New statistical resilience or how to survive in the data ecosystem

Ana Carmen Saura Vinuesa
Yolanda Gómez Menchón
Ana Cánovas Zapata

First draft: June 2024

This draft: June 2024

New statistical resilience or how to survive in the data ecosystem

Abstract

As declared by the Commission "*the European data strategy aims to make the EU a leader in a data-driven society. Creating a single market for data will allow it to flow freely within the EU and across sectors for the benefit of businesses, researchers and public administrations*", it means that this Strategy necessarily crosses its path with the statistical functions. This could be a threat or an opportunity for the ESS.

Aware of this situation and taking into account all the previous initiatives already carried out by the ESS, the revision of the Statistical Law has been boosted to tap the potential of the new data sources for official statistics available in the current digitalised society.

This paper scrutinizes the opportunities, challenges and potential roles for the statistical community in the future implementation of the amended Statistical Law within the context of the new digital and legal ecosystem. It also looks for synergies with the European Data Strategy and related legal acts (Data Governance Act, Artificial Intelligence and Interoperable Europe Act).

Authors and Affiliations

Ana Carmen Saura Vinuesa
Yolanda Gómez Menchón
Ana Cánovas Zapata
National Statistics Institute (INE), Spain

New statistical resilience or how to survive in the data ecosystem

Ana Carmen Saura Vinuesa[1], Yolanda Gómez Menchón[2], Ana Cánovas Zapata[3]

[1, 2, 3] *National Statistics Institute (INE), Spain*

Abstract

As declared by the Commission "*the European data strategy aims to make the EU a leader in a data-driven society. Creating a single market for data will allow it to flow freely within the EU and across sectors for the benefit of businesses, researchers and public administrations*", it means that this Strategy necessarily crosses its path with the statistical functions. This could be a threat or an opportunity for the ESS.

Aware of this situation and taking into account all the previous initiatives already carried out by the ESS, the revision of the Statistical Law has been boosted to tap the potential of the new data sources for official statistics available in the current digitalised society.

This paper scrutinizes the opportunities, challenges and potential roles for the statistical community in the future implementation of the amended Statistical Law within the context of the new digital and legal ecosystem. It also looks for synergies with the European Data Strategy and related legal acts (Data Governance Act, Artificial Intelligence and Interoperable Europe Act).

Keywords: Statistical law, institutional environment, digital ecosystem, new roles, legal issues.

1. Introduction

Since 2013, timid steps have been taken in the European Union to regulate the data economy, focusing firstly on the amendment of the regulation of the re-use of public sector information (open data). In parallel and independently, the efforts in the European Statistical System (ESS) focused on the use of Big Data for official statistics and, in particular, on access to these data.

But 2020 marked a leap forward with the adoption by the European Commission of the *Communication 'A European Strategy for Data'* outlining a strategy on policy measures and investments to enable the data economy of the next five years. This strategy aims at creating a single market for data that will ensure Europe's global competitiveness and data sovereignty. This also leads the creation of Common European Data Spaces fostering that more data becomes available for use in the economy and society.

The Data Strategy was presented at the same time as the *European Commission's Communication 'Shaping Europe's Digital Future'* and the *White Paper on Artificial Intelligence*. As a result of the implementation of this policy, new European legal acts with a direct impact on official statistics have been adopted.

Firstly, the Data Governance Act aims to increase trust in data sharing, strengthen mechanisms to increase data availability and overcome technical obstacles to the reuse of data. It allows to make more data available and facilitate data sharing across sectors and EU countries in order to leverage the potential of data for the benefit of European citizens and businesses. This law also supports the set-up and development of Common European Data Spaces (CEDS) in strategic domains, which open new data sources for official statistics.

Secondly, the Data Act boosts the EU's data economy by unlocking industrial data, optimising its accessibility and use, and fostering a competitive and reliable European cloud market. It seeks to ensure that the benefits of the digital revolution are shared by everyone. It is easy to see that the Data Act lays the foundations for a market for data. As a consequence, from the ESS, the recent reform of Regulation 223/2009 on European statistics has aimed at regulating the most harmonised possible access to big data of private data holders (PDH) beyond the potential impact of Data Act on official statistics.

