Digital library activities in Germany: the German Digital Library Program GLOBAL INFO
Schmiede, Rudi

Preprint / Preprint
Arbeitspapier / working paper

gesis
Leibniz-Institut
für Sozialwissenschaften

Mitglied der
Leibniz-Gemeinschaft

Digital Library Activities in Germany
The German Digital Library Program GLOBAL INFO

Rudi Schmiede

Darmstadt University of Technology
Dpt. of Sociology
Residenzschloss
D-64283 Darmstadt
schmiede@ifs.tu-darmstadt.de

Abstract

Several digital library projects have emerged in Germany in recent years. The German Digital Library Programme GLOBAL INFO, which is funded by the federal ministry for education and research from 1998 to 2003, is about to become the most important of them. It has the aim to advance for the single scientist „optimal access to the worldwide electronic and multimedial information on full texts, literature references, factual databases and software" at every workdesk. The programme requests the close cooperation between all the parties taking part in the processes of provision of information and documents. A first wave of projects has recently started. Their areas of concern are tools and standards in the production of documents (along the publication chain), description and retrieval of documents and dealing with their heterogeneity (i.e. metadata and retrieval, distributed systems), and provision and payment systems for published documents in distributed systems. Some more projects in other fields (dynamic documents, large distributed systems, administration systems for users and integration of electronic business models and payment systems) have been or will be handed in and evaluated in the course of 1999.

1. Introduction

Electronic information and communication in the sciences have spread rapidly in recent years. Electronic catalogues of books and other documents, full texts, or databases of facts are becoming accessible on a national and international scale. World-wide retrieval and communication, electronic publishing and cooperation are widely used in natural and technological sciences, but also in medicine, whereas in humanities and social science they are just about to become a common medium. Recently, the implementation of dynamic documents has gained momentum.[1] This is the reason why in Germany in the 1990ies – as in other countries, too – several initiatives emerged in the field of digital library development. I first want to sketch briefly their scope and achievements.

2. The Institutional Framework of Digital Library Activities in Germany

To understand the diversity of digital library activities in Germany, it should be noted that in Germany the funding of research and development as well as that of technical applications has several sources. The universities are financed and run by the sixteen states (Laender). Private universities were founded only recently; they do not play any considerable role up to now. So each university is moving and locating itself within a network of local, state, federal and international influences and connections. For digital library activities at universities this is to say that many of them take place in the context of laender-specific initiatives and policies. In average, the outside finance of research contributes about 20 – 25% to a university's budget.

This additional funding of research and development is provided by several institutions. The most important one is the „Deutsche Forschungsgemeinschaft" (The German Research Association) which is financed by the federal government (about 60%) and the states (around 40%). This institution is traditionally responsible for a considerable funding of conventional and innovative activities in the field of libraries. Contemporarily, it supports three different programs in the field of digital libraries: (a) a program called „Modernization and Rationalization in Scientific Libraries" with 14 projects; (b) a program on „Electronic Publishing", which aims at library services for electronic publications and multimedia and comprises about 11 projects; (c) a program on restrospective digitilization with about 20 projects. In addition, in 1997 a program, essentially dominated by computer scientist, started with the title

„Distributed Management and Handling of Digital Documents" with about 25 projects was started. As a whole, the funding by the Deutsche Forschungsgemeinschaft covered (in 1998) around 63 projects on digital library development, amounting to a funding sum of more than 11 Million DM in 1998. The DFG currently is adapting its research and development funding towards new areas of innovation in digital libraries; the traditional field of funding of the so-called „Special Collection Fields" (Sondersammelgebiete) is now being converted into the development of specialised digital libraries [2].

The second important financier of digital library activities is the federal Ministry of Education and Research (Bundesministerium für Bildung und Forschung BMBF). The ministry is funding one of the electronic document delivery services in Germany, called SUBITO, in the meantime in its version SUBITO II. Another large project funded in the area of electronic publishing is the development of „multimedia books" in the sphere of teaching and education. In addition, this ministry has launched the program GLOBAL INFO, to which I will come back below. The German Network Association (Deutsches Forschungsnetz e.V., DFN), an association financed substantially by this ministry, is a third important player in this field. The DFN supports innovative projects in the sphere of network based scientific information systems (e.g. projects building MathNet and PhysNet have been supported; the author of this paper is currently engaged in projects accomplishing a SocioNet [DFN] and a virtual library of social sciences [DFG].

A further feature of the German library (and, consequently, also the digital library) scene is that scientific libraries, too, are organized by responsibility of the states (Laender). Historically, this meant that each state had its own library system. Fortunately, in the 1980ies and 1990ies several of the Laender library systems united into a limited number of associated systems (Verbuende). So today we have less than ten such associations still using different library systems. In the states of northern Germany and in the middle and parts of the east the Dutch PICA system is applied; in Bavaria cascaded Siemens softwares (SISIS, Sokrates) are used; in the rather successful South-West-Verbund considerable funding and development achtivities have been invested into a new Verbund software on the basis of the Dynix program system; the outcome, however, is dubious, yet; the same is true for the largest state of Germany, North-Rhine-Westphalia. In this state, furthermore, in 1998 a separate digital library initiative (the Digital Library NRW) was started. In general, however, it might be said that the systems applied are technology of the 1980ies, i.e. they depend upon a centralized database with strictly formalized data formats to which local information is added. In the Eastern German Verbund of Berlin-Brandenburg a novel generation of library administration is being developed and tested. On the basis in the first phase of the Dynix library software, then based on

own system developments, the project leader, the (mathematicians') Konrad Zuse Institute for Information Technologies at Berlin (ZIB) is developing an ambitious technology comprising some 200 odd distributed databases. The success of this project will be of crucial importance for the entire German library scene in terms of the breakthrough of new distributed technologies.

