European Commission

Study on the Open Data Directive, Data Governance and Data Act and their possible impact on research

Independent Expert Report

Wirtten by Mirelle van Eechoud

Research and Innovation

Study on the Open Data Directive, Data Governance and Data Act and their possible impact on research

European Commission
Directorate-General for Research and Innovation
Directorate A — ERA & Innovation
Unit A.4 — Open Science
Contact Alea López de San Román
Email RTD-OPEN-SCIENCE@*ec.europa.eu*
Alea.lopez-de-san-roman@*ec.europa.eu*
RTD-PUBLICATIONS@ec.europa.eu
European Commission
B-1049 Brussels

Manuscript completed in March 2022.

PDF ISBN 978-92-76-53631-4 doi: 10.2777/71619 KI-08-22-204-EN-N

Luxembourg: Publications Office of the European Union, 2022

EUROPEAN COMMISSION

Study on the Open Data Directive, Data Governance and Data Act and their possible impact on research

2022 Directorate-General for Research and Innovation

Table of Contents

Research conducted between December 2021 and March 2022 for the European Commission, Directorate-General for Research and Innovation (DG RTD), Open Science Unit. At the time of finalising the study, the legislative procedures for both the Data Governance Act and the Data Act were still ongoing.

Mireille van Eechoud is professor of Information Law at the Institute for Information Law (IViR) and director of the Amsterdam Graduate School of Law.

EXECUTIVE SUMMARY

This study analyses the possible impact of three major legislative instruments in the EU's general data strategy for the field of research, especially for research performing organisations and research funding organisations. **It does so against the background of the European Open Science policy pursued,** in which the development of a European Open Science Cloud (EOSC) is a major undertaking.

The *Open Data Directive* is a generic instrument that regulates re-use of data held by public sector bodies in Member States. Since 2019, it also covers research data. The Open Data Directive creates certain obligations for research performing organisations (in and outside universities) as suppliers of data, and contains rights for them as users of data.

The *Data Governance Act* final compromise text consists of three major building blocks. It seeks to expand principles of the Open Data Directive to a broader set of public sector data. In addition, it regulates data intermediation services by imposing requirements, including registration. It also sets up a registration scheme for organisations engaging in data altruism, i.e., the sharing of data for public interest purposes including academic research.

The proposed *Data Act* addresses disparate issues. It aims to promote the (re-)use of Internet of Things (IoT) data by empowering the users of connected products and services. Also, switching cloud and other data processing services should become easier once providers are subject to a series of obligations. Furthermore, data holders will be subject to conditions with respect to licensing terms, whether their obligation to provide access derives from the Data Act itself *or any other* Union legislation. A number of provisions are further aimed at the realisation of interoperability for data spaces (including EOSC). Also, public sector bodies get to obtain data from businesses in the event of public crises and other exceptional need situations. Researchers may benefit in modest ways from this new power.

Taken together, the instruments create a complex legal framework that will affect the development of eosc and its components in various ways. The impact is difficult to assess at this stage. Partly this has to do with the fact that the open data directive, the data governance act final compromise text and the proposed data act are generic instruments. They are not tailored to the specifics of publicly funded scientific research that takes place in a global hybrid environment, involving a complex web of public research performing organisations, support organisations and private providers of all manner of services. There will be formal and non-binding kinds of rulemaking and standard setting that will take place at the eu level in many areas, e.g., around standard contractual clauses for data sharing and use, requirements for data intermediation services and data altruism organisations, interoperability requirements for cloud services and data spaces. Potentially this can help realise eosc. Much will depend on how effectively the needs of research stakeholders are accommodated in these processes, including safeguards for academic freedom as protected under the eu charter of fundamental rights. Especially for research performing organisations, the open data directive and the data governance act final compromise text bring some legal uncertainty, notably because functions of public universities (education, research, library) are subjected to different rules.

1 INTRODUCTION

This introductory chapter sets the scene for a more in-depth discussion of three key pieces of (upcoming) legislation that affect public research data. **The central question addressed in this study[1] is: what is the impact on research performing organisations, research funding organisations, and research infrastructures** (together: 'research stakeholders') **of the Open Data Directive as well as the Data Governance Act final compromise text and the proposed Data Act?** The main characteristics of each instrument and their relevance to research are set out in **dedicated chapters** (2 through 4). This is done on the basis of desk research, consisting of legal texts including legislative preparatory materials, academic literature and policy documents. To properly contextualise the three instruments, **this chapter** explains how the EU's Open Science policy relates to the wider EU data strategy and digital principles, what roles research performing organisations fulfil in regards to data creation and sharing, and how the European Open Science Cloud (EOSC) is a cornerstone in the EU's policy agenda on open science. The **concluding chapter 5** synthesises findings and makes recommendations.

1.1 Open Science policy in relation to the EU-data strategy

The key instruments discussed in this study are part of a broader EU data strategy.[2] It is a generic strategy, aimed primarily at creating a single European market for data and (thereby) stimulating innovation and competition. Wider benefits extolled include improvements in health(care), the environment, and public service delivery. To date, important instruments in the realization of the data strategy are the General Data Protection Regulation (GDPR 2018), the Regulation on the free flow of non-personal data (FFD 2019), and **the Open Data Directive** (ODD 2019).[3] Two key upcoming regulations are **the Data Governance Act (cDGA)[4] and the Data Act** (pDA).[5] It is the latter three that **are the focus of this study**.

The legislative basis of the instruments discussed is Article 114 of the Treaty on the functioning of the EU (TFEU), which creates legislative competence for measures aimed at the establishment and functioning of the internal market. **On a normative level, the European Declaration on Digital Rights and Principles for the Digital Decade** (January 2022)[6] **sets out the public values that the EU institutions and Member States want to uphold** in the ongoing digital transformation of business, government and society at large. This political vision does not speak directly to research and the role of science in society. It does however provide a backdrop, by enumerating principles of digital

[1] Invaluable support was provided by A. Meiring LLM.

[2] European Commission. Directorate General for Communication, The European Data: Shaping Europe's Digital Future, Publications Office 2020.

[3] Regulation (EU) 2016/670, Regulation (EU) 2018/1807, Directive (EU) 2019/1024.

[4] Proposal for a Regulation of the European Parliament and of the Council on European data governance (Data Governance Act) - Analysis of the final compromise text in view to agreement, 10.12.2021 14606/21.

[5] 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), Brussels, 23.2.2022 COM(2022) 68 final.

[6] Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions – Establishing a European Declaration on Digital rights and principles for the Digital Decade, Brussels, 26.1.2022 COM(2022) 28 final.

sovereignty, privacy and individual control over data, inclusion, security, trust, and the importance of a "dynamic, resource-efficient and fair economy and society in the Union". Respect for fundamental rights is at its heart. Especially relevant to scientific research are freedom of expression, academic freedom, privacy and the protection of (intellectual) property. Specifically with respect to education fit for the digital age, the Declaration states the ambition to support "efforts to equip all education and training institutions with digital connectivity, infrastructure and tools", enable life-long learning and equip educators and learners with digital skills.[7]

The EU's activities in the field of research are regulated largely by Articles 179-189 TFEU. **The European Commission is tasked with establishing a European Research Area** (ERA) as a means to further the objective of strengthening the scientific and technological bases of the EU (Article 179 TFEU). **A key instrument is the funding of scientific research**; the main EU framework programme for R&I has a budget of approximately 95.5 billion euro. The concept of a European Research Area has been rolled out since 2000, with the ambition to create a single 'market' for research, and innovation and technology more broadly. **The core ambitions are expressed in the so-called Pact for Research and Innovation**, and include ensuring integrity of research and innovation, academic freedom, the free circulation of knowledge and value creation (also for societal and economic) impact. **The ERA-policy agenda for 2022-2024 sets out 20 actions** to promote a single market, some of which are especially relevant from the perspective of research data: to enable open science, including through the European Open Science Cloud (EOSC), and to ensure an EU-copyright and data legislative framework fit for research. Also, the protection of academic freedom is a specific action point; Article 13 of the Charter of Fundamental Rights of the EU (CFREU) guarantees that "arts and scientific research shall be free of constraint" and that "academic freedom shall be respected".

The EU's Open Science policy has eight pillars or themes, ranging from optimising research integrity and reproducibility of scientific results to promoting citizen science. **From the perspective of research** cycles – the circular process from research design through data collection, analysis, interpretation, and reporting of outcomes, on which other research can build – **the following pillars are particularly relevant:[8]**

- FAIR (Findable, Accessible, Interoperable and Re-usable) and open data sharing should become the default for the results of EU-funded scientific research;
- All publicly funded research in the EU should adhere to commonly agreed standards of research integrity;
- The general public should be able to make significant contributions and be recognised as valid European science knowledge producers;
- Develop the European Open Science Cloud (EOSC): "a trusted, virtual, federated environment that cuts across borders and scientific disciplines to store, share, process and reuse research digital objects (like publications, data, and software)."

The concept of FAIR and open data sharing also features in the Horizon framework programme. **FAIR represents the four key characteristics of well-managed research data which enable discovery, evaluation and downstream reuse:**

[7] Ibid., p. 3.

[8] See https://ec.europa.eu/info/research-and-innovation/strategy/strategy-2020-2024/our-digital-future/open-science_en#8- ambitions-of-the-eus-open-science-policy.

data must be Findable, Accessible, Interoperable and Reusable.[9] Building on earlier initiatives across academic disciplines, a set of eleven detailed FAIR guiding principles were drafted by a diverse group of academics, institutions and private sector stakeholders (such as publishers). The principles focus on (technical) qualities of (meta)datasets, resources (e.g., repositories), communication protocols and tools. Rich metadata, the use of persistent identifiers, use of open and standardised communication protocols which allow for authentication and authorisation procedures are some of these qualities. As shall be analysed in the next chapter, there are similarities with some of the principles laid down in the Open Data Directive. **Overall, however, FAIR is about the technical aspects of sharing data, not much about the legal aspects.** The only guiding legal principle of FAIR comes under the R: "(meta)data are released with a clear and accessible data usage license" (R1.1).[10]

Of separate lineage is the concept of open data, which takes its inspiration from multiple sources, including open source software principles and open content licensing schemes, notably Creative Commons (2002). In its broadest form the concept refers to data that "anyone can freely access, use, modify, and share for any purpose (subject, at most, to requirements that preserve provenance and openness)".[11] To enable this, not only must data be legally open – that is, unencumbered especially by intellectual property restrictions – it must also be technically open – that is, machine-readable and available in an open format. The open data- concept is strongly associated with the rise of public sector information policies promoting reuse. Since the Public Sector Information (PSI) Directive's revision of 2013, the principle that whenever possible, public sector data must be technically and legally open has become more visible in the Open Data Directive (see next chapter). Although the concept of open data does not prescribe the use of a specific type of license or declaration, Creative Commons have become popular in the field of research, partly as a result of their promotion by the European Commission.[12] Where Member States develop specific standard licenses, compatibility must be strived for.

Yet another concept relevant to the circulation of research data is FRAND. The acronym stands for Fair, Reasonable and Non-Discriminatory terms. It is best known from the (tele)communications field where standards to ensure

---

[9] These elements have been identified by various authors and institutions in various disciplines, but as an acronym and coherent set, FAIR has gotten particular traction following Wilkinson, M., Dumontier, M., Aalbersberg, I. et al. The FAIR Guiding Principles for scientific data management and stewardship. Sci Data 3, 160018 (2016), https://doi.org/10.1038/sdata.2016.18.

[10] Mark D Wilkinson and others, 'The FAIR Guiding Principles for Scientific Data Management and Stewardship' (2016) 3 Scientific Data 160018.

[11] https://opendefinition.org/.

[12] The European Commission recommends the use of CC in the context of the Open Data Directive (Commission Notice. Guidelines on recommended standard licences, datasets and charging for the reuse of documents, 24.07.2014 2014/C 240/01) and also uses CC for its own documents (Commission Decision adopting Creative Commons as an open licence under the European Commission's reuse policy, Brussels, 22.2.2019 C(2019)1655 final). See also: EC Reuse policy: a study on available reuse implementing instruments and licensing considerations, European Commission's Central IP Service, EUR 29685 EN, Publications Office of the European Union, Luxembourg, 2019, doi:10.2760/95373, JRC115947. Creative Commons licenses or equivalent licenses are also required to Horizon Europe beneficiaries under the Grant Agreements obligations on Open Science, see pages 110-112 of the Horizon Europe Model Grant Agreement https://ec.europa.eu/info/funding-tenders/opportunities/docs/2021-2027/common/agr-contr/general-mga_horizon-euratom_en.pdf .

interoperability of systems and devices are needed. FRAND is used as shorthand for requirements that licenses must meet in situations where holders of information or technology (protected by patents or other intellectual property) are required to share it. Parties collaborating in standard-setting organisations such as ISO and ETSI (telecommunications standardisation body) commit to licensing intellectual property that is essential to a given standard under FRAND terms. It is for the holder of intellectual property to agree to the specific terms including royalties.[13] In that sense, FRAND terms are not standardised licensing schemes the way Creative Commons licenses are. In the context of the proposed Digital Markets Act, FRAND obligations are set to be imposed on large platforms with respect to certain services where they are gatekeepers (e.g., app stores). The concept also features in the proposed Data Act.

1.2 Generic instruments relevant to open science

The landscape of laws affecting open science is wide and complicated to navigate. In large part, this is caused by the fact that **many instruments are not primarily drafted with universities, other research performing organisations and academic (support) institutions in mind.**

The Open Data Directive is the generic instrument that regulates re-use of data held by public sector bodies in Member States. First legislated in 2003, the current 2019 version includes for the first time a provision on research data. The Open Data Directive creates certain obligations for research performing organisations (in and outside universities) as suppliers of data, and rights for them as users of data.

The Data Governance Act final compromise text is expected to come into effect in the form of a regulation in the course of 2022. It consists of three main building blocks and some additional provisions. On the one hand, it seeks to expand principles of the Open Data Directive to a broader set of public sector data, without however obliging Member States to allow such re-use (Chapter II). On the other hand, it regulates data intermediation services (Chapter III) and sets up a registration scheme for organisations engaging in data altruism, i.e., the sharing of data for public interest purposes including academic research (Chapter IV).

The proposed Data Act sets out generic horizontal rules for the access and use of data generated by connected products and their related services. It aims to promote the (re)use of Internet of Things (IoT) data by empowering the users of connected products and services. Users have the right to access and share data generated by 'their' connected products and related services. To facilitate sharing intellectual property in IoT data will be limited. Furthermore, providers of cloud services and other data processing services will become subject to a series of obligations which should make enable users to switch more easily between service providers, including adherence to interoperability mechanisms. To facilitate data

[13] Méniére Y. 'Fair, Reasonable and Non-Discriminatory (FRAND) Licensing Terms. Research Analysis of a Controversial Concept', EUR 27333, Luxembourg: Publications Office of the European Union 2015, JRC96258. Note that SSO policies can request that its members do not claim copyright in standards documentation (e.g. on text or graphics), so as to enable SSO- members to be able to make copies of standards documentation. ETSI itself claims the copyright in standards agreed and accompanying technical documentation. See e.g. ETSI Guide on Intellectual Property Rights (IPRs), Version adopted by Board#133 (10 June 2021).

sharing, the European Commission will be tasked with developing (non-binding) standard contract clauses. More importantly, data holders will be subject to conditions with respect to licensing terms, regardless of whether their obligation to provide access derives from the Data Act itself or any other Union legislation, including implementing laws. A number of provisions are aimed at the realisation of interoperability and key requirements for data spaces (including EOSC). Another element of the proposed Data Act – albeit subject to strict limitations – involves the power of public sector bodies to obtain data from the private sector (businesses) in the event of public crises and other exceptional need situations, which could then also be shared with researchers. Finally, the Data Act proposes duties to keep data (services) secure and to prevent (international) parties from gaining unlawful access to data.

