

Applicable European Regulations for Data-driven Policy-making

Sonia Bilbao

TECNALIA, Basque Research and
Technology Alliance (BRTA),
P. Tecnológico Bizkaia, Ed. 700, 48160
Derio, Spain
Sonia.bilbao@tecnalia.com

Maria José López

TECNALIA, Basque Research and
Technology Alliance (BRTA),
P. Tecnológico Bizkaia, Ed. 700, 48160
Derio, Spain
MariaJose.Lopez@tecnalia.com

Sergio Campos

TECNALIA, Basque Research and
Technology Alliance (BRTA),
P. Tecnológico Bizkaia, Ed. 700, 48160
Derio, Spain
sergio.campos@tecnalia.com

ABSTRACT

Data-driven policy making aims to make optimal use of data and collaborate with citizens to co-create urban policies and, in general, to conduct a more reliable decision-making process. The European Commission considers data an essential resource for economic growth, competitiveness, innovation, and disposes as part of its strategy a set of regulations and guides, aiming to that more data becomes available for use while keeping the rights and trustworthy of the companies and individuals who generate and consume the data during the whole lifecycle. These regulations impact and open new challenges and opportunities when addressing the decision-making process in URBAN Transformation and specifically urban mobility as in the case of the URBANITE project.

KEYWORDS

Data, regulations, ethics, trustworthy, privacy, governance

1 INTRODUCTION

Urban mobility faces greater uncertainty and complexity in the long term generated by two main factors: the demand for growth in urban environments, the pressure for a more sustainable model of mobility in the face of the emergence of global warming. In general, we find that the social conscience is changing in favor of more equitable and sustainable ways, and the recovery of the space of the city for the people. On the other hand, the accelerated technological development in the transport modes and business models themselves, including innovations such as autonomous driving, micro-mobility, connected vehicles, electro-mobility, mobility as a service (MaaS), new vehicle ownership models, etc. mark specific challenges. in your deployment. These trends are changing the landscape of urban planning and mobility management in cities, incorporating new challenges. All of these require new advances in mobility planning processes and methods, with the aim of helping public administrations and policy makers to better understand this new context, supporting them in decision-making and policy definition. Policies should be discussed among the main actors in the new urban mobility scenario: citizens, service providers, public servants and political leaders.

This scenario can be built on two pillars: 1) co-creation sessions and 2) empirical analysis on stakeholder trust, attitude, impact, benefits and risks in the use of disruptive technologies. Now, traditional technological solutions are no longer valid in this situation, and disruptive technologies such as big data analysis or artificial intelligence emerge as a promising support to those responsible for formulating new policies. Data-driven policy making aims to make optimal use of existing heterogenous data and collaborate with citizens to co-create policy. This opportunity entails specific challenges to favor the acceptance by users of the results obtained through the application of these technologies and, first, to collect the relevant data from the different local stakeholders. These are some of the objectives of the URBANITE project, to face challenges, attitudes, confidence and opportunities in the use of disruptive technologies in public services in the context of urban mobility.

URBANITE identifies several key results: a Social Policy Lab – an environment to promote digital co-creation and methodologies and methods to support the co-design and co-creation for policies, a Data Management Platform – To provide automatic support to the whole data processing chain and its life cycle, starting with the collection process up to its use and a Decision-Making Support System – Powerful tools which combine multiple data sources with advanced algorithms, simulation, recommendation, and visualisation.

The project identifies different stages from the perspective of data and more specifically, its availability, openness and privacy:

- 1st Stage- Setup of participation labs and initial gathering:
 - Open data currently available
 - including identification and recruitment of participants, the preparation of an informed consent procedure to implement for individual participation
 - The register and use of the virtual participation platform as a complement of previous sessions
- 2nd Stage. The potential use of existing non-open data, personal and non-personal on the cities to the objectives of the project.
- 3rd Stage. The transfer of collected data from 3rd parties, defining a transfer agreement among both parties (company and city use case)

2 RELATED EUROPEAN REGULATIONS

On the other hand, the European Commission considers data an essential resource for economic growth, competitiveness, innovation, defining an European strategy for data aiming to

ensure Europe's global competitiveness, a data sovereignty, that more data becomes available for use, while keeping the rights of the companies and individuals who generate and consume the data during the whole lifecycle. As part of this support, the Commission has proposed some material as part of its data strategy, disposing several normative and guides to conduct a success and trustworthy data management.