As a whole, the funding of digital library activities in Germany does not seem to suffer from insufficient financeial resources. The DFG is spending about 11 Mill. DM; the DFN's own estimation of its support of DL activites amounts to about 10 Mill. DM. GLOBAL INFO, dependant upon the evaluation of new projects, will contribute about 5 – 15 Mill. DM p.a. to this initiative. In addition, one has to take into account the funding of the states, which will amount to a sum of significantly more than 10 Mill. DM p.a. Added up, the amount of support for digital library activities in Germany is substantial: Together with state fundings of the scientific information centers, of which at least part goes into digital library activities, it amounts to a sum of something between 40 Mill. and 60 Mill. DM (in 1998). This sum is not trivial; in the contrary, it is on a competitive level with funding in the USA, the UK or Australia, to name the most innovative initiatives.

3. Digital Library Projects in Germany

In comparison to the funding sums the achievements in the digital library sphere up to now are rather modest as might be guessed from the following, probably incomplete, listing of activites. [3] Electronic access to scientific information and publications is still limited and fragmented.

(a)   The traditional databases provided by the *scientific information centers* (Fachinformationszentren: FIZ Chemie, FIZ Technik, FIZ Karlsruhe, IZ Sozialwissenschaften) are in the meantime in most cases accessible on the Web using an ordinary browser. But their contents are still limited to bibliographic references and in many cases abstracts; they do not provide any substantial full text basis..

(b)   All important libraries and library systems (Verbuende) have created *OPACs* in the last years. But most of these OPACs are located within specific library systems. Only in a few cases (e.g. the national library Die Deutsche Bibliothek DDB at Frankfurt am Main and Leipzig, the national journals database at the Deutsches Bibliotheksinstitut DBI at Berlin, some regional systems) the Z.39.50 protocol has been implemented. So at least technically cross access is possible. But organizationally it is hardly realized. This is the aim of *virtual catalogues*: In the Karlsruhe Virtual Catalogue (KVK) the simultaneous research in several major catalogues of the DDB and regional systems (Verbuende) is possible; the Darmstadt Virtual Catalogue (DVK) comprises some 80 libraries of the regi-

on and has integrated Regex in a simple-to-handle Web form.

(c) There are some very useful special *bibliographies*, as e.g. in computer science collections for the topics „database systems" and „logic programming" on the dblp server at Trier which include more than 80.000 browsable journal references and more than 4.000 links to persons working in these fields. The collection of Computer Science Bibliographies at Karlsruhe CSBK contains about 1.200 bibliographies with more than 800.000 references and about 30.000 links to documents in full text. At the library of Regensburg university information on electronically available journals is collected and maintained systematically.

(d) In the field of *retrieval* some search engines with gathering and broking facilities are implemented and operative. The FreeWAIS system developed at Dortmund university is operating in several digital library systems. The GERHARD retrieval system at Oldenburg has integrated an UDC classification to provide more high quality research results for German language documents in the Internet. In the OSIRIS project at Osnabrueck university library computer linguistic techniques are implemented to generate knowledge bases from a given bibliographic or full-text database by cross-referencing free-text terms, classification codes or bibliographic data to authority files and so creating contexts of content related knowledge for text items.

(e) At practically all universities *local full-text servers* are in the process of being set up or already operating. There are several reasons for the acceleration of this development: Electronic publications of university members and institutions, especially dissertations, are made accessible at many places; the DFG funded project Diss-Online is creating rules and standards for dissertations. Since 1998, many universites and/or states have bought licences from suppliers of electronic journals or other scientific information (e.g. Elsevier, Academic Press, Springer, APS, IEEE, ACM, Current Contents, juris etc.) to make them available and accessible at university or state level. However, by and large access and availability of these sources is still chaotic: You either know where you have to look and research, or you will find it very difficult to retrieve anything systematically. The ELDORADO project at the university library of Dortmund is using the Austrian Hyper-G resp. HyperWave internet server technology to build an attractive hypertext and hyperlink structure.

(f) *Delivery of electronic documents* on a commercial or fee basis is still at its beginnings. The already mentioned SUBITO document delivery service of libraries is still in its starting phase and under development. The Technische Informationsbibliothek TIB Hannover has implemented the electronic searching and ordering of technological documents, but not yet their electronic delivery. The document delivery by publishers or journal agencies like Swets & Zeitlinger or EBSCO is usually working on a institutional subscription basis but not yet along models of pay per view or pay per download. Furthermore, the user has to enter the proprietory system of the supplier and to move there. The pre-competitive linking of various private content bases is tested for a few isolated suppliers.

