

# Urban Mobility Analytics

David Jonietz<sup>\*1</sup>, Monika Sester<sup>\*2</sup>, Kathleen Stewart<sup>\*3</sup>,  
Stephan Winter<sup>\*4</sup>, Martin Tomko<sup>\*5</sup>, and Yanan Xin<sup>†6</sup>

- 1 **HERE** – Zürich, CH. david.jonietz@here.com
- 2 **Leibniz Universität Hannover**, DE. monika.sester@ikg.uni-hannover.de
- 3 **University of Maryland – College Park**, US. stewartk@umd.edu
- 4 **The University of Melbourne**, AU. winter@unimelb.edu.au
- 5 **The University of Melbourne**, AU. tomkom@unimelb.edu.au
- 6 **ETH** – Zürich, CH. yanxin@ethz.ch

---

## Abstract

This report documents the program and the outcomes of Dagstuhl Seminar 22162 “Urban Mobility Analytics”. The seminar brought together researchers from academia and industry who work in complementary ways on urban mobility analytics. The seminar especially aimed at bringing together ideas and approaches from deep learning research, which is requiring large datasets, and reproducible research, which is requiring access to data.

**Seminar** April 18–22, 2022 – <http://www.dagstuhl.de/22162>

**2012 ACM Subject Classification** Applied computing → Transportation; Computing methodologies → Planning under uncertainty; Information systems → Geographic information systems

**Keywords and phrases** data analytics, Deep learning, Reproducible research, urban mobility

**Digital Object Identifier** 10.4230/DagRep.12.4.26

## 1 Executive Summary

*Monika Sester (Leibniz Universität Hannover, DE)*

*Martin Tomko (The University of Melbourne, AU)*

*Stephan Winter (The University of Melbourne, AU)*

**License**  Creative Commons BY 4.0 International license  
© Monika Sester, Martin Tomko, and Stephan Winter

Seminar 22162 addressed recent trends in urban mobility analytics that are shaping the information available to transport planners, operators, and travellers. Seminar participants were particularly discussing how information can be provided that supports the critical transformation of urban mobility towards climate neutrality and other sustainability goals, i.e. that supports to change mobility behaviour.

The trends identified for this seminar were, on one hand, the rise of deep learning methods for massive data analytics, and on the other hand the emerging digital divide between those having massive data and those who haven’t, which, in short, forms the challenges of academia for reproducible research. Massive data on urban mobility is collected by industry and transport authorities, with limited access outside, for various reasons. Also, the research and development capacity behind the closed doors of large transnational companies – especially in the platform economy – is arguably faster than the typical PhD process.

---

\* Editor / Organizer

† Editorial Assistant / Collector



Except where otherwise noted, content of this report is licensed under a Creative Commons BY 4.0 International license

Urban Mobility Analytics, *Dagstuhl Reports*, Vol. 12, Issue 4, pp. 26–53

Editors: David Jonietz, Monika Sester, Kathleen Stewart, Stephan Winter, Martin Tomko, and Yanan Xin



DAGSTUHL  
REPORTS

Dagstuhl Reports  
Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

These challenges and opportunities were shaping the discussions where participants split into working groups on (a) *ethics and the social good* – how can information trigger change in mobility behaviour; (b) *methods and explainability*; (c) *benchmarking and datasets*; and (d) *applications*.

The seminar had quite a diversity of participants, which was inspiring in all the discussions. Participants from industry gave talks about what happened behind their ‘closed doors’, and further tutorials were introducing datasets, the principle of reproducible research, and European funding opportunities.

The industry partners showed great interest in collaboration with academia, however, the problem of data sharing was still considered as paramount. There are trends to open certain kinds of data, e.g. in aggregated form, or simulated data, or only based on contracts with certain institutions. Still, open data sharing remains to be a challenge.

## 2 Table of Contents

### Executive Summary

<i>Monika Sester, Martin Tomko, and Stephan Winter</i> . . . . .	26
--	----

### Overview of Seminar

Urban Mobility Analytics Seminar	
<i>Monika Sester, Martin Tomko, and Stephan Winter</i> . . . . .	29

### Overview of Talks

Vehicle Data Democratization at Volkswagen Commercial Vehicles	
<i>Michael Nolting</i> . . . . .	31
Mobility Research at Swiss Federal Railway Company	
<i>Erik Nygren</i> . . . . .	32
Traffic4cast Data Intro	
<i>Moritz Neun and Christian Eichenberger</i> . . . . .	33
Deep Learning of Road User Behavior	
<i>Hao Cheng</i> . . . . .	35
Modeling the Interaction between Places and Human Mobility	
<i>Cheng Fu</i> . . . . .	36
Horizon Europe: Introduction to R&I Funding in the Field of Mobility	
<i>Stephan Winter and David Doerr</i> . . . . .	37
Reproducibility for Urban Mobility Analysis	
<i>Daniel Nüst</i> . . . . .	38

### Working Groups

Group on Ethics / Social Good	
<i>Alexandra Millonig, Ivan Majic, Edoardo Neerhut, Moritz Neun, Luca Pappalardo, Chiara Renso, and Stephan Winter</i> . . . . .	40
Group on Benchmarking and Datasets	
<i>Vanessa Brum-Bastos, Christian Eichenberger, Cheng Fu, Erik Nygren, and Maya Sekeran</i> . . . . .	43
Group on Methods and Explainability	
<i>Anita Graser, Hao Cheng, Tao Cheng, Ioannis Giannopoulos, Daniel Nüst, Martin Tomko, and Yanan Xin</i> . . . . .	46
Group on Applications	
<i>Martin Lauer, Andris Clio, Dirk Christian Mattfeld, and Monika Sester</i> . . . . .	50
Outcomes	
<i>Monika Sester, Martin Tomko, and Stephan Winter</i> . . . . .	52

<b>Participants</b> . . . . .	53
-------------------------------	----

<b>Remote Participants</b> . . . . .	53
--------------------------------------	----

## 3 Overview of Seminar

### 3.1 Urban Mobility Analytics Seminar

*Monika Sester (Leibniz Universität Hannover, DE), Martin Tomko (The University of Melbourne, AU), Stephan Winter (The University of Melbourne, AU)*

License  Creative Commons BY 4.0 International license  
© Monika Sester, Martin Tomko, and Stephan Winter

Transportation in cities is undergoing unprecedented change, such as by vehicle technology towards autonomous driving (disrupting mobility); massive real-time data and data analytics (smart cities, sensing cities, dashboards); sharing and integration platforms (ride-hailing, mobility-as-a-service) and urban logistics (changing shopping patterns). All this happens in parallel with an increasing willingness, and a sharp necessity, to change mobility behaviour in the face of human-induced climate change, where urban transport is a major contributor [1, 2, 3].

Critical to the success of transforming urban mobility towards climate neutrality is information provided to planners, operators, and travellers. Increasingly, this information can be produced based on data. This Dagstuhl Seminar on Urban Mobility Analytics addressed recent trends that are shaping the information derived from such urban mobility analytics:

- A prominent trend, not only in transportation research, is the rise of deep learning methods for massive data analytics [4, 5]. In the domain of urban mobility, this massive data emerges from a range of sensor platforms, from infrastructure (CCTV, induction loops, people counters, WiFi, smart cards, air quality) to vehicles (GPS, vision, LiDAR, radar) and smartphones (GPS, location-based apps, accelerometer, gyroscope, magnetometer), in volume, heterogeneity, velocity and veracity a prime application domain for deep learning.
- A second trend is the emerging digital divide between academia and industry and its challenges for reproducible research [6, 7], a trend that has been compared to digital feudalism [8]. While massive data on urban mobility is collected by industry and transport authorities, their access for academic research is limited by privacy concerns and also by commercial sensitivities. While reproducible research hinges on access to data (and the generation of benchmark datasets is costly and often limited to narrow use cases), the research and development capacity behind the closed doors of large transnational companies – especially in the platform economy – is arguably faster than the typical PhD process.
- Related to both trends above is the buzzword of Digital Twins (e.g., <https://muenchen.digital/twin/>). Since both data and data analytics become more often available sufficiently close to real-time, the information derived is less and less consumed in human decision making but in the self-regulation of cyber-physical-social systems. These systems will propel future urban mobility by autonomously driving vehicles and mobility-as-a-service [9, 10], however, their development and use involve many still-open research questions, such as reliability, trust, and the interaction of humans with these systems, let alone the bigger question of social or ethical engagement in data-driven solutions [11].

Accordingly, the seminar brought together researchers from academia and industry who work in complementary ways on urban mobility analytics such that they do not necessarily meet at the same conferences or refer to standardized discipline practices. Especially we aimed to bring together ideas and approaches from deep learning research, which is requiring large datasets, and reproducible research, which is requiring access to data.

Tuesday	Wednesday	Thursday	Friday
Welcome, aims, program, and short introductions	Session 1 of participants talks – relevant research	David Doerr: <i>Horizon Europe's calls on smart mobility</i>	From break-out discussions to outcomes
Industry stories: • Michael Nolting (VW) • Erik Nygren (SBB)	Session 2 of participants talks – relevant research	Break-out groups	Conclusions: plans ahead
Moritz Neun and Christian Eichenberger (IARAI): <i>IARAI data tutorial</i>	<i>Hike into the nearby woods</i>	Break-out groups, ctd.	
Data challenges and activities	Review of challenges and activities	Daniel Nüst: <i>Reproducible research</i>	

■ **Figure 1** Seminar structure.

