II The German perspective

Rafael Ball

Introduction

The range of services and information offered by modern libraries, especially in the scientific sector, is largely characterised by the opportunities provided by digital, mainly Internet-based information. Libraries frequently do not actually own the information but function as intermediaries, providing their customers with access to the digital resources of the World Wide Web. However, the inception of digital libraries dates back to the 1970s and 1980s of the last century. As long ago as the 1980s, the Central Library of Forschungszentrum Jülich began to catalogue its holdings in an electronically integrated library system based on mainframe computer technology while the searches were performed at terminals located in the library.
Approaches to defining a digital library

According to Hacker (2000), the term ‘digital library’ refers to a collection of ‘digital’ documents. In the strict sense, this means the digital holdings of a specific library that are accessible on its own server, while in a broader sense this is a comprehensive and complex search, reference and access system for distributed resources stored in both digital and non-digital forms. With this description, Hacker comes close to the definition given by Lagoze and Payette (1998) who describe a digital library as a ‘managed collection of digital objects and services that support the storage, discovery, retrieval and preservation of those objects’. In his definition of a digital library, Umstätter (in Fuchs-Kittowski et al., 2000) describes a digital library as an integral part of a quadripartite library whose traditional areas of work have been complemented by digital services. In this definition, the ‘digital’ aspect is reduced to the physical form of the new type of publication and ignores the basic paradigm change associated with the digitisation of information and the way it is presented (database, e-journal, electronic archive) in libraries and information centres.

Wilfried Enderle (1997) divides digital libraries into three sections: access systems, electronic document delivery services and digital documents. According to his definition, one can only speak of a digital library in a comprehensive sense when all these fields are covered and readers can satisfy their entire range of information needs via their PC workstations connected to the intranet/Internet, irrespective of whether they require print or digital documents. In her fundamental article ‘Digital Libraries – Informationsform der Zukunft für die Informationsversorgung und Informationsaufbereitung’, Rusch-Feja (1999) dates the origin of the term ‘digital library’ to 1994 and relates it to the call for project proposals by the National Science Foundation in the USA as part of the initiative for ‘Research on Digital Libraries’.

How did it start? Digital library initiatives in Germany

Actors and decision-makers

The description of a particular national situation also always requires an outline of the political and administrative circumstances to provide
background knowledge with regard to the respective form of cooperation, initiatives and projects. Knowledge of Germany’s federal structure and the associated distributed responsibility for issues of science and research is essential for an understanding of the organisation of information and library policy in the country. Responsibility for science, research and education is laid down in the constitution of the Federal Republic of Germany. In the spirit of cooperative federalism, the federal government is responsible, among other aspects, for funding scientific research (non-university research). This task is largely carried out by the Federal Ministry of Education and Research (BMBF). The ambitious funding programme ‘Information as a Raw Material for Innovation (1996–2000)’ included, among other projects, GLOBAL-INFO initiated and financed by the BMBF as ‘The German Digital Libraries Project’. In 2008, BMBF no longer had any in-house group working on issues of specialist information policy. The duties of the former working group had been delegated to other departments or ministries.

The individual federal states are responsible for university policy. On the basis of the joint agreement on research funding (Section 91, Constitution), the Federal Government and the federal states jointly fund non-university research institutions (Max Planck Society, ¹ Helmholtz Association, ² etc.) as well as the German Research Council (DFG). DFG, which grew out of the Emergency Association for German Science (founded in 1920), functions as an autonomous organisation funding German science. Its duties include initiating and funding research projects of national significance as model or pilot projects. In providing information for the sciences, DFG takes on a key role in implementing and linking information infrastructures for science and in preserving our cultural heritage. In this respect, DFG is a source of important initiatives in the field of digitisation.

Against this background, it becomes clear how difficult systematic planning is on a national level and that a protracted consultation process is necessary. As a consequence, the Federal and State Commission for Educational Planning and Research Promotion was set up and in January 2008 was reshaped as the ‘Joint Science Conference’ (GWK). The ‘Representative for Culture and Media’ acts on behalf of the federal government and has already commissioned important studies such as ‘On the Way to a German Digital Library’ ³ and an ‘Overall Evaluation of the Digitisation of Cultural Assets and Fields of Action’. ⁴

