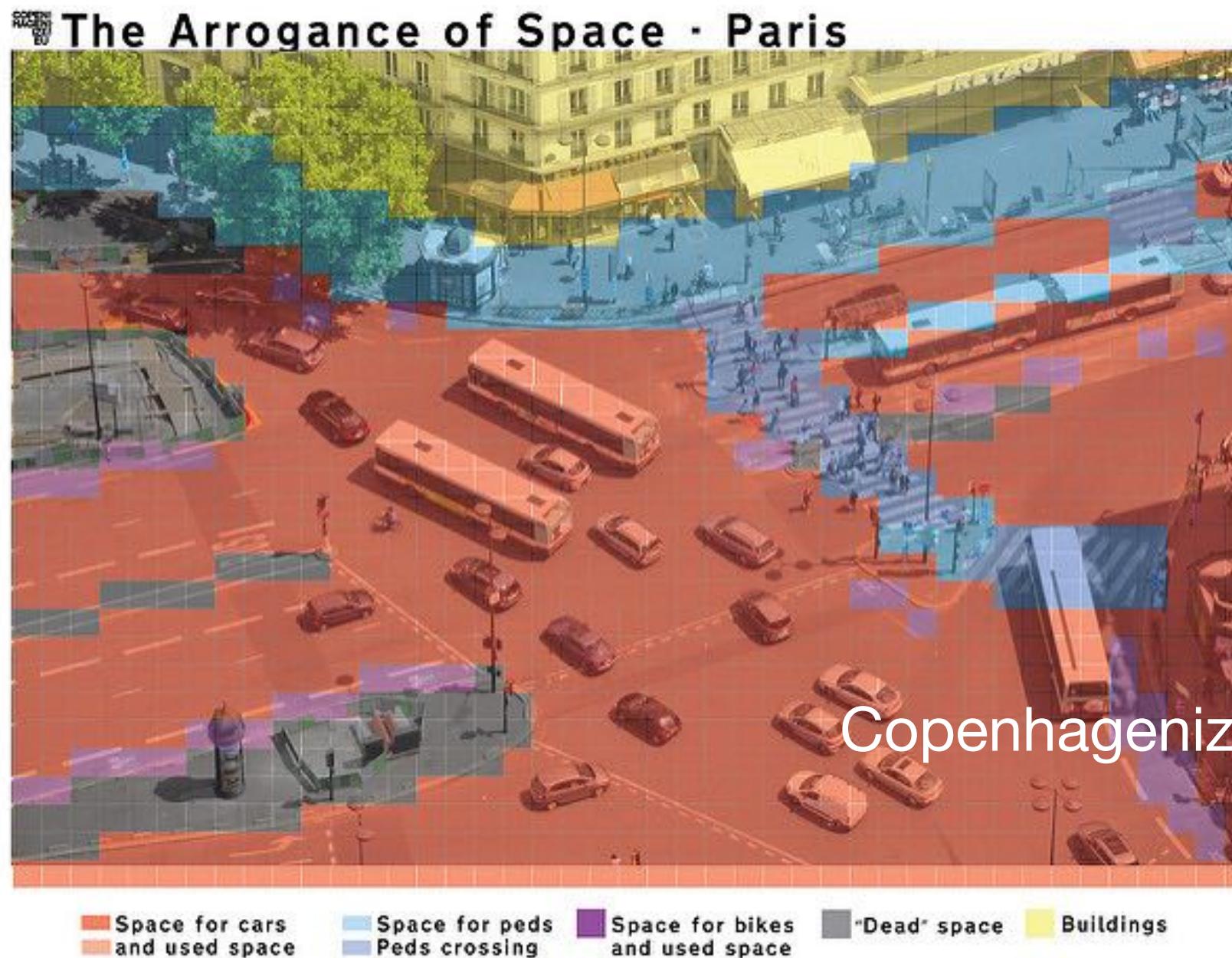
Da der Fahrradverkehr in Berlin stetig zunimmt, sollen Radwege stark ausgebaut werden so bis 2030 ein Netz von Radverkehrsverbündungen mit einer Gesamtlänge von 1000 km entstehen, welche Straßen an die Anforderungen des Fahrradverkehrs angepasst werden.

Die geplante Radverkehrsführung soll sich über die gesamte Stadtfläche erstrecken und dabei bestehende Radwege sowie Radwegabschnitte, welche nicht den Anforderungen des Fahrradverkehrs entsprechen, ausgebaut werden. Ein großer Vorteile ist, dass Radfahrer nicht auf asphaltierte Wege müssen, aber ebenfalls die Wünsche der Fußgänger und Radfahrer müssen berücksichtigt werden. Die Radverkehrsverbündungen sollen die Radfahrerstandards und nicht nur die Fußgängersicherheit erhöhen.

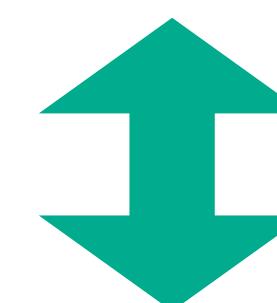
Die geplante Radverkehrsführung soll sich über die gesamte Stadtfläche erstrecken und dabei bestehende Radwege sowie Radwegabschnitte, welche nicht den Anforderungen des Fahrradverkehrs entsprechen, ausgebaut werden. Ein großer Vorteile ist, dass Radfahrer nicht auf asphaltierte Wege müssen, aber ebenfalls die Wünsche der Fußgänger und Radfahrer müssen berücksichtigt werden. Die Radverkehrsverbündungen sollen die Radfahrerstandards und nicht nur die Fußgängersicherheit erhöhen.



# Cycling is marginalized - in urban landscapes and in data collection

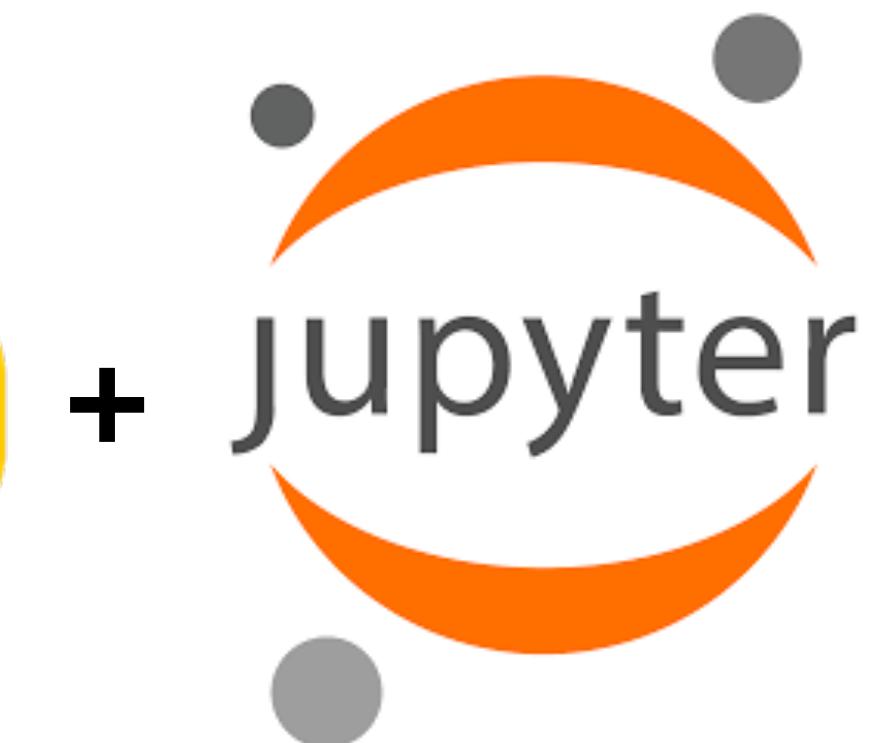
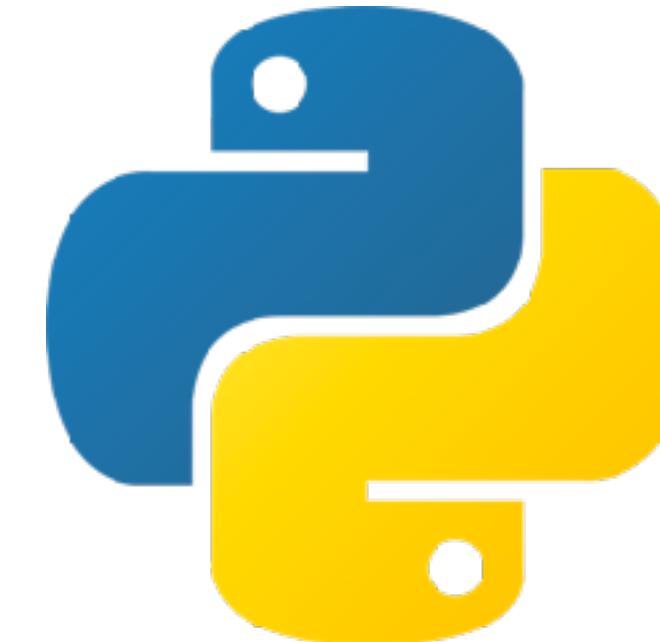
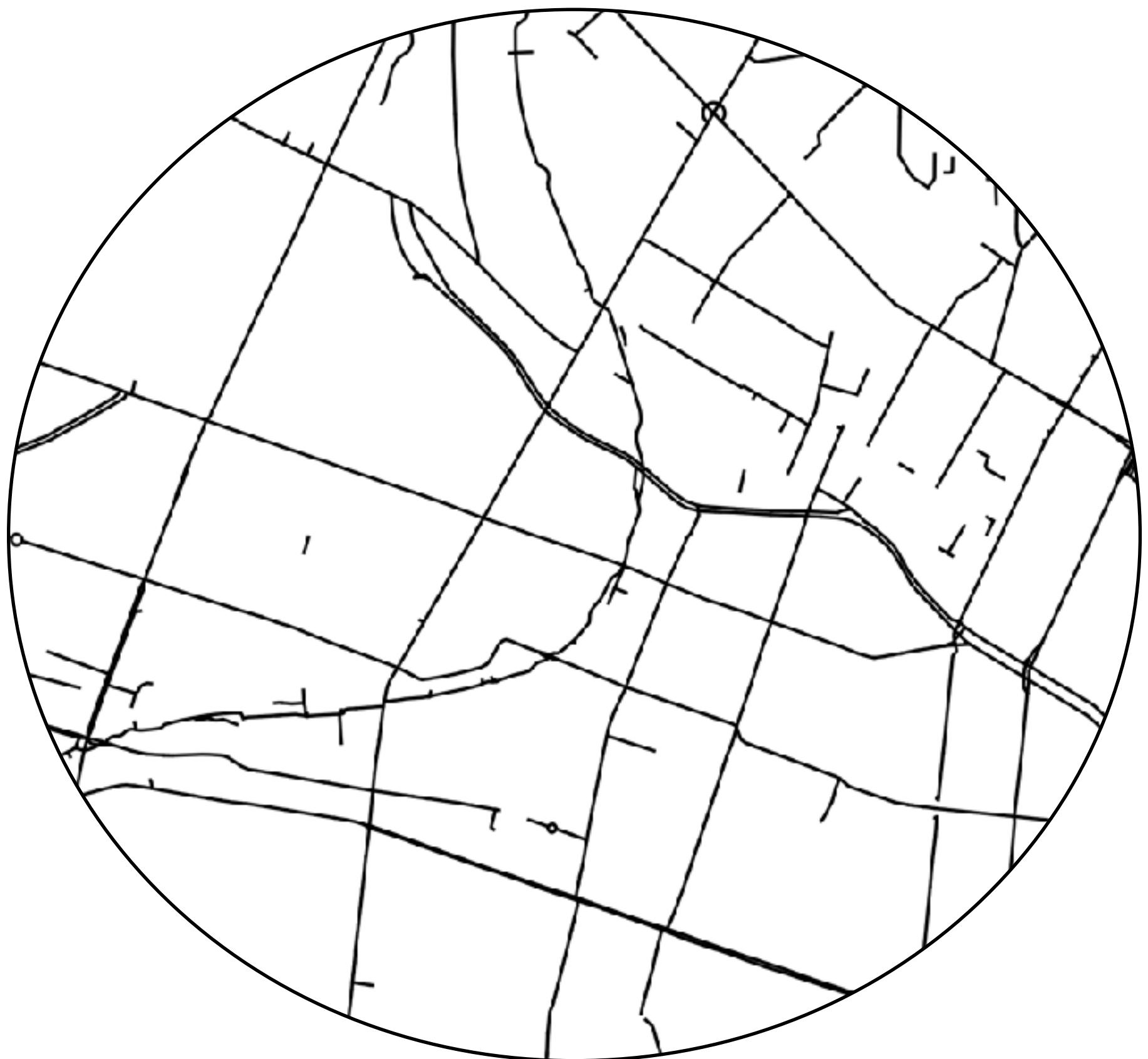


Data reflects priorities



Data influences priorities and decisions

# BikeDNA: Bicycle Infrastructure Data & Network Assessment



Funded by The Danish Road Directorate



# DATA QUALITY IS MULTI-FACETED

## ISO 19115

- Completeness
- Consistency
- Positional accuracy
- Temporal accuracy
- Thematic accuracy



# DATA QUALITY IS MULTI-FACETED

## ISO 19115

- Completeness
- Consistency
- Positional accuracy
- Temporal accuracy
- Thematic accuracy



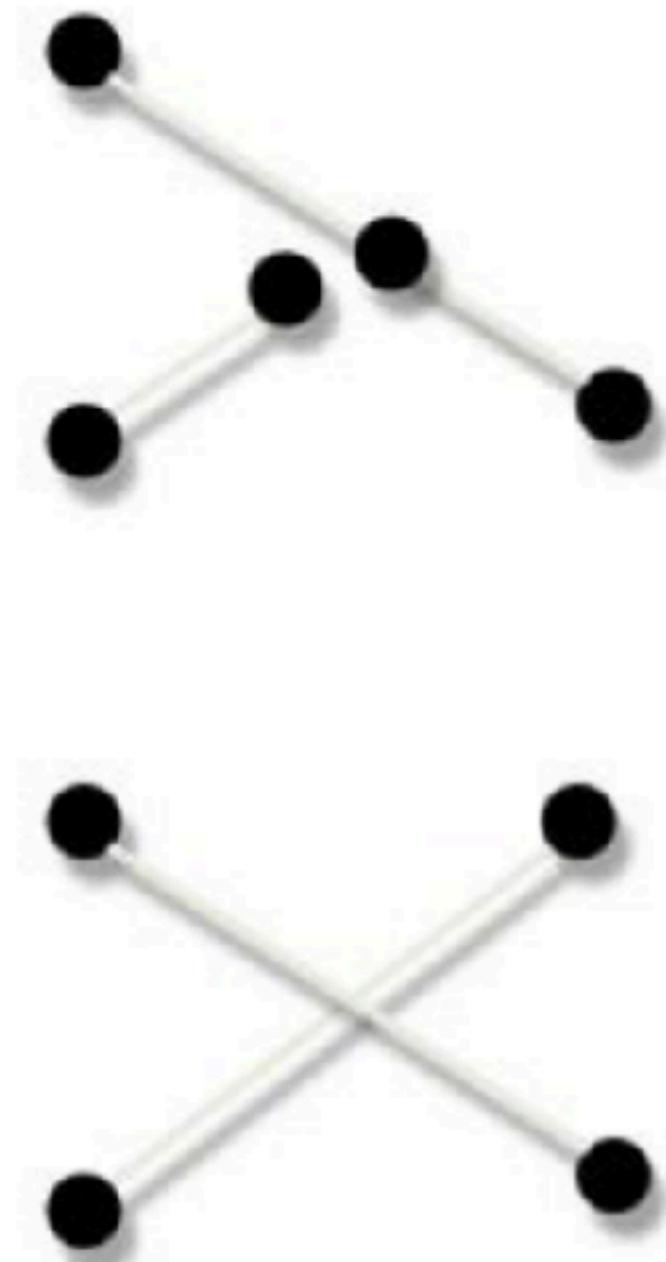
# 'Fitness for Purpose'

Are data good enough for my use case?



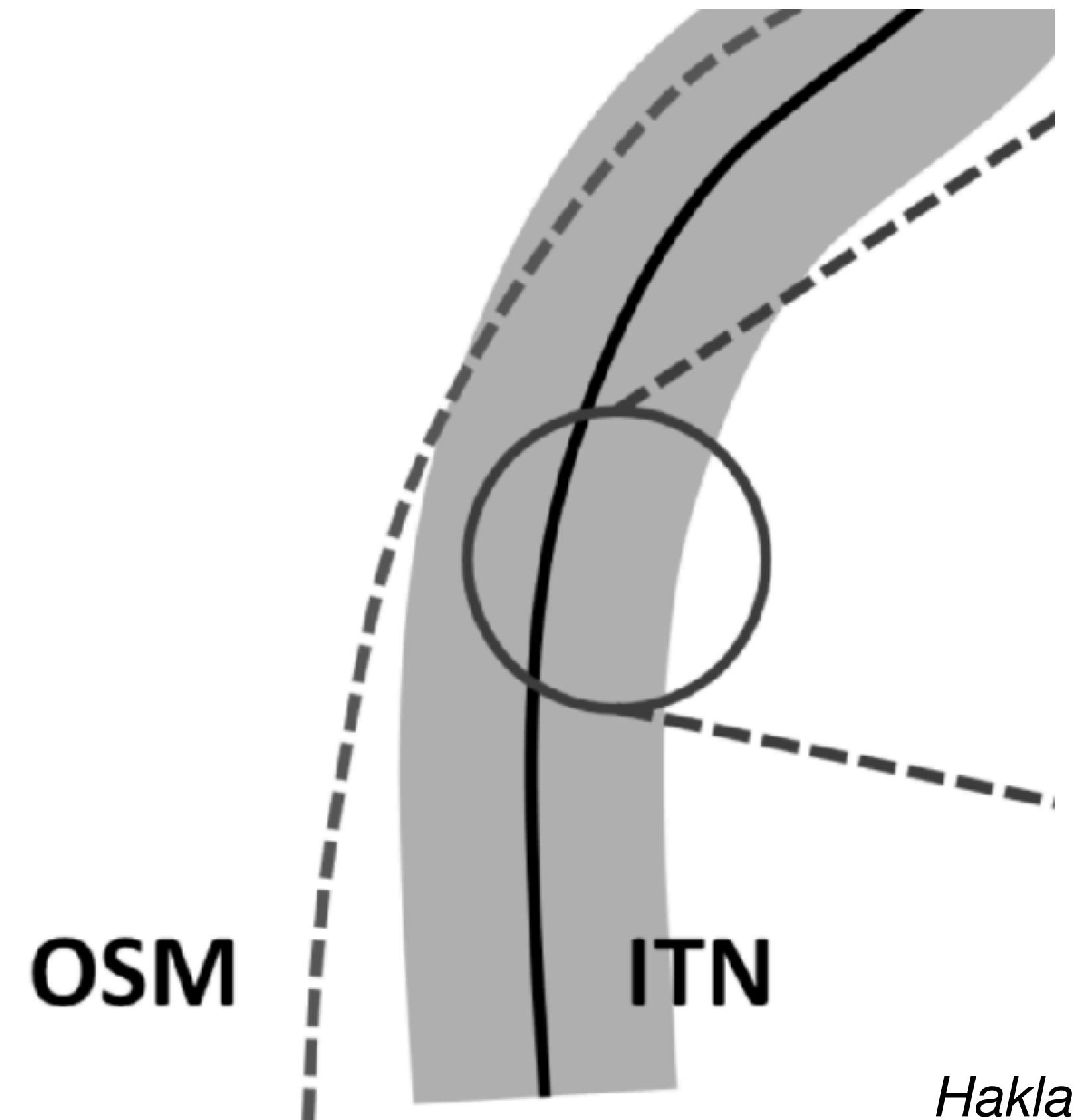
# We care more about topology than accuracy

## TOPOLOGY



*Neis et al. 2012*

## ACCURACY



*Haklay et al. 2010*

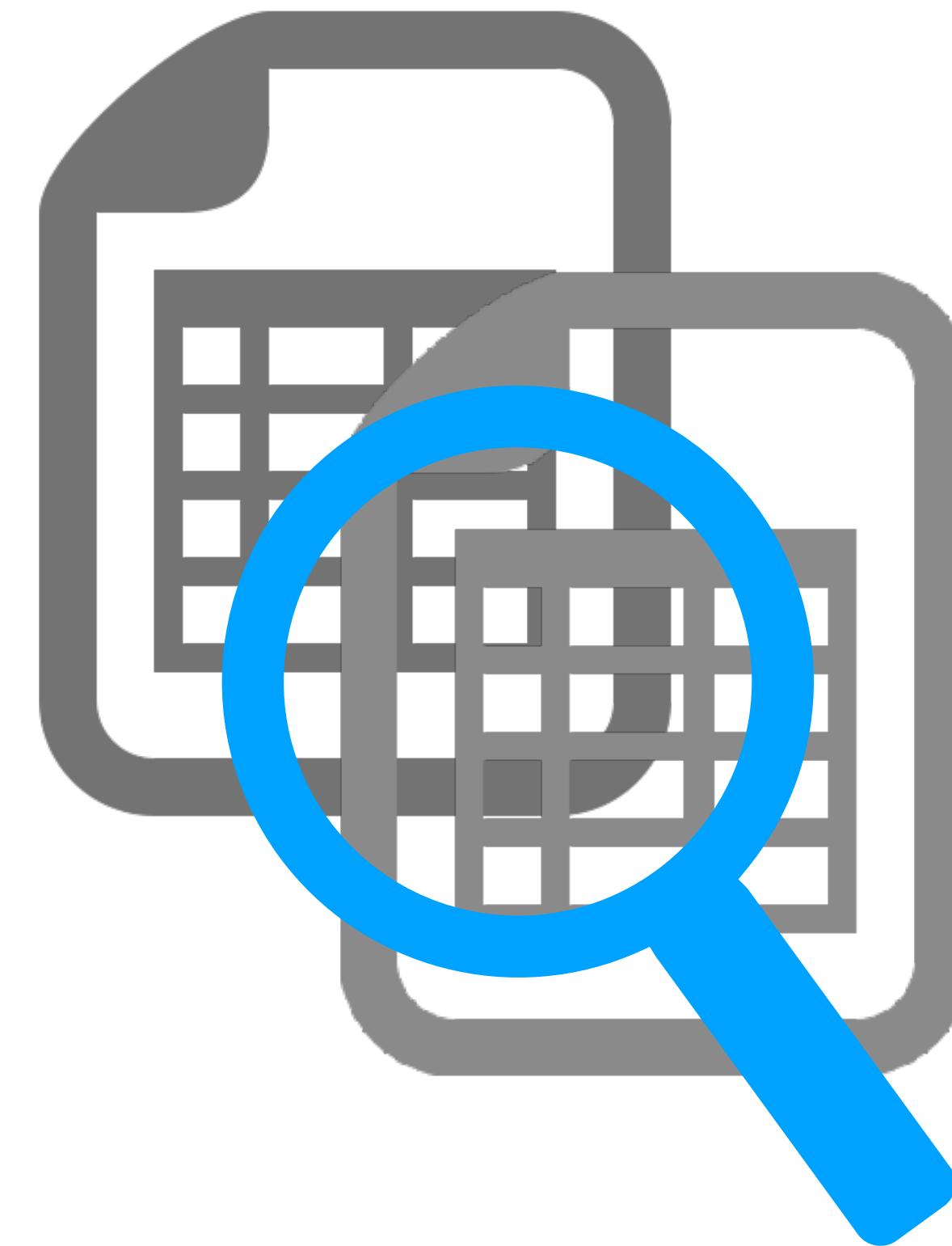
See also Hochmair et al. 2015, Barron et al. 2014

You can analyze one data set or compare two

INTRINSIC



EXTRINSIC



# BikeDNA: Bicycle Infrastructure Data & Network Assessment

README.md

## BikeDNA: Bicycle Infrastructure Data & Network Assessment

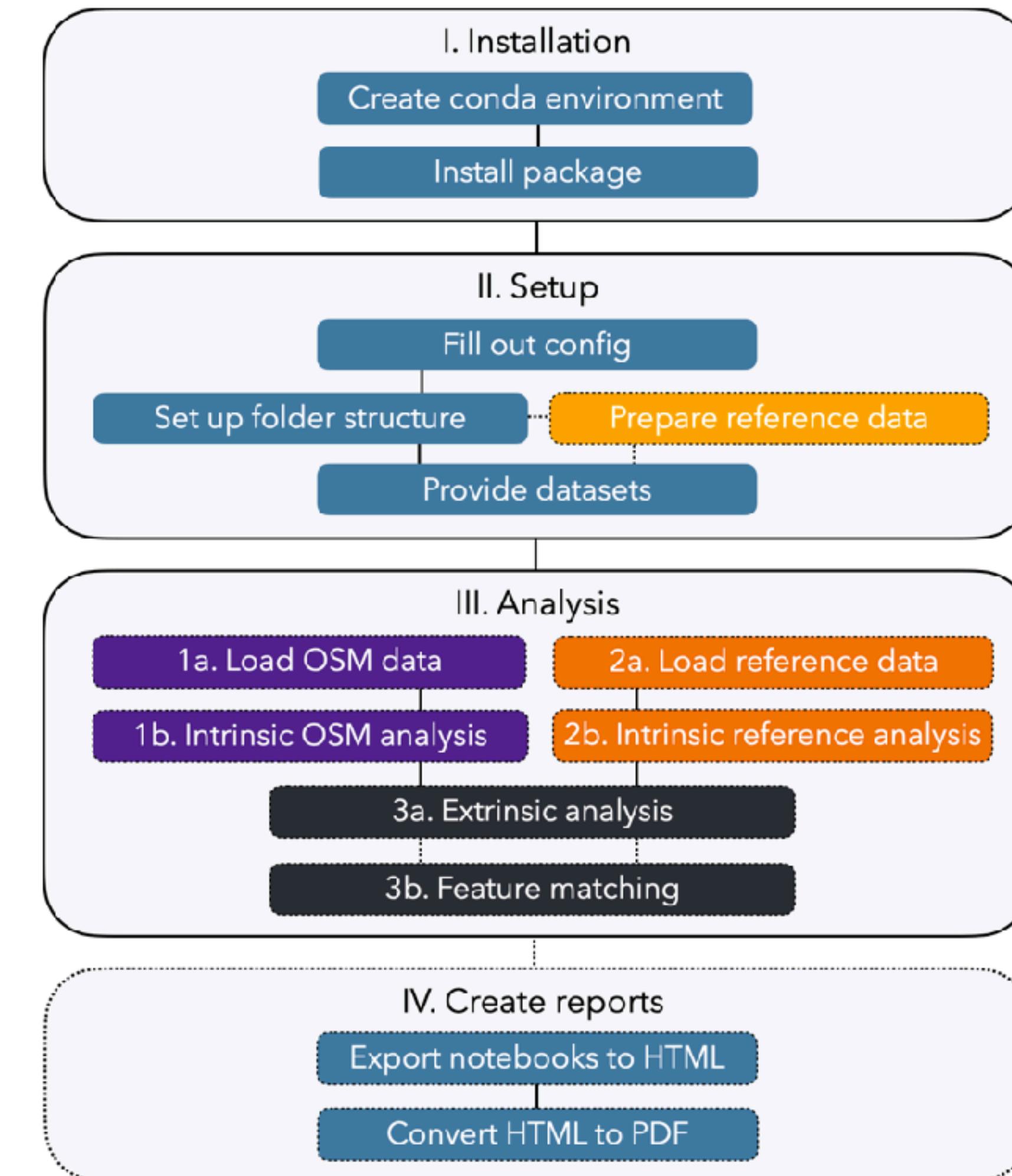
This is the repository of BikeDNA, a tool for assessing the quality of [OpenStreetMap \(OSM\)](#) and other bicycle infrastructure data sets in a reproducible way. It provides researchers and others who work with bicycle networks a detailed, informed overview of data quality in a given area.