Thirdly, the Interoperable Europe Act sets up a new cooperation framework for EU public administrations to ensure the unified delivery of public services across borders, and to provide for support measures promoting innovation and enhancing skills and knowledge exchange. This new law establishes an interoperability governance structure with a view to creating an ecosystem of shared interoperability solutions for the EU's public sector. This way, public administrations in the EU can contribute to and re-use such solutions, innovate together and create added value.

Finally, the Artificial Intelligence Act (AI) aims to provide AI developers and deployers with clear requirements and obligations regarding specific uses of AI, while reducing administrative and financial burdens for business, in particular small and medium-sized enterprises.

Having all that in mind, it is clear that the European Data Strategy and the new legal ecosystem cross their paths with the statistical functions and this could be a threat or an opportunity for the ESS.

2. The amendment of the EU Statistical Law in 2024: opportunities and challenges

A significant amendment of Regulation (EC) No 223/2009 has just been undertaken due to the increasing emphasis on digitalization and technological advancements in the EU, as it transits into a more data-driven society. The new text offers great opportunities for the European Statistical System but also some challenges to face.

The most relevant modification of the Statistical Law is the granting of access to privately held data to the institutions responsible for developing, producing and disseminating European

statistics. This requirement is established without prejudice to reporting obligations, data collections or any data access for the private data holders laid down in sectorial statistical legislation of the Union or the Data Act.

This is a great opportunity for the ESS, since it goes beyond the Data Act that only obligates data holders to make data available based on exceptional need and/or public emergency and does not establish any condition for the daily access to private data needed for the production of official statistics. "Privately held data" (PHD) are basically Big/Smart Data held by private data holders, it refers to the "*vast amount of data held by private entities obtained as a result of their activity, which could be used by statistical authorities to produce official statistics*" (recital 10).

Another advantage to highlight is that the data and metadata are obtained free of charge. This marks a difference with the Data Act that states, in article 20, that data holder shall be entitled to fair remuneration for making data available which shall cover the technical and organisational costs incurred to comply with the request plus a reasonable margin. It also establishes that data holders shall not be able to request compensation for making data in case the specific task in the public interest is the production of official statistics and where the purchase of data is not allowed by national law. Therefore, if in cases of exceptional need "a fair remuneration" is established, this rule would also be applied in all cases, with the exception of those countries whose national law does not permit the purchase of data.

However, the "free of charge" rule is compatible, as the Statistical Law mentions, with a possible compensation from the NSIs to the PDH when the access to the data requires specific processing. Moreover, a differentiation about the compensation is set up depending on who request the data. When data are requested by Member States, they may decide whether there is a compensation or not, while when they are requested by the Commission, it will propose reasonable compensation whenever there is a data specific processing service.

Nevertheless, the Statistical Law has also included some conditions to provide data free of charge. Firstly, the data requested is strictly necessary for the development, production and dissemination of European statistics (excluding, then, the access for the production of national official statistics). Secondly, and what is more difficult to demonstrate, is that the data cannot be obtained by other means or their reuse will result in a considerable reduction in the response burden on data holders and other businesses. This will imply that NSIs will have to justify case by case why the data could not be obtained by other means. Besides, there is a contradiction in this final text, because "purchasing" would be another mean to obtain data, and this is against the rule of access free of charge of all kind of data for European statistical purposes. Then, the challenge will be to show that the access to private data will result in a considerable

reduction in the response burden. In general, any access to PHD reduces the burden to the overall interviewers but may increase the burden on the concrete private data holder to whom the data are required.

The third condition is voluntary, data collections or data accesses may be included by the Commission in the annual work programme. The idea here is to be more transparent to the private data holders mentioning when this kind of data could be necessary for certain statistics.

Fourthly, as regards the private data holders, the NSIs or the Commission could only address to small and micro enterprises, in duly justified cases when the data held by them are of specific interest for official statistics because of the nature and volume of those data at national level. Here also, the burden of proof is placed on the statistical authorities.

Concerning the data requested, there are also obligations and challenges to be met by Member States and the Commission. First, the data requested shall primarily concern non-personal data, and only in specific circumstances, personal data of the categories specified in sectoral legislation. Here, the first challenge is to understand what "personal data of the categories specified in sectoral legislation" are. For example, are the data requested under the IESS Regulation[1] and its implementing and delegated acts considered as personal data specified in sectoral legislation? Moreover, the secure infrastructure adopted shall be based on technologies specifically designed to comply with GDPR[2] requirements, in addition, the NSIs and the Commission (Eurostat) shall establish appropriate safeguards regarding the processing of personal data for statistical purposes particularly to ensure respect for the principle of pseudonymization. To finalise, the data shall only be used in conformity with statistical principles for statistical purposes (obviously), and will not be shared outside the ESS unless the data holder has agreed to do so. These conditions on the data that follows a restrictive interpretation on GDPR further complicates the access to these kind of data, setting up differences between the access to PHD and the access to personal data for statistical purposes by "traditional" means that are difficult to explain and understand.