2.1 Data Governance Act

This initiative refers to the management of personal as well as non-personal data, therefore being linked at the legislative level with the General Data Protection Regulation (GDPR)[1] and the Directive on privacy and electronic communications [2]. The European Commission has implemented a solid and trustworthy legal framework for the protection of data, in order to promote a single data market, for which it must guarantee that data from the public sector, companies and citizens can be available and used in the most efficient and responsible way possible, while companies and The Data Governance Act [3] is the first of a set of measures announced in the 2020 European data strategy, aims to promote the availability of data for its use, increasing trust between the parties and strengthening data collection mechanisms throughout the European Union. The DGA will also support the establishment and development of common European data spaces in strategic domains, involving both public and private actors.

The framework addresses the following scenarios:

- The transfer of public sector data for reuse, in cases where such data is subject to the rights of third parties. It establishes a mechanism for the reuse of certain categories of protected data from the public sector, which is subject to respect for the rights of third parties.
- The transfer of personal data with the help of intermediaries, whose work will consist of helping providers to exercise the rights conferred by the General Data Protection Regulation (GDPR). The objective is to strengthen trust in the exchange of personal and non-personal data, and reduce the costs of transactions linked to the exchange of data between providers and their consumers, with neutral facilitators,
- The transfer of data for altruistic purposes (making data available to the common good, on a voluntary basis, by individuals or companies). Establish a registration and consent in order to reduce costs and facilitate data portability.
- The exchange of data between companies in exchange for some type of remuneration.

2.2 ETHICS GUIDELINES FOR TRUSTWORTHY AI

Despite the fact that AI technologies are mature enough, their adoption by companies is very uneven, and in general, much lower than one would expect. There are obstacles that hinder the widespread extension of AI technologies, both cultural and technical. AI technologies will not spread massively until the scientific community is able to develop reliable technology from the user and from the different data providers. On the other hand, the use of these technologies involves risks that must be managed

appropriately. To ensure that we are on the right track, it is necessary to abide by a human-centered approach to AI, without losing the goal of improving human well-being. The concept of trusted AI addresses reliance on technology as a first step. The new guidelines are aimed at all parties involved who develop, apply or use AI, encompassing companies, organizations, researchers, public services, institutions, individuals or other entities.

According to the regulations “Reliable AI has two components: 1) it must respect fundamental rights, current laws and essential principles and values, so as to guarantee an« ethical purpose », and 2) it must be reliable and technically sound , since a little technological mastery can cause involuntary damages, although the intentions are good ”.

Therefore, these Guidelines establish the framework of a reliable AI, guiding in three levels of abstraction, from the most abstract to the most concrete of aspects to be evaluated:

- Guarantee of the ethical purpose of AI, establishing fundamental rights, as well as the essential principles and values, which it must comply with.
- A series of guidelines, addressing both ethical purpose and technical soundness, listing the requirements for reliable AI, and providing a summary of technical and non-technical methods that can be used for their application.
- A concrete, but not exhaustive, list of aspects that must be evaluated in order to achieve reliable AI.

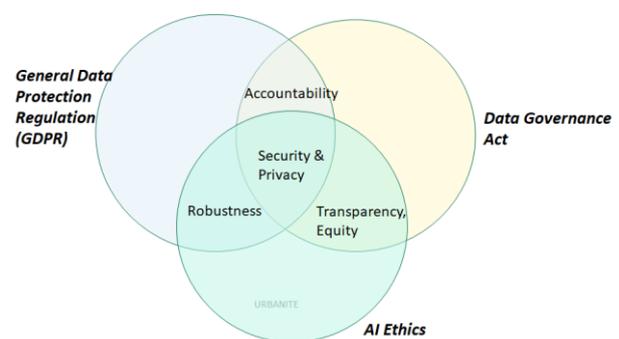


Figure 1: Relation and overlapping among regulations

According to these principles, the work package in charge of the definition, design and adaption of artificial intelligence and data analytics is ensuring that the applied methods meet the seven key requirements for Trustworthy AI:

- 1) human agency and oversight. URBANITE proposes a decision support methodology and supporting tool for policy creation, that combine and carefully balance different methods:
- 2) harvested historical data, GIS, expert knowledge, outputs of decision models, and others. The last word will remain in the hands of municipal experts, the platform being a tool to facilitate their decision. On the other hand, one of the pillars of the project is the implementation of a thoughtful space of discussion among the main actors of the new urban mobility scenario: citizens, service providers, public servants and policy makers.
- 3) technical robustness and safety. The work focused on the algorithms and simulations to be deployed, defines for as

objective metrics a pair of KPIs, refereed to the precision of predictions and the quality of recommended policies or procedures.

