Issue 1 | September 2018

MASTER

Multiple ASpects TrajEctoRy management and analysis

Inside this issue

This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Sklodowska-Curie grant No 777695 Duration: March 1 2018 | February 28 2022

Editorial

Welcome to the first issue of the **MASTER newsletter.** The objective of this newsletter is to keep the consortium updated with the main activities that are going on during the project, summarizing the main interesting results in terms of secondments, training, scientific results, vision and dissemination activities. We also want to communicate our main outcomes to a wider audience, find more opportunities for collaborations and exchange of ideas with interested readers outside the consortium.

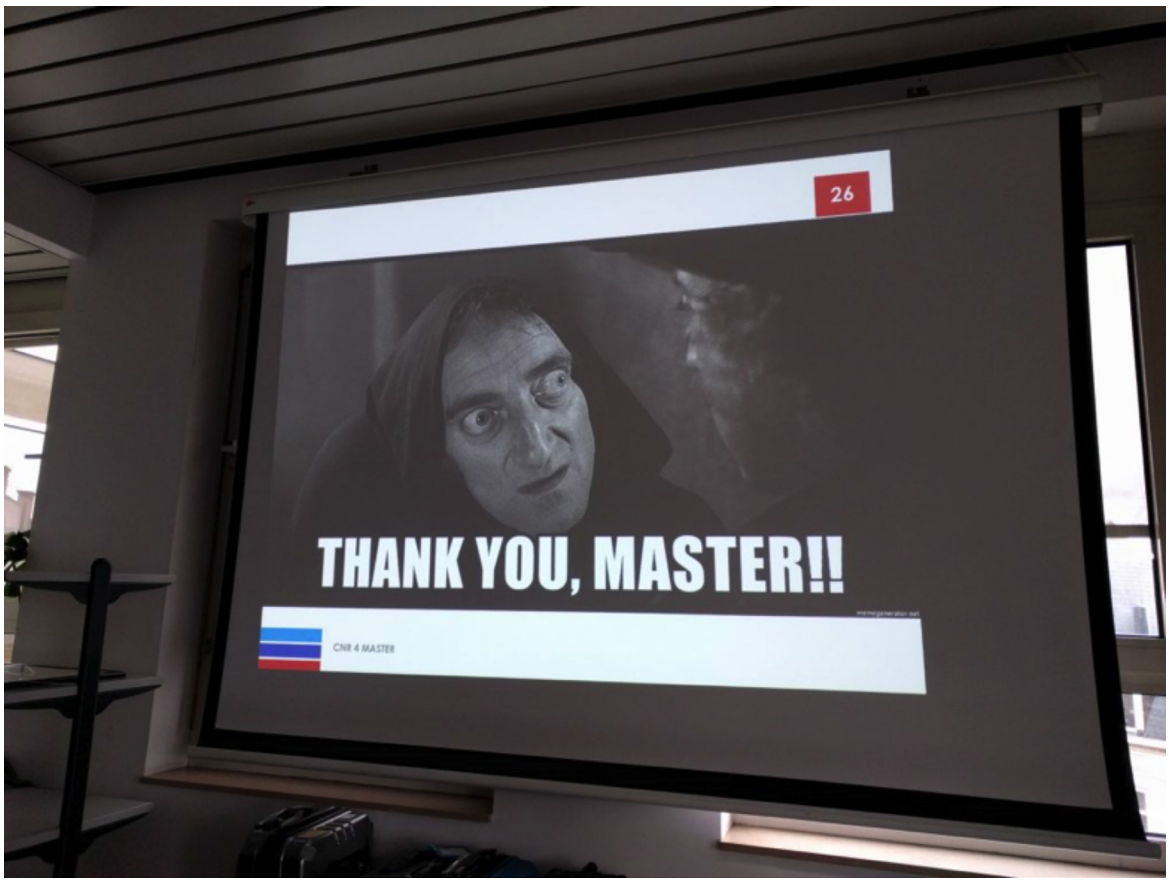
Our newsletter will be published every 6 months. Each issue will be available for download from the MASTER web site at http://www.master-project-h2020.eu

In this first issue we introduce **the project, the consortium and report about our kick off event held the last March 2018.**
We are very happy and proud that our **Independent Ethical Advisor Prof. Bettina Berendt** accepted our invitation to prepare a very stimulating article expressing her vision about the need of ethical consideration in data science for holistic trajectories. Prof. Berendt shares her thoughts on two case studies: the well known New York taxi dataset and the controversial topic of the analysis of ships traces for the detection of search and rescue operations in the Mediterranean Sea, linked to the migration phenomena. Her vision will be our starting point for reflecting on the ethical issues that our research is posing.

The next issue will be published at the end of February 2019: we will report results from our secondments, selecting highlights on the major research results and the plan for activities like the first MASTER Workshop and the planned Summer School.

Stay tuned and happy reading!

