SSOAR

Open Access Repository
www.ssoar.info

Advancing digital libraries in Germany and creating distributed scientific resources - the telos working group at Darmstadt University of Technology

Schmiede, Rudi

Veröffentlichungsversion / Published Version
Zeitschriftenartikel / journal article

Empfohlene Zitierung / Suggested Citation:
Schmiede, R. (2004). Advancing digital libraries in Germany and creating distributed scientific resources - the telos working group at Darmstadt University of Technology. *International Journal on Digital Libraries*, *4*(1), 48-51. https://doi.org/10.1007/s00799-003-0056-6

Nutzungsbedingungen:
Dieser Text wird unter einer CC BY Lizenz (Namensnennung) zur Verfügung gestellt. Nähere Auskünfte zu den CC-Lizenzen finden Sie hier:
https://creativecommons.org/licenses/by/4.0/deed.de

Terms of use:
This document is made available under a CC BY Licence (Attribution). For more Information see:
https://creativecommons.org/licenses/by/4.0

gesis Leibniz-Institut für Sozialwissenschaften

Mitglied der Leibniz-Gemeinschaft

Diese Version ist zitierbar unter / This version is citable under:
https://nbn-resolving.org/urn:nbn:de:0168-ssoar-255790

Int J Digit Libr (2004) 4: 48–51 / Digital Object Identifier (DOI) 10.1007/s00799-003-0056-6

Advancing digital libraries in Germany and creating distributed scientific resources – the *telos* working group at Darmstadt University of Technology

Rudi Schmiede

Darmstadt University of Technology, Fachbereich 2, Residenzschloss, 64283 Darmstadt, Germany
e-mail: schmiede@ifs.tu-darmstadt.de

Published online: 27 July 2004 – © Springer-Verlag 2004

Abstract. This text gives an overview of the work of the author and his working group ***telos*** over the past 20 years and sketches the main focus of this work, i.e., creating technologies for distributed scientific resources and their intelligent retrieval.

Keywords: Digital libraries – Distributed resources – XML – Ontologies – Germany

The ***telos*** working group,[1] which is headed by the author of this article, is located in the Department of Sociology at the Darmstadt University of Technology. It might be surprising to readers of this journal to find an editor and a working group doing research and development in digital libraries in the field of sociology. This is easily explained: The developing access to worldwide scientific information is going to change not only the content and conditions of work in the field of scientific study, but also the general way in which society produces and reproduces itself and people. The transition to modern information, communication, and knowledge technologies is in no way confined to academia but is increasingly shaping the structures of industry and administration.

Thus, coming from industrial sociology with its analysis of work and of its social, economic, and political conditions and motivated by an interest in a critical theory of advanced capitalism, the author began at the end of the 1970s to engage in research on the impact of the new information technologies on work and society and in the implementation of this technology for practical research purposes. The first practical result was, in the pre-PC era, the digital documentation of relevant German literature and sources on the historical development of working time and its conditions as part of an empirical study on this subject (realized by Ulrike Stein, who then extended the documentation subject to another study on internal labor markets in German industrial history). And this research interest was one of the reaons to move, after 15 years at an institute of social research, to a university of technology where, although sociology is one of the less-favored scientific disciplines, cooperation with engineering and information scientists is possible in everyday contact.

In Darmstadt, one of the first tasks we set before ourselves in the early 1990s was to develop the Darmstadt Virtual Union Catalogue (Darmstaedter Virtueller Gesamtkatalog) [1, 2]. This union catalog was inspired by the basic conviction that various information sources should be unified and made accessible on a metalevel that abstracts standardized metadata. Practically speaking, the project team created interfaces for all library administration systems available in the marketplace at the time, so that all bibliographic data sets could be harvested in one database; thus, the data of the central library of Darmstadt University, of some 80 decentralized department and institute libraries, and in addition of neighboring institutions like Darmstadt Polytechnic University and Darmstadt Public Library (and, for some time, several leading public libraries in the German state of Hessen) could be made available in a day-to-day resp. night-to-night actualized version. We learned the basic technological concept at the time from our successful project to transpose the object-oriented architecture of the Tinlib system (of the British IME company that cooperated

