MDPI

A European Approach to the Establishment of Data Spaces

Marco Minghini [1,†,‡] , Alexander Kotsev [1,*,†,‡] and Carlos Granell [2,†]

1 European Commission, Joint Research Centre (JRC), 21027 Ispra, Italy
2 Institute of New Imaging Technologies, Universitat Jaume I of Castellon, 12071 Castellon, Spain
* Correspondence: alexander.kotsev@ec.europa.eu
† These authors contributed equally to this work.
‡ The views expressed are purely those of the authors and may not in any circumstances be regarded as stating an official position of the European Commission.

1. Introduction

Within a context defined by the rapid increase in the availability of data, combined with the complexity of data sources, infrastructures, technologies and actors involved in data sharing flows, the European Union (EU) is devising approaches that can reap the benefits of data-driven innovation. The policy vision defined in the European strategy for data [1] back in 2020 aims to make the EU a leader in today's data-driven society through an improved use of data across actors and sectors with the ultimate goal of making better decisions in business and the public sector. This is to be achieved in practice by using a set of interdependent legal instruments (see Figure 1), most notably the following: (i) the Data Governance Act [2], creating processes and structures to facilitate voluntary data sharing by companies, individuals and the public sector; (ii) the Data Act [3], establishing rules and conditions for a fairer access to and reuse of data from industry, including Internet of Things (IoT); and (iii) the upcoming Act on high-value datasets, complementing the Open Data Directive [4] with a list of technical requirements and datasets that public sector bodies are required to publish in machine-readable format, for free, and under open licenses. In addition, the European strategy for data foresees the establishment of a common European data space by combining sector-specific data spaces in domains such as agriculture, mobility, finance and environment. This data space will act as a European single market for data and create value by incentivising digital innovation at scale.

Technical frameworks and specific approaches that are based on Europe's shared societal values should complement the legal provisions. As those are still in the making, it is important that scientific evidence relating to data sharing from an organisational and technical and also from a legal perspective informs different stages in the implementation of the European strategy for data [5,6]. Accordingly, this Special Issue, titled "A European Approach to the Establishment of Data Spaces", aimed to welcome multidisciplinary and multi-domain submissions that would contribute to shaping the research agenda and providing best practices for data-driven innovation that are in line with European values and the already presented legal initiatives. The contributions to the Special Issue fall into two main groups. First, a set of *research articles* covers empirical aspects and methodological practices for improving data-driven innovation. Second, *data description papers* emphasise the need of having relevant data sets along with complementary data collection methods to shape and enrich the upcoming European data space. All articles included in the Special Issue inform the implementation of the European strategy for data [1,7] and can not only play a role in defining blueprints, addressing specific functional and non-functional requirements for European data spaces but also provide fertile ground for relevant research in the years to come.
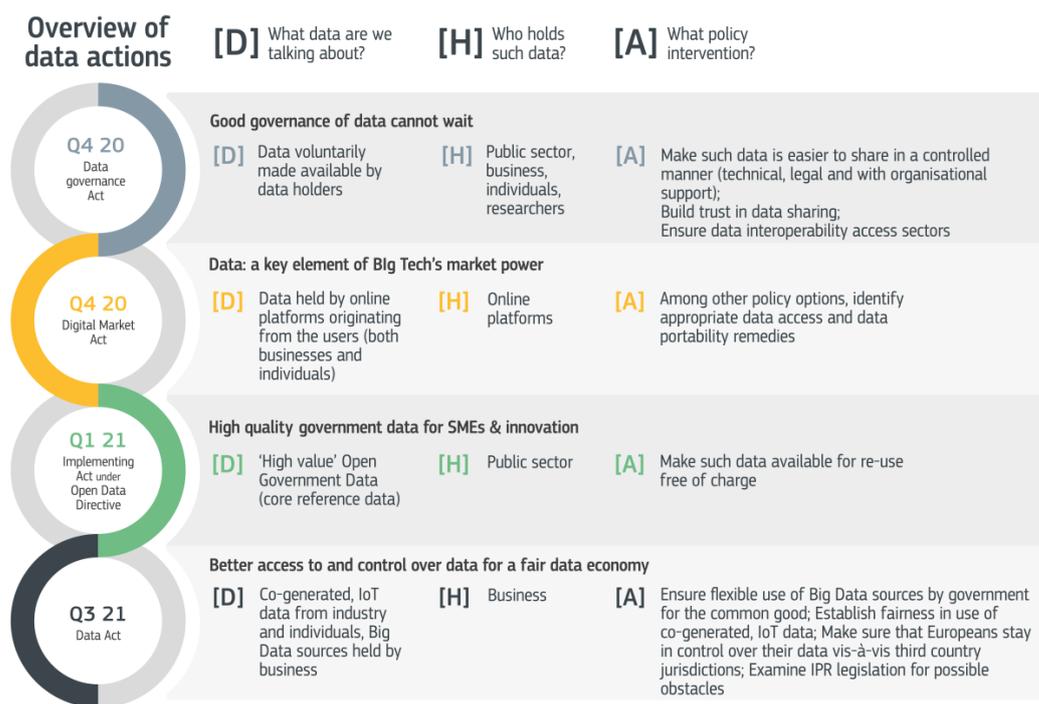
Figure 1. Legislative initiatives defined within the context of the European strategy for data. Source: European Commission, 2020.

