

January 2021

Statement on the proposal for a Data Governance Act (DGA) by the EU Commission

The EU Commission's plan to set a Europe-wide legal framework for the establishment of data sharing providers and thereby stimulate data sharing is very welcome. It focuses on new players to help share data more effectively in economy and society. In order to create a legally secure and trust-building framework to enforce data sharing in the common market, the Commission sets up comprehensive requirements for new intermediary organisations. With the following four recommendations the German Council for Scientific Information Infrastructures (RfII)¹ has commented on the Commission's proposal. The council expresses its views from the perspective of an advisory body for the German government (federal and *Länder* level) which accompanies the digital transformation in science and its enabling information infrastructures. The recommendations of the RfII for the further development of the proposal on the DGA aim to enhance the current draft and give more emphasis to practical issues of fostering data sharing via data service providers. Likewise, in the view of the RfII, essential concerns of research and higher education should be taken into account in the legal establishment of new intermediaries.²

The European Data Governance Act (DGA) seeks to put the sharing and re-use of protected data - including public sector data - on a legally reliable and economically attractive basis by creating new forms of fiduciary data intermediaries (data sharing providers, hereinafter DSPs), which do not exploit the data themselves, but take stock of these data, prepare them for appropriate use and distribute them.

- On the one hand, this will affect the (re-)use of data for scientific purposes - in this respect, the RfII considers the introduction of a research clause in the DGA to be indispensable (see point A of this statement).

¹ The statement was prepared by the working group on data trusteeship and the Chair of the RfII in January 2021 and was submitted to the EU Commission's participation procedure on the DGA on 29 January 2021. The RfII adopted this statement in its 20th plenary session on 18 March 2021. Thereby, the RfII acknowledged the compromise proposal of the Portuguese Presidency of the Council of the EU on the DGA from 22 February 2021. However, regarding this proposal the RfII sees no need to make substantive changes to its own statement formulated in January. The recommendations articulated in this statement remain urgent and necessary to address the requirements of science in the further elaboration of the DGA proposal.

² Cf. RfII (2020) – Designing Trusted Intermediaries for Data Exchange. On Experiences from Science, <http://www.rfii.de/?p=4318> (last accessed on 26.01.2021).

- On the other hand, some research institutions can seek a particular role as DSPs, in terms of the DGA. Doing so, they can act either as a profit-oriented company, or as a non-profit or so-called “data altruism organisation recognised in the Union”. The Rfll comments on this in points B and C of this statement and recommends clarifications in the according norm text.
- Finally, the current DGA draft hardly addresses aspects of data quality and corresponding tasks of the DSP. However, data quality is central both for the motivation to share data and for the effective development of a European data ecosystem. In point D of this statement, the Rfll states that quality requirements in the DGA must go beyond the reference to the FAIR principles to boost the sharing and exploit the innovative features of real high quality data.³

A. Granting data access for research purposes

If protected public sector data are made accessible and usable in Europe by non-exclusive data intermediaries, these new DSPs or data sharing services (DSS) should also be obliged to grant scientists an indispensable right of access to the data held and establish appropriate access regimes. Access is, however, limited to the purposes of genuine research projects. The Rfll strongly recommends such a research (or science) clause because research is always an innovative force to improve quality and linkage possibilities for data.

The introduction of a research clause is urgently needed from the perspective of scientific actors. The digital transformation should not deteriorate the conditions for scientific research beneath a level that was already reached in access to analogue public (and other) data. This applies specifically when access management for data is shifted from data producing entities to the new intermediaries. Publicly funded research has been privileged in the past in most areas of public sector information (medicine, public archives, labour market data, environmental and climate data, art and cultural assets) for good reason in terms of data access. Similarly, the DGA must also ensure the innovative capacity of universities and publicly funded research institutions. Regulatory guidelines are required, but these can be agreed within the framework of registration or accreditation procedures for DSP/DSS. Depending on the sensitivity of the data, different access regimes can be arranged between DSP/DSS and researchers. The role of research actors – particularly as co-operation partners for DSP/DSS – is of great importance for the successful development of the European data ecosystem.

