

Comments on the European Commissions' EU Data Strategy

Franziska Boehm, Oliver Vettermann

Summary

The following is a response to the European Commissions' EU Data Strategy and an addition to the web-based survey. The comments address the Pillars in the strategy and refer to:

- **Pillar I:**
 - **The one-sided concept for the cross-sectoral governance framework: the strategy omits to involve research**
 - Concrete example: the “Implementing act on high-value data sets” should **consider and involve research institutions** that deal with public data or are in need of easier access to these data.
- **Pillar II:**
 - **The difference of first-line and second-line enablers leads to an imbalance or drift between research institutions and economy**
 - Further, **to protect the rights and freedoms of data subjects, the definition of pseudonymization and anonymous/anonymized data** should be addressed in more detail.
- **Pillar III:**
 - Compared to pillars I and II, the portfolio of measures in this Pillar is noticeably short.
 - **There is a lack of legal empowerment for individuals and a more concrete relation to the GDPR and privacy related rights is absolutely necessary.**
- **Pillar IV:**
 - **The term data spaces urgently needs a detailed explanation**
 - The support for existing data infrastructures needs to be mentioned and explained

1. Introduction

First of all, the approach of the EU to condense a strategy for pushing forward the transformation of economy and society by data-driven innovation should be positively received. “Data is at the centre of this transformation and more is to come.”¹ At the same time, the several individual strategies in EU Member States should also work towards the common goal set out in the strategy. Aiming for the position of the “leading role model for a society empowered by data”² through diverse regulatory models – including both diversity in discipline and measure. In the light of the societal focus on data and innovation, it is also mandatory to discuss the role of data as a public good and the possible change in legal concepts from a personal, sensitive information model to a (mostly) public property model. Further, the multilateral interests in (personal) data have to be weighted by the legislation or government to protect (and empower) individuals³, but also to avoid possible market imbalances⁴. Also important, although easy to overlook, is the need to provide standards in information and data security to protect against technical influences on the relationships between government, business and customer.

Reflecting this, the following comments on the EU Data Strategy will outline missing perspectives to facilitate a more streamlined strategy that will both remove obstacles to the strategy’s goals as well as to include scientific interests to ensure success beyond the planned period 2020-2030. In this regard, every Pillar in the strategy will be analysed and enriched with perspectives from research and development by national research data infrastructures.

2. Pillar I: The one-sided concept for the cross-sectoral governance framework

The first Pillar of the strategy outlines the necessity of a horizontal, cross-sectoral governance framework for data access and use. In terms of the strategy, cross-sectoral means measures active across sectors or disciplines and across the Member States. “Such measures should nonetheless take into account the specificities of individual sectors and of the Member States.”⁵ For example, this is reflected by the legislative framework for the governance of data spaces in general and their accompanying mechanisms – “prioritise

¹ EU Data Strategy, pp. 1.

² Ibid.

³ See EU Data Strategy, pp. 10 f.

⁴ See EU Data Strategy, pp. 8.

⁵ EU Data Strategy, pp. 12.

standardisation activities”; “principles on Findability, Accessibility, Interoperability and Reusability of data”; ‘data altruism’ as an individual right⁶ – and in the evaluation of the need for legislative action on issues that affect relations between actors in the data-agile economy. One of the main elements of this Pillar is the Data Act (2021) which focuses on the latter by supporting and building up data-driven cooperation between business and government.

Having said this, whilst a broadened spectrum of business and government interests is shown in the strategy, the spectrum does not extend much wider than these sectors. A closer look reveals that the strategy is almost entirely focused on the economic impact of data and its positive consequences for the society.

One indicator for this is the aforementioned exploration of possible legislation for the data-agile economy. Another main element of this Pillar is the Implementing act on high-value data sets (Q1 2021) where the regulatory framework will focus on Application Programming Interfaces (APIs) and their standardisation for “making more high-quality public sector data available for re-use”.⁷ This approach in general is not to be judged negatively; the possibility for easier access to public data is to be welcomed. However, it is noticeable that the foundation for the law is based on the needs of SMEs.⁸ Instead, this **evaluation process should be opened also to research institutions** that deal with public data or are in need of easier access to these data. This would support the multilateral approach of the strategy and could prolong or positively influence the standardisation of the APIs as well as their development in the future. This is also essential from the perspective of information and data security, which has to be ensured actively to protect both data and the (bilateral) API-using institutions.

