Funding Information Infrastructures for Research

A DFG Strategy Paper

Bonn, 15 March 2018

Deutsche Forschungsgemeinschaft
Kennedyallee 40 · 53175 Bonn / Germany · postal address: 53170 Bonn
phone: + 49 228 885-1 · fax: + 49 228 885-2777 · postmaster@dfg.de · www.dfg.de

Funding Information Infrastructures for Research

Executive Summary

This strategy paper discusses the funding activities of the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) in the area of scientific information infrastructures. It has two functions. First, it reflects on the basic structural framework of DFG funding activities in relation to information infrastructures in the sciences and humanities – both at an overarching level and with regard to three selected topics – and discusses the necessary changes based on the current situation. It also identifies specific measures as a response to the analysed challenges and action required. Second, it thus provides a guideline for funding activities in the area of Scientific Library Services and Information Systems (LIS).

As the self-governing organisation representing the interests of the research community in Germany, the DFG believes that it has a responsibility to actively participate in shaping the digital transformation in the sciences and humanities. Funding in the area of information infrastructures aims to respond to requirements articulated by the research community, to keep pace with constantly changing dynamics, to be open to unconventional project ideas and projects at the exploratory stage, to provide impetus and to stimulate the development of structures and standards.

In **Section A**, the funding activities of the DFG are analysed according to these criteria, and, at a broader level, three topic areas are identified on which the DFG, or more specifically the relevant body of the Joint Committee, the Committee on Scientific Library Services and Information Systems (AWBI), will focus over the next several years. Issues requiring clarification include:

- how the impacts of the digital transformation on work practices in research – for instance in relation to the creation of standards or the introduction of policies – can be better integrated into funding activities
- how the self-organisation processes essential to infrastructure funding can be supported and stimulated both within individual disciplines and between infrastructure facilities
- how a productive discussion can be initiated with a view to developing long-term funding prospects for nationally relevant infrastructures established using project funds, such as specialised information services.

Section B contains a more in-depth examination of the three core topics of Indexing and Digitisation, Open Access Transformation and Research Data. These are the areas of funding in which, in the DFG's view, there is currently the greatest need for action. Specific measures are described for all three areas. They reflect, for the individual funding areas as well, the DFG's

aim of responding to needs articulated by the research community, of participating in shaping the digital transformation and, through funding, of providing impetus and stimulating the development of structures and standards.

In the **Indexing and Digitisation** programme (B.I.), the current limitation of funding to manuscripts and printed materials should be lifted to open up the programme to all objects of potential scholarly interest. To lay the foundations for this opening up of the programme, funding will initially concentrate on

- developing interoperable minimum data sets
- promoting the standardisation of indexing and digitisation processes for multidimensional objects.

In terms of **Open Access Transformation** (B.II.), the DFG sees the greatest need for action in the areas of funding, monitoring and policy. The DFG will therefore

- review its practice of fixed-rate publication fees with a view to open access and if appropriate link it to the further development of institutional support for literature provision (evaluation of the Open Access Publishing and Licences for Digital Content programmes)
- establish monitoring mechanisms to obtain more accurate data about publication output from DFG-funded projects and the associated costs
- adapt its open access policy such that the DFG *requests* researchers in receipt of DFG funding to publish on an open access basis.

The key areas to be supported through targeted initiatives in the area of **Research Data** (B.III.) in the years ahead include policies and regulations, digital skills and the federation of infrastructures. Specifically, it aims

- to support the further development and negotiation of subject-specific policies and regulations – focusing on research communities with a lower degree of organisation – and to reflect on whether the request to develop and use data management plans in project proposals to the DFG should be made more explicit
- to develop funding options to improve research data management skills
- to design funding initiatives for the coordination and networking of existing community-specific infrastructures to ensure the connectivity and interoperability of these different systems.

In this strategy paper, the DFG reflects for a third time – following on from 2006 and 2012 – on the ramifications of the digital transformation for its funding activities in the area of scientific

information infrastructures.[1] The paper was drawn up by the Committee on Scientific Library Services and Information Systems (AWBI), the body responsible for the funding of this type of infrastructure.[2] As part of its remit as a body of the Joint Committee, the AWBI is responsible for designing the funding portfolio and individual funding initiatives relating to scientific information infrastructures.

[1] See DFG strategy paper (2006): Scientific Library Services and Information Systems: Funding Priorities Through 2015 http://www.dfg.de/download/pdf/foerderung/programme/lis/pos_papier_funding_priorities_2015_en.pdf (DFG 2006); DFG strategy paper (2012): Taking Digital Transformation to the Next Level – The Contribution of the DFG to an Innovative Information Infrastructure for Research http://www.dfg.de/download/pdf/foerderung/programme/lis/strategy_paper_digital_transformation.pdf (DFG 2012).

[2] For information about the AWBI http://www.dfg.de/en/research_funding/programmes/infrastructure/lis/awbi.

A. DFG Funding for Scientific Information Infrastructure

An effective, modern and secure information infrastructure is essential to excellent research. In view of this, infrastructure funding has always been a fixed component of the funding offered by the DFG. From the beginning, the purpose of this funding has been to create a coordinated system of information infrastructures for research in partnership with other stakeholders.

With the advent of digital technology, the requirements and basic framework relating to information infrastructures in the sciences and humanities have changed enormously. This is revealed by countless analyses,[3] as well as a large number of national and international initiatives, all designed to create an adequate framework for digital information infrastructures. These include the recommendations to establish federated data infrastructures and legislative efforts to legally safeguard the use of text and data mining, to name just two examples.

These changes – which have only been touched on here – impact on DFG funding activities in the area of scientific information infrastructures. These activities are initiated and take place within a complex structure influenced by the DFG's aim of functioning as the self-governing organisation of the German research community, by the conditions of project-based funding, and by specific requirements relating to the largely systematic and coordinated funding of digital information infrastructures.

In line with the definition used by the German Council for Scientific Information Infrastructures (RfII), scientific information infrastructures are understood as "technically and organisationally networked services and offerings for accessing and preserving data, information and knowledge. The RfII defines them as primarily serving research purposes; they are frequently an object of research and always have an enabling function."[4] The DFG's funding activities are focused on the benefits of information infrastructures for researchers.

---

[3] Among others, Gesamtkonzept für die Informationsinfrastruktur in Deutschland. Empfehlungen der Kommission Zukunft der Informationsinfrastruktur, [Concept for Information Infrastructure in Germany. Recommendations of the Future of Information Infrastructure Committee], April 2011 (available only in German); Empfehlungen des Wissenschaftsrates zur Weiterentwicklung der wissenschaftlichen Informationsinfrastrukturen in Deutschland bis 2020 [Recommendations of the German Council of Science and Humanities on the Further Development of Scientific Information Infrastructures in Germany Between Now and 2020] (Drs. 2359-12), July 2012 (available only in German); Enhancing Research Data Management: Performance through Diversity. Recommendations regarding structures, processes, and financing for research data management in Germany, Göttingen 2016 www.rfii.de (RfII 2016); KE Report "Knowledge Exchange approach towards Open Scholarship", https://repository.jisc.ac.uk/6685/1/KE_APPROACH_TOWARDS_OPEN_SCHOLARSHIP_AUG_2017.pdf; Europe's Future: Open Innovation, Open Science, Open to the World. Reflections of the RISE Group. European Commission 2017 (doi:10.2777/79895).

[4] RfII 2016, p. A-15.

1. Basic Framework and Types of Funding

Objectives and funding conditions

Through its Scientific Library Services and Information Systems programme, the DFG funds projects at research institutions and other research service and information centres in Germany. This funding helps to create optimum conditions for the provision, dissemination and processing of information for research purposes. The aim is to provide users with free, comprehensive access to analogue and digital research information and to link up knowledge and data. As noted in the DFG strategy paper of 2012, "Taking Digital Transformation to the Next Level", this also requires a suitable legal, technical, financial and organisational framework that enables optimum scientific work to take place. The establishment of a coordinated system of information infrastructures for research is a dynamic process in which the technical options and conditions pertaining to research work and usage requirements are mutually dependent, as well as being subject to constant change and the need for regular adaptation.[5]

The funding is characterised by the aim of achieving improvements in information provision across locations and thus better basic conditions for research. In other words, funding should not solely benefit the location where a particular information infrastructure is operated; its impact should extend beyond it. Hence, in large part this type of funding also constitutes structural funding. The funding achieves its effect by providing targeted incentives to take up emerging structures at a variety of institutions or in different research communities and further develop them.

Another objective of this type of funding is to provide impetus, that is, to support research institutions in responding to the permanent change processes to which scientific information infrastructures are exposed, and not just through the digital transformation. Scientific information provision has always been subject to a constant process of adaptation, both to a changing historical and technical framework and to the needs and expectations of research users. The DFG has always seen its role as supporting this adaptation process, providing additional resources to enable the system as a whole to take up innovations and put them to good use in the interests of research. In view of this, it is possible to define the criteria for the funding of scientific information infrastructures by the DFG. These are listed in Overview 1.