The seminar made also a deliberate effort to invite people from both sides of the digital divide (i.e., from academia and industry) to share their experiences, their approaches, and their challenges, and to explore more future collaboration. These dynamics at the seminar were also inspired by two available real-world, large traffic datasets, one sponsored by IARAI (<https://www.iarai.ac.at/traffic4cast/>), and one by SBB and others (<https://flatland.aicrowd.com>).

Since the seminar took place in the Easter week, it was only a four-day seminar, starting with Tuesday. The week had a recognizable structure (Figure 1):

- **Tuesday**, after short introductions of the participants, belonged completely to the industry. Two industry speakers, Michael Nolting from VWN and Erik Nygren from SBB, spoke about data capture in their commercial environments, and the use of this data for management and planning. Erik Nygren introduced also the Flatland challenge. Afterwards Moritz Neun and Christian Eichenberger (IARAI) presented the Traffic4cast dataset in detail. In the last session, participants discussed themes for the week, including a data challenge.
- **Wednesday** was filled with short talks by the participants, introducing their research interests and embedding them into the context of the seminar. The afternoon, after the traditional hike, was completed by a plenary session to rank the identified themes and to plan for break-out groups.
- **Thursday** was filled with these break-out groups, framed by an introduction to current and future Horizon Europe calls on smart mobility in the morning by David Doerr, and an introduction to reproducible research by Daniel Nüst.
- **Friday** collected the discussions of the break-out groups, discussed reporting, and ended with collecting concrete commitments for next steps.

In the following, the report will first give an overview of selected talks, followed by the summaries of the break-out groups. It concludes with a brief summary of the outcomes.

Due to the ongoing pandemic, the seminar took place in hybrid mode, with about two thirds of participants on site. While Dagstuhl's conferencing system and our two technical assistants Maya Santhira Sekeran and Ivan Majic were providing a smooth interaction, time zones were causing challenges for the online participants. Unfortunately, two of the original organizers – Kathleen Stewart and David Jonietz – were not able to participate at all, such that the remaining two original organizers – Monika Sester and Stephan Winter – were grateful for Martin Tomko to come on board in the last minute. The organizers acknowledge every person's contributions and commitment.

## References

- 1 Boschmann, E. & Kwan, M. Toward Socially Sustainable Urban Transportation: Progress and Potentials. *International Journal Of Sustainable Transportation*. **2**, 138-157 (2008)
- 2 May, A. Urban Transport and Sustainability: The Key Challenges. *International Journal Of Sustainable Transportation*. **7**, 170-185 (2013)
- 3 OECD Decarbonising Urban Mobility with Land Use and Transport Policies: The Case of Auckland. (OECD Publishing,2020)
- 4 Li, S., Dragicevic, S., Anton, F., Sester, M., Winter, S., Coltekin, A., Pettit, C., Jiang, B., Haworth, J., Stein, A. & Cheng, T. Geospatial Big Data Handling Theory and Methods: A Review and Research Challenges. *ISPRS Journal Of Photogrammetry And Remote Sensing*. **115**, 119-133 (2016)
- 5 Varghese, V., Chikaraishi, M. & Urata, J. Deep Learning in Transport Studies: A Meta-analysis on the Prediction Accuracy. *Journal Of Big Data Analytics In Transportation*. **2**, 199-220 (2020), <https://doi.org/10.1007/s42421-020-00030-z>
- 6 Ivie, P. & Thain, D. Reproducibility in Scientific Computing. *ACM Comput. Surv.* **51**, Article 63 (2018), <https://doi.org/10.1145/3186266>
- 7 Nüst, D. & Pebesma, E. Practical Reproducibility in Geography and Geosciences. *Annals Of The American Association Of Geographers*. **111**, 1300-1310 (2021), <https://doi.org/10.1080/24694452.2020.1806028>
- 8 Jensen, J. The Medieval Internet: Power, Politics and Participation in the Digital Age. (Emerald Publishing,2020)
- 9 Ibrahim, M., Rassölkin, A., Vaimann, T. & Kallaste, A. Overview on Digital Twin for Autonomous Electrical Vehicles Propulsion Drive System. *Sustainability*. **14** (2022), <https://www.mdpi.com/2071-1050/14/2/601>
- 10 Mahmoud, E., Darwish, A. & Hassanien, A. The Future of Digital Twins for Autonomous Systems: Analysis and Opportunities. *Digital Twins For Digital Transformation: Innovation In Industry*. pp. 187-200 (2022), [https://doi.org/10.1007/978-3-030-96802-1\\_10](https://doi.org/10.1007/978-3-030-96802-1_10)
- 11 Charitonidou, M. Urban scale digital twins in data-driven society: Challenging digital universalism in urban planning decision-making. *International Journal Of Architectural Computing*. pp. 14780771211070005 (2022), <https://doi.org/10.1177/14780771211070005>

## 4 Overview of Talks

### 4.1 Vehicle Data Democratization at Volkswagen Commercial Vehicles

Michael Nolting (*Volkswagen Nutzfahrzeuge – Hannover, DE*)

License © Creative Commons BY 4.0 International license  
© Michael Nolting

In the near future, cars will become more than just mechanical objects bringing customers safely and quickly from point A to B. Currently, the focus is on mechanical improvements to safety features such as airbags or reducing fuel consumption and emissions. Soon, however, the focus will switch to the electrical and IT aspects of cars. The automobile will become a rolling computer, providing added value via autonomous driving and an infotainment system which makes the resulting leisure time more productive and enjoyable. Thanks to improvements in electric motor technology, the complexity regarding manufacturing has decreased significantly. This means that enterprises such as Google or Apple – companies that develop sophisticated computer, smart phone and infotainment platforms – will be more involved in important aspects of car production than traditional car manufacturers such as Volkswagen. As often stated, based on valuable information stemming from big data analytics, future cars will be electric, autonomous, connected, on the whole smart.

You can regularly read in newspapers about the progress car manufacturers and their suppliers are making regarding the first three points. There is not much conversation, however, regarding the coming smart car revolution. It is obvious that the industry has recognized the risk of losing future market share by failing to evolve into big data enterprises based on recent acquisitions, investments and partnerships they have made, and their hiring spree of data scientists. BMW, Audi, and Mercedes, for example, just bought Nokia's mapping service, and BMW has undertaken a massive recruitment of data scientists. Nevertheless, time is running out. If car manufacturers don't start to take the reins now, they will miss out on the smart car revolution of tomorrow.

This is the reason why car manufacturers have to transform into a software company. In order to achieve this, they have to increase their daily deployment frequency and reduce the overall lead time. This can be achieved by coping with the integration issues, which currently still exist. In his talk, Dr. Michael Nolting has shown ways how to overcome these integration issues and so paving the way how to transform into a software-driven company. In addition to this, the next steps would then be to become a data-driven company by democratizing data and AI within the company and find ways and solutions how to share this data with research institutes, which is still a challenging task.

## 4.2 Mobility Research at Swiss Federal Railway Company

*Erik Nygren (Schweizerische Bundesbahnen – Bern, CH)*

License  Creative Commons BY 4.0 International license  
© Erik Nygren

The Swiss Federal Railway Company (SBB) is an integral part of the Swiss public mobility network and transports over 1 million passengers and 200 thousand tons of goods each day. Current projections indicate that the demand for public urban mobility will increase by up to 40%. To be able to continue to offer a reliable public service many technical, social, and operational challenges must be overcome. SBB has identified 7 main topics of interest where applied research together with research institutes and other partners is being conducted.

1. Customer Oriented Railway: We focus on enhancing the attractiveness of railway through different incentives and improvements to our service. Social and behavioral research is needed to better understand the decision making of our customers.
2. Simplify Access to Railway: The focus lies in the seamless integration of railway into other modes of transport. Research focus is both on technical issues related to the interaction of different modes of transportation as well as social and behavioral aspects to understand the customer needs.
3. Flexible offers and production models: The change in mobility behavior both in passenger and freight transportation towards on-demand systems requires dynamic and real-time planning. Research is mainly conducted around network wide optimization and planning in real-time.
4. Resilience and Efficiency: Railway industry is asset heavy, and its reliable operations depend on well planned and efficient maintenance. We research the use of novel technologies and algorithms to improve both planning and execution of maintenance work on the infrastructure and rollingstock.
5. Long-term strengthening of SBB: What will the future of mobility look like and what will the role of railway be? This and many other questions are being investigated while considering both technological and social developments in society.

6. Environment and Sustainability: Railway today is already one of the most sustainable forms of mobility, we aim to further lower the impact of mobility by improving the efficiency of material usage and lower the energy consumption per travelled kilometre.
7. Optimized freight logistics: To facilitate the move from road-based transportation systems to railway, more dynamic freight logistics need to be offered. Together with academia, we are looking for more efficient and reliable planning and coordination algorithms as well as new business models to improve railway freight attractiveness.

SBB uses different forms of applied and academic research to tackle open questions around these 7 topics of interest. Research projects conducted in close collaboration with academic partners allow us to get a good grasp of future possibilities and plan our services accordingly. In addition, we use Open Data and Open Research to reach a larger group of experts from many different fields. Open Research allows us to compare different research results with each other and monitor the progress and performance of novel approaches.