A minor indicator for the lack of inclusion of the research perspective can be found in an issue which will be relevant for the mentioned Data Act. Here, the strategy paper plans to evaluate “the IPR framework with a view to further enhance data access and use (including a possible revision of the Database Directive [...])”⁹. Reading this with the stated economic focus in mind, the legal exceptions for research and science in Art. 6(2)(b) Directive 96/9/EC do not fit within this framework even if it fits to the wording of the strategy paper. To **take into account the interests of research and science institutions** in this respect, an analysis is required as to whether, and how, the

⁶ Ibid.

⁷ EU Data Strategy, pp. 13.

⁸ Ibid.

⁹ EU Data Strategy, pp. 13.

IPR framework also needs to be adapted. Otherwise, research and economy will drift further apart, making the intended standardisation needlessly difficult. Instead, **existing particularities on both sides should be considered** when implementing the strategy in order not to endanger it. The apparent limitation of research institutions to the European Open Science Cloud (EOSC) in the strategy's closing example¹⁰ at least indicates that the European Commission is willing to integrate this thought.

3. Pillar II: The difference of first-line and second-line enablers

The second Pillar of the EU Data Strategy details the measurement of the endeavour. The so-called enablers consist in the “convening power [of the Commission] as well as EU funding programmes to strengthen Europe’s technological sovereignty for the data-agile economy. This will be done through standard setting, tool development, best practices collection on how to deal with personal data (especially around pseudonymization) as well as build-out of next-generation infrastructures for data processing.”¹¹ Again, and similar to Pillar I, the elements of this Pillar have a strong link to the economic perspective. In this regard, the Commission plans to invest in High Impact Projects in European data spaces and federated cloud infrastructures which deal with “data-sharing tools, architectures and governance mechanisms for thriving data-sharing and Artificial Intelligence ecosystems”, in new technologies as part of its industrial strategy and in establishing EU-wide common, interoperable, data spaces.¹² Only in passing does the strategy paper mention clear technical or legal measurements as discussed in the beginning of this paper.

Analysing this Pillar, there seems to be a **division between ‘first-line’ and ‘second-line’ enablers**. Reading the introductory definition, above, in the context of the overall alignment of the section, the funding approach is clearly the first-line Enabler. This is not only indicated by the long-term project funding between 2021-2027. The economic perspective of the strategy also supports this argument in relation to the Commission’s plans to fund High Impact Projects, new technologies as part of an industrial strategy and common data spaces. Subordinate enablers (or second-line enablers), in contrary, are mentioned less frequently: the ‘cloud rulebook’ as a compendium of existing cloud codes of conduct and certification and the elaboration of European standards; requirements for the public procurement of data processing services. Funding for research in these fields is possible

¹⁰ EU Data Strategy, pp. 15.

¹¹ EU Data Strategy, pp. 16.

¹² Ibid.

through the Horizon Europe programme, mentioned quite briefly next to the EOSC.¹³

Outlining this division, there is a great **imbalance between first-line and second-line enablers** to the disadvantage of the latter; the second sentence of the definition above is not as relevant in the planned measurements of the Commission. Also, there is no clear connection between enabler categories – even if they will influence each other both technically and practically. This structure could lead to an imbalance or drift between research institutions and economy and could push away the technical measurements of the second-line enablers. Mirroring this ratio to a data-driven reality and society would risk technical insecurities, missing guidelines in processing public and/or personal data and funding projects without significant results. In short: the funding will not help to overcome the aforementioned obstacles.

To avoid this, the imbalance should be corrected. The division can be overcome by providing further details concerning second-line enablers. For instance, the **type of standards and their development** – installed/controlled by government or established through practice – should be concretized.

Further, **to protect the rights and freedoms of data subjects, the definition of pseudonymization and anonymous/anonymized data** should be addressed in more detail. Although the CJEU has discussed the differentiation in several cases¹⁴, it is a problem for researchers to divide these terms and transfer these legal understandings into technical practice. But this problem is solely reflected in the Strategy by mentioning, in passing, the word „pseudonymization” in the definition above. This cannot be solved with the general approach of the Data Strategy to be a guideline for the differentiation of personal and non-personal data.¹⁵ What is needed instead is to address the issue clearly and integrate it actively into project funding. It could also be addressed through its own research projects, supported by the Commission. For example, advanced anonymization algorithms could be researched or existing procedures could be strengthened in order to prevent encryption from being decrypted.

¹³ See EU Data Strategy, pp. 19.