(g) In some scientific disciplines, *electronic pre-publishing, publishing and networking* have been started. In physics, the world-wide pre-print network PhysDoc with some mirrors in Germany is the most important actual information network. In mathematics, the German MathNet project is just being finished. It has furnished a user-driven information system with a network of information servers at the mathematical departments in Germany and secondary homepages, a pre-print server MPRESS, a persons directory in mathematics (Persona Mathematica), an index of mathematical software (Elib), a directory of relevant resources in the Internet (MathNetLinks) and a search engine for mathematical resources in Germany (German Broker). This disciplinary information, communication and networking system is now being extended by the IMU's Committee on Electronic Information and Communication CEIC to become a world-wide infrastructure in mathematics. On a sub-disciplinary level a similar system in computer science is realized in form of the Electronic Colloquium on Computational Complexity ECCC at the university of Trier. This system includes a reviewing and publishing system for scientific papers, access to experts and a discussion forum. The international Flybrain project on the nervous system of Drosophila as a world-wide model project in biology is well-known.

(h) Last but not least: In the fields of infrastructure and basic techniques relevant for digital libraries various activities are carried on. To mention only the most important ones: Several German working groups are taking part in international standardization commitees such as the W3C commitees or the Dublin Core working group. The areas of Metadata, of new document description languages and formats like XML with its derivates and RDF, and of new document identifiers like DOI are covered here. The question of long-term archiving of electronic documents is unsolved yet. The German Research Association (DFG) is funding a program on „Distributed Digital Documents", in which computer scientists are developing basic technologies for distributed library systems. The metatagging of documents and databases, intelligent agent technologies and the like will be necessary pre-conditions for successful retrieval in large distributed systems.

It can be seen from this listing of relevant activities in Germany in the field of digital libraries that various projects, efforts and initiatives in all areas relevant are under way. There is a considerable amout of funding of these activities, too. At the same time, however, it must be said that most of these efforts are more or less particularized, not part of a general concept or at least an agreed division of labour. The reasons for this fragmentation are the institutional barriers and segmentations between the different

funding lines mentioned earlier, the political limitations rooted in the federal structure of the political system of Germany, and, last but not least, the different interests of the various players in this arena: Authors, learned societies, universities, publishing houses, scientific information centers, libraries, and state agencies. It seems that for future development of digital libraries these institutional, economic, political and social problems have to be tackled by new cooperation and moderation structures among all parties taking part in this process. It is this view of the problems which is at the basis of the German GLOBAL INFO program, the forms of its organization, its funding conditions, and its perspectives.

4. The German Digital Library Program GLOBAL INFO

4.1. Backgrond

For several years in the 1980ies and in the beginning of the 1990ies electronic information and communication was introduced and funded in the form of „information counselling" (Informationsvermittlung), i.e. persons were educated and paid for doing researches in specialized databases each of which had its own syntax and retrieval language. Only with the spread of the internet and especially its popular offspring, the WWW with its graphical user interface, the idea of the priority of access of every scientist in teaching, research and administration began to become realistic.

This is the background why in Germany in January of 1995 a cooperation agreement between some important learned societies – the German National Mathematical Society (Deutsche Mathematiker-Vereinigung DMV), the German Physical Society (Deutsche Physikalische Gesellschaft DPG), the Society of German Chemists (Gesellschaft Deutscher Chemiker GDCh), and the German Computer Society (Gesellschaft fuer Informatik GI) – was agreed upon and resulted in the so-called „common initiative of German scientific societies for electronic information and communication" (Gemeinsame Initiative der wissenschaftlichen Fachgesellschaften zur elektronischen Information und Kommunikation, in short IuK-Initiative). The societies agreed to cooperate in the field of electronic information and communication and to build a decentralized and distributed infrastructure to provide scientific information as complete as possible and well-structured at minimal costs. World-wide standards and introduced methods shoud be used. The different scientific cultures should cooperate to learn from each other and to support their common aims. Organizationally, they agreed to install locally information delegates in the faculties and departments, on the national level to cooperate in the IuK-Commission into which representatives of the learned societies and of universities are delegated and to hold a yearly workshop on status and perspectives of electronic information and communication. In this year the workshop took place at Jena in March of 1999 on „Dynamic Documents". [1] The work is supported by work groups, e.g. on metadata and on electronic journals, and by one general and several disciplinary mailing lists.

In 1996 and 1997, other scientific societies – the German Society of Educational Sciences (Deutsche Gesellschaft fuer Erziehungswissenschaften DGfE), the German Society of Sociology (Deutsche Gesellschaft fuer Soziologie DGS), the Association of German Biologists (Verband Deutscher Biologen VDBiol), the German Society of Information Technology Engineering (Informationstechnische Gesellschaft ITG), and the German Psychological Society (Deutsche Gesellschaft für Psychologie DGPs) – have joined this agreement so that a considerable part of German science is taking part in the initiative. This is the background from which – together with parallel activites of publishing houses and others – the idea of the funding programme GLOBAL INFO was developed in 1997.