Chiara Renso, Project Coordinator



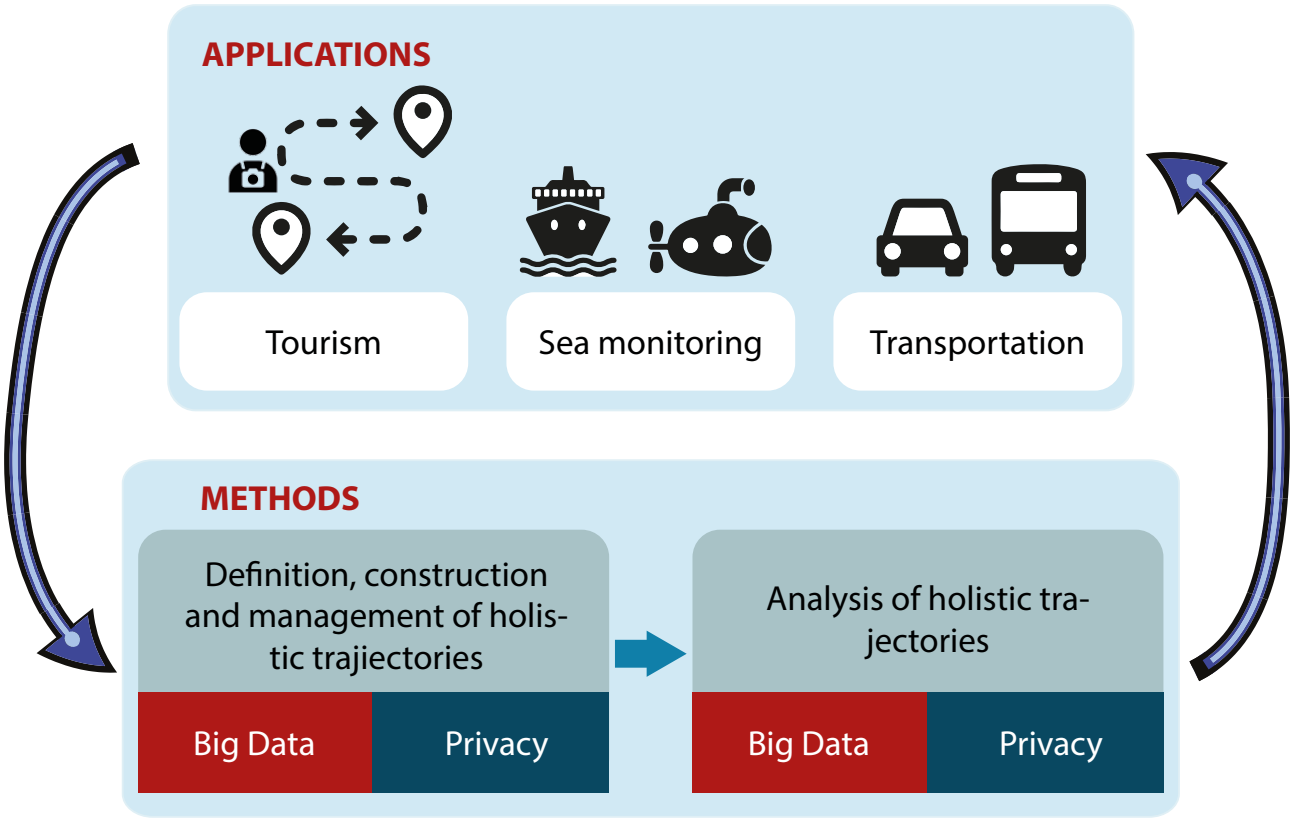
The MASTER project

MASTER is a Horizon 2020 project under the programme Marie Sklodowska-Curie Actions - Research and Innovation Staff Exchange (RISE) for boosting the carrier perspectives of researchers through staff exchange

Chiara Renso, Project Coordinator, HPC Lab, ISTI-CNR, Italy

In our everyday life we interact with a number of diverse, real-life applications that produce massive amounts of spatio-temporal data representing trajectories of moving objects. For example, we can considering mobile phone calls, GPS enabled tracking applications, (geotagged) social media but also land, sea, and air surveillance systems with smart sensors and public transportation like buses and metro. In all these areas, positioning data is produced, collected, stored and analysed together with additional contextual data. These trajectory data is therefore evolving from pure sequences of timestamps and position coordinates to more comprehensive and semantically significant objects representing the multiple aspects of movement data.

In the MASTER (Multiple Aspects Trajectory Management and Analysis) project we call them holistic trajectories, meaning trajectories characterized by the fact that space, time and semantics are all different aspects that are intimately correlated and should be considered as a whole. Current state-of-the-art methods do not provide methods "ready for use" for these multiple aspects trajectories.



The concept of the MASTER research programme

THE MASTER PROJECT IN 5 QUESTIONS:

How can we create and manage holistic trajectories?

We investigate methods for managing and building semantically rich trajectories from heterogeneous and multidimensional data.

