Deutsche Forschungsgemeinschaft
Strategy Paper

DFG Committee on Scientific Library Services and Information Systems

Taking Digital Transformation to the Next Level

The Contribution of the DFG to an Innovative Information Infrastructure for Research

Bonn, 3 July 2012

DFG

Taking Digital Transformation to the Next Level
The Contribution of the DFG to an
Innovative Information Infrastructure for Research

In its 2006 strategy paper *Scientific Library Services and Information Systems: Funding Priorities through 2015* the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) presented its analyses and proposals on the realignment of information infrastructures in Germany. In the spirit of this strategy paper, innovative and targeted funding measures have since been implemented with the aim of creating a sustainable and well-coordinated system of information infrastructures for research. The purpose of the system is to provide users with free and comprehensive access to digital scientific information and to build a network of knowledge and data. This implies the necessity to create an appropriate legal, technical and financial framework to optimise research conditions.

The development of a coordinated system of information infrastructures for scientists and scholars should be understood as a dynamic process in which researchers' technical working environments and their needs as users are both interdependent and subject to continuous change and adaptation. Against this backdrop, the present strategy paper *Taking Digital Transformation to the Next Level: The Contribution of the DFG to an Innovative Information Infrastructure for Research* further develops the funding strategies for scientific library services and information systems which the DFG has been pursuing since 2006. It discusses current challenges to scientific information infrastructures, takes on new developments, and identifies areas to be enhanced with targeted funding initiatives.

From the perspective of the DFG, the following principles are crucial for the success of a coordinated system of information infrastructures:
First, every stage of development must be closely aligned with the interests of researchers. This requires intensive dialogue between all stakeholders.

Secondly, a coordinated system of information infrastructures for research aims not only at providing information; it also offers the possibility to reuse digital content on a broad scale and facilitates new forms of scientific collaboration, for example in virtual research environments. It is therefore essential to make digital contents broadly accessible and reusable in terms of both their technical and their legal properties.

Thirdly and finally, a powerful and forward-looking system of information infrastructures extends beyond regional and national boundaries. Interoperability with and integration into international structures must therefore always be considered.

The DFG supports the development of an innovative and internationally oriented system of information infrastructures on several levels:

It promotes dialogue at the interface between science and information infrastructure. With its grant programmes it contributes to developing and testing innovative forms of information gathering and processing, of information exchange and networking. Moreover, it advocates the importance of comprehensive and networked information services as an integral part of German and European research policy in the national and international — especially European — political discourse and consensus-finding processes.

The need for developing a coordinated system of information infrastructures poses a continuous challenge to the DFG to adapt its grant programmes to the latest technical developments and the current needs of researchers. Similarly, the bodies that sustain these information infrastructures are faced with the necessity to continually review their core tasks, adapt them to changing demands, and direct financial flows accordingly. The DFG's funding mission in the field of information infrastructure is complementary to the basic mission of the academic and scientific institutions. The success of its funding programmes and activities depends crucially on the efforts of universities and research institutes, as well as of the federal and state governments as financing bodies, to ensure the sustainability of activities and structures initiated by DFG grants.

The ability to manage a growing amount of information and information sources as well as increasingly complex communication structures requires a degree of media and information literacy that extends beyond pure research skills. This competency should be a mandatory objective of graduate and postgraduate curriculums to improve the way in which modern digital information infrastructure is used, both generally across disciplines and specifically within each subject area.

National and international cooperation between relevant stakeholders is essential for the establishment and development of a coordinated digital information space. Against this backdrop — and triggered by the DFG's 2006 strategy paper — the Priority Initiative *Digital Information* was launched in 2008 by the Alliance of German Research Organisations. This initiative brings together the activities of the partners — universities and research

organisations — in key areas such as national licensing and national hosting strategies, and it coordinates the Alliance's actions in political processes like the debate on copyright reform. The DFG's funding activities on information infrastructure are a major driver of the Priority Initiative *Digital Information*. They also gave impetus to the masterplan for information infrastructure in Germany, which was developed on behalf of the Joint Science Conference (Gemeinsame Wissenschaftskonferenz) under the guidance of the Leibniz Association (Leibniz-Gemeinschaft) and with significant participation on part of the DFG.