4.2. Characteristics of the program

This program – its full name is „Global Electronic and Multimedial Information Systems for Natural Science and Engineering" (Globale Elektronische und Multimediale Informationssysteme für Naturwissenschaft und Technik) – is to run for six years (1998 – 2003) and has a planned financial volume of about 60 Mill. DM. With a funding rate of 50% in effect 120 Mill. DM of investments into digital library activities should be mobilized. The program includes all the disciplines whose learned societies are active in the IUK initiative. The most important aims of the IuK initiative have been integrated into GLOBAL INFO. The program shall – as it is said in the ministry's press release –

- advance „optimal access to the world-wide electronic and multimedial information on full texts, literature references, factual databases and software" for the single scientist.
- The access to bases of information „which are stored in distributed information systems, shall be accessible from their work-place computer. Keyword is the 'digital library'."
- The program aims at cooperation of all parties acting in the process of providing, distributing and using information; it is to integrate the producers (i.e. the authors, represented by the learned societies or the scientific publishing houses), the distributors (publishing companies, book sellers, scientific database providers, and scientific libraries) and the consumers (i.e. readers or users, represented again by the learned societies and representatives of university departments).
- Cooperation of these parties in project consortia in the precompetitive sphere is the condition of funding for projects.

This last condition in the meantime has been specified by some rules: In a project at least two universities, two publishers and two other distributors have to participate in order to fulfill the prerequisites for funding. In addition, at least four of the scientific disciplines organized in the IuK-Initiative, have to take part. The emphasis of the program is on development projects which result in visible improvements of access and supply of scientific information for the single user, not on research projects. It was the high threshold created by these conditions for the design of project proposals which was responsible for some delay in the initial year of the program in 1998 and resulted in the creation of rather large and complicated project clusters, as will be shown below.

Institutions and companies taking part in GLOBAL INFO have to provide half of total project costs as their own contribution. For universities, other funding is accepted as part of their contribution, for normally these institutions because of their public finance funding have hardly any possibilities to redefine budgets and persons for exclusive use in GLOBAL INFO. Because of a principal political decision of the ministerial department responsible for GLOBAL INFO, institutions and organizations which are totally financed by the state and want to take part in the program, in the first two years of its course have to finance their participation by themselves. The rationale of this political line is founded in the intention to show up the real costs of supply of information whereever possible and to economize scientific information processes by that means; the inroad for market forces – even if they are monopoly or oligopoly force – into the scientific information culture might to a certain extent be paved by this feature of GLOBAL INFO. However, the relation between public provision of a basic corpus of scientific information and purchase of additional information value which in one or the other way has to be financed privately, is open yet. But, to my knowledge, this is the case internationally. In any case, here a latent conflict potential between the parties participating in GLOBAL INFO, but also e.g. between publishers and libraries or the scientific community and great publishing companies in general is visible. In GLOBAL INFO this problem appears also in the form of different perspectives of those disciplines who themselves try to organize and control a substantial part of their information infrastructure (e.g. physics, mathematics) in comparison to those whose information sources are to a great extent private property (chemistry) or those who are not big enough or not tightly enough organized to constitute an interesting market for the big players (humanities and social sciences, but also computer science). The future development of digital libraries will decide whether substantial revirements of functions and influence between the traditional institutions and players are one of the results of the digitization of scientific information and communication.

If one considers the technical, economic, social and political frame within which GLOBAL INFO is gaining shape it seems to be very clear that this program will either adjust to the international standards and development in distributed information infrastructure (i.e. have real global dimensions) or it will perish. The basic global medium of the existing distributed information infrastructure is the internet and especially the WWW. GLOBAL INFO will have to be compatible to this structure, i.e. to develop elements and single systems being connective to the Web. Within these modular single systems the dimensions of protocols (e.g. HTML, Z.39.50 etc.), of metadata (e.g. DC standards, thesauri, classifications of various kinds etc.), and of heterogeneity (relevance, consistence, scope) are not independent of each other, but are one system structure. These systems should be simply structured, lean and flat (with few abstraction layers). Because of the interrelated character of the dimensions of this modular systems, they may be proprietary as long as the condition of connectivity is fulfilled. Within these systems distributed objects should support a multidimensional functionality. Since the distributed system at large is the Web there will not be *the* system architecture of GLOBAL INFO. A minimum number of protocols, interfaces, and requirements for interoperability has to be agreed upon (but they, too, are essentially determined by global standards).

The general aim of GLOBAL INFO, therefore, can be defined in the following way: The program is to advance – in the context of a global, principally open system of information provision and brokering – the development and implementation of modular user driven single systems for various functionalities.

Because of the criteria of interdisciplinary cooperation of all parties and of the reliance upon international standards and methods an independant evaluation of project proposals by internationally recognized experts is of crucial importance. Therefore, the first wave of project proposals – after a first evaluating vote by the GLOBAL INFO CONSORTIUM (see below) – has been examined by several national and international experts. In the near future these experts will be organized in the form of a standing experts board which will have regular meetings.

