

Participatory smart cities: a crowdsourced data integration model for sustainable and safer mobility

Manuel Merolla

Phd Student in Infrastructures and Transport (XXXIX Cycle)
Curriculum: Transport and Land-use Planning

Supervisor: Prof. Luca Persia



SAPIENZA
UNIVERSITÀ DI ROMA

Tutti i diritti relativi al presente materiale didattico ed al suo contenuto sono riservati a Sapienza e ai suoi autori (o docenti che lo hanno prodotto). È consentito l'uso personale dello stesso da parte dello studente a fini di studio. Ne è vietata nel modo più assoluto la diffusione, duplicazione, cessione, trasmissione, distribuzione a terzi o al pubblico pena le sanzioni applicabili per legge

Who I am

- **Manuel Merolla**, PhD Student in Infrastructures and Transport at CTL – DICEA, Sapienza University of Rome (Italy)
- Supervisor: **Prof. Luca Persia**
- In December 2024 I won the Sapienza's call for mobility periods abroad (**3-month scholarship**)
- Host Institution: **DTU (Denmark)**
- Supervisor: **Prof. Filipe Rodrigues**
- Time: **23 February-24 May 2025**



Outline

1. Introduction

- Topic
- Objectives
- Hypotheses and Research Questions

2. Review Activity

- Literature, Technology, and Legislation Review

3. Crowdsourcing in Denmark

- Analysis of CS Platforms in Denmark and Identification of Case Studies
- Definition and Implementation of Data Integration Model (step-by-step) on Chosen Case Studies and Results

4. Discussion and Conclusion

- Issues and Critical Aspects
- Further Work

INTRODUCTION

Background: from tangible to the intangible



Source: Maptionnaire

What is crowdsourcing?

Definition

“A type of participative online activity in which an individual, an institution, a non-profit organization, or company proposes to a group of individuals of varying knowledge, heterogeneity, and number, via a flexible open call, the voluntary undertaking of a task” (Estellés and González, 2012)

Strengths

- Elaborate information which is not evincible from objective data
- Collect proposals from collective intelligence
- Create a community of users who are responsive to sustainable mobility and road safety issues
- Focus on the acceptability of shared strategies

Weaknesses

- Reliability of information provided by users: true or not?
- Legislative vacuum about CS
- Lack of homogeneity in adoption of CS tools (too many different techniques and platforms)
- Lack of continuity in CS projects (they might be quitted after some months)