For the Bundesrat (the Upper House of the German Parliament), Dr Gerald Meier was appointed Digitisation Commissioner as part of the
EU action plan ‘eEurope Digital Heritage and Cultural Content’. In 2007, the title was changed to ‘Federal Commissioner for Digitisation and Online Access to Cultural Material and Its Digital Preservation’. As part of this function, he is also a member of EUBAM (European Issues for Libraries, Archives, Museums and Preservation of Historic Monuments). This inter-ministerial committee coordinates German activities and positions in a European context. It is composed of representatives of the Conference of Ministers of Education (KMK), various federal and state ministries, experts from the sectors of libraries, archives, museums and preservation of historic monuments and also from the German Research Council. Because of its institutional cross-sectoral function, as a national agency EUBAM could in future develop strategies for conserving the scientific and cultural heritage. Apart from the government agencies or scientific institutions funded by the state, academic societies and individual initiatives, such as the German Initiative for Network Information (DINI e.V.) and the Max Planck Society, are also of great significance in initiating innovative projects in the field of digital libraries.

**Digital library initiatives – first projects**

The start of involvement in the field of digital libraries is closely linked to the project on ‘Global Electronic and Multimedia Information Systems for Science and Technology’ (GLOBAL-INFO).

The report by the Federal Government 'Info 2000 – Germany’s Road to the Information Society' and the programme associated with it 'Information as a Raw Material for Innovation by the Federal Government' for the years 1996 to 2000 demanded an active transformation of German society into an information society. One of the objectives outlined in the report was the further development of the scientific and technical infrastructure into a 'global digital library'. No fewer than 52 partners from academia, science, publishing and libraries took part in this ambitious, internationally oriented project. Existing knowledge was to be optimally exploited by all scientists as 'raw material for innovations' and the 'basis for productivity and employment in Germany'. As part of GLOBAL-INFO and numerous other associated projects, integrated scientific information systems were to be developed for use worldwide covering all aspects of a digital library from the generation of electronic scientific information and the dissemination of this information up to and including the way in which searches for information can be made.
Project funding focused on the following topics:

- supplementation and processing of content: document types, methods and tools for electronic publishing, transfer, storage, conversion and indexing;
- networking of teaching materials;
- formal description, identification and retrieval, metadata, networking;
- use of content: alerting, awareness, information networks, dissemination of information, evaluation of results, etc.;
- economic efficiency models, billing, statistics (Rusch-Feja, 1999).

As the first step, specifications for special funding measures were to be drawn up in working groups. In the special funding measures, the specific projects would then be implemented in order to develop and test models for the provision and utilisation of electronic and multimedia full texts, and for quality assurance and networking. The participants would interact through the special funding measures. This also required the reconciliation of (in some cases) divergent interests. GLOBAL-INFO was to function as a platform for the necessary adjustment of interests and for providing opportunities for funding – also for libraries. The participants were to contribute their common interests to GLOBAL-INFO and to profit from the transfer of know-how and technology.

DFG funding ceased in 2000 and the project was consequently concluded, or rather renamed ‘digital library concepts’ and turned into a forum for designing an interdisciplinary and globally oriented digital library for science (Schöning-Walter, 2003). In retrospect, it must be said that the whole GLOBAL-INFO project needs to be regarded as a failure. The reasons for this can probably be found in the complexity of the project framework as well as in the large number of partners involved and their diverging interests. The individual projects proved to have little sustainability.

**Digitisation**

The involvement of the DFG is not only related to the creation of information structures for the virtual information supply but also to the creation of specific digital content for research at German universities and research institutions. In addition to print holdings in libraries and archives, unique historical documents which can be preserved by digitisation and thus made available for research purposes are also part
of digitisation projects. Apart from the mere provision of digital documents on the Internet, DFG activities are also intended in the long term to create an Internet-based, virtual information infrastructure for research in the humanities. Digitisation activities and the provision of academically relevant digital documents is, in the view of the DFG, part of the standard range of library services in the twenty-first century.

The funding programme for ‘Retrospective Digitisation in Libraries’, launched by the DFG in 1997, pursued the following central goals:

- improving the provision of literature;
- direct access to important holdings for research and teaching;
- multiple use of frequently requested literature;
- networking to create a ‘Distributed Digital Research Library’ (Dorr, in Benkert et al., 2003).

In selecting documents for digitisation, the DFG always attaches great significance to the aspect of use.6 ‘High acceptance in a small community is much more desirable than a lower level of familiarity in a very large group of potential users that are largely unreachable.’ In terms of open access, the results of DFG-funded projects will be made available for worldwide use free of charge. Against this background, commercial interests are of little or no significance. This does not affect fees being charged for the production of hard copies or other media forms. This also includes the further use of digitised documents by publishing houses.

