

An integrated digital research environment – DFG perspectives

Anne Lipp

Scientific Library Services and Information Systems (LIS), German Research Foundation (DFG), Bonn, Germany

Abstract:

Purpose:

The article focuses on the vision that the Deutsche Forschungsgemeinschaft (German Research Foundation, DFG) pursues with its funding programmes in the field of digital information provision.

Design/methodology/approach:

The first section of the article sketches out the strategic decisions which determine the funding policy of the DFG in the field of digital information, while the second section describes in more detail the central funding schemes of the DFG in the field of digital information.

Findings:

The funding policy of the DFG seeks to build an integrated digital research environment that includes scholarly publications and primary research data as well as new forms of communication in virtual research and work environments. But it will be only by the common effort of scholars, libraries and providers of scholarly information, funders, publishers and fee collecting agencies that the vision of an integrated digital research environment will come true.

Originality / value:

The outline of the funding programmes of the DFG in the area of digital information provision will be of interest to librarians and information professionals seeking information about library funding policies and strategies in Germany.

Paper type:

General review

Keywords:

National Licences, Open Access, Digitisation, Research Data, Virtual Research Environments, National Hosting Strategy

An integrated digital research environment – DFG perspectives

In the ideal digital world, every researcher would be able to access all the information needed for scholarly work from his or her workplace – wherever that may be. This includes scholarly publications, primary research data, and new forms of communication in virtual research and work environments.

This is the vision that the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) pursues with its funding programmes in the field of digital information provision. In this digital environment, each researcher can compile his or her own library, using modules offered nationwide and available free of charge to all who require them.

One comment before focusing on digital developments: it should be made clear that the DFG does not focus exclusively on digital content in its information infrastructure funding. Digital information provision is *one* of several funding priorities, albeit a very central one. But it in no way fully reflects the funding spectrum of the DFG in the field of scientific library services and information systems. The DFG also supports libraries and other centres for information services in Germany in acquiring conventional forms of the most relevant scholarly literature in a given field via the system of so-called special subject collections, and in making them available to all interested researchers nationwide. In 2008, the DFG committed around 10 million euros to special subject collections alone, representing 20% of the total funds available for supporting scientific library services and information systems.

Now back to the focus of this paper: the provision of scholarly information in an ideal digital world. The first section will sketch out the strategic decisions which determine the funding policy of the DFG in the field of digital information. Then, the second section will describe in more detail the central funding schemes of the DFG in the field of digital information.

1. Strategic decisions

The DFG is Germany's largest research funding organisation. It promotes research projects in all fields of science and the humanities. Furthermore, the DFG supports projects for funding scholarly infrastructure. The two areas of infrastructure funding are scientific equipment funding, such as major research apparatus, and the funding of scholarly information infrastructure.

In the current year, 2009, the DFG's annual budget is around 2.2 billion euros, of which 200 million euros is allocated to research infrastructure, and of this amount, around 63 million euros goes to funding scholarly information infrastructure. The remaining funds for infrastructure are spent on major research apparatus, on research vessels, but also on

facilities such as the European Liaison Office of the German Research Organisations (KoWi) or the Institute for Research Information and Quality Assurance (IFQ).

In the area of scholarly information infrastructure, the DFG funds projects and schemes that go beyond the primary responsibilities of a given information providing institution, or which could not be carried out by any individual institutions, but rather require cooperation. To do this, the DFG's funding portfolio is divided into four major areas: (1) Nationwide Library Services – (2) Indexing and Digitisation – (3) Electronic Publications – (4) Information Management. The body responsible for taking strategic decisions in the area of information infrastructure is the Committee on Scientific Library Services and Information Systems (Ausschuss für Wissenschaftliche Bibliotheken und Informationssysteme, AWBI), the members of which come from libraries, archives and other centres for information services as well as from the academic community itself.

The current funding priorities are based on a strategic review completed by the AWBI in 2005. In it, the AWBI examined how the DFG's funding programmes can better take into account the changing methods of working among researchers, as well as the changed conditions in general for scholarly publication and communication in the digital (research) age. The AWBI formulated as its goal the creation of an integrated digital research environment set forth in the 2006 DFG position paper "Scientific Library Services and Information Systems: Funding Priorities Through 2015" (DFG, 2006).