The existing and new instruments mentioned above are all intended to contribute to the realisation of a European data space. However, it is not always easy to establish how the instruments relate and which rules take precedence in case of contradiction. The GDPR – with its limitations on who can process personal data for which purposes, rights of access and portability for data subjects, and an intricate system of rules for the transfer of data outside the EU/EEA – has priority over all three. This means, for example, that where a dataset contains personal and non-personal data, its use will have to conform to the GDPR. The Data Governance Act, on the other hand, will take a back seat to sector specific regulation, as well as to the Regulation on the free flow of non-personal data. The latter restricts Member States' ability to impose geo-localisation rules, e.g., by demanding that certain cloud services are based in a specific territory.

1.3 Roles and legal challenges for research organisations

Research performing organisations (also: "RPOs") **are the key players in the implementation of the EU's Open Science policy and research infrastructures.** Since the impact of the instruments analysed here is probably greatest for them, **the focus of this study is mainly on RPOs.** Where relevant, however, the positions of research funding organisations (also: "RFOs") and of the broader EOSC community are discussed as well. RFOs are a major driver of open science, not only because they contribute to the funding of infrastructures, but also because they can and do impose conditions regarding data sharing on researchers, for instance the use of specific open licences for research outputs generated with their funding.

Unique to research performing organisations is that they typically wear multiple hats. On the one hand, they are sources of data collected or produced during research activities, and on the other hand, they are users of data that could serve as input for those activities. **With these different roles come different challenges and needs.** As data users, RPOs have an interest in a clear legal framework regulating access to data held by other entities. As data holders, however, RPOs may have more of an interest in reducing the administrative efforts that come with making research data publicly accessible and/or available for re-use and complying with all sorts of laws (e.g., data protection, intellectual property, contracts, export controls), in securing public private partnerships, and in maintaining autonomy vis-à-vis data processing service providers. **The next chapters will use these two roles as lenses to analyse the impact of key EU data legislation on RPOs.**

Importantly, **the concepts of 'research' and 'research performing organisations' are not defined in the Open Data Directive, the Data Governance Act final compromise text or the proposed Data Act.**[14] The Open Data Directive, however, does contain a definition of 'research data'. Research data are understood as "documents in digital form, other than scientific publications, which are collected or produced in the course of scientific research activities and are used as evidence in the research process or are commonly accepted in the research community as necessary to validate research findings and results" (article 2(9) ODD). According to recital 27 of the Directive, statistics, results of experiments, measurements, observations resulting from fieldwork, survey results, interview recordings, images, meta-data, and other digital objects that are not scientific articles are considered research data. Taking the definition of research data as a starting point, the present study considers RPOs as "legal entities, public or private, in which individuals carry out scientific research activities that involve the collection and/or production of research data within the meaning of article 2(9) of the Open Data Directive."[15] Commercial parties performing research aimed at the development, enhancement or optimisation of products and services therefore fall outside the scope of this report.[16] Still, according to this broad interpretation, many organisations may qualify as RPOs, such as research units in universities (as opposed to educational branches), independent research centres and state research institutes. However, as will be seen in chapters 2-4, **the current legal framework leaves these organisations somewhat in the dark as to their rights and obligations due to a lack of clear definitions.** For universities this is particularly true, since their activities – education, research and maintaining university libraries – fall in different categories distinguished in the legal framework, i.e., 'educational establishments', 'research performing organisations' and 'libraries', to which different rules apply.

The objective of this study is to analyse whether the aforementioned regulatory instruments adequately respond to the (legal) challenges and needs as experienced by research stakeholders. In the context of open science, individual researchers and RPOs have voiced various concerns stemming from their role as data users as well as their position as data holders.

The main issue from a data user-perspective is the need for broader access rights to data (sources) from the public sector but also from the private sector, as data held by big technology companies are crucial for the advancement of academic research in many disciplines (e.g., media studies,

---

[14] Recital 20 of the Data Governance Act final compromise text does also provide a definition of 'scientific research purposes': "any type of research related purpose regardless of the organisational or financial structure of the research institution in question, with the exception of research that is being conducted by a company aiming at the development, enhancement or optimisation of products". However, this interpretation is explicitly limited to "the specific context" of charging fees for re-use of certain categories of protected data (see chapter 3).

[15] Compare the definition provided by the European Science Foundation (ESF), Evaluation in Research and Research Funding Organisations – European Practices, November 2012, p. 23: "An institute or other organisation which is itself realising research and which employs active researchers". See also the definition provided by the European Research Infrastructures in the International Landscape (RISCAPE) project, International Research Infrastructure Landscape 2019 – A European Perspective, 2019, p. 90, 94-97 (Appendix 2 – RISCAPE Questionnaire): "Organisation making research activities and products directly by their staff. Typical examples are universities and research centres."

[16] In line with the definition of 'scientific research purposes' set forth in recital 20 of the Data Governance Act.

behavioural studies, law, economics). Researchers would thus benefit from a clear legal framework through which they can obtain access to public and private data.

The challenges and needs associated with the data holder-position, on the other hand, are rather heterogeneous. First, it appears that researchers have a strong sense of 'ownership' towards their research data and also value trust in data sharing[17] (for example, universities want to prevent commercial parties to build a service based on research data for the use of which universities later have to pay). **Second**, researchers have criticised the lack of a clear and specific legal framework for opening datasets, particularly in countries with regional governments that each have different open data policies and regulations.[18] **Third**, the lack of (guidance on) common standards regarding structures, formats, languages, security, etc. of data has been identified as an obstacle for sharing.[19] **Fourth**, researchers have indicated that they do not always know how and under what conditions they may share confidential data, such as business-critical data or privacy-sensitive data.[20] **Finally,** researchers report that they experience difficulties in choosing appropriate data licences and in assessing which licence applies when multiple datasets are combined.[21] In fact, licensing systems are thought to be inconsistent and not universally oriented towards favouring re-use.[22]

1.4 European Open Science Cloud (EOSC)

As mentioned in paragraph 1.1, **one of the substantial efforts made by the EU to enable open science is the creation of an EU-wide platform for the exchange of publicly funded research data across the scientific community: the European Open Science Cloud** (EOSC). The EOSC is a long-term initiative of the European Commission in collaboration with the European research community to develop what was initially termed "a virtual environment with free at the point of use, open and seamless services for storage, management, analysis and re-use of research data, across borders and scientific disciplines."[23] In a nutshell, the EOSC aims to facilitate access to, FAIR management of, and reliable re-use of research data and other digital research

---

[17] T. T. Chan, I. Meijer, The Netherlands' Plan on Open Science (Open Science Monitor Case Study), October 2018, p. 10; M. streefkerk, 'onderzoeksdata delen onder eigen voorwaarden', Amsterdam Economic Board, 17 June 2021, https://amsterdameconomicboard.com/nieuws/onderzoeksdata-delen-onder-eigen-voorwaarden/.

[18] See e.g., J. Berends e.a., Analytical Report n5, Luxembourg: Publications Office 2020, p. 16-17; R. Morais and L. Borell- Damián, 2017-2018 EUA Open Access Survey Results, April 2019.

[19] See e.g., PwC EU Services, Cost of not having FAIR research data: Cost-Benefit analysis for FAIR research data, 2019, p. 6-7; OECD, Recommendation of the Council concerning Access to Research Data from Public Funding, OECD/LEGAL/0347, 2021, para. IV, IX; KNAW, Storage and availability of research: From intentions to implementation, 2021, p. 6.

[20] See e.g., Data Sharing Coalition, 'SURF: sharing (confidential) data for the benefits of research', 16 September 2020; European Commission, Horizon 2020 Monitoring Report, 2017, p. 72.

[21] J. Berends e.a., Analytical Report n5, Luxembourg: Publications Office 2020, p. 16-17.

[22] Deloitte, Study to support the review of Directive 2003/98/EC on the re-use of public sector information, 2017/006, 2018, p. 259; J. Berends e.a., Analytical Report n5, Luxembourg: Publications Office 2020, p. 16-17.

[23] European Commission, COM(2016) 178 final, p. 6.

objects (e.g., methods, software and publications).[24]

One of the earliest references to the creation of a "research open science cloud" can be found in the 2015 Commission's Communication on shaping Europe's digital future.[25] Since then, the project has grown exponentially both in terms of finances and stakeholders involved. Under Horizon 2020, the EOSC was granted €600 million[26] to support its advancement, and the financial support has been continued under Horizon Europe 2021-2027. **The many stakeholders the EOSC has attracted over the years are represented in two entities that, together with the European Commission ('tripartite governance'), support the implementation of the EOSC ecosystem in Europe:[27]**

- **The EOSC Association** – a legal entity (AISBL under Belgian law) set up in July 2020 with over 200 members and observers from the wider EOSC stakeholder community.[28] Among the members and observers are (associations of) universities and research institutes/laboratories, research funding organisations,[29] [30] service providers for research (including commercial parties such as publishers and IT-companies that are university spin-offs) 30 and what the EOSC Association classifies as 'other organisations' connected to research;[31]
- **The EOSC Steering Board** – an expert group set up in 2021 consisting of representatives of Member States and Associated Countries to Horizon Europe.

Membership of the EOSC Association is very diverse. Associations may have members that are direct members of the EOSC Association themselves.[32] Sometimes associations both have public and private partners.[33] The members'

---

[24] European Commission, European Open Science Cloud (EOSC): What the cloud is, how it was developed and being implemented, https://ec.europa.eu/info/research-and-innovation/strategy/strategy-2020-2024/our-digital-future/open- science/european-open-science-cloud-eosc_en; ssee also the general and specific objectives as formulated in the Memorandum of Understanding of the Co-programmed European Partnership on the European Open Science Cloud, available at: https://ec.europa.eu/info/sites/default/files/research_and_innovation/funding/documents/c_2021_4113_f1_annex_en_v3_p1_1213802.pdf, p. 3-4.

[25] European Commission, COM(2015) 192 final, p. 15.

[26] https://digital-strategy.ec.europa.eu/en/policies/open-science-cloud

[27] https://ec.europa.eu/info/research-and-innovation/strategy/strategy-2020-2024/our-digital-future/open-science/european- open-science-cloud-eosc_en

[28] https://ec.europa.eu/info/research-and-innovation/strategy/strategy-2020-2024/our-digital-future/open-science/european- open-science-cloud-eosc_en; for an overview of all the members and observers see https://eosc.eu/members

[29] E.g., Deutsche Forschungsgemeinschaft and Science Europe.

[30] E.g. Elsevier (large publisher, increasingly becoming a data analytics company), Premotec GmbH (a spin-off from EHT Zürich). Many members are non-profit organisations/partnerships such as key player OpenAire (initially created to share research outputs resulting from the EU's Framework Programmes) and LERU (association of research-intensive universities).

[31] E.g., LIBER (Association of European Research Libraries).

[32] Elsevier B.V., for example, is a member of EOSC but is also a member of euroCRIS (which in turn is a direct member of EOSC). Interesting in this regard is also that a separate service of Elsevier, Elsevier-PURE (a research information management system), is a sponsor of euroCRIS. See: https://eurocris.org/community/members

[33] For example the Dutch Techcentre for Life sciences (DTL): https://www.dtls.nl/community/partners/list/

interests in and potential contribution to EOSC thus differ according to their legal form and nature of activities. **To what extent the EOSC Association and its members (will) qualify as addressees of the Open Data Directive, the Data Governance and/or Data Act must be assessed on a case-by-case basis.** For instance, the EOSC Association as such is not considered a provider of 'data intermediation service' within the meaning of the Data Governance Act. However, some individual entities active in the virtual environment may in fact qualify as offering data intermediation services and be subject to the corresponding legal obligations (see chapter 3).

The first step of the EOSC is to federate existing scientific data infrastructures that are scattered across disciplines and across Member States[34] in order to connect digital research objects deposited in domain-specific or institutional repositories within a pan-European governance structure. From the perspective of the Open Data Directive, the EOSC offers an organisational and technical infrastructure for making increased amounts of re-usable research data effectively available for re-use beyond RPO's public tasks.

Having the technical infrastructure for sharing is one thing, but to what extent research data will actually be shared also depends on the (legal) conditions under which access and use are allowed. Licensing is therefore an important aspect. If incompatible licensing models are used, it is difficult to combine data from different sources. License 'interoperability' is a matter of concern, which is reflected in the recommendation of the EOSC Interoperability Task Force (2021) to draft a list of EOSC-recommended standard licences accompanied with information about the compatibility of these licenses with Member States' recommended licences.[35] This effort also connects well to the ambitions of the Open Data Directive to encourage the use of standard open licences for the re-use of public data. However, a number of Member States have developed dedicated licenses for public sector data.

The EOSC will not operate in isolation. In the European data strategy of 2020, **the European Commission recognised the EOSC as the basis for a "science, research and innovation data space" to complement and be connected with the nine so-called 'Common European Data Spaces'.[36]** Each of these sectoral data spaces are currently being set up to overcome legal and technical barriers to data sharing across organisations.[37] The next step in unlocking the potential of data would be to eliminate barriers between domain-specific data spaces. The proposal of the Data Act aims to contribute to this objective by providing a legal framework for the development of interoperability standards for data sharing across European data spaces – including the EOSC – and between other data that are not within the scope of a Common European Data Space (see chapter 4).[38]

---

[34] European Commission, COM(2016) 178 final, p. 7.

[35] European Commission, EOSC Interoperability Framework: Report from the EOSC Executive Board Working Groups (WG) FAIR and Architecture, February 2021.

[36] European Commission, COM(2020) 66 final, p. 33. The other data spaces are related to the following areas: industry (manufacturing), Green Deal (climate change/circular economy/biodiversity), mobility, health, finance, energy, agriculture, public administration and skills (education and training for the labour market).

[37] Ibid., p. 16.

[38] Explanatory Memorandum to the 2022 Commission Proposal, p. 3.

2. OPEN DATA DIRECTIVE

2.1 Key aspects

The current Open Data Directive ('ODD') entered into force on 16 July 2019. It is a continuation of the Directive on the re-use of public sector information ('PSI Directive' 2003/98/EC, amended by 2013/37/EU). **Member States have been slow implementing it.** The vast majority missed the July 2021 deadline and the Commission had to announce infringement proceedings.[39] By March 2022, eight Member States had still not taken transposition measures. Not much material on the implementation of the new article 10 on research data was at hand during the drafting of this study.