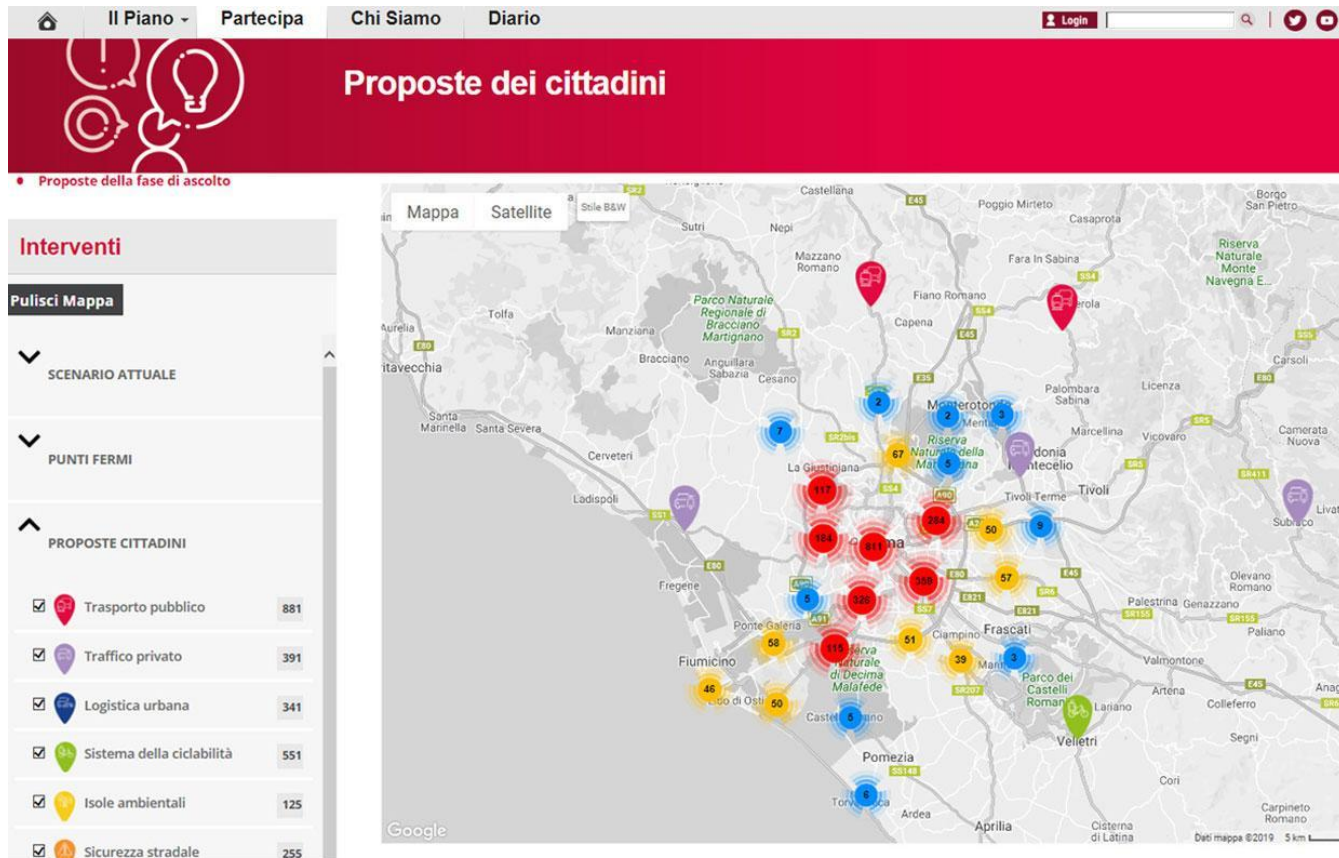
Overall objectives

- Propose a **data integration model** to find the “common ground” between subjective (**crowdsourced**) and objective data regarding sustainable mobility and road safety issues
- Make available to scientific community and local administrators a **reliable tool** based on CTL crowdsourcing platform
- Test the model on **specific case studies** and **use cases** at local, national, and **international** level (e.g., municipalities, regions, EU projects)

Expected outcomes: enable public administrations in more effective data collection, analysis, and exploitation to:

- understand **citizens’ needs** and potential criticalities in transport and road safety management
- plan and adopt the **right interventions** (e.g., SUMP)

The CTL Crowdsourcing platform (SUMP Rome)



Hypotheses and research questions

By comparing **crowdsourced** and **official** data, the proposed model should:

- Identify points of **convergence** and **divergence** between citizens' perceptions and official statistics
- Extract **actionable insights** regarding both spatial and thematic trends (e.g., traffic congestion, road safety concerns);
- Support public administrations in translating these insights into tailored, evidence-based **recommendations** for sustainable mobility and road safety planning

Q1: How research conceptualizes CS in mobility and road safety policies?



Q2: Which regulatory frameworks influence CS adoption?



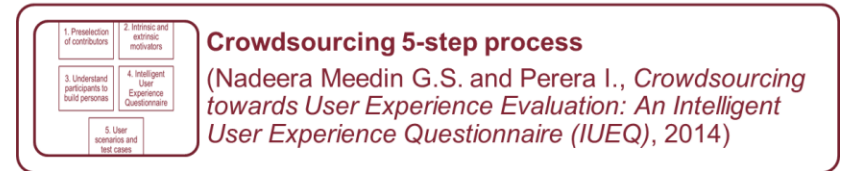
Q3: Which innovations enhance citizens with CS?

REVIEW ACTIVITY

Crowdsourcing: Literature review



1. Whitaker R. et al., *New frontiers for Crowdsourcing: the Extended Mind*, 2015
2. To H. et al., *An Empirical Study of Workers' Behavior in Spatial Crowdsourcing*, 2016
3. Feng W. et al., *A Survey on Security, Privacy, and Trust in Mobile Crowdsourcing*, 2017
4. Karimi H. et al., *Personalized accessibility map (PAM): a novel assisted wayfinding approach for people with disabilities*, 2014
5. Phuttarak J. and Loke S., *A Review of Mobile Crowdsourcing Architectures and Challenges: Toward Crowd-Empowered Internet-of-Things*, 2019
6. Marzano G. et al., *Crowdsourcing solutions for supporting urban mobility*, 2019
7. Brunswicker S. et al., *Taming wicked civic challenges with an innovative crowd*, 2019



Mobility

- **Advanced Infomobility:** exchange of information in two-way interaction with users to improve service (Filippi F. et al., *User empowerment and advanced public transport solutions*, 2013)
- **MaaS (Mobility as a Service):** mobility solutions as a field by considering the point of view of the final users (Volpi V. et al., *Improving Quality of Interaction with the Mobility Services Through the Gamification Approach*, 2018)
- **Smart Mobility at All (SMALL):** includes user data management in data-side support services (Mirri S. et al., *A Service-Oriented Approach to Crowdsensing for Accessible Smart Mobility Scenarios*, 2016)