In a second step, in June 2008 the partners who together make up the Alliance of German Science Organisations (Allianz der deutschen Wissenschaftsorganisationen) issued a joint strategy entitled "Digital Information". This is currently the major policy statement coordinating funding and other activities in relation to information infrastructure in Germany. It begins with the notion that an integrated and innovative system of information provision can be set up most effectively through cooperation in the national and international context. In the national context, the Alliance organisations have agreed to more intensively coordinate the activities of the individual partners in the field of digital scholarly information systems, and to further develop them through a joint "Digital Information" priority initiative in the years 2008 to 2012. The goal of this initiative is to concentrate competencies and resources as well as to further improve the coordination of current and future activities (Alliance of German Science Organisations, 2008).

The central elements of an integrated provision of digital information as funded by the DFG come under the headings of National licensing, Open Access, Digitisation, Research data, Virtual Research Environments and – particularly in the context of the Alliance's "Digital Information" initiative – National hosting strategy. The nature of the DFG's contribution to shaping these individual elements is described in more detail below.

2. Central funding schemes

National licences / National licensing

When so-called national licences were first acquired with DFG funds in 2004, all the participants in this process – the publishing houses providing the content, the libraries obtaining the licences, and the DFG as funding provider – entered into new territory. Products purchased with a national licence are accessible free of charge to all academic institutions in Germany. Many of the products are also available for private individuals in Germany to register for and use. Since 2004, 124 products have been purchased in all – single databases and periodicals archives – for around 68 million euros. A list of the products currently accessible under the national licences scheme can be found at the website of this funding scheme (<http://www.nationallizenzen.de/>). The national licences for databases and journal archives were funded 100% through the DFG.

As a variant of this so-called “classic” national licence – i.e. where the licensed products are accessible to all academic institutions – the licensing of current periodicals is also being tested in a pilot project running from 2008 to 2010. This is being financed jointly by the institutions taking out the licence and the DFG, as a so-called “opt-in” model. The DFG has made available approximately 18.5 million euros for a total of 12 license packages with various publishing companies. While the current issues of the journals are initially only available to the institutions participating in the opt-in procedure, the updated archives are made available to all academic institutions after one year, under the *moving wall* principle. A third variant of national licences has been developed within the framework of Knowledge Exchange as a *Multinational Licensing Tender*. In this procedure the DFG issued a joint tender for multinational licences with the partner organisations DEFF (Denmark’s Electronic Research Library, DK), JISC (Joint Information Systems Committee, UK) and SURFfoundation (NL). Publishers were invited to submit plans for common international licensing, which simultaneously combine a “classical” national licence with an opt-in model. This allowed each country to take its “own path”, while still allowing common negotiations and activities to take place. The DFG has committed around 2 million euros for this project. After four years of funding for national licences, the scheme is now beginning to bear fruit. National licences have allowed a funding and acquisition model to be successfully installed, which takes into account both the challenge of digital information provision (national provision) as well as its opportunities (potentially unlimited access for all scholars from anywhere in Germany) in an appropriate way. The DFG sees national licences as a significant improvement in access to scholarly information across the board. They have enabled the development of new structures and forms of cooperation that bring Germany

closer to the vision of a digital research environment – one in which all researchers have the greatest possible digital access to the information they need.

Obviously these structures need to be further developed and established as a fixture within the academic landscape. At the same time however, it should also be emphasised that this is a task which neither any single institution nor the DFG can take on as part of its remit. The AWBI has therefore decided to modify the existing model of “classic” national licences (100% DFG funding), and, as of 2011, to replace it with a new model that treats the financing of national licences as a task to be undertaken jointly by all scholarly information institutions together with the DFG.

Indeed, one of the primary objectives of the Alliance's "Digital information" initiative is to find ways of jointly licensing electronic media such as current periodicals, complete databases, dynamic databases and e-books.

Open Access

The knowledge arising from academic research - constantly being extended, modified and documented in scholarly publications - must be made available in digital form, ideally without financial, technical or legal barriers. To facilitate this, a forward looking digital research environment must guarantee well-organised, financially sustainable, and uncomplicated access to all forms of published knowledge. Together with the Alliance of German Science Organisations, the DFG is emphatically committed to the idea that research findings from projects supported with public funds should provide the greatest available benefit to the public. The DFG therefore expects funded projects to make their research findings available either via repositories (the so-called "green road" of Open Access) in addition to a peer reviewed published journal – usually in a time-delayed manner – or to publish them immediately in suitable Open Access journals (the so-called "golden road").