How can we infer interesting knowledge from large amounts of holistic trajectories?

We study data analysis methods capable of taking into account the different aspects of holistic trajectories. We focus specifically on: similarity analysis, clustering, graph analysis, prediction and recommendation.

Which applications benefit from the analysis of holistic trajectories?

We assess the impact of the holistic trajectories analysis methods in the tourism, sea monitoring and public transpor-

tation domains.

Is the privacy preserved?

The project develops methods based on privacy-by-design.

How can we deal with the Big Data characteristics?

The storage and analysis methods will pose emphasis on efficiency and large scale data management.

plementary domains: tourism, sea monitoring and public transportation. On the other hand, the methods that are being developed can stimulate new requirements in the application domains. Big Data principles have to be considered as part of the methods. Privacy is a pervasive ethical aspect when dealing with positioning data, therefore, in MASTER, we have specific tasks for privacy pre-

one is European non academic and four are international academic.

The **European academic partners** are:

National Research Council of Italy (CNR) participates with the ISTI institute (Istituto di Scienze e Tecnologie dell'Informazione) located in Pisa in Italy and it is the coordinator of the project. The main expertize of CNR researchers are

The consortium at the kick-off meeting

MASTER

MASTER is a H2020 Marie Sklodowska-Curie Actions RISE (Research and Innovation Staff Exchange) project. The overarching objective of MASTER is to form an international and inter-sectoral network of partners working on a joint research programme to answer the above research challenges. The concept of MASTER is depicted in the figure. We see that the research objective of the project is to develop methods to build, manage and analyse holistic trajectories. These methods are driven by application scenarios from three different and com-

serving methods for holistic trajectories. Furthermore, the consortium is supported by an Independent Ethical Advisor (Prof. Bettina Berendt) and an Ethical Committee (Dr.ssa Rosaria de Luca, Prof. Tommaso Piazza, Prof. Celia Zolynski). A precious first contribution from Prof. Berendt giving her ethical view on two typical cases of privacy in trajectory data is reported later in the newsletter.

THE CONSORTIUM

MASTER consortium is formed by ten synergic and complementary partners: five partners are European academic,

in mobility data and semantics, machine learning, data mining for spatio-temporal data. The principal investigator and coordinator of the project is Dr. Chiara Renso. Members of the research group are Raffaele Perego, Cristina Ioana Muntean, Beatrice Rapisarda, Ida Mele, Vinicius Monteiro de Lira, Andrea Michienzi.

University of Ca' Foscari Venice (UNIVE) is located in Venice in Italy. The main expertise is in Trajectory Data warehousing, efficient and scalable methods for storage, mining and quantitative mod-

elling of agents' behaviours. Principal Investigator is Dr. Alessandra Raffaeta' and members of the research group include Marta Simeoni, Elisabetta Russo, Salvatore Orlando, Claudio Lucchese, Fabio Pranovi, Andrea Marin and Claudio Silvestri.

University of Pireaus Research Center (UPRC) is located in Piraeus in Greece. Main expertize is in spatio-temporal databases and analysis methods, privacy, big data analysis for movement data. Principal investigator is Prof. Nikos Pelekis. The research group includes Prof. Yannis Theodoridis.

University of Versailles Saint Quentin (UVSQ) is located in the west of Paris, and its science campus in Versailles city. UVSQ main expertize is trajectory data modeling, indexing, mining, as well as privacy, with a focus on trajectories following a predefined network. UVSQ also investigates time dependent graphs either by aggregation of trajectories or as their underlying context. Principal investigator is Prof. Karine Zeitouni. The research group includes Prof. Yehia Taher, and Prof. Iulian Sandu Popa, Alexandros Kontarinis and Jingwei Zuo.

Harokopio University (HUA) is located in Athens in Greece. The main expertize is in sea monitoring analysis from vessels data, machine learning and big data analysis. Principal investigator is Prof. Konstantinos Tserpes. Members of the group include Iraklis Varlamis.
The European non academic partner is the **Thira Municipality (Thira)** in the Santorini island in Greece. Thira is represented in MASTER by Christallia Papaoikonomou who is supporting seconded researchers to better understand the tourism issues of the island, the data that can be collected and the kinds of analysis that can be done on such data.

The **international partners** are:

The Federal University of Santa Catarina (UFSC) is located in Florianopolis in the state of Santa Catarina in the south of Brazil. UFSC experize is on semantic trajectory modelling and similarity measures. Principal investigator is Prof. Vania Bogorny. Members of the group are Prof. Ronaldo Melo, Prof. Luis Otavio Alvares and student Lucas May Petry.

The Federal University of Ceara' (UFC) is located in Fortaleza in the state of Ceara' in the north of Brazil. UFC expertize is on mobility patterns and prediction, trajectory reconstruction and graph modelling. Principal investigator is Prof. Jose Fernandes de Macedo. Members of the group include Francesco Lettich, Igo Brilhante and the students Livia Almada, Ticiana Linhares.