4.3. Organizational Aspects of the program

The mobilization of independent expertise is one of the organizational features of GLOBAL INFO. In order to make sure that this and the other principles listed in section 4.2 of this paper will be realized early in 1997 the *GLOBAL INFO CONSORTIUM* was installed as board of GLOBAL INFO to guide the programme as a whole and to decide upon the essentials of its direction and organization.

The GLOBAL INFO CONSORTIUM (GIC) which later in 1997 was re-eclected by the assembly of participants in GLOBAL INFO consists of 5 representatives of learned societies, 4 representatives of publishing houses, 1 representative of scientific information centers and 1 represen-

tative of scientific libraries. Contemporarily, the learned societies of mathematics, chemistry, computer science, education sciences and sociology are represented in GIC. The publishing companies are represented by one delegate of a small publisher (Teubner), one delegate of a medium-sized house (Spektrum) which is, however, part of one of the international companies in publishing (Holtzbrinck), and a representative of the largest German scientific publishing company (Springer) which recently has been taken over by one of the big players in the world market (Bertelsmann). Until recently, a delegate of one of the big international publishers (Elsevier) was member of GIC; he left for personal reasons and remains to be replaced. Representatives of the Ministry (BMBF), the programme organizing institution (PTF) and the German Research Association (DFG) have guest status at the GIC. The *project administration* of GLOBAL INFO, consisting of a project manager, supported by an assistant and a secretary, resides with the Society of German Chemists (GDCh) at Frankfurt am Main; the *GLOBAL INFO server* is located physically and administered technically at the university of Oldenburg [*http://www.global-info.org/*]. Responsible for this organization structure is the *chairman* of GLOBAL INFO, the author of this paper, supported by the project administration and especially by the assistant with part of his work. In GIC the constituting parties of GLOBAL INFO are cooperating in one board. The cost of this project administration including the substantial travel expenses of its members is about 600.000 DM p.a.

To create a basis for cooperation of the various participants in the program locally a number of *pilot projects* (called pre-projects, Vorprojekte) have been and are still funded by the ministry, each of them endowed with 100.000 DM and aiming at the sollicitation of cooperation between different scientific disciplines locally and the integration of publishers into this cooperation. 16 of these projects have been granted to universities, 13 to publishing houses for the period 1998/99, i.e. a sum of 2,9 Mill. DM is spent on these pilot projects. Above that, several institutions with special interests in this field (mainly foreign publishing companies, German science database providers and scientific libraries) have been invited to join the program. In the first wave of project development 1998 and 1999 only these groups of participants (universities and publishing companies with pre-projects and acknowledged interest institutions) are intitled to hand in project proposals (under the conditions stated above). In the 16 university pilot projects 82 different university departments are involved. Among them are 16 departments of computer science (i.e. at every university the department of computer science is taking part), 15 departments of each of the other funding members of the IuK-Initiative, mathematics, physics, and chemistry, 5 departments of biology and 3 departments of electrical engineering; the remaining departments represent various scientific disciplines. In addi-

tion, 13 university computing centers and 11 university libraries are involved. [5]

A first framework for contents and the work of GIC was established by the definition of five „focal points" (Schwerpunkte) in the preamble of GLOBAL INFO [http://www.global-info.org/]. These focal points were

- Organization of content: document types, procedures, and tools for electronic publishing, transfer, storage, conversion, and indexing.
- Networking of educational material.
- Formal description, identification and retrieval, metadata, networking.
- Usage of contents: alerting, awareness, information pooling, information mediation.
- Economy models, billing (micro-billing) and accounting, statistics.

This provisional definition of focal points at the same time constituted the first internal organization structure of GLOBAL INFO: Each focal point was coordinated by two members of the GLOBAL INFO CONSORTIUM, one from universities and learned societies, the other one from publishers, scientific database providers or libraries. In addition, the emerging working parties along these focal points named representatives to keep contact with the other working parties and to coordinate the various activities and project proposals.

These working parties in 1998 have been the framework within which project generating activities and cooperation relations were sollicitated and organized. The five working parties oriented towards the five focal points have been the organizational focus of work in the past year.[4] The limitations of this kind of organization, however, have also been shown up, for, of course, the reality is not segmented into five clean focusses of work. There was a tendency to subsume various projects to one umbrella defined loosely by a focus. Since in their efforts in project generation, in the external evaluation and the consequent final decision by the ministry, the various working parties were not equally successful, the future internal organization of GLOBAL INFO is subject of intensive discussion. Probably, project focusses will be created as the way to organize the main subject areas of GLOBAL INFO. The focal points will continue to have an important guideline function.

4.4. Development projects in GLOBAL INFO

Four projects clusters with a considerable number of sub-projects have been approved in February and March 1999 and are now in their starting phase. Before commenting on them it has to be mentioned that already since the beginning of 1998 the project „Interdoc" has been funded in GLOBAL INFO. This project was a continuation of the „*MeDoc*" project in which from 1995 to 1997 an ambitious digital library system for the computer science has been conceptualized and in parts been developed. This

system combines a full text store and retrieval, reference databases, access control, licensing and payment facilities and a broker to handle these various parts. Unfortunately, only parts of the functionalities planned could be realized in the system. But its full text basis (apart from a large number of documents without fees) comprises some 70 – 80 electronic books provided by several publishers and licensed to user institutions. Since it was a vital interest of the publishing companies to maintain this full text supply, in „*Interdoc*" which is ending these days, some features were improved, licensing activities and payment handling were continued, and the system was tested and evaluated by representatives and user institutions of other scientific disciplines (physics, sociology, pedagogics) as to its usability for these science branches.[6] The result of this evaluation was not very positive. Especially it became evident that the basic technology of the system is no longer up to contemporary state-of-the-art so that it has to be thought over thoroughly which parts of the system should be continued and which ones have to be innovated.