This approach has proven valuable for traffic management systems research, where SBB has hosted an optimization competition called Flatland from 2018 until today. The competition consists of a simple railway traffic simulation and the objective to optimize traffic flow towards high punctuality while stochastic disturbances force frequent large-scale replanning. Over the years many novel and surprising solutions have been submitted by participants around the world and been evaluated. We observe that new algorithms from the field of Deep Learning still struggle to outperform classical optimization algorithms, but that their quality and performance is advancing each year. Throughout all different research projects conducted at our company we have learned that a close and bilateral collaboration on research topics is necessary to shorten the time to market for new ideas. Inspiration for new solutions often comes from academic research and industry partners can provide valuable insights into real-world challenges faced by mobility providers. We, therefore, invest in more interdisciplinary exchange and collaboration between industry and academia.

### 4.3 Traffic4cast Data Intro

*Moritz Neun (IARAI – Zürich, CH) and Christian Eichenberger (IARAI – Zürich, CH)*

License © Creative Commons BY 4.0 International license  
© Moritz Neun and Christian Eichenberger

Expectations on AI in mobilities are immense, targeting climate-neutral & smart cities. To achieve this goal there is an increasing need not only for models but also for metrics that help steer and influence these complex urban mobility systems. To give priority to people and not to transportation vehicles, cities must learn how to understand traffic as a whole for shaping the performance of a transportation system and for evaluating and reflecting changes in the real-world through an iterative data-driven setting (i.e. online model calibration; model, data fusion and metrics innovation; real-world change detection). Such a learned digital twin would help close the gap between traffic control systems and model-based planning tools by providing an integrated holistic feedback loop with more diverse data and a deeper embedding of traffic dependencies.

The Traffic4cast competition aims at providing such an understanding of traffic rules in a data-driven way, pushing the latest methods in modern machine learning to model complex spatial systems over time. The competition was part of the NeurIPS competition track in the last 3 years and provides an industrial-scale dataset with a high-resolution privacy-preserving

view of urban traffic. The dataset is derived from GPS trajectories (floating car data) of a large fleet of probe vehicles and covers 10 culturally diverse cities around the world in a time span of 2 years.

Floating car data from GPS is a typical source to be used in trajectory or origin-destination analysis. On the other hand, in traffic detection and prediction other data sources such as stationary traffic counters are used. While traffic counter data only observes traffic in certain locations (spatial bias), floating car data can observe traffic everywhere but usually sees only a tiny part of the total traffic “population” (car fleet bias of industrial providers and participation bias of crowd-sourced initiatives). With the Traffic4cast dataset we now do have a floating car dataset that has a sufficiently large volume to study those effects.

In Traffic4cast the floating car data is aggregated into the traffic map movie format using spatio-temporal cells – this facilitates processing and preserves privacy. Each cell corresponds to the area of approximately 100m x 100m. The GPS data is aggregated as vehicle count and average speed per cell, temporally in 5 minute time intervals and axially in 4 heading quadrants (NE, SE, SW, and NW). While this 4-level aggregation and compression scheme gives away some details in the data, at the same time it also allows it to handle much larger data volumes over longer periods in time. It also allows and encourages the direct use of the latest advances in analytics and ML, as has been successfully shown in the results of the past three competitions.

In 2022, Traffic4cast is moving from the grid based movies format to a graph based format which allows to combine loop counter and floating car data. The floating car data in the gridded map movie format will serve as input for deriving the speed and travel time ground truth labels on the road graph. Therefore the 5 minute time bins in our usual spatio-temporal data format (see [1]) are getting aggregated to 15 minutes time bins taking the average speed to our loop counter data. From the dynamic data, we then derive the ground truth labels, namely CongestionClass (CC; red/congested, yellow/warning, green/uncongested) for each segment in the road graph and Travel Time (ETA) for each super-segment.

We focus on the three cities London, Madrid and Melbourne where both floating car data as well as large and open traffic counter datasets are available. The underlying raw data is similar to the input for commercial traffic maps and routing products. Similar to the previous years, the data will be made available for download from HERE Technologies.

## References

- 1 Christian Eichenberger, Moritz Neun, Henry Martin, Pedro Herruzo, Markus Spanring, Yichao Lu, Sungbin Choi, Vsevolod Konyakhin, Nina Lukashina, Aleksei Shpilman, Nina Wiedemann, Martin Raubal, Bo Wang, Hai L. Vu, Reza Mohajerpoor, Inhi Kim, Luca Hermes, Andrew Melnik, Riza Velioglu, Markus Vieth, Malte Schilling, Alabi Bojesomo, Hasan Al Marzouqi, Panos Liatsis, Jay Santokhi, Dylan Hillier, Yiming Yang, Joned Sarwar, Anna Jordan, Emil Hewage, David Jonietz, Fei Tang, Aleksandra Gruca, Michael Kopp, David Kreil, and Sepp Hochreiter. Traffic4cast at NeurIPS 2021 – temporal and spatial few-shot transfer learning in gridded geo-spatial processes. In Hugo Jair Escalante and Katja Hofmann, editors, *Proceedings of the NeurIPS 2021 Competition Track*, volume forthcoming of *Proceedings of Machine Learning Research*. PMLR, 2022. URL <https://arxiv.org/abs/2203.17070>.

## 4.4 Deep Learning of Road User Behavior

Hao Cheng (Leibniz Universität Hannover, DE)

License © Creative Commons BY 4.0 International license  
© Hao Cheng

Learning how road users behave is essential for developing many intelligent systems, such as traffic safety control, self-driving cars, and robot navigation systems. However, automated and accurate recognition of road users' behavior is still one of the bottlenecks in realizing such systems in urban traffic that is – compared to other types of traffic – especially dynamic and full of uncertainties. Some urban environments make detecting and predicting road users' behavior particularly challenging, e.g., temporarily shared spaces of intersections for vehicle turning or shared spaces as a traffic design. The former allows vehicles to turn and interact with other crossing road users, the latter is intended to make different types of road users share the space, therefore reducing the dominance of vehicles and improving pedestrian movement and comfort. Direct interactions between vehicles and vulnerable road users (VRUs, e.g., pedestrians and cyclists) lead to high uncertainty of traffic behavior. Their dynamic movements and mutual influence make their behavior multimodal, such as stopping, accelerating, and turning in different directions. A road user may have more than one option to choose such maneuvers even in the same traffic situation. Moreover, ambiguous traffic situations (e.g., road users negotiating usage of the road) make their behavior difficult to predict.

With the development of deep learning techniques and the availability of large-scale real-world traffic data, there is a high chance to automatically and accurately learn road users' behavior. The core question is how to leverage such traffic data captured by e.g., camera or LiDAR, or vector data like trajectories and GPS tracks, to train a deep learning model that can mimic the multimodality of road users' behavior in various traffic situations, especially at places as mentioned above where they inevitably confront each other?

This research project aims to investigate building smart intersections and shared spaces, with the ability to predict how different types of road users move and interact with each other. The following steps are a conceptual pipeline of learning road users' multimodal intent. First, traffic in the area of interest at a vehicle turning intersection or in a shared space is captured by using stationary camera sensors. Then, the image pre-processing steps, i.e., camera calibration, object detection, projection transformation, and object tracking, are carried out to extract trajectories from the image data. The well-established approaches from, e.g., benchmark multi-object tracking algorithms<sup>1</sup> will be applied for trajectory tracking. The main contributions of this research lie in the step of trajectory forecasting and intent prediction. The predicted behavior in space and time, e.g., intended trajectories, are analyzed for three major tasks: a) path planning of the ego agent [1, 2], b) safety analysis, including collisions and conflicts of different severity [3], and c) anomaly detection, classifying behavior patterns that do not conform to a well-defined notion of normal behavior [4]. In parallel to the predicted behavior, the further extracted behavior will serve as an observed reference for evaluating the performances of the prediction tasks.

---

<sup>1</sup> <https://motchallenge.net/>

## References

- 1 Cheng, H., Liao, W., Tang, X., Yang, M. Y., Sester, M., and Rosenhahn, B. (2021). Exploring dynamic context for multi-path trajectory prediction. In *2021 IEEE International Conference on Robotics and Automation (ICRA)*, pages 12795–12801.
- 2 Cheng, H., Liao, W., Yang, M. Y., Rosenhahn, B., and Sester, M. (2021). Amenet: Attentive maps encoder network for trajectory prediction. *ISPRS Journal of Photogrammetry and Remote Sensing*, 172:253–266.
- 3 H. Cheng, H. Liu, T. Hirayama, F. Shinmura, N. Akai, and H. Murase, “Automatic interaction detection between vehicles and vulnerable road users during turning at an intersection,” in *Proceedings of the 31st IEEE Intelligent Vehicles Symposium, Las Vegas, NV, USA*, vol. 19, 2020.
- 4 Koetsier, C., Fiosina, J., Gremmel, J. N., Müller, J. P., Woisetschläger, D. M., and Sester, M. (2022). Detection of anomalous vehicle trajectories using federated learning. *ISPRS Open Journal of Photogrammetry and Remote Sensing*, 4:100013.

## 4.5 Modeling the Interaction between Places and Human Mobility

Cheng Fu (Universität Zürich, CH)

License  Creative Commons BY 4.0 International license  
 © Cheng Fu

Place and human mobility are the two sides of the coin in an urban system: Human activities, including mobility, reflect the actual usage of places. On the other hand, the spatial organization of places determines the origin and destination of the daily trips, e.g., commuting flows start from residential areas to commercial areas. The purposes of the trips are also associated with the amenities provided by the destination places. Such dual relationships between place and human mobility can happen at different spatial, temporal, and behavioral scales.