¹⁴ CJEU C-582/14 *Patrick Breyer v Bundesrepublik Deutschland* [2014] ECLI:EU:C:2016:779, Par. 45 f.; C-210/16 *Unabhängiges Landeszentrum für Datenschutz Schleswig-Holstein v Wirtschaftsakademie Schleswig-Holstein GmbH* [2018] ECLI:EU:C:2018:388, Par. 38 f.; C-673/17 *Bundesverband der Verbraucherzentralen und Verbraucherverbände - Verbraucherzentrale Bundesverband e.V. v Planet49 GmbH* [2019] ECLI:EU:C:2019:801, Par. 66 f.

¹⁵ EU Data Strategy, pp. 1, 4.

To conclude, it is worth mentioning the risks associated with a lack of correction of the imbalance. **If the focus of the strategy paper is predominantly economics, the implementation of the EU Data Strategy may end up as an uncertain technical and legal construct.** This legal uncertainty collides with the general principle of certainty laid out in Art. 2 TEU.¹⁶ Further, the technical uncertainty risks the fundamental rights in Art. 7 and 8 CFREU – in particular the technical aspects of data protection in Art. 25, 32 GDPR. Promoting both fundamental values with a long-term data strategy might lead to proceedings at the CJEU if these aspects are not sufficiently protected by the intended measurements in fulfilling the Strategy.

Pillar III: The lack of legal empowerment for individuals

This Pillar turns away from the economic approach of the EU Data Strategy by analysing the rights and freedoms of individuals to respect/protect their data. Therefore, individuals “could be supported by enhancing the portability right [...] under Article 20 of the GDPR”. Besides that, an update to the “Digital Education Action Plan will reinforce better access to and use of data [...], in order to make education and training institutions fit for the digital age”.¹⁷ However, SMEs and start-ups will profit from the support of data literacy as well.

Compared to pillars I and II, the portfolio of measures in this Pillar is noticeably short. This could be due to the fact that the Data Strategy is not primarily aimed at influencing Data Protection and Privacy law. But this does not prevent data subjects, as suppliers of the data that are the subject of the Data Strategy, from being protected. Indeed, the broad data concept of the Strategy argues explicitly for this. With reference to the shortcomings already mentioned, reference can therefore only be made, again, to clarify the uncertainties in Data Protection and Privacy law. This Strategy deals with personal data and so should guarantee the rights and freedoms of individuals which serve to protect them from any harm that may result from new technologies and extensive data processing and storage. **A more concrete relation to the GDPR and privacy related rights, as well as in elaborating compatible safeguards, is absolutely necessary.**

Pillar IV: The inadequate definition of common data spaces

Finally, the last Pillar in the strategy of the Commission explains the term “data spaces”. These are “large pools of data in these sectors and domains of

¹⁶ E.g. CJEU C-169/80 *Administration des douanes v Société anonyme Gondrand Frères* ECLI:EU:C:1981:171 [1981], Par. 17.

¹⁷ EU Data Strategy, pp. 21.

public interest.”¹⁸ To qualify as a data space, as storage infrastructure, the infrastructure must facilitate cross-sectoral use of data as well as to facilitate public interest and interoperability. Putting all the pieces from the other Pillars together, these data spaces are implemented in the horizontal framework discussed above.

First, the definition of the term is welcome in providing clarity in understanding the approach of the Data Strategy. However, **in order to understand the different versions of data spaces, the term should have been explained earlier and in more detailed manner.** For example, in the description of Pillar III, “personal data spaces” are discussed¹⁹. In Pillar IV a difference between common and uncommon data spaces is evident – this is not only stated in the title, the list of common data spaces which will be supported by the Commission also makes this differentiation clear. In contrary, the opposite ‘uncommon data space’ is never mentioned and can barely be seen in the general definition of data spaces: “While not having a one-size-fits-all-approach, common governance concepts and models can be replicated in the different sectors.” In this regard, it seems that uncommon data spaces could be funded depending on the public interest, while the listed common data spaces definitely will be supported without the need to fulfill requirements.

Further, the fact that a reference to a common European research data space (or data space for science) is missing might be due to the existing plans for the EOSC. Still, a data space for science and the support for existing data infrastructures besides the listed infrastructures should be mentioned. As a long-term strategy, national plans for interdisciplinary and cross-sector research data infrastructures – like the National Research Data Infrastructure initiative in Germany – should have been considered and implemented to ensure an interdisciplinary dialogue aimed at making research data broadly available. The absence of such approaches unfortunately contributes little to building a future-oriented data infrastructure that supports research and development.

¹⁸ EU Data Strategy, pp. 21.

¹⁹ EU Data Strategy, pp. 20.