The Open Data Directive's main objective is to maximise the re-use of public data to further stimulate digital innovation in products and services, and thus to maximise social and economic benefits within the Union.[40] To that end, the Directive has expanded its scope from traditional public sector information – i.e., documents produced and collected by public sector bodies – to also cover documents held by certain public undertakings and data resulting from publicly funded research (together: 'public data'). The notion of 'document' is central to the ODD, and covers all manner of recorded information/data. Compared to its forebears, the Directive promotes the concept of 'open data' (as explained in chapter 1) more forcefully.[41]

The original PSI Directive left Member States much discretion as to which information they make available for re-use, how and on what terms. The rules were tightened in 2013.[42] A subsequent evaluation showed that a number of issues still had to be addressed in order to better exploit the potential of public sector information for the economy and society.[43] The European Commission's proposal[44] of April 2018 was adopted without much controversy. **The provisions of the 2019 Directive reflect the following key principles:**

- **Re-use obligations apply to public data in a broad sense.** The ODD covers documents held by all manner of public sector bodies, as well as documents held by specified public undertakings and certain research data (article 1(1)). There are some exceptions; documents in which third parties own intellectual property, for instance, are not subject to re-use rules (article 1(2)(c)), nor are personal data where data protection and privacy laws limit access and use (articles 1(2)h, (4)).
- **Re-use obligations only apply to documents that are already publicly accessible**. The Directive builds on, and is without prejudice to, Union or national public access regimes, e.g., laws on public registers or freedom of information acts. Article 1(2) also lists a number of exemptions for documents to which access is commonly excluded or restricted by virtue of (national) access regimes.
- **Documents are reusable by default.** All documents falling within the Directive's scope must be made available for re-use (article 3); only in

[39] European Commission Press Corner, Daily News 30/09/2021, https://ec.europa.eu/commission/presscorner/detail/en/mex_21_4962.
[40] See recitals 3-4 and article 1(1).
[41] See recitals 16-18 ("open by design and by default").
[42] European Commission, Explanatory Memorandum on the Proposal for a Directive of the European Parliament and of the Council on the re-use of public sector information (recast), COM(2018) 234 final, p. 1.
[43] Ibid., p. 5; see also Recital 2-4 of the Open Data Directive.
[44] Ibid.

limited cases this is not mandatory but optional (for example, for university libraries).

- **Conditions for re-use are non-discriminatory for comparable categories of re-use**. Any conditions for re-use must be the same for comparable categories of re-use and reusers. Different policies are allowed for e.g., commercial and non-commercial re-use (article 11, recital 46).
- **Re-use is open to all potential actors in the market (non-exclusivity).** The grant of exclusive rights is not allowed, except in cases where it is necessary for the provision of a service in the public interest (article 12) and only for a limited period.
- **Information is preferably made available as open data.** Various provisions signal that 'technical openness' and 'legal openness'[45] of data is preferred. Technical openness is mostly reflected in articles 5, 6 and 7 of the ODD. Article 5(1) requires that wherever possible, public data are made available in formats that are open, machine-readable and compliant with formal open standards. Article 5(2) further encourages public sector bodies and public undertakings to act in accordance with the principle of 'open by design and by default' when facilitating access and re-use. Article 6 determines that the re-use of public data should as a rule be free of charge. Where charges are however necessary, they should be limited to the marginal costs incurred for the reproduction, provision and dissemination and/or anonymisation of documents. Legal openness, on the other hand, is mostly reflected in article 8(1). It stipulates that the re-use of documents should in principle not be subject to conditions, unless such conditions are objective, proportionate, non-discriminatory and justified on grounds of a public interest objective. Any permissible licence for the re-use of public sector information should place as few restrictions on re-use as possible, for instance by limiting restrictions to an indication of the source (attribution). Recital 44 emphasises that open licences "should play an important role" in this respect. Another clause referring to legal openness is the prohibition for public sector bodies to exercise their sui generis database right with the purpose of preventing the re-use of documents or restricting reuse beyond the limits set by the Directive (article 1(6)).[46] Although the Open Data Directive does not explicitly prohibit public sector bodies to exercise any copyright to prevent or control re-use, it can be inferred from the obligation to allow re-use by default (article 3(1)) that their room for manoeuvre is also limited in this regard.[47]
- **Practical arrangements are made to facilitate effective processing of requests for re-use.** Requests are handled through electronic means and as soon as possible (in principle within 20 working days), with adequate information about means of redress against a decision to refuse re-use or against terms and conditions (article 4).
- **High-value datasets will be identified and made available free of charge in machine-readable form via API's or as bulk download (article 14).** The Commission is empowered to specify for a number of enumerated themes (e.g., statistics and earth observation data) and other themes, which data sets will be made available, and how.

[45] Open Data Handbook, https://opendatahandbook.org/glossary/en/terms/open-data/

[46] See also European Commission, SWD(2018) 45 final, p. 38.

[47] See also M. van Eechoud, 'A Serpent Eating Its Tail: The Database Directive Meets the Open Data Directive', ICC 2021, No. 52, p. 376.

2.2 Scope of application to research data

Article 10 of the current Open Data Directive contains specific rules for the re-use of research data. Member States must develop national open access policies which must be compatible with the FAIR-principles (article 10(1) ODD) and ensure that such research data becomes available for re-use (article 10(2) ODD). In the Directive, FAIR is a concept used explicitly in the context of research data, but not for other public sector data. Many general provisions, however, fit well with the key elements of FAIR. Public data must be made available, together with their metadata, in formats that are open, machine-readable, accessible, findable, re-usable, and compliant with formal open standards (Article 5(1)). The licensing requirements of article 8 also resonate with the guiding principle R.1.1: '(Meta)data are released with a clear and accessible data usage license'.[48] Article 9 further contains a requirement for Member States to make practical arrangements facilitating the search for documents, such as assets lists of main documents with relevant metadata.

Universities and other educational establishments, including their libraries, are often organised as public sector bodies. **To understand the scope of application of the re-use obligations to research data, it is important to see how the position of universities and other research performing organisations** (RPOs) **has developed under the Open Data Directive.**

2.2.1 PSI Directive 2003: exclusion of education and research

The original PSI Directive of 2003 followed what could be referred to as an institutional/task- oriented approach to determine whether information was subject to the re-use regime, that is, by looking at the institution's legal status and the nature of its tasks: documents had to be held by public sector bodies and be produced in the performance of public tasks.[49] **By contrast, the current Open Data Directive also considers the nature of the data and how the data are funded** (data type/funding-oriented approach).

The definition of public sector body (PSB) has always been broad, covering states, regional and local authorities and bodies governed by public law. **Many publicly funded educational and research organisations therefore qualify as PSB. However, documents held by public universities and the like were explicitly excluded from the PSI Directive 2003.** The exclusion applied to "documents held by educational and research establishments, such as schools, universities, archives, libraries and research facilities including, where relevant, organisations established for the transfer of research results" (article 1(2)(e)). The exemption was justified for various reasons: the benefits for re-users would not outweigh the administrative burden of these specific public sector bodies; as carriers of culture and knowledge they had a special function in society; and much of the information would anyhow fall outside the scope of the Directive because of

[48] See the specification of FAIR criteria in: M.D. Wilkinson e.a., 'Comment: The FAIR Guiding Principles for scientific data management and stewardship', Nature.com 2016.
[49] That is, without having to prove a particular interest under the access regime in order to obtain access, see article 1(3) old.

third-party copyrights.[50] **Documents from** public service broadcasting organisations and **cultural establishments were excluded too.** The latter category in the context of the Directive includes theatres, art houses, museums, libraries and archives.

2.2.2 PSI Directive 2013: inclusion of university libraries

With the 2013 revision of the PSI Directive, university libraries were included in its scope (article 1(e) old jo. article 1(f) old jo. article 3(2) old). Cultural establishments remained excluded, except for publicly funded museums, archives and libraries. The idea was that the collections held by these institutions would be of great value for re-use (recital 18 old). The EU legislator recognised that resources held by libraries are often protected by intellectual property rights, either owned by third parties or by the (university) libraries or museums or archives themselves. Documents in which third parties owned (copyright and related rights, database rights) remained exempt. **Materials in which public sector (university) libraries, museums and archives themselves owned rights, or in which no intellectual property existed, were however subjected to the re-use rules.**[51]. In 2013 **it became mandatory for public sector bodies to allow re-use** (article 3(1)), **but for libraries, museums and archives allowing re-use was optional**, not an obligation (article 3(2) old).

2.2.3 Open Data Directive: inclusion of research data

During the review process of the PSI Directive in 2017-2018, it became clear that the general (legislative) opinion on the exemption of certain public data had further changed over time. Based on evaluation reports and stakeholder consultations, the Commission concluded that more public data should become widely available, including high-value public data generated by public undertakings and data resulting from publicly funded research.[52] This was not only

---

[50] European Commission, Explanatory Memorandum on the Proposal for a Directive of the European Parliament and of the Council on the re-use and commercial exploitation of public sector documents, COM(2002) 207 final - COD(2002) 123. Given the fact that educational establishments and research establishments were both excluded from the Directive by the same provision, it can be assumed that the aforementioned arguments in favour of the exemption for public sector bodies "in the cultural and educational area" also applied to research establishments and research data (research area).

[51] Recital 9 of the 2013 amending Directive seems contradictory to article 3(2), stating that "if a third party was the initial owner of the intellectual property rights for a document held by libraries, including university libraries, museums and archives and the term of protection of those rights has not expired, that document should, for the purpose of this Directive, be considered as a document for which third parties hold intellectual property rights [emphasis added]." The recital could be interpreted as excluding from the scope of the Directive any document in the possession of a library, museum, or archive for which a third party originally held intellectual property rights that would still be valid today. According to P. Keller e.a., however, this reading contradicts with article 3 as well the legislative history and overall objectives and principles enshrined in the Directive to open up public knowledge for re-use. See: P. Keller e.a., 'Re-use of public sector information in cultural heritage institutions', International Free and Open Source Software Law Review, 2014, Vol. 6, No. 1.

[52] European Commission, Explanatory Memorandum on the Proposal for a Directive of the European Parliament and of the Council on the re-use of public sector information (recast), COM(2018) 234 final, p. 5-7; see also recitals 4 and 27-28 of the Open Data

motivated by economic arguments but also by the normative argument that it is right to repurpose data for other uses because they have been funded with public (taxpayers') money.[53] Moreover, the idea of open science – open access of publications, FAIR sharing of research data – had gained traction in academia and received political recognition internationally.[54]

In light of these developments, **it was decided to include research data within the scope of the Open Data Directive** (article 1(1)(c) jo. article 10(2)), regardless of whether they are held by public or private entities. **The previous exemption of the PSI Directive 2013 for educational and research establishments was split into two separate exemptions.** Documents held by educational establishments (article 1(2)(k)) are still exempt, except where it concerns research data held by institutes of higher education (education above secondary level). Documents from RPOs, RFOs and research transfer organisations are still exempt, but again, not research data held by them (article 1(2)(l)). To give some examples: research data generated by universities fall within the Directive's scope, but their educational materials do not; and research data generated by RPOs are subject re-use rules, but administrative data on operational activities (e.g., budgeting) are not.

Significantly, the re-use obligations do not apply to all research data. Article 10(2) stipulates that research data shall be re-usable "insofar as they are publicly funded[55] and researchers, research performing organisations or research funding organisations have already made them publicly available through an institutional or subject-based data repository"[56] (for instance, because publication is mandated by national law or individual grant agreements).[57] As will be discussed below, there are many types of repositories and other publication modes for research data, of which repositories operated by a public university – or universities jointly – are just one example. Where exactly research data have been made publicly available does not seem to matter; it could have been in a repository operated (indirectly) by academic publishers or by other institutions that are not public sector bodies subject to the Open Data Directive. It does seem however, that in case of publication of research data outside of repositories, there must be some legal basis in Member State law to bring such research data within scope of article 10 (see recital 28).

Directive.

[53] See also recital 5 of the Data Governance Act final compromise text.

[54] Demonstrated by e.g., the OECD Declaration on Access to Research Data from Public Funding (2007); the European Commission Recommendation on access to and preservation of scientific information (2012); the draft of the foundations of the FAIR principles in 2014 (see M.D. Wilkinson e.a., 'Comment: The FAIR Guiding Principles for scientific data management and stewardship', Nature.com 2016, p. 3); and the Open Research Data Pilot initiative in Horizon 2020 (Article 29.3 of the Horizon 2020 Annotated Model Grant Agreement).

[55] Or, according to recital 28, if they follow from research that is co-funded by public and private-sector entities.

[56] Or, if explicitly allowed by the Member State in national law, when they have been made available by other data infrastructures than repositories, through open access publications, as an attached file to an article, a data paper or a paper in a data journal (recital 28).

[57] European Commission, Explanatory Memorandum on the Proposal for a Directive of the European Parliament and of the Council on the re-use of public sector information (recast), COM(2018) 234 final, p. 4, 10.

Another result of the review of the PSI Directive was the introduction of a legal obligation for Member States to support the availability of research data by adopting national open access policies. These policies are to be addressed to research performing organisations (RPOs) and research funding organisations (RFOs) (article 10(1)), and must set limits to researchers' capacities to define restrictive terms of usage and ensure that data are searchable and discoverable. It is for the Member States to define the specificities of the open access obligations (e.g., scope, level of obligations, embargos, rules on opt-outs from open access obligations, enforcement), but these should be in line with the Commission's revised Recommendation on access to and preservation of scientific information (2018/790).[58]

The EU legislator has left unaddressed whether the public interest in accessibility and re-use could be in conflict with individual researchers' right to academic freedom. Furthermore, intellectual property is also at play here. Datasets may in certain circumstances be subject to copyright or database rights, or contain materials protected by copyright or related rights such as text, images or audio-recordings. Both academic freedom and intellectual property are protected under the Charter of Fundamental Rights of the EU. Assuming that in many cases the initial owner of a copyrighted work will be the researcher(s) that created it, there are boundaries to what obligations national/EU legislators and RPOs can impose on individual researchers with respect to accessibility and allowing re-use.

2.3 Research data in the Open Data Directive: ambiguities & challenges

The Open Data Directive is very generic in that it addresses all manner of public sector bodies, including universities and (other) research performing organisations. **In its attempt to do justice to the specifics of scientific research** – with all its varied domains **– the legislator left a number of ambiguities and unanswered questions.** One ambiguity concerns the nature of research data as opposed to other public sector information. The second is the inclusion of publicly funded research, regardless of whether the recipient of funding is a private institution or whether research is partially privately funded. Third, it is not entirely clear when research data are considered 'publicly available' and therefore subject to re-use obligations. Finally, ambiguity results from the treatment of universities in light of their various functions. **These ambiguities are discussed below in connection with various risks and challenges that arise under the framework.**

2.3.1 Nature of research data

As noted in paragraph 2.2.1, the original PSI Directive only applied to public sector information in a narrow sense: documents produced or collected by public sector bodies in the performance of their public tasks (institutional/task-oriented approach). Now, the specific nature of data (research data) and the source of funding (publicly funded) have also become relevant criteria.[59] **This new data type/funding-oriented approach sparked criticism among various stakeholders.**

---

[58] European Commission, Commission Staff Working Document - Impact Assessment accompanying the Proposal for a Directive of the European Parliament and of the Council on the re-use of public sector information, SWD(2018) 127 final, 25 April 2018, p. 38-39.
[59] See also recital 3: "public sector and publicly funded information".

According to research associations, research data are fundamentally different from administrative data due to their specificity and complexity. They should therefore not be treated in a similar way and/or by the same legal framework designed for executive, judicial and administrative data originating from the public sector.[60]

Fundamental to the notion of re-use in the ODD is the idea that data created for a particular public task should be repurposed for different activities elsewhere, primarily outside the public sector. This is made explicit in the definition of re-use. **However, for scientific research this distinction does not make a lot of sense,** because the re-use of research data takes place primarily in research organisations (universities, research centres). What open science policy seeks to achieve by sharing and allowing re-use is primarily to further research itself. Access to research data allows better quality controls (review, falsification through replication studies), as well as more efficient use of data by pooling and avoiding duplication of data collection. Wider access can of course also create opportunities for civil society, media and businesses working with and benefiting from research data, but these are secondary purposes. **This suggests that it will remain a challenge to accommodate research data in the general re-use regime**, with its obligations that rest on public access under national administrative law and its request procedures tailored to administrative, judicial and legislative bodies.

