Rat für Informations Infrastrukturen

March 2019

Statement of the Council for Scientific Information Infrastructures (RfII) on current developments concerning Open Data and Open Access

PREFACE

The Council for Scientific Information Infrastructures (RfII) was established in 2014 to advise on the development of a contemporary and sustainable infrastructure for access to scientific information in Germany. RfII recommendations are directed at academia and at the Federal Government and the governments of the *Länder*. They also serve to support the development and communication of German positions in international debates.

In this context, the RfII has discussed current European efforts to give new political impetus to the "transition to openness" proclaimed some 20 years ago. This includes the founding of the European Open Science Cloud (EOSC) and related regulatory initiatives of the European Commission[1] as well as initiatives of Science Europe, the association of European Research Funding Organisations.[2]

The transition to openness in digital policy offers opportunities to science. In 2016 the RfII acknowledged the open science paradigm and the FAIR principles as a basis for the re-usability of research data, but without underestimating the issues and challenges they raise.[3] In this respect, the RfII unequivocally welcomes the push towards open science. However, the RfII also reiterates that the prerequisites for gaining added value from freely available resources and services need to be examined, assessed, and discussed in more detail as well as consciously designed.

In particular, it must be a key concern of science to practice openness in combination with quality assurance, and therefore to ensure the data to be used is of high quality. For the RfII an understanding of quality based on scientific methods is an essential prerequisite for realising the high expectations associated with open access to scientific results and data.

In 2019, the RfII will deal in greater depth with the issues of "quality of data" and "re-use and valorisation of scientific outputs". In view of the new regulations at European level, which will also need to be implemented in Germany, the RfII already points out some essential points.

---

[1] EC (2018) – Proposal for a Directive; EC (2018) – Commission Recommendation EU 2018/790; Legislative Observatory (2019) – Copyright in the Digital Single Market (website).

[2] Science Europe (2018) – cOAlition S (press release).

[3] Cf. RfII (2016) – Enhancing Research Data Management, Stance on fundamental issues, Ch. 3 and Recommendations, section 4.1. and 4.9.

1 RESEARCH DATA AS PUBLIC SECTOR INFORMATION

In its new PSI Directive[4] (Directive on Open Data and Public Sector Information) the European Union considers research data to be public sector data, and therefore re-usable in principle in a non-discriminatory manner and according to transparent conditions set out in the Directive. Specifically the PSI Directive aims to enable publicly-funded research data that is already accessible to the public in repositories to be re-used, and in particular to be re-used commercially. Furthermore, the Directive also requires the member states to develop national strategies for enabling open access to research data.

The RfII recognises the intention of achieving harmonisation between research data repositories, which have become quite common services throughout Europe. However, it must be taken into consideration that resulting data collections are often closely linked to a particular research question and the specific scientific methods applied. In addition, research not only collects primary data and numerous secondary products, but also yields data on the research process itself. These can be explicitly assigned "metadata" or automatically generated digital footprints, for example. Furthermore, in the humanities, documents and publications are also considered to be “research data”. Even with good documentation, many of these data require a well-founded, scientific evaluation of their usability/usefulness if they are to be analysed or used beyond their context of origin. In Germany, the National Research Data Infrastructure (NFDI) is currently opening up avenues for the development and scientific use of data that are particularly suited to the typology of research data. The intention is to form partnerships between data creators and data users in order to develop the best possible services for science's own needs. Within this framework, data products for business and society can also be made available that are quality-assured and suitable for re-use by parties from other (potentially non-scientific) domains.

From the point of view of the RfII, the Federal German strategy for free access to research data called for by the EU should focus on scientifically adequate models of data sharing and, instead of quantitative growth, above all pursue the qualitative goal of high-quality data collections and services. It is clear that researchers have to work as transparently and verifiably as possible, and naturally have to publish their findings. However, general and indiscriminate “obligations to publish” – including for example the publication of all interim results produced along the way to a final result – do not meet academic performance standards or fulfil the responsibility for scientific quality.[5]

[4] Public Sector Information (PSI).

[5] Regarding prerequisites for added value, cf. section 4 of this paper.