By modeling the influence of places on driving behaviors and trajectories, our research team found that certain types of places are good predictors for estimating car accident risks for drivers [1]. We also applied places (e.g., POIs) as the semantic context of trajectories for adaptive simplification so that the segments of a trajectory near dense POIs are simplified less while segments near sparse POIs are simplified more [2].

Our recent interests mainly focus on how places may influence older adults. Many countries are experiencing fast aging. Infrastructures and facilities in cities however are primarily planned for working-age commuters. To achieve healthy aging, we thus need to understand the mobility patterns of the older adults and model their interactions with the places. The Mobility, Activity, and Social Interaction Study (MOASIS) project overcomes the digital gap of older adults by collecting mobility data from customized GPS loggers. The collected data are small in terms of the number of participants but rich in their demographic details and their physical and mental health status. We supervised several MSc theses regarding whether older adults’ visitation patterns to certain types of places can be good indicators to infer their physical and mental health.

For place modeling with big human mobility data, the conventional workflows model the activities in the place per se as the features for learning. The recent development of neural networks, and particularly deep learning networks, shows the capacity to embed complex contextual information. That provides more powerful computational models for more complex conceptual models on place modeling with mobility data.

## References

- 1 Brühwiler, L., Fu, C., Huang, H., Longhi, L. & Weibel, R. Predicting individuals' car accident risk by trajectory, driving events, and geographical context. *Computers, Environment And Urban Systems*. **93**, 101760 (2022), <https://doi.org/10.1016/j.compenvurbsys.2022.101760>
- 2 Fu, C., Huang, H. & Weibel, R. Adaptive simplification of GPS trajectories with geographic context – a quadtree-based approach. *International Journal Of Geographical Information Science*. pp. 1-28 (2020), <https://doi.org/10.1080/13658816.2020.1778003>

## 4.6 Horizon Europe: Introduction to R&I Funding in the Field of Mobility

Stephan Winter (The University of Melbourne, AU) and David Doerr (TÜV Rheinland – Köln, DE)

License © Creative Commons BY 4.0 International license  
 © Stephan Winter and David Doerr  
 URL <https://ec.europa.eu/research/pdf/horizon-europe/annex-5.pdf>

David Doerr, from *Nationale Kontaktstelle Klima, Energie, Mobilität für das EU-Rahmenprogramm für Forschung und Innovation “Horizont Europa”*, presented an introduction to mobility research funding available through Horizon Europe, the EU's key funding programme for research and innovation.

The EU Framework programmes are complementary to national funding and require European added value. They promote cooperation between researchers and innovators (such as companies, research centres). The priorities of the current Horizon Europe programme are *green – digital – innovative – open – resilient – participative*. Of the three pillars of the current framework, this talk focused on *Global Challenges and European Industrial Competitiveness: Climate, Energy and Mobility*, specifically Cluster 5: *Climate, Energy and Mobility*<sup>2</sup>. Actions 5 and 6 are about mobility.

Relevant for this seminar is *Destination 6: Safe, resilient transport and smart mobility services for passengers and goods*, where especially two topics are related to mobility data analytics: connected, cooperative, automated mobility (CCAM), and multimodal and sustainable transport systems.

Calls open in the Work Programme for 2021-2022 are:

- HORIZON-CL5-2022-D6-02-02: Urban logistics and planning: anticipating urban freight generation and demand including digitalisation of urban freight (Innovation Action)
- HORIZON-CL5-2022-D6-02-04: Accelerating the deployment of new and shared mobility services for the next decade (Innovation Action)
- HORIZON-CL5-2022-D6-02-05: Advanced multimodal network and traffic management for seamless door-to-door mobility of passengers and freight transport (Research and Innovation Action)

These calls are closing on 6 September 2022.

Related, with the same deadline, is a call in the *Missions Work Programme 2021-22 on Climate-neutral and Smart Cities*:

<sup>2</sup> <https://ec.europa.eu/research/pdf/horizon-europe/annex-5.pdf>

- HORIZON-MISS-2022-CIT-01-01: Designing inclusive, safe, affordable and sustainable urban mobility (Innovation Action)

For this action, a consortium must contain at least four cities as living labs, and additional four follower cities.

## 4.7 Reproducibility for Urban Mobility Analysis

*Daniel Nüst (Universität Münster, DE)*

License  Creative Commons BY 4.0 International license  
© Daniel Nüst

Main reference Daniel Nüst: “Reproducibility for Urban Mobility Analysis”, Zenodo, 2022.

URL <https://doi.org/10.5281/zenodo.6477034>

Open and reproducible research is a prerequisite for a sustainable and meaningful science. As I research in the context of geospatial sciences [1], the problems stemming from irreproducibility are substantial and solutions to improve reproducibility, albeit practical both regarding technology and culture, are too rarely implemented. If research is not reproducible it cannot be inspected, reused, or extended and thereby such works slow innovation and hinder the advancement of science. In my experience, challenges but also solutions from other disciplines can and should be transferred to mobility research to ensure the urban mobility researchers can help to solve societal challenges in the most transparent, effective and collaborative way. That is why I applaud the organisers of the Dagstuhl Seminar 22162 “Urban Mobility Analytics” [2] to provide time for the participants to learn about and discuss the foundations and challenges of reproducible research.

The session material is published at [3]. It covers theoretical and practical basics of reproducibility and draws from the large amount of open educational resources provided by the reproducible research and open science communities, e.g., [4, 5]. In the talk, I give advice on working reproducibly as individuals and present how communities can change their practices. For *individuals* to create reproducible research, I postulate the minimalistic motto “*Have a README: all else is details.*”, which is inspired by Greg Wilson’s first Rule of *Teaching Tech Together* and intends to make clear that reproducibility is an ideal to strive for with best efforts, not a binary property of a paper. I further stress the aspects of *craftpersonship* needed to realise computational reproducibility. The motto is extended to comprise different practices for integrating reproducible workflows into personal habits, such as good filenames, consciously managing the computational environment, using notebooks, publishing & citing code and data, and creating research compendia (cf. [6]). In *research groups or labs*, a very effective means to improve reproducibility-related skills and increase the quality of work are mutual reproductions amongst colleagues before submissions of manuscripts. For *community practices*, I argue that the establishing of code execution and reproductions as part of peer review is crucial to ensure adoption and proper credit. Two examples of successful initiatives for reproducibility reviews illustrate the feasibility of this approach [7, 8]. To convince the seminar participants to adopt these practices, to provide literature for further self-study, and to give material for sharing and convincing others, the session material includes an overview of relevant literature with a special eye on the individual benefits of working reproducibly.

Finally, I see the following *relevant topics for the urban mobility and transportation disciplines* with respect to open reproducible research. It can be noted that the topic is not widely discussed yet, with few mentions in the literature and no comprehensive studies on

the reproducibility of papers. The specific challenges are (i) the importance of (tailored, novel) hardware for research, e.g., for autonomous vehicles, which can not be shared digitally and is often expensive or unique, (ii) the high complexity of analyses, e.g., in the context of routing and networks, which often requires high performance computing methods, (iii) the lack of open, not proprietary, datasets and the bias of existing data towards specific modes of transport, i.e., cars, and (iv) the widespread use of machine learning and artificial intelligence (AI) approaches due to the complexity of the problems, which often are only shared as black boxes. Solutions for some these challenges do exist, such as synthetic data sharing and compiling new open datasets, or are an active field of research, such as explainable AI, but they to complicate matters and are likely to be used as an excuse to not conduct the steps for open reproducible research, which are still wrongly perceived as “extra work”. The close collaboration with industry partners and public authorities, not the least because they have the (closed) data, also comes with opportunities. The skill set required for reproducible computational workflows, such as research software engineering (RSEng, [9]), is highly relevant for individuals to transition to industry jobs. Furthermore, reproducible data analyses have a proven functionality and high reusability that can greatly reduce the efforts for putting them into practice.

For the next steps, I invite the seminar participants to initiate discussions on the importance of reproducible research across their different roles in academia, be they authors, reviewers, or editors, as much as I encourage them to try to improve their own habits. As a next step, understanding the state of reproducibility in urban mobility research seems necessary. However, the community may need to have a discourse about which of the goals for data-driven research are most relevant for them: *replicability*, albeit needing new data and being more work, or *robustness* checks through applying multiple methods on the same data? All of these goals help to unhide the data and code underlying computational and data-driven research in urban mobility and eventually ensure a recognition of their important contributions to science.