The Pontifical Catholic University of Rio de Janiero (PUC) is located in the city of Rio de Janeiro in Brazil. PUC expertize is in semantic web, social media analysis and in transportation applications. Principal investigator is Prof. Marco Antonio Casanova.
The Dalhousie University (DAL) is located in the city of Halifax in the Nova Scotia province in the eastern Canada. DAL expertize is in machine learning, trajectory classification, privacy, big data. Principal investigator is Prof. Stan Matwin. Members of the group include Amilcar Soares Junior and Mohammad Etemad.

KICK-OFF MEETING

The Kick off meeting has been a great opportunity to meet "live" the European Academic partners, our Project Officer Simona Losmanova from REA and the Independent Ethical Advisor Prof. Bettina Berendt from University of Leuven.
The meeting was held in Brussels thanks to the kind hospitality of the CNR Brussels office at Rue du Trone, 98 on March 22, 2018. A total of 13 people participated in the meeting from CNR, UNIVE, UPRC, UVSQ and HUA. Partner Organizations from Brazil (UFC, UFSC, PUC) and Canada (DAL) could participate via remote connection.
We had productive discussions and two interactive Question Answering sessions with the Project Officer and the Independent Ethical Advisor.
Indeed, particularly useful has been the clarifications of Simona to our questions on how to implement secondments and the Prof. Berendt's vision about ethical issues that might arise when dealing with mobility data.

www.master-project-h2020.eu

Facebook: Master Project H2020

Twitter: Master Project H2020

Holistic trajectories,

privacy, and the strategic concealment/ disclosure of information

A reflection on ethical questions in the mapping of taxi rides and Mediterranean migrant rescue operations, and the implications of these questions for data-science modelling

Bettina Berendt, Independent Ethics Advisor of MASTER, Department of Computer Science, KU Leuven, Belgium

As the Independent Ethics Advisor (IEA) of MASTER, I am tasked with "provid[ing] support to the consortium in reflecting on the broader ethical considerations related to holistic trajectories". The Grant Agreement (from which this task is taken) further describes that this role complements that of the Ethics Committee, whose task is to "guarantee the good governance of data and the research integrity and academic ethics" as well as to ensure "compliance with ethical and legal framework in case additional datasets will become available during the lifetime of the project". To fill the challenging role of IEA, I have worked out, together with the consortium, several concrete steps, in particular documentation and interaction provisions, which are described in an internal Deliverable. Our plan is to describe experiences in and insights from this process in publications, with the aim of sharing lessons learned with future projects. Before the (necessarily ex post) publication(s), this line of activities is likely to be of interest primarily to project members themselves.

So what could be contributions of the IEA during the project that are interesting to a wider public and therefore fit a newsletter? I believe this should be reports from inspiring discussions in and around MASTER's topics that illustrate selected ethical questions around data science and the activities of data scientists, and that highlight why and how the task of reflecting on broader ethical considerations differs from those around data governance, research integrity, academic ethics, and compliance. The present text is the first in this planned series. As often happens with ethical considerations, it presents more questions than results – the open-endedness of this text is therefore a feature and an invitation for carrying the discussion further.

Specifically, I will share some observations on data science for improving humanitarian rescue operations in the Mediterranean, a timely topic area with lots of spatiotemporal data and even more options for creating, analysing and predicting holistic trajectories.

BACKGROUND: HOLISTIC TRAJECTORIES, PRIVACY, DATA SCIENCE, AND THE MASTER APPROACH

Many holistic trajectories involve data and inferences on human mobility, and human mobility is well-known as a field that raises many legal and ethical issues. The MASTER Grant Agreement summarizes the challenges and the approach of the project: "The location of a mobile user represents very sensitive information and it has been shown … that the simple obfuscation of the location data or the user identifier do not grant privacy: an attacker could correlate the coarse location provided with background knowledge to try to identify the location with higher precision. … [With] contextually enriched trajectories, the privacy problem may become even more challenging. … [W]e plan to cope with this challenge by studying management and analysis methods for holistic trajectories that are compliant to the privacy-by-design principle."

> MASTER will certainly contribute to the emerging field of data privacy for rich spatio-temporal data

Anonymization and other methods for privacy-aware data sharing are certainly key for the goal of privacy by design, and MASTER will certainly contribute to the emerging field of data privacy for rich spatiotemporal data. Linked to this and equally undisputed is the necessity to observe data protection law and be one of the first projects to demonstrate how to effectively do data science under the European Union's General Data Protection Regulation (GDPR), which came into effect two months after the start of MASTER. (MASTER faces the additional interesting challenge that some of the consortium members operate under other jurisdictions, but this is not the topic of the present text.)

But are data privacy and GDPR compliance all there is to an ethically reflected approach to holistic trajectory analysis? What other questions surface when one considers different use cases?