2. Contributions to the Special Issue

2.1. Research Articles

In the first paper published in the Special Issue, a group of Italian researchers from multiple institutions (Marco Minghini, Alessandro Sarretta, and Maurizio Napolitano) analysed the impact of OpenStreetMap (OSM)—a popular geospatial crowdsourcing project—to help address the first wave of the COVID-19 pandemic in Italy [8]. Several activities initiated by the Italian OSM community, such as mapping the areas that were first hit by the pandemic, updating information on pharmacies and adding details about commercial activities offering delivery services, demonstrated OSM's huge potential as an emerging source of information in the vast European data landscape. The study assesses the OSM project from a data ecosystem perspective in light of the European strategy for data, building on the Italian COVID-19 experiences to reflect on opportunities and challenges for the project's success in the European context from a technical, organisational and legal point of view.

Sharing data and knowledge datasets by public and private actors are cornerstones of making data spaces real. Traditionally, private companies and public organisations from different sectors produced and collected a wide variety of data and stored it in data silos, which have largely impeded the creation and sharing of knowledge, which is in turn a central pillar of a data-driven society. In their study [9], Jean-Quartier and colleagues from the Graz University of Technology and Know-Center GmbH, Graz, investigated the positive and limiting factors influencing the enhanced cooperative utilisation of data between producers and users of data. The authors explored the challenges of participating in cross-organisational data ecosystems in a regional area in Austria by conducting interviews to public and private organisations to shed some light on data sharing, technical infrastructure requirements, practices and expert knowledge (skills) on collaborative data use. The results of the qualitative study suggest that while there is a willingness from all stakeholders to participate in collaborative data spaces, the organisational changes, synergies and improvements required to reach the necessary standards that would make a true collaborative data ecosystem possible are currently the main impediments. Studies with a regional perspective such as [9] are necessary for identifying local strengths and

weaknesses, as well as lessons learned, which can be exported to other similar regions or even to wider contexts.

While the creation of cooperative data spaces is absolutely necessary, improving methodological practices for data publishing and sharing is equally crucial to providing best practices for data-driven innovation and decision making. In their paper [10], a Belgian team of academics and government officials was tasked to investigate which methods of publishing linked open-data time series are suitable in a sustainable and cost-effective manner. The authors focus on smart cities and sensory data to demonstrate the cross-domain applicability of their data publishing approach by examining two different use cases: (i) air-quality data series captured by the delivery vans of a postal operator, and (ii) railway infrastructure data shared in accordance with Linked Open Data principles. Experimental results show that the linked data approach applied at public endpoints for both use cases substantially reduces the cost of data publication while increasing data availability and interoperability.

Further elaborating on the importance of data sharing for making the European data spaces a reality, an international group of researchers (Lorenzino Vaccari, Monica Posada, Mark Boyd and Mattia Santoro) investigated the role of Application Programming Interfaces (APIs) from a science and policy perspective [11]. Their work analyses the role of APIs as a catalyst for digital innovation for EU governments. The contribution of the colleagues is based on a thorough landscape analysis and draws insights from (i) prominent policy documents (both European and national), (ii) the investigation of available web APIs and standards and (iii) the analysis of almost 4000 relevant documents. The analysis of this rich knowledge base identifies the opportunities, enablers, challenges and bottlenecks for the adoption of APIs by governments from several interrelated angles: legal, organisational, technical and semantic. While government strategies are still in their early stage, the results show that there is a rich and mature ecosystem of solutions and standards that can boost the public-sector uptake of APIs.