[1] The acronym *telos* stands for: **t**echnology for **e**lectronic **l**ibraries and the **o**rganization of the **S**emantic Web; in Greek, it expresses the orientation toward an aim, a telos. In contrast, the more theoretical activities of the research unit (work, technology, and society) are organized in the working group *kairos* (**k**ritical **a**nalysis of the **i**nfor**m**atization **o**f **s**ociety). The research and development work is complemented by the third working group *transfer*, putting forward the practically oriented consequences and proposals of research activities.

then with Unisys; the author of this software, Peter Noerr, has been working for several years with SUN and has written a notable introduction to the technology of digital libraries) to the German library world; unfortunately, Unisys and IME could not agree on a common business model. The basic idea we learned from that work was to conceive of information systems as virtual systems comprising various data sets (categories) put together on the fly. In a world of diversity of resources, only such a concept of virtuality, i.e., the integration of modules on an intellectually defined metalevel, can meet the challenges of reality.

It is worth noting that, in the author's opinion, the fact that the above-mentioned and still operative union catalog was realized by sociologists, and not by information scientists or librarians, is not accidental: we approached this field of technological development from the point of view of users, albeit those reflecting the development of their scientific discipline, and not from that of systems analysts or developers. This experience engendered the author's conviction that an operative program system cannot be designed and implemented solely by information specialists but has to be realized by real users in cooperation with the information systems specialists. The consequences for IT systems development are obvious: only the rare species of informationally self-educated scientific users or, which will be the model of the future, cooperative teams of discipline-oriented scientists and information specialists will be able to develop and further the scientific information systems of the future. New combinations of knowledge and qualification in information systems and scientific disciplines will have to be developed and consolidated to realize prospective and accepted scientific information systems.[2]

This theoretical, practical, and political background caused the author to get, in the autumn of 1996, the German Association of Sociology (Deutsche Gesellschaft für Soziologie, DGS) to join the Initiative of Learned Societies on Information and Communication Technologies (IuK-Initiative der deutschen wissenschaftlichen Fachgesellschaften), founded in 1995 originally by the German Association of Mathematicians, the German Phycical Society, the Association of German Chemists, and the German Society of Informatics; the initiative now comprises some dozen learned societies, including several in the humanities, representing over 120,000 scientific staff [3]. Besides establishing an information resource on German sociology [4], which has in the meantime been moved to the administrative board of DGS, the author was selected to chair a federal funding program that was an offspring of joint activities between the IuK-Initiative and an innovative concept of public advancement of information technologies in the sciences of the Federal Ministry of Education and Research; the program had the programmatic name *Global Info* which was coined being aware that the long-term task was to make available worldwide scientific information "at the fingertips" of the researching, teaching, or learning scientist [5]. The special and unique profile of Global Info was to bring together producers and users (i.e., scientists, as represented by the IuK-Initiative), publishers (represented by large international and national publishing companies), representatives of established distributing institutions (libraries and central scientific database providers), and federal authorities; a representative of the German Research Association, the most important funding organization of scientic research in Germany, was present at the Global Info Consortium, the governing board of the program, as well [6–8].

In the program some 20 million deutschmarks (resp. 10 million euros) were spent on funding digital libraries in Germany. Several tools and modules for publishing, handling metainformation, interoperability, retrieval, and rights management were developed. The program had an important impact on mobilizing and enhancing the German digital library movement; for some years it brought together in projects, working groups, and conferences most of the researchers and developers active in this sphere. However, its weak side was that it was too thinly spread and too heterogeneous; thus after some time project groups dispersed into their special fields of work. The large publishers followed their own international strategies and largely withdrew from the common activities. The desired effect of learning from each other and from the most advanced groups subsequently diminished. In addition, the funding ministry decided to have an in-between phase of basic strategic thinking where the development of scientific information should go in the next 5–10 years and for that purpose commissioned several strategic studies that were carried out in 2001 and 2002.