I will investigate two examples of vehicle trajectory data. The first is the New York taxi rides dataset. The reception of this dataset represents a standard example of today's discussions in data science around privacy-sensitive mobility datasets, and it was, at least for me, the initial "blueprint" for identifying key questions to be asked of a research project dealing with vehicle trajectory data. The second is an analysis of AIS ship location data to study rescue operations at sea, and a brief description of the wider data ecosystem

related to migrant movements in the Mediterranean. This example illustrates that even "privacy" takes on a different meaning in this context, that legal and technological questions relevant in the first example do not appear applicable, and that instead the example points to many further questions that researchers should ask. While migrant phenomena and sea rescue operations are no doubt a particularly politically charged environment, I believe that many of these questions are relevant for other contexts of applied research too.

CASE STUDY 1: NEW YORK CITY TAX RIDES DATASET

In 2014, the City of New York released, in response to a Freedom of Information request, data about all 173 million taxi rides in New York in 2013, with the taxi identifiers pseudonymised, and exact spatiotemporal data about start- and endpoints, as well as fares, given. This dataset provided a rich real-life dataset for a wide range of data mining studies, such as "optimization of the revenue of NYC Taxi Service using Markov Decision Processes" (Li, Bhulai, & van Essen, 2017). At the same time, the publication of the dataset was soon criticized on privacy grounds. For example, the taxi pseudonyms could easily be re-identified to their actual medallion numbers (Pandurangan, 2014). It was also argued that the data allowed inferences towards sensitive attributes of the taxi drivers, such as the patterns of breaks during the day indicating that someone is a devout Muslim (uluman, 2015). Finally, with some background knowledge, inferences can be made towards the identity of taxi customers, and based on that, details about their whereabouts learned (Atockar, 2014). The futility even of better pseudonymisation/anonymisation approaches was demonstrated by Douriez et al. (2016). Medallion and driver license IDs were removed from NYC's taxi datasets released in subsequent years.
(https://data.cityofnewyork.us/browse?q=taxi)

The taxi rides represent a typical case of personal data in the sense of the GDPR. Personal data are "any information relating to an identified or identifiable natural person ('data subject')" (Article 4). In the sense that at least some, and likely many, taxi drivers and taxi customers are easily identifiable, the dataset contains personal data. Taxi customers (and conceivably also taxi drivers) had not been asked to give their consent to these data being published online for unspecified purposes, nor are other grounds for such processing (Article 6 GDPR) present. This is textbook privacy violation by data (more accurately in the EU context: a violation of data protection law), and therefore, data protection / ethics boards in EU universities strongly discourage the use of this dataset for any kind of data mining. (This is anecdotal evidence, gathered not at the universities of the MASTER partners.)

This perception of a dataset assumes that the population of data subjects consists of informed individuals, who exercise their autonomy among other things by travelling in vehicle passages they pay for, and who have a reasonable expectation of privacy in doing so that requires that the data about their movements remain confidential. The main question for the responsible data scientist appears to follow from the observation that the removal of taxi identifiers "would adversely impact certain types of analysis on the data" (Douriez et al., 2016, p. 148) and the need to find different analysis types.

CASE STUDY 2: AIS DATA FOR DESCRIBING MIGRANT RESCUE OPERATIONS

The second case study is based on a paper published in a report by the IOM (Hoffmann et al., 2017), the UN International Organization for Migration, and illustrated in an interactive and multimedia online presentation (see: http://rescuesignatures.unglobalpulse.net/mediterranean/). As in case study 1, the base data are in principle publicly accessible. They are data from the Automatic Information System (AIS), a maritime communications system through which vessels regularly broadcast information, including their identifier, vessel type, latitude and longitude, speed, course and destination. The information is used by maritime authorities and ships to locate nearby vessels and avoid collisions. Based on these spatiotemporal data and enriched with textual and pictorial data from other sources, the authors generate a type of holistic trajectories, manually label them as representing (or not) a rescue operation, and use clustering and machine learning with a view to classification and prediction (The enriching data in this case include broadcast warning data produced by WWNWS, a global service managed by the UN Maritime Organization IMO, and the tweets issued by NGO vessels).

Other researchers, including from the MASTER project, have investigated how to model and detect such trajectories. Based on AIS data, complex events, including but not limited to SAR (search and rescue) missions, and involving one or several vessels, can be modelled and detected efficiently and in real time using combinations of exploratory, machine learning, and logics-based (event calculus) techniques (Patroumpas et al., 2017; Varlamis, Tserpes, & Sardianos, 2018).

Hoffmann et al. mention several limitations of their method, mainly with regard to data quality, including the fact that as circumstances change, so do the data and patterns (thus, the analysis of timely data is crucial).