Smart city

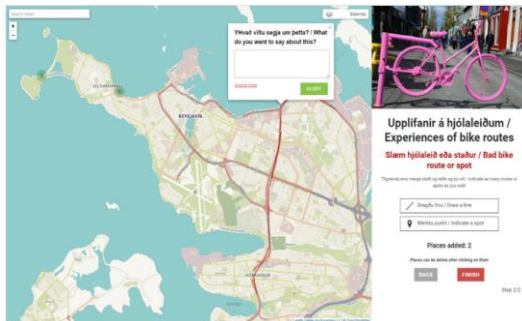
- **Smart People:** dealing with affinity to life long learning, socio-ethnic plurality, flexibility, creativity, open-mindedness, and participation in public life (Giffinger R., et al., *Smart cities - Ranking of European medium-sized cities*, 2006)
- **Participating Citizens:** actively participate in the realization of smart cities and in gathering crowdsourced information to create a starting dataset (Poslončec-Petrić V. et al., *Voluntary Noise Mapping for Smart Cities*, 2016)
- **Community dimension:** organizations, governing institutions, businesses, and residents who are working together, through ICT-based strategies, to create a "smart governance" (Chouikh A., *HandYwiN: A Crowdsourcing-Mapping Solution Towards Accessible Cities*, 2018)

Other applications

- **City Event Management System:** crowdsourcing data as useful source to classify events in data processing step (Chunjiao X. et al., *City Event Management System Based on Multiple Data Source*, 2015)
- **Emergency Management:** crowdsourcers or initiators (local authorities) and crowdsorcees (citizen/foreigners, volunteers, victims) meet each other on a specific platform (Ernst C. et al., *Collaboration and crowdsourcing in emergency management*, 2017)
- **AI in Crowdsourcing Solutions Spaces:** connecting semantic clusters to idea metrics such as likes and comments (Just J. et al., *Catching but a glimpse? - Navigating crowdsourced solution spaces with transformer-based language models*, 2024)

Crowdsourcing: Case Studies in Literature

Reykjavik (Iceland)



Panek J. and Benediktsson K., *Emotional mapping and its participatory potential: Opinions about cycling conditions in Reykjavik, Iceland*, 2016

Milan (Italy)

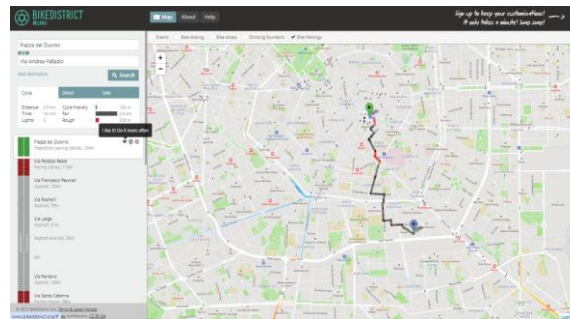
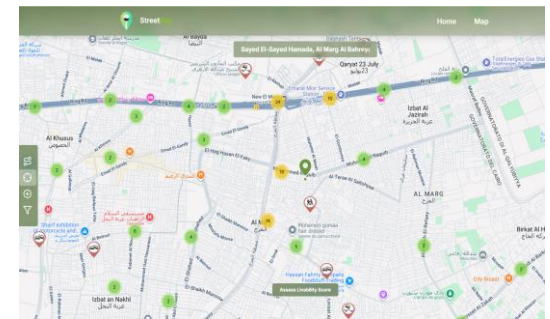


Fig. 4 - Assessment of high impact infrastructural actions. Highlighted in red are high-connectivity roads that are systematically avoided by cyclists due to inadequate conditions.

Quaggiotto M., *Crowdsourced Digital Systems for Sustainable Mobility: Navigation, Data-gathering and Planning*, 2020

Eastern Cairo (Egypt)



Farrag, M. A., Heikal, A. Z. E., Ahmed, M. S., and Osama Amer, A., *Using crowdsourced data to assess safety in developing countries: the case study of Eastern Cairo, Egypt*, 2025

Crowdsourcing: Technology review

A comparison between **crowdsourcing platforms** based on the following key features:

- An **interactive map** for placing reports/proposals
- A **form** to include all the relevant information of reports/proposals
- A tool for **surveys/polls**
- **Sentiment insights** monitoring
- Upload **strategic documents** (e.g., plans)
- An **internal staff** to manage reports/proposals and support users' requests
- **Implementation of AI** to drive monitoring and answering to users' questions
- **Compliance with Data Processing Agreements** (e.g., GDPR)
- A **mobile app** designed for both front and back-end usability

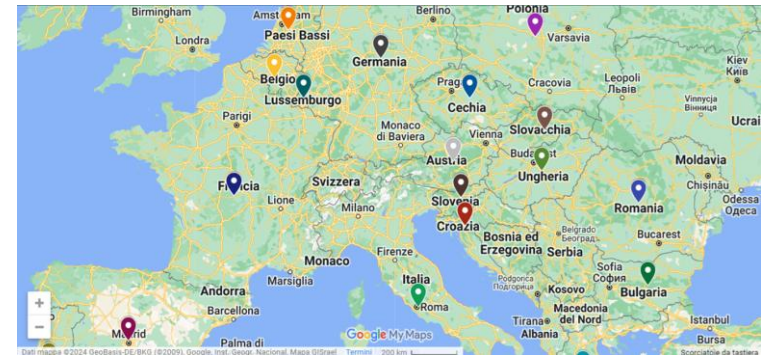
	Interactive map	Form for proposals	Conducting surveys	Tools for insights	Tools for docs	Internal staff
CTL Crowdsourcing	✓	✓	✓		✓	✓
ArchGIS Hub	✓	✓	✓	✓	✓	✓
Maptionnaire	✓	✓	✓	✓	✓	✓
Decidim	✓	✓	✓	✓	✓	✓
Commonplace	✓	✓	✓	✓	✓	
Go Vocal (formerly CitizenLab)	✓	✓	✓	✓	✓	✓
Konveio	✓	✓	✓	✓	✓	
ForeAction (by Civicentric)	✓	✓		✓		✓
SimpliCity	✓			✓		✓
Ushadidi	✓	✓	✓	✓		✓