## References

- 1 Nüst, D. 2022. Infrastructures and Practices for Reproducible Research in Geography, Geosciences, and GIScience. Zenodo. <https://doi.org/10.5281/zenodo.4768096>
- 2 Dagstuhl Seminar 22162. Urban Mobility Analytics. D. Jonietz, M. Sester, K. Stewart, S. Winter (Organizers), M. Tomko (Coordinator). <https://www.dagstuhl.de/22162>
- 3 Nüst, D. 2022. Reproducibility for Urban Mobility Analysis. Zenodo. <https://doi.org/10.5281/zenodo.6477034>
- 4 The Turing Way Community. 2021. The Turing Way: A handbook for reproducible, ethical and collaborative research (1.0.1). Zenodo. <https://doi.org/10.5281/zenodo.5671094>
- 5 Reproducible Research Support Service (R2S2) Knowledge Base. <https://confluence.uni-muenster.de/display/r2s2/>
- 6 Nüst, D. & E. Pebesma. 2021. Practical reproducibility in geography and geosciences. *Annals of the American Association of Geographers*, 111(5), 1300–1310. [tps://doi.org/10.1080/24694452.2020.180602](https://doi.org/10.1080/24694452.2020.180602)
- 7 Reproducible AGILE. <https://reproducible-agile.github.io/>
- 8 Nüst, D., & S. J. Eglén. 2021. CODECHECK: An Open Science initiative for the independent execution of computations underlying research articles during peer review to improve reproducibility. *F1000Research*, 10, 253. <https://doi.org/10.12688/f1000research.51738>; <https://codecheck.org.uk/>
- 9 Cohen, J., D. S. Katz, M. Barker, N. Chue Hong, R. Haines and C. Jay. 2021. The Four Pillars of Research Software Engineering. *IEEE Software* 38(1), 97-105. <https://doi.org/10.1109/MS.2020.2973362>

## 5 Working Groups

### 5.1 Group on Ethics / Social Good

*Alexandra Millonig (AIT – Austrian Institute of Technology – Wien, AT), Ivan Majic (TU Graz, AT), Edoardo Neerhut (Meta – Burlingame, US), Moritz Neun (IARAI – Zürich, CH), Luca Pappalardo (CNR – Pisa, IT), Chiara Renso (ISTI-CNR – Pisa, IT), and Stephan Winter (The University of Melbourne, AU)*

License  Creative Commons BY 4.0 International license  
 © Alexandra Millonig, Ivan Majic, Edoardo Neerhut, Moritz Neun, Luca Pappalardo, Chiara Renso, and Stephan Winter

#### 5.1.1 Background

Going by paper titles at relevant conferences in both the intelligent transportation sector and the computing sector, much of current research on urban mobility analytics is spent on efficiency improvements of current mobility solutions: improving the efficiency of individual decisions on mobility choices such as modes (how to go), routes (where to go), or times (when to go), or transport network optimizations. All this research engages in predicting demand, providing supply, or managing systems. There are a number of issues with such thinking:

- It is treating human decision making as rational. But behavioral economists characterize human decision making as intrinsically tied to overconfidence, loss aversion, limited attention, and cognitive biases, all leading to systematic errors in judgment [1, 2].
- It is treating the accumulation of optimal decisions of individuals as the best achievable, which is a fallacy [3, 4].
- Its promised efficiency gains are just too little or too slow to respond to the current climate crisis and climate goals [5].

Therefore, the group discussion returned to the widely promoted three lines of action towards sustainable urban mobility, *avoid – shift – improve*<sup>3</sup>, which have been introduced in the early 1990s but have failed to succeed in the coming decades, particularly regarding the avoidance of transport and a significant shift from motorized to non-motorized forms of mobility. Therefore, the group questioned the solutions which have been developed in the past under this principle and asked whether we, as a scientific community, could (or even should), do more urban mobility analytics research on avoiding and shifting (motorized) mobility. For example, the *where to go* question seems to have ignored the option of choosing alternative destinations for certain activities that are closer in space or travel time, with few exceptions such as [6].

#### 5.1.2 Challenges and Opportunities for Urban Mobility Analytics

If change of mobility behavior is required, it should be incentivized, either by material or immaterial recognition, to trigger the human internal reward system and reinforcement learning [7]. More generally, the group asked whether negatively framed expressions, such as emissions-based or climate-based (reduction of a threat) might be too vague to trigger loss aversion biases, but positively framed expressions such as gains in quality of life (and rewards for contributions) are more tangible. One tested way of rewards in transport demand management is gamification [8], although incentives can also be linked to intrinsic motivations. However, the group is also aware of the limitations that gamification and

<sup>3</sup> e.g., <https://en.wikipedia.org/wiki/Avoid-Shift-Improve>

nudging approaches have in inducing behavior change across modes or even avoiding trips, as habits are particularly difficult to break if the alternative is not well known or is subject to prejudice. Positive reinforcement therefore demands bespoke messages to individual behavior and consequently very good insight into current behavior.

In this regard, the group identified one challenge: To monitor (prove) a change in behaviour of an individual, in order to reward, for example in cases when (a) people choose to walk rather than take motorised forms of transportation, or (b) people deliberately swap a longer distance trip for a nearer one, sacrificing some original intentions or rewards. The group sees here a potential application for urban mobility analytics. While gamification approaches reward individual decisions from a system's perspective (e.g., a public transport ride outside peak time), they typically do not link their rewards to a change in an individual's behavior. These changes could happen in each category: avoid, shift, or improve. Incentivization of individuals, while contributing to the societal good, raise significant ethical questions of privacy though.

The second area identified by the group where urban mobility analytics can make a difference is in supporting sustainable living in cities. Information derived from urban mobility analytics (Geo AI, mobility AI [9]) can be used to:

- Track, predict, and manage the impact of mobility on climate, with an immediate goal of climate neutrality. Air quality, temperature (health), and coping of infrastructure with weather events (urban resilience) are main factors of livability. Examples of this approach are [10, 11]. In this regard, current AI systems are optimised to satisfy the need of the individual only, without caring about the collective effects on the city. We should design AI systems that optimise for both individual needs and collective societal goods. For example, we should ensure that routing suggestions from navigation apps (e.g., Google Maps, Waze) optimise for societal good, and that this routing is advantaged over AI suggestions that do not.
- The spatial distribution of venues (places for activities) in a city significantly impacts citizens' mobility needs. The reduction of the impact of human mobility on climate and city well-being should be achieved not only by improving the efficiency of our transportation means, but also by reshaping our cities in terms, e.g., of how we distribute venues within them. Modern cities, in which neighborhoods contain clusters of a variety of economic or social activities, may reduce mobility demand, equalise transportation disadvantages, open public space for forms of use other than transport and can hence decrease the negative impacts of motorized traffic [12]. The 15-minute city<sup>4</sup> – a city of nearly self-sufficient neighborhoods ('superblocks') – should be tested more rigorously by means of data analytics, and what-if analyses about the spatial distribution of venues and their impact on mobility should be conducted and integrated in urban planners' body of knowledge [13, 14] and tools.
- Support the resilience of a city to epidemics. The COVID-19 pandemic made evident how fragile our cities are in facing epidemics [15, 16]. Indeed, limiting human mobility was needed to reduce the diffusion of the COVID-19 pandemic in cities all around the world. Policies for the reduction of impact of mobility on climate and well-being should take into account also the epidemics dimension, so to avoid mobility solutions that are potentially dangerous in cases of epidemics spread.

---

<sup>4</sup> [https://en.wikipedia.org/wiki/15-minute\\_city](https://en.wikipedia.org/wiki/15-minute_city)

### 5.1.3 Potential Contributions

In this regard, urban mobility analytics can be further developed to form a body of knowledge about the sensitivities of urban designs, and thus inform urban planners through novel decision-support systems.

#### Further references

- Carlos Moreno: The 15-minute city (<https://youtu.be/TQ2f4sJVXAI>)
- Andreas M. Dalsgaard, Jan Gehl: The Human Scale (<https://youtu.be/oA2eAQKkr-k>)

#### References

- 1 Gigerenzer, G. & Goldstein, D. Reasoning the Fast and Frugal Way: Models of Bounded Rationality. *Psychological Review*. **103**, 650-669 (1996)
- 2 Kahneman, D. Thinking, Fast and Slow. (Farrar, Straus,2011)
- 3 Tversky, A. & Kahneman, D. Judgement under Uncertainty: Heuristics and Biases. *Science*. **185**, 1124-1131 (1974)
- 4 Roughgarden, T. Selfish Routing and the Price of Anarchy. (MIT Press,2005)
- 5 IPCC Climate Change 2022: Impacts, Adaptation, and Vulnerability – Contribution of Working Group II to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change. (Intergovernmental Panel on Climate Change,2022), <https://www.ipcc.ch/report/ar6/wg2/>
- 6 Wang, Y., Winter, S. & Tomko, M. Collaborative activity-based ridesharing. *Journal Of Transport Geography*. **72**, 131-138 (2018)
- 7 Schultz, W. Neuronal Reward and Decision Signals: From Theories to Data. *Physiological Reviews*. **95**, 853-951 (2015)
- 8 Yen, B., Mulley, C. & Burke, M. Gamification in transport interventions: Another way to improve travel behavioural change. *Cities*. **85** pp. 140-149 (2019)
- 9 Luca, M., Barlacchi, G., Lepri, B. & Pappalardo, L. A Survey on Deep Learning for Human Mobility. *ACM Computing Surveys*. **55**, Article 7 (2021)
- 10 Nyhan, M., Sobolevsky, S., Kang, C., Robinson, P., Corti, A., Szell, M., Streets, D., Lu, Z., Britter, R., Barrett, S. & Others Predicting vehicular emissions in high spatial resolution using pervasively measured transportation data and microscopic emissions model. *Atmospheric Environment*. **140** pp. 352-363 (2016)
- 11 Böhm, M., Nanni, M. & Pappalardo, L. Gross polluters and vehicles' emissions reduction. (arXiv,2021)
- 12 Millonig, A., Rudloff, C., Richter, G., Lorenz, F. & Peer, S. Fair mobility budgets: A concept for achieving climate neutrality and transport equity. *Transportation Research Part D: Transport And Environment*. **103** pp. 103165 (2022), <https://www.sciencedirect.com/science/article/pii/S1361920921004600>
- 13 Christaller, W. Die zentralen Orte in Süddeutschland. (Gustav Fischer,1933)
- 14 Jacobs, J. The death and life of great American cities. (Random House,1961)
- 15 Kraemer, M., Yang, C., Gutierrez, B., Wu, C., Klein, B., Pigott David, M., Group, O., Plessis, L., Faria, N., Li, R., Hanage, W., Brownstein, J., Layan, M., Vespignani, A., Tian, H., Dye, C., Pybus, O. & Scarpino, S. The effect of human mobility and control measures on the COVID-19 epidemic in China. *Science*. **368**, 493-497 (2020)
- 16 Lucchini, L., Centellegher, S., Pappalardo, L., Gallotti, R., Privitera, F., Lepri, B. & De Nadai, M. Living in a pandemic: changes in mobility routines, social activity and adherence to COVID-19 protective measures. *Scientific Reports*. **11**, 1-12 (2021)