Under the heading of "privacy", they raise several points. The first is a reference to concerns over port security as a consequence of AIS data being publicly available. The second is the possibility that rescue organisations may not want the full details of their operations to be publicly known, because they are facing opposition and threats (a European far-right group threatening to attack rescue vessels is mentioned). Both concerns are not privacy concerns in the sense of European law (in particular because the agent requesting the confidentiality is not a natural person). As a third reason, the authors mention that "adversarial users could take advantage of the data to track the location of individual refugees [identified by record linkage with data such as photos or statements, or other background knowledge], attack rescue boats or guide piracy operations" (p. 40). Presumably, the attacks and piracy operations are security/safety concerns of the rescue vessels, and these concerns could arise from the public availability of the data as well as from

possible predictors learned from them, i.e. the data scientists' work.

It also appears, from the sentence, that the possible tracking of individuals is considered a security/safety risk (because it could lead to attacks) rather than a typical privacy risk (by which an individual migrant would want to keep their identity or properties hidden). It is difficult to say what role such expectations of, or wishes for, privacy in our usual sense, play in this extreme situation. Also, it has been observed increasingly over the past years that rather than trying to hide their voyage, "migrants from Libya facilitated their traceability by national authorities and monitoring systems, anticipating in space and time border patrols by sending an SOS as soon as they entered the international waters" (Tazzioli, 2016, p. 576). In other words, along their journey, migrants deal strategically with visibility and invisibility, with information disclosure and hiding/confidentiality. This is quite probably a very rational strategy given the fact that a successful and invisible journey to Europe is by now nearly impossible for many reasons, including that traffickers severely overload and under-equip their vessels, and that due to the high-resolution sensors employed in the European Border Surveillance System EUROSUR (Deibler, 2015), even very small vessels are likely to be spotted and monitored. Strategic information disclosure (in addition to strategic information hiding) by individuals can also be observed in many other contexts that are less dramatic than the life-or-death situations faced by migrants on the Mediterranean, and it has been pointed out that strategic information disclosures too can be privacy-related behaviour (Gürses & Berendt, 2010).

The Alexander Maersk trajectory rescuing refugees in the Mediterranean Sea

A second question related to privacy is related to the referent of the data. Technically a ships' trajectory could be considered personal data in the same sense as a taxi's trajectory. (This concerns the ships provided by the traffickers as well as rescuing ships once they have been boarded by migrants.) However, the chances appear slim that individuals would be identified like in the dedicated attacks in the taxi example, and it is unclear whether they would even be identifiable. In this sense, then, AIS-based trajectories of rescuing ships may not count as personal data. In fact, as has been argued in this context (Tazzioli, 2016) as well as in connection with other applications of big data analyses to humanitarian causes (Taylor, 2017), there is a temptation to focus on migrants as a *group* defined only by one feature (here: to be in need of rescue).

The fact that big data constitute new risks in the profiling of groups has been lamented often in connection with data protection laws such as the GDPR (which focus on the protection of *individuals'* rights and freedoms); in the humanitarian realm, it creates additional and different challenges (Taylor, van der Sloot, & Floridi, 2017).

For the data scientist, this means that also the response to these risks and threats may need to be very different, because traditional approaches to (for example) anonymisation are focussed on the protection of individuals from threats against persons as individuals. It is an open research question what could constitute effective measures of group protection.

Data privacy, viewed technically, does not need to make a clear distinction between protecting information and control over it related to individuals (a concept rooted in human rights) and protecting information and control over it related to other entities (such as organisations, the NGOs in the current example and in the argument made by Hoffmann et al., a concept rooted in IT security). In data privacy, a different and independent dimension becomes relevant when one asks "whose privacy" should be protected. A useful distinction is that between data owner, data respondent (the data subject, although not always in its legal sense of an individual person), and data user; and this distinction has implications for the choice of data-privacy protection methods (Domingo-Ferrer, 2007). In the present example, one assignment of these roles that follows the argument about risks above could be: the NGO as the data owner, the migrant (or migrant group) as the data respondent, and various (potential) data users: the public, politicians, pirates, ...

Moving beyond privacy and data privacy, many other questions, technical as well as ethical, arise about information disclosure and hiding. The study and visualisation of "rescue patterns" can have different objectives. Hoffmann et al. mention operational objectives (e.g., supporting coordination of rescue operations), analytic objectives (e.g., deter-

mining conditions under which rescues are most effective), and reporting objectives. The latter are described as follows: "supplement the large amount of qualitative, descriptive coverage already produced by NGOs and the news media", "help external observers … obtain a high-level picture of what is happening in the region over time. An overview of these patterns is critical for coordination and advocacy purposes; it enables stakeholders to see the true magnitude of rescue operations, and to quantify costs, shortcomings and future needs." (Hoffmann et al., 2017, p. 30)

Concentrating on the reporting objective, it can be argued that rescue patterns constitute a counter-mapping practice: in the EUROSUR monitoring system, selected migratory events are produced from the sensed data and mapped in time and space (Tazzioli, 2016). The website watchthemed.net, initiated and run by a network of NGOs, activists and researchers, maps events to monitor deaths and violations of migrants' rights. In the SAR-centric applications described here, rescue events are produced and mapped. EUROSUR is run by Frontex, and its data and analytics are not available to NGOs and other external partners, whereas the rescue patterns are mined from data available publicly (AIS data) or to partners of the research (the broadcast warnings), and enriched with further aspects from public data (such as tweets).