B. Designing practical requirements for data service providers – incentives are needed

The proposal for the DGA outlines the new role of a DSP - be it commercial or data altruistic - by fixing (minimum) requirements for companies that want to fulfil this function: For example registration, waiver of own efforts in data exploitation, use of certain standard contracts etc. (cf. Chapters III and IV DGA). Monitoring by national authorities is also provided for. The planned requirements describe rules that limit the field of action for DSP in terms of securing public welfare orientation. On the one hand, these limiting requirements create legal certainty

³ With regard to the FAIR principles see <https://www.go-fair.org/fair-principles>. Cf. Rfll (2020) – The Data Quality Challenge. Recommendations for Sustainable Research in the Digital Turn, p. 95ff., <http://www.rfii.de/?p=4203> (last accessed on 26.01.2021).

and are therefore to be welcomed. On the other hand, the current DGA draft lacks positive legal benchmarks for the development of economically viable business models for DSP/DSS. Only if DSP/DSS have real chances for a successful formation in the common market in a long-term perspective, they can consistently foster data sharing. Firstly, the RfII advocates that the DGA should also set incentives that facilitate potential providers to enter this market and promote competition for innovation. Innovation can only be generated on the basis of a verifiable, top notch database that meets the highest possible standards. Secondly, the DGA must be supplemented with more concrete specifications for quality assurance. This is the only way to establish reliable trust in the role of the new intermediaries in economy and society.

In general, the RfII suggests a specification of business models for intermediaries, which should be operated at least cost-covering, for example within the framework of an existing institution under public law, on a non-profit basis or also commercially profitable. Incentives should also be provided for the establishment of DSPs or DSSs. The faith of data holders or data producers in a DSS cannot be based solely on its neutrality. Rather, its sustainable data handling and its economic stability are also important issues. The role of the new intermediaries must therefore open up entrepreneurial room for manoeuvre. This applies equally to commercial and altruistic DSPs.

Since the purposes of collecting and using data are very diverse in economy and society, a Europe-wide DGA should not aim for a one-size-fits-all solution. Rather, a legal framework should be established that serves the needs of different sectors and actors. Hence, the RfII sees two levels of requirements for DSP/DSS:

All data, data providers and data services covered by the DGA should be subject to a generally binding legal framework (see below under 1.) which provides minimum standards for data-related trust protection (in the sense of the German General Data Protection Regulation [DSGVO]) and enables the establishment of sustainable business models. The legal requirements must be set to motivate a large number of services to enter the market. These requirements should also be easy to implement for small and medium-sized enterprises. On a second level (see below under 2.), DSP and DSS commit themselves to further obligations in order to demonstrate superior trustworthiness. For example, they can be certified as complying with particularly high standards in data security and data quality. Examples of such a "premium segment" are the accredited research data centers in certain scientific fields. With such a stage model, broad European competition can emerge depending on the type of data, consumer interest and intended use.

1. In order to enable the establishment of DSP in a sustainable form and on a fundamental level, the RfII recommends that the following specifications be taken into account as supplementary requirements for all services to be registered within the framework of the DGA:

- The European Commission and Member States are providing impetus to promote the framework conditions for the development of appropriate insurance solutions that cover the risks for DSPs. The regulation of liability issues via insurance models promotes public faith in the new market players. At the same time, solutions must be found here that can also be implemented pragmatically and cost-effectively for small and medium-

sized enterprises, making it as easy as possible for them to establish sustainable business models. In the medium term, the EU Member States are asked to examine the prerequisites and operating conditions for the possible introduction of compulsory insurance solutions for DSPs.

- In case of insolvency of a DSP as well as business closure and transfer of a service to another non-fiduciary type of company, the data shall be transferred to the public authorities. If the public authorities decide to delete the data, they are not liable to compensate damages to data holders/producers.
- If a DSS is operated by a public institution that relinquishes its role as a DSP, liability regulations for data loss (analogous to the corporate insolvency case) are limited.
- If a DSS is operated by a scientific institution for scientific purposes Article 11, No. 1 does not apply (the use of metadata may also go beyond the further development of the DSS).
- The national authorities will additionally be provided with an advisory function to enable providers of data sharing services to give assistance, particularly with regard to the clarification and interpretation of legal issues arising in the course of their activities as DSPs.
- In order to ensure a largely coherent/harmonised implementation and supervision of these requirements, the national authorities should be provided with guidelines. The European Data Innovation Board could take over this task of moderating such an agreement process and ensuring the formulation and application of such guidelines.
- Given the driving force of publicly funded science in building a European data ecosystem, representatives of academia (from different scientific communities) should be members in the European Data Innovation Board.

2. For DSPs aiming to establish their business model at the second superior level (see above), more extensive regulations must apply. This indicates to data holders/producers as well as data users that these intermediaries meet the highest possible standards of security and quality in terms of data sharing. For this category of DSP, the RfII recommends the following supplements to the DGA:

- The supervisory authorities have not only a control function, but also carry out certification duties, which the RfII considers essential for the question of quality assurance of data and services. This can also - and from the perspective of the science system: preferably - be developed within the framework of an accreditation procedure that is still to be developed. An according accreditation system should be pragmatically aligned with the requirements of the different business models of DSP.
- Providers of the second level are listed in a register that is publicly maintained. These registers can serve as an orientation for consumers and other user groups with regard to particularly trustworthy and high-quality services. This is an additional incentive for DSPs to seek certification.
- DSPs offer an extensive insurance coverage, e.g. for critical or particularly high-value data, which can serve both data providers/producers and data users as an orientation mark for the sensitivity of data and for extraordinary efforts in data security and quality control.