## 5.2 Group on Benchmarking and Datasets

*Vanessa Brum-Bastos (Wroclaw University of Environmental and Life Sci., PL), Christian Eichenberger (IARAI – Zürich, CH), Cheng Fu (Universität Zürich, CH), Erik Nygren (Schweizerische Bundesbahnen – Bern, CH), and Maya Sekeran (TU München, DE)*

License  Creative Commons BY 4.0 International license  
© Vanessa Brum-Bastos, Christian Eichenberger, Cheng Fu, Erik Nygren, and Maya Sekeran

Movement is a fundamental characteristic of life [1]. Humans move on a daily basis for the most diverse reasons, to multiple places and using varied transportation modes. The understanding of human mobility behaviour, i.e., “why”, “where”, “when” and “how” people move, is extremely relevant for urban planning, traffic engineering, policy making and other many applications [2]. Moreover, in face of climate changes and the recent COVID-19 pandemic, it even became pivotal to further understand human mobility in order to manage its related carbon footprint and epidemiological role.

### 5.2.1 Challenges and Problems

The recent technological advances in location-based services and devices have been providing researchers with unprecedented amounts of data to study human mobility [3]. However, the same vast amounts of data that brought new research opportunities in urban mobility analytics have also created a series of new challenges that still need to be addressed by the research community. Our group discussion focused on trying to identify and propose potential solutions to these problems.

1. **Benchmark datasets:** benchmark datasets are used for both training and testing of methodologies, but can also be helpful for evaluating the performance of unfamiliar datasets for implemented methods. The desire for benchmark datasets became clear during the discussions in breakout groups and the plenary. However, the multidisciplinary aspect of urban mobility analytics seems to call for not one but multiple benchmarking datasets that can attend the demands of specific applications within urban mobility analytics, such as traffic planning, routing or programming self-driving vehicles.
2. **Bias and representativeness of the data:** the massive amounts of data provided by modern location-based technologies come along with an increasing public and legal concern with individual’s privacy. These privacy concerns have led to data anonymization, meaning that researchers do not have any metadata describing the population sample (e.g., gender, age, income) that generated the mobility datasets and therefore are not aware of how representative of the total population these data are. Understanding the bias in modern mobility datasets is especially critical for applied research on urban mobility analytics for solving real world problems.
3. **Data availability and access:** Even though massive amounts of data on human mobility are produced daily, that does not mean that these data are accessible to researchers and policy makers. Many of these datasets are produced, owned and sold by private companies, such as Near and Strava. The high fees for such data may pose a barrier to many researchers, especially the ones working at countries with less advantaged economies. There are also datasets owned by governmental agencies and public companies, however, even when those are available, many of them are not widely known by or not readily accessible to the research community.

### 5.2.2 Potential Paths for Future Solutions

We believe that a unified international data catalog could be a good starting point for addressing the aforementioned issues. The development of a unified international data catalog has been a strategy used by movement ecology (see [4]), a discipline that has also recently benefited from the new opportunities and challenges brought by location-based services and devices [5]. In fact, a discipline that has been working for decades on challenges that urban movement analytics researchers are just now facing.

We believe that unified international data catalogue would support and foster a growing research community as well as potentially bring the following benefits:

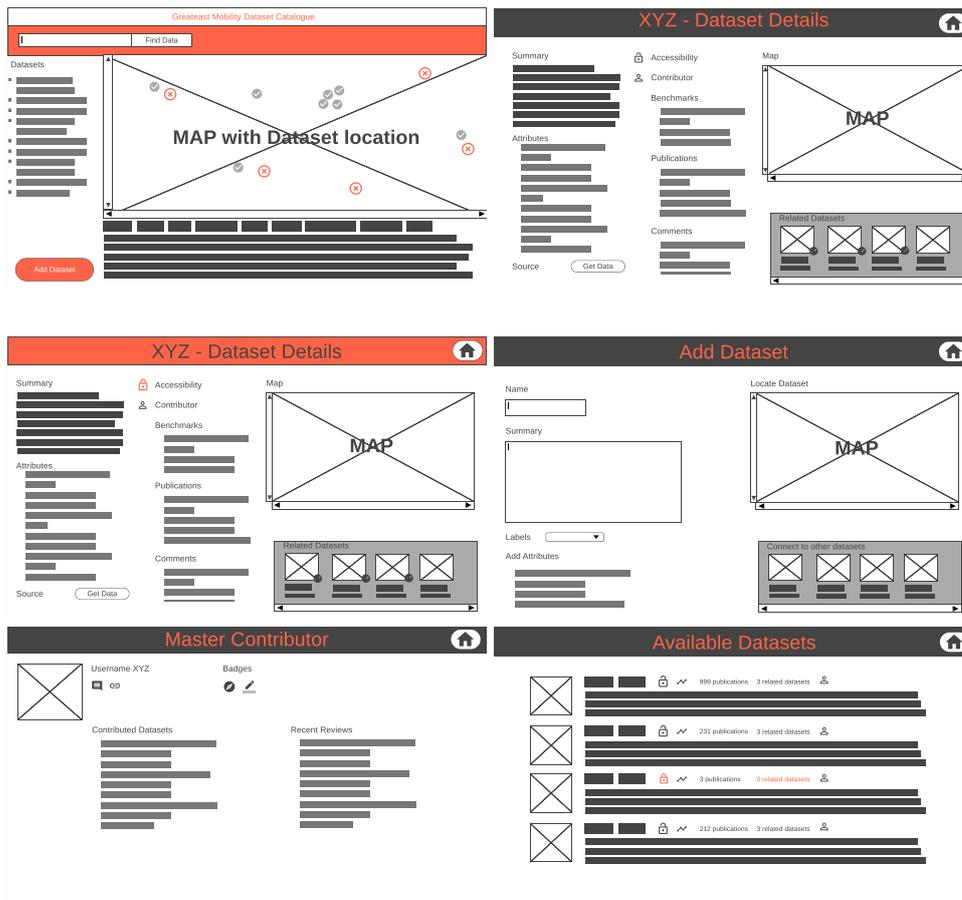
1. **Federated cataloguing data sets on human and human-related movement** as a centralised international index. The goal is to facilitate the access to and exploration of data sets that would otherwise be unknown, while supporting research collaboration at global, regional and local level. Moreover, this is also an opportunity to maximize the financial resources applied for data purchase by avoiding buying duplicate data from companies.
2. **Data reviewing across disciplines** The catalogue should be a forum for the community from different fields to put data sets into perspective with first-hand experience and related contextual data such as census data, fostering transparency on openness of data sets. It supports best practices for cross-research work. Data sets come with papers and code they have been used in, linking to contextual and supplementary material.
3. **Data collection waste reduction** The catalogue helps to foster research collaboration and to establish networks in order to reduce data collection redundancy and improve the use of the limited financial resources.
4. **Data fusion and multi-modality** The catalogue is a channel for data compilation where data sets on different transportation modes can be combined to answer specific research questions, while also encouraging multi-modal transportation research and planning.
5. **Methodological development** In its core, the catalogue is data and benchmark centered. However, it is supposed to become a central starting point for developing methodologies and state-of-the-art benchmarks and enhance reproducibility in human movement analytics studies across disciplines, in particular for bias detection. It supports discovering transferability opportunities from solutions derived from the data shared.

### 5.2.3 Mock-Ups

From the discussions held on how a mobility data catalog should look like, the overall motivation is to provide a mobility data platform which encourages open research and reproducibility.

Figure 2 show mock-up interfaces with different functionalities. Other than the typical search bar, a map with dataset locations and accessibility levels helps users to immediately have a sense of where mobility data is being collected, how to access them and the intensity of data collection activities in a certain location.

Upon clicking the dataset of interest, the next page offers a detailed description of the dataset which includes a summary, dataset attributes, benchmark, publications, comments and related datasets. In this way, users are able to understand how the dataset has been used for research and improvement opportunities from the publications that used the particular dataset. The related dataset offers possibility to validate findings and implement models or algorithms using other sources of data as well.



■ **Figure 2** Mockup of Mobility Dataset Catalogue.

To encourage data contributions, contributor profiles include badges that acts as acknowledgements for offering the data and follow up badges for open research, etc.

Submitting data is easily done by a click of a button and a simple submission interface and finally an overview page provides a list of datasets with accessibility levels, number of views, number of publications, related datasets and a link to the contributor profile.

#### 5.2.4 Discussion and Possible Extensions

Obviously, a catalogue is only as good as its users and curators. If it does not provide any insights or does not suit the way people work, it will be a non-starter. However, data repository and engineering workflow (“github for datasets”) or an integrated analysis platform (see e.g. [6]) would mean a huge conceptual effort that requires engineering and platform resources, which is beyond our scope, hence, the reason why we envision the catalogue to be a more special-purpose combination of <https://paperswithcode.com/datasets> and <https://openreview.net/> known in the ML community.