[5] DFG 2012, p. 2.

Overview 1: Funding criteria

The funding of scientific information infrastructures by the DFG

- is governed by the *response mode* principle, i.e. funding responds to a need articulated by the research community.
- is explicitly nationwide in its aim, i.e. the results should not just benefit those carrying out the project; instead, the focus is on the benefit, which extends beyond local or individual solutions to problems.
- relates on the one hand to projects at an exploratory stage, i.e. funding is available for projects in which novel solutions to current challenges are being developed, explored and implemented by way of example.
- on the other hand, aims to stimulate the development of structures, i.e. funding is available for projects involving the establishment, optimisation and consolidation of information infrastructures.
- provides incentives to tackle challenges identified as new and to develop solutions.
- is quality-oriented additional funding intended both to further develop the knowledge base through innovative and user-friendly information infrastructures and to stimulate and support the formation of new structures. This enables research institutions to address, develop and test new fields and topics in addition to their ”routine business”.
- is designed to complement the basic tasks of the institutions. All DFG grants therefore presuppose stable basic funding for the institution.

Funding is offered to applicants from scientific information infrastructure facilities such as libraries, archives, museums, and computing, data and media centres, and also to individual researchers.

The current funding portfolio: programmes and calls

The DFG funding portfolio for Scientific Library Services and Information Systems currently consists of seven funding programmes. Broadly speaking, these programmes can be divided into two groups:

- first, programmes designed to be open and thus offer a funding option for all types of projects, which may relate to various information provision subtasks or different development phases of information infrastructures. They also offer scope for projects that address unconventional and fundamentally new questions or innovations relating to information infrastructures.[6]

[6] Currently two programmes fall within this category: Infrastructure for Electronic Publications and Digital Scholarly Communication, and e-Research Technologies.

– second, programmes which are geared as a whole – meaning all projects funded through the programme – towards a specific goal, or which have a strategic focus and are designed to achieve a clearly defined objective of creating structures or standards.[7]

The content framework within which funding is available encompasses all areas of information provision and can be divided into phases covering the development, implementation and consolidation of information infrastructures. The following matrix diagram (Overview 2) illustrates this framework. It schematically links the process of developing and establishing infrastructures (information infrastructure development phase) with important information provision tasks, thus illustrating the maximum scope of DFG funding for infrastructure projects. DFG funding clearly reaches its limit at the point where attention shifts to the establishment of permanent operation and maintenance of information infrastructures.

[7] Currently five programmes fall within this category: Specialised Information Services; Licences for Digital Content; Indexing and Digitisation; Open Access Publishing; and Information Infrastructures for Research Data.

Overview 2: Stages of development of information infrastructures

Please note: Each column contains only one illustrative example

Information provision tasks	Research & development	Implementation & networking	Consolidation	Permanent operation
Generate & procure information	– *Develop pilot projects on new licence models* – *(…)*	– *Successfully apply licence models* – *(…)*	– *Integrate models into a new or existing programme* – *(…)*	
Make information accessible & usable	– *Automatic indexing* – *(…)*	– *Make automatic indexing tools ready for use (e.g. recommender, visualisation)* – *(…)*	– *Integrate search technique into "n" portals* – *(…)*	
Provide & disseminate information	– *Develop hosting models* – *(…)*	– *Implement pilot applications* – *(…)*	– *Establish hosting structure* – *(…)*	
Safeguard information long-term	– *Research / development on open archiving tools* – *(…)*	– *Successfully apply archiving tools and make them more effective by means of networking* – *(…)*	– *Establishment of a national archiving structure* – *(…)*	

In all funding programmes, calls for proposals can be formulated to specify certain focal areas, stimulate developments and position specific topics. Fifteen calls have been published since 2012, nearly all of them in line with the key topics identified in the 2012 strategy paper. Since 2012, three pilot measures have also been implemented in the area of Indexing and Digitisation. In all programmes it is also possible to fund projects that contribute to the self-organisation of a community by promoting networking and infrastructure-related professionalisation, as well as projects for networking and coordination at the international level.

Overview 3: Calls between 2012 and 2017

Specialised Information Services
- Establishment of Competence Centres for the Licensing of Electronic Resources – A Cross-Disciplinary Task in the DFG-Funded System of Specialised Information Services (2013)
- Outstanding Research Libraries (2012)

Licences for Digital Content
- Open Access Transformation Contracts (2017)

Indexing and Digitisation
- Digitisation of Archival Sources (2016)
- Digitisation and Indexing of Printed Material Published in the 18th Century in German-Speaking Areas (VD 18) (2013 and 2014)
- Creation of Standards for the Indexing and/or Digitisation of Object Classes in Scientific Collections (2013)
- Pilot Projects for the Digitisation of a. Historical Newspapers, b. Archival Sources and c. Medieval Manuscripts (2012)

Information Infrastructures for Research Data
- Research Data in Practice (2015)

Infrastructure for Electronic Publications and Digital Scholarly Communication
- Open Access Transformation (2014)
- Scientific Monographs and Monograph Series in Open Access (2013)
- National Hosting (2012)

e-Research Technologies
- Scalable Techniques of Text and Structure Recognition for the Full-Text Digitisation of Historical Printed Material (2017)
- Research Software Sustainability (2016)
- Coordinated Funding Initiative for the Further Development of Optical Character Recognition (OCR) Processes (2014)
- NEH/DFG Bilateral Digital Humanities Programme (2012 and 2014)
- Restructuring of National Information Services (2012)

2. Funding Activities in the Context of Digital Transformation in Research

The strategy paper "Taking Digital Transformation to the Next Level: The Contribution of the DFG to an Innovative Information Infrastructure for Research", prepared by the AWBI in 2012 and approved by the DFG Senate, defines a number of thematic focuses for funding which were subsequently implemented. The measures focused primarily on shaping the transformation of physically available information into digital information systems. Now, just a few years later, the starting point is different, being more noticeably characterised by genuinely digital research. Although the transformation was triggered by technical developments, it extends far beyond the technical level. Stimulated by digital technology, new areas of competence have developed which lead to different roles and responsibilities, different business models and funding requirements, and different forms of communication and reputation mechanisms. The changes are so far-reaching and lasting that neither the actual state of affairs nor the terminology used to describe it has stabilised as yet. Hence, many debates and publications amply demonstrate that linguistic negotiation processes and attempts to authoritatively define meanings are as ubiquitous as the varying terminology, including digitisation, digitality, digital science, Science 2.0 and open science.

The current situation within the research community is that, in many areas, both information provision (access, search, reception, dissemination, sharing) and research work (analysis, processing, data, software, methods, publication) are partly or entirely digital. These far-reaching changes in scientific working affect the entire research cycle.[8] At every phase of the research endeavour, we are therefore seeing needs and challenges relating to information infrastructures that are no longer shaped by the demands of a transformation, but by those of a genuinely digital work process. In addition to the familiar tasks of scientific information provision – primarily the acquisition of research publications and data, making them accessible and usable, and systematic networking structures (identifiers, standard data, standards, interfaces) – there are new tasks relating to research methods (development of methods and software, content mining) and collaborative work practices, including the exploitation of artificial intelligence for the research process. However, the availability and (algorithmic) analysis of data are becoming particularly vital to research in the digital era.

The challenges arising as the result of a genuinely digital work process affect every aspect of research – epistemic, scholarly, organisational, technical, financial, political, legal and cultural. It was against this background that the DFG Senate initiated a process in 2016 intended to establish the DFG's basic position on the digital turn in the sciences and humanities. Establishing this position will make it possible to represent the interests of the research community

[8] Example of the 21st Century Digital Scholarship Cycle: https://library.ucf.edu/about/departments/scholarly-communication/overview-research-lifecycle/.

even more effectively on all issues relating to digital transformation in the sciences and humanities – not just infrastructure.[9]

Digital transformation in research presents infrastructure facilities not only with the challenge of a considerable expansion and modification of their responsibilities, but also large-scale adaptation and negotiation processes, all interacting with each other. Digitisation makes it *possible* to use services provided locally from any location, even globally. But it also makes it *essential* to approach tasks cooperatively, because this is necessary and appropriate from both a scholarly and an economic perspective. This applies, for instance, to standardised cataloguing and media indexing systems; cross-referencing and coordinated cataloguing and, building on this, indexes of printed and digital media; an electronic resource management system; a national hosting system and long-term availability.[10]

As it is only possible to create a coordinated system of information infrastructures in partnership with many other stakeholders and responsible bodies, the DFG participates in national and international partnerships such as the Digital Information Initiative of the Alliance of Science Organisations in Germany, Knowledge Exchange, and Science Europe.[11]

The question that immediately arises as to coordination and the allocation of responsibilities among information infrastructure facilities is one of enormous relevance. This question first became the subject of science policy debate through the joint statement issued by the DFG and the German Council of Science and Humanities on the development of library networks. On the funding policy side, it was accompanied by the DFG call for proposals for the "Restructuring of National Information Systems".[12] The statement called for the replacement of the established arrangement of regionally based task allocation by a functional allocation of tasks adapted to the challenges of digital information provision. Now, seven years after the statement

---

[9] A position is to be adopted on three levels: (1) Scientific discourse: A science-driven position will be developed on opportunities and risks facing science and scholarship in the digital era. (2) Funding activities: Current activities in all DFG funding areas with a connection to the topic of digital science will be systematically identified and new funding areas and opportunities developed; (3) Policy advice: The DFG's role as the voice of the research community in national and international processes of standards creation will be examined.