The procedure to access PHD is divided in three phases. First, the request that consists of informing on the data and metadata requested, the statistical need, the frequency, the

---

[1] Regulation (EU) 2019/2152 of the European Parliament and of the Council of 27 November 2019 on European business statistics, repealing 10 legal acts in the field of business statistics.

[2] Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation) and Regulation (EU) 2018/1725 of the European Parliament and of the Council of 23 October 2018 on the protection of natural persons with regard to the processing of personal data by the Union institutions, bodies, offices and agencies and on the free movement of such data, and repealing Regulation (EC) No 45/2001 and Decision No 1247/2002/EC.

deadlines and the operational arrangements. Second, a dialogue with the concerned data holder to discuss and agree on the measures needed for making data available. Third, a possible second request issued to the data holder to ensure the effective enforcement of the first requests, which could be followed by enforcing measures, including fines (only if the procedure has been started by the Commission).

Figure 1: Opportunities and challenges on the access to PHD under new Regulation 223/2009.
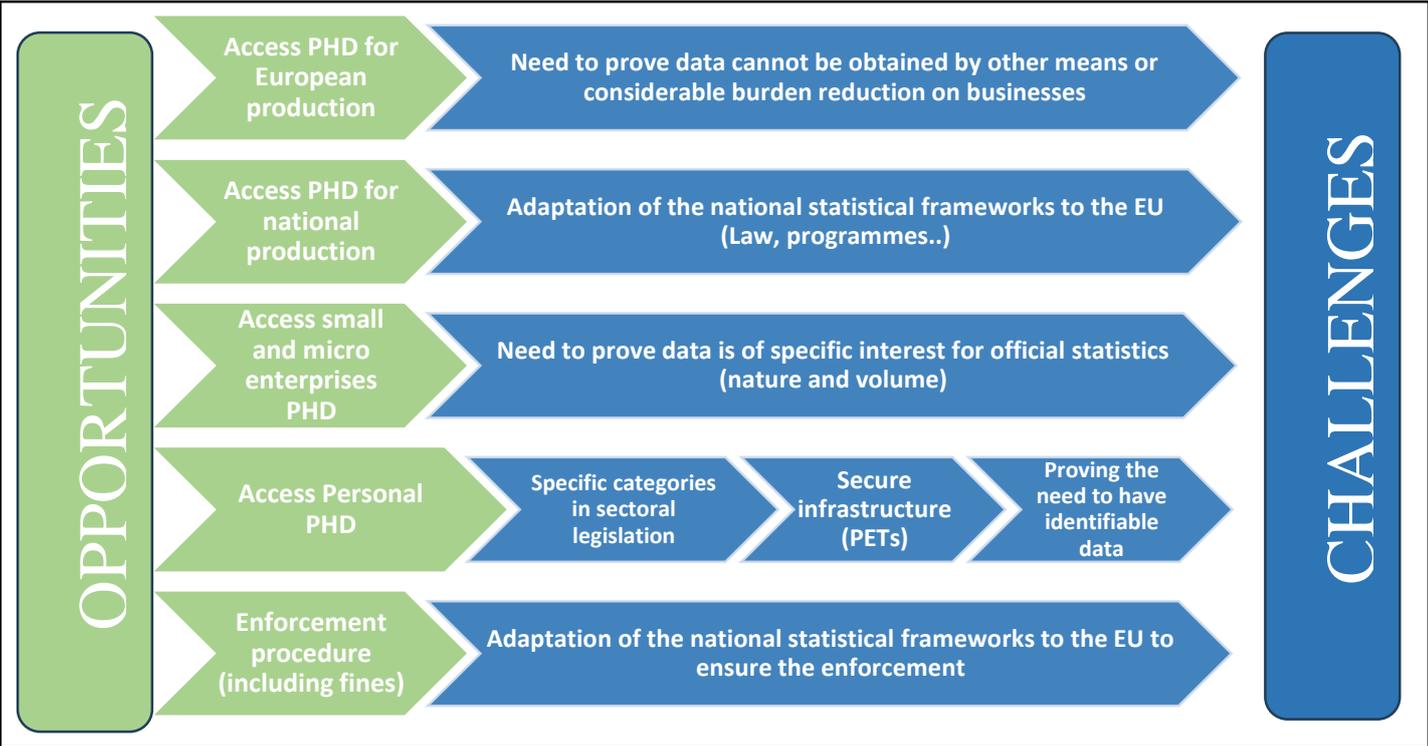
3. The EU Data Ecosystem