► Background

### Workflow

BikeDNA consists of Jupyter notebooks that analyze and compare bicycle infrastructure data sets. Therefore, to install and run BikeDNA, an installation of [Python](#), including tools for [Jupyter notebook](#), is required.

The [installation](#), [setup](#), [analysis](#), and [export](#) steps are illustrated in the figure and described in detail below. Dotted parts are optional.



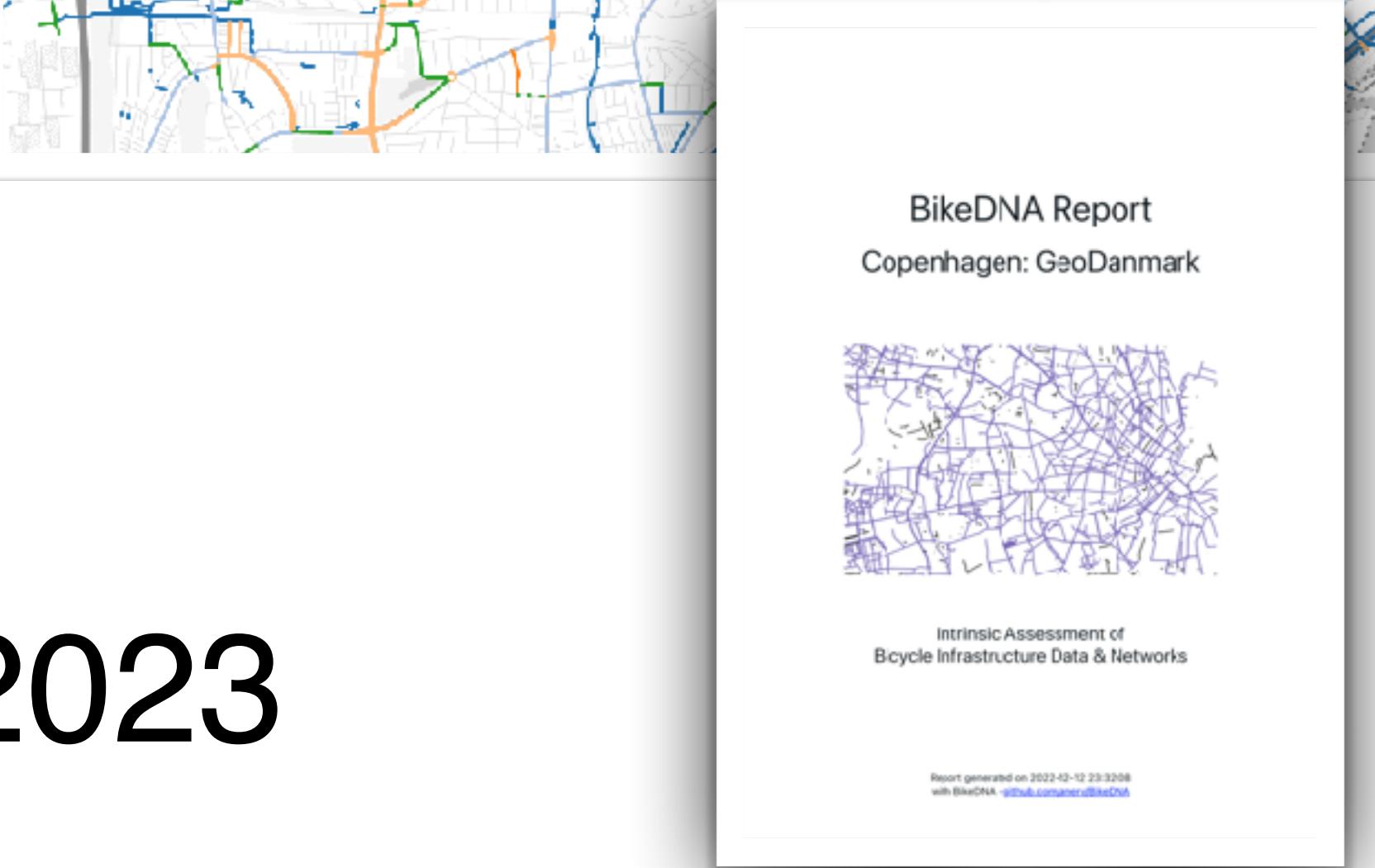
# Features of BikeDNA

Data + network quality (completeness, consistency, accuracy)

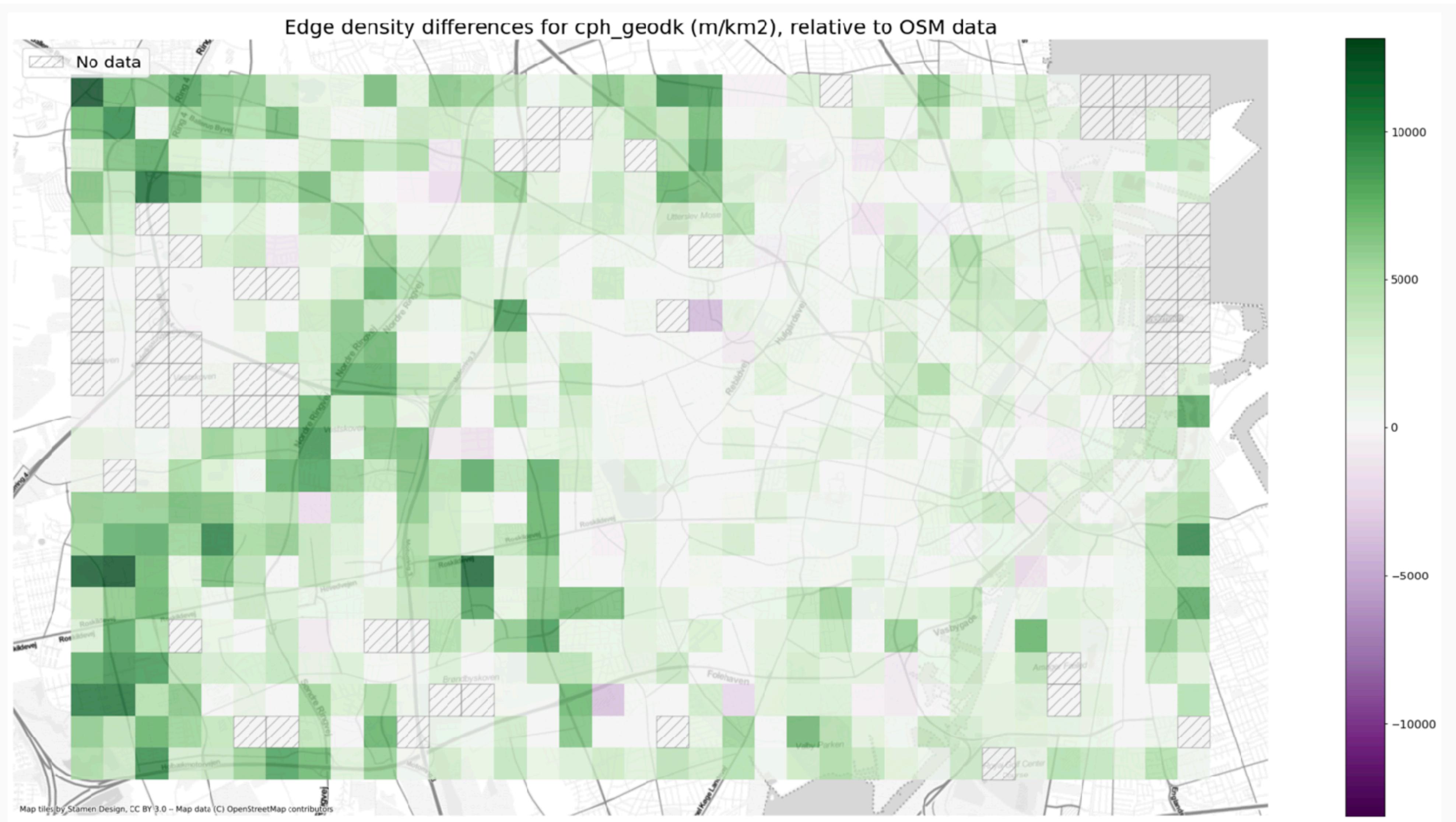
Both intrinsic and extrinsic, comparing reference data to OSM

Export reports: HTML (interactive), PDF

Under development, release early 2023



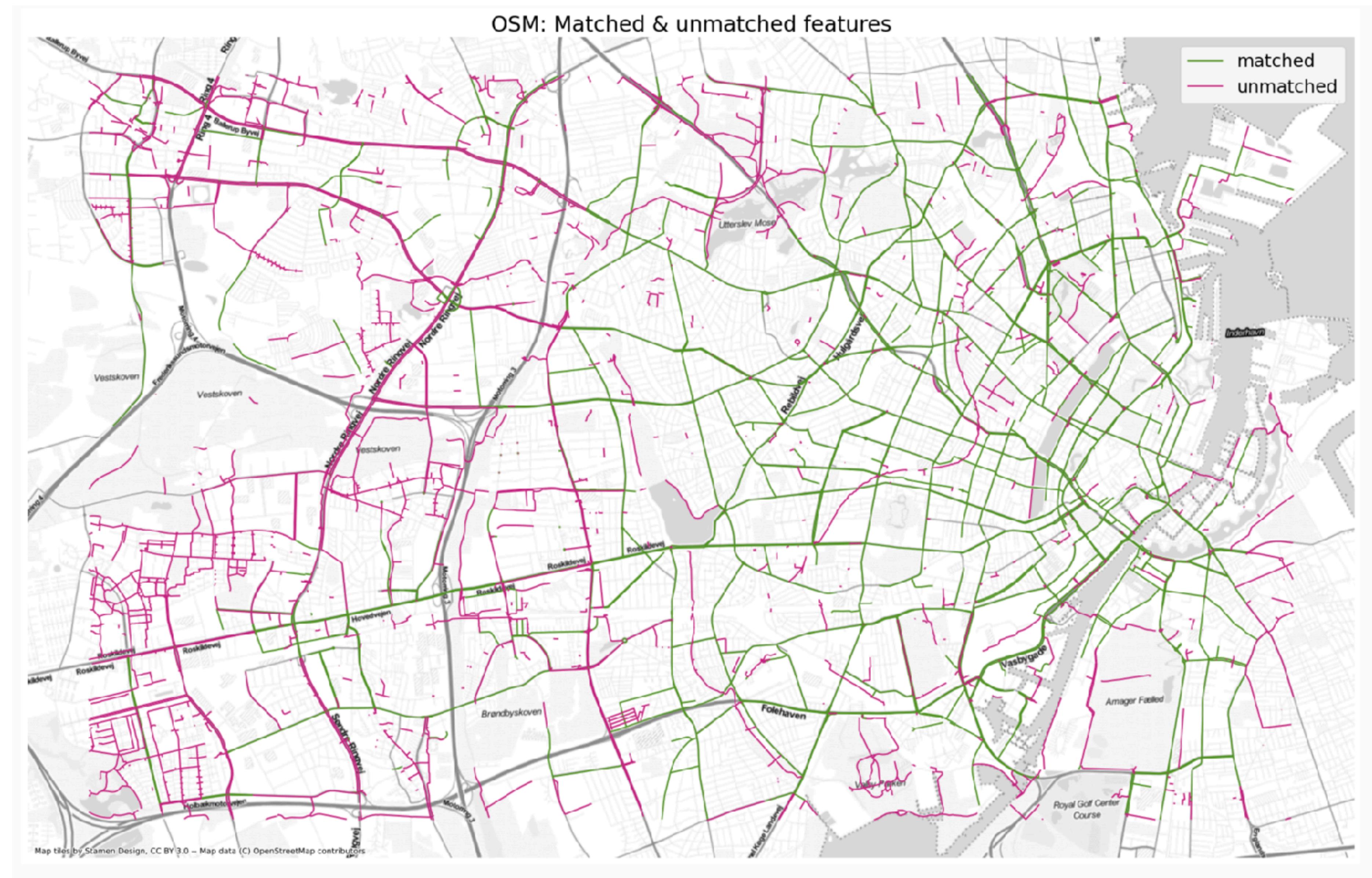
# Data completeness is the first step



...but comparing data completeness is tricky



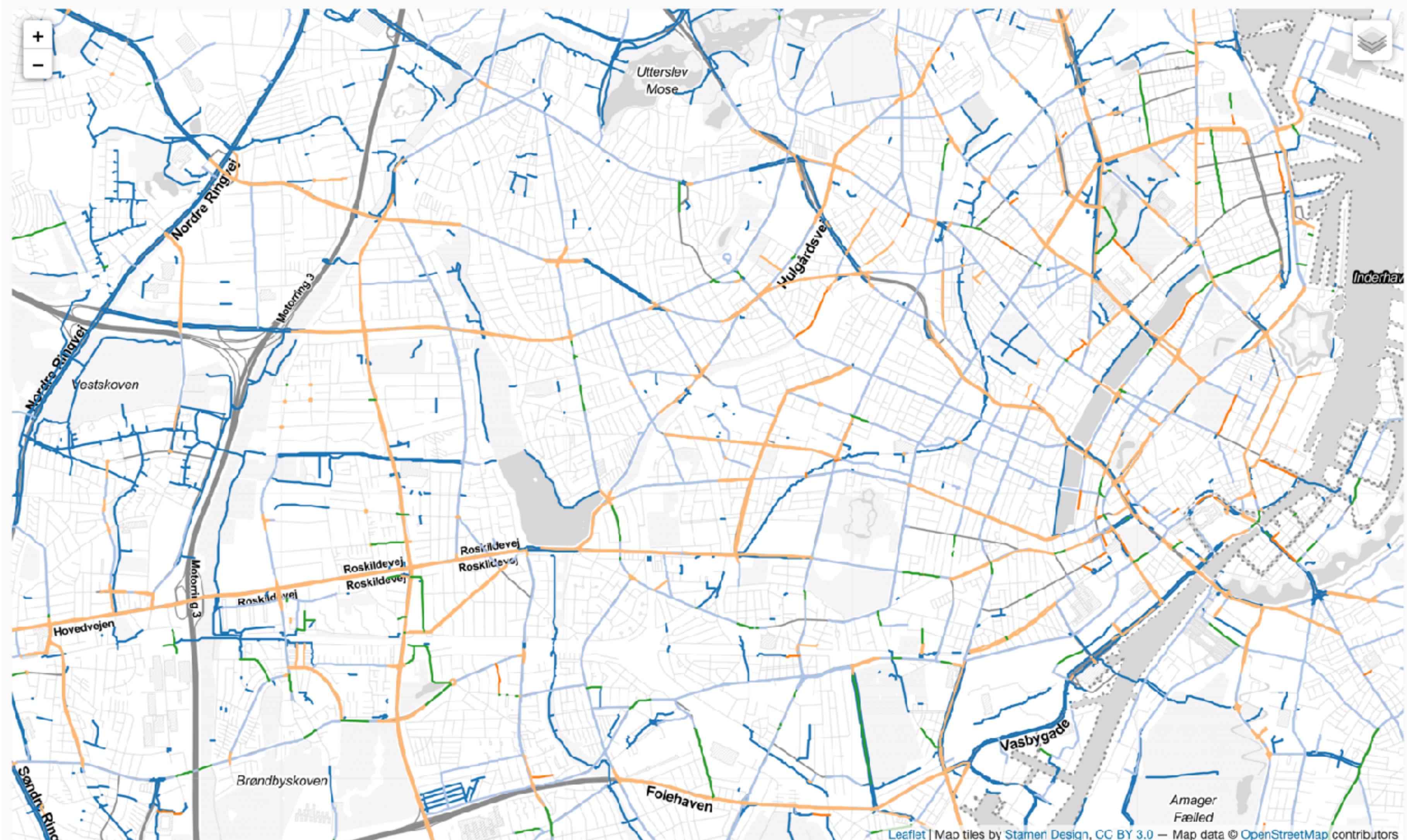
# Feature matching reveals inaccuracies & missing data