[10] DFG-Positionspapier (2011) zur Weiterentwicklung der Bibliotheksverbünde als Teil einer überregionalen Informationsinfrastruktur [DFG Strategy Paper on the Development of Library Networks as Part of a National Information Infrastructure (available only in German) http://www.dfg.de/foerderung/programme/infrastruktur/lis/veroeffentlichungen (DFG 2011); Stellungnahme der Sektion 4 des Deutschen Bibliotheksverbandes (dbv) zur Neuausrichtung der überregionalen Informationservices [Statement by Section 4 of the Deutscher Bibliotheksverband (dbv) on the Restructuring of National Information Services], published on 14 October 2015 (available only in German) http://www.bibliotheksverband.de/fachgruppen/sektionen/sektion-4/publikationen (dbv Section 4 2015).

[11] For more information see http://www.allianzinitiative.de/en/, http://www.knowledge-exchange.info, http://www.scienceeurope.org.

[12] See Gemeinsame Erklärung der Deutschen Forschungsgemeinschaft und des Wissenschaftsrates zur Zukunft der Bibliotheksverbünde als Teil einer überregionalen Informationsinfrastruktur in Deutschland [Joint Declaration of the DFG and the German Council of Science and Humanities on the Future of Library Networks as Part of a National Information Infrastructure in Germany] (available only in German) http://www.dfg.de/foerderung/programme/infrastruktur/lis/veroeffentlichungen (DFG/WR 2011).

was published, this is still the desired goal.[13] At the same time, the question of coordination and the allocation of tasks – and thus a possible redefinition of role – is now raised not only for organisations such as the networks but also for the libraries themselves.[14]

The new and extensive possibilities for measuring usage are another effect of the digital transformation. Because every item of digital information is countable, the use of information services and facilities can be quantified. This data, and in particular its capacity to provide meaningful information, must be examined critically; it must not be understood solely as an indicator of success for information facilities or used to track and document individual user behaviour. From the perspective of the research community and information facilities, critical debate is therefore urgently needed on the potential consequences of quantification, tracking and, if relevant, monitoring – for example if personalised authentication, authorisation and access systems are implemented by commercial providers, and personal access and usage data are therefore collected.

3. Challenges in Funding Activities: Analysis and Proposed Actions

As the self-governing organisation representing the interests of the German research community, the DFG aims to be actively involved in the shaping of the digital transformation in the sciences and humanities. Funding activities relating to information infrastructures are designed to (a) respond to requirements articulated by the research community, (b) keep pace with the constantly changing dynamics described above, (c) be open to unconventional project ideas and projects at the exploratory stage, (d) create impetus and stimulate the development of structures and standards. The DFG can achieve these goals in part through its own actions, while in other respects it relies on partnership with other stakeholders.

The aim of responding to needs articulated by the research community requires both subject-specific differentiation and an understanding of the unique characteristics and different working practices of different disciplines. To align the DFG's funding programmes for information infrastructure more appropriately to the needs of research, it may be helpful to survey and analyse subject-specific requirements and working practices. This forms part of the internal DFG project "The Digital Turn in the Sciences and Humanities". However, surveys and analyses presuppose that the research community has already achieved a degree of self-organisation and

[13] No evaluation of the DFG call has yet been carried out.

[14] The AWBI is currently in the process of examining the impacts of digitisation on Germany's system of academic libraries and is expected to issue a statement on this in a separate publication.
See also the strategy paper Wissenschaftliche Bibliotheken 2025 [Academic Libraries 2025], available only in German), published by Section 4 (Wissenschaftliche Universalbibliotheken) of the German Library Association.

that individual disciplines are able to agree on and articulate common needs. While this process is already well advanced in some disciplines,[15] other disciplines are still very much in the early stages and are (as yet) unable to articulate their needs and expectations in relation to information infrastructures.

In view of the current situation, it would seem appropriate when designing the funding portfolio to give greater consideration to a genuinely digital research process. This would entail, for example, creating scope for the testing of novel solutions based on digital information infrastructures, promoting connectivity and the creation of standards in all programmes and giving greater consideration in funding activities to the growing importance of policies that enable specific objectives to be agreed and implementation measures to be defined.[16] Only through such an approach can the DFG achieve its aim of actively shaping the development of scientific information infrastructures. This active process also includes analysing the impact of funding programmes, calls for proposals and individual projects which may be considered to have definite structural potential of their own. Analysis of this type is essential to the identification of funding gaps and potential for the reuse of existing services and structures.

The aim of being open to projects from across the whole spectrum of scientific information provision, including unconventional and fundamentally new research questions, is covered by the e-Research Technologies programme. This programme allows applicants to propose and receive funding for projects relating to different information provision tasks and different development stages of information infrastructures. In particular, it provides a funding framework for projects designed to address current questions and problems relating to scientific information infrastructures through an investigative and experimental approach.

The aim of providing impetus and stimulating the creation of standards and structures involves more prerequisites and is therefore much more difficult to implement. It presupposes the identification of appropriate topics for which standard-setting and structural measures are judged to be necessary. Naturally enough, these are often topics – and approaches associated with them – which are not yet generally known and accepted, with the result that people must first be convinced of their relevance before implementation can go ahead. These topics must gain the approval of the specialised disciplines, on the one hand, that will benefit from the measures, and, on the other hand, of the institutions involved in proposing a project and hence developing and implementing standards and establishing structures, and last but not least the funding providers, which may be required to guarantee long-term funding. This requires a process of negotiation between the DFG and applicants as well as self-organisation processes on the part

[15] This is the case in astrophysics, biodiversity, social sciences and archaeology, for example.

[16] This is especially true with respect to research data and open access.

of applicants themselves. Forums and processes for the self-organisation of academic or infrastructure-related stakeholders in relation to these questions are not yet sufficiently established.

In this respect it is occasionally noted that libraries, in particular, lack a central coordinating body.[17] However, well-established and efficient self-organisation processes in academic and information science communities have the potential to compensate for a lack of central institutions. Compared with central institutions, which may develop agendas of their own over time, they also have the advantage that they can ensure more direct feedback to scientific users and infrastructure providers.

The question is then how it is possible to stimulate the self-organisation processes necessary for the design of a funding programme and its general acceptance. It would be appropriate to consider to what extent the Scientific Networks format established in research funding programmes could be transferred to scientific information infrastructures, if such networks are designed not only to explore new topics but also to coordinate and network existing projects or support particular communities in their professionalisation efforts. Such a format could also make it easier to pick up on impetuses and signals from the communities that are relevant from a funding perspective, or indeed identify those areas where existing offerings need to be more effectively networked.

While self-organisation is a prerequisite for achieving the creation of standards and structures through funding activities, and the DFG can influence this process at least to a certain extent, for example through suitable funding formats, the success of funding activities is also subject to other factors which are outside the DFG's province to date.