Mapping practices generate a narrative around their real-world phenomenon. The current data models and visualizations of rescue patterns, maybe for technical reasons (because the EUROSUR data are not available), maybe to avoid visual clutter, display these patterns in an otherwise "empty" space. Is it possible, and is it advisable, to at least represent that far more data exist (even if one does not have access to them)? In other words, should the "known unknown" data be modelled and represented too, and if so, how? These data are important for technical reasons as much as for narrative reasons – how can and should these two motivations be addressed, and how can the choices made be made in a transparent and accountable way?

In the following paragraphs, I will illustrate three examples of these considerations.

First, sometimes trajectory data illustrate very directly the influences of context and the uncertainty and the "unknowns" of vessel operators. As an example, consider the recent case of a commercial cargo ship that took on 113 people saved by an NGO rescue ship and then spent four days in a political standoff on a zig-zag trajectory between ports before being allowed to dock in Sicily (Al Jazeera, 2018; Borghese et al., 2018) (I thank Konstantinos Tserpes for mentioning this example and making available the visualization of the trajectory). Can and should holistic trajectories measure and visualize the enormous costs caused by such decisions, as well as the incentives and influences this may have on further behaviour by vessel operators?

> Although data can help citizens demand accountability, ultimately, the inferences that can be drawn from the data are only as valuable as the actions they induce.

Second, many of the existing, but not accessible data have strong effects on the rescue events modelled. For example, the Libyan coastguard now has indirect access to EUROSUR data (Monroy, 2018); thus, their rescue actions, including those in cooperation or competition with European actors, may be planned based on data that are not modelled in the rescue patterns system, and which therefore can co-determine the "coordination" and "effectiveness" of a rescue. Can and should these data (or at least the fact of their existence and possible influence) be modelled?

Third, further questions concern which aspects are important to judge the legal and ethical dimensions of a rescue operation (cf. Medina, 2018, resp. Cancellato, 2018): in a recent case in which a commercial towboat under Italian flag rescued migrants and then handed them over to the Libyan coastguard, a key legal question revolves not around the spatiotemporal data of the rescue operation, but on whether it was instructed by the Italian or the Libyan authorities (Ziniti, 2018). Can these aspects be modelled as part of multi-aspect trajectories, and how could this be done if the datum itself is still being contested?

In sum, an ethical consideration of the modelling and reporting of vehicle data and patterns, even if restricted to what data are to be included and how, what information is to be kept confidential or disclosed, can reach far beyond the traditional questions discussed under data protection and data privacy. It requires a critical examination of the sociopolitical background of the mobility that these vehicles afford, support, or impede, and of the goals of the data-science project undertaken. And although "data can help citizens demand accountability", "ultimately, the inferences that can be drawn from the data are only as valuable as the actions they induce. … this problem will not be solved with data alone" (Hoffmann et al., 2017, p. 42). This added complexity may be what makes data science projects "for social good" so fascinating and worthwhile. I look forward to feedback on this essay and to future work in the area!

REFERENCES

Al Jazeera News (2018). Danish cargo ship carrying refugees allowed to dock in Italy. 26 June 2018. https://www.aljazeera.com/news/2018/06/danish-cargo-ship-carrying-refugees-allowed-dock-italy-180626081632471.html

Atockar (2014). Riding with the Stars: Passenger Privacy in the NYC Taxicab Dataset. https://research.neustar.biz/2014/09/15/riding-with-the-stars-passenger-privacy-in-the-nyc-taxicab-dataset/

Borghese, L., Vandoorne, S., & Vonberg, J. (2018). Migrant rescue ship Lifeline to dock in Malta after being stranded for five days in the Mediterranean. CNN News, 26 June 2018. https://edition.cnn.com/2018/06/26/europe/migrant-ships-maersk-lifeline-intl/index.html

Cancellato, F. (2018). Poche palle, i migranti della Asso 28 li abbiamo respinti noi: e la Libia è solo la foglia di fico della nostra ipocrisia. Linkiesta, 1 August 2018. https://www.linkiesta.it/it/article/2018/08/01/poche-palle-i-migranti-della-asso-28-li-abbiamo-respinti-noi-e-la-libi/39019/

Deibler, D. (2015). EUROSUR – A Sci-fi Border Zone Patrolled by Drones? In J. Camenisch, S.