3.1. Potential roles for the statistical community

The DGA Regulation establishes the conditions for the reuse, within the Union, of certain categories of data held by public sector bodies (beyond the Open Data); the notification and supervision framework for the provision of data intermediation services; the framework for voluntary registration in a registry of entities that collect and process data transferred for altruistic purposes, and the creation of a European Data Innovation Board (EDIB).

In particular, as regards the statistical confidential data the DGA Regulation does not establish new obligations for the NSIs to allow access to information and its subsequent re-use, but rather incorporates a series of novel mechanisms that aim to make access to information compatible, as far as possible, with respect for the confidentiality requirements. In fact, the need to respect statistical confidentiality means that the application of the DGA is subject to the provisions of our national and European statistical laws for access to confidential data for

scientific purposes. Besides this, the DGA provides for some roles for Member States public bodies that could have implications for the NSIs.

A first role is mentioned in article 5, that establishes that Member States shall designate the competent public sector bodies under national law to grant or refuse access for re-use of data. It is clear that, by law, NSIs are competent statistical public sector bodies regarding the access to confidential data for scientific purposes, so for that case they already play this role by default.

A second role is defined in article 7.1, establishing that Member States shall designate one or more new or existing competent bodies, which may be competent for certain sectors, to provide assistance to public sector bodies granting or denying access for data reuse, with state-of-the-art techniques, including on how to best structure and store data to make data easily accessible, in particular through application programming interfaces, as well as make data interoperable, transferable and searchable. In this respect, the national Data Protection authorities may be designated as national competent authorities without prejudice to the supervisory powers and competences of data protection authorities under GDPR law. Some NSIs have been also designated as national sectorial authorities under article 7.1, which implies an additional advisory role to the users and the possibility of being part of the EDIB through the subgroup of "competent national authorities", and participating in the establishment of harmonised rules.

Finally, a third role is mentioned in article 8, which states that, among these authorities, the Member States shall establish a new body or designate an existing body or structure as the single information point (SIP). An NSI could also be the SIP, but this will implies to assume the role of national datastewardship and the need to separate clearly from its statistical functions.For the Common European Data Spaces, the statistical offices could play a relevant role. For instance, Eurostat and NSIs could contribute to define the technical and non-technical norms that shape the functioning of the CEDS, ensure interoperability, alignment of classifications and definitions adopted by the CEDS, or at least coherence with those in use in the ESS, to maximize synergies between the CEDS and the ESS.

As regards the Data Act, Chapter V deals with the access to PHD when there is a exceptional need. As for the implementing of the B2B (business-to-business) rules, article 37 foresees that each Member State shall designate one or more new or existing competent authorities to be responsible for the application and enforcement of the Data Act. If there are more than one, the Member State shall designate a data coordinator. For some tasks the competent authority shall have experience in the field of data and electronic communications services. Anyway, this is without prejudice of the competence of sectoral authorities for specific sectoral data access and uses, that shall be respected (as the statistical tasks of the statistical authorities).

If the NSI wants to be national authority under this regulation, it also will imply assuming datastewardship roles beyond the statistical functions that could report some benefits (as being part of the governance bodies and influence in the decision-making) but also a clear extra workload (so, it would need more financial and human resources).

As for the Interoperable Europe Act, that entered into force in April this year, it offers solutions and tools for citizens, businesses, and public administrations to be used in interconnected digital public services that require cross-border exchange of data. The Regulation applies to Union entities and public sector bodies that regulate, provide, manage or implement trans-European digital public services and there is no exception for EU sectorial legislation. Therefore, it is in principle applicable to the exchange among ESS members and also ESCB members.