	AI tools	Data compliant	Mobile app	Other features
CTL Crowdsourcing		✓		
ArchGIS Hub	✓	✓		• Create events and websites
Maptionnaire		✓		• Webpage builder (CMS-based) • Participatory budgeting • Data compatibility & integration
Decidim		✓		• Modules to extend the platform (official, community, authorisation)
Commonplace	✓	✓		• Multilingual
Go Vocal (formerly CitizenLab)	✓	✓		• Participatory budgeting • Follow function to receive updates
Konveio	✓	✓		• A private data access to workshop content
ForeAction (by Civicentric)		✓	✓	• Specific questions for each category • Mobile app is only back-end
SimpliCity	✓	✓	✓	• CityExplore mobile app tool to engage with museums, parks, etc.
Ushadidi	✓	✓	✓	• Multilingual • SMS support

Crowdsourcing: Legislation review

- Almost all the countries of the world suffer the **lack of regulation** on crowdsourcing
- The **United States Code** is the only normative source that explicitly refers to crowdsourcing in “[15 USC 3724: Crowdsourcing and citizen science](#)”
- At global and European levels, we find some laws concerning the **civic participation** in decision-making and policies (e.g., [Aarhus Convention](#), [EU Directives 2003/4/EC](#) and [2003/35/EC](#))
- A list of **guidelines**: [White Paper on Governance](#) (EU), [Guidelines for Citizen Participation Processes](#) (OECD), [Crowdsourcing: The Citizens Highway to Policy-making in the EU](#) (ECAS)

EU countries



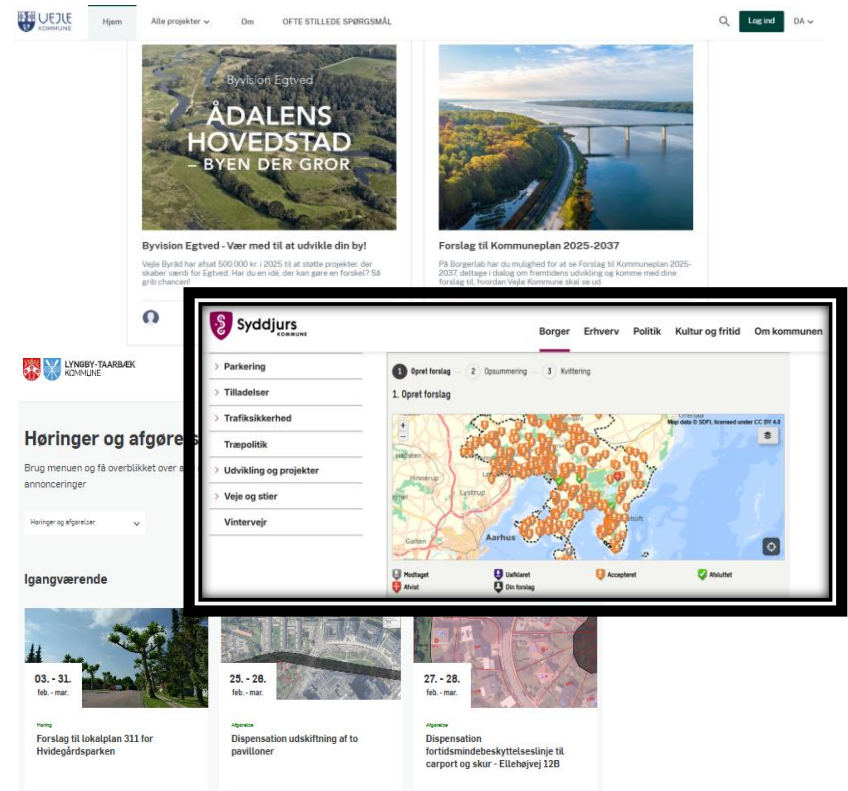
Italian regions



CROWDSOURCING IN DENMARK

CS and Civic Participation in Denmark: most common examples

- *Borgerforslag* (“Citizen proposal”): a platform to create, view, and support proposals for something that citizens believe should be changed in legislation or in society
- **Giv et praj** (“Give a tip”): a platform (mainly map-based) to report issues related with roads, signals, lighting, etc. (**the best and more shared solution**)
- *Høringer og afgørelser* (“Hearings and decisions”): directly submitted by the municipality, requires the citizens’ involvement for their approval/denial