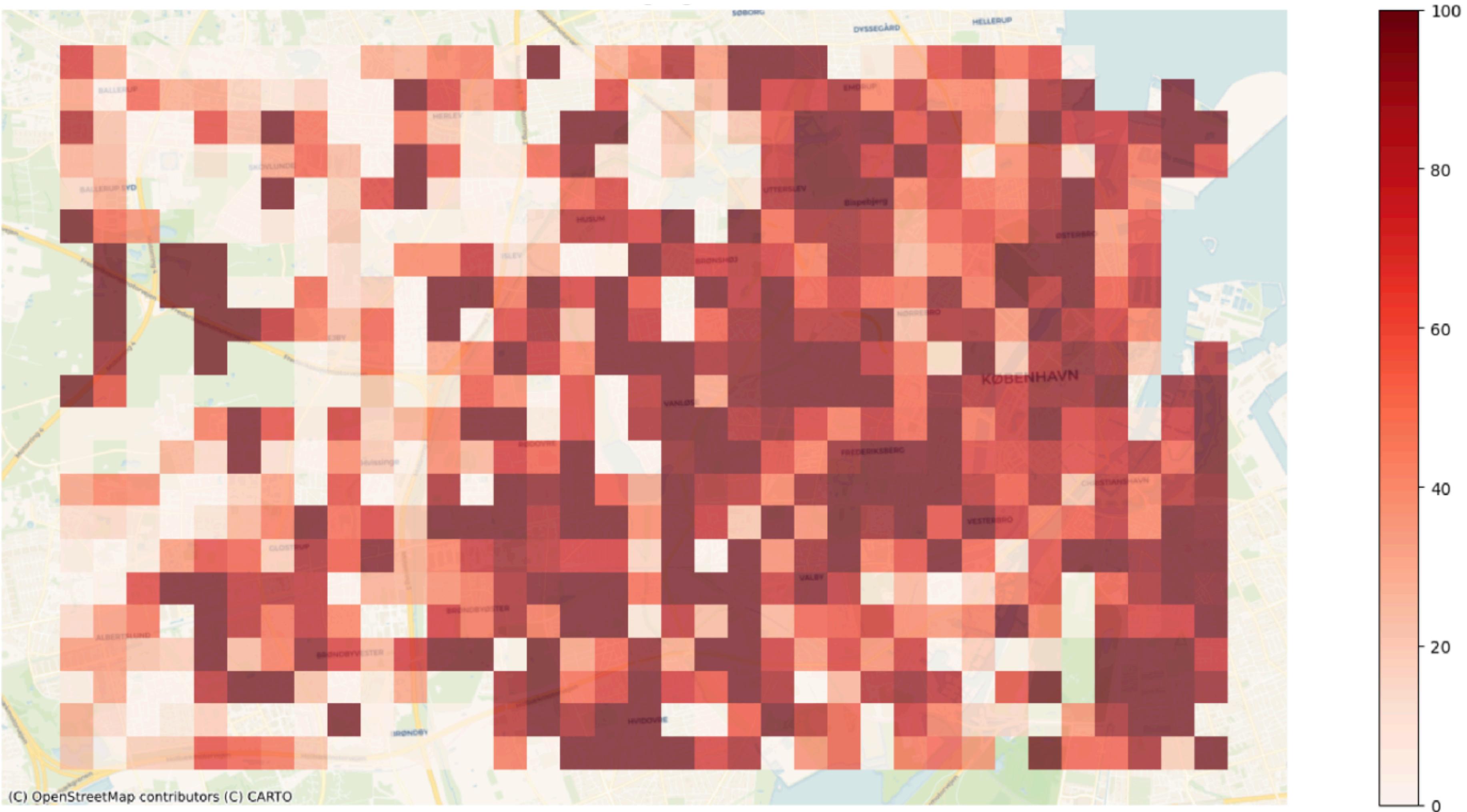
See also Koukoletsos et al. 2012

© OpenStreetMap contributors, © GeoDanmark

# We map the tagging patterns to help new users understand OSM practices



# OSM tags are added inconsistently



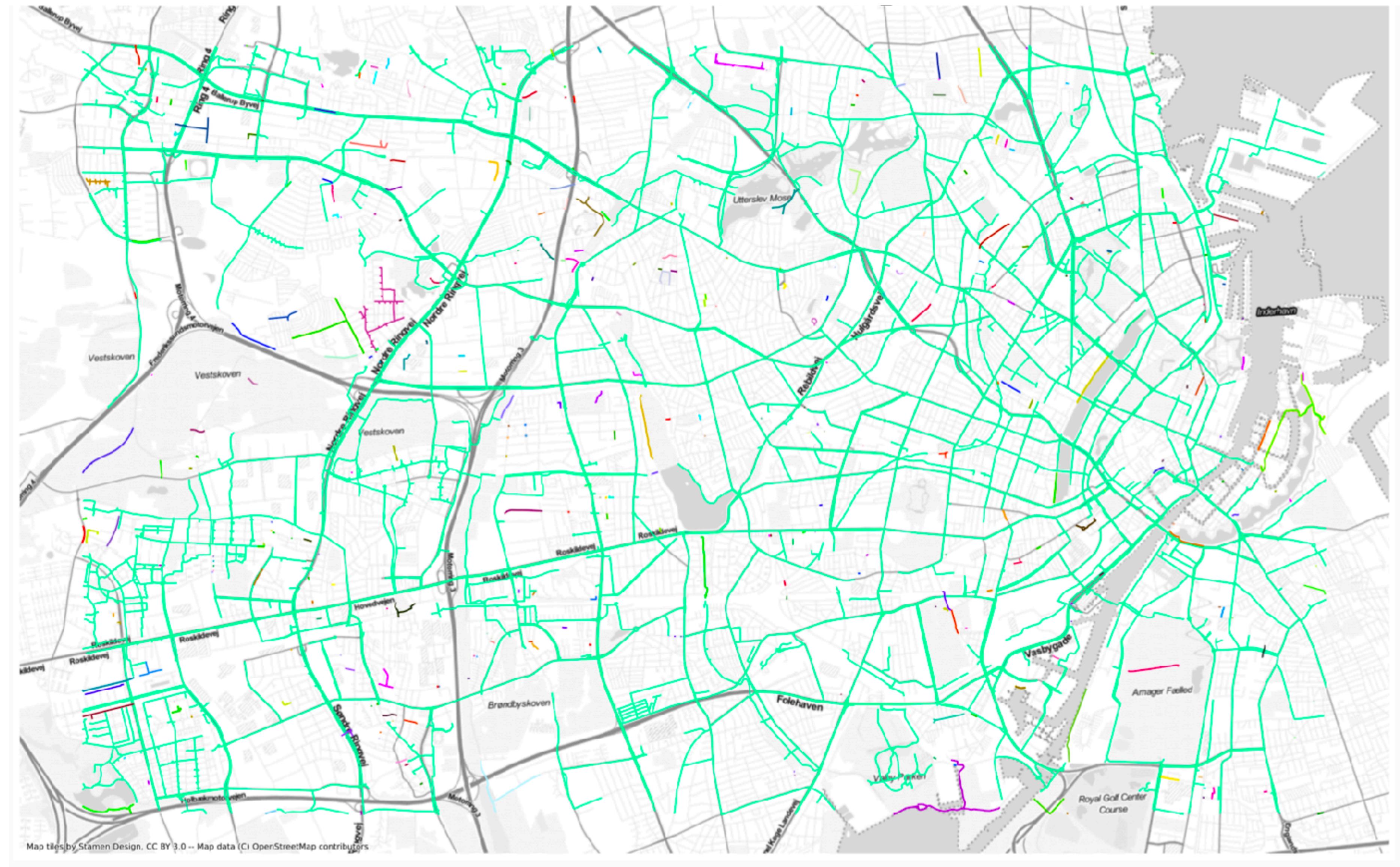
% edges without 'surface' tag

© OpenStreetMap contributors

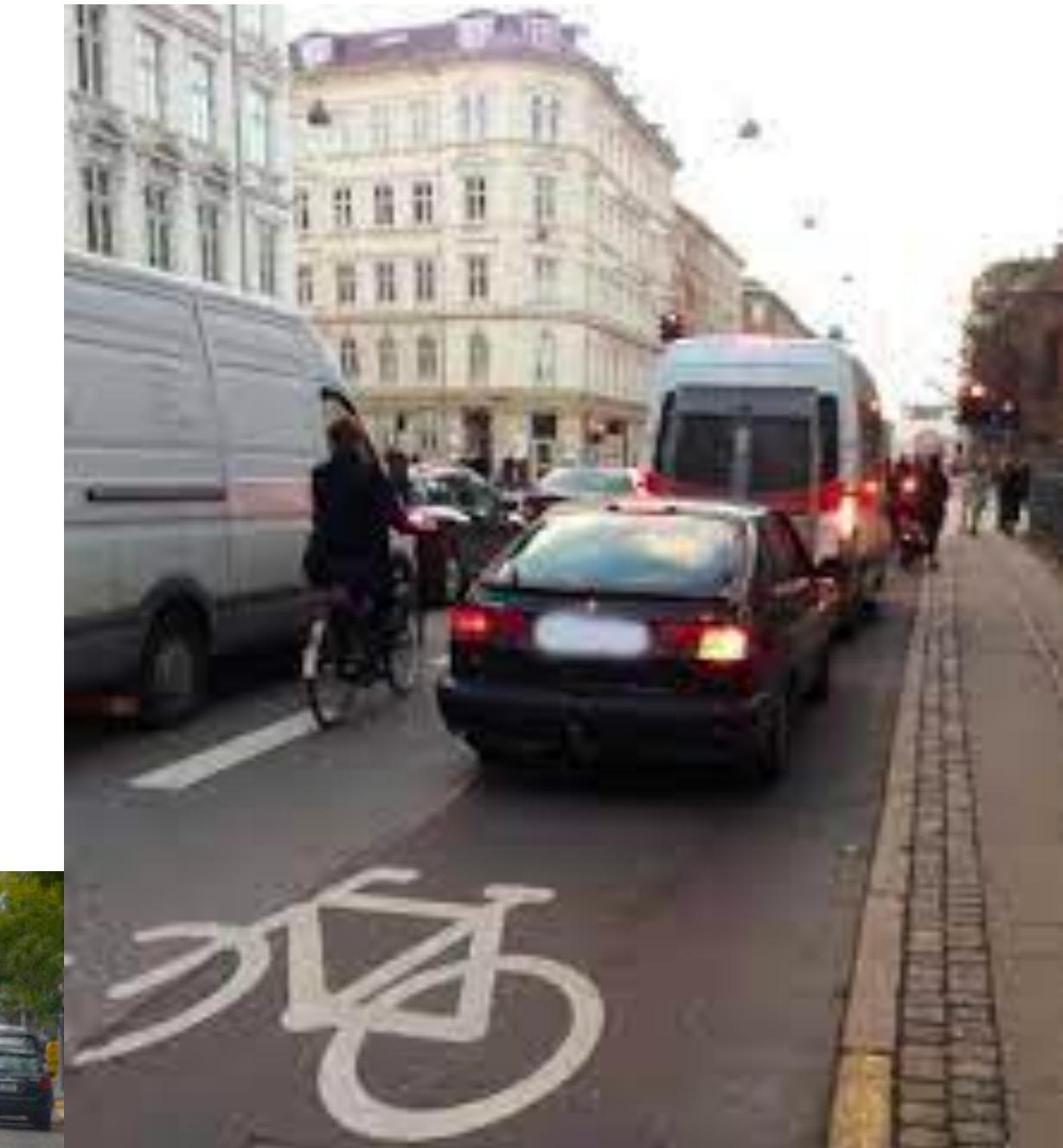
# Actual bicycle networks are often not connected....



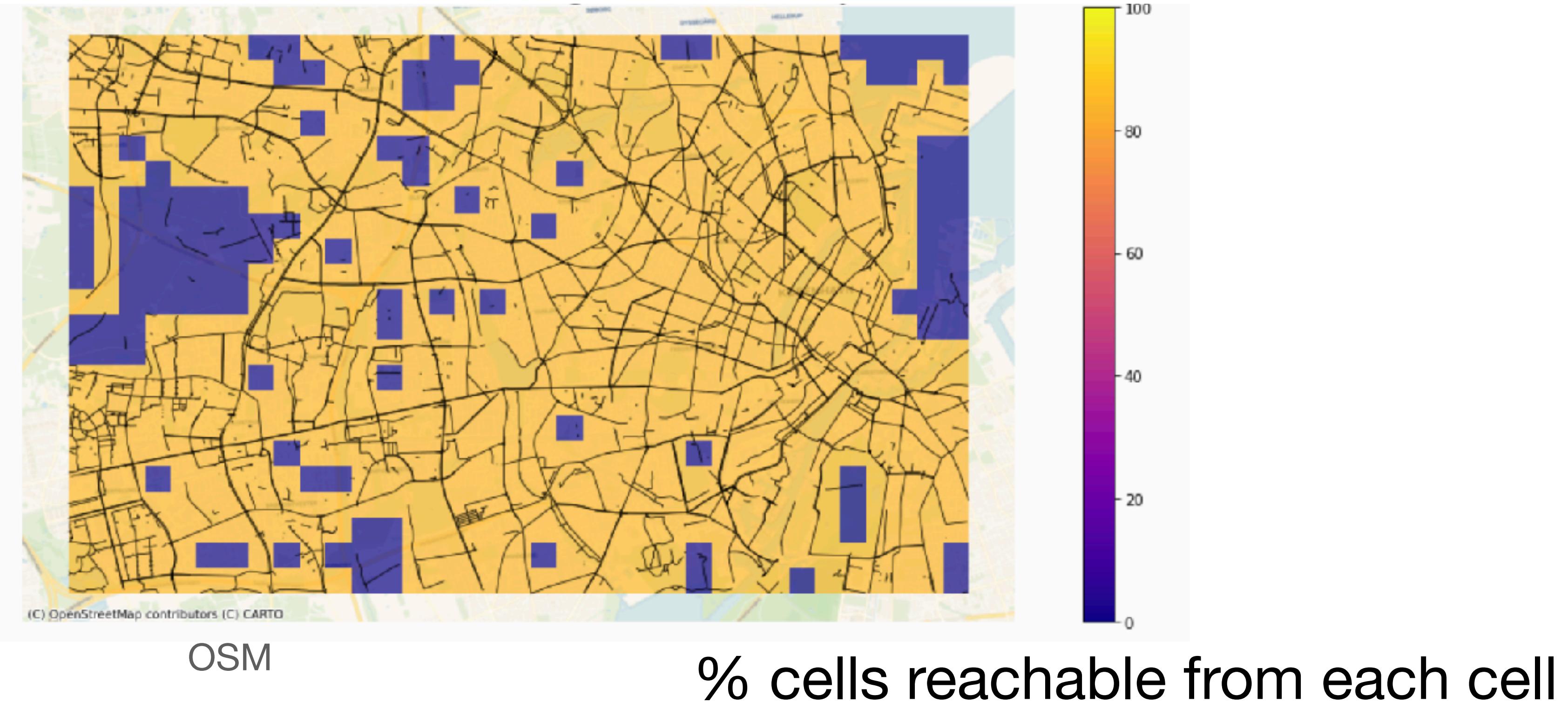
...but even less so when mapped



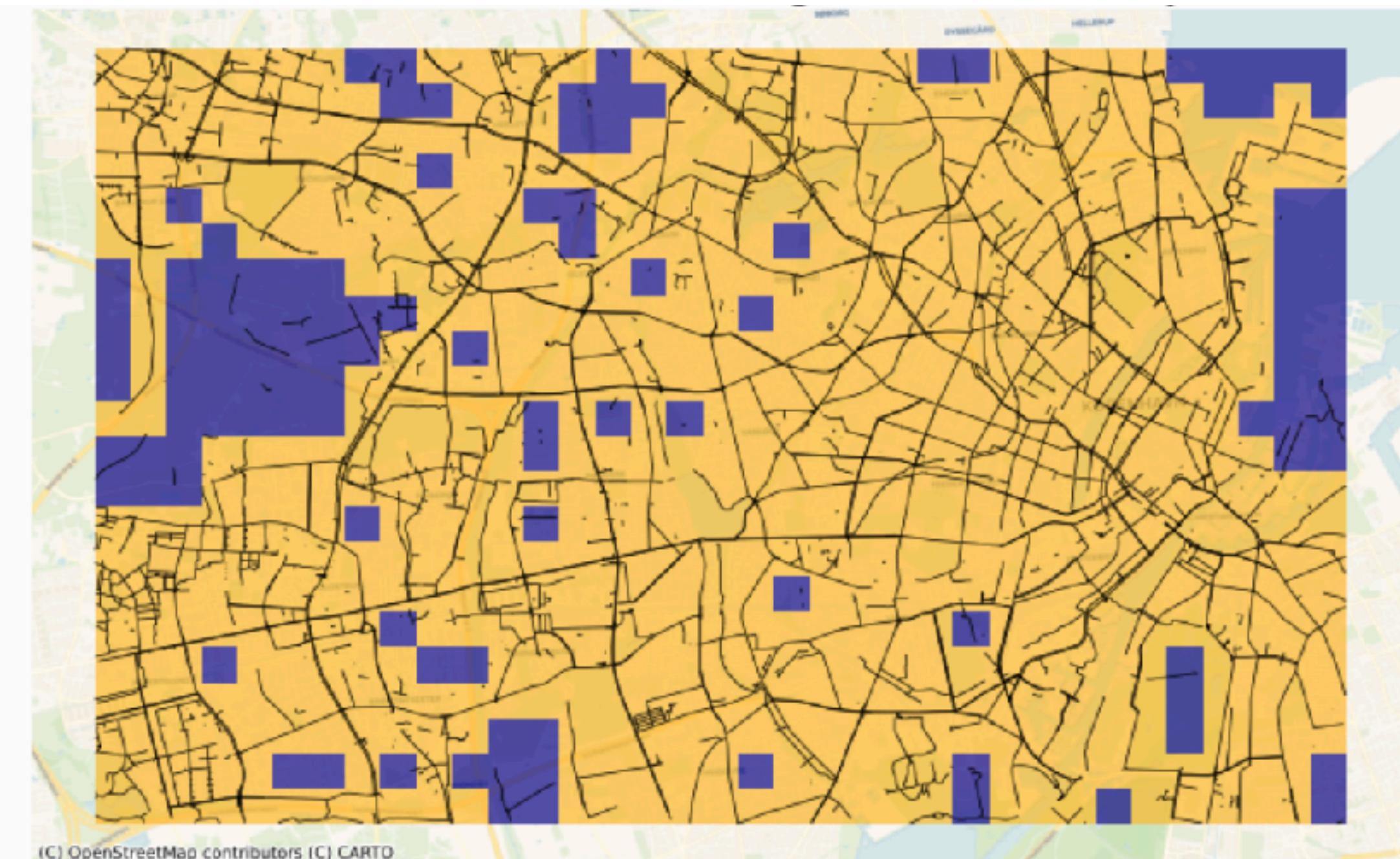
# Inconsistent mapping methods results data gaps



# Errors and omissions have real effects

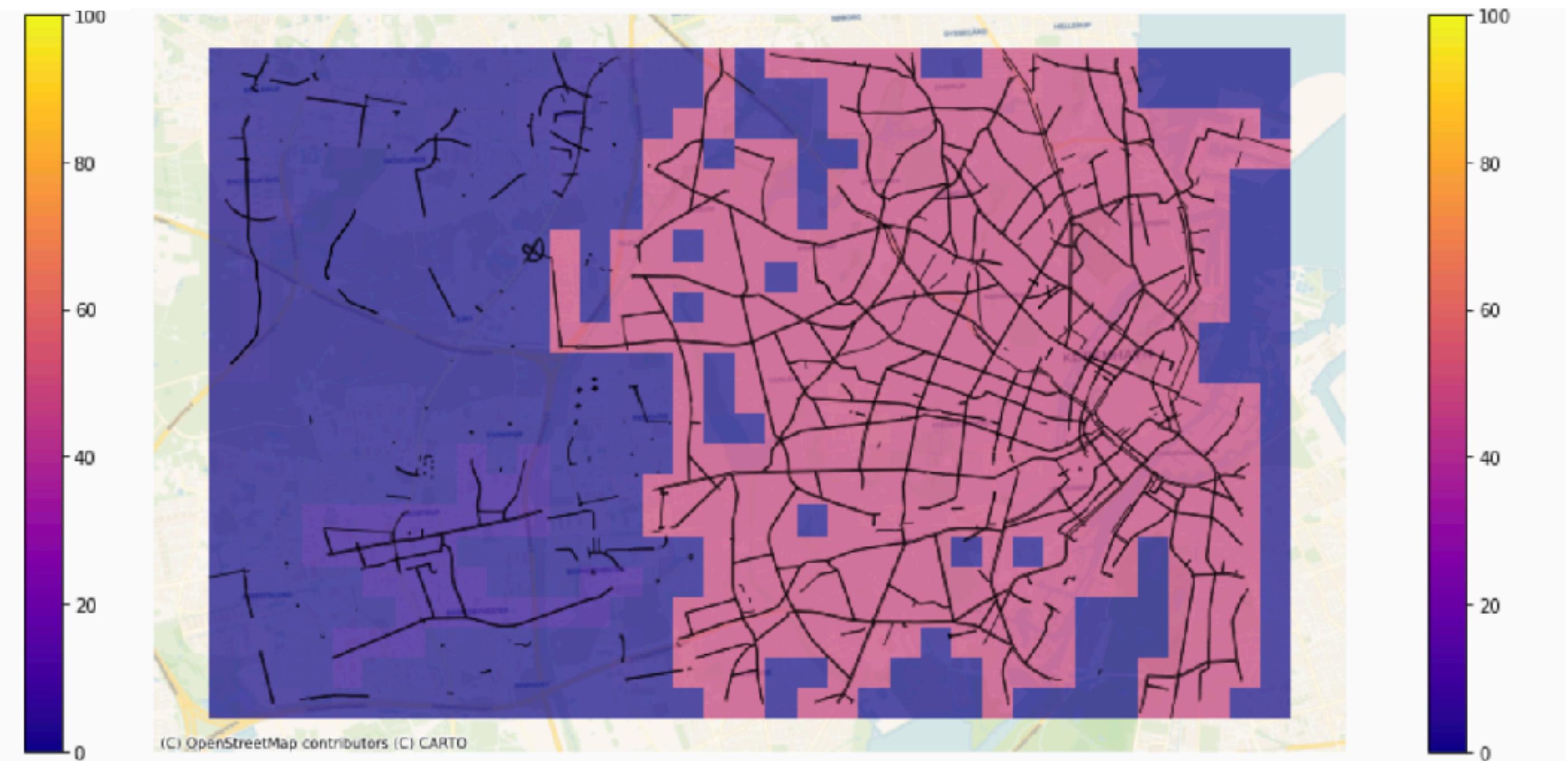


# Errors and omissions have real effects



OSM

% cells reachable from each cell



GeoDanmark

# Use cases of BikeDNA

Urban / regional / national planners

Researchers

OSM maintainers

Cycling advocates

# There are still unsolved challenges

No ground truth

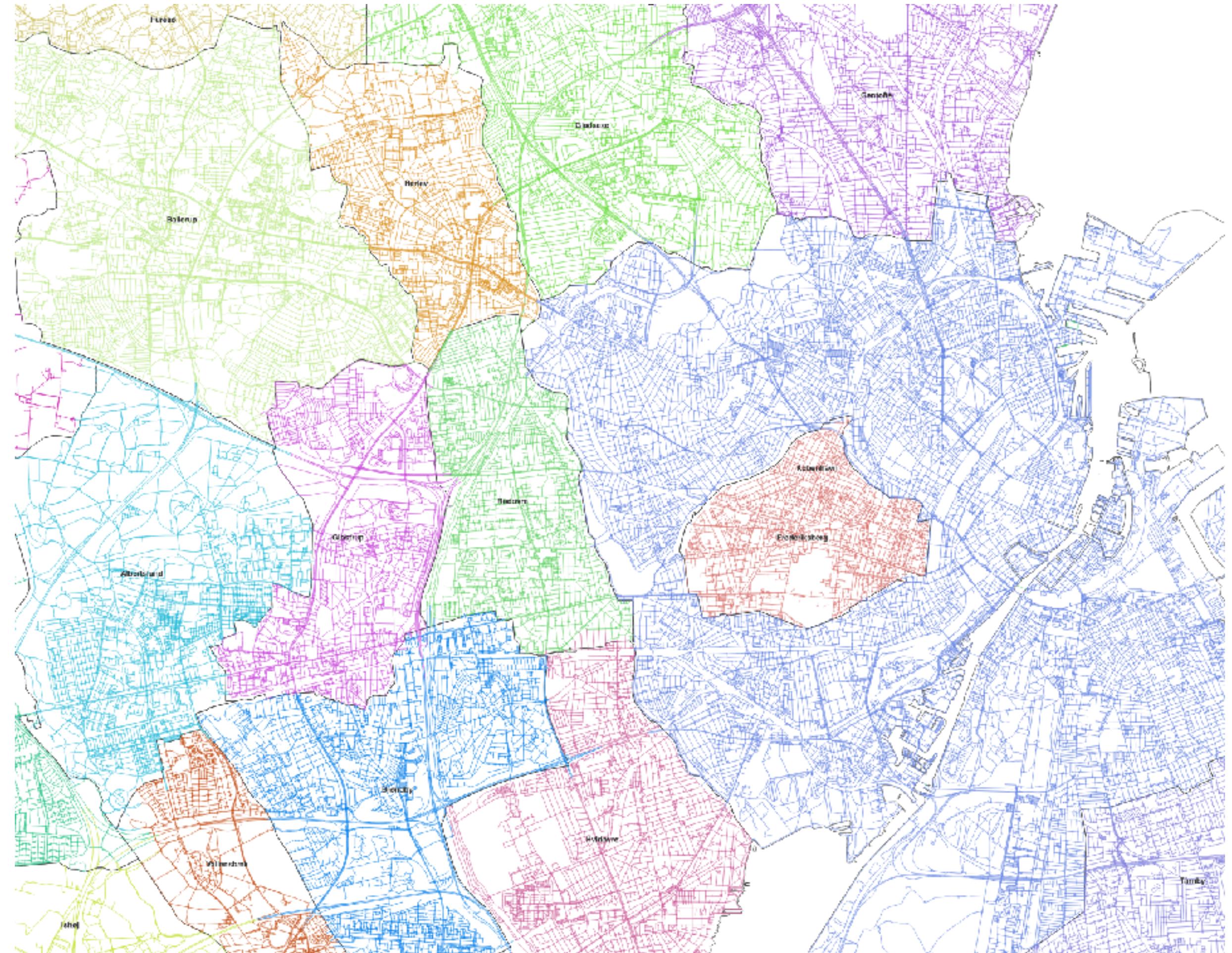


# There are still unsolved challenges

No ground truth

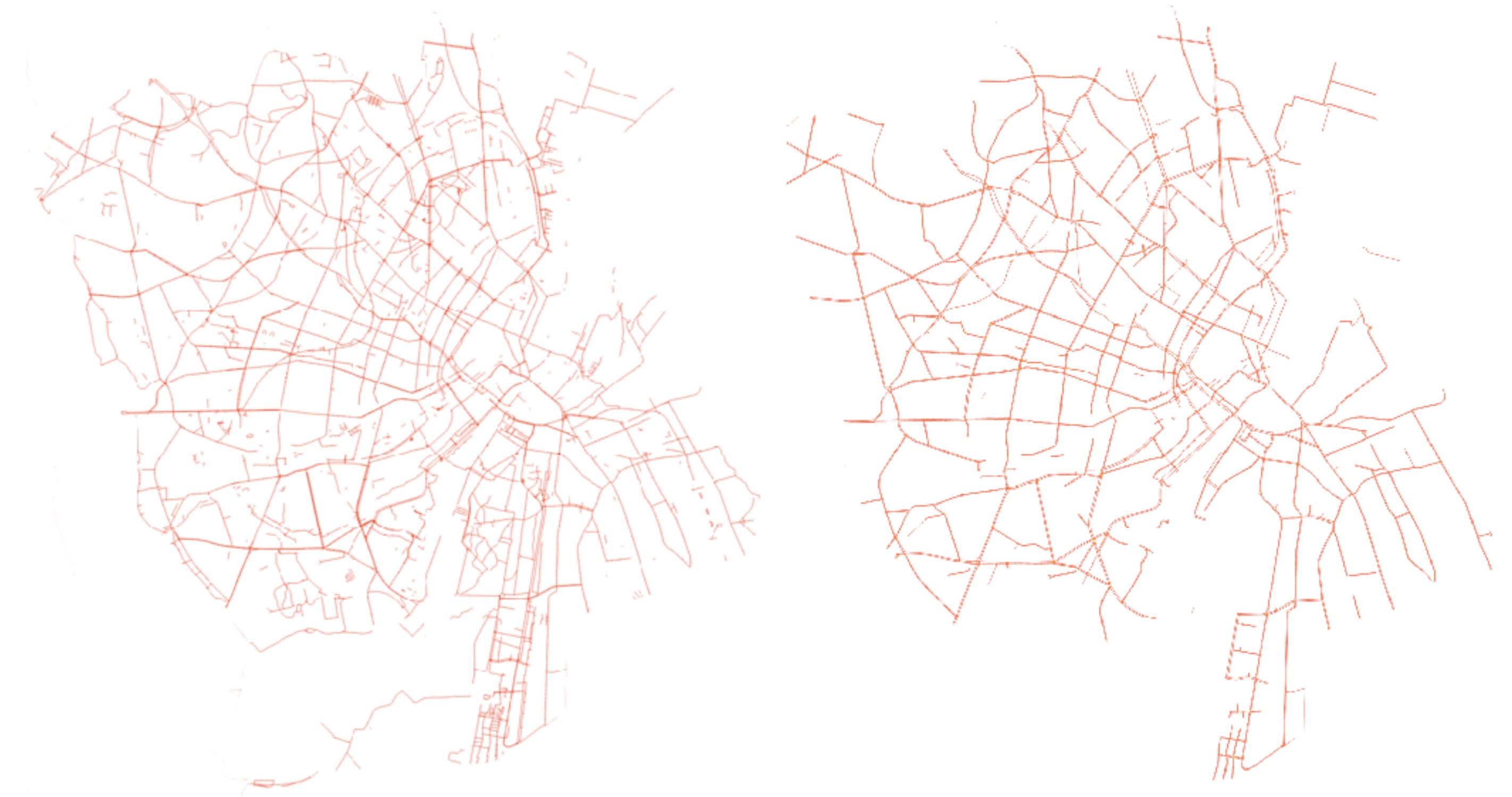


How do we benchmark and compare?