One key success factor in the establishment of a coordinated system of information infrastructures for research, and hence for the success of funding efforts, is that research institutions are both motivated and able to co-design and develop structures and services for national use and make them available on a long-term basis. This demands a willingness to cooperate and coordinate efforts. However, to achieve sustainable services which are not limited to a particular location, this also requires a set of basic conditions: organisational (to establish accepted and effective cooperation and coordination), legal (to enable cooperation and funding across different federal states) and financial (to progress from project funding to structural funding).[18]

In 2015, following on from the recommendations of the DFG and the German Council of Science and Humanities and the call issued by the DFG in 2012, the Scientific Libraries Section of the German Library Association published a statement identifying a range of services that

[17] See, most recently, Michael Knoche, Die Idee der Bibliothek und ihre Zukunft [The Idea of the Library and Its Future], Göttingen 2018, p. 90 ff. (available only in German)

[18] The AWBI is currently in the process of examining the impacts of digitisation on Germany's system of academic libraries and is expected to issue a statement on this in a separate publication.

should be offered within a nationally coordinated framework.[19] There are already initiatives and funded projects for most of the action areas listed in this statement, including electronic resource management, national hosting, long-term availability, a nationally distributed system for the archiving and availability of research data and nationwide licence procurement.[20] However, the challenge lies in transitioning from individual initiatives to a strategic process, the goal of which is to create an organisational, legal and financial framework for long-term restructuring. The central question which arises, for instance in the context of restructuring national information systems, is the long-term operation of services across the boundaries of individual federal states. This cannot be resolved by individual institutions, nor by short-term funding. It thus remains to be determined how – with the involvement of individual initiatives – a strategic process can be initiated that creates the necessary financial and legal framework to enable services provided on a "division of labour" basis to be offered at the national level with secure funding.[21]

The solution to this problem is also important for the success of DFG funding for scientific information infrastructures to the extent that the funding is specifically intended to contribute to structure creation. One specific application with regard to DFG funding is the question of permanent funding prospects for specialised information services. These are a special example of infrastructures which are provided locally but used at a national level; their continued funding after the end of the project phase must be an urgent consideration. Specialised information services complement basic information provision with services that exceed the basic mandate of an institution, for example the improvement of literature searches through innovative search functions or the central procurement of literature from regions with unstable markets. In an ideal scenario, these services would become indispensable to the areas of scholarship served by a specialised information service, thus raising the issue of reliable funding prospects which cannot, in this case, be left to local institutions alone. While the establishment of a specialised information service can be financed with DFG project funding, the problem of

[19] dbv Section 4 2015 http://www.bibliotheksverband.de/fachgruppen/sektionen/sektion-4/publikationen.html.

[20] The DFG funds projects on the topics of electronic resource management, national hosting and distributed research data infrastructures. The concept of a National Research Data Infrastructure (NFDI) is a topic of discussion in the Joint Science Conference (GWK). The national acquisition of licences is organised jointly by all German universities and research organisations as part of DEAL (a project to conclude nationwide licensing agreements for electronic journals).

[21] The German Council of Science and Humanities identified this deficit in its Empfehlungen zur Weiterentwicklung der wissenschaftlichen Informationsinfrastrukturen in Deutschland bis 2020 [Recommendations on the Further Development of Scientific Information Infrastructures in Germany Between Now and 2020] (Drs. 2359-12), July 2012 (available only in German).

its continued funding after the project-funded phase remains to be resolved. Of the 38 specialised information services[22], 21 are financed exclusively by universities and ten others are supported jointly by universities and other partners.

This example shows that universities play an important role in the non-location-dependent provision of information infrastructures. The expertise of universities in this area can only be maintained and expanded if they are able to rely on funding options that extend beyond basic university funding and project funding. Specialised information services are thus another application for the request formulated by the German Council of Science and Humanities in 2011 to the federal and state governments to find ways of "increasingly basing information infrastructures at universities or making them jointly funded by universities and non-university research institutions to avoid dysfunctions in the research system".[23]

[22] As at January 2018.

[23] German Council of Science and Humanities: Übergreifende Empfehlungen zu Informationsinfrastrukturen [Comprehensive Recommendations on Information Infrastructures] in: German Council of Science and Humanities: Empfehlungen zu Forschungsinfrastrukturen, Cologne 2011, p. 5–70, p. 42. (available only in German).

B. Funding Policy Considerations: Indexing and Digitisation – Open Access Transformation – Research Data

Indexing and Digitisation, Open Access Transformation and **Research Data** are the areas for which, from the AWBI's point of view, the need for discussion is the greatest. This does not reflect any prioritisation of these programmes compared with others in the area of scientific information infrastructures. In the sections that follow, the current funding situation for each of these three programmes is outlined, the need for action to develop the funding portfolio is analysed and starting points for specific measures are explored. The aims defined in the first section – to actively participate as the self-governing organisation in the digital transformation of the sciences and humanities while acting on needs articulated by the research community itself, to keep pace with constantly changing dynamics and to create impetus, and stimulate structures and standards through funding – are also important considerations in the examination of individual funding areas.

All other funding programmes, which are not considered in detail in this strategy paper, have either undergone recent extensive restructuring and already satisfy the requirements of digital transformation in research, or detailed road maps are in place for an in-depth and systematic examination of these programmes. The first group includes the Infrastructure for Electronic Publications and Digital Scholarly Communication programme (IePdW), introduced in 2014, which incorporated the previous programmes Electronic Publications and Scientific Journals, and the e-Research Technologies programme, launched in 2016, which incorporates both the funding options Scientific Information Management Tools and Methods and Virtual Research Environments. The second group includes the Specialised Information Services programme, which is currently being evaluated in order to provide a basis for discussing the positioning of this programme within the DFG funding portfolio from a funding policy perspective. Also in this group are the Licences for Digital Content and Open Access Publishing programmes, which will be evaluated together in 2018 and 2019.

I. Indexing and Digitisation

1. Current Funding Situation

The DFG has been providing financial support for the indexing of manuscripts and printed works since the 1950s. Since 1997, this has been supplemented by funding for the digitisation of these materials. The aim of this funding continues to be to make holdings of national importance to research visible and accessible. Since 2006, in addition to individual projects on holdings in all scholarly disciplines – with a recognisable focus on the humanities – the funding area of Indexing and Digitisation has included structure-building measures, which are accelerating the move towards the mass digitisation of texts. This includes the digitisation of titles in the lists of printed material published in the German-speaking areas in the 16th, 17th and 18th centuries (referred to as VD), the retroconversion of archival finding aids and pilot phases for the preparation of the large-scale digitisation of medieval manuscripts, historical newspapers and archival sources.

The DFG Practical Guidelines on Digitisation, first introduced in 2007 and updated several times since then, play an important role. Because they define technical standards and provide practical organisational recommendations, these guidelines have become the *de facto* standard, influencing text digitisation practice both within and beyond DFG-funded activities as well as internationally. Together with the DFG Viewer for the presentation of image files and structural data for digitised material, the Practical Guidelines have been a major step forward in the standardisation of written materials.

The extensive collections of digitised images of historical printed material and manuscripts which funding has made possible require the use of technical methods such as full text recognition to exploit their full research potential as text corpora. In 2014, the programme produced a funding initiative for the further development of full text recognition through Optical Character Recognition (OCR-D). Focusing on method development, it forms part of the e-Research Technologies programme.

In 2011 and 2013, anticipating and following the recommendations published by the German Council of Science and Humanities in 2011 on Wissenschaftliche Sammlungen als Forschungsinfrastrukturen [Scientific Collections as Research Infrastructures], available only in German), the programme was extended to include items other than printed materials and manuscripts for the first time, with two calls for the indexing and digitisation of academic collections of objects.

2. Funding Activities in the Context of Digital Transformation

Researchers have traditionally regarded the ability to access digital versions of analogue documents from any location and without time restrictions, with the associated convenience, as the greatest advantage of the digital transformation of our cultural heritage. Digital research is characterised by both the ability to consult material on the researcher's own computer screen and the processing of data[24] with the aid of software. For example, quantitative analyses, qualitative annotations, interdisciplinary networking and pattern recognition can all help to answer research questions. Just the growth in the volume of data alone makes the development and use of (partly) automated methods relevant. The large reservoir of data also makes it possible to use exploratory data analysis techniques. Research questions, hypotheses and indexing concepts can be continuously adapted as data is analysed or actually developed in the first place as a direct result of the exploration of what are usually large volumes of data. In particular, comparisons with previously validated values, indexing methods and data and the resulting application to other holdings offer considerable potential.

The existence of digital versions also affects the type and depth of indexing. For instance, the mere existence of key pages for VD 17 led to the modification of bibliographical descriptions for written materials. Depending on the nature of the object and the scholarly discipline in question, the digital version may do away with the need for more detailed descriptive metadata that is created in advance, so long as the object is capable of being adequately cited for research purposes with minimal and interoperable description and persistent addressing.

In the context of digital research, indexing should be understood not as a completed, one-off attribution of metadata by employees of the cultural institutions that hold materials, but rather as a stage within the collaborative scholarly endeavour and generation of knowledge and thus a part of the research cycle. This implies the possibility of planned and unplanned distribution among different stakeholders, each with different competences, and a decoupling of indexing from location and time limitations. The relevance of digital content indexing that is standardised and can be analysed automatically increases with the volume of data. However, indexing can also take place downstream, in an iterative process, and in part through the research community itself. It may range from the indexing of individual objects to the contextualisation of entire collections.

[24] The term "data" is used here in a broad sense. In the area of Indexing and Digitisation, it refers to both the metadata for the analogue document or object and the digital version, and the digital version itself (sometimes referred to as the "digitised item" or "digitised material" in the case of written materials). The term "metadata" in turn includes descriptive, structural, administrative and technical metadata. No distinction is made between these and research data.