In relation to the so-called "green road" of Open Access, the DFG also supports the establishment of discipline-based and institutional repositories. In the Social Science Open Access Repository, for example, more than 3,400 publications have so far been made openly available over the internet. In such repositories, publications are released partly directly by the authors, and partly by the repository operator in cooperation with the publishers. Value-added services – such as tools for citation analysis, usage statistics or plagiarism detection – are also offered to authors as additional incentives for making their work available via repositories.

The funding of the "golden road" is at present limited to publication funds, which can be applied for as a fixed sum in every DFG-funded research project and can be used to fund the author fees of an Open Access Gold journal.

Free access for scholars and researchers to all published knowledge via one of the two types of Open Access is certainly a very worthwhile goal. However, both Open Access Green and Open Access Gold have currently reached their limits, and these need to be overcome. The "green road" of Open Access demands additional effort on the part of scholars to store a copy of their publication in a repository. Moreover, few publishers allow a copy of the publication which is identical in layout to the printed version to be stored in a repository, so that authors fear a loss of reputation if the repository version is cited rather than the printed version. Solutions to this dilemma are currently being sought in the framework of the Alliance's "Digital information" initiative. Such a solution could involve an "Alliance-License", jointly negotiated by the Alliance partners, which includes the right for authors to enter their papers as they appeared in the licensed journals into an (institutional or discipline-specific) repository of their choice in the same format as the published printed PDF, and thus make it available under Open Access.

In terms of Open Access Gold publication, it must be pointed out that, at present, the structural conditions are not yet in place which would enable this model of publication to be established as a genuine alternative to journals based on the subscription model. In particular, there is a lack of reliable funding mechanisms that allow not only authors but also publishers, as the service providers, to carry out reliable financial planning. An early important initiative in this area is the international SCOAP3 initiative (Sponsoring Consortium for Open Access Publishing in Particle Physics), which is being supported in Germany by the Max Planck Society, the Helmholtz Association and the DFG. This has the objective of converting the leading scientific journals in particle physics into Open Access journals in cooperation with their publishers, so that all publications in this research field are accessible free of charge [1]. Only if Open Access succeeds in creating the structural conditions at the institutional level for supporting scholars and researchers specifically in the area of Open Access publication, while at the same time relieving them of the task of needing to find additional funding for their Open Access publications, will this form of publication have a serious chance of establishing itself in the publications market and competing on equal terms with the dominant subscription model. The next necessary step would seem to involve supporting research institutions in establishing stable financing structures that enable all authors who want to publish in an Open Access journal to do so. This offers the only prospect of success in the long term in setting in motion a partial and inevitable transformation process from the subscription model to the Open Access model. In turn, it would involve a partial reallocation of the funds currently used for subscriptions to journals. To start this restructuring process, additional funds will be necessary at the outset, since no direct correlation can be expected in the short term between an increasing number of Open Access publications and a decrease in current subscriptions. In the medium to long term

however, this will allow Open Access publication to become a more attractive model for scholars alongside the continuing subscription model.

Digitisation

While the challenge in *born digital* materials consists of finding licensing and publication forms that allow unrestricted access to the digital content, the challenge in relation to the majority of the already existing knowledge base is one of transferring them into digital form in the first place. Even if retro-digitisation sounds antiquated in an integrated digital research environment, the monumental tasks which remain to be done in this area should not be underestimated. First of all, the sheer quantity of material that must be considered for retro-digitisation is overwhelming. From the hand-written and printed heritage in academic libraries, via the scholarly relevant holdings in archives to collections of objects in cultural, art history and natural history museums – theoretically all of these collections are candidates for retro-digitisation. Secondly, almost all material from the 20th century is copyrighted, and this will mean working with the rights holders and the fee collecting agencies so that larger retro-digitisation measures can be set into motion. Thirdly and finally, there is still a considerable need for further research and optimisation in this area. Just one example would be the large-scale digitisation of three-dimensional objects.

The main priority of DFG funding currently involves the digitisation of out-of-copyright hand-written and printed material from the 16th to the 18th century – with the aim of quickly setting up a digital library containing the most important literature of the pre-digital era. An important aspect of all funding programmes that support retro-digitisation is that standards are observed – be it in the technical area or in the presentation of the digitised material –, such as those laid down in the recently updated "Practical Guidelines on Digitisation" from the DFG (DFG, 2009).