This is the starting point for one of the new project clusters with the somewhat clumsy German title „*Testing of forms of electronic supply, business models and payment methods on a test platform for GLOBAL INFO*". It derived from the work on focal point 5. Four sub-projects will be funded under the preconditions that a definitive decision on the technological perspective of the Interdoc system will be achieved and that a positive interest of the publishing companies, underlined by their financial engagement, is given. The sub-projects deal with

- the improvement of MeDoc-Software: A new database, introduction of CORBA, new JAVA applications, improvment of GUI; extended tools for user administration and for management of distributed databases, extended retrieval.
- statistics and a survey on user interests and user administration, a new database protocol on the basis of JDBC, improvement of license administration and new licensing models, acquisition of new contents from publishers.
- the integration of (non-anonymous) debit accounts and of (anonymous) EC-Moneycard functionalities for electronic payment, especially for a model of online transactions with micro-payment.
- the implementation of the banking SET standard and the integration of electronic money payment systems (CyberCash, Millicent or Ecash); implementation of statistical functionalities relevant for these facilities.

If this project is successful it should accomplish a substantial step forward in the area of business models and payment systems.

A second project cluster approved carries the attractive name *CARMEN* which is the acronym for „*Content Analysis, Retrieval and Metadata: Effective Networking*". It comprises the essential contents of focal point 3 and has its focusses on metadata and retrieval and on dealing with heterogeneity of data and databases. In this broad area nine sub-projects have been approved:

- Digital Signatures and a „Metamaker" with tools for document authors to provide metadata as part of their documents; establishing of a PICS Labelling agency
- Development and implementation of RDF definitions for metadata concerning terms and conditions of documents and for archiving
- Development and implementation of persistent identifiers and metadata for different document types, definition of their relationship, administration of name spaces, developing a prototype for the management of metadata and persistent identifiers.
- Metadata based indexing of scientific resources: Conversion of MAJOUR metadata (Springer) into DC/RDF metadata, analysis of document URNs to automatically detect relations among documents, implementation in a testbed.
- A Document Referencing and Linking System: Development of a metadata extractor (automatic registering of metadata in documents), improvement of Harvest to create an integrated retrieval and hypertext system (including XML).
- Interdisciplinary Information Systems: Making the contents and metadata in MathNet and PhysNet compatible and testing of retrieval under these conditions. (This sub-project is to start half a year later to be able to rely upon the first results of some other projects.)
- Dealing with heterogeneity of textual information in different data types and content description methods: Development of transfer components between different qualities of document description, comparison of automatic methods to create metadata.
- Development of concordances of classifications and thesauri in mathematics, physics and sociology (including the Dewey Decimal Classification DDC, the German RSWK system of Die Deutsche Bibliothek and the other important German system, the Regensburg Classification). In this project a German-English concordance will emerge, by this way, too.
- A project with a special character is the support of and participation in international cooperation in W3C, DC and other commitees in the creation and definition of standards. These activities will be supported, but they will be located with the central project administration of GLOBAL INFO, since they are relevant for all parts of the program.

The project cluster CARMEN covers a broad range of relevant problems and developments in the area of the infrastructure of digital libraries insofar as the description, structure and retrieval of documents is concerned. It is to be hoped that the international standardization efforts will proceed here so that solutions can be developed which will last for some time and enhance world-wide compatibility of document formats.

A third project cluster with the poetic title „*Blue Rose*“ (acronym for „*Building Libraries Unifying Enhanced Retrieval-Oriented user SErvices*“) is also related to problems of digital libraries' infrastructure, but to its technical side. These sub-projects resulted from one of the projects which have emerged from focal point 4; only this one has been voted positively by the GIC. Here, the research component of the proposed projects is most pronounced. This is the reason why the ministry decided to support sub-projects which have been approved principally, not directly out of GLOBAL INFO, but from funds of a neighbouring department in the ministry which directly supports software development, or to support them in GLOBAL INFO only when they cooperate under project leadership of a publishing company. Under these auspices the following sub-projects have been approved:

- Alerting and Profiling Services: Definition of a system independent protocol for and development and implementation of these services together with some publishing companies, support of different systems of profile description.
- Retrieval in distributed databases with facilities dealing with heterogeneity and vagueness, definition of client and server protocols with XML and RDF, merging of retrieval results.
- Federation Services and Document Servers: Recognition and administration of identical objects, building and mediating local data pools, usage of XML for document recognition.
- Development of a broker/trader being able to deal with the specification of various protocols, contains modules with the basic services in this field, and contains a universal type specification mechanism for these services.
- A fifth sub-project (extending technically and functionally the earlier mentioned dblp-system in computer science at Trier university) was positively voted upon, but the ministry for principal reasons declined to fund this fee-less service of Trier university. It shall be examined, however, if and how far other information systems in the area of computer science could be integrated; this should be supported.