“Giv et praj”: how it works

Syddjurs
KOMMUNE

Borger Erhverv Politik Kultur og fritid Om kommunen

> Parkering

> Tilladelser

> Trafiksikkerhed

Træpolitik

> Udvikling og projekter

> Veje og stier

Vintervej

1 Opret forslag — 2 Opsummering — 3 Kvittering

1. Opret forslag

Map data © SDFI, licensed under CC BY 4.0

Modtaget Uafklaret Accepteret Afsluttet

Afvist Din forslag

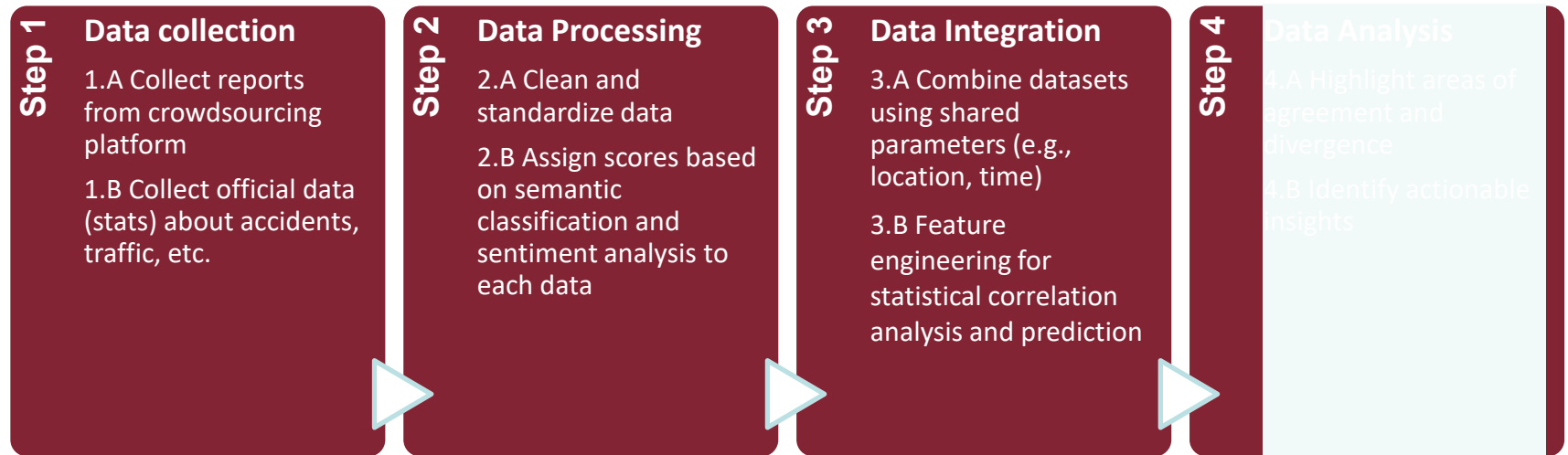
“Giv et Praj”: Review and Case Studies Selection for Data Integration

- 95 out of 98 Danish municipalities rely on “Giv et praj” platforms (commonly developed by external companies)
- DriftWeb Citizen (by Sweco) is the most used CS solution
- First case study: **Aarhus** (Central Jutland, big-sized municipality, 373,388 inhabitants, subdivision into 28 postal districts)
- Other case studies (medium-sized):
 - **Odense**: 4th largest municipality per population (210,083 inhabitants), located in Fyn Island
 - **Gentofte**: 75,076 inhabitants, municipality of the Greater Copenhagen Area)
- Topic: **road accidents**

Platform	Service Provider	N. of municipalities using it
DriftWeb Citizen	Sweco	38
GeoNote	IT34	14
Synchronicer	Soft Design	12
Borger Tip	Skyhost	10
In-house platforms (e.g. "Giv et Praj KBH")	Municipalities (e.g. Copenhagen)	5
Intergraph Field Inspector (Giv et Praj)	Intergraph	4
Rapport Fra Sedet	ADDIN	3
EG Selvbetjening	EG	3
Not available	None	3
PRIOPTI Public	Torgius	3
NemLog-in	Digitaliseringsstyrelsen (The Danish Digital Agency)	2
SmartTip	Elbek & Vejrup	1
TOTAL		98



Data Integration Model: a 4-Step Process



Step 1: Data Collection

- **Subjective Data:** Reports collected from “Giv et praj” platforms of each municipality and translated from Danish to English:
 - area (e.g., Aarhus C, Åbyhøj, etc.)
 - description (the text of the report)
 - year (from 2023 to 2025)
- **Objective Data:** 2023 and 2024 road accidents data derived from the official stats (source: Statistics Denmark):
 - type of accident (all accidents, involving persons influence by alcohol, other accidents): not taken into consideration
 - urban area (if inside or outside urban area)
 - accident situation (e.g., one-vehicle accidents, accidents involving pedestrians, etc.)
 - year