C. Clarifying the concept of data altruism

The proposal of a DGA introduces the concept of data altruism into European legislation on data services and data use - on the one hand to describe a motive for data subjects or primary data providers to pass on data to DSP/DSS, but on the other hand also as a characteristic of a DSP, as they can register as a “data altruism organisation recognised in the Union”.

The DGA draft describes the registration process as the central criterion for data altruism. The characteristic features that distinguish an organisation engaging in data altruism from other organisations are not particularly named. The wording of the law offers a large margin for interpretation. Accordingly, the meaning of “altruism” in the DGA must be clarified

- regarding the purpose for which the data is collected and made available for sharing (e.g. use for medical research, in sports or culture etc.),
- with respect to the way data are handled (e.g. high quality standards regarding green IT/climate neutrality of their data storage),
- corresponding to charitability in the sense of a non-profit business model.

The Rfll recommends a clarification also in this aspect. Data altruism should first be characterised as a concrete economic criterion (non-profit or refraining from profit interests, if necessary also on the basis of a company/corporate form intended for this purpose). Secondly, this clarification should ensure a clear distinction between commercial and “altruistic” DSP/DSSs. If the DGA intends to promote data altruism, then the benefits to society as a whole – that arise from this particular type of sharing and re-using data – should be described more precisely (e.g. orientation towards charity, public benefit, promotion of civil society innovations etc.).

Furthermore, Article 22 of the DGA draft should be enriched with clarifications for the (minimum) requirements for a consent form. A useful complement to the research clause called for above is to establish a transparent consent model to enable certain data-intensive and data-linking research procedures. This could be supported by technology solutions which should be standardized.

D. Quality of data and data intermediaries: Thinking beyond FAIR

The Rfll is aware of the challenges that come along with the creation of a European legal framework for the re-use of protected public sector data and laying the foundations for public faith in the development of DSP/DSS. The Council suggests that questions of the quality of data and data-related services should be considered more comprehensively in the DGA. Based on the recommendations given above, further quality assurance mechanisms should be envisaged. In particular, guidelines are needed for the quality control and classification of data that the DSP receives, prepares and makes available for (re-)users. Data providers/producers should be appropriately involved in this process. If, for example, research data are involved, the data-providing scientific institutions are primarily responsible for data quality. They must be involved in questions of input control, classification and preparation of the data by the DSP.

The Rfll has stressed in several of its publications, that a mere reference to the FAIR principles is not sufficient if real quality improvements on data should be achieved. Not only from the perspective of science or the scientific usability of data, it would make sense for the DSP/DSS

to at least include information on the quality of the data they provide. Even data that are not used for research purposes only have added value for innovation chains in economy and society when they provide information about their provenance, other contexts of their creation or "production" and their previous stations of transfer. Such information should be given at least within the attached metadata. Transparently designed communication between the DSP and the data providers/producers is also necessary at this interface.

With regard to responsible and transparent data governance in general and trustworthy handling of personal data in particular, the scientific system can contribute best practices as orientation marks for a DGA that aims to address the entire European society. Such examples for best practice can for instance be found in the data protection concepts applied and established in the field of medical research. In the current DGA draft, the RfII has so far missed explanations which, for example, contribute to ensure that the risk of re-identification in the provision of personal data is sufficiently taken into account and minimised by the future DSP/DSS. Only such precautions create the faith on which a flourishing market for data – able to promote public welfare and innovation - can be built. With regard to anonymisation techniques, the RfII supports plans to work towards standards that are valid throughout Europe.

The German Council for Scientific Information Infrastructures (RfII) was established by the Joint Science Conference (GWK) of the Federal Republic of Germany to provide advice on the enhancement of scientific information infrastructures and related topics of the digital turn in science and humanities. In putting forward its reflections on these issues, the RfII attaches great importance to a well-balanced consideration of the sometimes overlapping needs of science, public administration, and industry, as well as relating aspects of the international co-operation.

Imprint

Courtesy translation

German Council for Scientific Information Infrastructures (RfII)

Head Office

Papendiek 16, 37073 Göttingen, Germany

Phone +49 551 392 70 50

Email info@rfii.de

Web www.rfii.de

This work is licensed under a Creative Commons Attribution-No Derivatives 4.0 International License (CC BY-ND).