At the international level, the DFG contributes its experience with developing a system of coordinated information infrastructures to the design of the European Research Area. In addition to the DFG's engagement as a member of Science Europe, this takes place especially in the Knowledge Exchange network, which brings together the activities of European partners in the field of information infrastructures and advocates a research-friendly framework.

The areas in which the DFG identifies a special need for funding in the coming years are presented in the following chapters. All proposed measures aim to optimise access to scientifically relevant information, to research data, and to work and communication platforms, and make it as open as possible. They are based on the DFG's commitment to the open access paradigm.

In addition to the areas for which this paper proposes new, decidedly strategic funding actions, the overall objectives will also be promoted by continued funding in the established grant programmes for Nationwide Library Services, Cataloguing and Digitisation, Electronic Publications, and Information Management, which are not discussed in detail in this paper. In particular, it is important that, in addition to strategic funding measures, grants will continue to be available for bottom-up driven projects in all funding areas.

Nationwide Library Services

Funding for Nationwide Library Services aims to enable researchers across disciplines in Germany to access relevant publications as quickly and broadly as possible, regardless of where their institution is located. To this end, the DFG supports projects aiming at the national licensing of digital publications and databases. Moreover, it funds the so called Special Subject Collections, a system of distributed and comprehensive acquisition of specialised literature in all subject areas. It also supports projects to build specialised information portals (virtual subject libraries) for searching and accessing these holdings.

Development 2007 – 2011

By funding national licenses the DFG has supported libraries in testing and implementing new licensing models as well as approaches to bundle licensing activities in electronic-media acquisition.

To ensure better integration of digital media in the Special Subject Collections system, the development of virtual subject libraries was continued persistently. Of central importance for the further development of this funding area are the findings of the evaluation of the Special Subject Collections Programme conducted in 2010 / 2011.

Funding for specialised libraries of nationwide significance was replaced in 2009 by a new funding scheme for research libraries.

Continuing and new challenges

Funding for Nationwide Library Services will continue to focus on improving the supply of information — with an emphasis on digital media — and thus on the above stated goal of comprehensive and rapid provision. This should entail nationwide licensing of pivotal research literature as well as providing scholars with specialised literature, which has traditionally been the responsibility of Special Subject Collections.

National Licences. The current funding policy will be continued. Additional funding measures are planned to support targeted adjustments and improvements to national licensing.

Special Subject Collections. Based on the results of the evaluation in 2010 / 2011, the mission of the Special Subject Collections system is being redefined to allow greater responsiveness to subject-specific interests as well as to enhance access to digital publications substantially by making it faster and more direct. The core task of the system is to competently provide the diverse research communities with print and electronic resources

and all relevant types of media, as well as with comfortable retrieval tools. Whereas this task is no longer subject to the primacy of completeness, anticipatory collection building based on the requirements of particular research communities is still an option. The principles of serving a subject area will therefore no longer be uniform across all disciplines but defined independently by each library in dialogue with the respective research community. To effectively meet the challenges of this modified mission, the system will undergo significant restructuring, and the name "Special Subject Collections" will be adjusted in order to highlight the information service aspect.

This new orientation will most of all lead to significantly greater flexibility in funding. An important aspect is the extension of funding to support projects that develop and maintain individual services (such as information portals and digitisation services), which accordingly are to be designed to align with the interests of the respective research community.

Research Libraries. To strengthen further the alignment of funding criteria with scientific concerns, the ongoing funding for excellent research libraries should be continued in an open format and, contingent on positive evaluation, made permanent.

Implementation steps

► *National Licenses will become a permanent part of the funding portfolio. The focus of funding will be the further development of business and financing models as well as the connection with objectives in other funding areas, such as the promotion of open access.*

► *For the Special Subject Collections, a new, more flexible funding model will be developed to respond to the findings of the evaluation. This model should make it easier to reflect consistently the interests of the research communities and should include the individual development of digital information services as an integral part. The increased flexibility should also extend to the proposal and review process. For the negotiation and administration of nationwide licenses for specialised literature, grant opportunities will be created for the development of a new comprehensive support.*

► *The funding opportunity for research libraries will be assessed after the outcome of the first two calls of proposals has been completed and evaluated. If considered appropriate, it should then be converted into a regular programme.*

Cataloguing and Digitisation

The objective of the funding programme for Cataloguing and Digitisation is the digital conversion and provision of outstanding scholarly and cultural-heritage collections of nationwide importance for researchers. Digital cataloguing as well as image and full-text digitisation adhere to standards developed specifically for the respective type of material. The results are made available for the long term in national and international networks to researchers of all disciplines.