Aarhus

Year	Number of Reports
2023	203
2024	770
2025 (1-3)	686
TOT.	1659

Odense

Year	Number of Reports
2023	235
2024	456
2025 (1-6)	316
TOT.	1007

Gentofte

Year	Number of Reports
2023	172
2024	201
2025 (1-6)	136
TOT.	509

Step 2: Data Processing

1. Preprocessing (bert-base-uncased)	<ul style="list-style-type: none">• Lowercase description text for all the reports
2. Semantic Classification (facebook/bart-large-mnli)	<ul style="list-style-type: none">• Assign an accident situation (objective dataset) as forecasted label to each report (subjective dataset)• Assign a confidence score (from 0 to 1) to each report (relating to the assigned forecasted label)
3. Sentiment Analysis (cardiffnlp/twitter-roberta-base-sentiment)	<ul style="list-style-type: none">• Assign sentiment label (positive, negative, or neutral)• Assign sentiment score (from -1 to 1) to each report (relating to the assigned sentiment)

Semantic Classification (from Aarhus Case Study)

Accident Situation (Forecasted Label)	Count
Accident involving animals, obstacles, etc., on or above roadway	486
Vehicles on different roads meeting in T-junction, Y-junction, crossroads, etc. turning	459
Vehicles on different roads meeting in crossroads without turning	254
Accidents involving pedestrians	114
Accidents involving parked vehicles	106
Vehicles on same road going in same direction without turning from road	88
Vehicles on same road going in opposite directions without turning from road	68
None of the above	52
Vehicles on same road going in same direction, turning into T-junction, Y-junction, crossroads	18
One-vehicle accidents	14
Vehicles on same road going in opposite direction, turning into T-junction, Y-junction, crossroads	0
Total	1659

Issues:

- Semantic Text Classification missed one accident situation/forecasted label (see the table)
- Forecasting of some reports was not so performing (e.g.: “Electronic sign out of order, the sign is located next to Møllesvinget 8541, Skødstrup. The sign should indicate the speed of cars”: assigned to “None of the above”)

Sentiment Analysis (from Aarhus Case Study)

Sentiment Labels (with example of report)	Assigned Forecasted Label	Score	Year
<p><u>Negative</u>: "It's very annoying for us who live here that it's noisy when the trucks drive in the hole in the roadway, probably also dangerous for their tires"</p>	"Accident involving animals, obstacles, etc., on or above roadway"	-0,97	2024
<p><u>Neutral</u>: "Hole right in the bike lane"</p>	"One-vehicle accidents"	0*	2024
<p><u>Positive</u>: "A lot of trees have been planted in the median strip on the ring road - and it looks great! It would be really nice if you removed all the old fence remnants that are still standing in the verge. It would look really nice"</p>	"Accident involving animals, obstacles, etc., on or above roadway"	0,95	2024

Observations:

- Negative tone triggers appropriate sentiment but doesn't guarantee correct category: label may fit a word but not the real context
- Non-emotional safety risks are labeled as neutral: sentiment analysis misses critical issues without emotional language
- Positive tone can mark issues: words like "nice", "great" skew sentiment high, even when a request for improvement is present
- Labeling based on keywords, not comprehension: NLI model may misinterpret mentions of physical elements as active safety issues

Step 3: Data Integration

1. Data Alignment	Match subjective and objective data on shared dimensions: space, time (2023 and 2024 reports-accidents)
2. Feature Engineering	Create features from both datasets: <ul style="list-style-type: none">• sentiment, confidence (min, max, avg, and std)• comparative metrics (accident to report ratio, accident ratio, report ratio, discrepancy score)
3. Exploratory and Pattern Analysis	Statistical Correlation Analysis (Pearson and Spearman Coefficients and p-values)
4. Supervised Learning	Use ML algorithms (XGBoost, Random Forest) to predict objective metrics based on combination of objective data, subjective reports, and engineered features

Pearson and Spearman key findings

Feature Pair	Type	Pearson r	Pearson p	Spearman ρ	Spearman p	Interpretation
avg_confidence – std_confidence (Odense)	Same nature	0.94	1.2e-09	0.93	8.4e-09	Very strong: higher avg. confidence implies greater variability
avg_confidence – max_confidence (Odense)	Same nature	0.87	5.2e-07	0.87	5.9e-07	High avg. confidence aligns with higher peaks
accident_to_report_ratio – discrepancy_score (Aarhus)	Comparative	-0.65	0.002	-0.88	2.7e-07	Strong inverse relation: areas with more accidents per report have lower agreement
report_ratio – discrepancy_score (Gentofte)	Comparative	0.68	0.0004	0.79	1.0e-05	Higher reporting aligns with greater discrepancy
min_sentiment – report_ratio (Aarhus)	Cross-type	-0.54	0.01	-0.83	5.7e-06	Strong negative perception occurs in less reported areas

Supervised Learning

Machine Learning algorithms were explored to predict **number of accidents for 2025**, using engineered features, number of reports, and historical total accidents following train-test-predict approach:

- 2023 data were used to train the model
- 2024 data were used to test and validate the model's predictive reliability
- 2025 data (engineered features and number of reports) were then used exclusively for prediction