In this project cluster some principal problems of contemporary computer science are thematical. The GLOBAL INFO program has limits how far it can support and fund basic research activities. The dominant view was that the program should limit its scope to more user-oriented areas. There are other funding possibilities where basic innovations have a more appropriate context.

The fourth project cluster has the acronym *WEP* for its German title „*Werkzeuge für Elektronisches Publizieren*“ (tools for electronic publishing). It is one of the results of the working group on focal point 1. These projects are located along the production chain of documents from the author to the printer (resp. electronic publisher) and long-term archiving. Here, for reasons easy to understand, many publishing houses are involved. This broad range of items, however, contributes to a certain heterogeneity of this project cluster, too. The following eicht sub-projects which differ in volume and scope and, therefore, were grouped into some projects groups have been approved:

- Text oriented authoring tools independent of software systems as a working environment for articles and books on the basis of SGML and XML.
- Authoring tools on the basis of system independent document description languages (XML and its derivates), rules, tools for conversion and formatting including development of document type specifications for different documents along XML rules and of converters to and from LaTex and HTML..
- Rules for authoring founded on media and instruction psychological considerations, implementation as prototype by producing two multimedial books.
- Standards, authoring rules and layout formula on the basis of thesauri and metadata definitions in pedagogics and sociology for print, mixed and digital media.
- Software environment to enable informal teamwork in the production of teaching and learning media, i.e. a groupware solution for the creation of various media products.
- Multimedia Authoring and publishing tools, implementation in prototype production of a multimedia book on databases.
- A conversion workbench between the formats of print and multimedial media.

A project proposal on individualized print modules and tools to provide these individualized prints has been approved in principal but was postponed in order to find publishers to take part actively in the development.

It can be seen from the number, heterogeneity, but also the similarity of some of the sub-projects that this field is very near to product development on the one hand, but that on the other hand an evident basis is not yet existent. Clear lines and structures have not yet developed. Closer cooperation of groups of sub-projects has still to be found and realized.

Two more projects proposals – one from focal point 1 on the creation of multimedial documents in CML including the recourse to experimental data in chemistry, the other one from focal point 2 on new developments electronic in learning and teaching tools and documents – are in the pipeline. Others will follow in the course of this year.

So, after all, GLOBAL INFO is just about to start into the first period of substantial work on development projects. A considerable number of publishers, several libraries and scientific information centers and many universities and disciplines are involved in the projects starting now. The broad range of activities, of contents and themes is providing evidence that a large field of unsolved questions and problems still exists. As a whole, one can be confident that GLOBAL INFO will deliver a substantial contribution towards their solution.

5. Conclusions and Perspectives

So after all it seems that GLOBAL INFO is now about to accomplish the transition into a nation-wide networking system of interdependent projects with intensive international contacts and cooperation towards distributed and dynamically developing digital library systems. One can realistically hope that within a few years the chances of the single scientist to use the advantages of large integrated scientific information systems will be amended dramatically. Nevertheless, there are some crucial questions the answer to which is not yet visible. It seems to me that they have to be voiced in order to look at least for the directions of answers.

(a)   For the German digital library scene to me it seems to be a crucial step to seek and to find public attention, organizational ways and fora to overcome traditional barriers between the various parties and state agencies taking part in DL activities. As mentioned above the situation in Germany is especially complicated because of the federal structure of funding of universities and of parts of research. A more or less accentuated competition between DFG, BMBF and DFN-Verein, between the federal government's level and the states, and between traditional library organization, scientific information centers, universities and the WWW are difficult to overcome. An articulate digital library information culture, supported by a server information basis and organized around an electronic (and additionaly printed) journal on digital libraries – as it exists in the USA and in Great Britain – could provide a new focus to deal with an decide the relationship between the different initiatives. A „round table" on modern scientific information systems could be a focus for such a clearing mechanism. It would have to be, I guess, to be initiated and supported by the top political level of government. For the national Government, but also on state level the idea seems to be worthwile to be deliberated seriously whether the national information and communication structure in sciences is not one of the crucial conditions of contemporary and future development, competitiveness of the national economy and economic growth, and if it should not be a central area of state attention, support and guidance.

(b)   One may hope that a new „division of labour" between the traditional players in this field – the scientists themselves, their universities, the libraries, the database providers, the publishing companies and the book sellers – will be achieved; one should not underestimate this economic and political dimension as one of the central problems and tasks in the construction of complex scientific information systems. The solution of organizational questions and problems is equally important – it might be even more urgent – as the technological development in this sphere. But all participants in this process should continu-

ously have in mind that they can only succeed together: The *scientists* will not be able (and should not be interested) to compete the publishing houses out of the market; and the other way around the great *publishers* will have to avoid to dry out their customers' demand potential. The *database providers* will have to find a rapid and smooth way to integrate their offer into an increasingly integrated – Web-mediated – information landscape; proprietary retrieval forms and languages will have to be substituted by visualized, intuitive forms of research to be handled without special retrieval knowledge. The *libraries* and the *librarians* will have to accomodate with their changing role as local or regional information broking institutions and information and knowledge mediators instead of mere book storage institutions and ordering and catalogueing persons; from a self-understanding to be primarily archiving units they will have to move to that of specialists for the provision and brokering of information. The same is – mutatis mutandis – true for the *booksellers* and *bookstores* in the market sphere. But it should not be underestimated that the way to such a new division of labour is paved with ground mines, made difficult by differing interest and narrowed by existing power relations. Plenty of good will and – as we say in German – jumping beyond one's own shade will be requested to find a „king's way" here.