Fischer-Hübner, & M. Hansen (Eds.), Privacy and Identity Management for the Future Internet in the Age of Globalisation (pp. 87-109). Berlin etc.: Springer. Retrieved from https://hal.inria.fr/hal-01431600/document

Domingo-Ferrer, J. (2007). A three-dimensional conceptual framework for database privacy. In Proc. SDM 2007 (pp. 193–202). Berlin etc.: Springer, LNCS 4721. Retrieved from https://crises-deim.urv.cat/webCrises/publications/lncs/lncs4721.pdf

Douriez, M., Doraiswamy, H., Freire, J., & Silva, C.T. (2016). Anonymizing NYC Taxi Data: Does It Matter? In Proc. 2016 IEEE International Conference on Data Science and Advanced Analytics (DSAA), Montreal, QC, 2016 (pp. 140-148). Retrieved from https://vgc.poly.edu/~juliana/pub/taxi-privacy.pdf

Gürses, S.F. & Berendt, B. (2010). The Social Web and Privacy. In E. Ferrari & F. Bonchi (Eds.), Privacy-Aware Knowledge Discovery: Novel Applications and New Techniques. Boca Raton, FL: Chapman & Hall/CRC Press, Data Mining and Knowledge Discovery Series. Retrieved from https://www.esat.kuleuven.be/cosic/publications/article-1304.pdf

Hoffmann, K., Boy, J., Leon-Dufour, J., Breen, D. Earney, C. & Luengo-Oroz, M. (2017). Using big data to study rescue patterns in the Mediterranean. In: Fatal Journeys. Volume 3, Part 1: Improving Data on Missing Migrants (pp. 24-46). Geneva, Switzerland: International Organization for Migration. https://publications.iom.int/system/files/pdf/fatal_journeys_volume_3_part_1.pdf

Li, J.P.K., Bhulai, S., & van Essen, T. (2017). Optimization of the revenue of the New York City Taxi Service using Markov Decision Processes. In Proc. of DATA ANALYTICS 2017: The Sixth International Conference on Data Analytics. Retrieved from https://www.thinkmind.org/download.php?articleid=data_analytics_2017_4_10_68005

Medina, J. (2018). U.N. says migrants' return to Libya by Italian boat could be illegal. Reuters, 31 July 2018. https://www.reuters.com/article/us-europe-migrants-libya/migrants-return-to-libya-by-italian-boat-could-breach-international-law-u-n-idUSKBN1KL1K4

Monroy, M. (2018). Durch die Hintertür: Anschluss Libyens an europäische Überwachungssysteme. CILIP Blog, 19 January 2018. https://www.cilip.de/2018/01/19/durch-die-hintertuer-anschluss-libyens-an-europaeische-ueberwachungssysteme/

Pandurangan, V. (2014). On Taxis and Rainbows Lessons from NYC's improperly anonymized taxi logs. https://tech.vijayp.ca/of-taxis-and-rainbows-f6bc289679a1

Patroumpas, K., Alevizos, E., Artikis, A., Vodas, M., Pelekis, N., & Theodoridis, Y. (2017). Online event recognition from moving vessel trajectories. GeoInformatica, 21(2), 389-427. https://link.springer.com/article/10.1007/s10707-016-0266-x

Taylor, L. (2017). Safety in Numbers? Group Privacy and Big Data Analytics in the Developing World. In L. Taylor, B. van der Sloot, & L. Floridi (Eds.), Group Privacy: the Challenges of New Data Technologies. Berlin etc.: Springer. Available at SSRN: https://ssrn.com/abstract=2848825

Taylor, L., van der Sloot, B., & Floridi, L. (Eds.) (2017). Group Privacy: the Challenges of New Data Technologies. Berlin etc.: Springer. https://www.springer.com/de/book/9783319466064

Tazzioli, M. (2016).Eurosur, humanitarian visibility, and (nearly) real-time mapping in the Mediterranean. Acme, 15 (3), 561-579. https://acme-journal.org/index.php/acme/article/view/1223/1201

uluman (2015). Identifying Muslim cabbies from trip data and prayer times. https://www.reddit.com/r/dataisbeautiful/comments/2t201h/identifying_muslim_cabbies_from_trip_data_and/

Varlamis, I., Tserpes, K., & Sardianos, C. (2018). "Detecting Search and Rescue Missions from AIS Data," 2018 IEEE 34th International Conference on Data Engineering Workshops (ICDEW), Paris, 2018 (pp. 60-65). http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=8402020&isnumber=8402003

Ziniti, A. (2018). Migranti, inchiesta sul comportamento del rimorchiatore italiano Asso 28. La Repubblica, 8 August 2018. http://www.repubblica.it/cronaca/2018/08/08/news/migranti_presentato_un_esposto_sul_comportamento_del_rimorchiatore_italiano_asso_28-203654671/

This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Sklodowska-Curie grant No 777695
Duration: March 1 2018 | February 28 2022

Privacy statement
The personal data (names, email addresses...) and the other information entered in MASTER Newsletter will be treated according with the provision set out in Legislative Degree 196/2003 (known as Privacy Code) and subsequently integration and amendment.

Coordinator of the project: Chiara Renso | chiara.renso@isti.cnr.it

MASTER Newsletter is not for sale but it is distributed for purposes of study and research and published online at http://www.master-project-h2020.eu