In addition to the well-documented digitisation of text, where considerable progress has been made, the digitisation of multidimensional objects[25] and of sound and images, and the integration of networking with genuinely digital data are all becoming more important. But dynamic and promising as advances in information technology may be, particularly in 3D digitisation, very few scientific or academic requirements or subject-specific quality requirements have been identified. There is also a lack of standards and best practices for digitisation processes and methods, and no standard has yet become established for the documentation of existing results. Standardised, non-software-dependent data formats are rare, with proprietary formats being the norm, especially in the case of audiovisual objects. This inevitably results in a lack of options for connecting results. There is no guarantee that researchers can freely reuse the data or that it is reliable.

However, irrespective of these hurdles, scholarship does not exclude any object a priori. Furthermore, it is only by combining and networking a wide array of physical and digital objects that the possibilities of research-relevant collections can be fully exploited. The transformation of analogue collections through digitisation is thus only one step in a larger process. The option of flexible supplementation with genuinely digital objects or assembling a collection of exclusively digital objects giving due consideration to the relevant metadata represents another step. This openness reflects an inherent trait of scientific collections – the fact that they are works in progress – and provides good connectivity for digital research methods as well.

3. Action Required and Challenges

Expanding the range of objects

The current limitation of the Indexing and Digitisation funding programme to manuscripts and printed material is based on pragmatic considerations, yet it contradicts the interests and basic premise of scholarship and often reproduces the structures of the analogue world due to category- and material-specific standards. Scholarship does not exclude any object in advance. Researchers may refer to individual objects and existing collections of objects, but also turn their attention to relationships and generate networks and new collections through the exploration of available data and objects.

This need can only be met by opening up the programme for the funding of indexing and digitisation to all objects of potential relevance to research. This absolute expansion of the range of objects covered would require an entirely new conceptual design for the Indexing and

[25] An “object” is understood as physical or non-physical and may be a carrier of data, for example in the form of text, images, videos or sound, and of interest for archiving or research purposes due to its characteristics.

Digitisation programme. The positive aspects of the current funding model will need to be retained and supplemented by suitable funding instruments so as to produce reliably archived data that can be networked, reused and cited.

Digitisation and indexing

The expansion of the range of objects requires a funding concept for the digitisation of all objects which both keeps pace with the dynamism and potential of digitisation and also takes account of the need for long-term infrastructure funding that stimulates the creation of structures. The aim of digitisation funding is to develop methods that can be used in the digitisation of many different types of objects. However, the lack of standardisation in 3D digitisation and the preponderance of proprietary data formats in the case of audiovisual objects shows that starting out with an unrestricted approach to funding would be premature from an infrastructure perspective.

In the case of text digitisation, the established standards set out in the DFG Practical Guidelines on Digitisation continue to apply. Where there are recognised standards for the indexing and digitisation of specific objects, funding can be similar to that for written materials. In all other respects, funding should focus on the expansion of knowledge and the networking of experts as well as exploratory projects. This approach may be expected to contribute to the development of standards and reusable methods and procedures. The development of suitable equipment or the adaptation of equipment for the purposes of digitising cultural assets is also regarded as relevant, meaning that the funding of scientific infrastructure is closely linked to DFG funding for infrastructure and equipment.

Indexing, which provides contextualisation and structuring for individual objects and collections of objects through the creation of metadata, is undoubtedly relevant. However, current funding needs to be adapted, particularly due to the following three factors. First, the relationship between indexing and digitisation is reciprocal, and digital versions influence the depth of indexing. Second, because of the possibilities offered by working digitally, indexing becomes easier to allocate among various parties and can take place in a downstream process. Third, it can be advantageous to develop, improve and validate indexing concepts where there are already large volumes of data on which to base the concepts. Funding programmes should respond to this by enabling the digitisation of analogue objects even without the availability of detailed indexing data.

This also means that more emphasis is needed on the development of automated structuring and indexing methods based on representative samples and training data in order to respond to the likely growth in data volumes. While the e-Research Technologies programme focuses

on the development and implementation of processes, the Indexing and Digitisation programme will incorporate the application of these processes to index the content generated.

Standardisation and data management

The use of standards has always been the foundation for the provision of content, providing the ability to connect different research disciplines. The larger the data volumes and the more heterogeneous the objects being studied, the more relevant both the use of standards and the funding of standards development become for scientific infrastructure funding. Standards may emerge from the community itself in a bottom-up process and/or be initiated on a top-down basis for policy and strategic reasons. Existing negotiation processes and approaches to developing discipline- and material-specific standards must receive continued funding support and the interoperability of subject-, category- and material-specific exchange formats and vocabularies must be further developed. Expert workshops and networking opportunities should also be offered to encourage the development of standards in areas which are still underdeveloped or for objects for which recognised standards do not yet exist. Consideration should be given to both international connectivity and existing international initiatives. Since standards only remain applicable through use and dynamic ongoing development, when designing funding programmes the approach should always be to gauge, based on the results, at what stage a standard has established itself to the point where extensive funding for indexing and digitisation can be based on it.

The development of an interoperable minimum data set is essential, especially if digitisation is possible without the creation of deep indexing metadata. The combination of a greatly reduced set of metadata and persistent identification will facilitate findability and reliable citation as well as the networking of heterogeneous data both in different disciplinary contexts and in cross-disciplinary and cross-category contexts. At the same time, standardised and clearly comprehensible documentation of rights should create added transparency for the reusability of data. This minimal level of exchange will constitute the minimum requirement for the creation of digital versions.

In terms of the anticipated complexity of the data and the varying degree to which standards have been established for different objects, disciplines and categories, well-designed data management will be extremely important. This must take into account the data life cycle, integration into the research process, indexing and archiving concepts including the legal status of the data, the time dimensions and recognised challenges and planned approaches. It follows that there can be no doubt as to the importance of data management plans for indexing and digitisation projects. The content and degree of differentiation must be developed with reference to existing concepts, for example those for research data.

Scientific information infrastructure

The more heterogeneous the objects with relevance to research are which are processed, not only *for* but increasingly *by* the research community, the more essential close interplay between researchers and scientific infrastructure providers becomes. In addition to the current approach of the funding programme, based on holdings and collections, a research-motivated approach will become increasingly important. Hence, it is not enough to simply convert research collections from analogue to digital form. In fact, dynamic research collections are created by the ability to combine and network a wide range of different objects and data.

Along with these requirements for standardisation and the design of suitable collaborative strategies, it will become increasingly important to create and further develop flexible and interactive systems. These reference and presentation systems must be flexible enough to represent existing heterogeneous data sources with diverse quality levels and differing provenance, but they must also support data input by researchers and employees of information infrastructure facilities in processes that are in part independent of time and location. Connectivity with existing international systems is a prerequisite for this. The close interlinking of DFG funding formats in the area of scientific infrastructure, but also in research funding, is clear to see. While the Indexing and Digitisation programme is primarily concerned with making research-relevant content available, the e-Research Technologies programme offers funding for the establishment or expansion of national digital information infrastructures. The intersection with research funding arises from the intensity with which the objects are processed and the research questions that may be associated with the actual indexing and digitisation process. Although current funding models already provide opportunities for linking research projects and information infrastructure projects, the aim is to achieve a more open funding concept to satisfy needs relating to the combined processing and study of objects.

4. Measures

The challenges described that arise from the reorientation of the funding programme are far-reaching. It follows that funding options will evolve gradually, especially since funding opportunities for archival sources, medieval manuscripts and historical newspapers are still being initiated on the basis of the results of the pilot phases for the digitisation of these materials.

The measures named above that accompany the opening up of the funding programme to include the indexing of all potential research objects are as follows:

- funding for the development of an interoperable minimum data set to enable the funding of digitisation without prior indexing

- funding for the use of automated indexing and structuring methods (funding for the development of techniques in the e-Research Technologies programme)
- funding for expert workshops and networking opportunities for the creation of standards, relating to
    - the indexing of objects for which no recognised discipline- or material-specific standards yet exist
    - the digitisation of objects, especially for multidimensional objects and where there is a preponderance of proprietary formats
- funding for exploratory projects on the digitisation of objects if they have the potential to contribute to the development of standards
- development of the necessary degree of differentiation of data management plans for indexing and digitisation projects, drawing on existing concepts with the help of interviews or meetings with experts
- further development of flexible and interactive systems appropriate to the generated data and research working methods (funding for the development of systems in the e-Research Technologies programme)
- consideration of the interaction between scientific information infrastructure funding and DFG infrastructure and instrumentation funding for digitisation technology, with a view to developing suitable equipment for the digitisation of objects or to funding the acquisition of such equipment
- facilitation of cooperation between research projects and information infrastructure projects by raising awareness of available opportunities to link up funding opportunities.