2.3.2 Public funding criterion and the role of the private sector

The fact that the research data regime is 'blind' to the legal status of the entity carrying out research which is funded by public money means that **private RPOs may also be subject to reuse obligations.** This suggests that in practice, private institutions must differentiate between research datasets following from (semi-)publicly funded research and datasets following from privately funded research. There are practical difficulties in making that distinction, for public and private organisations alike. For instance, the provenance of datasets and the types of projects in which they have been generated or used must be clear. The administrative burden involved could potentially discourage private institutions to engage in publicly funded research, or discourage public RPOs from seeking collaboration. Another important question is what level of public funding would trigger the application of article 10. Any amount, half, or more?[61] **Uncertainty could make private partners risk averse.**

A last concern regards intellectual property rights. These, as well as legitimate commercial interests (also of RPOs), have to be taken into account while crafting open science policies. It remains unclear, however, how these interests are to be reconciled with pursuing maximum reusability of research data. Here, too, unpredictability may lead to risk averse behaviour.[62] For example, in

---

[60] European Parliamentary Research Service, 'Briefing EU Legislation in Progress: Re-use of public sector information', July 2019, p. 8; see also Science Europe e.a., 'Joint Statement on the Revision of the PSI Directive', 12 November 2018.

[61] Compare the Explanatory Memorandum to the Dutch proposal for implementation of the Directive, stating that "the amount of public funding that determines whether research data can be considered as 'resulting from publicly funded research activities' should be determined for each institution due to the large differences in the legal forms of research institutions and educational institutions", p. 7.

[62] Deloitte, Study to support the review of Directive 2003/98/EC on the re-use of public sector information, SMART 2017/0061, 2018, p. 296 (https://op.europa.eu/nl/publication-

the context of public-private research collaborations where private parties (mainly) provide research data, researchers could agree to transfer intellectual property to the private partner to reassure the latter that commercially sensitive or valuable data remain secure. It is difficult to assess how much of a barrier the research data regime of the ODD might be to university-industry collaboration; there is no hard evidence yet.

2.3.3 Status of research data as public

As noted, **the obligation to allow re-use applies to publicly funded research data that have already been made publicly available through an institutional or subject-based repository** (or in case Member States permit it in their national laws, through other infrastructures mentioned in recital 28). This regime is meant to reduce the administrative burden on research organisations and individual researchers that could otherwise be confronted with "requests for access" to research data.[63]

At this point, it is important to reiterate that **an essential principle of the ODD is that it builds on national access regimes.** The right to access data (and the corresponding obligation of a PSB to provide it) must be sharply distinguished from the right to re-use. There exists no such thing as a generic right to access public sector information; such rights are regulated in many different instruments, mostly at Member State level (e.g., access to official documents acts, public registries law, statistics law). To what extent research data held by public bodies (universities and other public RPOs) are subject to national public access laws is unclear. What is clear, however, is that at EU-level there are no harmonised rights to access research data held by public bodies, or duties to publish research data. The ODD tries to overcome this by obliging Member States to have open access policies in place, directing RPOs and funders to enhance access and re-use, and by relying on repositories. **But when are research data made 'publicly available' in a repository?** Does this imply that the data must have been deposited in an repository to which anyone has unencumbered access,[64] without having to go through authentication or authorisation procedures? What level of openness is required? For example, do all repositories identified in well-known registries of research data repositories as 'open' (as opposed to 'restricted')[65] meet the requirement that data is publicly available, or must an assessment be made for each repository?

It also remains unclear from article 10(2) when RPOs and RFOs should make their research data available for re-use. Research data are not subject to the right of anyone to request re-use. The time frame of 20 days mentioned in article 4(2) in the context of requests for re-use does not apply to RPOs, RFOs and research data (article 4(5)(b) jo. article (10(2), see also below). It is thus for the Member States to provide clarity in their national laws. The German transposition, for example, explicitly links the fulfilment of the obligation to make research data available for re-use to the completion of the research project and the fulfilment of

---

detail/-/publication/45328d2e-4834-11e8-be1d-01aa75ed71a1).

[63] European Commission, Commission Staff Working Document - Impact Assessment accompanying the Proposal for a Directive of the European Parliament and of the Council on the re-use of public sector information, SWD(2018) 127 final, 25 April 2018, p. 39; see also recital 28.

[64] See European Commission, Improving the availability of data and information on species, habitats and sites, 2021, p. 81: "New rules will also facilitate the re-usability of research data that is already contained in open repositories."

[65] Re3data is a global registry of research data repositories, initiated by German academic institutes, https://www.re3data.org/

the research purposes. The German legislator has taken into consideration that making research data FAIR requires more effort and time than meeting the requirements for (other) public sector information. A balance needs to be struck between the needs of open science and the burden for 'data providers' (Datenbereitsteller) to raise resources to make their research data re- usable.[66] **It is likely that Member States make different choices in this regard**, depending on how publicly funded research is organised, which shared facilities are already in place (e.g., repositories), how well data management practices are established in research communities, and other factors.

In light of the concerns raised by research communities about data sharing, **it is important that the consequences of making data available in a repository are clear and foreseeable** (i.e., if a certain standard license is used, what this means for control the researcher or RPO retains or relinquishes over the use by others). **If they are not, this may well have a chilling effect on researchers' willingness to deposit research data.** The national open access policies that Member States have to adopt play an important role here, but for cross-border sharing it is important that national policies align. **The current text of article 10 is quite ambiguous and leaves Member States a lot of discretion.**

2.3.4 Different roles/functions of universities

Scientific research to a large degree takes place in universities. A university can be a private institution (supported with public funding or not), or a body governed by public law (with or without some private funding). Universities typically engage both in research and education and run libraries to support these activities. They are also expected to engage in knowledge transfer and 'valorisation' of especially research. It is necessary in light of the applicable rights and obligations under the relevant legal framework to distinguish between all these roles.

It seems that the ODD treats the branches of universities in which research is conducted as a species of research performing organisations.[67] This allows the Directive to lay down the same rules regarding research data for universities and other (public) RPOs which exist in many Member States. **However, the distinction between research, education and library presupposes that there is in fact a clear separation of these functions, and crucially, of data or other information resources in the day-to-day workings of universities.**

The treatment of university libraries as separate institutions ('libraries') rather than as parts of universities also seems quite artificial given that

---

[66] See Gesetzenentwurf der Bundesregierung. Entwurf eines Gesetzes zur Änderung des E-Government-Gesetzes und zur Einführung des Gesetzes für die Nutzung von Daten des öffentlichen Sektors (Proposal for an Act amending the e-Government Act and introducing the Act on the use of public sector data), 09.03.2021, 19/27442. Available at: https://dserver.bundestag.de/btd/19/274/1927442.pdf, p. 29, 30-31.

[67] The EC regards them as such, see: https://ec.europa.eu/commission/presscorner/detail/de/MEMO_12_564 and https://ec.europa.eu/info/research-and-innovation/funding/funding-opportunities/funding-programmes-and-open-calls/horizon-europe/reforming-and-enhancing-european-research-and-innovation-system_en; EOSC also treats universities as RPOs, see: https://eosc.eu/members.

they normally constitute a single legal entity.[68] While educational establishments are excluded from the ODD, **university libraries are placed within the category 'cultural establishments'.** Cultural establishments are exempted from the ODD, except for libraries, museums and archives. Being considered as libraries, university libraries do not seem to have a legal obligation to allow re-use of their resources on the basis of article 3(2) (although they are expected to do so, under largely the same conditions as other public sector bodies). But if a university library hosts a data repository, which rules apply? Those for cultural establishments, libraries, or only the regime of article 10 since the university is handling research data? Another example: if university students engage in research as part of their degree programme, do the resulting data which may be stored in a library-run service come within the scope of article 10, or are they to be regarded as 'documents' that belong to the educational establishment?

Of course, Member States could in their implementing laws cure these uncertainties, but this would create a risk of divergent interpretations that ultimately the Court of Justice EU might have to resolve. That would take a long time. A glance at the current transpositions of article 10 also shows that Member States have not paid much attention to universities and the distinction between university libraries, educational establishments and research performing organisations.[69]

2.4 Impact on research stakeholders

2.4.1 Obligations for research stakeholders as data holders

It follows from Chapters III and IV of the Open Data Directive that in Member States' implementations, public sector bodies[70] and public undertakings must be the main addressees of the re-use obligations set out in these provisions. **With regard to research data, the Directive is less straightforward as to which entities are responsible for permitting re-use.** Article 10(2) merely states in general terms that publicly funded and publicly deposited research data "shall be re-usable" in accordance with the obligations laid down in Chapters III and IV, seemingly irrespective of the entity holding the data (a public sector body, a private institution, or an individual researcher).

---

[68] European Parliament, Report on the proposal for a directive of the European Parliament and of the Council amending Directive 2003/98/EC on re-use of public sector information, (COM(2011)0877 – C7-0502/2011 – 2011/0430(COD), 7 December 2012, p. 19: "This amendment is intended to clarify that the directive does not apply to documents held by a library which forms part of the university which holds the intellectual property right (IPR) in the document. A university and its libraries may constitute a single legal entity [emphasis added]. Without amendment, the exclusion of documents subject to third-party IPR would not apply where a library holds the document but the IPR is held by the university because the university would not be a separate (i.e., third) party."

[69] An exception is Germany, which at the federal level explicitly distinguished universities (Hochschule) from RPOs (Forschungseinrichtungen) and RFOs (Forschungsfördereinrichtungen), see Article 2, Section 2, Para. 2(3) of the Gesetz zur Änderung des E-Government-Gesetzes und zur Einführung des Gesetzes für die Nutzung von Daten des öffentlichen Sektors, Bundesgesetzblatt 2021, Nr. 46 vom 22.07.2021. The Dutch implementation proposal also names onderwijsinstellingen (universities and polytechnics / universities of applied sciences), see Concept wetsvoorstel Wet implementatie open data richtlijn, 23.12.2021, available at: https://www.internetconsultatie.nl/wetimplementatieopendatarichtlijn/document/7680.

[70] Including public universities and publicly funded (university) libraries, museums and archives.

It is clear from article 10(1), however, that the open access policies developed at national level must be directed towards RFO's and RPO's. It is by imposing conditions on funded research projects and organisations, that funders have so far played a key role in the sharing of research data and publications by RPOs and their researchers. The funders carry the proverbial stick. But they also play a role in funding shared facilities and services needed to enable data sharing at a larger scale, as can also be seen from their involvement in EOSC.

Where article 10(2) re-use obligations are concerned, it would be consistent with scientific practice and the logic enshrined in the Directive to assume that the data holders mandated by national open access policies as referred to in article 10(1) **to make their data openly available, are also the ones responsible for making the data re-usable.** The Dutch proposal for the transposition of article 10(2), for example, places the burden of re-use on RPOs and RFOs, as well as higher educational establishments, including universities, while explicitly excluding "data held by other organisations that could be regarded as research data".[71] Also interesting in this regard is the German transposition,[72] which appoints individual researchers as 'subsidiary' data providers to make publicly accessible research data available for re-use in case institutions have failed to do so. Although the German legislator still considers the institutions as the 'lead' data providers – since they are usually better equipped in terms of organisation and personnel to meet the re-use obligations with a reasonable amount of effort[73] – the provision appears to be the single national transposition explicitly imposing a re-use obligation on individual researchers. Admittedly, article 10(2) of the Directive includes a reference to individual researchers, but that reference is made, strictly speaking, to indicate which entities could (or should) make research data publicly available in repositories and not to indicate which entities are (consequently) subject to re-use obligations. Imposing a legal obligation on researchers to make their data available for re-use is arguably problematic given the burden it entails for individuals who, compared to the institutions, do not, to the same extent, have the facilities and knowledge to do so.

Once it is clear that an entity, as a data holder, is subject to the re-use regime, the following obligations apply:

- Data is made available as is, or where possible and appropriate in (technically) open formats, and is accessible, findable and re-usable, together with their metadata (article 5);
- Practical arrangements are made facilitating the search for and preservation of documents available for re-use;
- No conditions on re-use are imposed, unless such conditions are objective, proportionate, non-discriminatory, justified on grounds of a public interest objective and not unnecessarily restrictive; where licenses are used, standard licenses (electronically processible) are to be made available (article 8);
- Comparable categories of re-use(s) are treated equally (article 11);
- Re-use of research data is free of charge for users (article 6(6)(b));

---

[71] Article 5c; see also Explanatory Memorandum tot the Proposal amending the Wet hergebruik overheidsinformatie, p. 35, available at: https://www.internetconsultatie.nl/wetimplementatieopendatarichtlijn/document/7681.

[72] Article 2, Section 2, Para. 2(3)(b) of the Gesetz zur Änderung des E-Government-Gesetzes und zur Einführung des Gesetzes für die Nutzung von Daten des öffentlichen Sektors.

[73] Explanatory Memorandum to the Gesetz zur Änderung des E-Government-Gesetzes und zur Einführung des Gesetzes für die Nutzung von Daten des öffentlichen Sektors, p. 35.

- Applicants for re-use of research data are informed of the available means of redress relating to decisions or practices affecting them (article 7(3)).

The rules of the Directive that set out **requirements for handling re-use requests** do not apply to educational establishments, RPOs and RFOs (see article 4(6)(b)). University libraries are however subjected to these rules. RPO's and RFOs may decide for themselves if and how they handle requests for research data.

Of particular relevance for the sharing of research data is also the promotion of the use of standard licenses. As was set out in chapter 1, there is a difference between open licensing and FAIR. The ODD does not contain a hard legal obligation to make data technically and legally open, but does expect public sector bodies to make an effort towards this. The licensing terms most compatible with the ODD are open licenses such as Creative Commons Zero (no rights reserved), Creative Commons-BY (re-use is free, for anyone world-wide, for any purpose, irrevocably, subject only to attribution of the source) and the Open Data License (OdL) or Open Database License (OdbL), although the latter obliges re-users to also license adapted databases under the OdbL ('share alike' obligation).[74] The research data article refers to FAIR, promotes 'open by default' but does not impose the use of open licenses. On the contrary, by stipulating that "legitimate commercial interests, knowledge transfer activities and pre-existing intellectual property rights shall be taken into account", the provision suggests that a variety of licensing types can be used depending on the particulars of a dataset.

2.4.2 Rights for research stakeholders as data (re-)users

In their role as re-users of public data for scientific research, **universities, RPOs and RFOs strongly benefit from the re-usability of public data as promoted by the Directive.** It follows from the legal obligations imposed on public sector bodies, public undertakings and entities holding research data that when properly implemented, re-users can request certain public data (article 4) and enforce the conditions under which the data are offered for re-use (Chapters III and IV).