**Digitisation centres**

In 1997, two German digitisation centres were put into operation, funded by the DFG. As centres of excellence, the Göttingen Digitisation Centre (GDZ) at Göttingen State and University Library and the Munich Digitisation Centre (MDZ) at the Bavarian State Library in Munich were to provide advice on digitisation projects and acquire and actively pass on expertise in a number of fields (Dörr, 1999). These centres have considerable experience and expertise in the high-quality digitisation of valuable documents with regard to:

- digitisation techniques;
- systems for providing and presenting content on the Internet;
- standards and best practice;
• links to existing integrated library systems;
• ensuring the long-term availability of the documents.

Since the end of start-up financing by DFG in 2001, the two centres have provided their own funding.

The GDZ is an integral part of the Göttingen State and University Library. The services of the GDZ range from advice, planning and implementation of digitisation projects up to software development. A major focus at GDZ is the programming of the generic open source workflow software GOOBI (Göttingen online-objects binaries), with the aid of which it is possible to programme individual project-specific steps in the digitisation process.

The Russian Digital Mathematics Library (RusDML)\(^7\) is a cooperative project from the GDZ concerned with setting up an archive of digitised Russian publications on mathematics. In the first phase of the project, the most important Russian-language journals from 1866 to the present day will be digitised. The information will be accessible online via the Zentralblatt MATH portal. As part of an international collaboration between the Technical Information Library in Hanover (TIB), the Russian National Public Scientific and Technical Library in Moscow (GPNTB) and Berlin University of Technology, it is intended to make the entire literature on mathematics available to researchers all over the world in a digital form.

Another GDZ project, the Digital Journal Library (DigiZeitschriften),\(^8\) offers students and scientists access to key German research journals. Access to approximately 130 journals is provided by libraries and scientific institutions, who have to pay a fee for this service. The key measurement is the number of researchers. Since its foundation in 1999, 13 libraries with special collections have joined together, thus providing access to more than 3.4 million journal pages. The service is used by approximately 180 libraries.

The Munich Digitisation Centre (MDZ) (Schäffler et al., 2005) was established in 1997 with financial support from the DFG. The Centre was intended to undertake digitisation projects in cooperation with an external service provider. After a test phase of approximately one year, the decision was taken in 1998 for the Centre to build up its own know-how. In 2003, MDZ was incorporated into the ‘Digital Library’ section of the department for collection building and cataloguing of the Bavarian State Library.\(^9\) In contrast to the GDZ, the MDZ does not function as a service provider for external institutions but primarily concentrates on realising internal and cooperative projects. In view of
its function as a special history collection, MDZ has mainly focused on making historical source material available. Since the start of the digitisation projects in 1997 up until 2005, according to its own figures, MDZ generated and archived 2.4 terabytes of digital master data.

The Bavarian State Library has the only existing medieval manuscript containing the almost complete text of the Babylonian Talmud. Digitisation by the MDZ means that a valuable manuscript has been preserved and made available for historical research throughout the world.

More than 400 volumes of shorthand Transcripts of the German Reichstag (1867 to 1895) and its predecessors (the North German Confederation and the Customs Parliament) document the early phase of the German Empire. However, the digitisation only made use of 165 volumes with approximately 100,000 pages. These index volumes were manually recorded in SGML (Standard Generalised Markup Language) by the MDZ and refer to the corresponding image pages of the individual volumes of transcripts via inserted links.

Selected digital library projects and initiatives

From the large number of projects already completed and those still in progress, a selection of central projects will be discussed in the following providing interesting insights into the range of digital library projects in existence in Germany.

Virtual special libraries

As part of the national provision of literature in Germany, since their establishment in 1949, the special collections and special libraries, as well as the three central special libraries of the Leibniz Science Association, have had the mission of collecting, holding and making available via interlibrary loans and document delivery services the most comprehensive possible collection of important scientifically relevant literature. In the early period, this was restricted to the field of print media and grey literature, but as the electronic provision of information became established new tasks were acquired in the dissemination of electronic information systems and the digitisation of library holdings.
As a consequence, since the 1990s the DFG has funded a network of virtual special libraries (ViFa), which comprise the special collection libraries in cooperation with other specialised information centres. In accordance with their mission as central specialist portals, they will focus the special holdings, information provision and services of the libraries with special collections of national importance in a user-friendly manner. Their advantages in comparison to general Internet search engines are:

- integration of digital and conventional information resources;
- qualitative indexing and selection of Internet resources in accordance with the library-related demands of quality, stability, authenticity and long-term availability of the information;
- widest possible incorporation of the 'hidden web' into the information supply;
- seamless interlinkage of record of and access to (distributed) information via various access channels both free of charge and fee-based.