XGBoost Performance

Aarhus: MAE (3.17), RMSE (4.79), R^2 (0.61)

Random Forest Performance

- Odense: MAE (0.84), RMSE (1.19), R^2 (0.85)
- Gentofte: MAE (0.78), RMSE (1.30), R^2 (0.45)

Step 4: Reports vs. Accidents Analysis (2023-24)

1. High Reporting, Low Accidents (Over-Perceived Risk)

Category	Reports	Accidents	Insight
<i>Accident involving animals, obstacles, etc., on or above roadway – both Inside and Outside (Odense)</i>	179 / 47	0 / 2	People often report obstructions, but they rarely lead to accidents — high perceived risk, low actual risk
<i>Vehicles on different roads in junctions/crossroads turning – Inside (Aarhus)</i>	210	27	Perceived as dangerous due to complexity; more reports than accidents

2. Low Reporting, High Accidents (Under-Perceived Risk)

<i>Vehicles on same road going opposite direction, turning into junctions – Inside (Aarhus)</i>	0	35	No reports, but 35 accidents — a major under-perception of a high-risk scenario
<i>Same road and direction, turning into junctions – Inside (Gentofte)</i>	3	16	Very few reports, more accidents. This may call for attention, as people might not perceive the danger

3. High Agreement Areas (Perception ≈ Reality)

<i>Same direction, no turning – both Inside and Outside (Aarhus)</i>	43 / 11	35 / 10	Closer match — this is a valid perception
--	----------------	----------------	--

Step 4: Accidents data prediction for 2025 (Aarhus)

Objective Area Category	Accident Situation	TR (2023)	TR (2024)	TR (2025)	2023	2024	2025 (xgb)	DIFF (2025-2024)
Inside Urban Area	One-vehicle accidents	1	6	4	7	5	1	↓ 4
	Vehicles on same road going in same direction without turning from road	12	31	25	19	16	23	↑ 7
	Vehicles on same road going in opposite directions without turning from road	11	29	19	4	5	1	↓ 4
	Vehicles on same road going in same direction, turning into T-junction, Y-junction, crossroads	2	9	5	28	27	26	↓ 1
	Vehicles on same road going in opposite direction, turning into T-junction, Y-junction, crossroads	0	0	0	17	18	18	0
	Vehicles on different roads meeting in crossroads without turning	33	97	80	23	15	27	↑ 12
	Vehicles on different roads meeting in T-junction, Y-junction, crossroads, etc. turning	43	167	153	15	12	21	↑ 9
	Accidents involving parked vehicles	16	48	30	6	6	3	↓ 3
	Accidents involving pedestrians	8	33	53	41	21	34	↑ 13
	Accident involving animals, obstacles, etc., on or above roadway	44	175	160	0	0	1	↑ 1
None of the above	3	25	16	0	0	0	0	
Outside Urban Area	One-vehicle accidents	1	1	1	4	2	2	0
	Vehicles on same road going in same direction without turning from road	3	8	9	6	4	1	↓ 3
	Vehicles on same road going in opposite directions without turning from road	2	4	3	2	4	3	↓ 1
	Vehicles on same road going in same direction, turning into T-junction, Y-junction, crossroads	0	1	1	0	1	0	↓ 1
	Vehicles on same road going in opposite direction, turning into T-junction, Y-junction, crossroads	0	0	0	4	2	1	↓ 1
	Vehicles on different roads meeting in crossroads without turning	4	19	21	3	8	3	↓ 5
	Vehicles on different roads meeting in T-junction, Y-junction, crossroads, etc. turning	7	46	43	4	2	2	0
	Accidents involving parked vehicles	1	5	6	1	1	3	↑ 2
	Accidents involving pedestrians	2	8	10	4	3	2	↓ 1
	Accident involving animals, obstacles, etc., on or above roadway	9	53	45	0	0	2	↑ 2
None of the above	1	5	2	0	0	0	0	

Step 4: Accidents data prediction for 2025 (Odense)