# Key take aways

Data quality matters



# Key take aways

Data quality matters

Quality data not a given

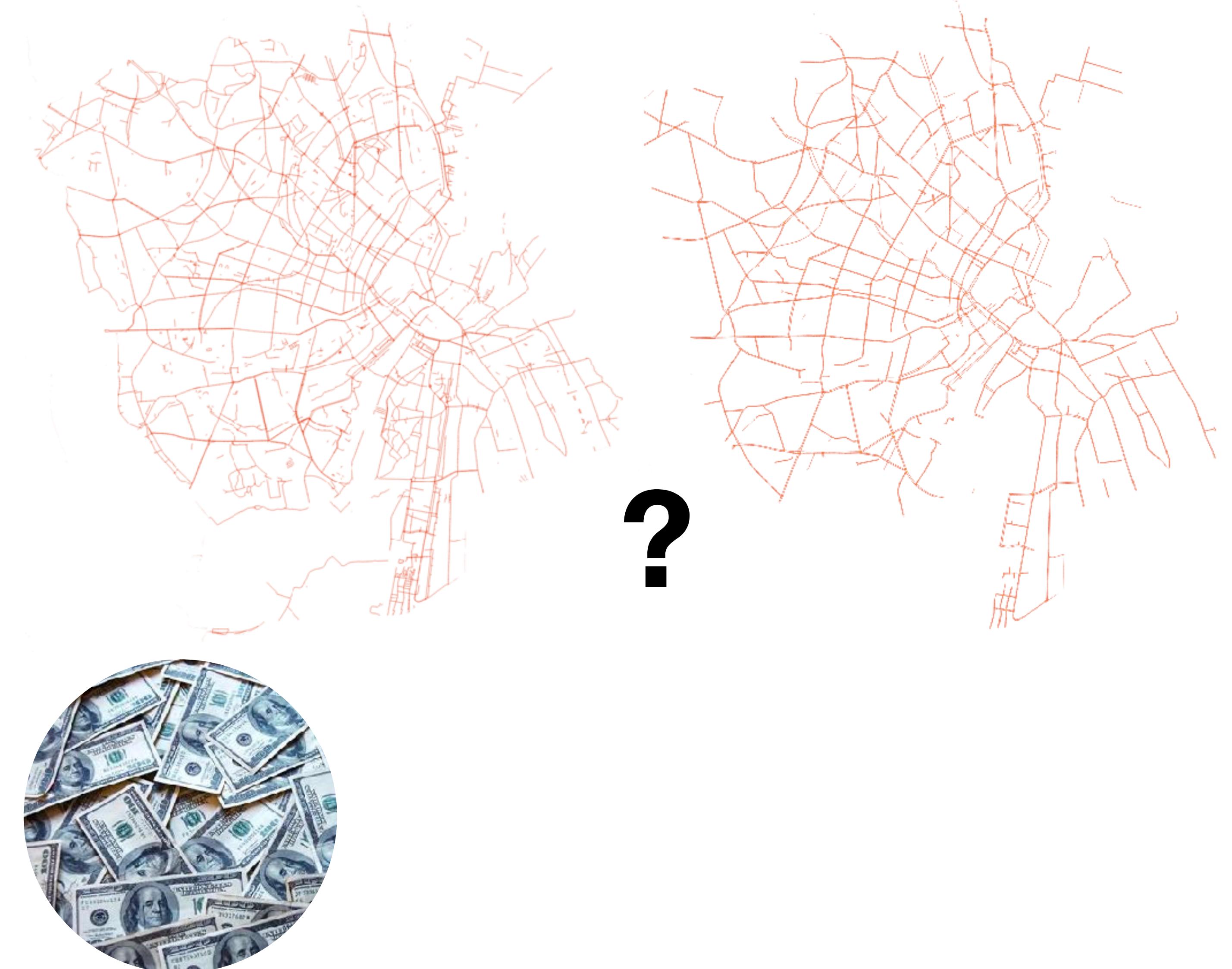


# Key take aways

Data quality matters

Quality data not a given

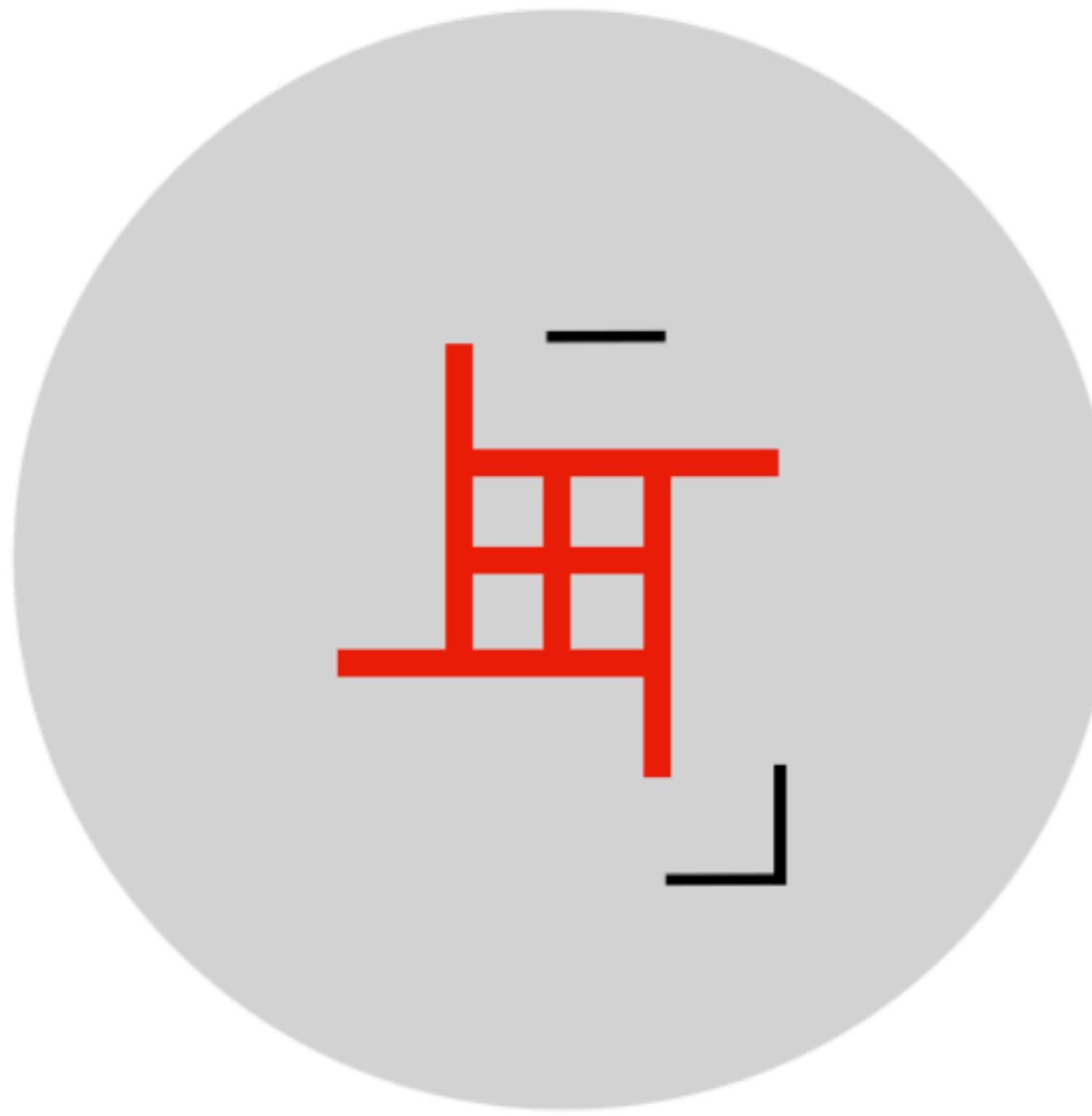
Data (quality) is political



# Different cities need different strategies

Most cities

Not developed

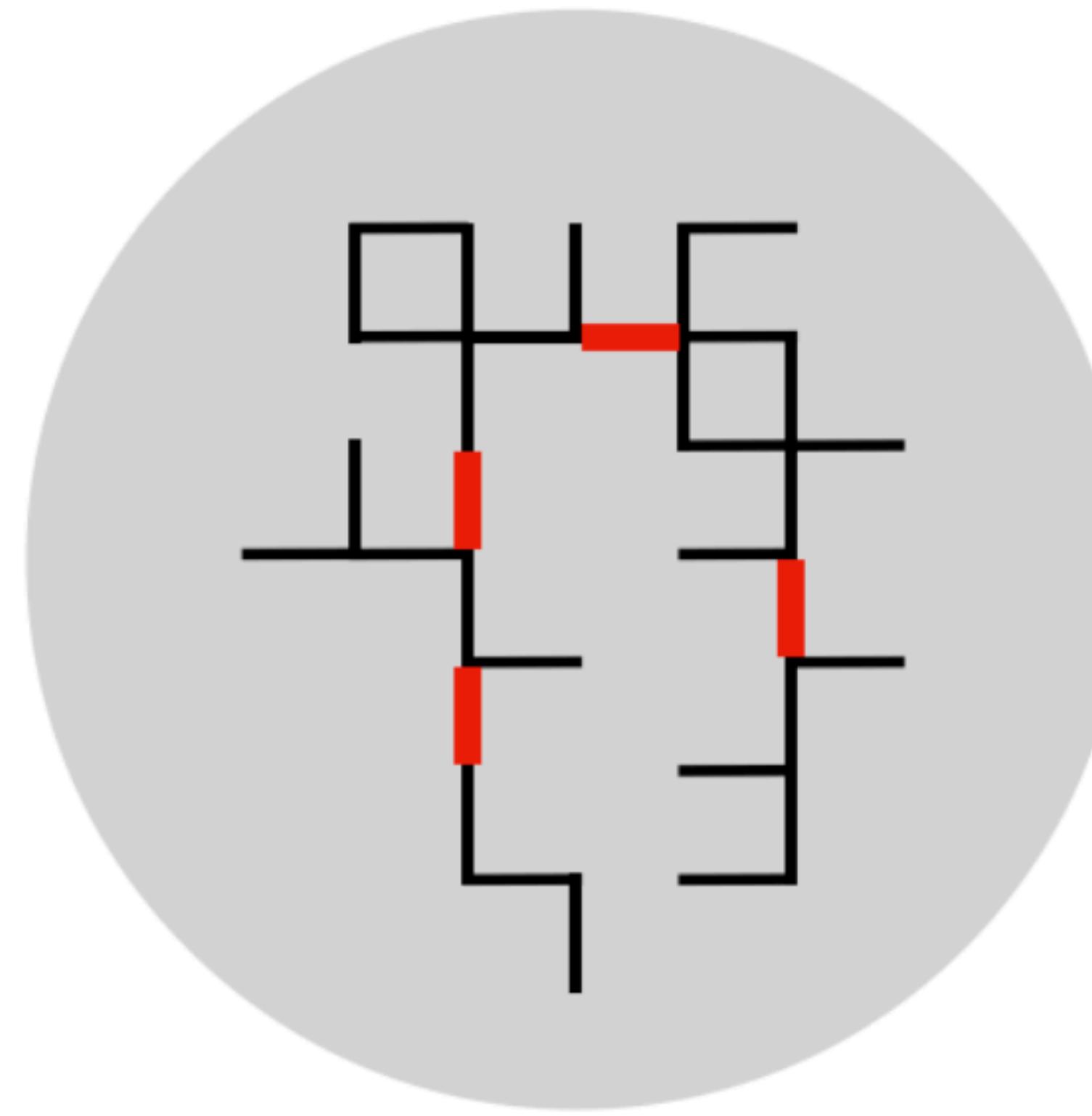


**Grow from scratch**

GrowBike.Net

Some cities

Developed but  
disconnected

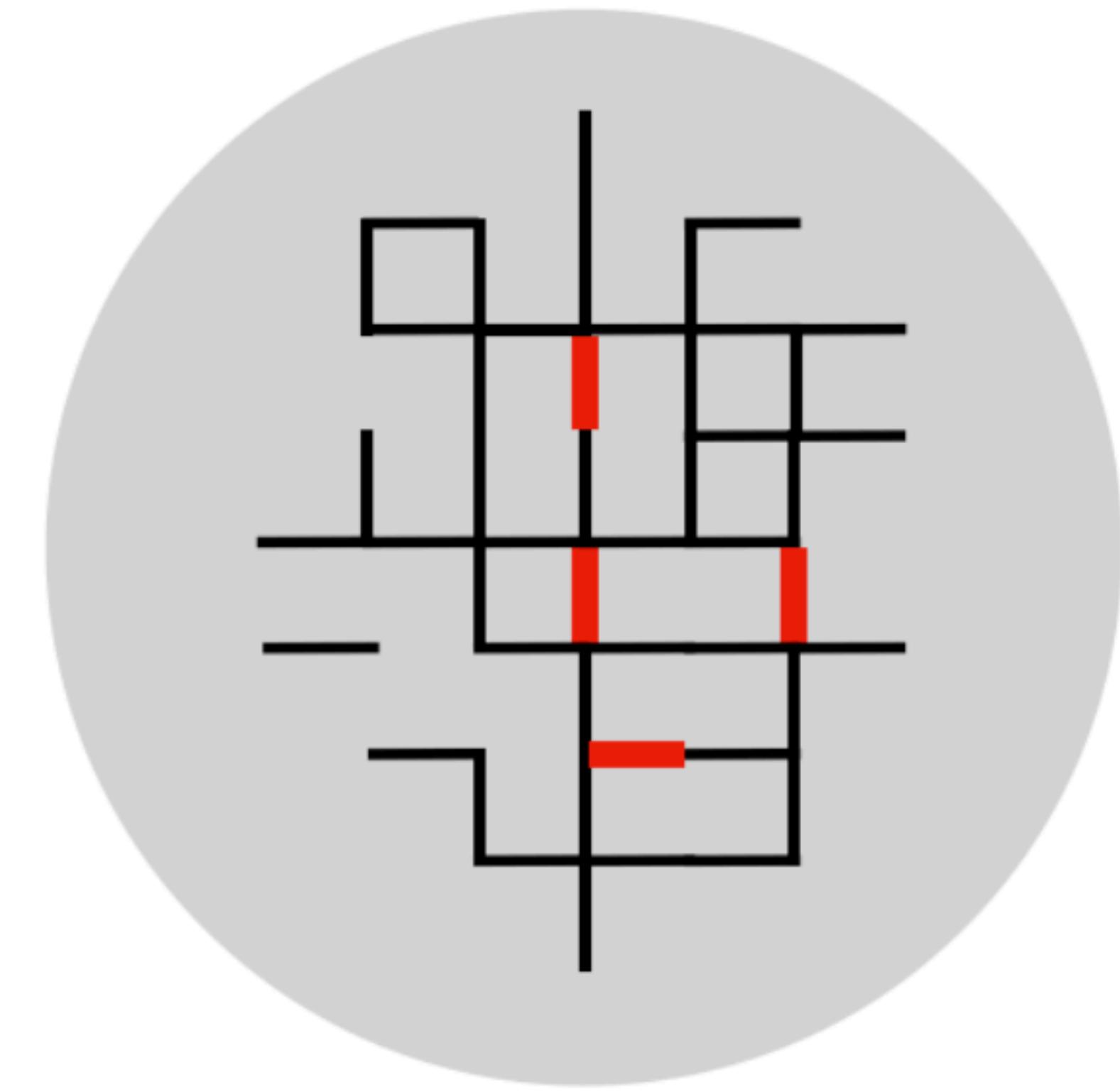


**Connect components**

LinkBike.Net

Few cities

Developed and  
mostly connected

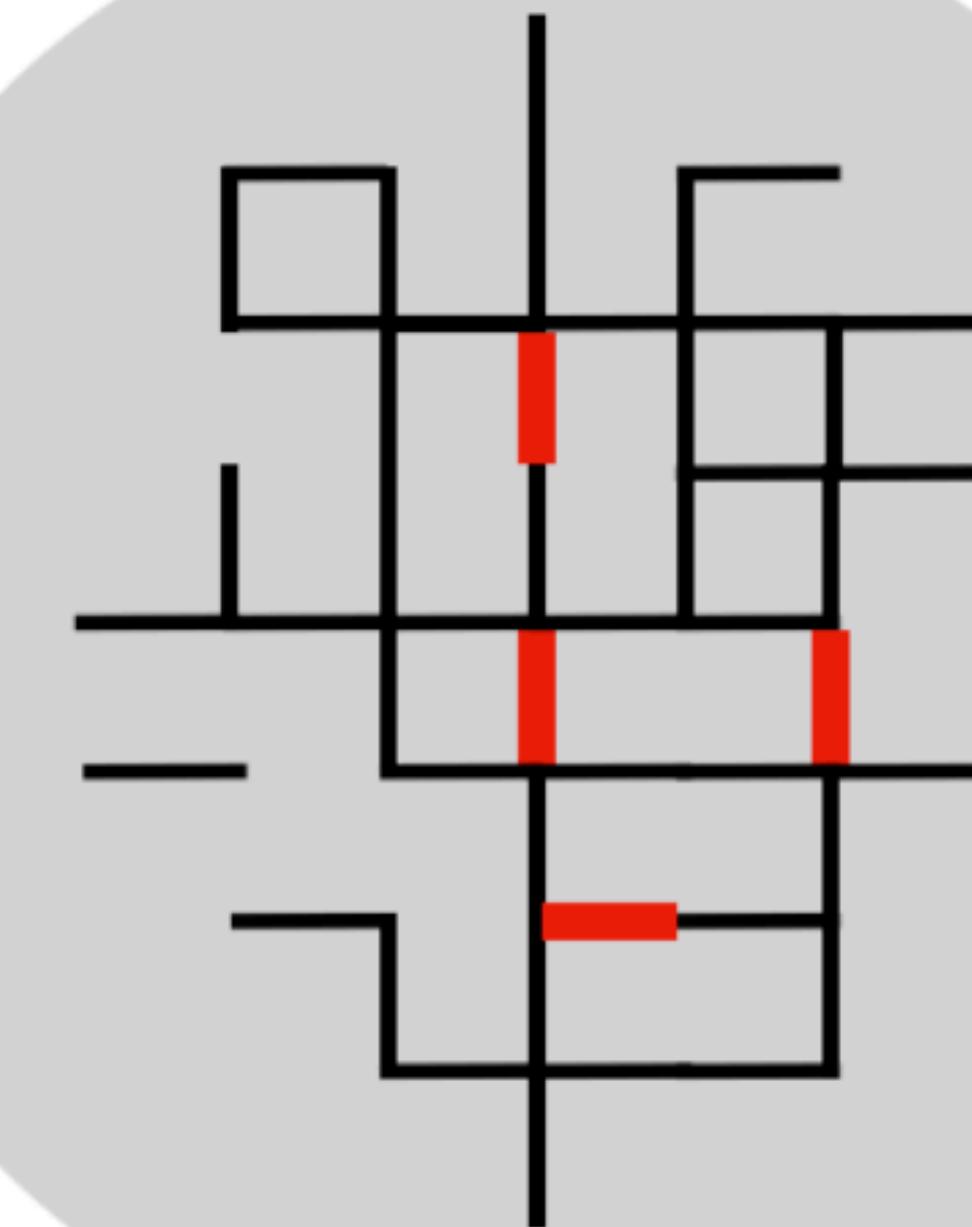


**Find missing links**

FixBike.Net

# Fixing bicycle networks

Few cities  
Developed and  
mostly connected



Find missing links  
FixBike.Net

# How to find the missing links in well-developed networks?

In Copenhagen, most of the network is 1 connected component.



# How to find the missing links in well-developed networks?

In Copenhagen, most of the network is 1 connected component.

Still, there are a lot of "missing links".

How to find them?  
How to prioritize them?



# From map to gap: IPDC

- 1) **I**dentify gaps
- 2) **P**rioritize gaps
- 3) **D**ecluster gaps
- 4) **C**lassify gaps



# 1) Identify: We need a formal definition of “gap”

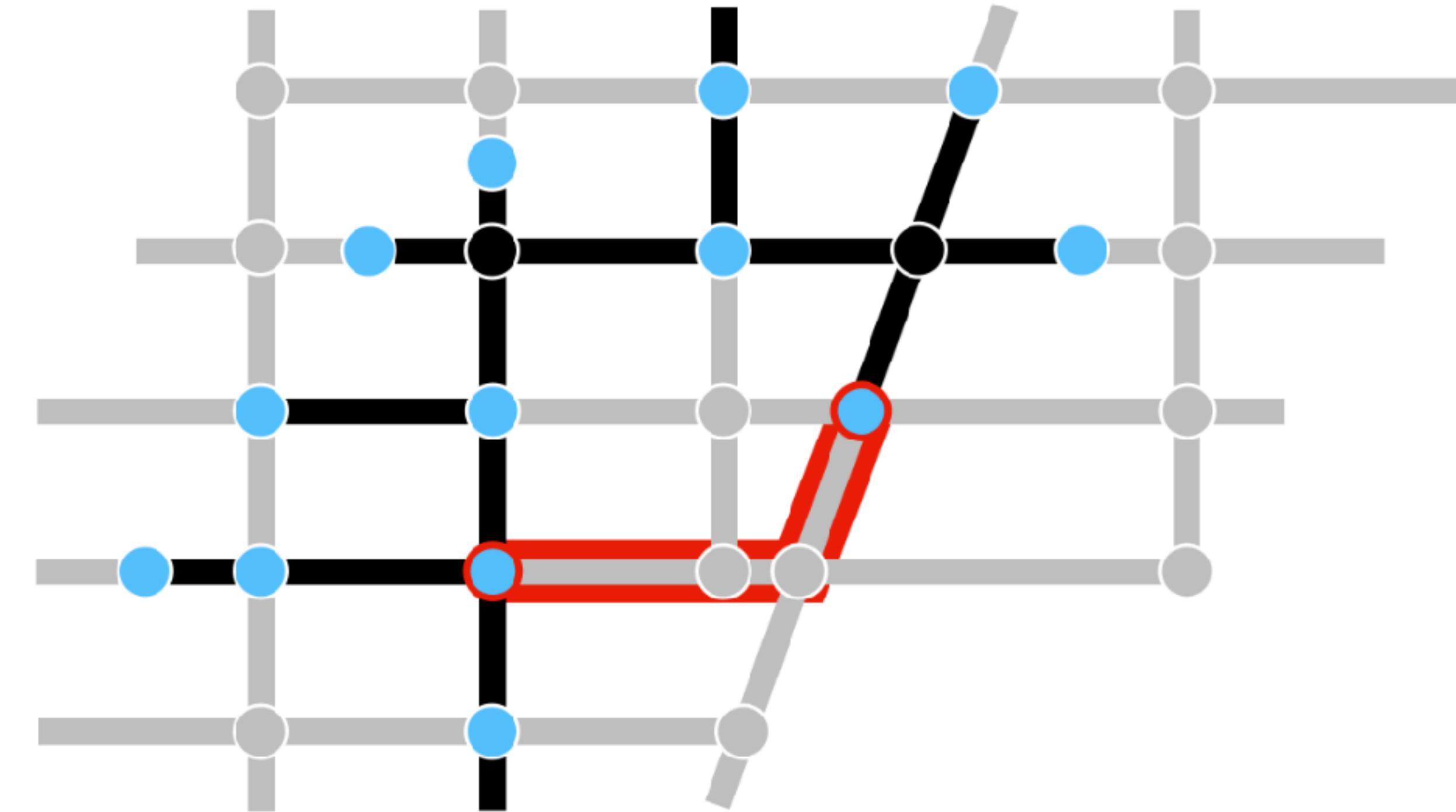
## Multiplex network

### Links

- 1) unprotected
- 2) protected

### Nodes

- 1) unprotected
- 2) protected
- 3) contact



A **gap** is a shortest path between two **contact nodes** that consists only of **unprotected links**

## 2) Prioritize

We could find millions of gaps...

We need a metric to prioritize them.