#### References

- 1 Nathan, R., Getz Wayne, M., Revilla, E., Holyoak, M., Kadmon, R., Saltz, D. & Smouse Peter, E. A movement ecology paradigm for unifying organismal movement re-

- search. *Proceedings Of The National Academy Of Sciences*. **105**, 19052-19059 (2008), <https://doi.org/10.1073/pnas.0800375105>
- 2 Brum-Bastos, V., Long, J. & Demšar, U. Weather effects on human mobility: a study using multi-channel sequence analysis. *Computers, Environment And Urban Systems*. **71** pp. 131-152 (2018)
  - 3 Demšar, U., Long, J., Benitez-Paez, F., Brum-Bastos, V., Marion, S., Martin, G., Sekulić, S., Smolak, K., Zein, B. & Siła-Nowicka, K. Establishing the Integrated Science of Movement: bringing together concepts and methods from animal and human movement analysis. *International Journal Of Geographical Information Science*. **35**, 1273-1308 (2021)
  - 4 Kranstauber, B., Cameron, A., Weinzerl, R., Fountain, T., Tilak, S., Wikelski, M. & Kays, R. The Movebank data model for animal tracking. *Environmental Modelling And Software*. **26**, 834-835 (2011)
  - 5 Cagnacci, F., Boitani, L., Powell, R. & Boyce, M. Animal ecology meets GPS-based radiotelemetry: a perfect storm of opportunities and challenges.. *Philosophical Transactions Of The Royal Society Of London. Series B, Biological Sciences*. **365**, 2157-62 (2010,7)
  - 6 Tomko, M., Bayliss, C., Galang, G., Greenwood, P., Koetsier, J., Mannix, D., Morandini, L., Nino-Ruiz, M., Pettit, C., Sarwar, M., Voorsluys, W., Widjaja, I., Stimson, R. & Sinnott, R. The AURIN e-Infrastructure: design, development and delivery. (2012)

### 5.3 Group on Methods and Explainability

Anita Graser (AIT – Austrian Institute of Technology – Wien, AT), Hao Cheng (Leibniz Universität Hannover, DE), Tao Cheng (University College London, GB), Ioannis Giannopoulos (TU Wien, AT), Daniel Nüst (Universität Münster, DE), Martin Tomko (The University of Melbourne, AU), and Yanan Xin (ETH Zürich, CH)

License © Creative Commons BY 4.0 International license  
 © Anita Graser, Hao Cheng, Tao Cheng, Ioannis Giannopoulos, Daniel Nüst, Martin Tomko, and Yanan Xin

#### 5.3.1 Introduction

Today's research and analytical methods, such as machine learning, rely heavily on tooling (software and scripts) capturing analysis as *code*, and on data. Any computational research poses strong challenges for effective communication and collaboration since both data and code are not well shared through the traditional method of scholarly communication, the scientific paper. If research outcomes cannot be shared, they can not transparently inform practices to improve mobility management in society.

The framework presented in Figure 3 framed the discussion in this group. Applicable to any computational research, this framework ties together the requirements on data (further discussed by a separate breakout group) and analysis.

Currently, many computational mobility analysis methods are de-facto *black boxes*. This may be either because the methods are only vaguely described in papers (often to an insufficient level of formality or because these methods are non-trivial to be then implemented in a computational environment) or because of the nature of the methods themselves, (i.e., the hard to explain and interpret computational procedures learned from specific datasets, such as in the case of deep neural networks). In cases where no human directly implemented assumptions and models, but a computer derived (learned) patterns from datasets, *explainability* must be especially considered to provide openness and transparency. Lack of explainability can result in erosion of trust from users who may question the results.

		Data	
		Same	Different
Analysis	Same	Reproducible	Replicable
	Different	Robust	Generalisable

■ **Figure 3** The reproducibility decision space.

This issue is also related to ethics because resulting recommendations may discriminate against certain groups due to biased training data or other issues. Furthermore, without explainability, theoretical frameworks cannot advance and therefore, scientific progress is slowed down.

We argued that a comprehensive publication of methods and their explainability is required to enable understanding, reuse, and extension of pieces of research so that societal challenges can be tackled effectively. Methods and the explainability of methods were discussed in this group along with possible avenues for improvement as a move towards better transparency: improvement in generalisability and robustness of methods, the ability to re-use the state of the art, and thus advance the discipline in a transparent manner.

### 5.3.2 Mobility Analytics Methods Research Agenda

The group collected topics broadly reflecting on current methods underpinning mobility analysis. This list naturally shows the breadth of the methodological toolkit needed to model and analyse mobility:

1. Explainability and transferability in deep learning models for human mobility [1];
2. Data integration, incl. geospatial encodings and embeddings for AI and challenges around vector vs. raster data;
3. How to separate *outlier* patterns of interest in large datasets from large amounts of *standard* situations;
4. Integrating physics-based models with data-driven models for analyzing mobility data to achieve more interpretable results;
5. Investigation of trajectory analytical methods for modeling travel behaviours using network-based trajectory data rather than grid-based mobility data;
6. The challenge of analysing interactions between modelled individuals, who are often assumed to be independent by mobility predictors and generators;
7. Models that capture the complexity of mobility more comprehensively, such as the complexities of dynamic structures of transport networks and multi-modal routing;
8. How to comprehensively model urban mobility across modalities (road traffic, public transport, pedestrian/bike), incl using deep learning (DL); and
9. Multi-modal networks performance optimization.

To identify missing methods that need to be developed, domain knowledge and context information are of vital importance. Overall, the design of methods should be further improved to satisfy the theoretical constraints of mobility patterns, instead of just relying on training complex models and large amounts of data to lead to usable outcomes.

### 5.3.3 Advancing the Scientific Mobility Analytics Toolbox

Participants with a background in scientific software development emphasized the need for a *consolidation in tool development* based on replicable and reusable software principles, thus enabling *reference implementations* of critical partial methods. This would enable us to compare, consolidate, and advance the discipline that is currently fragmented across tools and approaches<sup>5</sup>.

A consolidation of tools and their open publishing is a prerequisite for more collaborative and effective development of tools, where tool extension and improvement to fix real world problems are valued more than marginal methodological improvements. Understanding and theory building should be the goal, i.e., scientific progress, not merely addressing an engineering challenge.

The group also discussed possible incentives to contribute to joint efforts enabling *methodological explainability*. Participants were interested in inspecting DL models and increasing the understanding of model behaviour. A key question is: what methods exist to explain the inner workings of DL models but also their inputs. Potential explainability approaches include: transfer of methods from computer vision, model comparison (especially for simple/simplified models), or visualisation of intermediate layers for understanding outputs.

*Understanding and interpretability* were also discussed on a more abstract level, that is: What kind of interpretability do we want for which type of analysis? Do we want to understand the mechanisms for generating the predictions, or make the actual model understandable to humans?

The participants share a quite critical view of the current state of methods in urban mobility analysis. There are concerns about the usefulness, usability, and relevance of methods, especially modern yet complex and possibly questionable methods based on AI and DL. These shared concerns motivate the following definition of challenges and recommendations.

### 5.3.4 Challenges in Developing Methods for Mobility Analysis

#### ■ **Challenge 1: defining appropriate objectives to evaluate machine learning models**

High accuracy alone is often inadequate to ensure accountable and reliable deployment of models in practice. We also need to offer explanations to end-users as to why the model produces a certain prediction (such as a routing suggestion) to guarantee transparency, fairness, and reliability [2]. In these cases, explainability should also be considered as one of the objectives besides accuracy. But explainability is a fuzzy concept and can mean various things in different contexts. This makes it an extremely challenging objective to optimize for [3]. Whether an explanation achieves its desired goal also depends on who receives it, how the person perceives it, and what actions a person can take based on the explanation.

#### ■ **Challenge 2: bridging a theory-driven approach with a data-driven approach**

It is still unclear how to effectively combine the data-driven approach with the theory-driven approach in mobility analysis. A key barrier to the integration is a lack of shared data, tools, and terminologies across different research communities working on mobility analysis (e.g., city planning, transportation engineering, geography, computer science, etc.).

---

<sup>5</sup> See a list at <https://github.com/anitagraser/movement-analysis-tools>

### 5.3.5 Recommendations

To address the above mentioned challenges, we have identified the following steps revolving around the development of shared resources for mobility analysis:

- **Recommendation 1: joint terminology**  
Fundamental terminological glossaries are needed to enable communication within the community. This terminological glossary would also relate to best practice about *methodological* translation of a mobility/transport concept (stop, home location, activity, behaviour) to the computational methods realising it, and an argument about their appropriateness in a given setting.
- **Recommendation 2: reference implementations**  
Standard, community-verified implementations ( cross-linguistic, possibly, to enable implementation in the main programming languages) would enable contributors to build robust solutions.
- **Recommendation 3: standard dataset**  
A set of simple to more complex shared datasets (e.g., basic trajectories, semantic trajectories, related urban structure data) are needed to test and evaluate computational methods – see recommendations of Group 2.
- **Recommendation 4: standard data formats**  
A definition of a common data representation for movement data would enable an interoperable (incl multi-language) interface to the data (see, e.g., the OGD MovingFeatures standard or the standardisation on input formats in geoparquet for python/geopandas and R).
- **Recommendation 5: shared ML models**  
Reference implementations of basic ML methods and a platform for sharing benchmarks or established models built on top of these methods could become a community convergence point (similar to *Hugging Face*<sup>6</sup> in the ML community).