Objective Area Category	Accident Situation	TR (2023)	TR (2024)	TR (2025)	2023	2024	2025 (RF)	DIFF (2025-2024)
Inside Urban Area	One-vehicle accidents	1	3	2	8	7	6	↓ 1
	Vehicles on same road going in same direction without turning from road	12	9	15	2	6	2	↓ 4
	Vehicles on same road going in opposite directions without turning from road	9	13	9	3	3	1	↓ 2
	Vehicles on same road going in same direction, turning into T-junction, Y-junction, crossroads	1	6	3	6	8	6	↓ 2
	Vehicles on same road going in opposite direction, turning into T-junction, Y-junction, crossroads	0	0	0	8	7	6	↓ 1
	Vehicles on different roads meeting in crossroads without turning	33	72	44	8	6	6	0
	Vehicles on different roads meeting in T-junction, Y-junction, crossroads, etc. turning	44	75	44	8	4	6	↑ 2
	Accidents involving parked vehicles	14	17	12	1	0	1	↑ 1
	Accidents involving pedestrians	18	41	27	9	8	7	↓ 1
	Accident involving animals, obstacles, etc., on or above roadway	52	127	73	0	0	1	↑ 1
	None of the above	4	6	7	0	0	0	0
	Outside Urban Area	One-vehicle accidents	0	1	2	0	4	1
Vehicles on same road going in same direction without turning from road		1	3	4	1	6	1	↓ 5
Vehicles on same road going in opposite directions without turning from road		1	2	4	0	0	0	0
Vehicles on same road going in same direction, turning into T-junction, Y-junction, crossroads		2	0	1	0	0	1	↑ 1
Vehicles on same road going in opposite direction, turning into T-junction, Y-junction, crossroads		0	0	0	0	1	1	0
Vehicles on different roads meeting in crossroads without turning		6	15	16	0	0	0	0
Vehicles on different roads meeting in T-junction, Y-junction, crossroads, etc. turning		17	19	18	1	1	1	0
Accidents involving parked vehicles		1	3	6	0	0	0	0
Accidents involving pedestrians		8	6	11	0	0	0	0
Accident involving animals, obstacles, etc., on or above roadway	10	37	16	1	1	1	0	
None of the above	1	1	2	0	0	0	0	

Step 4: Accidents data prediction for 2025 (Gentofte)

Objective Area Category	Accident Situation	TR (2023)	TR (2024)	TR (2025)	2023	2024	2025 (RF)	DIFF (2025-2024)
Inside Urban Area	One-vehicle accidents	0	0	0	2	1	2	↑ 1
	Vehicles on same road going in same direction without turning from road	5	5	3	4	1	2	↑ 1
	Vehicles on same road going in opposite directions without turning from road	3	6	8	1	0	0	0
	Vehicles on same road going in same direction, turning into T-junction, Y-junction, crossroads	1	2	1	9	7	2	↓ 5
	Vehicles on same road going in opposite direction, turning into T-junction, Y-junction, crossroads	0	0	0	4	2	3	↑ 1
	Vehicles on different roads meeting in crossroads without turning	29	32	28	2	2	2	0
	Vehicles on different roads meeting in T-junction, Y-junction, crossroads, etc. turning	26	28	23	3	0	2	↑ 2
	Accidents involving parked vehicles	7	21	8	2	2	2	0
	Accidents involving pedestrians	25	31	18	5	1	3	↑ 2
	Accident involving animals, obstacles, etc., on or above roadway	54	40	32	0	1	0	↓ 1
	None of the above	2	5	1	0	0	0	0
	Outside Urban Area	One-vehicle accidents	0	0	0	2	1	2
Vehicles on same road going in same direction without turning from road		0	2	0	0	1	0	↓ 1
Vehicles on same road going in opposite directions without turning from road		0	0	1	0	0	0	0
Vehicles on same road going in same direction, turning into T-junction, Y-junction, crossroads		0	0	0	1	0	0	0
Vehicles on same road going in opposite direction, turning into T-junction, Y-junction, crossroads		0	0	0	0	0	0	0
Vehicles on different roads meeting in crossroads without turning		5	4	1	0	0	0	0
Vehicles on different roads meeting in T-junction, Y-junction, crossroads, etc. turning		3	9	6	0	0	0	0
Accidents involving parked vehicles		0	3	2	0	0	0	0
Accidents involving pedestrians		3	8	1	0	0	0	0
Accident involving animals, obstacles, etc., on or above roadway		7	5	3	0	0	0	0
None of the above		2	0	0	0	0	0	0

DISCUSSION AND CONCLUSION

Issues and Critical Aspects

Data-related issues

- Heterogeneous and incomplete reports, with inconsistent formats and vocabularies
- Uneven citizen participation: smaller municipalities and urban vs. suburban imbalance produce sparse, less representative data
- Irregular sentiment and confidence may reduce data reliability

Methodological challenges

- Complex semantic alignment between free-text reports and detailed accident types
- Limited geographic granularity in Danish case studies (urban/non-urban distinction)
- Predictive accuracy affected by class imbalance, temporal aggregation, and reporting lag

Conceptual considerations

- Divergence between perceived and actual risks
- Road safety focus: strong foundation, but extensions needed for broader mobility contexts
- Socio-cultural biases may influence reporting patterns
- Translating model outputs into actionable policy insights remains challenging

Further Work

Methodology advancements

- Use specialized, domain-adapted language models for improved semantic classification
- Introduce image analysis to strengthen textual evidence of citizens reports
- Analyze topics and behavioral factors driving user participation and report quality

Practical development

- Introduce Step 5: “Feedback and Action” to complete the framework
- Test the model in pilot projects (e.g., SUMP, EU/regional initiatives)
- Extend application beyond road safety to other mobility domains

Expected Scientific Impact

- Strengthen evidence-based and participatory planning through data fusion
- Provide a scalable framework adaptable to diverse contexts
- Deepen understanding of socio-spatial and behavioral dimensions in crowdsourcing

THANK YOU FOR YOUR ATTENTION!

manuel.merolla@uniroma1.it

 [Manuel Merolla](#)