The 2019 Open Data Directive has been beneficial research stakeholders for a number of reasons. First, the broadening of the Directive's scope from traditional public sector information to documents held by certain public undertakings and research data has enlarged the pool of public data available for re-use for scientific purposes. **Second,** the newly introduced provisions on the re-use of high-value datasets will enable researchers and RPOs to access more sources and larger amounts of real-time data provided via APIs. Of note, when RPOs use high-value data sets for research, it may well be modified and/or integrated with other sources, and if this is done with (partial) public funding, the data becomes research data for the purposes of article 10 ODD. **Third,** the increased use of standardised licenses by public sector bodies will bring down transaction costs for researchers. **Finally**, provided that such licenses are compatible with licenses used in academic communities more broadly, it will be easier to combine data from different sources and share it for further (re)use.

[74] See https://opendefinition.org/licenses/ for an overview of various compatible ('interoperable') open licenses.

3 THE DATA GOVERNANCE ACT FINAL COMPROMISE TEXT

3.1 Key aspects

The legislative procedure of the Data Governance Act is in its final stage, with an agreement on the final compromise text (cDGA) reached in November 2021 and the DGA to be enacted in 2022.[75] **The Regulation aims to "improve the conditions for data sharing in the international market**, by creating a harmonised framework for data exchanges and laying down certain basic requirements for data governance (…)".[76] In light of this objective, **the cDGA consists of three main regulatory pillars:**

- Rules to ensure wider availability of data held by public sector bodies for re-use;
- A notification system and requirements for data intermediation services; and
- A registration system for data altruism.

In addition, **a European Data Innovation Board** (EDIB) **will be established**, tasked with advising and supporting the Commission in developing a consistent practice and guidelines with respect to the issues regulated by the cDGA (article 27). EDIB will involve expertise of the European Data Protection Board (EDPB) and others (article 26(1) cDGA). In its final provisions, the Regulation lays down a set of **general rules on international transfers of non-personal data** performed by public sector bodies, re-users of public data, data intermediation service providers or data altruism organisations. These are aimed at ensuring no such transfer conflicts with EU or Member State laws.

3.1.1 Wider re-use of protected public sector information

Chapter II of the cDGA sets out the conditions for re-use of certain public data, i.e., data held by public sector bodies which are protected on grounds of (i) commercial confidentiality, (ii) statistical confidentiality, (iii) protection of intellectual property rights of third parties, or (iv) protection of personal data insofar as such data fall outside the scope of the Open Data Directive. Typically these categories of information are not publicly accessible and therefore not subject to re-use obligations under the ODD (see recital 7 cDGA). In this regard, **the Data Governance Act final compromise text complements the Open Data Directive.** The difference between the regimes is that the proposed Regulation does not oblige public sector bodies to make the data available for re-use (article 3(3) cDGA). Its provisions are without prejudice to national, Union and international law on the protection of these categories of data. Like the ODD, the cDGA has a broad definition of public sector bodies, but excludes certain cultural establishments and educational establishments. **The impact on research stakeholders, including universities, is discussed further below.**

At the heart of the regime is article 5, which requires public sector bodies competent to grant or refuse access for re-use to set out conditions for re-use while implementing the provisions laid down in article 5(3)-(12) cDGA. These conditions must be "non-discriminatory, transparent, proportionate and objectively justified" (article 5(2)). The provisions of article 5 involve, among

---

[75] The text used here is the version published of 10.12.2021, following (informal) agreement between Council and European Parliament in November 2021. Available at: https://data.consilium.europa.eu/doc/document/ST-14606-2021-INIT/en/pdf.

[76] Recital 3 of the latest proposal of 10 December 2021, 14606/21.

other things, the obligation for public sector bodies allowing re-use to ensure that the protected nature of the data is preserved, for example through anonymisation, aggregation or secure environment processing (article 5(3)); the obligation to impose confidentiality requirements on re-users (article 5(5a)); the obligation to respect intellectual property rights, i.e., to obtain authorisation from the right holder (article 5(7)); the prohibition to invoke the sui generis database right with the aim of preventing re-use or restricting the re-use beyond the limits set by the cDGA (article 5(7)); and the obligation to ensure that non-personal confidential data or data protected by intellectual property rights are not transferred to third countries unless certain safeguards are put in place (article 5(8a)-(12)).

Other rules regarding exclusive arrangements, fees, competent bodies, information points, and the processing of requests for re-use are laid down in article 4 and articles 6-8a cDGA. **In contrast to the Open Data Directive** – which in principle prohibits public sector bodies to charge fees for re-use with the exception of the marginal costs incurred for reproduction, provision and dissemination of documents, or for the measures taken to anonymise personal data and protect commercially confidential information – **the cDGA leaves more room for fee-charging** (article 6). **At the same time, however, the cDGA requires public sector bodies to encourage the re-use of data for non-commercial purposes such as scientific research** by making data available at a discounted fee or free of charge (article 6(4) jo. recital 20).

3.1.2 Data intermediation services

Chapter III of the cDGA deals with a whole different topic, introducing a notification and supervision framework for (certain) data intermediation services. Data intermediation services are understood as services aiming to "establish commercial relationships for the purpose of data sharing between an undetermined number of data subjects and data holders, on the one hand, and data users on the other hand, through technical, legal or other means, including for the exercise of data subjects' rights in relation to personal data" (article 2(2c)). Given this relatively broad definition, many services may be subject to the regime of Chapter III, including even public sector bodies (recital 22) and – potentially – research stakeholders (see paragraph 3.3.2). Data cooperatives are subject to the rules as well (article 9(1)(c) cDGA).

It follows from article 2(2c)(a) in conjunction with recital 22a that **data intermediation services must essentially operate as neutral services**. They must not alter the data being shared (and the data inferred or observed from the activity of users) or use the data for other purposes than providing or facilitating the intermediary service. **To qualify as a registered service and to be allowed to operate across the EU a number of other requirements need to be met**. Only if the competent authority (which the Member States will have to establish) declares that the service complies with the requirements, it is allowed to call itself 'provider of data intermediation services recognised in the Union' (a 'certification' logo will be created by the EC). The requirements are aimed at giving data subjects, data holders and data users better control over access to and use of their data in order to increase trust in intermediary services. Ultimately, this trust should contribute to the voluntary exchange of substantial amounts of relevant data (recital 25 jo. 22).

3.1.3 Data altruism organisations

Separate from the data intermediaries regulated in Chapter III cDGA, Chapter IV regulates so- called data altruism organisations by establishing a registration system. Such organisations intend to promote the "voluntary sharing of data based on consent by data subjects to process personal data pertaining to them, or permissions of other data holders to allow use of their non¬personal data without seeking or receiving a reward that goes beyond a compensation related to the costs they incur making their data available, for purposes of general interest, (….), such as healthcare, combating climate change, improving mobility, facilitating the establishment of official statistics, improving public services, public policy making or scientific research purposes in the general interest" (article 2(10) cDGA). The EU legislator sees "a strong potential" in the use of data made available voluntarily by data subjects for objectives of general interest (recital 35). Because of this potential, the proposed Regulation aims to also "contribute to the emergence of pools of data made available on the basis of data altruism that have a sufficient size in order to enable data analytics and machine learning" (recital 35).

Member States are encouraged to facilitate data altruism through national policies (article 14a), **and must inform the Commission of such policies. There will also be a system of public registers.** Eligible organisations may seek registration and if that gets approved, they are allowed to call themselves 'data altruism organisation recognised in the Union' (articles 15¬22). Important requirements are that the organisation operates as a non-profit and is a formally and functionally independent entity. **The Commission may adopt delegated acts establishing a 'rulebook'** regarding information, technical and security requirements and recommendations on interoperability standards.

3.2 Scope of application to research data

Unlike the Open Data Directive, the DGA final compromise text does not distinguish between research data and non-research data. The voluntary re-use regime does not contain a separate regime for research data, and data intermediation services and data altruism organisations are agnostic as to the nature of the data shared.

In theory, RPOs and their researchers could be the buyers or sellers of data share through data intermediation services, but considering the non-profit nature of most publicly funded research, **RPOs are not the most likely customers of intermediation services, especially not as providers of data.** What could potentially occur, is that entities start collecting publicly funded research data that is released under the Open Data Directive and seek to commercialise the data as a provider through a data intermediary service. Intermediation services might even be set up to facilitate the mining and commercial exploitation of research data by providers and users/customers.

Where it concerns data altruism, it is clear that enabling the use for research purposes of donated data is a major driver for the scheme. Where research is done in organisations that qualify as public sector bodies, the donated data could be regarded as research data subject to both the Open Data Directive and the re-use provisions of the Data Governance Act. **To what extent research data are within the scope of the latter is however not clear.**

There is some ambiguity in the proposal with respect to the position of RPOs and the research data and non-research data they hold. Article 3(1) cDGA seems to suggest that public RPOs, including universities and university libraries, fall – as 'hybrid' public sector bodies – under the re-use regime of Chapter II irrespective of whether they hold data that have been produced or collected during research or during other activities. Recital 8, on the other hand, signals that a 'separate' re-use regime exists for research performing organisations and research funding organisations. It states that the re-use principles apply to such hybrid organisations "only in their capacity as research performing organisations" [emphasis added]. Whether this means that the re-use principles solely apply to research data generated by RPOs and not to data generated through other activities (e.g., administrative data) does not become fully clear from the text but would be consistent with article 1(2)(l) of the Open Data Directive. The recital further 'clarifies' that the re-use regime is meant to exclude data that are held "as part of a 22 specific public-private association with private sector organisations or other public bodies, bodies governed by public law or hybrid research organisations with the main purpose of pursuing research".

Recital 8 further makes explicit that the exchange of data in the pursuit of public tasks among public sector bodies or between public sector bodies and public sector bodies in third countries or international organisations is also outside the scope of the re-use rules. This would also include "in particular the exchange of data between researchers for non-commercial research purposes". The latter phrase seems to suggest that research data exchanged between researchers employed at public sector bodies (e.g., national state centres for disease control and the World Health Organisation) are exempt from Chapter II of the cDGA, but that data shared with researchers from private institutions fall within its scope.

Overall, it seems that most research and non-research data generated by public RPOs are largely left outside the cDGA but the exact scope of the exception remains difficult to interpret. The fact that research data, RPOs and RFOs are dealt with in recitals instead of in the actual provisions of the cDGA does not contribute to legal certainty.

3.3 Impact on research stakeholders

3.3.1 Re-use of certain categories of public data (Chapter II)

In their role as (re-)users of public data for scientific research, universities, other RPOs and (indirectly) RFOs could benefit substantially from the cDGA. The Regulation does not provide specific legal rights for re-users, but it follows from the legal obligations imposed on public sector bodies who decide to allow the re-use of certain protected data that re-users can request these types of data (article 8a). Most likely, public sector bodies will expand or create facilities to enhance access, but with restrictions as to who can access them at what terms. An example are micro-statistics which due to privacy protections are made available to vetted researchers in a secure environment.

The DGA final compromise text explicitly considers the interests of scientific research. Recital 2d of de cDGA states that "in order to facilitate and encourage the use of public sector data for the purposes of scientific research, public sector bodies are encouraged to develop a harmonized approach and

processes to make public sector data easily accessible for the purposes of scientific research in the public interest". Through e.g., standardised data formatting, informative metadata and other measures, the use of "publicly funded and produced data for the purpose of scientific research" should be promoted in accordance with the principle 'as open as possible, as closed as necessary'. Recital 11 encourages that "the conditions for re-use should be designed in a manner promoting scientific research", stating that privileging scientific research should be considered non-discriminatory.

In their role as data holders, public universities, other public RPOs and public RFOs may also fall under the re-use regime and be subject to the obligations laid down therein. The cDGA does not distinguish between universities' various functions (education, research and library). As a result, it is somewhat unclear whether and to what extent the re-use rules apply to public universities. It seems that data held by the educational branches of universities are exempt from Chapter II (see article 3(2)(c) cDGA). Although the term 'educational establishments'[77] is not defined anywhere in the cDGA, an interpretation consistent with the ODD means it encompasses the educational branches of universities. If, as in the ODD, university libraries are to be considered cultural establishments, these would be excluded too under article 3(2)(c), regardless of whether they fulfil research data management functions. The research branches of universities, however, seem – according to recital 8 – to fall within the scope of Chapter II with respect to (certain exchanges of) their research data.

In the event that public RPOs, including (the research branches of) universities, public RFOs and public university libraries hold protected categories of data listed in article 3(1) cDGA, they may decide for themselves whether they permit the re-use (article 3(3) cDGA). If they decide to, they must adhere to the principles laid down in articles 4-8a of the proposed Regulation.

3.3.2 Data intermediation services (Chapter III)

In their role as re-users of data for scientific research, universities, other RPOs and RFOs may benefit from the notification and supervision framework for data intermediation services. Of note, the framework's main focus in on promoting the sharing of data between commercial data holders and data users (see recital 4).[78] Much of the data sharing that RPOs engage in takes place in a not-for-profit context.

Recital 22a cDGA contains some examples of data intermediation services, including "data marketplaces on which companies could make available data to others" and "orchestrators of data sharing ecosystems that are open to all interested parties, for instance in the context of common European data spaces". An important element in the definition of data intermediation service is the establishment of 'commercial relationships' for the purposes of data sharing. According to recital 22a, cDGA the provision of cloud storage, analytics, web browsers, e-mail and so on is not considered a data intermediation service if it is limited to providing technical data sharing tools but is neither aimed at establishing commercial relationships between data holders and data users nor

[77] According to recital 8, the exception for educational establishments applies because "the works and other documents they hold are predominantly covered by third party intellectual property rights".

[78] It follows from recital 4 that the DGA's main focus is on individuals and businesses. However, recital 22 explicitly notes that public sector bodies may also qualify as data intermediation services.

allows the services provider to acquire information on the establishment of such commercial relationships. Recital 22b cDGA explicitly notes that "other services that do not aim to establish commercial relationships, such as repositories aimed at enabling the re-use of scientific research data in accordance with Open Access principles should not be considered data intermediation services" [emphasis added].

Even though scientific data repositories are explicitly mentioned as not being considered data intermediation services, research stakeholders could in theory offer such services and on that account be subject to the obligations laid down in Chapter III. As noted above, the services must seek to establish commercial relationships between data holders and data users. **However, it seems unlikely that RPOs or RFOs would routinely engage in offering data intermediation services,** either by themselves or as part of EOSC.

EOSC does also not seem geared towards fostering commercial relationships between data holders and data users. Individual partners of and actors in the EOSC environment could however be focused on commercial parties. Plus, there may be data exchanges that facilitate both commercial and non-profit exchange of data. An example is AMdEX, a joint initiative of universities, the city of Amsterdam and others for an open, reliable and fair data market for exchanging data, which plans to run smart mobility data market but also a research data exchange.[79] It is unclear how such hybrid initiatives are situated in the legal framework of the cDGA.

What is certain, is that data intermediation service providers must notify the competent authority (Article 10) before they may start their activities and are subject to certain conditions related to, e.g., purpose limitation, pricing, formats/interoperability, the provision of additional tools and services, transparency, continuity, international transfers, security, information provision/logging and consent from data subjects (Article 11).

3.3.3 Data altruism (Chapter IV)

If the mechanism to promote data altruism is effective, research stakeholders stand to benefit because it will lead to the emergence of big 'pools' of data that can be used for research purposes across borders. **Where RPO's take a passive role and wait what data altruism organisations emerge, it remains to be seen whether data is of interest to researchers. Data altruism organisations could be set up by or in cooperation with RPO's, also in the context of EOSC**. This comes with a number of requirements, an important one being that they the data altruism organisation is a separate legal entity, operating on a non-profit basis. In that case, the organisation needs to be registered in a national register of recognised data altruism organisations in the Member State in which it is established (article 17), it must deliver annual activity reports (article 18), and fulfil specific requirements to safeguard rights and interests of data subjects and data holders as regards their data (article 19).