#### References

- 1 Luca, M., Barlacchi, G., Lepri, B. & Pappalardo, L. A Survey on Deep Learning for Human Mobility. *ACM Computing Surveys*. **55**, Article 7 (2021)
- 2 Doshi-Velez, F. & Kim, B. Towards a rigorous science of interpretable machine learning. *ArXiv Preprint ArXiv:1702.08608*. (2017)
- 3 Lipton, Z. The Mythos of Model Interpretability: In machine learning, the concept of interpretability is both important and slippery.. *Queue*. **16**, 31-57 (2018)

---

<sup>6</sup> Hugging Face: <https://huggingface.co/>

## 5.4 Group on Applications

*Martin Lauer (KIT – Karlsruher Institut für Technologie, DE), Andris Clio (Georgia Institute of Technology – Atlanta, US), Dirk Christian Mattfeld (TU Braunschweig, DE), and Monika Sester (Leibniz Universität Hannover, DE)*

License  Creative Commons BY 4.0 International license  
© Martin Lauer, Andris Clio, Dirk Christian Mattfeld, and Monika Sester

### 5.4.1 Goal

The goal of this discussion group was to generate ideas in which results of urban mobility analytics can be used in applications and how the respective application areas can be linked to urban mobility analytics. After a brainstorming phase the group concentrated on the fact that the knowledge of mobility behavior is of very much use in the domain of autonomous driving, in which autonomous vehicles interact with other traffic participant in urban traffic. To achieve an optimal behavior, the autonomous vehicles require knowledge about the behavior of other traffic participants to adapt their own behavior and to interact seamlessly with others. It is expected that for a long time human driven vehicles, autonomous vehicles, bicyclists, pedestrians and other traffic participants will share the same roads; therefore, models of pedestrians' behavior, bicyclists' behavior, and the behavior of human drivers is very much relevant. Furthermore, the behavior of traffic participants cannot be assumed to be the same worldwide but it differs from country to country. E.g. the driving style in Italy obviously differs from the driving style in Germany, the U.S. or in India or China. Similarly, pedestrians behave differently in those country. Hence, the autonomous vehicle must be able to understand the country specific behavior of other traffic participants, and also adapt their behaviour accordingly.

In the following, the aspects data, analysis methods, and necessary stakeholders were discussed, which would be needed to push forward such an application.

### 5.4.2 Data

The country-specific adaptation of behavior models can be considered as a typical application of machine learning methods, which are based on revealing this information from large amounts of observed data. So the question is how these data can be generated. Several data sources have been discussed in the group, including

- the usage of existing datasets like *highD*<sup>7</sup>, *INTERACTION*<sup>8</sup>, *LUMPI*<sup>9</sup>
- datasets for environmental perception (e.g. Kitty, nuScenes)
- recording own datasets from top view perspectives using drones or mounting cameras on high buildings
- recording own datasets from test fields, e.g. Testfeld Karlsruhe<sup>10</sup>

Besides these open data sets, commercial data would be very beneficial. Data should be available in different European cities and potentially in further cities outside of Europe.

Since it is impossible to record and model the behavior of traffic participants in all possible traffic situations it would be necessary to start with limiting the recordings to some most interesting, and critical, scenarios. These could include

<sup>7</sup> highD: <https://www.highd-dataset.com/>

<sup>8</sup> INTERACTION: <http://interaction-dataset.com/>

<sup>9</sup> LUMPI: [https://data.uni-hannover.de/cs\\_CZ/dataset/lumpi](https://data.uni-hannover.de/cs_CZ/dataset/lumpi)

<sup>10</sup> Testfeld Autonomes Fahren Baden-Württemberg <https://taf-bw.de/>

- merging into highway ramps
- shared spaces
- crossing intersections
- vehicle following

To model these scenarios it would be necessary to extract relevant features from the data including

- trajectory data (position over time, velocity, acceleration)
- relationship data (relative distances, relative velocities between traffic participants, interactions)
- movement patterns (e.g. groups)
- typical patterns at certain locations
- geometric data (shape of traffic participants, layout of road infrastructure, lanes, curbs)
- intentional data (face and body orientation of pedestrians, roll angle of bicyclists, indicator lights)

### 5.4.3 Data Analysis Methods

The analysis of the data requires automatic methods and processes, e.g. the extraction of the features mentioned above from the sensory data, the creation of behavior models for traffic participants, and the integration of those models into the decision making and trajectory planning system of the autonomous vehicle. Urban mobility analytics comes in play especially for the second category of methods that analyze the recorded mobility data. Various methods might be suitable, including deep learning methods, transformers, recurrent neural networks, inverse reinforcement learning. In addition, also methods from computational geometry, e.g. trajectory pattern analysis, are relevant. An important aspect is to include the notion of cooperation into the process, which allows the autonomous vehicle to also exploit information from other traffic participants, which enlarges its own perception range.

For the integration of behavior models to decision making and trajectory planning of the autonomous vehicle it would be useful to have a powerful simulation environment available, that also should be adapted to country specific behavior. Finally, for a real live demonstration the implementation of the behavior models in a real autonomous vehicle would be desirable.

### 5.4.4 Stakeholders

A joint project to extract and exploit country specific behavior patterns for autonomous vehicles could be relevant for various partners in research and industry. The following stakeholders could be included

- researchers in urban mobility analytics: analyze data, build behavior models, extract country specific parameters; develop models and concepts that allow for exploiting shared data, without the need to make it publicly available (e.g. using federated learning ([1]))
- researchers in perception: record sensory data and extract relevant features
- researchers in traffic psychology: analyze and explain country specific behavior patterns
- researchers in autonomous driving: provide autonomous vehicle, integrate behavior models in decision making of autonomous vehicles
- automobile industry: integrate country specific behavior models into their vehicles, provide data from onboard sensors
- data and map companies: collect and analyze country specific data; specify map interfaces for country specific behavior patterns
- municipalities: provide access to infrastructure, equip certain areas with sensors, implement certain regulations (e.g. temporal speed limits)

## References

- 1 Koetsier, C., Fiosina, J., Gremmel, J., Müller, J., Woisetschläger, D. & Sester, M. Detection of anomalous vehicle trajectories using federated learning. *ISPRS Open Journal Of Photogrammetry And Remote Sensing*. 4 pp. 100013 (2022), <https://www.sciencedirect.com/science/article/pii/S2667393222000023>

## 5.5 Outcomes

*Monika Sester (Leibniz Universität Hannover, DE), Martin Tomko (The University of Melbourne, AU), and Stephan Winter (The University of Melbourne, AU)*

License  Creative Commons BY 4.0 International license  
© Monika Sester, Martin Tomko, and Stephan Winter

The outcomes of the breakout group discussions were collected and shared on Friday, and further discussed.

In this session, the industry partners rated the cooperation and exchange with universities as very relevant to them. The caveat is usually, that collaboration attempts within industry are not considered as mainstream by all employees yet, but heavily depends on convinced individuals. This often leads to only ad-hoc projects – unless there is already an established long-standing cooperation based on contracts. It was also pointed out that industry partners could be very relevant for tool development, or that a Dagstuhl Seminar focusing on an innovative tool could be very attractive for industry.

As a way forward, the participants committed to:

- Extracting a vision paper out of this report, for a venue such as the ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems 2022
- Developing tools and teaching material for a Summer School in Spring 2023
- A potential follow-up Dagstuhl Seminar closer focusing either on the core of mobility analytics or on the application of the emerging body of knowledge for climate-neutral urban mobility
- In addition, an number of individual collaboration themes were identified (internships, co-supervisions, visits).

## Participants

- Vanessa Brum-Bastos  
Wroclaw University of  
Environmental and Life Sci., PL
- Hao Cheng  
Leibniz Universität  
Hannover, DE
- Tao Cheng  
University College London, GB
- David Doerr  
TÜV Rheinland – Köln, DE
- Christian Eichenberger  
IARAI – Zürich, CH
- Cheng Fu  
Universität Zürich, CH
- Martin Lauer  
Karlsruher Institut für  
Technologie, DE
- Dirk Christian Mattfeld  
TU Braunschweig, DE
- Alexandra Millonig  
Austrian Institute of Technology –  
Wien, AT
- Edoardo Neerhut  
Meta – Burlingame, US
- Moritz Neun  
IARAI – Zürich, CH
- Daniel Nüst  
Universität Münster, DE
- Erik Nygren  
Schweizerische Bundesbahnen –  
Bern, CH
- Maya Sekeran  
TU München, DE
- Monika Sester  
Leibniz Universität  
Hannover, DE
- Martin Tomko  
The University of Melbourne, AU
- Stephan Winter  
The University of Melbourne, AU
- Yanan Xin  
ETH Zürich, CH



## Remote Participants

- Andris Clio  
Georgia Institute of Technology –  
Atlanta, US
- Ioannis Giannopoulos  
TU Wien, AT
- Anita Graser  
Austrian Institute of Technology –  
Wien, AT
- Ivan Majic  
TU Graz, AT
- Michael Nolting  
Volkswagen Nutzfahrzeuge –  
Hannover, DE
- Luca Pappalardo  
CNR – Pisa, IT
- Chiara Renso  
ISTI-CNR – Pisa, IT
- Piyushimita Vonu Thakuriah  
Rutgers University – New  
Brunswick, US