## 2) Prioritize



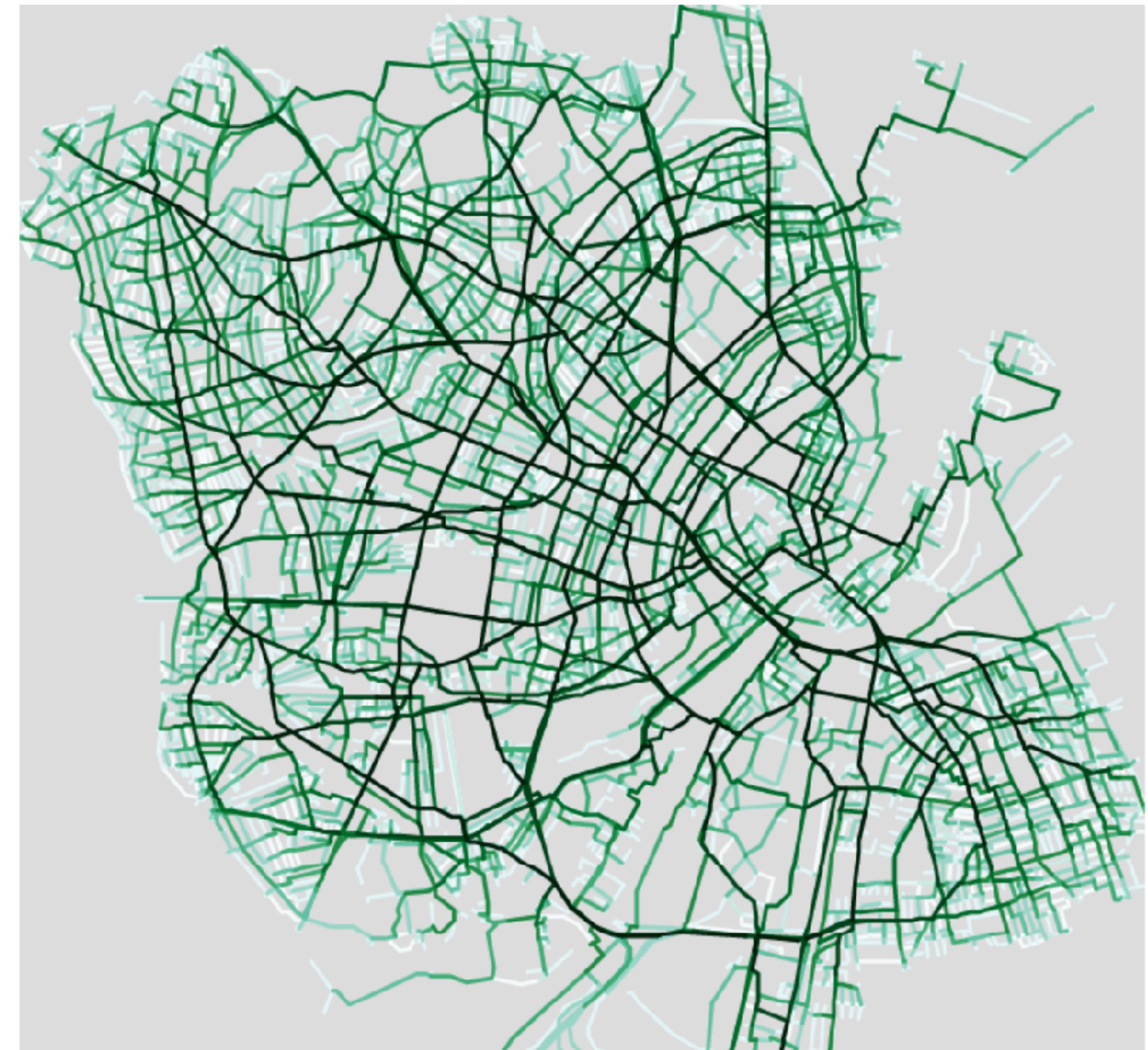
“If this gap was closed, how many meters cycled in mixed traffic would be avoided per investment unit?”



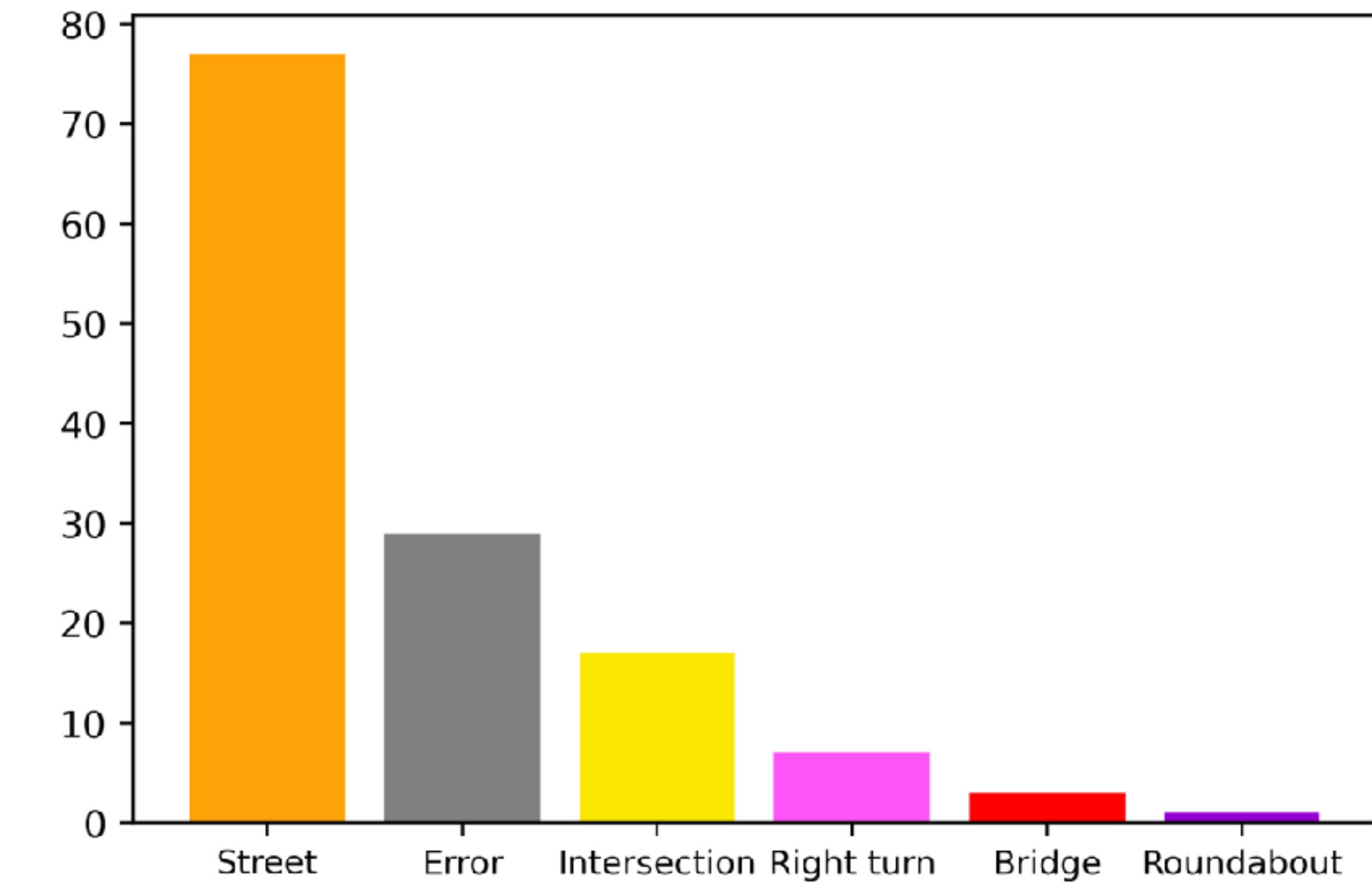
We can use betweenness centrality as a proxy for flow

$$C_B(i) = \sum_{i \neq j \neq k} \frac{\sigma_{jk}(i)}{\sigma_{jk}}$$

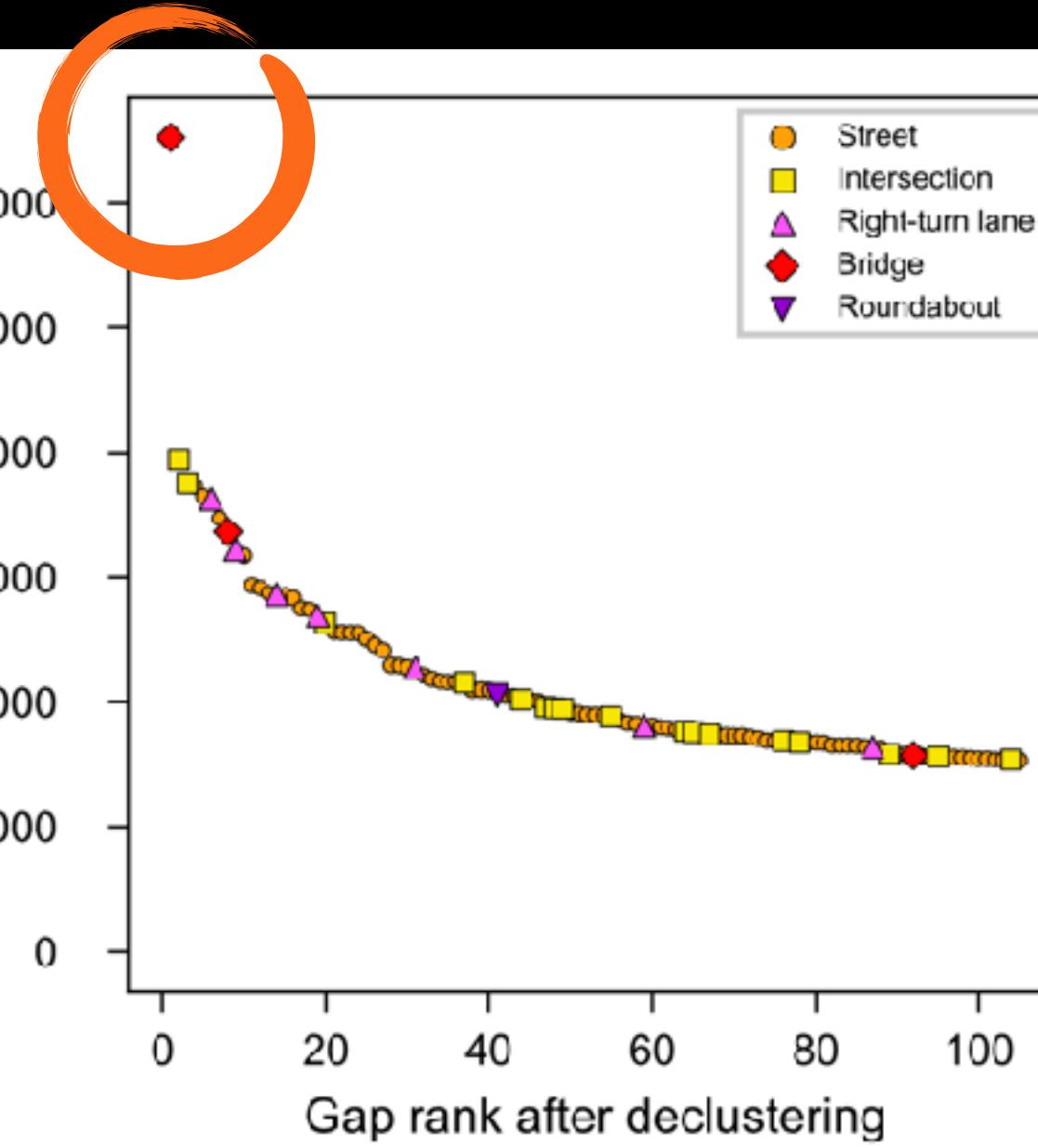
Cyclist flow data  
is hard to get



#### 4) Classify: Our top 105 gaps

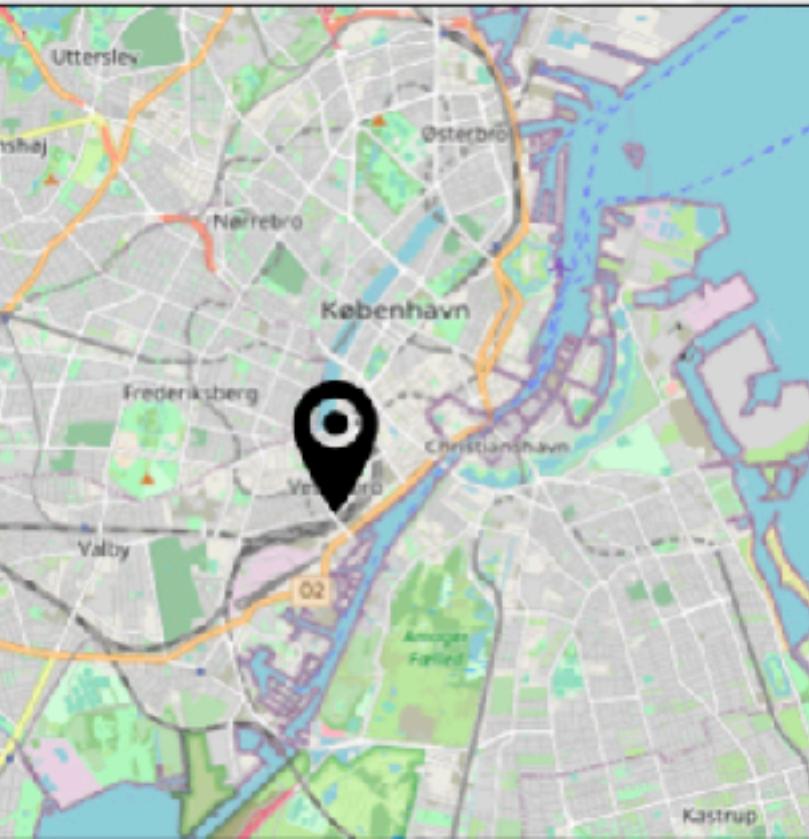
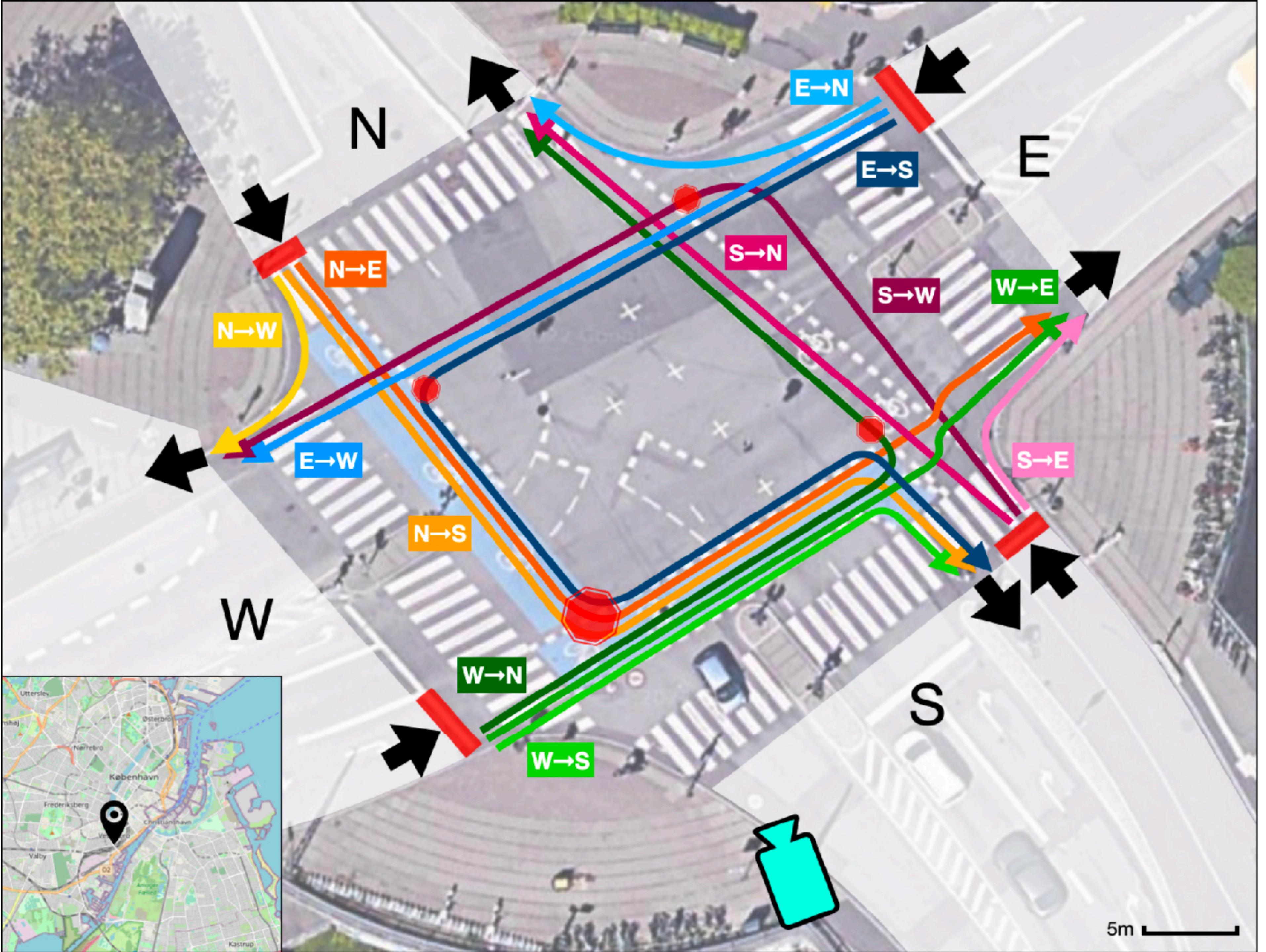