Clearly, recognised data altruism organisations that engage in the processing of personal data must also adhere to the rules laid down in the General Data Protection Regulation (GDPR), **and will most likely rely on the consent of data subjects**. In the context of data altruism activities involving personal data for scientific research purposes, referring to the GDPR recital 38 of

[79] See https://amdex.eu/usecases/

the cDGA states that "scientific research purposes can be supported by consent to certain areas of scientific research when in keeping with recognised ethical standards for scientific research or only to certain areas of research or parts of research projects" and that "further processing for scientific purposes should not be considered to be incompatible with the initial purposes". The precise interplay between the GDPR and the data altruism requirements of the cDGA is complex. The interplay is important for research stakeholders to be aware of, because it may directly affect how data can be used, what rights data donors have, and which measures research stakeholders as beneficiaries or operators of data altruism pools need or must take to keep data secure and ensure compliance.

4 THE PROPOSED DATA ACT

4.1 Key aspects

Of the three instruments discussed in this report, the proposed Data Act (pDA) is the most cloud-oriented piece of legislation, and the most directed to the private sector. In terms of obligations, the Open Data Directive primarily addresses public sector bodies. The Data Governance Act final compromise text to an important degree targets data resources of public sector bodies, but also seeks to enhance the role of certain types of service providers: commercial data intermediation services and (non-profit) data altruism organisations. The proposed Data Act, by contrast, is very much focused on data access rights and corresponding obligations for data holders and cloud service providers in mainly the commercial sphere (denoted as 'data processing service providers'). Running at 46 pages of recitals and provisions, **the proposed Data Act is a complex piece of legislation. It regulates a number of divergent themes especially relevant to research:**

- Access obligations imposed on mainly businesses in case of exceptional (public interest) needs;
- Access to certain Internet of Things (IoT) data for users of connected products (including a clarification of the intellectual property status of IoT data);
- FRAND licensing for data access, with specific protections for SMEs;
- Switching of cloud service providers and associated data portability;
- Interoperability for data spaces.

To support its main ambitions, the proposal also addresses security of data and international transfers. Data processing providers have a duty to keep data secure. The international transfer of data is subject to strict rules, laid down in article 27. Providers must take all reasonable measures to prevent access by governments and international transfers in violation of EU and Member State laws. Even in cases where the courts or administrative authorities of a third country order that access should be given to data, a set of further conditions must be met for the processing provider to be allowed to give access. In any case, the provider must seek to provide the minimum amount of data permissible. Note that these obligations are addressed to data processing service providers, not to data holders, although the latter must be informed about any requests.

Lastly, the proposed Act contains provisions on enforcement and the role of supervisory authorities. The Data Act would not change existing legislation that deals with access to specific types of data (e.g., in the field of payment services or energy) **but its principles are supposed to be horizontal and to be applied also in future** (revision of) **sector specific legislation,** e.g., in the field of mobility.

4.1.1 Business-to-government data provision in case of exceptional need

Across different areas of law, many information duties exist. For example, businesses must supply statistics offices with data, register certain company information, report for purposes of environmental protection (e.g., emissions) or comply with agricultural subsidy systems. Such duties for companies and others to provide information to public authorities exist towards EU institutions but mostly at Member State level, including regional and local levels. **The proposed 'Exceptional need obligation' would add to these and fill a purported gap** where currently public authorities would benefit from gaining access but have no proper legal instrument as basis.

The material scope is limited in two important ways. There must demonstrably be a public emergency situation, or threat of one, i.e., there must be 'risk of serious and lasting repercussions on living conditions or economic stability, or the substantial degradation of economic assets.' **In other cases, where a lack of data prevents a public sector body from exercising its public interest task as laid down by law, the public sector body can claim access too.** A major caveat is that it must not be possible to (timely) secure access by other means, such as buying the data (including at market prices), introducing a specific legislative duty to supply data, or where the procedure foreseen in the Data Act is actually less of a burden for the data holder. Small, Micro and Medium sized enterprises (SMEs) are explicitly excluded from the scope of Exceptional need access obligations (article 7(2) pDA).

The recipient public sector body is allowed to share the data with other public sector bodies, or other third parties including research organisations, when this is necessary to fulfil the task for which access is needed. An important condition is that the data holder is informed and in case of outsourcing (e.g., data analysis), that the agreement with the third party is made publicly available (article 17(4) pDA). Data obtained under the Exceptional need provisions do not on that basis become subject to the Open Data Directive (article 17(3) pDA). It arguably would come within the scope of the Data Governance Act final compromise text, but that Act does not impose a duty to allow re-use.

4.1.2 Product user access to IoT data and ownership

The proposed Data Act sets out generic horizontal rules for the access and use of data generated by connected[80] products and their related services (articles 3-7). It aims to promote the (re)use of Internet of Things (IoT) data by empowering the users of connected products and services. To achieve this, rights and duties are laid down for data holders (e.g., providers of IoT products), for users (of IoT products) and for third parties with access to IoT data. The obligation to provide access to IoT data does however not apply to SMEs (article 7 pDA); on the concept of SME and further special provisions, see below under 4.1.3.

Users are given the right to access and share data generated by 'their' connected products and related services. This covers products they have

[80] The proposal in its definitions speaks of products 'that are able to communicate that data via a publicly available electronic communications service' such as via mobile or land telephone networks, satellite, cable or near-field communication networks. Communication devices such as PCs, tablets, phones, tv's, smart speakers are not covered by the proposed Regulation.

bought, leased or rented. Under the GDPR, natural persons (data subjects) already have the right to access and port some types of personal data about them, and the IoT access right would further broaden this. The access right does not extend 'downstream', i.e., to information derived or inferred from data generated by the product or service directly.

Article 3 sets outs the obligation on the data holder to make data generated by products or related services accessible for the user (customer). It promotes access by design, i.e., that the product and related service are set up in such a way that the product user has (direct) access to the data easily and securely. If the data is not accessible directly and by default, the data holder must ensure it can deliver it on request. **Article 4 provides users with a right to access data via a simple electronic request**. It also limits the use that a data holder can make of data generated by the product or related service: this must be based on contractual agreement with the user in case of non-personal data. The data cannot be used to gain insight in the user's commercial position. For personal data, the GDPR applies.

There are limits to what the user can do: the data may not be used for the purposes of directly competing products, nor may the user allow others to use data for such purposes. **Article 5 details under what conditions data can be shared with third parties.** Users can share the data obtained with third parties, whether commercial, non-profit, for research purposes or other purposes. The data holder may not charge the user for this, but the third party may be charged. Users can allow a third party to obtain the IoT data directly from the data holder. The latter may not undermine the commercial position of the third party on markets where they are in direct competition (recital 29). **The third party itself is subject to further restrictions, elaborated in article 6**. The most important restrictions are that it can only use the IoT data for the purposes agreed on with the user, that it cannot prevent the user from sharing with others, and that it cannot use the data to develop a competing product. The beneficiary of the data may not use the data for analytics which would contribute to gaining insight in the data holders' commercial position. It may also not provide the data to platforms that are designated as gatekeepers under the upcoming Digital Markets Act.[81]

Although the provisions are contained in a Chapter named 'Business to consumer and business to business sharing', their scope are broader: the relevant actors are 'user' (of IoT products), 'data recipient' (of IoT or other data, not being a user) and 'data holder'. The definitions of these terms[82] provided in article 2 pDA seem broad enough to encompass not-for profit entities, public sector bodies and individuals.

Connected to the access right for IoT data is the issue of 'ownership' of such data. Data as such are not subject to intellectual property, but datasets may attract protection especially of the so- called sui generis database right. The element of the proposed Data act with probably the least impact concerns the 'clarification' of the scope of intellectual property rights in data. The proposed

[81] Proposal for a Regulation of the European Parliament and of the Council on contestable and fair markets in the digital sector (Digital Markets Act), COM(2020) 842 final. The study of Lundqvist for EC DG for Research and Innovation analyses the impact of the proposed DMA on research.

[82] Article 2(5)-(7): 'user' means a natural or legal person that owns, rents or leases a product or receives a service; data holder' means a legal or natural person who has the right or obligation... to make available certain data; 'data recipient' means a legal or natural person, acting for purposes which are related to that person's trade, business, craft or profession.

Data Act states that article 7 Database Directive (which lays down the sui generis right) 'does not apply to databases containing data obtained from or generated by the use of a product or a related service'. This provision seems intended as a mere clarification of existing law. Its phrasing is problematic however, because following CJEU's interpretation of the sui generis database right's criteria for protection, databases of IoT generated data can in certain cases attract protection, even if the data generated by or obtained from the device as such will not attract protection.[83]

4.1.3 FRAND licensing for data access

Arguably the most generic part of the Data Act proposal is Chapter III. It sets out obligations for data holders with respect to licensing terms, regardless of whether their obligation to provide access derives from the Data Act or any other Union legislation, including implementing laws. **The Data Act will however not have retroactive effect**: its Chapter on obligations for data holders to make data available will not apply to any pre-existing obligations to give access to data (article 12(3)).

Article 8 sets out the principal obligations tied to Fair, Reasonable and Non-Discriminatory terms (FRAND). Although the data holder is free to set the terms of data access, it is not allowed to contract around the provisions on data access to the detriment of the data recipient (article 12 pDA). The obligation to provide access does not go as far as to force data holders to divulge trade secrets. Exclusivity is only permitted in exceptional cases. Comparable companies must be treated equally. The data holder must also be transparent about how compensation is calculated. These are reminiscent of key principles of the Open Data Directive and Data Governance Act, although the former is more radical in the way it seeks to establish rights of re-use.

In keeping with FRAND, any compensation required by the data holder must be 'reasonable' (article 9). This is a broader criterion than the one laid down in the Open Data Directive, which puts remuneration at maximum of marginal costs of dissemination centre stage. Where re-use of data is subject to the ODD, the latter's charging principles (and those of implementing legislation) trump the Reasonable compensation of the Data act. This follows from article 9(3) pDA. It is important to note, however, that since the ODD itself does not create access obligations but regulates re-use of (already) public data, the interplay between the pDA and ODD seems limited.

Small, micro and medium sized enterprises enjoy a better position in

[83] For sui generis protection, substantial investments must have been made, but investment in the creation of data as such does not count. Whether there is substantial investment will depend not on the costs associated with having IoT devices generate data, but on other investment in the collection and presentation of data. See e.g. Karanikolova, K., Chicot, J., Gkogka, A., et al., Study in support of the evaluation of Directive 96/9/EC on the legal protection of databases : final report, Publications Office, 2018, https://data.europa.eu/doi/10.2759/04895. See also the study for DG Science and Innovation by M. Senftleben (not yet published).

many areas of EU law. This is also true under the proposed Data Act.[84] It applies the more favourable principle of the ODD: SMEs may not be charged more than the costs directly related to making the data available (article 9(2) pDA). More importantly, a dedicated set of rules bolsters the position of SMEs vis-à-vis data holders. In a nutshell, the SME that acquires data under unfair contractual terms unilaterally imposed by the data holder is not bound by those unfair terms. What constitutes an unfair term is defined as one that 'grossly deviates from good commercial practice in data access and use, contrary to good faith and fair dealing'. The article lists specific terms as unfair (so-called black listed) and as presumed unfair (grey listed), e.g. on limitations of liability and on termination of contracts at short notice. Grey listed are also terms that limit the data recipient from using data that itself has contributed or generated, or from making copies of such data (article 14(4) pDA). The data holder is allowed to use technical protection measures and smart contracts to ensure that the data user complies with the licensing terms.

In case of conflict over the terms and conditions, the data holder and user must have recourse to certified dispute settlement bodies (article 10). Next to this, the general chapter on enforcement also means Member States must designate a competent authority where those who seek to secure their rights under the Data Act can file complaints (Chapter IX).

Article 34 mandates that the EC develop non-binding model contractual terms on data access and use. Considering that this article is part of the general chapter on implementation and enforcement, these contractual terms can arguably cover not just FRAND for data access, but also agreements on IoT data, contractual arrangements for switching (as these may include data access for portability, see below), etc.

4.1.4 Facilitating cloud switching

A separate Chapter centres on improving the possibilities for customers to switch cloud service providers and on the necessary interoperability[85] **to facilitate switching**. The industry- developed 'Switching Cloud Providers and Porting Data (SWIPO)' Codes of Conduct, after EC was mandated to promote these under the Free Flow of Non-Personal Data Regulation (2018). These self-regulation instruments set out principles meant to ensure compliant service providers have processes in place for porting and switching, are transparent about what data can be ported and under which conditions, what the technical requirements are, etc. To date not many cloud service and software-as-a-service (SaaS) providers take part, and research organisations do not seem to be involved.[86] The EC suggests the Codes of Conduct are ineffective.[87]

Under the proposed Data act, providers of cloud services and other data

---

[84] The concept of SME in the proposal refers to definitions given by the EC earlier. They are enterprises –i.e., any entity engaged in an economic activity regardless of legal form– that have a maximum of 250 staff headcount, less than 50 mio turnover or 45 mio balance sheet total. See Commission Recommendation of 6 May 2003 concerning the definition of micro, small and medium-sized enterprises, OJ L 124, 20.5.2003, p. 36–41.

[85] For an analysis of interoperability aspects in data sharing, see Jörg Hoffmann, Begoña Otero, 'Demystifying The Role Of Data Interoperability In The Access And Sharing Debate', 11 (2021) JIPITEC 252.

[86] See https://swipo.eu/ for SaaS and IaaS offered by cloud service providers.

[87] Explanatory Memorandum Data Act Proposal, p. 9.

processing services, are subject to a series of obligations which should make it easier for users to switch service providers. They are instructed to remove certain contractual (article 26), commercial, technical (also in article 29) and organisational obstacles. In addition, the EC is given the authority to instigate development of common specifications to enhance interoperability for inter alia the common European data spaces (including EOSC), see the next section.

The concept of data processing service is very broad. It covers a wide range of services that 'allow on-demand and broad remote access to a scalable and elastic pool of shareable and distributed computing resources' (recital 71). These cloud resources can be networks, servers, operating systems, software, development tools, storage, applications and services. **The provisions impose different obligations on different actors:** cloud infrastructure providers must ensure that users who switch data processing services to a different provider, retain functional equivalence in the use of the new service. Such data processing providers in turn, must make 'open interfaces publicly available and free of charge'. These interfaces must comply with open interoperability specifications or European standards (which the EC can request standardization bodies to set, article 29(4)). Where no such standard exists, the provider must ensure it can comply with a customers' request 'for export all data generated or co-generated, including the relevant data formats and data structures, in a structured, commonly used and machine-readable format' (article 26(4) pDA). Characteristics and requirements of open interoperability specifications and European standards are detailed in article 29. The regulation would set dates aimed at ensuring the phasing out of switching charges; and empowers the EC to monitor this (article 25).