More than 40 virtual special libraries or subject portals give access to library catalogues, subject-specific databases and full-text servers, provide references to relevant fora and mailing lists, and offer supplementary services such as diaries and newsletters. These virtual special libraries are very heterogeneous in design, functionality and content. The scope and features of the respective portal are the responsibility of the respective library. Apart from a core set of content and functionalities, there are optional sectors which are provided at the discretion of each special library. Several evaluations in 2005 and 2008\(^{11}\) of the virtual special libraries have revealed an obvious need for optimisation and professionalisation in the fields of standardisation, functionality, adaptation and information. Consequently, proposals are being developed for a uniform presentation and common corporate design of the different portals. As of August 2008, more than 40 virtual libraries are listed in Vascoda, the portal for virtual special libraries, and their contents are searchable. The website of the Common Library Network (GBV)\(^{12}\) lists about 80 virtual special libraries and the online contents of sections of special collections (Online Contents Special Subject Collections (OLC-SSG)). Vascoda provides comprehensive search facilities in library catalogues, subject-specific databases and Internet resources from a shared search interface. The special libraries listed in Vascoda, complemented by the services of external partners such as Econ or MedPilot, represent a model for the cost-effective provision of scientific information.
GetInfo

Since May 2002, the databases and catalogues of the science and technology information services FIZ Chemie, FIZ Karlsruhe, FIZ Technik and also the Technical Information Library (TIB) in Hanover have been available via GetInfo. In combination with the delivery service of TIB Hanover, GetInfo can provide more than mere access to digitised information, which is the case, for example, with the subject information portals. Throughout Germany, GetInfo is regarded as the central portal for the reliable provision of information for science and technology.

By means of a metasearch in sci-tech databases, GetInfo provides access to books, journal articles, reports, university publications, conference proceedings and 'grey literature' that is not available from bookshops. The content focus of the databases is defined by the subject orientation of the project partners involved. The major emphasis is on science, technology, engineering and related areas. GetInfo received initial funding from the BMBF from the start of the project in 2001 up to 2007. Searches and brief bibliographical information on the documents found are free of charge. Abstracts and the provision of the electronic or print documents by TIB Hanover are fee based. Charges are structured according to the desired speed of service and the way in which orders are placed and documents delivered, and a clear distinction is made between academic and commercial users. For commercial users, GetInfo also provides a full service comprising both searches and also the provision of the required literature.

eSciDoc – a scientific information, communication and publication platform for research

As part of the e-Science initiative ‘e-Science – Using the Internet for Science’ of the federal government, virtual science environments are to be set up in cooperation between science and industry within the framework of the ‘IT Research 2006’ funding programme. The ‘eSciDoc’ project is part of the programme and is being implemented by the Max Planck Society (MPG) in cooperation with FIZ Karlsruhe. The project, which began in 2004 and will continue until mid-2009, is being funded by BMBF to the tune of about €6 million. Both the Max Planck Society and
FIZ Karlsruhe are involved as cooperation partners. The Max Planck Society is one of the group of non-university research institutions in Germany which also includes the Helmholtz Association of National Research Centres as well as the Leibniz Science Association and the Fraunhofer Society. The Max Planck Society has more than 80 research institutes in Germany and abroad. Since the establishment of the Max Planck Digital Library in 2007, MPG has had at its disposal a central service facility for all issues concerned with literature and information provision in its institutes and thus plays a pioneering role within the non-university research institutions. The Special Information Centre Karlsruhe (FIZ Karlsruhe) is one of the service facilities of the Leibniz Science Association (WGL). It regards itself as a service partner for science and research. FIZ Karlsruhe provides access to Sci-Tech Online information (STN International) and with FIZ AutoDoc offers a delivery service for academic and commercial users.

The e-Science project focuses, on the one hand, on access to scientific information for the general public and also for individual research groups and, on the other hand, on the provision of a flexible infrastructure for scientific work. The Max Planck Society makes the content available and serves as a test bed for applying the project findings. Karlsruhe Special Information Centre (FIZ Karlsruhe) organises the data storage and operation of the open-access publication platform while the service partner takes over the central development work and guarantees the long-term availability and long-term archiving.\textsuperscript{13}

eSciDoc is regarded by the project partners as an open and flexible infrastructure coupled with a productive system. The project planning includes a consideration of and close cooperation with other projects in the e-Science field such as KOPAL, D-GRID and Shibboleth. The goals of the eSciDoc project are:

- to ensure permanent access to the research results and research materials of the Max Planck Society and seamless integration within eSciDoc as well as integration into an emerging, global scientific knowledge space;
- to provide effective opportunities for access to information for scientists of the Max Planck Society and work groups;
- to support scientific collaboration in future e-science scenarios (Dreyer et al., 2007).