- 4) privacy and data governance, a fundamental right particularly affected by AI systems. Prevention of harm to privacy also necessitates adequate data governance that covers the quality and integrity of the data used. All the algorithms will work on the gathered data, just (pseudo) anonymised. If applied during the research stage on the algorithms, all the GRPR measures will be analysed and adopted to use personal data.
- 5) Transparency [7], the principle of explicability and encompasses transparency of elements relevant to an AI system: the data, the system and the business models: This implies traceability of decision along the whole cycle of data, from datasets, gathering, labelling and process, Explainability concerns will be considered for the methods applied, ensuring a better understanding of the underlying processes and related human decisions (e.g. xAI approach). In any case, the simulation and rules-based reasoning approaches, are well sited from the explainability point of view.
- 6) diversity, non-discrimination and fairness. URBANITE gather existing open data portals, geographical information systems (GIS), data coming from private data providers, the basis is any, comes in origin. SoPoLab sessions and use case evaluation support the feedback from stakeholders who may directly or indirectly be affected throughout its life cycle.
- 7) dedicated assessment of the algorithms during their design and use case deployment ensures the auditability.
- 8) environmental and societal well-being. In the last term, URBANITE provides a new decision-support system for Planning Sustainable Mobility and the early evaluation of urban policies. A Sustainable Urban Mobility Plans (SUMP), defines strategic plans based upon a long-term vision of transport and mobility, guaranteeing technical, economic, environmental and social sustainability.

3 ALGORITHMS ACTIONABILITY

Taking into account the previous regulations and based on previous experience in the context of Intelligent Transportation Systems [4], it is confirmed that aspects such as trust, precision and reliability, among other non-functional properties, are essential for predictive and analytical techniques to be practices in its use. We present the term Actionability, as the characteristic that any system based on data analysis or artificial intelligence must present to be implemented and used successfully in a real operating environment. This concept, in turn, identifies a series of desirable characteristics, which in URBANITE are contextualized in the field of urban mobility planning.

Data-based models are usually subject to uncertainty, involving non-deterministic stochastic processes, both in the learning, execution or training mechanisms / input data, and also present in the results. Once deployed, it is essential to provide an objective measure of the reliability and precision of the results, winning in terms of Trust. The need to explain and render the underlying analytical models interpretable is undoubtedly one of the research fields with the greatest impact, being considered under the concept of Explainable Artificial Intelligence [5][6] (xAI). This field of study comprises different techniques and methods, taking into account three fundamental factors: the

nature of the model to be explained, from intrinsically transparent to completely opaque and unintelligible; the user of the algorithms; and finally, the way in which said explanation must be prepared and presented to the decision maker, which will depend on their degree of knowledge, as well as the intrinsic possibilities offered by the model to be explained in one way or another.

Adaptation is the reaction of a system, model or process to new circumstances, with the idea of maintaining its performance or reducing its loss, compared to the ideal conditions that were taken into account in its design and initial adjustment. The main problem in scenarios whose underlying phenomena change over time, without being addressed by the model itself, is that the conclusions, predictions or categorizations will not be reliable. This phenomenon is called concept drift [8][9].

Robustness refers to the ability of a system to maintain service when external incidents occur. In the case of urban planning, it will not be so critical, since the decisions to be made will not be made in real time; However, the data ingestion of the different data sets and available stores, if it must be operational, to minimize the loss of input data, in addition to being robust data algorithms in such loss situations, fluctuations in the frequency of the themselves, poor quality data, etc. In the URBANITE project, data quality is explicitly addressed through the implemented components associated with data preservation.

Stability means ensuring that there are no surprises for the user in terms of functionality. In general, the algorithms are worked in a specific geographic area and according to the available data sets. However, for their deployment in a real environment, it is necessary to project them to larger areas and volumes of data. This issue must be taken into account from the design stage of the algorithms, to optimize their algorithmic complexity, which represents the amount of resources (temporal, execution time and space, required memory) that an algorithm needs to solve a problem. This characteristic allows to determine the efficiency of this algorithm, not in terms of absolute measures but measures relative to the size of the problem. Currently, the availability of new technologies and paradigms of parallel and distributed processing of massive volumes of data, allows an escalation of the methods, obtaining adequate response times. However, its exploitation requires the adequate implementation and adaptation of the algorithms according to the architecture in which they will be deployed, as well as optimizing this deployment of analytical workloads in the different layers.