2 ON THE COMMERCIAL RE-USE OF RESEARCH DATA

Article 10 of the PSI Directive states that the research data included in the scope of the Directive can be re-used for commercial und and non-commercial purposes. The implementation of the Directive is therefore intended to increase the potential for innovation for the benefit of the economy and society.

The PSI Directive clearly shifts the open data discourse within the scientific community into the context of a science and technology transfer, i.e. the desired valorisation of scientific results for the economy and society. This type of use requires, even more urgently than before, that researchers as well as their research facilities make conscious decisions regarding the publication and licensing of data. At the level of science policy, the terms "sharing" and "using" of research data are not equivalent. The RfII has already pointed out the disadvantages that can arise from an uncontrolled or unwanted outflow of research data and has spoken out in favour of active development of the interface between science and the economy.[6] A recently published brief assessment by legal experts illustrates the large regulatory gaps, e.g. regarding the power of decision over the data produced in a collaborative project. Legal risks can arise from violations of the General Data Protection Regulation, for example.[7] From the perspective of researchers, risks emerge from the effect that the willingness to participate in surveys/scientific surveys declines and the trust of test subjects changes when data are not used solely for scientific purposes. This effect may seriously distort study results. In this regard, the RfII welcomes provisions in the PSI Directive allowing to restrict openness in certain cases, which reflects the needs of science in particular. Nevertheless, it must be emphasised that re-use of data in the sense of the PSI Directive is a form of knowledge and technology transfer that extends the scope of scientific publication and data sharing. The PSI Directive also stipulates that data must be provided for free even when used for commercial purposes. If research organisations expect revenue from the transfer of knowledge and technology, then it is also necessary to create rules pertaining to the costs.

The EU has left the implementation of access policies for research data from the public sector in the hands of the member states. The concerns regarding indiscriminate duties to publish should be obviated as soon as possible. The decision whether or not to publish research data needs to remain with researchers and their institutions. The RfII recommends that commercial demand for published data should be documented in a suitable manner in order to showcase the value to society that has been created. In view of the investments in artificial intelligence currently being discussed and the large amounts of data required for this purpose, political strategies should also consider coordinated public-private efforts to develop the large collections of data needed.

[6] Cf. RfII (2016) – Enhancing Research Data Management, Recommendations, section 4.9.

[7] Cf. Expert report on the legal framework for research data management: Lauber-Rönsberg et al. (2018) – Rechtliche Rahmenbedingungen FDM (German only).

3 ON OPEN ACCESS TO SCIENTIFIC PUBLICATIONS

With its stronger support for free access to scientific publications, science policy has taken up a goal formulated by the scientific community itself.[8] This is due to the fact that expected cost reduction effects for scientific literature have not been fulfilled to date, that commercial license models in the digital sector often constrain research and that the criticised excessive profit margins of some publishers remain unchanged.

Within the framework of the German DEAL project, a comprehensive licensing and financing model is currently being negotiated to achieve Open Access objectives. It intends to provide researchers in Germany with full and unlimited access to the respective publisher's portfolio as well as continuous activation of their publications in Open Access.[9] The RfII is following with interest the first publisher agreement signed and the progress of negotiations with the other publishers, hoping to see a positive impact in the overall system. This also applies to the German Research Foundation's (DFG) funding line for the transformation of agreements, with small and medium-sized publishers. The RfII sees one strength of the selected approaches in the fact that they give scientists the freedom to choose the medium to which they want to submit their publications.[10]

In view of the RfII, the national strategies and action plans for open access required by the EU should above all promote structural changes and reduce financial burdens to scientific institutions. The RfII also advocates further efforts to counteract the advanced monopoly formation in the field of academic publishing. The large scientific societies and their journals have an important responsibility in in this regard. Furthermore, science, grant providers, and science policy must work consistently towards changing the scientific reputation systems in order to overcome the negative impact of excesses in the current publication system.

Regarding open access to publications, the RfII commends the success of the negotiations on the European Copyright Directive leading to a harmonisation of rights to mine text and data from information published in the Internet for scientific purposes throughout Europe.[11] The Directive also formulates a harmonised legal framework for the use of text and data mining for commercial purposes.[12] Both will foster the application of machine learning technologies and innovative methods of data analysis.