Development 2007 – 2011

With the *DFG Practical Guidelines on Digitisation*, nationwide rules and standards have been created for the digitisation of prints, the cataloguing and citation of digitised prints, as well as the exchange and dissemination of data via defined technical interfaces. With the DFG Viewer, minimum standards for the digital presentation of historical sources in Germany have been put into practice. The open-source DFG Viewer has proven to be user-friendly and is continually being enhanced in collaboration with the academic community. With the Central Index of Digitized Prints (Zentrales Verzeichnis Digitalisierter Drucke, ZVDD), a central access point and powerful media-specific aggregator of national and international portals has been created.

The creation of digital catalogues of prints has been complemented with other categories of material, including medieval manuscripts, archival finding aids, autographs and literary remains. Since 2006, bibliographies of prints published in the German-speaking areas (VD 16, VD 17, VD 18) have been systematically linked up with digitised images of these works. This digital library, which is funded by the DFG according to scientific criteria, aims for optimum availability of historical sources and retrospective full-text generation for the individual disciplines.

In addition to funding for primarily text-based media, grants have been awarded on a pilot basis for indexing and digitising non-textual materials (e.g. images and audio-visual materials). To develop and establish appropriate standards, pilot projects are funded on "Indexing and Digitisation of Object-Based Scientific Collections". The goal is to solve methodological, organisational and technical issues of object-based indexing and digitisation, to develop adequate solutions for the digital indexing of scientific collections at the national level, and to open up new opportunities for linking text-based documents with non-textual objects.

Continuing and new challenges

This programme continues to pursue the goal of the widest possible digital access to scholarly relevant materials. In its funding the DFG focuses on developing and establishing standards as well as on providing scholarly relevant core holdings. Versatile scientific reusability entails options such as automated retrieval, quantitative analysis through text or data mining, semantic analysis, pattern recognition in non-textual materials, data enrichment, contextualisation and further processing — also in virtual research environments — and is based on data mobility, appropriate access rights, and availability of the digital full text. Wherever possible and sensible, DFG funding in the field of textual materials will therefore promote the provision of digital full texts.

Digitisation of newspapers and unique materials. To further implement the digital transformation of analogue heritage, the focus of funding in the field of digitisation will gradually shift from prints to historic newspapers and unique materials (archival records, medieval manuscripts, non-textual objects). For printed materials from the period 1450 to 1800, DFG funding should enable the digitisation of a substantial portion of at least 50 per cent. Funding for the digitisation of historic newspapers and unique materials — some of which has already been granted — will primarily focus on pilot phases to define the methodological, organisational and technical framework for the various categories of material. Grants thus target the development and establishment of minimum standards for each category as well as the digital provision of scholarly relevant core holdings. The definition of selection criteria for scholarly relevant materials and their prioritisation is the task of representatives of the various scientific disciplines who are to be involved in the pilot phases of the respective measures.

Optimised digitisation methods and automated indexing models. On the basis of previous grants and progress made, average prices will have to be determined for each category of materials. Cost reductions and efficiency improvements will also be achieved through optimised digitisation methods. Automated models should be developed and applied for indexing (also of materials already digitised).

Improved visibility of project results. The visibility of all project results should be significantly increased through standardised collection-level description as well as the optimisation of retrieval-friendly access points and the establishment of interactive tools for research-relevant text enrichment and contextualisation.

Virtual merging of collections. To support the interests of researchers, topic-based access should be facilitated, also by virtually merging collections. Accordingly, funding should also extend to projects for the development and / or digitisation of materials that are distributed across several institutions.

Digitisation of copyright-protected literature. The development of models for the digitisation of copyrighted literature and other scholarly-relevant protected materials should be tested. The aim is to find, wherever there is a scientific need, suitable models for providing copyrighted or otherwise protected materials that are not yet available in a born digital format.

Consideration of public–private partnership projects. The increasing importance of public–private partnership projects needs to be taken into consideration when designing the framework for funding activities. In line with the needs of researchers, DFG funding aims to complement, not replace, the potential and expertise of relevant private-sector partners in mass digitisation projects as well as in the development of methods for optical character recognition. Even in public–private partnership projects, the DFG strives for full transparency in terms of the selection criteria for materials to be digitised and of reuse rights. Both the results of publicly funded and of public–private partnership projects should be openly accessible and fully available for reuse.