Here, Member States shall also designate one or more national competent authorities responsible for its implementation (if there are already entities in charge of interoperability issues, they should take over the role of competent authority). Where there is more than one national competent authority, a single point of contact should be designated from among them.[3]

The Interoperability Act foresees the development by the Board of the draft of an European Interoperability Framework (EIF) that will be adopted by the Commission, it may also propose other Interoperable Europe solutions and propose to the Commission to set up innovation measures. There are also rules for implemeting Interoperability regulatory sandboxes. Anyway, the Commission may adopt other interoperability frameworks (specialised interoperability frameworks) targeting the needs of specific sectors or administrative levels, that shall, however, be based on the EIF. It is not compulsory for the development of national interoperability framework and other relevant national policies, strategies or guidelines but the EIF shall be taken into utmost account. It is clear that this could affect and jeopardize the ESS work included in the Inovation Agenda and even tools we have already developed. Here, the NSIs could try to have some role as advisor, but the normal situation is that the entities that are currently dealing with interoperability issues in the Member States continue with this role (that normally are not the NSIs).

The interconection with the Artifitial Intelligence Act might come when the statistical offices or the ESS develop Artificial Intelligence tools to collect information because IA Act applies to providers or deployers that put into service AI systems. The exception for the ESS is AI systems or AI models, including their output, specifically developed and put into service for the

---

[3] https://joinup.ec.europa.eu/interoperable-europe

sole purpose of scientific research and development (excluding testing in real world). The ESS should take it into account for the implementation of the Innovation Agenda and in particular, the proyect "One-Stop-Shop for Artificial Intelligence/Machine Learning for Official Statistics for the ESS". As for the roles, here the national competent authority means a notifying authority or a market surveillance authority, both are very far from the NSIs tasks.

Table 1: European Bodies and groups and National authorities implied in the data governance

	DGA	Data Act	Data Spaces	Artifitial intelligence Act	Interoperability Act
EU Bodies or Groups	European Data Innovation Board *(Subgroup on standardisation, portability and interoperability & Stakeholder engagement sub-group & Subgroup of national competent authorities)*	Advisor of the EDIB	DSSC (Strategic Stakeholder Forum) SIMPL[4] Sectorial Data Spaces Bodies	European Committee on AI (*Standing Subgroup on Market Surveillance & Standing Subgroup on Notifications*) Advisory Forum AI	Interoperable Europe Board Interoperable Europe portal Interoperable Europe Support Centre
Competent National Authorities	SIP (art.8) Competent public sector bodies (Art.4) Nat. Auth. Art.7.1	Competent National Authority (Data Coordinator)	National Authority	Notifying Nat. Auth Market Surveillance Nat. Auth.	National Authority
Data Protection Bodies	Data protection authority Nat. Auth (Art.7.1)	EDPS Data Protection Auth.	EDPS Data Protection Auth.	EDPS Data Protection Auth.	EDPS Data Protection Auth.

3.2. New resilience of the statistical authorities: looking for synergies in a potential European Statistical Strategy

Once the first years of uncertainty and confusion about the development of the data economy policy have passed, it is the moment for the ESS to take position and make decisions, keeping in mind the subsidiarity principle. Then, the fact that some NSIs have decided to assume some tasks and roles in the data governance beyond their statistical competence and other don't does not change the European scenario.

Nevertheless, what it is clear is that when Member States act and work together in the ESS we are stronger. The ESS has already expressed their position[5] on the Data Strategy letting clear that *"Statistical offices (…) have a longstanding experience in handling large amounts of*

---

[4] Smart Open-source Middleware.

[5] Position Paper on the Future data act proposal (June 2021) and Position paper on Future governance framework for the common European data spaces (November 2020).

data for the common good and with a particular attention to the preservation of confidentiality. They are natural partners when it comes to building trust in data and building trust between those actors who generate or hold the data and those who want to use them". This message has been "brought" by the legislator in the new article 26a of Regulation 223/2009. This article recommends Member States assign to the NSIs at national level functions laid down in the national data governance frameworks with the goal of promoting data integration and interoperability, metadata description, quality assurance and standard setting, data sharing and re-use of data, as well as other tasks and functions laid down in DGA. These data steward functions shall be compatible with the exercise of the statistical functions performed in conformity with the statistical principles.