# The most important gaps are bridges

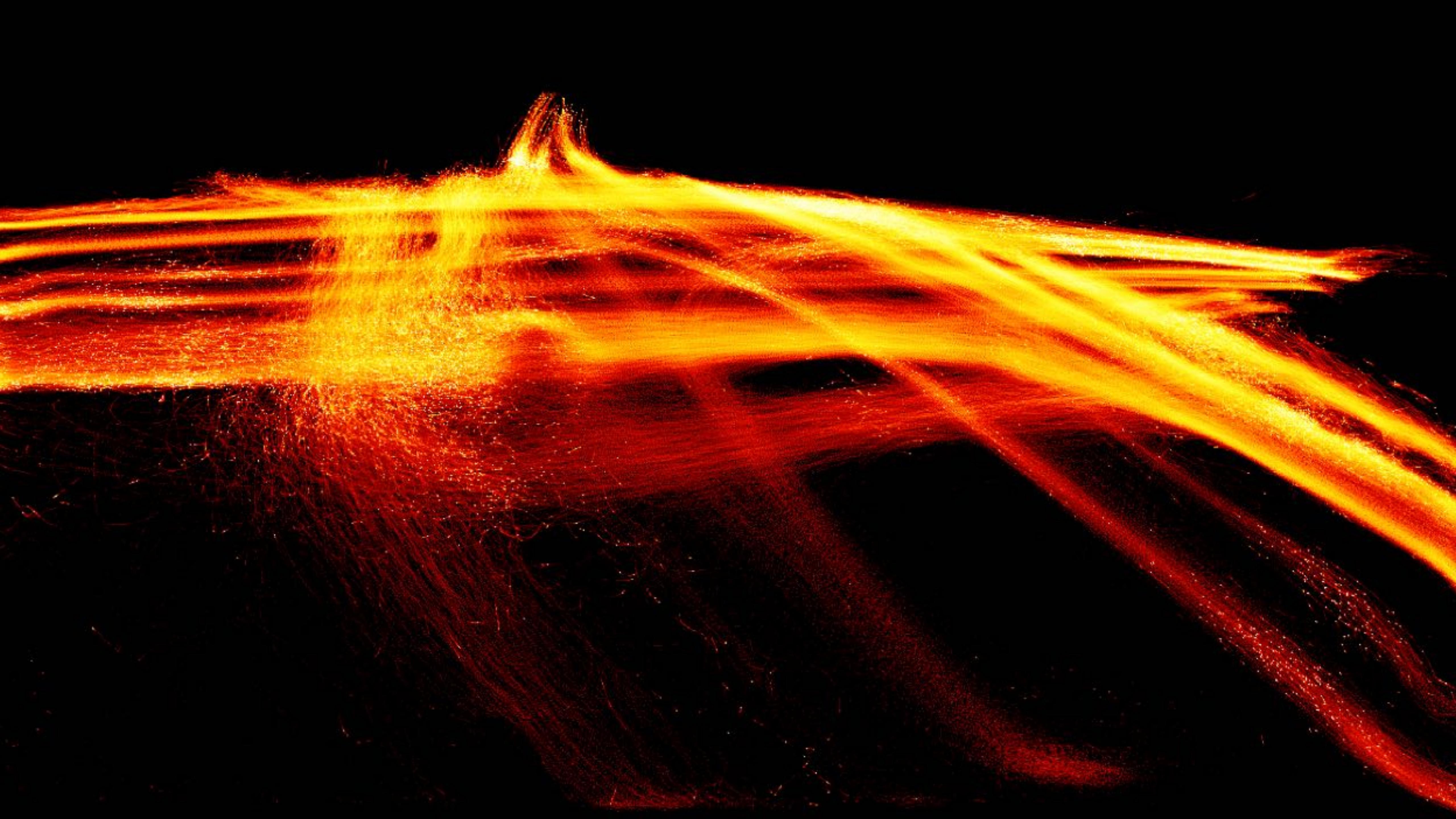


Knippelsbro

We lack data also  
on the microscopic level







# Automated methods have some problems

Heavy-weight



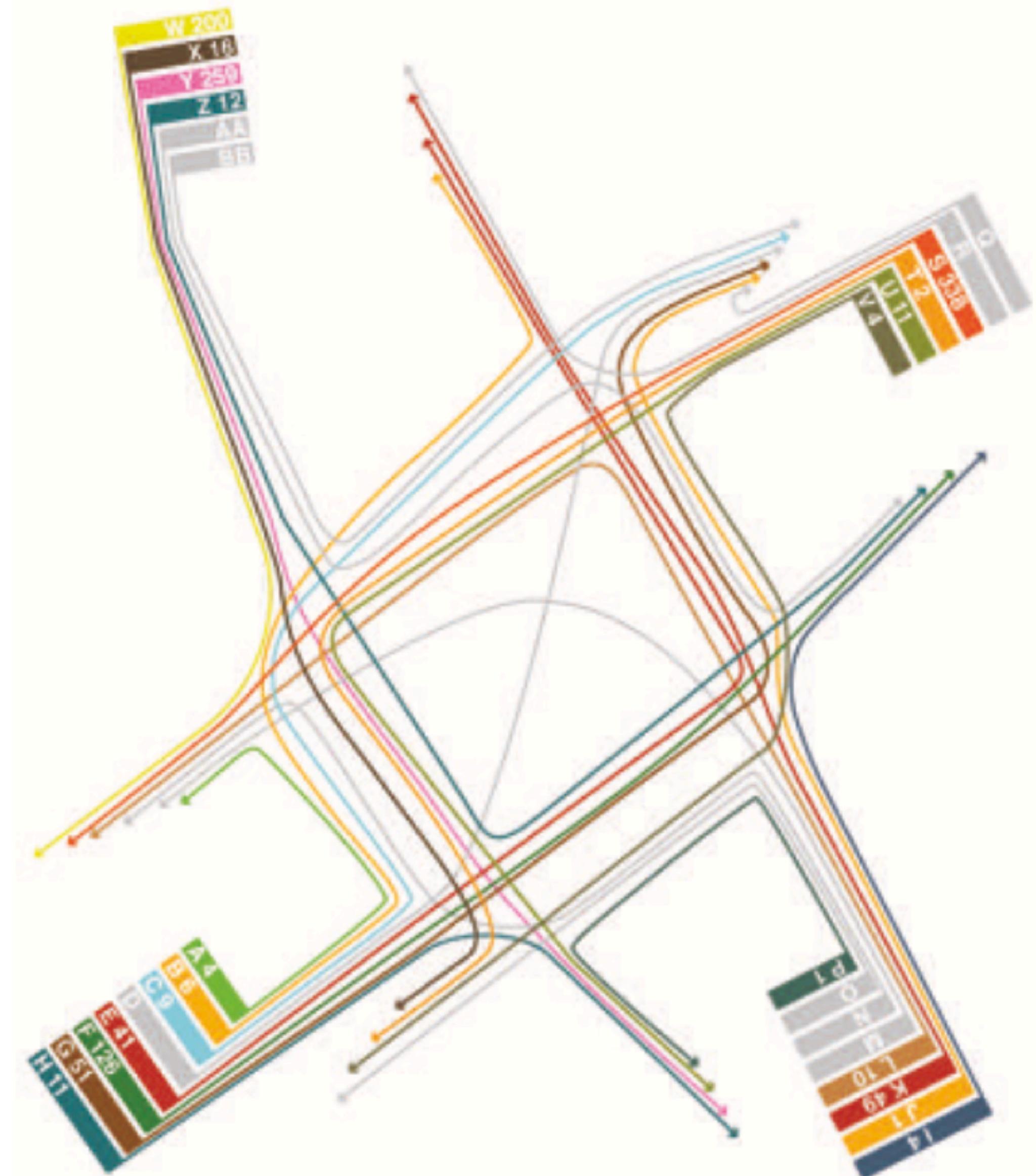
Car-centric



Focus on counting

# Desire line analysis is important

Uncovers discrepancy between cyclist behaviour and design

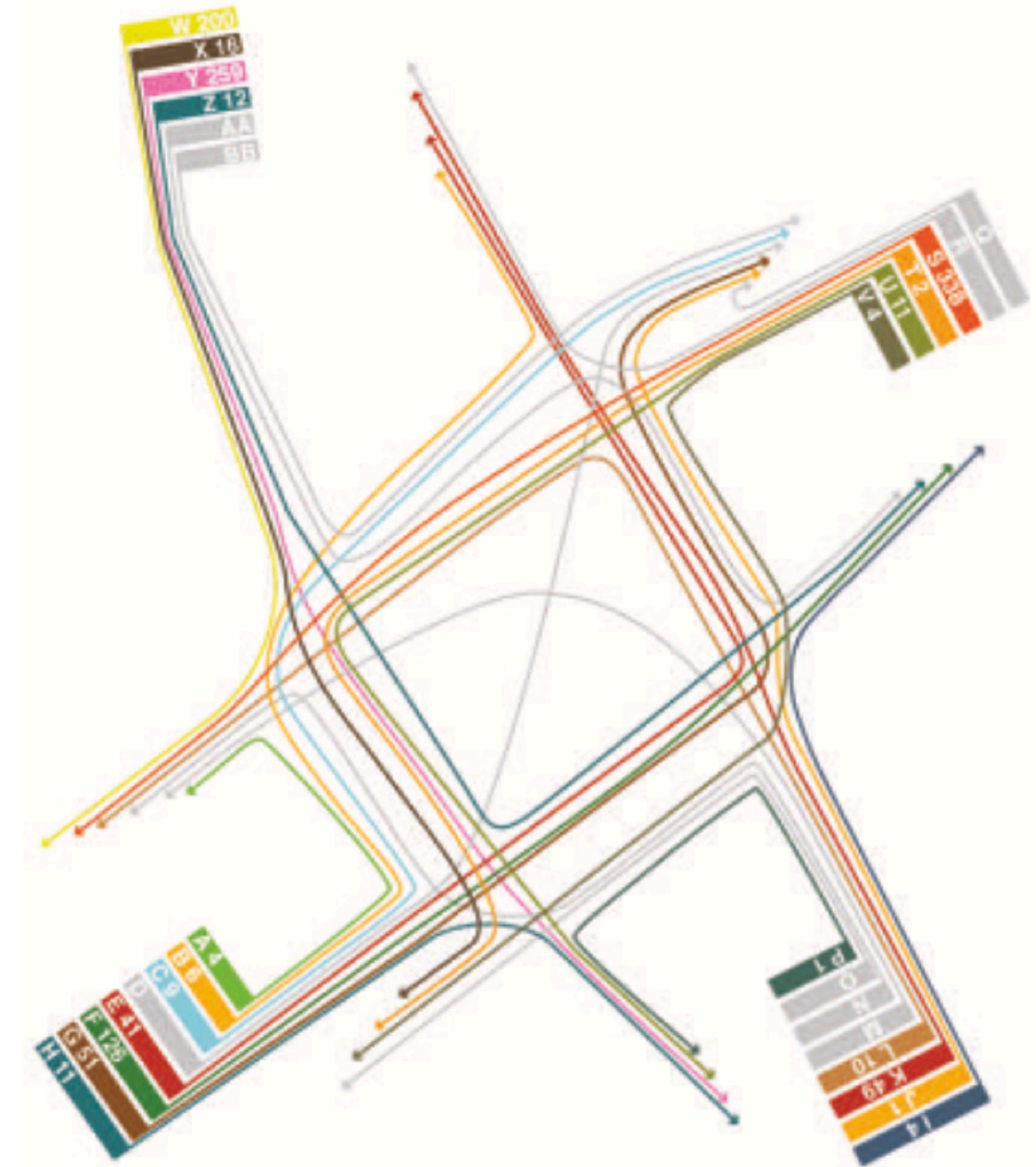


# Desire line analysis is manual, therefore **costly**

*What I learned from staring  
at 106,000 cyclists*



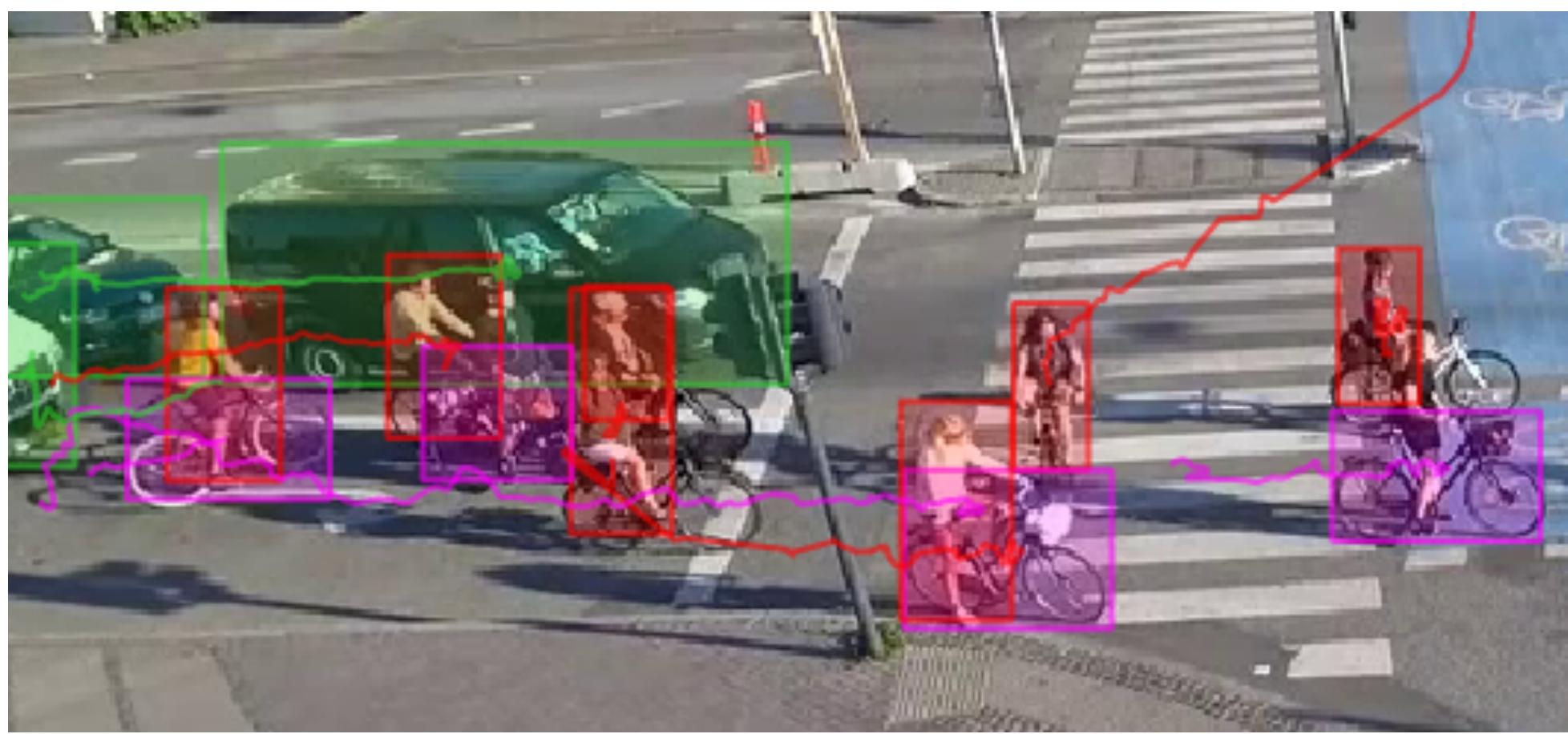
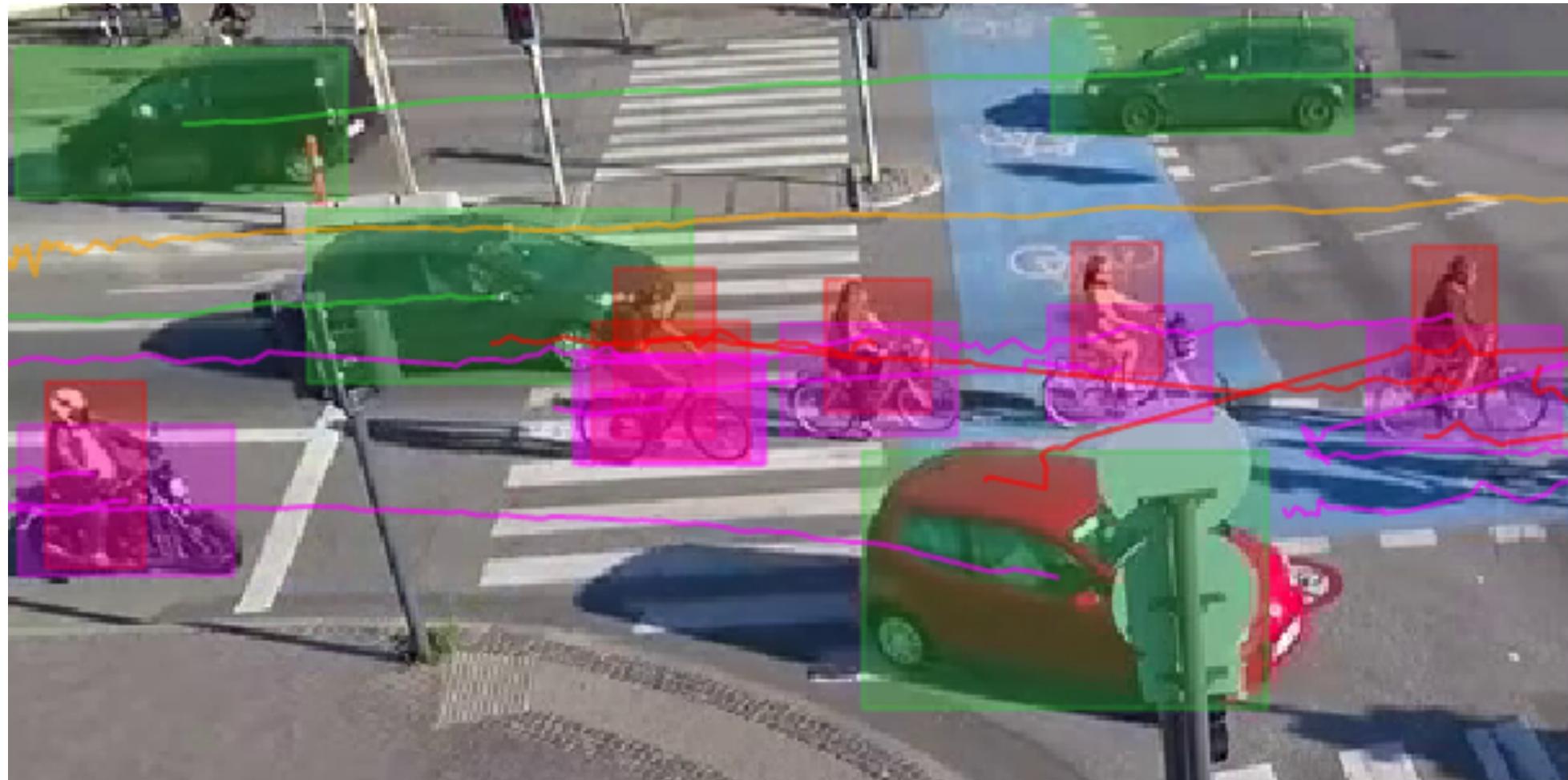
Can count paths, but not  
quantify them rigorously



Copenhagenize

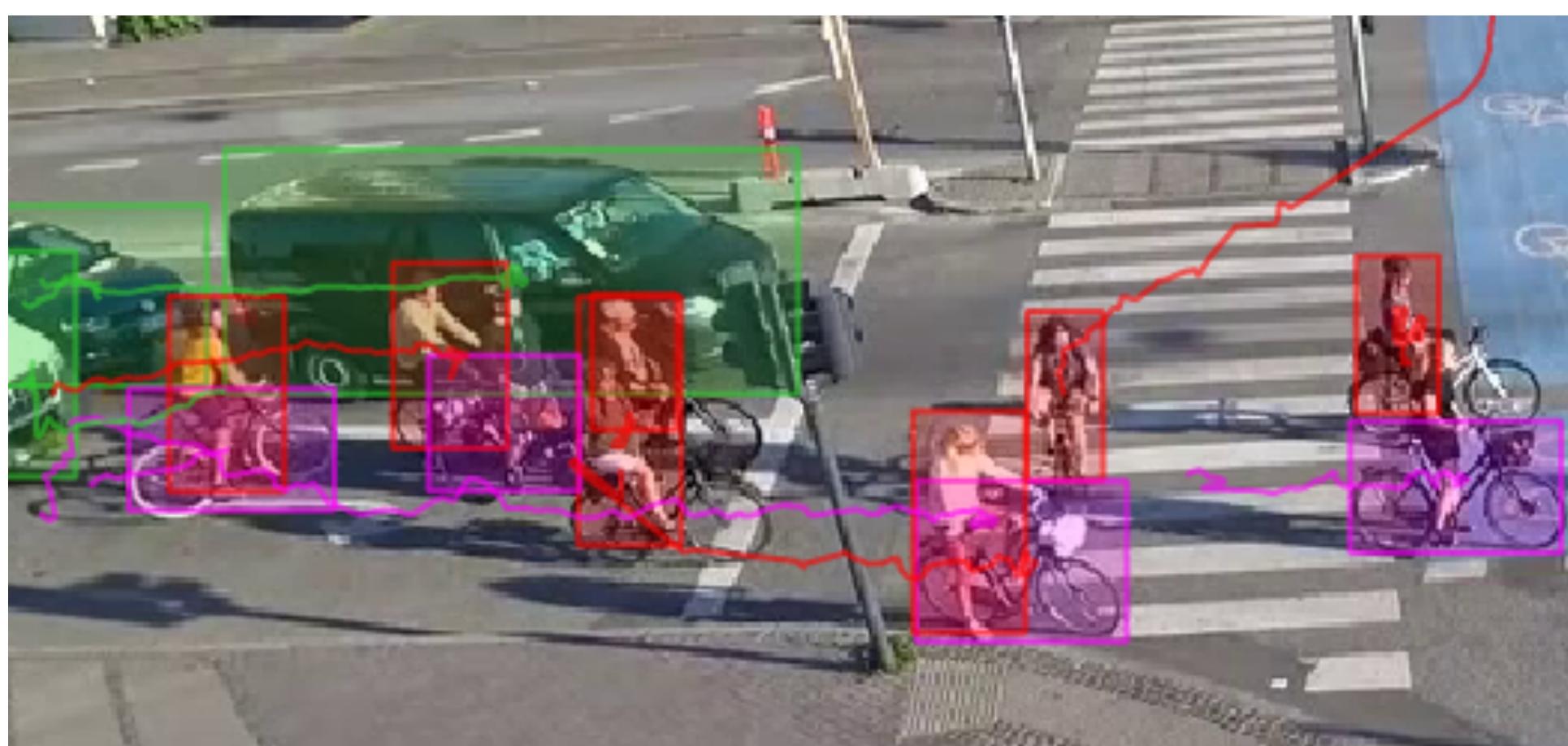
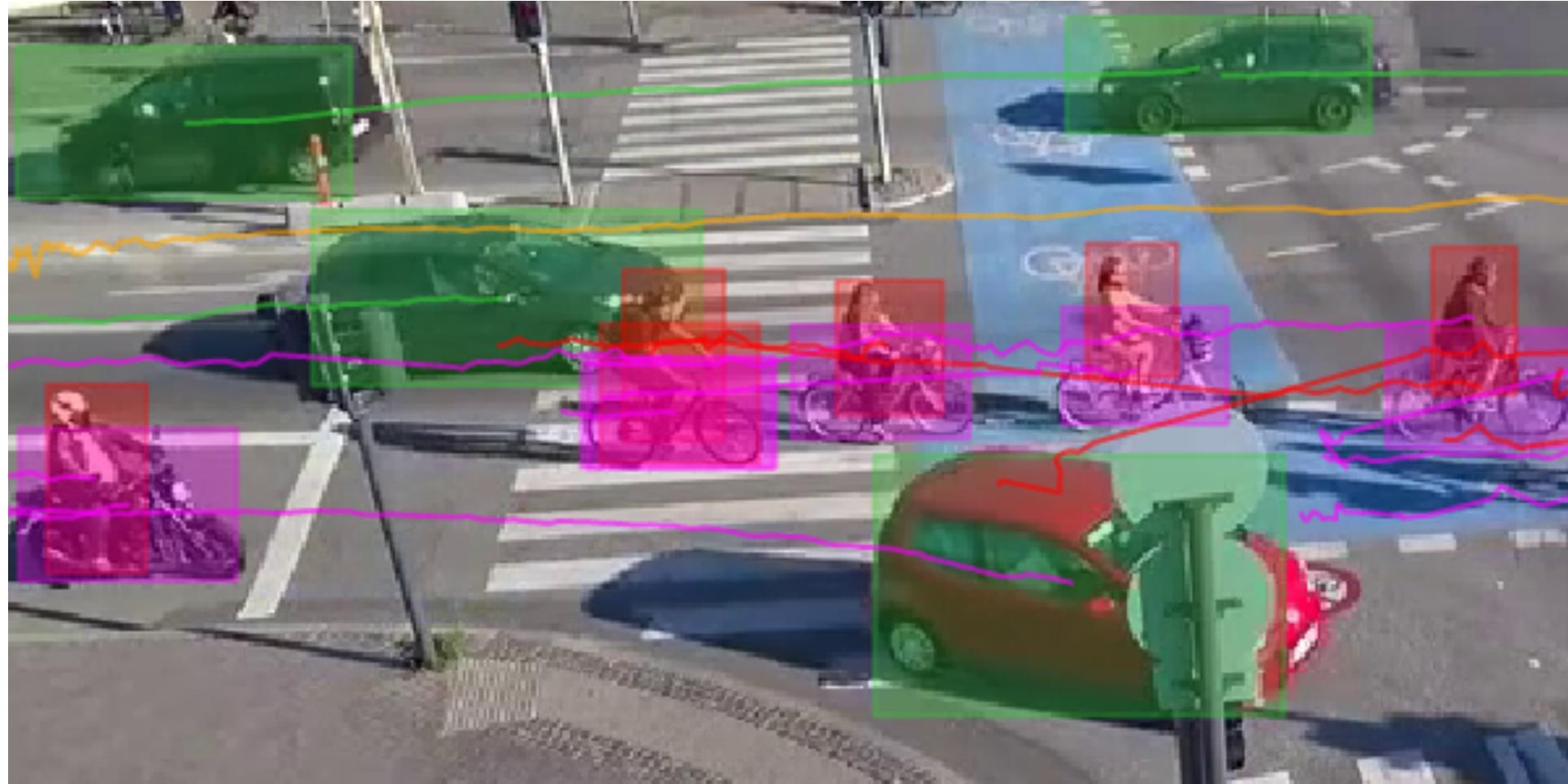
# To fix this, we trained a detection algorithm for cyclists

Good detection → easier tracking

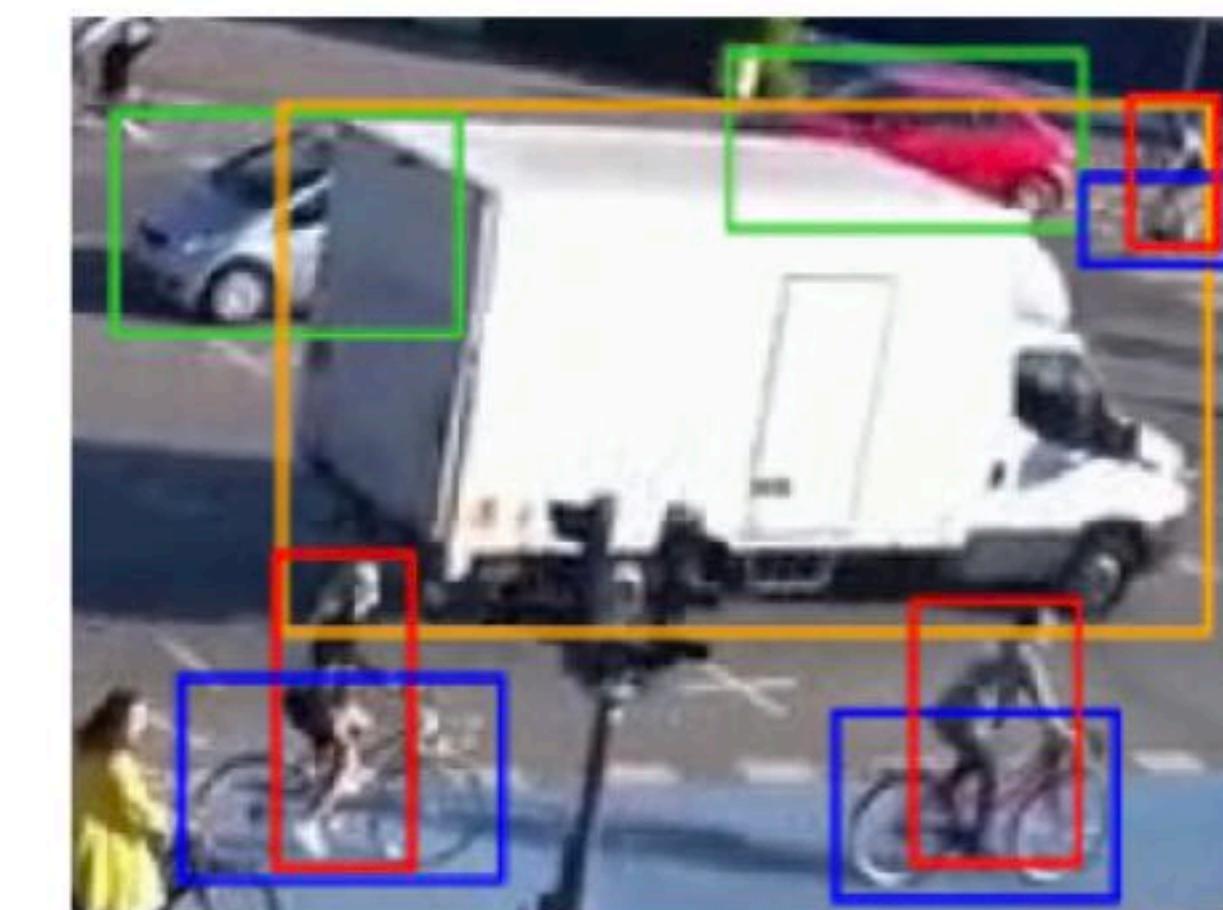


# To fix this, we trained a detection algorithm for cyclists

Good detection → easier tracking

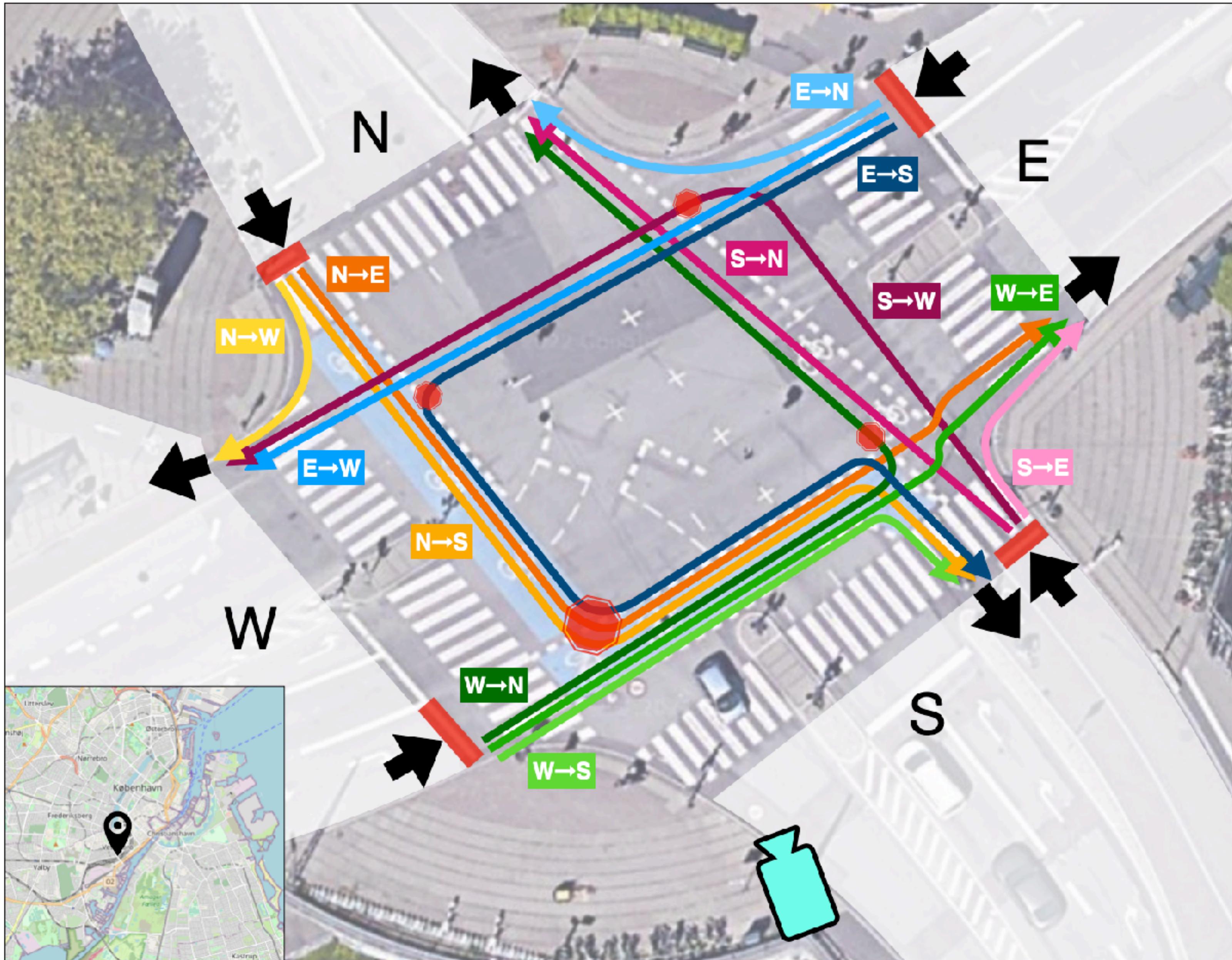


Bad detection → difficult tracking



There are still many issues...