Research data

In an integrated digital research environment, it is not just the published knowledge that must be digitally accessible, but above all the foundations of knowledge production: the primary data relevant to the research project concerned.

The challenges in this still comparatively young branch of scholarly infrastructure funding relate both to the material itself and to the structures that are necessary to be able to incorporate research data into the scholarly communication and publication process – and thus into a digital research environment.

On the material: If primary research data is understood to mean the raw data of research, the foundations of knowledge production, then primary research data must be understood to refer equally to measurement data and archaeological finds, to objects from natural history

collections and historical architectural plans, archival materials and art works of all kinds, mineralogical collections, DNA sequences, drill cores from the international oceanography programmes or measurement data from large-scale exploration projects in the field of biodiversity research. The list could easily be continued.

With a view to digital availability, this material breaks down into two categories: firstly there is the legacy material, for which no digital version as yet exists, such as herbariums, insect collections, mineralogical collections, archaeological finds or archival materials. For these materials the first challenge consists of creating any sort of digital copy at all. While there is a clearly defined DFG funding programme for the hand-written and printed material held in libraries and archives, this is not (yet) the case for relevant museum collections that also represent raw research data for academic study. In a second category we can place all the primary research data already documented in digital form, such as measurement data from the natural sciences, or the survey results from sociological studies, and much more.

One task which is on the way to being undertaken is formulating the framework within which the DFG will offer a funding programme related to relevant collections in museums.

The greater challenge in relation to research data, however, is the setting up of appropriate infrastructures in order to integrate these datasets into a unified digital environment. This includes providing backup and archiving systems at suitable locations for the datasets – which in many cases require considerable resources to compile – to furnish them with comprehensive metadata, and finally to make them freely accessible in a suitable form to all interested researchers.

While there is a centuries-old infrastructure of publishers and libraries for the distribution and provisioning of published knowledge, a functioning infrastructure for primary research data only exists in individual disciplines, for example geology, marine geology, climate research and also, to some extent, in sociology. A large-scale infrastructure, however, covering all academic disciplines, is a challenge which the academic community and the funding bodies worldwide are only confronting now. For a funding organisation such as the DFG this is an important area of activity. The Committee on Information Management – a sub-committee of the AWBI – published a statement on this issue only recently, encouraging the development of discipline-specific organisational concepts that can regulate the sustainable storage of data.

The funding objective of establishing structures that are suitable for implementing a practical and sustainable system of handling research data while guaranteeing reusability and adherence to quality assurance principles, is much welcomed in all disciplines. There is also unity on the point that there cannot be, and must not be, a prescriptive approach on this issue. Therefore, it is to be left to the individual disciplines to define the framework and requirements necessary for handling research data according to the needs of the respective

disciplines, to store them securely in the long term, and to enable their efficient reuse. Only such an approach can guarantee that a given structure will work, both in terms of encouraging scholars to supply their data and enabling its reuse.

Future funding programmes in this area must therefore focus on the intersection between discipline-specific and infrastructure funding. An example of a successful model is the Centre for Psychological Information and Documentation (Zentrum für Psychologische Information und Dokumentation, ZPID) in Trier, a nationwide subject information centre for psychology that collects and prepares relevant literature, test methods, audiovisual media and high-quality resources, and makes them available online as quality-controlled materials.

Further work is being carried out to better integrate research data into the scholarly process and into the information structures. For example, the German National Library of Science and Technology in Hannover (TIB) has set up a registration agency for scientific data that can assign persistent and guaranteed digital identifiers to research data, i.e. these codes enable the relevant dataset to be accessed, and more importantly, cited at any time. This is already being very actively exploited by the earth sciences and environmental research communities, who foster an approach to handling research data that is oriented towards reusability as part of their scientific practice.

The DFG currently supports various initiatives in the fields of biodiversity, water research and archaeology aimed at establishing a sustainable strategy for primary research data. The stated aim of the funding is to extend the range of disciplinary expertise related to research data and to support all academic disciplines in setting up structures that fit their specific needs, provide long-term curation of research data, technical systems for long-term preservation, and the means to keep data accessible and reusable.