4.1.5 Interoperability Data spaces

For providers of data processing services, the proposal sets out further obligations in order to facilitate data sharing, mechanisms and services. **These obligations also apply to 'operators' of data spaces;** what type of actor qualifies as operator is not articulated. Providers must meet a number of 'essential requirements', and make an EU declaration of conformity to them. Among the essential requirements (article 28(1)) are the sufficient and consistent description of data structures, data formats, vocabularies, classification schemes, taxonomies, code lists, and the technical means to access the data. Article 28(2) empowers the Commission to adopt delegated acts to supplement specifications. According to article 28(4), the Commission is allowed to mandate the development of 'harmonised standards' that satisfy the essential requirements. The Commission is also tasked with adopting 'common specifications' where harmonised standards do not exist or where it considers them insufficient to ensure conformity with the essential requirements (article 28(5) pDA).

Specifically to aid the functioning of data spaces, article 28(6) **empowers the Commission to adopt guidelines** laying down interoperability specifications for data and data sharing such as architectural models. Notably, the Commission can also issue guidelines on technical standards that implement legal rules and (contractual) arrangements on data access and permitted uses between parties.

Smart contracts (self-executing contracts) are expected to play a large role in data spaces. Article 30 lays down essential requirements regarding smart contracts for data sharing. Parties selling or deploying smart contracts must ensure e.g., robustness and safe access controls, and draw up an EU declaration of conformity. The Commission has authority to ask European standards bodies to draft harmonised standards that satisfy those essential requirements.

With respect to data spaces, or needs of certain sectors or public interests, the EU could legislate specific requirements e.g. (further) limiting what data holders are allowed to do with data, **or technical aspects of data access.** These would be unaffected by the Data Act (article 40(2) pDA).

4.2 Scope of application to research data

The proposed Data Act is a generic piece of legislation, so from that perspective it is not surprising that in the preparatory phases, the position of science has not been a real issue under consideration. The studies commissioned in support of the proposal (the Explanatory Memorandum mentions several) are either silent on data sharing in science, or mention research data only in passing.[88] One notable exception is the evaluation study of the Database Directive, where the barriers that sui generis database rights cause for (academic) research and the sharing of research data is highlighted.[89] **If we consider more narrowly the extent to which the proposed Data Act would apply to research data**, i.e., to data collected or produced in the course of scientific research activities, used as evidence or to validate findings, **three things stand out.**

First, IoT and 'exceptional need' data will typically not be research data, but can become so when put to research purposes. They would then be subject both to the Open Data Directive's regime and to the Data Act. It requires a case-by-case evaluation of datasets to work out what the legal ramifications are. Take the example of users of certain smart home devices that allow researchers as 'third' parties to obtain IoT data directly from data holders, to be used for say behavioural research. Each data holder can impose restrictions and seek compensation for data access and use (under FRAND conditions) from the researcher. Depending on the terms, researchers will likely be prevented from sharing same data (perhaps modified or enriched) under open policies as promoted by the Open Data Directive.

Second, **a similar complication arises for 'exceptional need' for data to which researchers have access because they have been commissioned to assist a public sector body, in say the management of a public health emergency by a Ministry of Health.** But this time, the potential clash is between Data Act, Data Governance Act, and Open Data Directive. The pDA is clear about the position of the Ministry: even though it holds the data, they will not become subject to the ODD so there is no obligation for it to allow re-use (in all likelihood, this would not arise even if one would apply the ODD, because the data in question will likely not be publicly accessible under national law, so not subject to re-use obligations under the ODD anyway). Under the cDGA, however, it seems the Ministry needs to make an effort to ensure it can allow re-use. Arguably, when researchers obtain the data, article 10 ODD is not triggered

---

[88] See e.g. Study on provisions applying to B2G data sharing for the purpose of the production of statistics in the context of the European Statistical System (B2G4S); 'Study on the switching of cloud service providers, which included a cross-sectorial workshop in Q2 2017' (EC 2018); 'Study on the economic detriment from unfair and unbalanced cloud computing contracts. Deloitte et al, Impact Assessment on enhancing the use of data in Europe Report on Task 1 – Data governance (SMART 2020/694 | D2), (EC 2020).

[89] JIIP et al., Study in Support of the Evaluation of Directive 96/9/EC on the Legal Protection of Databases Final Report (EC 2018).

either. But strictly speaking, the pDA only sets aside the ODD in case of "...data held by public sector bodies." So, clarification is desirable.

Third, when data have a 'dual' nature, i.e., they qualify as research data but also as IoT data, or as research data and data falling within the scope of access and use obligations for which the pDA prescribes FRAND, **the question becomes which body takes the lead on developing standard licenses, and how these could be made compatible with FAIR based models of research data licensing** that Member States might develop on the basis of article 10 ODD? The division of responsibilities in this space is complex, involving various parts of the EC, Member States, research funding organisations and research performing organisations. Remember also that the Data Act explicitly leaves room for the EU to legislate specific requirements and norms for data holders when the development of data spaces (including EOSC) necessitates this.

4.3 Impact on research stakeholders

With respect to 'exceptional need data', research performing organisations, funders and providers of EOSC infrastructure components and data processing services can in principle be subjected to a duty to provide data under article 14 pDA. Initially presented as covering B2G (business-to-government) information obligations, the duty would seem to exist for all data holders. This because under article 2(6) Data Act proposal, a data holder is any 'legal or natural person who has the right or obligation, in accordance with this Regulation, applicable Union law or national legislation implementing Union law, or in the case of non-personal data and through control of the technical design of the product and related services, the ability, to make available certain data'. On that reading, this would include all manner of public sector bodies including universities and research performing organisations.

The exclusion of SMEs from the provisions on Exceptional need data is relevant to academic institutions. An SME can be an enterprise (partially) controlled by universities/research performance organisations such as a technology start-up. Such SMEs cannot be forced to provide data.

As a rule, a research performing organisation will not be in a position itself to demand access to data according to the provisions in Article 14 and 15 pDA. Although an institution may well qualify as a public sector body within the meaning of the Act (Recital 56 pDA)[90], it will typically not itself be tasked with addressing (impending) emergencies. This might be different for e.g., centres for disease control that also engage in research. With respect to other public interest data, the bar for obtaining access is high. The university or research centre will have to show the data is needed to fulfil a specific task in the public interest that has been explicitly provided in law. It will also have to show there is no way to acquire data, e.g., by concluding an agreement with the data holder. Researchers might be given access via an indirect route. Under Article 21 pDA public sector bodies (e.g. when responsible for emergency response to natural or man-made disasters) are given the right to share the acquired data with organisations (or individuals) for carrying out scientific research. Such research must be however be compatible with the purpose for which the data was

[90] Per Recital 56 pDA "In situations of exceptional need, it may be necessary for public sector bodies or Union institutions, agencies or bodies to use data held by an enterprise to respond to public emergencies or in other exceptional cases. Research-performing organisations and research-funding organisations could also be organised as public sector bodies or bodies governed by public law."

initially requested[91]. The pDA further limits eligible recipients to organisations and individuals operating on a non-profit basis or under a public interest mission recognized by law.

With respect to IoT data, it seems doubtful that research performing organisations as data holders will be subject to access obligations themselves, simply because they typically are not manufactures or sellers of IoT products and related services. **On the other hand**, staff of research performing organisations will use a degree of IoT products such as lab equipment and medical devices. **Researchers thus stand to benefit from the data access rights.** The access obligations only target large companies. So only researchers that use their IoT products stand to benefit from IoT data.

Where the framework for (cloud) data processing services and data spaces is concerned, here the impact of the Data Act can be expected to be substantial. Where the functioning of EOSC relies on (commercial) cloud service providers, the rules enabling switching may reduce dependency. In addition, the new powers of the EC to set norms and achieve standardisation with respect to dataspaces, dataspace interoperability and smart contracts may also substantially impact EOSC. In addition, there is the role of the EC as regards developing model contract clauses on data access and use (article 34).

Another issue that would benefit from clarification concerns the model contract clauses on data access and use that the EC must develop (Article 34 pDA). How these would apply to research data is somewhat unclear. Where there is no obligation to give access to research data, the provision does not seem to kick in. The EU, however, also has the power to develop licensing models for data spaces, which is something the EOSC association is working on. How, in turn, this relates to internationally supported licensing schemes for specific domains, e.g. in the biomedical field, is another piece of the puzzle that needs solving.

It is also unclear what types of licensing models will function in which spheres, and how the concurrent use of models affect the exchange of research data in EOSC. As was set out in chapters 1 and 2, the Open Data Directive's general thrust is towards open content licensing (i.e. technically and legally open), whereas for research data specifically FAIR is considered the more appropriate frame, also in the ODD. The proposed Data Act however centres on the more indeterminate FRAND obligations which are potentially much more restrictive.

The key challenge seems to be to establish where further development and standardisation of licensing terms is needed, and how a workable level of license interoperability can be guaranteed. Given the incredible diversity of data and of data processing services, this seems a formidable task which will require close cooperation of all stakeholders. An added complication is that data sharing takes place in a global context, involving many commercial parties in the field of e.g. data analytics, tool development, and publishing. So developing standards may be difficult.

What does seem to be clear is that the leading idea of the Open Data Directive – that research data should wherever possible be legally and technically open – **is one that cannot be applied in the context of the data sharing**

[91] Article 21(2) "Individuals or organisations receiving the data pursuant to paragraph 1 shall act on a not-for-profit basis or in the context of a public-interest mission recognised in Union or Member State law. They shall not include organisations upon which commercial undertakings have a decisive influence or which could result in preferential access to the results of the research."

mechanisms regulated in the Data Governance Act final compromise text and the proposed Data Act. Harnessing the power of data altruism organisations for research, and making use of data intermediary services to provide access to and use research data, can be a means to ensure legitimate interests are protected, including commercial interests of private partners or university itself, and data donors' privacy and integrity.

5. SYNTHESIS AND RECOMMENDATIONS

This short concluding chapter recapitulates the findings on the Open Data Directive, Data Governance Act final compromise text and proposed Data Act, and in broad strokes describes how these instruments contribute to realization of EU research data policy. **It identifies and makes recommendations about key legal issues that need to be resolved**. These have to do with ambiguities in the scope of application to research data, the interpretation of provisions, and the consistency between the instruments from the perspective of open science research policy.

To sum up: the Open Data Directive is the generic instrument that regulates re-use of data held by public sector bodies in Member States. Since 2019 it also covers research data. The Open Data Directive creates certain obligations for research performing organisations (in and outside universities) as suppliers of data, and contains rights for them as users of data.

The Data Governance Act final compromise text consists of three major building blocks. It seeks to expand principles of the Open Data Directive to a broader set of public sector data. In addition, it regulates data intermediation services by imposing requirements, including registration. It also sets up a registration scheme for organisations engaging in data altruism, i.e., the sharing of data for public interest purposes including academic research.

The proposed Data Act addresses disparate issues. It aims to promote the (re-)use of Internet of Things (IoT) data by empowering the users of connected products and services. Also, switching cloud and other data processing services should become easier once providers are subject to a series of obligations. Furthermore, data holders will be subject to conditions with respect to licensing terms, whether their obligation to provide access derives from the Data Act itself or any other Union legislation. A number of provisions are aimed at the realisation of interoperability for data spaces (including EOSC). Also, public sector bodies get to obtain data from businesses in the event of public crises and other exceptional need situations. Researchers may benefit in modest ways from this new power.

5.1 Main findings

As research performing organisations (whether as part of universities or independent organisations) **operate in an increasingly digitising and connected transnational environment, the complexity of the legal framework within which they must operate grows**. **These frameworks are not necessarily tailored to the specific functions of universities, other RPOs or funders.** It is clear, however, that when it comes to academic research, universities are seen both as beneficiaries of provisions (access to more data) and as subjects of obligations (giving more access to data). The Open Data Directive is the key instrument here. In addition, in particular the proposed Data Act and the Data Governance Act will play a role in the realization of the European Open

Science Cloud. There will be additional formal and non-binding kinds of rulemaking and standard setting at the EU level in many areas, e.g. around standard contractual clauses for data sharing and use, requirements for data intermediation services and data altruism organisations, interoperability requirements for cloud services and data spaces. Just how much this will facilitate the realization and operation of EOSC will depend on how effectively the needs of research stakeholders are accommodated in these processes.

An important point of attention for the future concerns the relation between EU law and the laws of other jurisdictions. EOSC is intended – as its name suggests – as a European science data space, one of several in the EUs data strategy. The question arises what position it will take within the broader international context. EOSC-designers recognize that the international dimension must not be overlooked,[92] research is a global endeavour. The international exchange of research data involves many legal issues and distinct areas of law, e.g. intellectual property, (contractual) rights of users of services, data protection law, export controls of sensitive knowledge, etc.

With respect to the sharing of research data specifically, the Open Data Directive remains the key instrument. This because it addresses directly when public sector bodies must allow re-use of data by researchers and others, but also because it mandates that Member States must develop open science policies and specifically for data made public in research repositories, mandates that re-use of research data.

For especially research performing organisations, the Open Data Directive and Data Governance Act bring some legal uncertainty, notably because functions of public universities (education, research, library) are subjected to different rules. The analysis of the Open Data Directive shows that different functions and constituent parts in education, research, and the changing role of libraries from collectors to managers of information assets are not clearly in view when shaping reuse policy.

It is a key ambition of the EU's open science policy that FAIR (Findable, Accessible, Interoperable and Re-usable) **and open data sharing become the default for EU-funded scientific research**. **Through the Open Data Directive's article 10, Member States are expected to emulate this at the national level**, through policies aimed at research funding organisations and research performing organisations. Data sharing requires clarity about the conditions under which data is made available, and what if any the terms of use are. This is where licensing systems play their role.

The three instruments use different templates for the terms and conditions under which access and re-use of data is allowed. The ODD does not mandate the use of specific licensing schemes, but breathes a preference of open licensing, i.e. data is technically open and there are no restrictions on its re-use for whatever purpose. It also asks that standard licenses are used, and that any restrictions are kept to a minimum. As such, this is consistent with EU's open science policy for data.

---

[92] 'As we co-create EOSC we must set our work in the broader international context. Research is global and the standards we deploy in a European context to create a web of FAIR data need to align, or at least interoperate with parallel initiatives in other countries, continents and sectors.' see: https://www.eoscsecretariat.eu/news-opinion/eosc-international-context

The research data provision within the ODD is at the same time more specific and more flexible by making FAIR a key concept. As described in chapter 1, the Findable, Accessible and Interoperable elements come with clear and specific guidance on what it means, in more detail than the provisions of the ODD that address 'technical openness'. The Data Governance Act final compromise text seeks to expand ODD principles, but is less prescriptive because it does not oblige public sector bodies to allow re-use and recognizes that to protect certain interests, terms and conditions may have to be imposed where access and re-use are allowed. The proposed Data Act, where it concerns access obligations, borrows the concept of FRAND licensing: fair, reasonable, non- discriminatory terms of use. It recognizes that technical protection measures and smart contracts may be needed to enforce FRAND. With its focus on remuneration and protection of commercial interests of the data holder against e.g. competition, it is quite removed from FAIR and open data.

Universities, research performing organisations and funders will have to navigate this complex space. To be able to do that, it will be necessary to keep track of the provenance of data used for and generated by research processes. Only then is it possible to ascertain on what conditions researchers can re-use data, and when research performing organizations (and researchers) are obliged to allow others re-use.

Also, where it concerns the terms under which research data are licensed, attention must also be paid to the academic freedom of researchers. For example, imposing the use of true open licenses means that academics are forced to accept that any person, anywhere in the world can do whatever they please with research data produced by them. To date there has been little attention to how the protection of the fundamental right to academic freedom (protected under the EU Charter of Fundamental Rights, international treaties and national constitutions) the limits the power of research funders, research performing organisations and especially also of (legislative) authorities to dictate how research data must be shared.