Coordination between funding activities by the DFG and the DDB. The DFG also advocates that digital materials in the German Digital Library (Deutsche Digitale Bibliothek, DDB) should be provided in a way that the data can be fully reused for education and research purposes. It is therefore in the interests of scholars that the funding activities of the DFG and the DBB be closely coordinated.

Implementation steps

► *Pilot projects should be launched to determine the methodical, organisational and technical framework for the digitisation of historical newspapers and unique materials (archival sources, manuscripts and collections of non-textual objects) and to help develop possible funding lines in close cooperation with the research communities.*

► *Taking into account technical developments as well as the need to serve researchers in the best possible way, it should be investigated how automated methods for in-depth indexing of previously digitised materials can be implemented. This should also entail support for the development of new methods, e.g. in machine learning. Moreover, the development of full-text provision needs to be supported.*

► *Through measures like the standardisation of collection-level description and the optimisation of retrieval-friendly access points, scholarly use and visibility of national indexing and digitisation projects should be increased.*

► *Depending on the needs of the relevant disciplines, pilot projects for the digitisation of copyright-protected materials should be developed in cooperation with publishers and scholarly societies.*

Electronic Publications

The purpose of the funding programme for Electronic Publications is to support the optimum creation, provision and dissemination of genuinely digital academic publications and to ensure their long-term availability at research libraries and other information service providers. Following through on the strategy paper Scientific Library Services and Information Systems: Funding Priorities through 2015, the DFG has supported the open-access provision of research results, and thus the establishment of this new publication paradigm, with targeted grants since 2006.

Development 2007 – 2011

With the network of certified open-access repositories, a point of reference has been created for the landscape of repositories, producing a standardising effect while allowing free access to publications stored in these repositories. At the same time, the network serves as a basic infrastructure for ensuring the long-term availability of digital content. This goal is the purpose of the "LOCKSS and KOPAL Infrastructure and Interoperability" project and of the cooperative reuse of KOPAL in the "DP4lib" project. The repository network, which is embedded in the European infrastructure (DRIVER), integrates institutional as well as discipline-specific repositories and offers value-added services based on the network. With CARPET, a comprehensive directory of electronic-publishing tools is available, both for publication via repositories and for the genuine publication of open-access journals. Under the Open-Access Publishing Programme, research universities raise grants to set up publication funds that can cover the fees required for publication in pure open-access journals.

Continuing and new challenges

Grants awarded under the Electronic Publications Programme will continue to focus on establishing open access as a publishing model. Preference should be given to the "gold road" to open access, i.e. the quality-controlled initial publication of scientific articles in an electronic medium that uses an open-access business model. It permits open-access provision of digital objects under legally protected conditions that also enable the comprehensive reuse of publications. Legal uncertainty has been seen as a significant obstacle in following the classic "green road", which entails the additional provision of a copy of a subscription publication.

Conversion of subscription journals to open access. The conversion of publication practices to the gold road of open access will result in organisational changes at universities and research institutes. Some of the resources for the purchase of scholarly literature and information will have to be shifted in order to set up publication funds. The DFG is aware that this change will impose an additional financial burden as long as journal subscription and licensing costs continue to coexist. To allow gradual reduction of these additional expenses, funding will not primarily encourage the inception of new open-access journals but rather provide targeted incentives for converting prestigious journals that are currently subscription-based into open-access publications.

Hybrid models and national open-access licensing. As the system transitions, the transparency of hybrid open-access publishing models, and thus the cost correlation between subscription contracts and open-access publication fees, must be analysed carefully and transparently in pilot projects. Project processes and results should be documented accurately for the scientific community in order to make successful models for converting publishing practices available as best practices. In addition, the DFG will combine its support for the national licensing of electronic journals with funding for open access in order to move existing approaches and models into the direction of national open-access licensing.

Monographs in open access. Depending on the respective scientific communities, open access will spread at different speeds and with different emphases. Transformation processes will therefore have to be organised jointly with the respective research communities. Different scientific disciplines prefer different publication formats. In order to ensure that open access will also benefit book-oriented disciplines, targeted support will go to projects that develop and test suitable models for access to monographic publications free of charge.