Therefore, the NSIs can now address to their governments to request for having its role in the data ecosystem (we consider that the minimum would be to be designated as competent national authority under article 7.1 DGA). In addition, the NSI should enhance the participation in other informal and technical groups, such as Data Spaces Strategic Stakeholder Forum, the IA Advisory Forum, Interoperability regulatory sandboxes, etc. For the ESS it is important to achieve a "statistical specialised interoperability framework" trying to follow/participate closely in the development of the European Interoperability Framework (EIF) and the IA implementing rules. What it is clear from the very beginning is that we cannot continue working isollated in our "statistical world". Even if we decided not to assume any additional role beyond the statistical one, we have to cooperate with the data governance bodies in order to be able to continue with our statistical tasks. Rules are changing for using data across the EU and we should be involved in the governance and technical functions.

Data Protection bodies are also authorities that we have to take seriously present and involved because they have an imortant and increasing role to play in the data ecosystem.

We believe that the ESS need to set up a statistical strategy including all possible synergies from the governance level to the technical one.
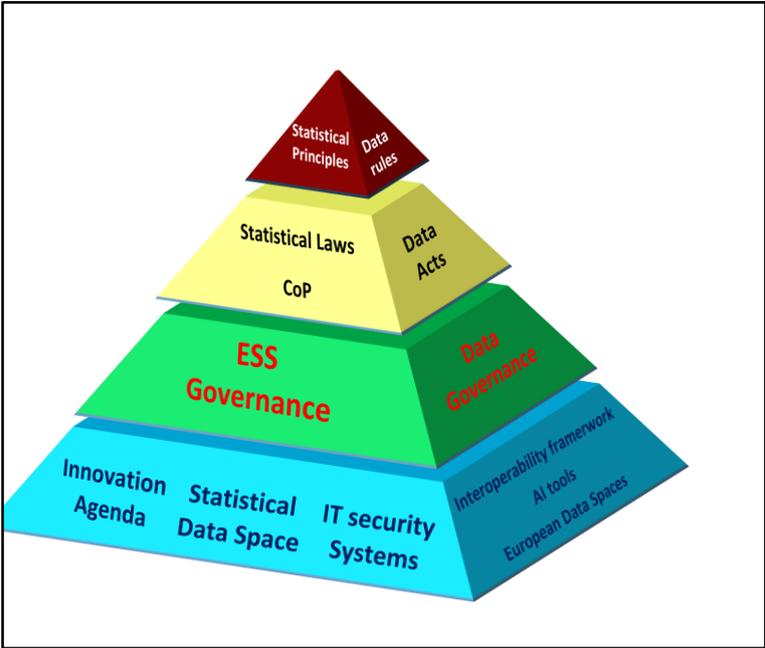

4.   Conclusions

The ESS adopted its own Innovation Agenda in 2023 and the European Statistical System Committee (ESSC) and Partnership Group (PG) are now analysing how design a strategy to take advantage of the data acts. The Innovation Agenda was elaborated, mainly, from a technical point of view. It foresees interaction with stakeholders and, specially with the academia. The approach of the Agenda is good to develop and test new sources and tools, but it should be complemented with a strategic governance as part of a European Statistical Strategy.

Therefore, the possible European Statistical Strategy should be merged with the Innovation Agenda. The political/strategic issues regarding our interaction with the data ecosistem governance should be linked to specific targets and possible projects of the Innovation Agenda. On the contrary, we would continue our issolated path[6]. Nevertheless, this Strategy should be wider, not limiting only to the data sphere, and cover all the relevant issue for the ESS over the next decade.

The different initiatives mentioned in this paper, with their opportunites and challenges, should be considered for drafting this Strategy. In particular, the statistical interoperability framework, the kind of IA tools we might use and how to adapt them (after the developing phase) to the IA regulation, how the ESS will interact with the different European Data Spaces and if it is necessary or not a European Statistical Data Space should be part of that Strategy. It should be designed having on top our statistical principles as well as the specific rules of the data acts and those of our statistical law. The ESS governance role could be addressed to the Partnership Group. Finally, the interaction with Data Governances (not only CEDS), and all the technical and methodological actions (including the Innovation Agenda) should also be included in the Strategy as Figure 2 shows.

Figure 2: Layers of a possible European Statistical Strategy.