Another key feature is the compliance of the new methods with transportation engineering. Existing traffic and mobility engineering practices are well established, with a powerful knowledge base. A better understanding of the hybridization of data analysis and simulation methods, data-driven and model-driven approaches, by combining the strengths of each side, will help us improve the models by identifying more complex underlying assumptions. In general, the results are closely linked to the experiments carried out; transferability is a desirable characteristic for algorithms and any model, in order to present adequate performance and functionality in other contexts and starting data, different from those used in learning.

Finally, we cannot forget the contextual aspects identified by the EU for the definition of sustainable mobility policies, measures and solutions, as part of SUMP and any support tool, with the aim of contributing to urban regeneration, transport sustainability, social inclusion and social empowerment through active participation.

4 OPPORTUNITIES AND NEXT STEPS

Additionally, to the actionability requirements for our methods and algorithms, the new regulation and especially the Data Governance Act presents a set of topics or opportunities to explore from the different action lines. The following table presents some of them, according to the type of data to explore on the project: public, personal or altruistic data

Table 1: Challenges and Research Opportunities

	Public sector data for reuse	Reuse of personal data	Altruistic purpose	Data Sharing among business
SoPoLab	Presentation of Data Governance Act among local stakeholders . Engagement of city departments, public transport service providers, collectives and final users on sharing data. Identification of potential local "Data Sharing Providers" .			
Data Management Platform	Support of Data Governance functionalities by the Urbanite Platform: - Data exchange policies . Establishment of policies regarding the exchange of data and events across the boundaries of an organization. - Data ownership . Opportunities or barriers to data and events between companies: IPRs, rights of access and exploitation, sensitivity, etc. - Security and privacy . Access and authentication, integrity, privacy, etc.			
Decision support tools	Explainable IA-techniques . Other algorithms avoiding the transfer of raw data between origin and the platform			

5 CONCLUSIONS

During the first period of the project, some relevant algorithms and data analysis have been identified based on the project's use cases. Having analyzed the different regulations around the

exploitation, data management and artificial intelligence technologies, machine learning and advanced processing, the requirements that these new algorithms must present in the future have been identified. Finally, it introduces the concept of Actionability as a key property of any data-based modeling and treatment process to generate knowledge of practical value for decisioning. All these aspects open challenges and, also opportunities for URBANITE project.

ACKNOWLEDGMENTS

These results have been achieved during funding project from the European Union's Horizon 2020 research and innovation programme under grant agreement #870338.

REFERENCES

- [1] [DO L 119 de 4.5.2016](#), p. 1.
- [2] [DO L 201 de 31.7.2002](#), p. 37.
- [3] <https://digital-strategy.ec.europa.eu/en/library/proposal-regulation-european-data-governance-data-governance-act>
- [4] Ibai Lana, Vlahogianni Elenni, Javier Del Ser, 2020, "From Data to Actions in Intelligent Transportation Systems: a Prescription of Functional Requirements for Model Actionability," arXiv, 2020.
- [5] Alain Barredo, N. Díaz-Rodríguez, Javier Del Ser, A. Bennetot, S. Tabik, A. Barbado, S. García, S. Gil-López, D. Molina, R. Benjamins, R. Chatila, Francisco Herrera, "Explainable Artificial Intelligence (XAI): Concepts, Taxonomies, Opportunities and Challenges toward Responsible AI. arXiv:1910.10045," 2020.
- [6] D. Doran, S. Schulz and T. R. Besold, 2017, "What does explainable AI really mean? a new conceptualization of perspectives. arXiv:1710.00794," 2017.
- [7] A. Datta, S. Sen and Y. Zick, 2014, "Algorithmic transparency via quantitative input influence: Theory and experiments with learning system," in 2016 IEEE symposium on security and privacy (SP), 598–617, 2016.
- [8] J. Gama, I. Zliobaite, A. Bifet, M. Pechenizkiy and A. Bouchachia, 2014, "A survey on concept drift adaptation," ACM computing surveys (CSUR), vol. 46, no. 4, p. 44, 2014.
- [9] Jesús Lobo, Javier Del Ser, Nekane Bilbao Cristina Perfecto, 2018, "DRED: An evolutionary diversity generation method for concept drift adaptation in online learning environments," Applied Soft