# We spatially cluster trajectories into origin-destination clusters

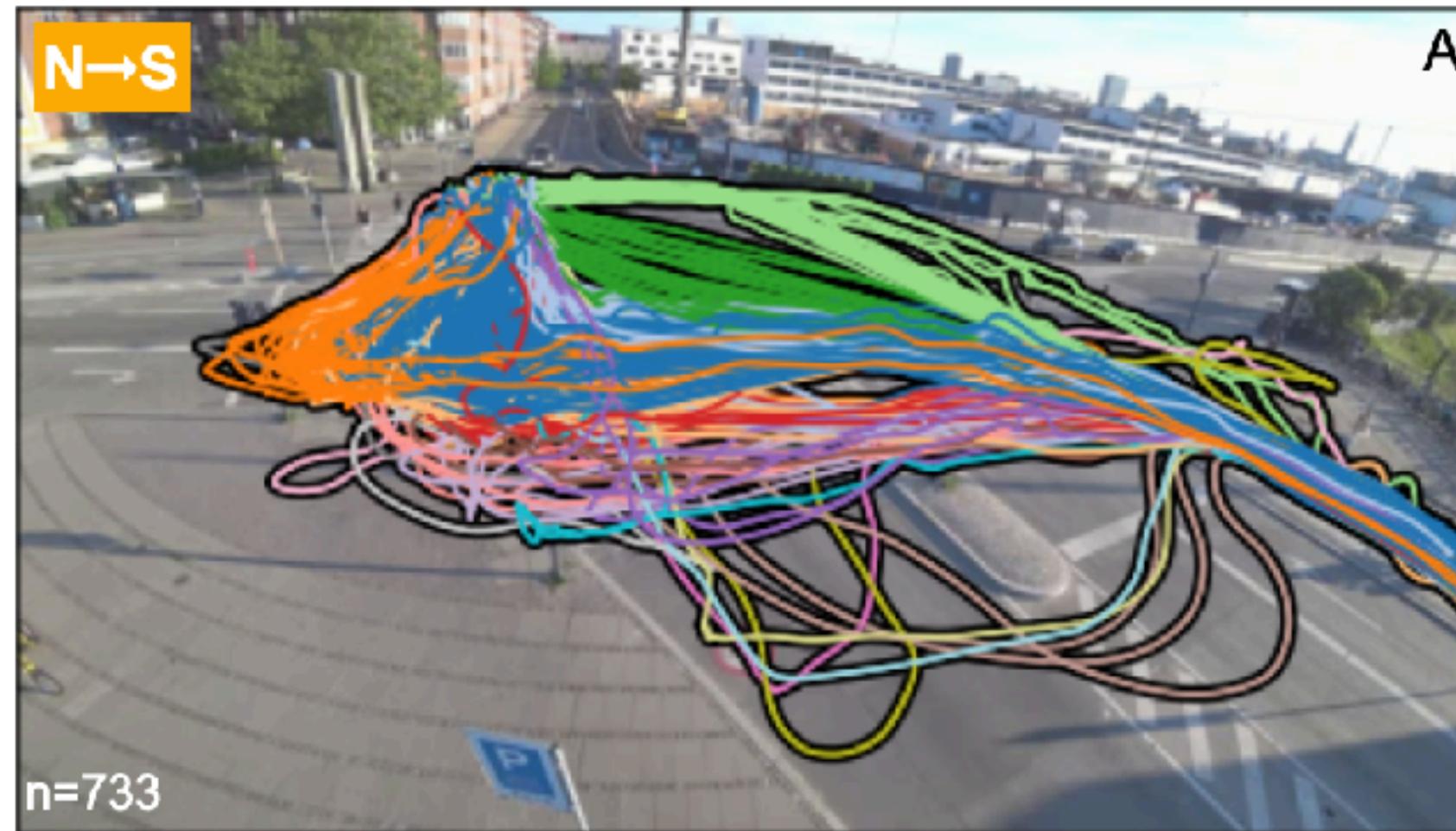


12 possible  
intended paths



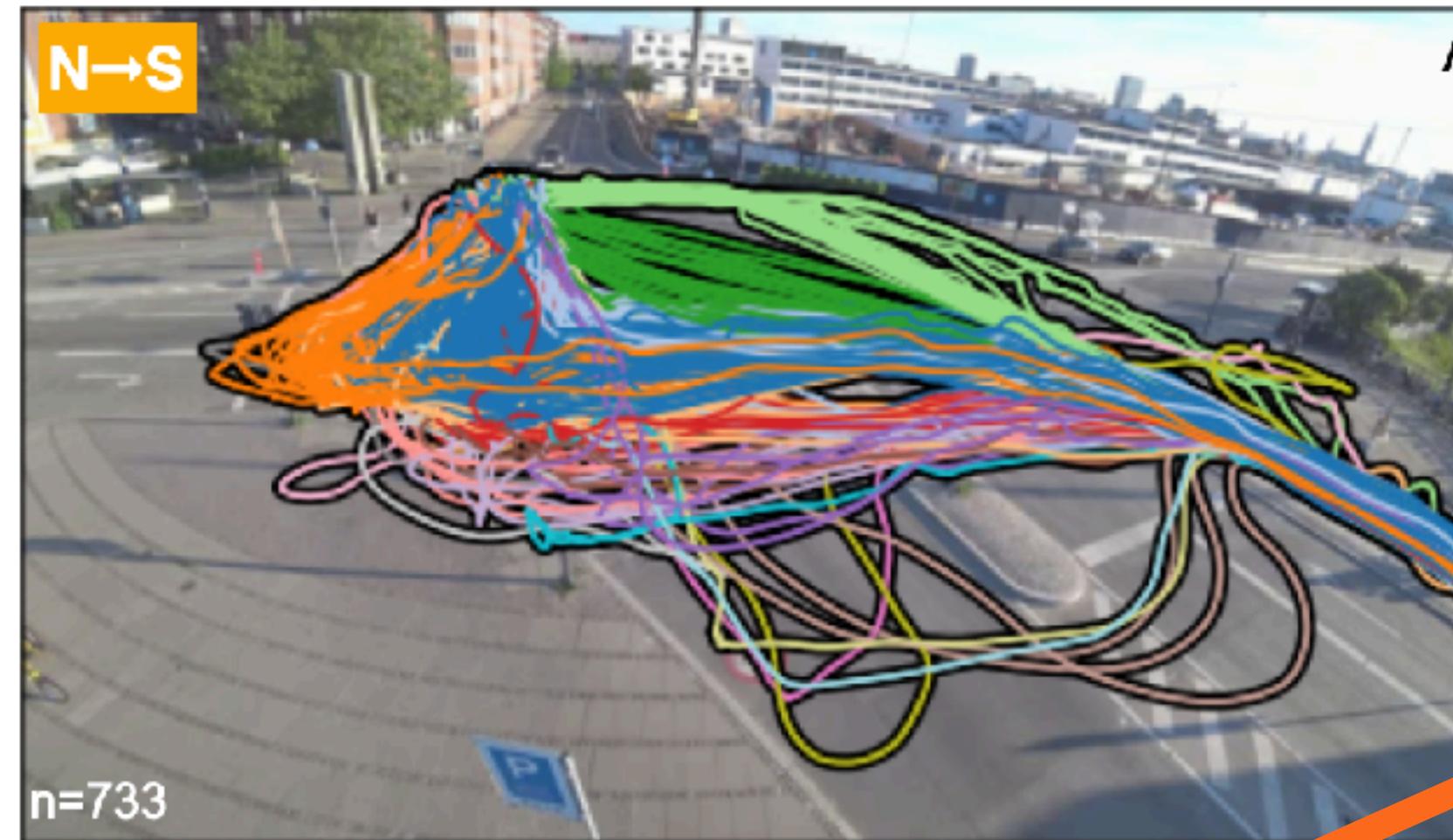
n=733

We separated trajectories into path-clusters with dynamic time warping



In just  
1 hour!

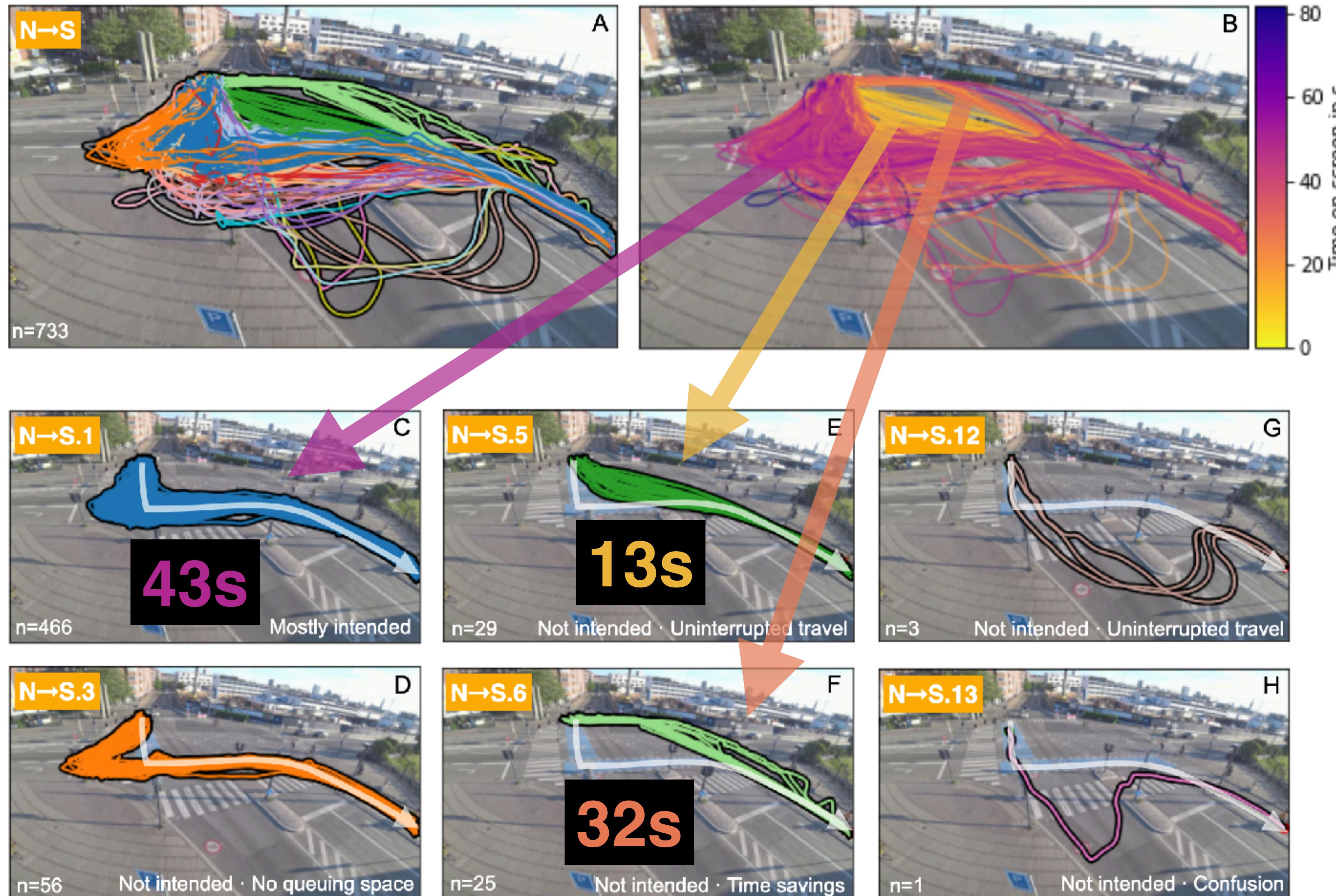
We separated trajectories into path-clusters with dynamic time warping



Only 466 are "mostly intended"  
Mismatch: 36%



# Cyclists prefer uninterrupted travel, which the intersection fails to provide





Breum, Simon Martin, Bojan Kostic, and Michael Szell. 2022. "Computational Desire Line Analysis of Cyclists on the Dybbølsbro Intersection in Copenhagen." *Findings*, December.

## TRANSPORT FINDINGS

# **Computational Desire Line Analysis of Cyclists on the Dybbølsbro Intersection in Copenhagen**

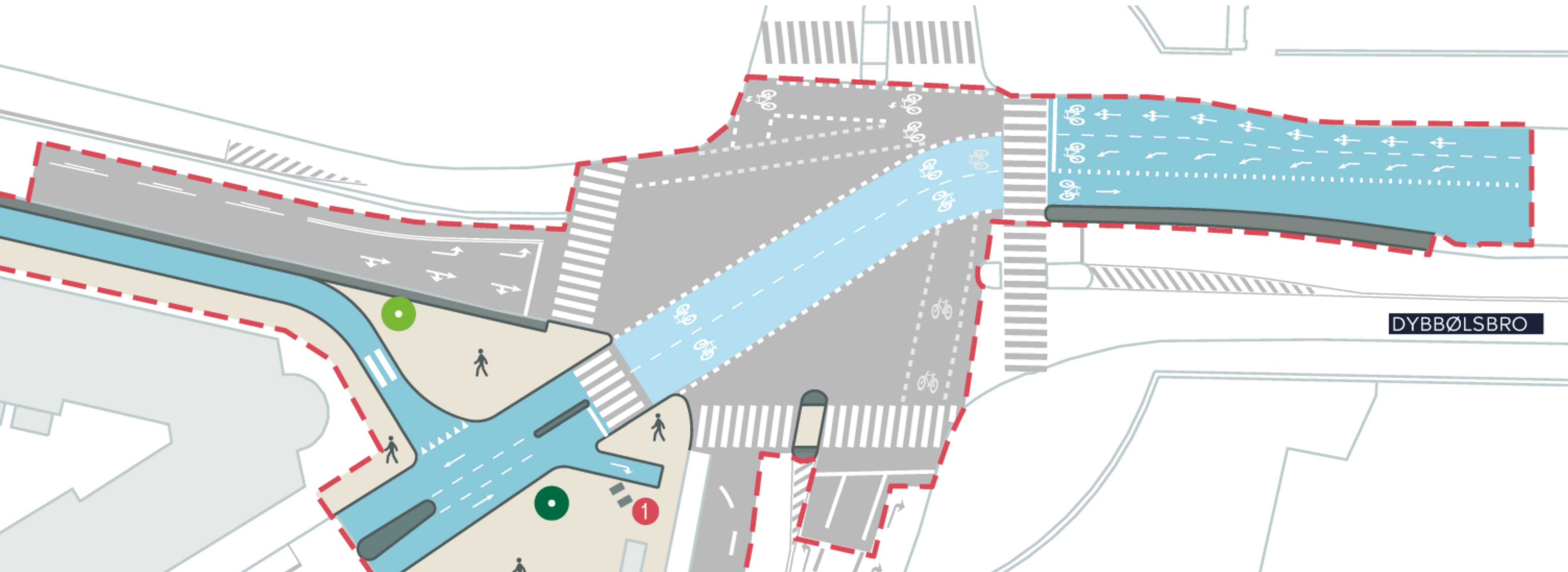
Simon Martin Breum<sup>1</sup> , Bojan Kostic<sup>1</sup> , Michael Szell<sup>1,2,3</sup>   <sup>a</sup>

<sup>1</sup> Computer Science, IT University of Copenhagen, <sup>2</sup> ISI Foundation, <sup>3</sup> Complexity Science Hub Vienna

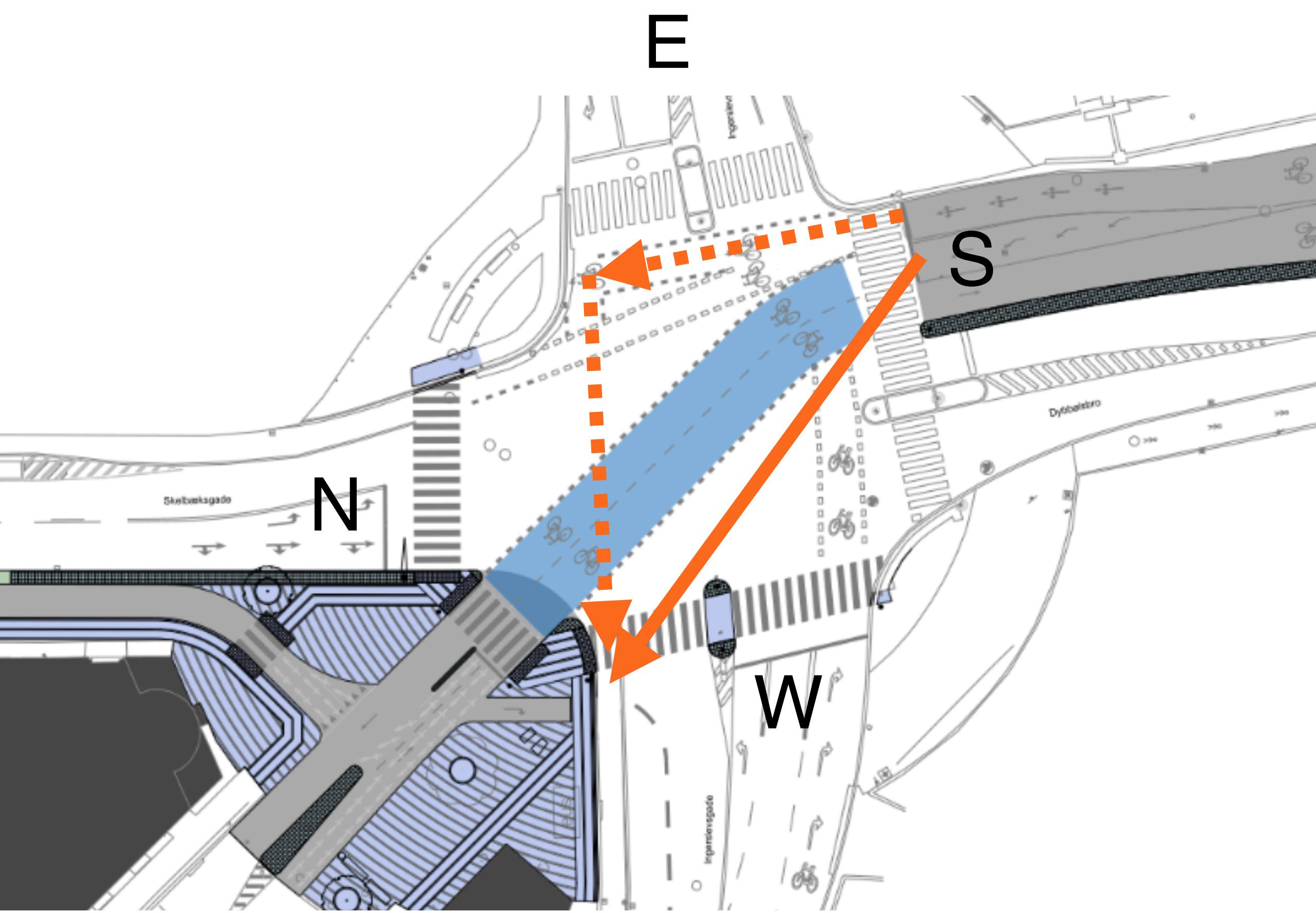
Keywords: urban data science, cycling, traffic behavior, intersection design, human-centric planning

<https://doi.org/10.32866/001c.56683>

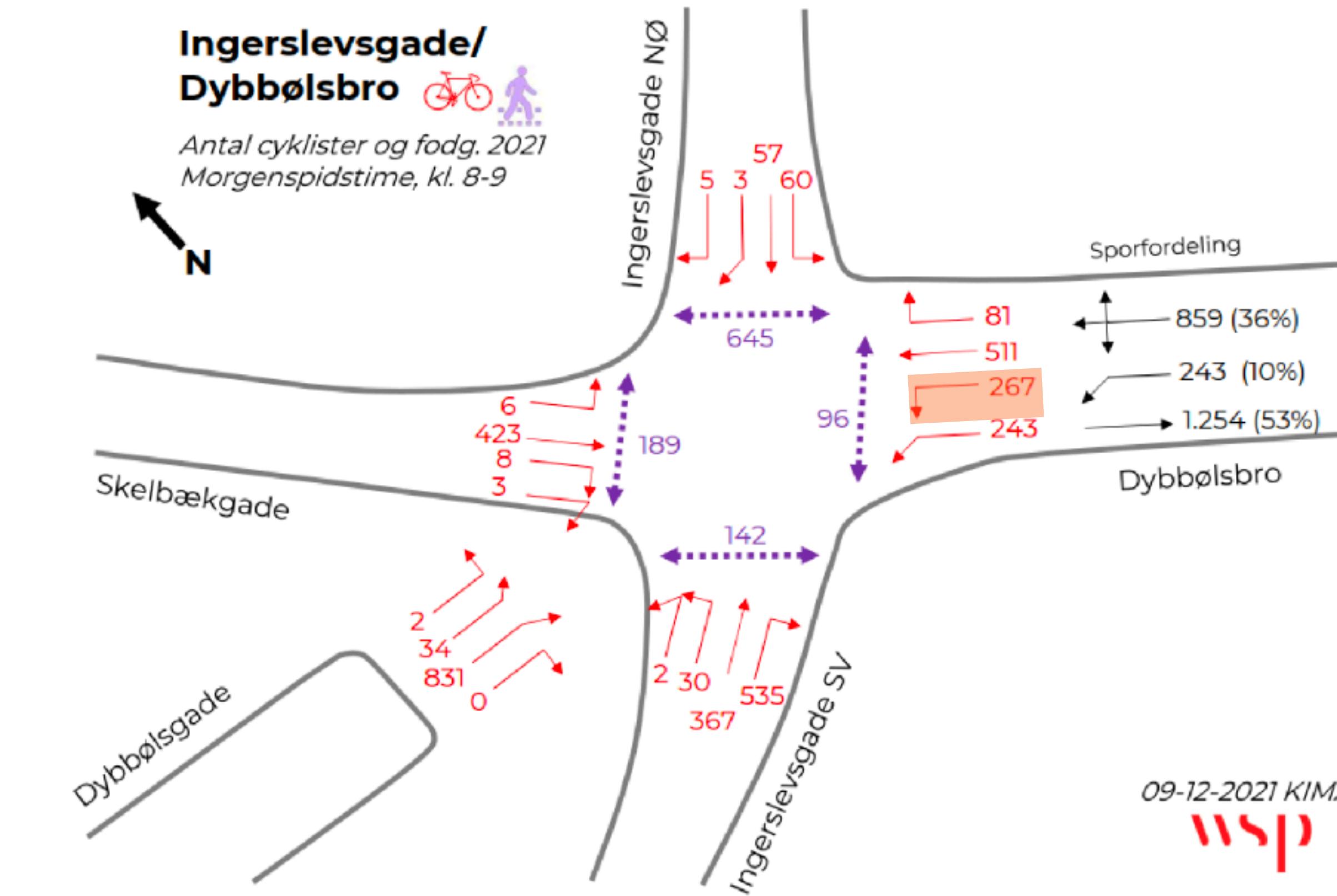
# The re-design



# Anticipated issues



- No solution for E→S
- New problem: S→W
- Increasing complexity (Yrsa plads)



Figur 11. Antal cyklister og fodgængere i morgenspidstimen. Trafiktælling 2021.