Implementation steps

► *In pilot projects with academic publishers as partners and / or in studies on hybrid open-access publishing, it will be explored how the accounting of subscription payments and publication fees can be made transparent.*

► *Grants should be awarded to pilot projects that correlate the concept of national licensing of scholarly journals with the assumption of open-access publication costs and move it persistently in the direction of national open-access licensing.*

► *The conversion of prestigious journals that are sponsored by scholarly societies should receive targeted support.*

► *In pilot projects with academic publishers as partners, business models for the open-access publication of monographs and / or monographic series should be developed (further) and tested.*

Information Management

Funding for Information Management provides the framework for developing and testing new tools and instruments, methods and organisational forms that help to improve retrieval and access services as well as the provision and reuse of scientific information. In terms of content, it includes bibliographic information services as well as virtual research environments and research data services, and thus covers a wide range. The Information Management Tools and Methods Programme supports primarily the development and testing of new instruments and methods. Separate funding opportunities are available for the development of new organisational structures.

Development 2007 – 2011

Regarding research library services in Germany, a significant need for restructuring was determined as early as 2006. This need has now been taken up and resulted in a call for proposals which aims towards a realignment of nationwide information services and is based on the recommendations of the DFG's 2011 *Position Paper on the Development of Library Networks as Part of the National Information Infrastructure*. The Virtual Research Environments Programme is dedicated to the establishment and development of virtual research environments as project-oriented work structures in all subject areas to optimise collaborative research activities. How to fund these structures in a sustainable way over the long term is still a largely unsolved problem.

Continuing and new challenges

The funding area Information Management is subject to constant dynamics due to the strong increase in the quantity of available digital information as well as the growing complexity of information provision. Furthermore, it is essential to reflect permanent technological change in qualitatively novel information services. A case in point are virtual research environments, which emulate new forms of collaborative research that were only made possible by the use of new technologies. The required close interaction between research and infrastructure, which is particularly prominent in the funding activities for developing virtual research environments and safeguarding research data, calls for more intensive cooperation between information infrastructure operators and the research community. This also applies to the realigned library information services. Other key issues in the Information Management funding area are international orientation and the sustainability of services. These challenges will most of all require increased commitment by funding bodies and decision makers.

Library information services. The suggestions put forward in the DFG's *Position Paper on the Development of Library Networks as Part of the National Information Infrastructure* are based on the perception that services need to be better adapted to the current requirements of modern information provision. A restructuring process is required that moves away from regionally organised structures and towards functional and nationwide oriented services that are easier to work with and benefit from for both researchers and infrastructure specialists. The DFG assumes the task to initiate the restructuring process and to support it in its initial phase. For this purpose, a call for proposals focussing on four topics aimed at the optimisation of intra-library processes as well as closer integration of library services into research processes is to be implemented.

Virtual research environments. In many research disciplines, virtual research environments have been established to varying degrees as theme-based structures and, particularly in the natural and life sciences, laid a foundation for collaborative research. Future funding measures should help establish them more among all specialties and intensify support for cooperation across disciplines. Increased reuse of individual components will become more important. A modular, overarching architecture that bundles activities should encourage the extensive reuse of existing services.

Research data. While some disciplines have long been in the practice of sharing and reusing research data and have appropriate structures in place and practical experience to draw on, other disciplines still face the challenge to develop and build the necessary organisational underpinnings and data repositories. Such endeavours should always work with a broad definition of research data that also includes information on objects and scientific collections. While there is great interest in and need for suitable solutions, there are also many open questions regarding organisational structures and responsibilities, legal frameworks and long-term sustainability. When it comes to research data, future funding activities should pay special attention to the unique needs of the different disciplinary communities. Given the high demand in this field, it is of crucial importance to design and offer appropriate grant opportunities soon.

Implementation steps

- ► *Funding opportunities are needed for the establishment of new and upgrades to existing library information services. They should mainly focus on the optimisation of library information infrastructure and on service improvements in regard to long-term availability, electronic resource management, and research-related infrastructure.*

- ► *Appropriate funding programmes are needed to establish virtual research environments both within and across disciplinary areas. Modular structures are likely to increase reuse, also across subject boundaries.*

- ► *Funding activities on information infrastructures for research data should be continued and perpetuated by establishing a regular grant programme. In addition to providing start-up aid, efforts should be made to expand existing structures that are heavily used by researchers, and to resolve open questions in overarching projects.*