II. Open Access Transformation

1. Current Funding Situation

"Open access" refers to the aim of making research results[26] widely available via the internet such that they can be accessed without financial, technical, organisational or legal restrictions and are available for productive reuse. The DFG has actively supported open access since 2003. The principle of open access represents a continuation of the digital transformation in publishing in a manner that meets the needs of research. Quality-assured open access publishing enables efficient research processes and novel research, and guarantees the verifiability of results. It can also contribute to enhancing public trust in expert discourse and the validity of scientific findings.

The DFG views the open access transformation as a restructuring of the scientific publishing system which includes aspects such as the transformation of publishing routes, business and finance models, processes and procedures, publication formats, evaluation and sharing mechanisms, and conventions and approaches. This transformation may be of equal relevance to all disciplines. The DFG therefore supports all open access models as long as they are conducive to scholarly communication. No privilege is given to any particular open access road within the funding framework.

At present there are multiple programmes for information infrastructure funding which support researchers and research institutions in the implementation of open access. The Open Access Publishing and Licences for Digital Content programmes and the call for proposals on Open Access Transformation Contracts – which also enables the transformation of e-books and other formats – are designed to support the improvement of basic structural conditions for access to scientific information across all disciplines. The programme Infrastructure for Electronic Publications and Digital Scholarly Communication enables institutions and individual researchers to establish and expand publishing infrastructures and repositories, convert journals to open access and implement other projects relating to metrics, monitoring and the development of new publishing formats.

It is also possible to apply for publishing costs in the majority of research funding formats (research grants, early career support programmes, Collaborative Research Centres, Research Training Groups, Research Units, Priority Programmes), which can be used to cover the costs of open access publishing.

[26] The broad term "research results" is used here in place of the term "publications" because today scientific findings are no longer documented exclusively in the "traditional" forms of journal article, contribution to an anthology, book chapter or monograph.

2. Open Access as Part of Academic Publishing: Dynamics of Current Trends

Academic publishing is experiencing sustained worldwide expansion. Quality-assured gold open access publications currently account for approximately 11% of output[27]. Large amounts of scientific literature are also accessible via the green road, personal websites or academic exchange platforms and can be found via search systems such as BASE (Bielefeld Academic Search Engine) as well as integrable tools[28].

Scholarly publishing is currently an extraordinarily dynamic area. There is an increasingly critical relationship between the interests voiced by the research community and the interests of market-leading publishers. Efforts are being made within the framework of international initiatives, such as OA2020, and national initiatives, including the DEAL project initiated by the Alliance of Science Organisations in Germany, the National Open Access Contact Point OA2020-de and the DFG call for proposals on Open Access Transformation Contracts, to achieve a large-scale transition from a subscription model, which is no longer adequate for academic requirements, to a system funded on the basis of open access publishing.

At the same time, publishers and infrastructure facilities are working towards offering researchers in a wide range of disciplines additional options that go beyond the kind of open access financed by article or book publication fees (for example publicly funded publishing infrastructures, funding from foundations, crowd funding and cooperative open access funding)[29]. Along with public institutions, smaller and new publishers are also taking advantage of the opportunity to provide services that meet the needs of academics and promote open scholarly communication while satisfying a range of needs, for example the continuing need for printed books. Forms of publishing and reception that function independently of publishing companies are also becoming established (for instance discipline-specific publishing servers run by the Center for Open Science).

In view of the ongoing expansion of open access publishing, it would appear that publishing and dissemination services are becoming increasingly disconnected from traditional mechanisms of filtering, quality assessment and reputation-building. The dynamic evolution of academic publishing can be seen in new and diverse publishing and reception preferences. In parallel to this, a new diversity and increasing complexity are emerging in relation to publishing

[27] Piwowar, Priem, Larivière et. al.: The State of OA: A large-scale analysis of the prevalence and impact of Open Access articles. This figure only relates to articles that appeared in 2015.

[28] E.g. oaDOI as the basis for services like Unpaywall (https://unpaywall.org/).

[29] See for example the Jussieu Call: http://jussieucall.org.

processes and funding models.[30] This dynamic within the open access transformation is also bringing about a shift in the established roles of stakeholders. Researchers are exploiting the possibilities of direct digital sharing and showing an interest in the rapid dissemination and perception of their results, whether interim or final. They are increasingly uneasy about traditional evaluation mechanisms inherited from the world of print journals, such as the journal impact factor.

Libraries are no longer focused on the task of acquiring and providing scientific information alone, but are also acting as publishing infrastructures, organisers of the open access transformation with a role in reallocating funds and providing advice. In an international framework, funding organisations themselves perform the task of establishing publishing infrastructures and linking them to evaluation mechanisms. By imposing open access obligations, they are setting policy guidelines and working on the development and harmonisation of their policies. Germany's federal states, individual countries and the European Commission are drafting open access strategies and defining targets for the large-scale transition to open access. These targets therefore also require monitoring mechanisms that supply comparable, contextualised data.

3. Challenges

The fundamental goal of open access is to improve scholarly communication; it is a form of publishing that is adequate to the needs of science and the humanities in the digital era. However, open access is value-neutral in itself, and its implementation may have both positive and negative effects on the overall system of scientific communication and its components.

In the DFG's view, open access should therefore always be based on research-friendly objectives and be implemented along corresponding lines. Determining the function of free access for the research community is key to the DFG's actions. A functional view of open access recognises not only the benefits of free access but also conflicts of interest regarding the definition of open access, sees path dependencies, and takes account of reinforcement effects and dialectic developments in the system as a whole. Regarding open access as an end in itself – and consequently imposing obligations to publish on an open access basis, for example – is not consistent with the DFG's role as the self-governing body of the research community. Instead, the DFG has an intrinsic interest in supporting research-friendly open access structures and engaging in a critical debate on the impacts of open access on the research system, both desirable and undesirable.

[30] Open access itself is becoming ever more complicated, at least in terms of the different roads that can be described: diamond, gold, bronze, green, blue, etc. Different directions are also emerging in relation to the implementation of open access.

With respect to the complexity of the situation, a clear position is needed as to the basic foundation on which open access can develop along research-friendly lines. The open access conditions defined in the Berlin Declaration of 2003 fundamentally apply, as do the principles based on these conditions defined by Science Europe[31]. In view of efforts to create definitions of open access that vary from the Berlin Declaration[32], further foundations are needed for a research-friendly open access policy.

Open access still means globally free accessibility and legally protected reuse of extensive information in return for appropriate remuneration of costs or as part of publicly funded and supported infrastructures. It is also generally in the interests of researchers and their institutions to retain rights of use over their work and to decide for themselves how it can be used, for example using CC licences without paying separately for them. The open access label should not be allowed to develop into a further complication in the supply of literature (for example through unclear rules on reuse and dissemination, differing embargo periods and different variants of open access). Different access modalities applying to the same content within individual countries or for different regions of the world are not adequate for the purposes of scientific research.

An instrumentalisation of open access and open science by commercial parties is also not in the interests of research. Platforms and services that were originally freely accessible and free of charge are increasingly being commercialised through takeover or a modification of the business model. There is a danger that researchers across the whole spectrum of science and scholarship may become dependent on digital services offered by dominant providers. The research community should reflect critically on this.

The ongoing expansion of academic publishing is also triggering phenomena that have a negative impact on the quality standards of published output. It does not serve the interests of science if strong pressure to publish leads to a weakening of quality assurance mechanisms. This applies to open access publishing just as it does to work that is not published directly on an open access basis. The DFG therefore supports the fundamental principles of good science publishing.[33]

It would also not be in the interests of science if an expansion of access to licensed content led to an attempt to enable personal traceability of reception (end-to-end traceability) even of freely accessible scientific content on publishing platforms, and if such access, usage and

---

[31] Science Europe Principles on Open Access to Research Publications, May 2015: http://www.scienceeurope.org/wp-content/uploads/2015/10/SE_POA_Pos_Statement_WEB_FINAL_20150617.pdf.

[32] This includes attempts by major publishing companies to redefine open access.

[33] The relevant basic principles or guidelines can be found in the Statement on scientific publications by three national Academies (Académie des sciences, Leopoldina and Royal Society) dated 24 October 2016, see https://www.leopoldina.org/en/de/press/press-releases/press-release/press/2451/.

where applicable dissemination data could be aggregated and analysed by commercial parties, states or other third parties. Finally, it is not in the interests of science that administrative and transaction costs at public institutions are increasing due to services that should generally be performed by publishers being outsourced to libraries and researchers.

The DFG believes that its future role in relation to open access will increasingly involve the adjustment of its funding mechanisms and standards in response to the systematic observation of positive as well as negative trends in academic publishing. Undesirable system effects should be avoided, the management of funding flows should be organised in a way that is beneficial to the research effort, and science-oriented open access goals should be achieved. This is all the more applicable in view of recognisable system trends which are harmful to science, for example trends towards further concentration among publishers, the possibility of conflicting objectives in open access[34] and an increase in different opinions[35] on the future restructuring of academic publishing.