Virtual Research Environments

For some time now we have observed how new communication technologies and methods of publication are changing the existing system of information infrastructures, altering the forms of scholarly collaboration and communication as well as the publication process. As a technical possibility, this is certainly not new. Indeed, the present challenge is less about further refining technologies and making them more powerful, but rather how to make what is technically possible more usable – to adapt it to the needs of research groups and individuals as they occur in practice. We need to find answers to questions such as how platforms can be designed for network-based collaborative work, supporting new forms of collaboration in a manner that is attractive to scholars. The bottom line is that it needs to make research, scholarly communication and publication easier. A call issued by the DFG in both 2008 and 2009 invited scholars and researchers to submit proposals jointly with information centres – such as libraries or computer centres – for discipline-specific and interdisciplinary networked

digital research infrastructures, so-called virtual research environments. The aim is to provide network-based support for working and communication processes, especially in research groups and associations.

National hosting strategy

A final activity to be highlighted here is one of the six activities within the Alliance's "Digital Information" initiative: a common strategy for hosting digital content. The starting point for considering this issue is that all academic institutions are faced with the challenge of making digital content – whether purchased printed publications or data resulting from retro-digitisation projects - permanently accessible via a suitable infrastructure. For economic reasons and for reasons of technical and organisational efficiency, it is clear that the partners within the Alliance of German Science Organisations need to develop a common strategy and to implement it without delay. A national hosting strategy will pave the way for an efficient hosting infrastructure for the storage of digital texts, guaranteeing sustainable access to licensed commercial publications as well as to retro-digitised inventories. A core component of such an infrastructure is a repository in which data can be reliably stored. This storage system will not come with its own user interface, but will instead function as a “back-end”, equipped with standard access interfaces and services. The Alliance partners are currently commissioning a study to tackle the question of a suitable architecture for the strategy. The results will be used to inform and support the Alliance of German Science Organisations in future decisions on a national hosting strategy. It is expected that this decision-making process will be able to begin in 2009.

This outlined vision of an integrated digital research environment represents a common task. It requires many partners who are ready to help shape it – scientists and scholars, libraries and providers of scholarly information, funders, publishers and fee collecting agencies. Centres for information services such as libraries and data repositories are the drivers of this process, gauging the needs of scholars, identifying and implementing innovation, and coordinating the process. The role of scholars should be to push for those technological developments that make scholarly communication and publication easier, and to solve the pressing issue of long-term preservation in the digital world. The organizations funding the scholarly system – including the federal government and the states (Länder), research funding as well as performing organizations, and foundations – must take charge in setting robust strategies and then financing their implementation. Publishers and similar information providers are called upon to work together with scholars, libraries and other information service providers, and funding bodies to enable a financially viable model for both roads to Open Access. Finally, fee-collecting authors' societies will play a key role in this process, ensuring that the interests of all stakeholders – the scholars on the one side, the copyright

holders in the broadest sense on the other – are adequately accounted for in the transition to a truly integrated digital research environment. It will only be through a common and coordinated effort that an integrated digital research environment will be set up successfully.

Notes

[i] For more information, cf.: <http://scoap3.org/> .

References

Alliance of German Science Organisations (2008), “Priority Initiative “Digital Information” by the Alliance of German Science Organisations”, available at:

http://www.dfg.de/forschungsfoerderung/wissenschaftliche_infrastruktur/lis/download/allianz_initiative_digital_information_en.pdf (accessed 20 May 2009).*

DFG (2006), “DFG Position Paper: Scientific Library Services and Information Systems – Funding Priorities Through 2015”, available at:

http://www.dfg.de/forschungsfoerderung/wissenschaftliche_infrastruktur/lis/download/pos_paper_funding_priorities_2015_en.pdf (accessed 20 May 2009).*

DFG (2009), “DFG-Praxisrichtlinien Digitalisierung, Stand April 2009”, available at:

http://www.dfg.de/forschungsfoerderung/wissenschaftliche_infrastruktur/lis/download/praxisregeln_digitalisierung.pdf (accessed 22 May 2009).*

* Due to restructuring of the DFG-website, the referenced papers may not be accessible under the given URLs in the future. Future accessibility is warranted under www.dfg.de/lis .

About the author

Anne Lipp is Head of the Scientific Library and Information Systems Division at Scientific Library Services and Information Systems (LIS), German Research Foundation (DFG), Bonn, Germany. Anne Lipp can be contacted at: anne.lipp@dfg.de