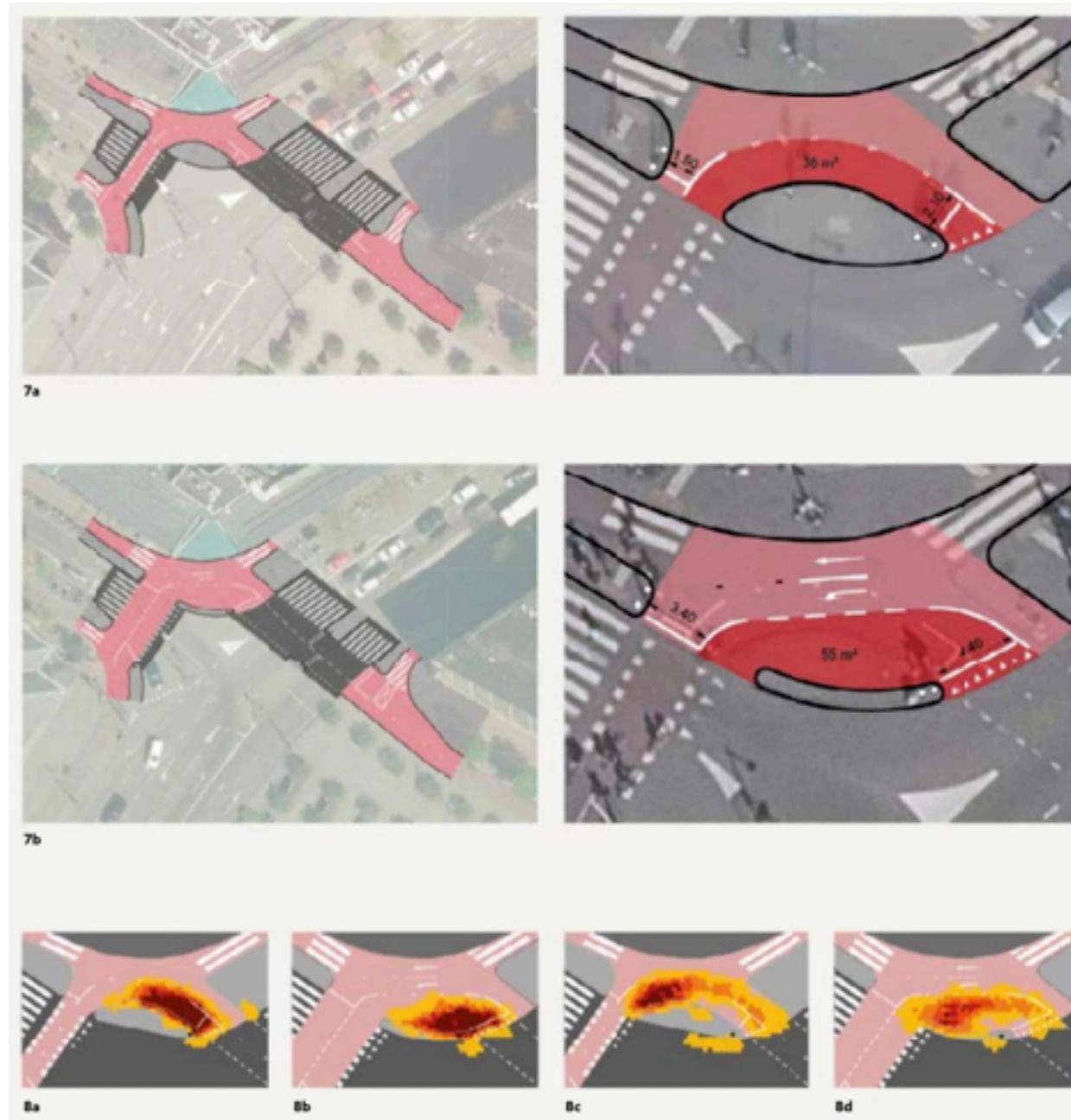
- Lack of queuing spaces
- Uni- to Bi-directional switch

The underlying issue is car-centric design

Why is vehicular traffic flow prioritized in CPH?  
Goes against science and international best practices

# The underlying issue is car-centric design

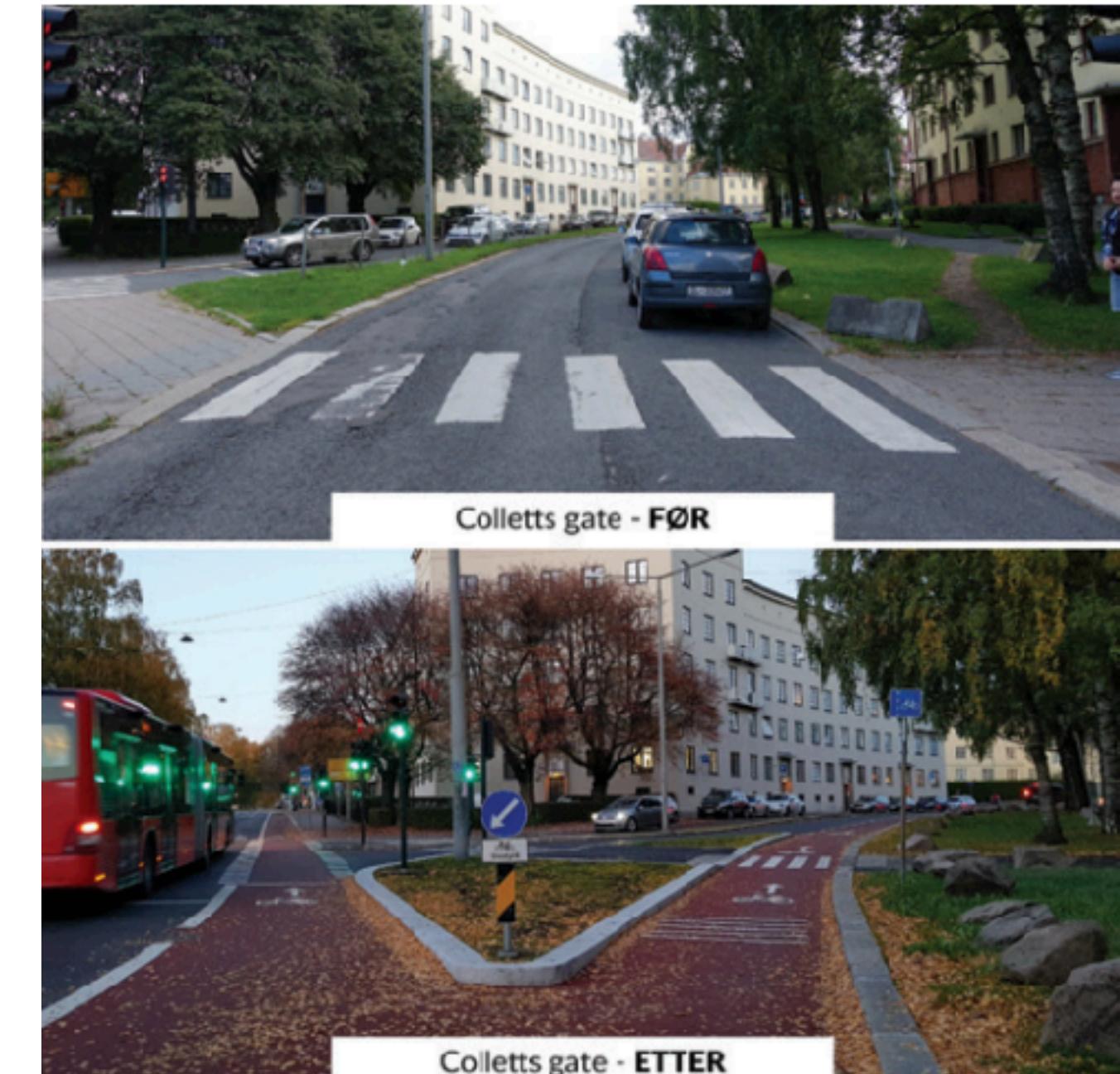
## Why is vehicular traffic flow prioritized in CPH? Goes against science and international best practices



Evidence-based "agile" design  
of crossing/queuing spaces



Removal of traffic lights on  
Alexanderplein



Prioritizing cycling  
desire lines (Oslo)

Drastic speed reductions to protect people  
from massive public health hazard of cars

Hahn & te Broemmelstroet, Transp Res Int P 9, 100289 (2021)  
Gemeente Amsterdam (2018)  
Hartmann & Abel, ITE J 90(5), 32-38 (2020)  
Klanjcic, Gavin, Tizzoni, Szell, EPJ Data Sci 11, 27 (2022)

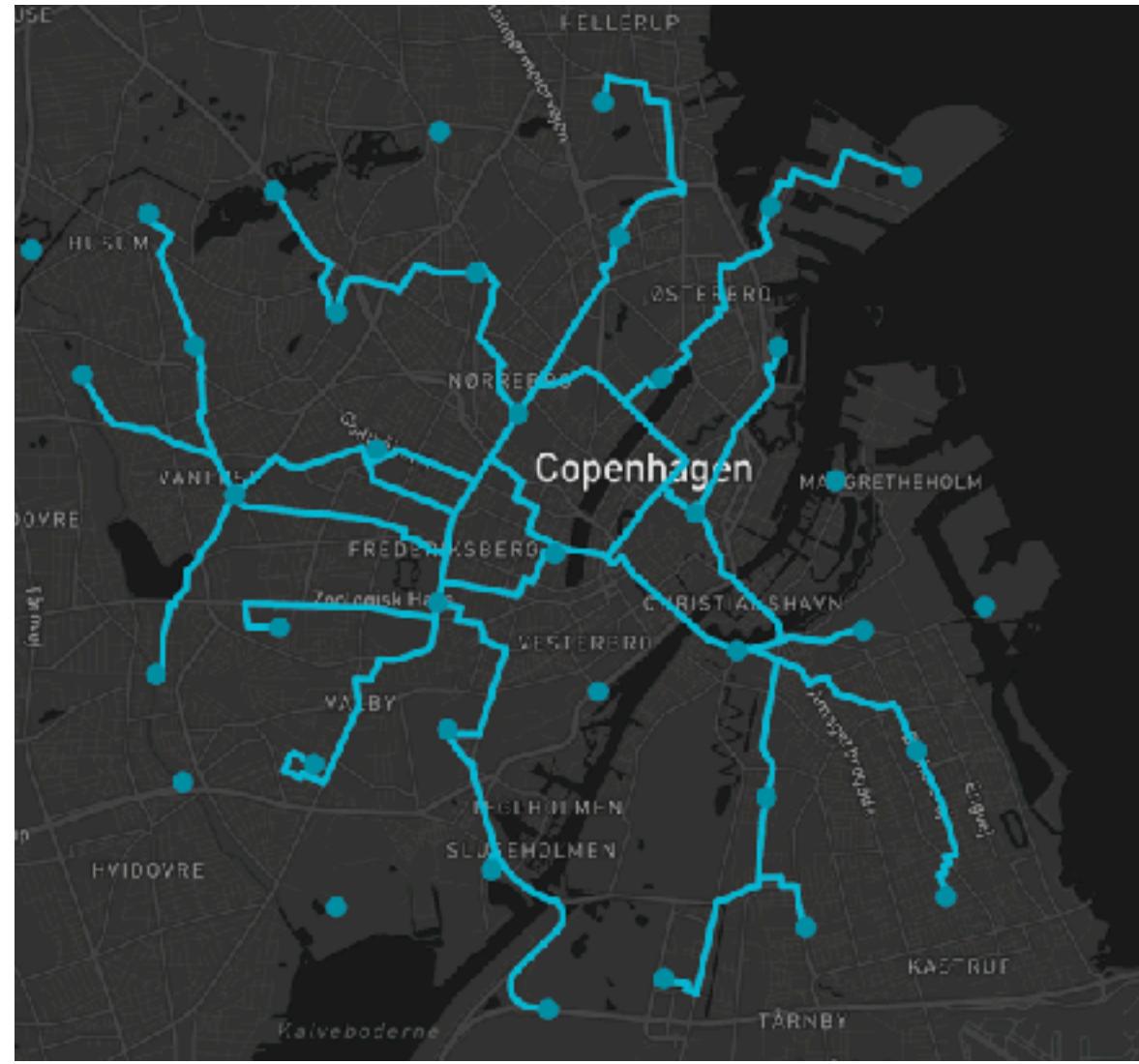
# We lack data+research on

- Cyclist traffic behavior
- Cyclist flows
- Bicycle networks
- ...

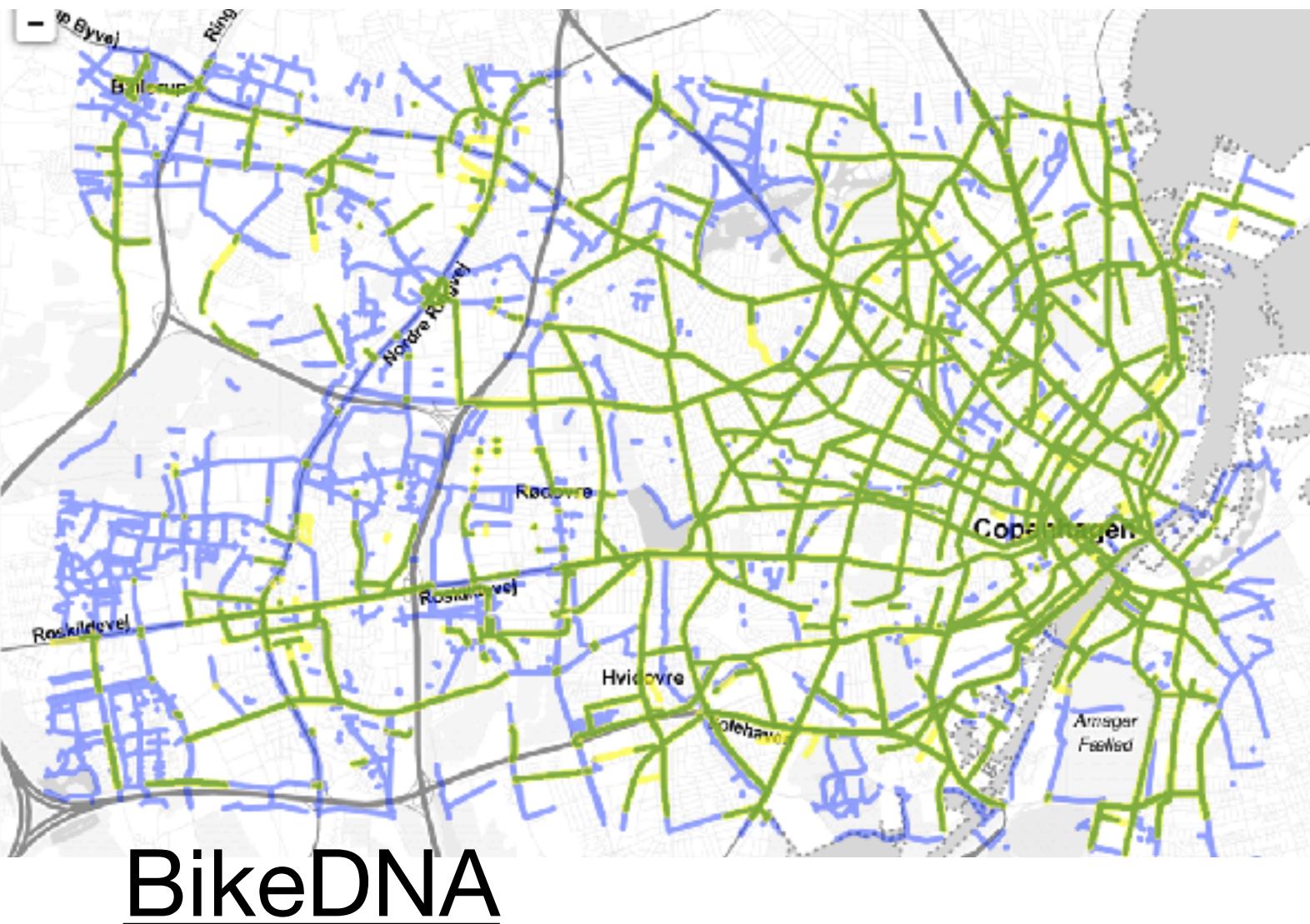
but..

..Building sustainable cities is a  
**political**, not a technical question

# Our work: Data-driven tools to help sustainable urban planning



Grow bicycle networks



BikeDNA

Geospatial Data Science (Spring 2022)

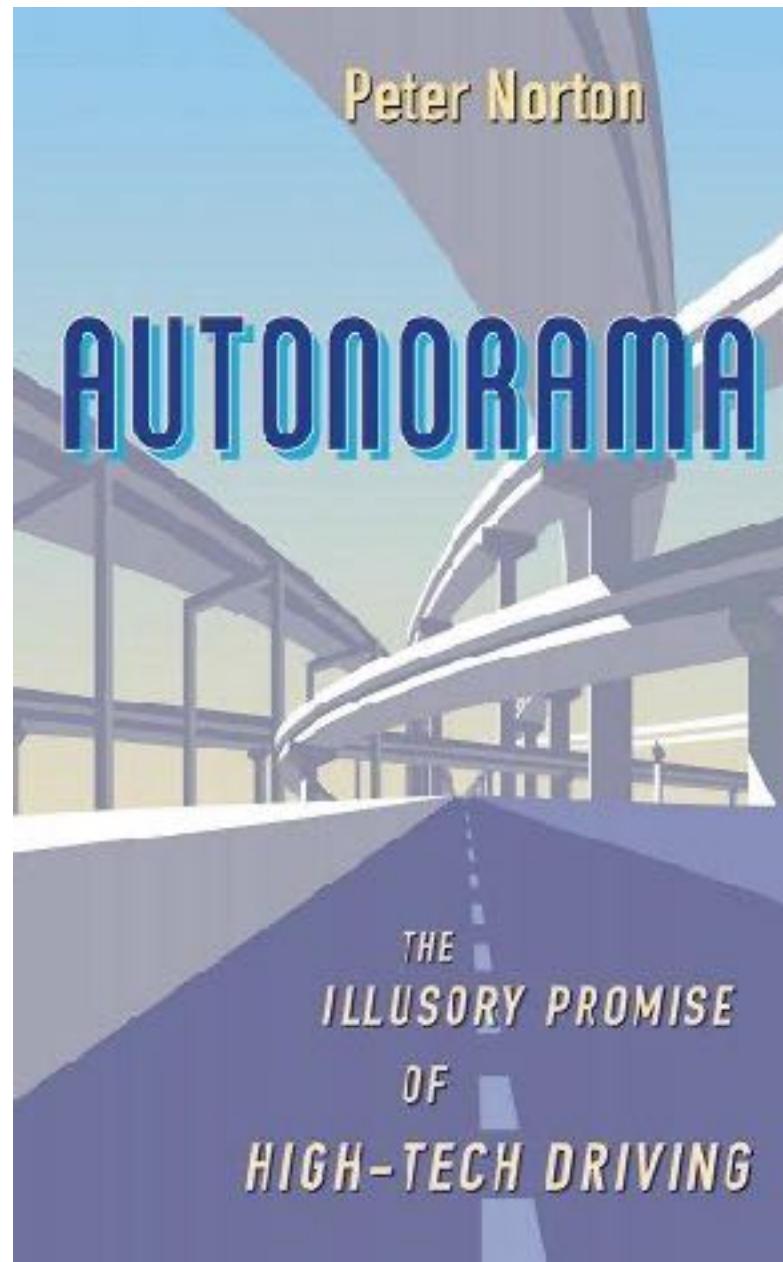
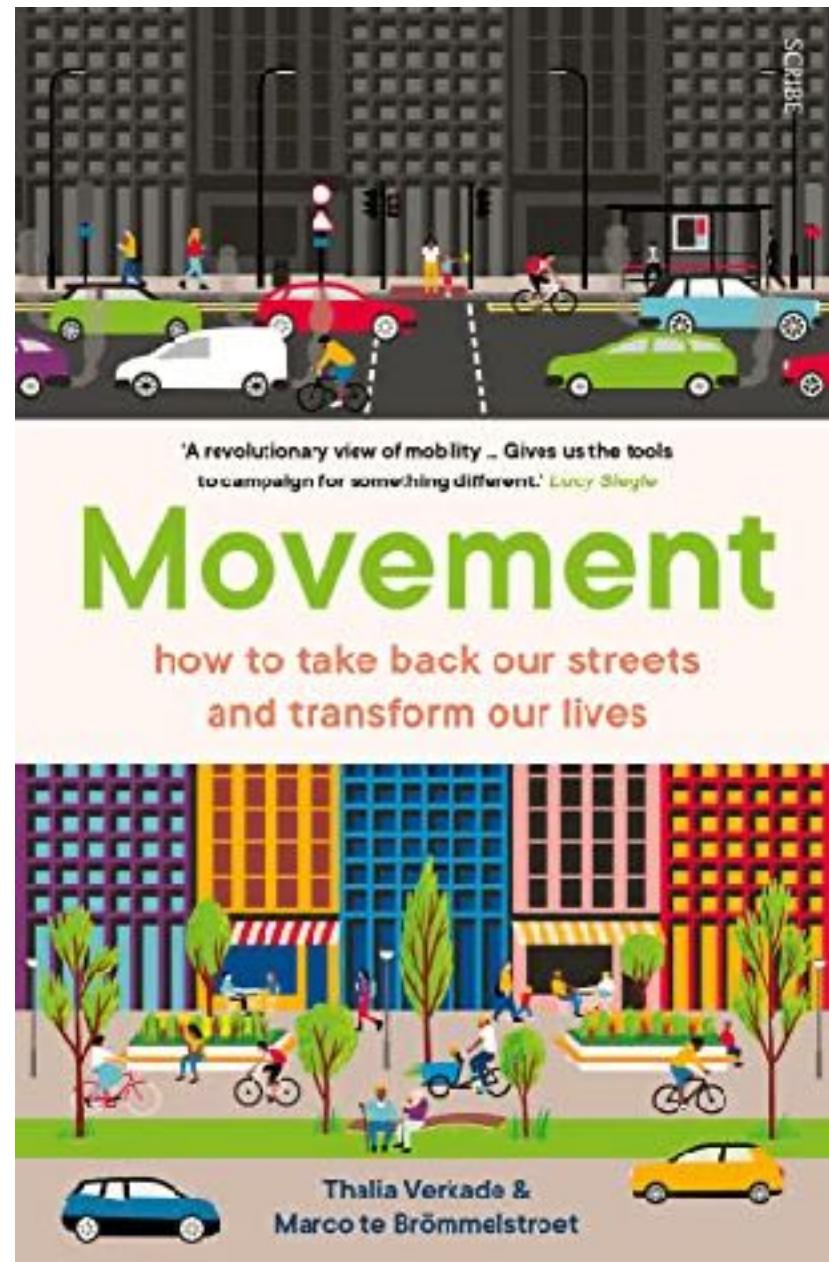


Geospatial Data Science



Missing Links

# Further materials



Henderson (2020): EVs are not the answer

Gössling (2016): Urban transport justice



[thewaroncars.org](http://thewaroncars.org)



Not Just Bikes



Youtube: Julia Steinberger:  
The importance of climate  
activism

# Our work: Data-driven tools to help sustainable urban planning

ROYAL SOCIETY  
OPEN SCIENCE

[royalsocietypublishing.org/journal/rsos](https://royalsocietypublishing.org/journal/rsos)

Research



## Data-driven strategies for optimal bicycle network growth

Luis Guillermo Natera Orozco<sup>1</sup>, Federico Battiston<sup>1</sup>,  
Gerardo Iñiguez<sup>1,2,3</sup> and Michael Szell<sup>4,5,6</sup>

Cite this article: Natera Orozco LG, Battiston F,

scientific reports

OPEN

## Growing urban bicycle networks

Michael Szell<sup>1,2,3</sup>, Sayat Mimar<sup>4</sup>, Tyler Perlman<sup>4</sup>, Gourab Ghoshal<sup>4</sup> & Roberta Sinatra<sup>1,2,3,5</sup>



TRANSPORT FINDINGS

## Computational Desire Line Analysis of Cyclists on the Dybbølsbro Intersection in Copenhagen

Simon Martin Breum<sup>1</sup>, Bojan Kostic<sup>1</sup>, Michael Szell<sup>1,2,3</sup> <sup>a</sup>

<sup>1</sup>Computer Science, IT University of Copenhagen, <sup>2</sup>ISI Foundation, <sup>3</sup>Complexity Science Hub Vienna

Keywords: urban data science, cycling, traffic behavior, intersection design, human-centric planning

<https://doi.org/10.32866/001c.56683>

geographical analysis

*Geographical Analysis* (2022) 0, 1–29

## Automated Detection of Missing Links in Bicycle Networks

Anastassia Vybornova<sup>1</sup> , Tiago Cunha<sup>1</sup>, Astrid Gühnemann<sup>2</sup> , Michael Szell<sup>1,3,4</sup>

Special Issue: Advances in Spatial and Transport Network Analysis

B Urban Analytics and City Science

## Data-driven micromobility network planning for demand and safety

EPB: Urban Analytics and City Science  
2022, Vol. 0(0) 1–16  
© The Author(s) 2022  
Article reuse guidelines:  
[sagepub.com/journals-permissions](https://sagepub.com/journals-permissions)  
DOI: [10.1177/2399808321135611](https://doi.org/10.1177/2399808321135611)  
[journals.sagepub.com/home/epb](https://journals.sagepub.com/home/epb)  
SAGE

Pietro Folco and Laetitia Gauvin  
ISI Foundation, Italy

Michele Tizzoni   
ISI Foundation, Italy  
University of Trento, Italy

Michael Szell   
IT University of Copenhagen, Denmark

[growbike.net](http://growbike.net)

[fixbike.net](http://fixbike.net)

[whatthestreet.com](http://whatthestreet.com)

Soon: BikeDNA



[nerds.itu.dk](http://nerds.itu.dk)