The current period of rapid change offers all publishing stakeholders different and far-reaching opportunities to participate in actively shaping the future of open access. The DFG therefore calls on researchers, infrastructure facilities and their financial supporters to make their respective areas of responsibility conducive to the research endeavour.

4. Action Required and Measures

The DFG's current approach, which it will continue to pursue, is to support all open access models which are deemed appropriate from a scientific perspective with an eye to their suitability for facilitating a research-friendly transformation. Wherever free access contributes to facilitating the sharing of information and knowledge transfer, enabling research processes or making them more efficient, improving quality assurance processes in scholarly communication or making publishing infrastructures and processes adequate for the needs of research, it will be supported by the DFG. Different subject areas use and develop different communication formats. DFG open access funding will therefore continue to take into account the heterogeneity of publishing practices that different subject cultures entail. Ultimately, the functional concept of transformation implies not only enabling the conversion of existing content in a restricted publishing mode to a freely accessible one, but also supporting the emergence of new

[34] The following situations may be considered as involving conflicting objectives: Are open access goals designed to increase the number of open access articles or reduce costs per publication? Should priority be given to supporting new publishing models and the creation of competition in publishing services or to transforming existing journals? Should preference be given to the green road or the gold?

[35] The reference is primarily to debates surrounding suitable open access variants and the way in which infrastructures are expanded (as public assets or on a commercial basis).

forms of scientific communication which are geared from the outset towards the open sharing of research results.

As far as information infrastructure funding is concerned, comprehensive support for open access projects on a bottom-up basis will continue in the years ahead through the Infrastructure for Electronic Publications and Digital Scholarly Communication programme. This will be accompanied by the more strategic/structural Open Access Publishing and Licences for Digital Content programmes, which will be evaluated and further developed in turn in 2018 and 2019.

The current proportion of freely accessible gold open access publications at German universities[36] raises the question of how, in future, DFG funding programmes and funding rules can optimally support the systematic and functionally based implementation of open access above and beyond this approach.

In the years ahead there will therefore be a need to design policies and programmes that go beyond information infrastructure funding and relate to research funding. This will present the challenge of harmonising, in strategic interventions, the heterogeneity of practice in individual subject areas and structurally uniform support by the DFG.

Further development of modalities for the funding of publication fees

Continued support for the open access transformation will require a new approach to publication costs. This calls for mechanisms to correlate expenditure by central information facilities with the organisation's level of research activity and to balance out temporary peaks in demand. At an institutional level, fees for open access publications are already frequently administered via central funds, which are usually managed by libraries and also draw on the library budget (without additional funding). Initially, universities in particular must have a central overview of *all* cost elements and expenditure for publishing in order to manage a sustainable transformation. In most cases, a research-intensive institution will also be very successful in attracting third-party funding and will face high acquisition and licence costs and, potentially, high publishing costs (colour charges, page charges, APC, membership fees and potentially contributions to consortium models). Future mechanisms for publication funding must be adapted both on an intra-university basis and between third-party funding providers and institutions. However, this should be undertaken in such a way that increased and reduced burdens are balanced throughout the system.

The DFG will review the practice of fixed-rate publication fees with a view to open access and link it to the further development of institutional support for literature provision (programme

[36] M. Wohlgemuth, C. Rimmert, N. Taubert: Nutzung von Gold Open Access auf globaler und europäischer Ebene sowie in Forschungsorganisationen [Use of Gold Open Access at the Global and European Level and in Research Organisations], July 2017, p. 36 and p. 41 (available only in German).

evaluation for Open Access Publishing and Licences for Digital Content). In any event, it is important to prevent the allocation of publication funds from contributing to an inappropriate rise in system-wide costs. The DFG is therefore calling on the federal states, individual institutions and researchers themselves to be aware of cost trends in open access charges when allocating and using funds and to introduce mechanisms to ensure cost transparency.

Open access monitoring within the DFG

Currently, the task at the national level and thus also for the DFG is the establishment of monitoring mechanisms with which to gather detailed, comparable information about publication output and the associated costs. Models for the transition to an open access publishing system, as envisaged by the OA2020 initiative and in the call for proposals on Open Access Transformation Contracts, must be developed on the basis of reliable data on publishing figures and costs. There is also a responsibility to maintain accountability in relation to open access publishing and cost trends in order to ensure that funds are being used economically. It should also be possible to correlate costs for different open access roads. The DFG therefore intends to establish processes to obtain more accurate data about publication output from DFG-funded projects and the associated costs. The systematic and comparable evaluation of relevant data will also provide an important basis for the ongoing development of DFG funding for open access publication fees. At the same time, the DFG will continue to participate in discussions on the methodology used to gather this type of data. The DFG does not limit the concept of monitoring to quantitative and empirical data. Monitoring also includes the analysis of system-wide trends of a qualitative nature. Quantitative data must be linked to qualitative objectives that go further than simple targets.

Further development of open access funding guidelines

In contrast to the approach adopted by other funding providers, the DFG will continue its policy of not requiring individual researchers to publish the findings resulting from DFG-funded projects on an open access basis. The DFG open access policy, adopted in 2006 and published as part of the usage guidelines, will, however, be adapted to the effect that the DFG will *request* researchers in receipt of DFG funding to publish through open access. The DFG will continuously monitor the extent to which its rules enable and support an academically driven evolution of open access.

III. Research Data

1. Current Funding Situation

Since research data is of fundamental importance to the quality of research work, the DFG supports the long-term safeguarding and provision of such data through policy statements, research funding regulations and the establishment of information infrastructures.

Following intensive discussions with the review boards, the Senate approved the Guidelines on the Handling of Research Data[37] in autumn 2015. At an abstract, cross-disciplinary level, these guidelines set out basic expectations of data-related project planning and the provision of data produced in DFG-funded projects. They also call on scientific communities to develop regulations for research data management that are adequate to the needs of science and scholarship. In the Proposal Preparation Instructions, researchers are also asked to describe exactly[38] how the data resulting from a project is to be made available to others working on the subject.

The funding programme Information Infrastructures for Research Data,[39] established in 2013, is designed to facilitate the development or ongoing professionalisation of information infrastructures for the archiving and provision of research data, suitable for specific subject areas and in response to specific research needs. The breadth of projects funded reflects the diversity of requirements in terms of suitable infrastructures and fulfils the intent behind the funding programme of making funding available to all disciplines with their quite distinct sets of needs.

A separate call[40] concentrated on the question of how research data already available in repositories can be used to answer specific research questions, and was also intended to demonstrate how the infrastructure would need to be optimised in order to help answer such questions.

---

[37] DFG Guidelines on the Handling of Research Data, 2015: http://www.dfg.de/download/pdf/foerderung/antragstellung/forschungsdaten/guidelines_research_data.pdf.

[38] See http://www.dfg.de/formulare/54_01/54_01_en.pdf, p. 5f.

[39] See http://www.dfg.de/formulare/12_14/12_14_en.pdf.

[40] See http://www.dfg.de/download/pdf/foerderung/programme/lis/ausschreibung_forschungsdaten_151130.pdf (available only in German).

2. Research Data: Fulfilling Scientific and Policy Requirements

Scientific communities are increasingly discussing the question of how to handle their research data. There is debate, for example, on what can be regarded as data in a given research area, how the technical handling of data can be supported, how research data should be described, documented and archived in specific subject areas and what usage regulations should be defined. The results of these discussions may yield recommendations and policies for research data management, such as are already available in certain research areas – in some cases making reference to DFG guidelines.[41]

In terms of DFG funding, the establishment and expansion of data-related information infrastructures has so far been concentrated in the social sciences and humanities and in the life sciences. It is evident that the funded projects are motivated by the prospect of more efficient organisation and division of labour, better links (both technical and information-related) between existing information systems, and a desire to satisfy the need to address interdisciplinary research questions. Against this background it is clear that disciplines in which the professional handling of research data is well established are those which are striving for information systems with an optimum federated structure.

Debate at the research policy level reflects this trend. At the national level, it is demonstrated by the publications and recommendations of the German Council for Scientific Information Infrastructures (RfII).[42] For example, the discussion paper presented by the RfII in April 2017 on a National Research Data Infrastructure (NFDI) contains essential elements for the efficient establishment of such an infrastructure,[43] namely the strengthening of self-governance processes within academic disciplines, a willingness on the part of the research community and infrastructure providers to engage in the relevant dialogue, and the need to develop data-related skills through new training programmes. There is also a working group focused on the implementation of the Digital Agenda which is concerned with the development of relevant skills. With regard to the efficient use of research data in academia and industry, this working group has designed a competition to identify good approaches to relevant skills training across the full spectrum of research data handling.[44]

There are clear parallels between the German and the international and European consideration of this issue. In particular, the discussions surrounding the European Open Science Cloud

---

[41] See http://www.dfg.de/en/research_funding/proposal_review_decision/applicants/research_data.