The final research paper, co-authored by Italian and Cypriot researchers, directly addresses the establishment of a common European sector-specific data space: the one on cultural heritage [12]. The study is focused on the semantic aspects of the data space and proposes a novel ontology to seamlessly organise and manage digital information about cultural heritage, which includes a varied and complex set of data ranging from scientific analyses to historical and archaeological interpretations. The ontology is grounded in the concept of a *digital twin*, introducing a so-called Heritage Digital Twin, and is based on existing domain-specific standards that ensure interoperability with data models and catalogues already in use. The authors offer the ontology as an initial effort towards the development of a cloud-based, distributed, federated and European-wide data space.

2.2. Data Descriptors

The ongoing COVID-19 crisis, particularly in its early stage, has seen an explosion of new digital technologies to track, monitor and inform about the pandemic. Among these, mobile apps have been most probably a unique example. A group of researchers from the European Commission's Joint Research Centre (JRC) analysed such apps in the period ranging from February to August 2020 [13]. They created a rich dataset composed of 837 mobile apps published across the world. The dataset includes both information retrieved from the Google Play and Apple's App Store and information manually collected by the authors—the latter including, e.g., the functionality of the apps (based on an articulated classification system), the type of app provider, the involvement of a public sector body and the presence of a clear privacy policy. While the paper already offers some basic descriptions and statistics on the apps, the authors have also released the full dataset under an open access license to stimulate its further use.

Platform economy models have recently gained attention as consumer and participatory data models for their potential to contribute to the sustainable development of society. However, cases of platform economy such as Uber, Airbnb and Deliveroo have created

great controversy about their socioeconomic impact, while other alternative models have been associated with a new form of cooperativism. An analysis by an international team of researchers from the United States, Spain and France [14] has created a new dataset based on the data collection of two European projects (DECODE and PLUS) that allows the comparison of different platform economy models and their connections with the United Nations Sustainable Development Goals. Value-added data sets such as the one provided by [14] can contribute to the development of the European strategy for data and the promotion of platforms that facilitate democratic data management and, consequently, produce a better social impact.

3. Concluding Remarks

The contributions to this Special Issue confirm that scientific evidence and technical demonstrations in support of facilitating data sharing and reuse [6], together with datasets generated by research activities, can play a central role in making the European strategy for data a reality. Armed with the instruments of the Horizon Europe research framework programme, which by default promotes the publication of open data, the EU is at the forefront in adhering to Finding, Accessible, Interoperable and Reusable (FAIR) data management principles. This is leading to unprecedented amounts of data being made available under open licenses. The implementation of the European strategy for data, in turn, provides a multitude of opportunities that can help scale and sustain scientific findings by fusing research data with other sources in a demand-driven manner. Furthermore, the ongoing Digital Europe Programme [15] and the EU Member State Recovery and Resilience strategies act as a catalyst for digital transformation and can scale and sustain innovative and scientifically sound data sharing practices.

Author Contributions: All authors contributed equally as Guest Editors of this Special Issue and in writing this editorial. All authors have read and agreed to the published version of the manuscript.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

The following abbreviations are used in this manuscript:

API	Application Programming Interface;
EU	European Union;
FAIR	Findable, Accessible, Interoperable and Reusable;
IoT	Internet of Things;
JRC	European Commission's Joint Research Centre;
OSM	OpenStreetMap.

References

1. European Commission. A European Strategy for Data. EC Staff Working Document, 2020. COM(2020) 66 Final. Available online: https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52020DC0066&from=EN (accessed on 16 February 2022).
2. European Commission. Proposal for a Regulation of the European Parliament and of the Council on European Data Governance (Data Governance Act), 2020. SWD(2020) 295 Final. Available online: https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52020PC0767&from=EN (accessed on 16 May 2022).

3. European Commission. Proposal for a Regulation of the European Parliament and of the Council on Harmonised Rules on Fair Access to and Use of Data (Data Act), 2022. SWD(2022) 34 Final. Available online: https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52020PC0767&from=EN (accessed on 16 May 2022).

4. European Parliament and European Council. Directive (EU) 2019/1024 of the European Parliament and of the Council of 20 June 2019 on open data and the re-use of public sector information (recast). *Off. J. Eur. Union* **2019**, *L 172*, 56–83.