(c)   There is yet another respect to which questions of content seem to be rather more complicated than purely technical ones: With the integration of scientific information systems the problem of identifying properly the adequate scientific culture, of developing intelligent description and retrieval systems which will guide me reliably into these areas of science which I am really looking for, will gain momentum. The relation of intellectual input and structuring on the one hand, the possible automation of processes of information provision and retrieval on the other hand are far from being solved. So as another big task we will have to be careful to build a scientific information reality which will allow for a balance that in every discipline of science teachers, researchers and students are helped, not restricted by or overflood with huge piles of senseless information by the systems we develop. What a German information scientist some ten years ago termed the ‚explosion of nonsense' has to be avoided. So questions of conceptual analysis, of the different meanings of terms, i.e. of the special knowledge frameworks and concepts in the differing scientific cultures will have to play their due role in the development of integrated scientific information systems.

(d)   Scientific information and communication in the various disciplines are rapidly gaining global range. In the natural sciences this is true since several years, and digital library systems here are more or less international on a world-wide scale. In engineering sciences and in humanities and social sciences this global extension is less developed. To a certain extent this is due to the differences in

central subjects of these branches of science: Societies, the institutional dimensions of economies, education systems and even political structures are still moulded by differing national conditions. But even here, the globalization of economics, politics, social trends, and cultures tends to level traditional differences. Therefore, efforts to extend the dimensions of scientific information systems beyond national and lingual boundaries seem to be crucial. Bi- or even multilingual lexica and thesauri, the ability of search engines to retrieve subjects and knowledge fields in more than one language, or the adherence of various national digital library initiatives to common international standards are important and, yet, not sufficiently realized dimensions of internationalization. In this respect, I expressively welcome the initiative of the American NSF to support the international cooperation of U.S. digital library projects with similar activities abroad. Although I shamefully must admit that to my knowledge this offer has not yet found a cooperating counterpart in Germany I am convinced that in the years before us this dimension of cooperation will rapidly gain momentum.

With all these questions, problems and perspectives in mind GLOBAL INFO – as the other digital library initiatives in various national contexts – will have to proceed in a tentative and pragmatic way. To find a due balance between systematic theorizing and conceptualization on the one hand and development of practically relevant and rapidly usable solutions and results on the other hand, is one of the most interesting, but also one of the most difficult and challenging tasks in science. The same is, obviously, true for the central aim of GLOBAL INFO and the world-wide development of digital libraries: To create access to world-wide and high-quality electronic information and communication integrating all kinds of relevant sources for every scientist from his or her workdesk.

6. References

[1] Cf. the recent conference of the German learned societies (The Initiative Information and Communication IuK) on „Dynamic Documents" at Jena in March 1999 at http://elfikom.physik. uni-oldenburg.de/IuK/IuK99.

[2] Juergen Bunzel: The Digital Library Program of the Deutsche Forschungsgemeinschaft, in: ABI-Technik. Zeitschrift fuer Automation, Bau und Technik im Archiv-, Bibliotheks- und Informationswesen, Muenchen: Verlag Neuer Merkur, vol. 18, 1998, No. 2, pp. 132 – 137.

[3] Cf. the collections of links and project descriptions at the DBI server at Berlin [http://www.dbi-berlin/projekte/d_lib/d_lib00. htm] and at the GMD IPSI at Darmstadt [http://delite.darmstadt. gmd.de/delite/UsefulLinks/]

[4] Rudi Schmiede: GLOBAL INFO – The German Digital Library Project. Development and Perspectives, in: ABI-Technik. Zeitschrift fuer Automation, Bau und Technik im Archiv-, Bibliotheks- und Informationswesen, Muenchen: Verlag Neuer Mer-kur, vol. 18, 1998, No. 2, pp. 147 – 153, also available at http://www.global-info.org/doc/980407-schmiede.html

[5] Erich J. Neuhold/Reginald Ferber: Scientific Digital Libraries in Germany: Global-Info, a Federal Government Initiative, in: Christos Nikolaou; Constantine Stephanidis (eds.): Research and Advanced Technology for Digital Libraries. Second European Conference on Digital Libraries (ECDL '98), Proceedings, Berlin; Heidelberg; New York: Springer (Lecture Notes in Computer Science; vol. 1513), pp. 29 - 39

[6] Additional information about the two projects can be obtained at http://medoc.informatik.tu-muenchen.de/ and at http://interdoc. offis.uni-oldenburg.de/.