In 2019, the RfII will deal again in greater depth with the transformation of the scientific publication system towards Open Access.

[8] The OA2020 initiative, a global alliance of scientific organisations, has been pursuing the transformation of the publications market since 2015, cf. Max Planck Digital Library (2018) – OA2020 (website).

[9] Publish and Read Model; https://www.projekt-deal.de/about-deal/.

[10] The restriction imposed on the freedom of choice is a major point of criticism expressed by scientific stakeholders on the transitional model recently proposed by cOAlition S, a consortium of European research funders ("Plan S").

[11] Already regulated in German law in Section 60d of the Act on Copyright and Related Rights.

[12] Cf. EC (2019) – Digital Single Market (press release).

4 ON THE PREREQUISITES FOR REALISING THE ADDED VALUE OF OPENNESS

The expected benefits associated with open access to data and publications include the more rapid development of science and stimuli for economic growth. Both are only possible when the data and publications are of high quality, i.e., they have gone through standardised quality assurance procedures. Also, the term "open" is often incorrectly equated with the term "online" (meaning only put on the Internet regardless of whether or not it is re-usable). The implementation of FAIR principles[13] is an important step to achieving genuine "openness" of data, which also includes machine readability.

The realisation of added value from open data originating in the scientific system is also highly dependent on which (possibly new) quality assurance pathways are found for these forms of publication. The quality of scientific data made available openly by publicly funded research can only be assured based on scientific and domain-specific standards. In light of the "oceans of data" already existing, strategies are needed for researchers to determine which data should be published and which conditions and requirements should apply to its publication. The RfII has strongly recommended the development of data services for the NFDI in joint responsibility with scientific users.[14] Strategies for open access to data must recognise these insights and – in addition to the "FAIRness" of data – emphasise in particular the requirements for scientific quality.

The RfII will publish an extensive position paper on the topic of data quality in 2019.

5 ON THE ROLE OF SPONSORS AND RESEARCH FUNDERS

The current trends indicate a strong willingness to support "openness" in science through specific and even binding measures. The RfII proposes to always carefully examine which undesirable side effects may also be associated with decisions. This applies in particular to the commercial re-use of research data. Good strategies are needed to provide all parties involved with the necessary confidence to organise the sharing and valorisation of data well.

In Germany, the NFDI can contribute to the establishment of good practices and to the creation of reliable and sustainable services in cooperation with trusted service providers. The NFDI will also – especially with regard to the EOSC – help ensure that these services and capabilities align internationally. In terms of science policy, it will be important to set good incentives that prevent negative effects or misdevelopments. Potential misdevelopments include in particular:

[13] The four principles for the provision of data that were formulated specifically with improved machine readability in mind are: Findable, Accessible, Interoperable, Reusable (https://www.force11.org/group/fairgroup/fairprinciples).

[14] RfII (2018) – Diskussionspapier: In der Breite und forschungsnah.

- Inadequately or unsustainably maintained service offers such as portals or apps that arise when good, regularly updated usage scenarios for the offers or long-term digital archiving capabilities are lacking[15];
- Making unusable information available under the banner of "openness" instead of implementing more important and scientifically appropriate measures (e.g. the establishment of professional data management);
- Market trends that make supplying information more expensive, e.g. exploding publication fees or other business models in which data, once it has been disclosed, is privatised, thus eliminating the ability of science to re-use the data or resulting in the data being hidden behind paywalls.

6 ON FURTHER DEVELOPMENTS

Researchers and their facilities are under pressure due to the rapid political and regulatory developments described here. In terms of the re-use of data in particular, regulations and binding agreements that would provide everyone involved with confidence to act are lacking. Assuming that even more far-reaching strategies for open research data will continue to spread throughout Europe and the world, little time is left for research organisations to professionalise their data management and grapple with effective data governance. To achieve this, a dialogue between researchers and their facilities is needed as well as the support of their sponsoring organisations and supervisory bodies.