[42] German Council for Scientific Information Infrastructures; http://www.rfii.de

[43] RfII Discussion Paper 2017: Zur Nationalen Forschungsdateninfrastruktur [On the National Research Data Infrastructure], available only in German); http://www.rfii.de/?wpdmdl=2269.

[44] See http://www.fz-juelich.de/ue/DE/Leistungen/Controlling_UE-C/WS_Wissenschaft/_node.html (available only in German).

(EOSC) proposed by the European Commission[45], which is intended to create an international network of initiatives, are important at present. The associated issues, such as standards, access regulations, business models and governance, largely reflect the issues being discussed nationally, and German experts are contributing to the debate.

In view of the highly dynamic nature of the debate, and given the clear differences between the needs and challenges of different academic disciplines where the handling of research data is concerned, the DFG should specifically focus its funding activities on those options that enable vital contributions to be made to the establishment of research data infrastructures coordinated at the national and European level. These options include policies and regulations, the development of research data management skills and the coordinated establishment of a federated research data infrastructure.

3. Action Required

Action area: Policies and regulations

The action area of policies and regulations is regarded as one of the most effective fields for establishing the professional handling of research data. Consideration must be given to both organisational and community-specific requirements. At a national level, in addition to the Alliance of German Science Organisations[46], many other scientific and academic organisations and institutions have drawn up regulations of a general nature and made them mandatory to varying degrees for their researchers. There is vigorous discussion, not least in the context of the EOSC, as to whether researchers submitting proposals should be required to submit binding data management plans designed to support the long-term archiving and provision of research data adequate to the needs of their discipline. The DFG does not require data management plans, but does emphasise the need for carefully planned data management by asking applicants to provide a detailed description in research project proposals of how they intend to handle research data.[47] The option of applying for a separate information management module in the case of Collaborative Research Centres also supports this effort.

The concrete implementation of the general recommendations and higher-level regulations for the handling of research data is especially successful where individual disciplines have defined in greater detail how the research data that is characteristic of their discipline is to be handled. Detailed regulations, agreed within individual scientific communities, do, however, tend to be

[45] The European Open Science Cloud Pilot Project https://eoscpilot.eu

[46] Principles for the Handling of Research Data (2010): https://www.allianzinitiative.de/fileadmin/user_upload/www.allianzinitiative.de/Principles_Research_Data_2010.pdf.

[47] Proposal Preparation Instructions – Project Proposals; DFG form 54.01 – 07/17; section 2.4 Data handling; http://www.dfg.de/formulare/54_01/54_01_en.pdf.

found in those disciplines which are already very well organised;[48] other research areas are only just beginning to discuss possible regulations or do not yet recognise a need to address this issue. It is emphasised that significant differences between the cultures of individual subject areas continue to pose a challenge and must always be taken into account in any consideration of these issues.

Action area: Establishment and development of research data management skills

The professional handling of research data requires sound knowledge of methods for gathering, analysis and (where relevant value-creating) use of such data as well as professional research data management. However, systematic training in these skills as part of the professional training of researchers and information specialists is still in its early stages. The German Council for Scientific Information Infrastructures, among others, is therefore calling for the targeted expansion of data-specific expertise, making a distinction between continuing and advanced training for researchers and professional training for specialists in data and information management.[49] To ensure the future success of the research endeavour, relevant training courses must be developed for early career researchers, and these must become a permanent feature of university curricula in the medium to long term. Although this is not part of the DFG's primary mandate, the organisation should consider where its own funding instruments provide suitable opportunities to introduce researchers, especially those at the early stages of their careers, to the professional management of research data. Relevant activities must be closely coordinated with appropriate stakeholders in the research system, such as learned societies, so that they can have a long-term structural impact.

Action area: Establishment of suitable infrastructure

Information infrastructures that provide long-term archiving and reliable provision of research data from different subject areas are already available in many forms and may be regarded as an effective basis which can be adapted to new requirements and further professionalised.

The fact that the infrastructures resulting from DFG funding are very different in character reflects the differing requirements and working practices of the scientific communities and the fact that most of these infrastructures have so far developed within a comparatively small number of research areas. It is important, when establishing and expanding data-related infrastructures, to avoid isolated solutions and instead, by linking new and existing data repositories and services through coordinated networking, to create a reliable and, where appropriate, cross-

[48] See the subject-specific information on handling research data (as note 1): http://www.dfg.de/en/research_funding/proposal_review_decision/applicants/research_data.

[49] See footnote 39.

disciplinary federation of information infrastructures. Funding should therefore be systematically designed to ensure the connectivity and interoperability of data-related infrastructure. It is necessary both to address the federation of community-specific infrastructures and to examine how institutional structures can be modified as required to reliably support scientific work on the basis of research data. So far, only isolated institutional concepts suitable for supporting research data management on the basis of strong IT architectures have been developed and implemented. The evaluation of the Information Infrastructures for Research Data funding programme scheduled in 2018 may be expected to provide guidance as to how best to support the specific needs of researchers, at least as far as community-specific infrastructures and services are concerned.

4. Measures

Action area: Policies and regulations

The further development and negotiation of subject-specific policies and regulations should be supported through DFG funding. This funding should be carried out in close consultation with subject area funding and should target scientific communities with a significantly lower level of organisation. There is a particular need for the development of research data guidelines and the associated self-organisation in research areas that work with volume data. The necessary measures, such as roundtable discussions, can be funded as part of the Information Infrastructures for Research Data programme in order to support these processes.

In particular, subject-specific data policies must address the question of what kind of data requires long-term storage and in which cases archiving can be deliberately waived. It is appropriate to establish a position on this point in order to formulate and implement any requirements for the provision of local storage capacities. This will also highlight how relevant it is to the implementation of subject-specific policies to consider and design the interfaces between local and national data infrastructures.

However, to implement the required policies it is necessary to specifically engage all stakeholders across the particular range of fields. For this reason, the DFG must also make a point of networking the many stakeholders involved and working towards achieving acceptance of the developed policies through the review boards. To enable the best possible research in digital infrastructures, the DFG must make a special effort to bring on board those disciplines which do not operate within a big data environment.

Finally, on the basis of such broad discussion, which also takes the circumstances of individual disciplines into account, it will be possible to examine whether the request to describe the development and use of data management plans in project proposals submitted to the DFG should be defined in more detail.

Action area: Establishment and development of research data management skills

Training in data skills is currently often provided in the context of extracurricular activities, such as summer schools. With regard to the DFG funding portfolio, it would therefore be appropriate to evaluate whether similar summer schools could be offered in DFG-specific funding formats such as Collaborative Research Centres, Research Training Groups or workshops for early career investigators, thus enhancing the level of expertise in research data management through research project funding. This evaluation could initially result in recommendations on systematic training in the relevant skills. It is also conceivable that the Information Infrastructures for Research Data programme could be specifically used to develop initial ideas for the design of such training or to identify at what places in the DFG funding portfolio the development of relevant skills could be systematically integrated.

This measure must be coordinated with the results of comparable initiatives in Germany and abroad. It will be important to cooperate with experts, especially at universities and other higher education institutions, and to widely publicise model examples of the quality-oriented handling of research data – for example on the part of university libraries.

Action area: Establishment of suitable infrastructure

Initially, the funding of information infrastructures for research data will continue to serve the needs of the scientific communities. In addition, however, funding is now also to specifically address institutions in a targeted manner. Both types of funding must ultimately be reconciled with due consideration of the DFG's contribution to the establishment of a National Research Data Infrastructure (NFDI).[50]

A funding option will be developed, drawing on the results of the evaluation of the funding programme in 2018, which is clearly oriented towards the coordination and federation of existing community-specific infrastructures so as to ensure the connectivity and interoperability of different systems.

At the same time, future funding will specifically create the impetus for research institutions to develop IT and data management concepts and, when implementing these concepts, to clearly orient their structures to support computer-assisted analysis and reuse of digital data in research. The national impact of the impetus envisaged will be achieved by initiating a dialogue among the funded institutions in parallel to actual infrastructure funding, for example to reach agreement on standards for these concepts.

[50] See footnote 39.

The DFG will make a significant contribution to the creation of a National Research Data Infrastructure, in particular through the measures envisaged to promote the networking of community-specific services. When designing future funding measures for federated infrastructures, a careful review will be necessary to establish how the latter respond to the needs and interests of researchers. To this end, close coordination will also be required between DFG funding initiatives and the funding measures associated with the National Research Data Infrastructure.