The functional requirements of the project are:\textsuperscript{14}
• the provision of community-specific solutions which serve the needs of a specific research group;
• the provision of standardised interfaces for other communities to re-use data and functionalities;
• the provision of user interfaces and viewing environments.

To this end, software tools based on open-source technology are to be developed and disseminated as an open source. When fully developed, eSciDoc will consist of several components:

• **Scholarly Workbench** – research-specific cataloguing, processing and evaluation of materials previously difficult to access by scientists;
• **publication management** – support for scientists in the dissemination and archiving of scientific research findings in accordance with the principle of open access;
• **eLib** – permanent access to the supply of electronic information (e-journals, databases);
• **eLab Journal** – permanent documentation of research findings (results, methodology, publications).

**BAM – joint internet portal for libraries, archives and museums**

The BAM portal is a very good example of a cross-sectoral project in which, as cultural partners, libraries, archives and museums provide digital content for the establishment of a unique information pool of cultural assets in Germany. For the first time, it is possible to retrieve subject-related material on different topics from the three cultural institutions via one unified search interface beyond institutional boundaries. Because of their different missions, materials and cataloguing methods, the creation of a prototype for a search across the various types of service providers represents a special challenge. The project started in 2001 with the following partners: Baden-Württemberg Library Service Centre (BSZ), Baden-Württemberg State Archives and the State Museum for Technology and Work. By 2006, data from 14 institutes and six associations had been integrated into the BAM portal and in 2007 the partners were joined by the Prussian Cultural Heritage Foundation and the Federal Archives – both important partners for the
project. Initial funding by the German Research Institution was terminated in 2007 and the project will be continued as a consortium of the participating institutions.

The BAM portal was set up in three project phases:

- **1st Phase: 2001–2002.** Creation of a central database (metabase) with materials on hydropower and industrialisation.

- **2nd Phase: 2002–2003.** Central database complemented by distributed searches in external data archives and semantic standardisation of the metadata used. This was accompanied by studies on the applicability of established library cataloguing systems for the purposes of the BAM portal and on computer linguistic methods for optimising retrieval.

- **3rd Phase: 2005–2007.** The inclusion of other cooperation partners and conversion of searches from a distributed search model to a data warehouse approach (search engine technology) incorporating the Apache Lucerne search software.

The technology of the BAM portal is based on the Lucerne search engine software with a central data index. Data integration was implemented via data-specific XML exchange formats.

Altogether, BAM provides access to more than 38 million datasets (as of October 2008). These datasets are processed for searches via a unified search interface. The search results are represented in an overall list (short-title entries) and can also be specially retrieved according to institute via the hit navigation. A link to Wikipedia was established in August 2007. This provides direct access to documents in the BAM portal from more than 500 Wikipedia articles.

**Conclusion**

The future success of the digital library movement in Germany will depend on two factors: sustainability and integration into European and international concepts. In the past, the lack of visibility and sustainability in digital library projects frequently led to the project as a whole being discontinued when DFG or other institutional funding came to an end so that no or no permanent transfer to the scientific community could be achieved.

What is crucial in the development of digital libraries is the question of concepts concerning content. The time has long since past when it was
simply fashionable to digitise collections and make them available to the community. The tension between quality and quantity has become a central question in this connection. Moreover, what must be criticised from the print age is the often ridiculed ‘ton-weight’ ideology of the librarians’ strategy of ‘all you can eat’. Either one has the vision that one day all contents in the world will be available digitally and accessible globally without barriers in a ‘world content pool’, or on the other hand one accepts the impossibility of such a vision and assume the duty of carrying out choices concerning content.

Besides this, the sense of single digital libraries and their portals must be scrutinised. The contents are already too strongly fragmented and will increasingly be found only in an incomplete state. If by the same token the search for digital contents calls for integrative search engines which must operate on several meta levels in order to index at least parts of the scattered digital inheritance, then nothing is gained and the user is once again referred back to his own resources.

Today digital libraries as an end in themselves have survived on their own. The contents and indexing must be integrated into the electronic information environment of libraries. They should no longer complement library contents separately with their own portal, but must be integral and structural parts of what is on offer. Digital libraries and digitisation have been for some time a long-term assignment for libraries, but they must distance themselves from project status and become an integrated part of collection development, indexing and use.

Notes

7. http://www.rusdml.de/rusdml/?page=rusdml

References