References

[1] ESS Position Paper on the Future data act proposal (June 2021). https://ec.europa.eu/eurostat/documents/13019146/13405116/main+ESS+position+paper+on+future+Data+Act+proposal.pdf/37f3b5c7-abfd-5a05-6be2-fdc4b87ee7d2?t=1631695372906

[2] ESS Position Paper on the Future governance framework for the common European data spaces (November 2020). https://ec.europa.eu/eurostat/documents/13019146/13405116/ESS+Position+paper+on+the+future+governance+framework+for+the+common+European+data+spaces.pdf/dfc6b351-a04c-f536-578d-d8bbe01f9ad1?t=1643724796324

[3] ESS Innovation Agenda. ESSC document (February 2023). https://cros.ec.europa.eu/dashboard/innovation-statistics

[4] ESSC Document (2019).: Trusted Smart Statistics Actions and Roadmap 2021-2027

[5] European Commission Communication "Artificial Intelligence for Europe" (COM/2018/237 final).

[6] European Commission Communication "Building a European data economy" (COM(2017) 9 final) 2017).

[7] European Commission Communication "Towards a common European data space" (COM/2018/232 final2018).

[8] European Commission Communication "Towards a thriving data-driven economy" (COM/2014/0442 final2014).

[9] European Commission Communication 'A European Strategy for Data' (COM/2020/66 final).

[10] European Commission Communication 'Shaping Europe's Digital Future' (COM/2020/67 final)

[11] European Commission White Paper on Artificial Intelligence (COM/2020/65 final).

[12] European Commission Communication of 19 May 2010 "A Digital Agenda for Europe" (COM(2010) 245 final).

[13] European Commission Communication of 5 July 2022 "A New European Innovation Agenda" (COM(2022) 332 final).

[14] European Council's conclusions on digital economy and innovation (2013),

[15] European Parliament legislative resolution of 13 March 2024 on the proposal for a regulation of the European Parliament and of the Council on laying down harmonised rules on Artificial Intelligence (Artificial Intelligence Act) and amending certain Union Legislative Acts (COM(2021)0206 – C9-0146/2021 – 2021/0106(COD).

[16] European Parliament legislative resolution of 13 March 2024 on the proposal for a regulation of the European Parliament and of the Council amending Regulation (EC) No 223/2009 on European statistics (COM(2023)0402 – C9-0246/2023 – 2023/0237(COD)

[17] EUROPEAN STATISTICS CODE OF PRACTICE. https://ec.europa.eu/eurostat/documents/4031688/8971242/KS-02-18-142-EN-N.pdf/e7f85f07-91db-4312-8118-f729c75878c7?t=1528447068000.

[18] Memorandum of Bucharest on Official statistics in a Datafied society (-Trusted Smart Statistics)- adopted by the ESSC on 12 October 2018. https://ec.europa.eu/eurostat/documents/13019146/13237859/The+Bucharest+Memorandum+on+Trusted+Smart+Statistics+FINAL.pdf/7a8f6a8f-9805-e77c-a409-eb55a2b36bce?t=1634144384767

[19] Memorandum of Scheveningen on Big Data and official statistics (2013), adopted by the ESSC on 27 September 2013. https://ec.europa.eu/eurostat/documents/13019146/13237859/Scheveningen-memorandum-27-09-13.pdf/2e730cdc-862f-4f27-bb43-2486c30298b6?t=1401195050000

[20] Regulation (EU) 2016/679 on the protection of natural persons with regard to the processing of personal data (GDPR).

[21] Regulation (EU) 2018/1725 on the protection of natural persons with regards the processing of personal data by the Union institutions, bodies, offices and agencies.

[22] Regulation (EU) 2022/868 of the European Parliament and of the Council of 30 May 2022 on European data governance and amending Regulation (EU) 2018/1724 (Data Governance Act).

[23] Regulation (EU) 2023/2854 of the European Parliament and of the Council of 13 December 2023 on harmonised rules on fair access to and use of data and amending Regulation (EU) 2017/2394 and Directive (EU) 2020/1828 (Data Act).

[24] Regulation (EU) 2024/903 of the European Parliament and of the Council of 13 March 2024 laying down measures for a high level of public sector interoperability across the Union (Interoperable Europe Act).

[25] Regulation (EU) 223/2009 of the European Parliament and of the Council of 11 March 2009 on European statistics.

[26] Regulation (EU) 2019/2152 of the European Parliament and of the Council of 27 November 2019 on European business statistics, repealing 10 legal acts in the field of business statistics.