5.2 Recommendations

- Work with research stakeholders, especially research performing organisations and funders, to clarify where **open content licensing and FAIR** is possible, and ensure that stakeholders are clear on the prerogative that the Open Data Directive and Data Governance Act final compromise text have in this respect over the potentially much more restrictive FRAND licensing that the proposed Data Act foregrounds.

- Establish where **specific development and standardisation of (open) licenses** is needed to enhance research data sharing, and how a workable level of license interoperability can be guaranteed.

- Clarify in guidelines (but preferably in the text of the law) precisely how **the university** (as a public sector body) **and its different constituent parts** (education, research, libraries) are subject to the Open Data Directive and to Chapter II of the Data Governance Act final compromise text, taking account of changing roles of libraries and of the interplay between research and education.

- Clarify through guidelines who are the **primary addressees of the obligation/responsibility to make certain research data available for re-use** in the Open Data Directive.
- Clarify when **research data** are to be regarded as **publicly accessible** and

inclusion in what **type of repository triggers the obligation to allow re-use** under article 10(2) Open Data Directive.

- Provide guidance on principles when **research data should be made available for re-use in repositories** (time frames, e.g. within a certain time period after a research project has been completed?)

- Clarify the relationship between Open Data Directive, Data Governance Act final compromise text and proposed Data Act **with respect to data obtained by the public sector under the exceptional need provisions** of the Data Act. For research performing organisations this would provide clarity with respect to the position of researchers that obtain access to such data, notably whether it can or cannot be re-used for other research purposes.

- Ensure that where the EC uses its delegated powers to ask for **interoperability standards,** the wider international dimension of data exchange is sufficiently taken on board so as to facilitate the (continued) sharing of research data beyond Europe.

- Ensure that in open science policy making, sufficient attention is given to **academic freedom** as guaranteed by the Charter of Fundamental Rights.

6 REFERENCES

Legislation and legislative proposals

Database Directive

Directive 96/9/EC of the European Parliament and of the Council of 11 March 1996 on the legal protection of databases. Available at: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex%3A31996L0009

Directive 2013/37/EU

Directive 2013/37/EU of the European Parliament and of European Parliament and of the Council of 26 June 2013 amending Directive 2003/98/EC on the re-use of public sector information. Available at: https://eur-lex.europa.eu/legal-content/FR/ALL/?uri=celex%3A32013L0037

Open Data Directive

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). Available at: https://eur-lex.europa.eu/legal-content/fr/TXT/?uri=CELEX%3A32019L1024

Proposal for a Data Act

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), Brussels, 23.2.2022 COM(2022) 68 final. Available at: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=COM%3A2022%3A68%3AFIN

Proposal for a Data Governance Act

Proposal for a Regulation of the European Parliament and of the Council on European data governance (Data Governance Act) - Analysis of the final compromise text in view to agreement, 10.12.2021, Brussels 14606/21 (published on 15.12.2021). Available at: https://data.consilium.europa.eu/doc/document/ST-14606-2021-INIT/en/pdf

Proposal for a Digital Markets Act

Proposal for a Regulation of the European Parliament and of the Council on contestable and fair markets in the digital sector (Digital Markets Act), COM(2020) 842 final. Available at: https://eur-lex.europa.eu/legal-content/en/TXT/?uri=COM%3A2020%3A842%3AFIN

Proposal for an Open Data Directive

Proposal for a Directive of the European Parliament and of the Council on the re-use of public sector information (recast), 25.04.2018, COM(2018) 234 final. Available at: https://eur-lex.europa.eu/legal-content/EN/HIS/?uri=CELEX%3A52018PC0234

Proposal for a PSI Directive

Proposal for a Directive of the European Parliament and of the Council on the re-use and commercial exploitation of public sector documents, 05.06.2002, COM(2002) 207 final. Available at: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A52002PC0207

PSI Directive

Directive 2003/98/EC of the European Parliament and of the Council of 17 November 2003 on the reuse of public sector information. Available at: https://eur-lex.europa.eu/legal-content/en/ALL/?uri=CELEX%3A32003L0098

Germany implementation Open Data Directive

Gesetzenentwurf der Bundesregiering. Entwurf eines Gesetzes zur Änderung des E-GovernmentGesetzes und zur Einführung des Gesetzes für die Nutzung von Daten des öffentlichen Sektors, 09.03.2021, 19/27442. Available at: https://dserver.bundestag.de/btd/19/274/1927442.pdf.

The Netherlands implementation Open Data Directive

Concept wetsvoorstel Wet implementatie open data richtlijn d.d. 23-12-2021. Available at: https://www.internetconsultatie.nl/wetimplementatieopendatarichtlijn/document/7680.

Spain implementation Open Data Directive

Ley 37/2007, de 16 noviembre, sobre reutilización de la información del sector público. Available at: https://www.boe.es/eli/es/l/2007/11/16/37/con.

Poland implementation Open Data Directive

Ustawa z dnia 11 sierpnia 2021 r. o otwartych danych i ponownym wykorzystywaniu informacji sektora publicznego. Available at: https://isap.sejm.gov.pl/isap.nsf/download.xsp/WDU20210001641/T/D20211641L.pdf.

Policy documents

EC/European Commission (2022). Horizon Europe General Model Grant Agreement, Version 1.1, April 2022. Available at: https://ec.europa.eu/info/funding-tenders/opportunities/docs/2021-2027/common/agr-contr/general-mga_horizon-euratom_en.pdf

EC/European Commission (2022). Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions – Establishing a European Declaration on Digital rights and principles for the Digital Decade, Brussels, 26.1.2022 COM(2022) 28 final. Available at: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:52022DC0027

EC/European Commission (2021). EOSC Interoperability Framework: Report from the EOSC Executive Board Working Groups (WG) FAIR and Architecture, February 2021. Available at: https://op.europa.eu/en/publication-detail/-/publication/d787ea54-6a87-11eb-aeb5-01aa75ed71a1

EC/European Commission (2020). A European strategy for data, 19.02.220, COM(2020) 66 final. Available at: https://ec.europa.eu/info/sites/default/files/communication-european-strategy-data-19feb2020_en.pdf

EC/European Commission (2019). Commission Decision adopting Creative Commons as an open licence under the European Commission's reuse policy, Brussels, 22.2.2019 C(2019)1655 final. Available at: https://ec.europa.eu/transparency/documents-register/detail?ref=C(2019)1655&lang=en

EC/European Commission (2018). Commission Staff Working Document, Impact Assessment Accompanying the document Proposal for a Directive of the European Parliament and of the Council on the re-use of public sector information, 25.04.2018, SWD(2018) 127 final. Available at: https://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=SWD:2018:0127:FIN:EN:PDF

EC/European Commission (2018). Commission Recommendation (EU) 2018/790 of 25 April 2018 on access to and preservation of scientific information. Available at: https://eur-lex.europa.eu/eli/reco/2018/790/oj

EC/European Commission (2016). European Cloud Initiative – Building a competitive data and knowledge economy in Europe, 19.04.2016, COM(2016) 178 final. Available at: https://eur-lex.europa.eu/legal-content/en/TXT/?uri=CELEX%3A52016DC0178

EC/European Commission (2015). A Digital Single Market Strategy for Europe, 06.05.2015, COM(2015) 192 final. Available at: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex%3A52015DC0192

EC/European Commission (2014). Commission Notice. Guidelines on recommended standard licences, datasets and charging for the reuse of documents, 24.07.2014 2014/C 240/01)

EOSC/European Open Science Cloud (2021). Memorandum of Understanding of the Co-programmed European Partnership on the European Open Science Cloud. Available at: https://ec.europa.eu/info/sites/default/files/research_and_innovation/funding/documents/c_2021_4113_f1_annex_en_v3_p1_1213802.pdf.

Literature

Deloitte, Study to support the review of Directive 2003/98/EC on the re-use of public sector information, 24.04.2018. Available at: https://op.europa.eu/nl/publication-detail/-/publication/45328d2e-4834-11e8-be1d-01aa75ed71a1.

European Commission (2020). Supporting the Transformative Impact of Research Infrastructures on European Research: Report of the High-Level Expert Group to Assess the Progress of ESFRI and Other World Class Research Infrastructures Towards Implementation and Long-Term Sustainability, June 2020. Available at: https://ec.europa.eu/info/sites/default/files/research_and_innovation/strategy_on_research_and_innovation/documents/ec_rtd_transformative-impact-ris-on-euro-research.pdf

EC/European Commission (2020). 'Study on the economic detriment from unfair and unbalanced cloud computing contracts. Deloitte et al, Impact Assessment on enhancing the use of data in Europe Report on Task 1 – Data governance (SMART 2020/694 | D2)'.

EC/European Commission (2019). Joint Research Centre, M. Iglesias, N. Malaquias, P. Triaille, e.a., 'Reuse policy: a study on available reuse implementing instruments and licensing considerations', Publications Office 2019. Available at : https://op.europa.eu/en/publication-detail/-/publication/971c48e2-51e3-11e9-a8ed-01aa75ed71a1/language-en

EC/European Commission (2018). 'Study on the switching of cloud service providers, which included a cross-sectorial workshop in Q2 2017'.

EC/European Commission, 'Open Science', https://ec.europa.eu/info/research-and-innovation/strategy/strategy- 2020-2024/our-digital-future/open-science_en#8-ambitions-of-the-eus-open-science-policy. Accessed 25 March 2022.

EC/European Commission, 'European Open Science Cloud', https://digital-strategy.ec.europa.eu/en/policies/open-science-cloud. Accessed 25 March 2022.

EC/European Commission, 'European Open Science Cloud (EOSC) - What the cloud is, how it was developed and being implemented', https://ec.europa.eu/info/research-and-innovation/strategy/strategy-2020- 2024/our-digital-future/open-science/european-open-science-cloud-eosc_en. Accessed 25 March 2022.

EC/European Commission Press Corner, Daily News 30/09/2021, https://ec.europa.eu/commission/presscorner/detail/en/mex_21_4962. Accessed 25 March 2022.

EC/European Commission, 'Reforming and enhancing the European Research and Innovation system: Policy, strategy, how to apply and work programmes', https://ec.europa.eu/info/research-and- innovation/funding/funding-opportunities/funding-programmes-and-open-calls/horizon- europe/reforming-and-enhancing-european-research-and-innovation-system_en. Accessed 28 March 2022.

EC/European Commission, Frequently asked questions on the European Research Area, Memo 17 July 2012, https://ec.europa.eu/commission/presscorner/detail/en/MEMO_12_564. Accessed 28 March 2022.

EOSC/European Open Science Cloud, 'Members', https://eosc.eu/members. Accessed 25 March 2022.

European Parliament (2011). 'Report on the proposal for a directive of the European Parliament and of the Council amending Directive 2003/98/EC on re-use of public sector information', (COM(2011)0877 – C7-0502/2011 – 2011/0430(COD)), 07.12.2012, A7-0404/2012.

EC/European Commission. Study on provisions applying to B2G data sharing for the purpose of the production of statistics in the context of the European Statistical System (B2G4S).

European Parliamentary Research Service (2019). 'Briefing EU Legislation in Progress: Re-use of public sector information', July 2019. Available at: https://www.europarl.europa.eu/thinktank/en/document/EPRS_BRI(2018)628312

European Research Infrastructures in the International Landscape (RISCAPE) (2019). 'International Research Infrastructure Landscape 2019 – A European Perspective', p. 90, 94-97, DOI 10.5281/zenodo.35392.

European Science Foundation (2012). 'Evaluation in Research and Research Funding Organisations – European Practices', November 2012.

M. van Eechoud, 'A Serpent Eating Its Tail: The Database Directive Meets the Open Data Directive', ICC 2021, No. 52.

Jörg Hoffmann, Begoña Otero, 'Demystifying The Role Of Data Interoperability In The Access And Sharing Debate', 11 (2021) JIPITEC 252.

JIIP et al., Study in Support of the Evaluation of Directive 96/9/EC on the Legal Protection of Databases Final Report (EC 2018). Available at: https://op.europa.eu/en/publication-detail/-/publication/5e9c7a51-597c-11e8-ab41-01aa75ed71a1

S. Jones, 'EOSC in an international context', https://www.eoscsecretariat.eu/news-opinion/eosc-international- context. Accessed 25 March 2022.

Karanikolova, K., Chicot, J., Gkogka, A., et al., Study in support of the evaluation of Directive 96/9/EC on the legal protection of databases: final report, Publications Office, 2018, https://data.europa.eu/doi/10.2759/04895.

P. Keller e.a., 'Re-use of public sector information in cultural heritage institutions', International Free and Open Source Software Law Review, 2014, Vol. 6, No. 1.

Y. Ménière, Fair, Reasonable and Non-Discriminatory (FRAIND) Licensing Terms: Research Analysis of a Controversial Concept, Publications Office 2015.

Open Data Handbook, 'Open Data', https://opendatahandbook.org/glossary/en/terms/open-data/. Accessed 25 March 2022.

Open Knowledge Foundation, 'Open Definition – defining open in open data, open content and open knowledge', https://opendefinition.org/. Accessed 25 March 2022.

Science Europe (2012). 'Joint Statement on the Revision of the PSI Directive', https://www.scienceeurope.org/media/q1hdge1r/20181112_joint-statement-on-the-revision-of-the-psi- directive.pdf. Accessed 25 March 2022.

M. D. Wilkinson, M. Dumontier, I. Aalbersberg e.a., 'The FAIR Guiding Principles for scientific data management and stewardship', Sci Data 3, 160018 (2016).

Getting in touch with the EU

IN PERSON
All over the European Union there are hundreds of Europe Direct information centres.
You can find the address of the centre nearest you at: https://europa.eu/european-union/contact_en

ON THE PHONE OR BY EMAIL
Europe Direct is a service that answers your questions about the European Union.
You can contact this service:
– by freephone: 00 800 6 7 8 9 10 11 (certain operators may charge for these calls),
– at the following standard number: +32 22999696, or
– by email via: https://europa.eu/european-union/contact_en

Finding information about the EU

ONLINE
Information about the European Union in all the official languages of the EU is available on the Europa website at: https://europa.eu/european-union/index_en

EU PUBLICATIONS
You can download or order free and priced EU publications from:
https://op.europa.eu/en/publications. Multiple copies of free publications may be obtained by contacting Europe Direct or your local information centre (see https://europa.eu/european-union/contact_en).

EU LAW AND RELATED DOCUMENTS
For access to legal information from the EU, including all EU law since 1952 in all the official language versions, go to EUR-Lex at: http://eur-lex.europa.eu

OPEN DATA FROM THE EU
The EU Open Data Portal (http://data.europa.eu/euodp/en) provides access to datasets from the EU. Data can be downloaded and reused for free, for both commercial and non-commercial purposes.

This study analyses the possible impact of three major legislative instruments in the European Sstrategy for data (the Open Data Directive, the Data Governance Act and the proposed Data Act) for the field of research, especially for research performing organisations and research funding organisations. It does so against the background of the European Open Science policy pursued, in which the development of a European Open Science Cloud (EOSC) is a major undertaking. Although the impact is difficult to assess at this stage, the study identifies and makes recommendations about key legal issues that need to be resolved. These have to do with ambiguities in the scope of application to research data, the interpretation of provisions, and the consistency between the instruments from the perspective of open science research policy.

Studies and reports

Publications Office
of the European Union