5. Kotsev, A.; Minghini, M.; Tomas, R.; Cetl, V.; Lutz, M. From Spatial Data Infrastructures to Data Spaces—A Technological Perspective on the Evolution of European SDIs. *ISPRS Int. J. Geo-Inf.* **2020**, *9*, 176. [CrossRef]

6. Granell, C.; Mooney, P.; Jirka, S.; Rieke, M.; Ostermann, F.; van Den Broecke, J.; Sarretta, A.; Verhulst, S.; Dencik, L.; Oost, H.; et al. *Emerging Approaches for Data-Driven Innovation in Europe*; EUR 30969 EN; JRC127730; Publications Office of the European Union: Luxembourg, 2022.

7. European Commission. Common European Data Spaces. EC Staff Working Document, 2020. SWD(2022) 45 Final. Available online: https://digital-strategy.ec.europa.eu/en/library/staff-working-document-data-spaces (accessed on 16 May 2022).

8. Minghini, M.; Sarretta, A.; Napolitano, M. OpenStreetMap Contribution to Local Data Ecosystems in COVID-19 Times: Experiences and Reflections from the Italian Case. *Data* **2022**, *7*, 39. [CrossRef]

9. Jean-Quartier, C.; Rey Mazón, M.; Lovrić, M.; Stryeck, S. Collaborative Data Use between Private and Public Stakeholders—A Regional Case Study. *Data* **2022**, *7*, 20. [CrossRef]

10. Buyle, R.; Van de Vyvere, B.; Rojas Meléndez, J.; Van Lancker, D.; Vlassenroot, E.; Van Compernolle, M.; Lefever, S.; Colpaert, P.; Mechant, P.; Mannens, E. A Sustainable Method for Publishing Interoperable Open Data on the Web. *Data* **2021**, *6*, 93. [CrossRef]

11. Vaccari, L.; Posada, M.; Boyd, M.; Santoro, M. APIs for EU Governments: A Landscape Analysis on Policy Instruments, Standards, Strategies and Best Practices. *Data* **2021**, *6*, 59. [CrossRef]

12. Niccolucci, F.; Felicetti, A.; Hermon, S. Populating the Data Space for Cultural Heritage with Heritage Digital Twins. *Data* **2022**, *7*, 105. [CrossRef]

13. Tsinaraki, C.; Mitton, I.; Minghini, M.; Micheli, M.; Kotsev, A.; Hernandez Quiros, L.; Spinelli, F.A.; Dalla Benetta, A.; Schade, S. Mobile Apps to Fight the COVID-19 Crisis. *Data* **2021**, *6*, 106. [CrossRef]

14. Fuster Morell, M.; Espelt, R.; Senabre Hidalgo, E. Data for Sustainable Platform Economy: Connections between Platform Models and Sustainable Development Goals. *Data* **2021**, *6*, 7. [CrossRef]

15. European Commission. The Digital Europe Programme, 2021. Available online: https://digital-strategy.ec.europa.eu/en/activities/digital-programme (accessed on 11 July 2022).

Short Biography of Authors

Marco Minghini Marco Minghini specialised in geomatics and Geographic Information Systems (GIS) and received a PhD at Politecnico di Milano in 2014. He first worked as a Postdoc at Politecnico di Milano and since 2018 he is a Scientific Project Officer at the European Commission's Joint Research Centre in Ispra (Italy), working in the Digital Economy Unit in support of European data sharing/interoperability policies, particularly in the geospatial field. He is an expert in geospatial data, standards and technology and a promoter of open source software, open data and open/reproducible science.

Alexander Kotsev Alexander Kotsev holds a PhD in GIScience and is a Fulbright senior scholar. He is currently Team Leader at the European Commission's Joint Research Centre, which he joined in 2012. His research focuses on the technological enablers for data-driven innovation across multiple domains and levels of governance, spanning across several areas such as IoT, citizen science, interoperability, standard-based web services, spatio-temporal data and GIS.

Carlos Granell Carlos Granell graduated in computer engineering in 2000. He received the PhD degree in computer science from the Universitat Jaume I, Castellón, Spain. He was a Postdoctoral Researcher at the European Commission's Joint Research Centre, Italy. He is currently an Associate Professor at the Universitat Jaume I. His research interests include multidisciplinary application of geographic information science, spatial analysis and visualization of new forms of spatial data, and reproducibility research practices.