The developments described also set the course for the EOSC. At the level of working groups and scientific, technological and innovation policy committees decisions are being made that will have a large impact on the ability of providers, creators, and users of research data to participate in the EOSC. Germany is currently investing considerable effort in the coordinating needs and investments at the level of the Federal Government and the *Länder* as well as at the level of universities and research organisations. In Germany, which is the largest European member state in regard to science, stakeholders have a special responsibility to participate actively and in a coordinated fashion in shaping the EOSC, especially in light of the opportunities associated with the NFDI on the national level. The RfII will also continue to play its part.

[15] "Ghost towns of abandoned pilots, outdated data portals and unused apps", The World Wide Web Foundation (2016) – ODB Global Report, p. 9.

BIBLIOGRAPHY

EC - European Commission (2018): Commission Recommendation (EU) 2018/790 on Access to and Preservation of Scientific Information, in: Official Journal of the European Union, L 134/12-L134/18, available online at: https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32018H0790&from=DE, last checked on: 09.05.2019.

EC - European Commission (2018): Proposal for a Directive of the European Parliament and of the Council on the Re-Use of Public Sector Information (recast), Brussels, 45 p., available online at: https://eur-lex.europa.eu/resource.html?uri=cellar:4e790e4c-4969-11e8-be1d-01aa75ed71a1.0001.02/DOC_1&format=PDF, last checked on: 09.05.2019.

EC - European Commission (2019): Digital Single Market: EU Negotiators Reach a Breakthrough to Modernise Copyright Rules, Strasbourg, available online at: http://europa.eu/rapid/press-release_IP-19-528_en.htm, last checked on: 09.05.2019.

Lauber-Rönsberg, Anne/Krahn, Philipp/Baumann, Paul (2018): Gutachten zu den rechtlichen Rahmenbedingungen des Forschungsdatenmanagements, 23 p., available online at: https://tu-dresden.de/gsw/jura/igetem/jfbimd13/ressourcen/dateien/publikationen/DataJus_Zusammenfassung_Gutachten_12-07-18.pdf?lang=de, last checked on: 09.05.2019.

Legislative Observatory (Hg.) (2019): Copyright in the Digital Single Market 2016/0280(COD), (website), available online at: https://oeil.secure.europarl.europa.eu/oeil/popups/ficheprocedure.do?lang=&reference=2016/0280(COD), last checked on: 09.05.2019.

Max Planck Digital Library (2018): OA2020. Learn More (website), available online at: https://oa2020.org/learn_more/, last checked on: 09.05.2019.

RfII – German Council for Scientific Information Infrastructures (2016): Enhancing Research Data Management: Performance through Diversity. Recommendations regarding structures, processes, and financing for research in data management in Germany, Göttingen, 92 p., available online at: https://d-nb.info/1121685978/34, last checked on: 09.05.2019.

RfII - Rat für Informationsinfrastrukturen (2018): In der Breite und forschungsnah: Handlungsfähige Konsortien. Dritter Diskussionsimpuls zur Ausgestaltung einer Nationalen Forschungsdateninfrastruktur (NFDI) für die Wissenschaft in Deutschland, RfII - Rat für Informationsinfrastrukturen, Göttingen, 6 p., available online at: https://d-nb.info/1172854858/34, last checked on: 09.05.2019.

Science Europe (2018): cOAlition S: Making Open Access a Reality by 2020, Brussels, 2 p., available online at: https://www.scienceeurope.org/wp-content/uploads/2018/09/cOAlitionS_Press_Release.pdf, last checked on: 09.05.2019.

The World Wide Web Foundation (Hg.) (2016): Open Data Barometer. ODB Global Report. Third Edition, n.p., 48 p., available online at: https://opendatabarometer.org/doc/3rdEdition/ODB-3rdEdition-GlobalReport.pdf, last checked on: 09.05.2019.

Imprint

Adopted at the 14th Council meeting of the RfII on 8.03.2019

Rat für Informationsinfrastrukturen (RfII) – Head Office
Papendiek 16, 37073 Göttingen
Germany
Phone +49 0551-3927050
E-mail info@rfii.de
Web www.rfii.de