One of these studies is called "digital library concepts". The project, together with the IuK-Initiative mentioned above, organized a forum on infrastructures for digital libraries and in this context working groups on digital library services, long-term preservation, and local access and information management; in addition, it subcontracted two expert studies on subject portals that were seen – together with local portals – as the most important unifying access possibilities for users. We advanced the importance of standardization in general and Web standards in particular in Germany and tried to get all funders to accept the insight that standardization work should be part and parcel of each development project. In the last 2 years the advancement of Semantic Web activities and of the new technology of Web services has been added. These activities led to the installation

[2] Thus the *telos* working group comprises at present, resp. in the recent past, people whose orginal academic qualification was philosophy, sociology, mathematics, chemistry, biology, information science, and computer science. One of the current activities of the *transfer* working group is to realize a new integrated qualification profile of practical and academic information competence at the academic level with the cooperation of industry, universities, trade unions, and the state.

of a new German "network of competence for new services, standardization, and metadata", which *telos* has the task of coordinating and advancing further. At the present time (summer 2003) a national project on long-term preservation and projects developing a distributed journals server, persistent identifier services, and generic portal modules are at work on this network; more are to follow.

On the level of concrete systems development, the *telos* working group conducted an extensive evaluation – in cooperation with physicists at the university of Essen – of a digital library publishing system originally developed in the mid-1990s for computer science (Medoc) and updated in the late 1990s to become a publishing and administration software application for small publishers (Interdoc); the question being addressed was whether the system could and should be used by other scientific disciplines like physics and sociology; the answer was rather critical [9]. The consequence of the basically negative evaluation was to develop, in cooperation with the German Information Center for Social Sciences (Informationszentrum Sozialwissenschaften) in Bonn, a distributed "virtual library of sociology" [10, 11]. In Germany, as a result of the destructive consequences of the Second World War, disciplinary collections of scientific literature are bundled in a decentralized fashion all over the country; thus the official library collection task for sociology lies with the library of the University of Cologne, whereas other large sociological collections can be found at research institutions in Bonn and Berlin. The project developed a distributed system aggregating bibliographic titles from these sources and from Darmstadt using the technology developed for the Darmstadt Virtual Union Catalogue and improving it substantially. The Z39.50-based system queries the distributed and heterogeneous information sources and ranks and merges the obtained results. As a prerequisite for merging, the different encoding formats (MARC, MAB, proprietary database formats) are transformed into a common, internal representation based on RDF/XML. Algorithms to deal with semantic heterogeneity (problems like UdSSR or USSR or Soviet Union or Russia for the same object) have been developed and tested. The results of this project are one of the building blocks of a future integrated social sciences portal in Germany.

Another building block will be access and availability of relevant Web resources in sociology. Developing the basic technology for these purposes and institutionalizing the relevant services and cooperation is the task of the ongoing project "SozioNet" [12], which aims at making available Web-based information on persons, institutions, projects, subjects, and publications in German sociology [13]. This project follows the basic direction of MathNet and PhysNet in mathematics and physics and has some similarities with the British Sosig subject gateway, but it tries to improve technology to hit the current state-of-the-art internationally. Concretely, SozioNet implements common metadata standards for all resources made available by participating institutions. Metadata schemes are modelled using RDF. While all common metadata elements are taken from qualified Dublin Core, domain-specific aspects are described by a backing ontology based on OWL. For example, the German-English thesaurus and the classification of social sciences (by the Social Science Information Center) are integrated into the ontology, so metadata records can directly reference the corresponding definition of a subject in the thesaurus. OWL clarifies many aspects that have been intentionally left open by RDF schema and offers additional language elements, for example, defining the cardinality of property relations. Also, many software tools today already offer support for OWL, thus providing a higher level access to the model.

The MetaWizard tool was designed to allow for user-friendly metadata markup. It offers distinguishing features like personalization, automatic extraction of existing metadata, and comfortable browsing of the existing social sciences thesaurus and classification for easy use. A native open-source XML database [14] is used in conjunction with a specially developed SozioNet (completely XML-based) harvester. The distributed resources are gathered and indexed by the harvester by crawling the Web locations known to the system. In addition to Web crawling, the OAI protocol is used to obtain metadata sets from some larger institutions that have already implemented their own metadata creation workflow.

As another service to users, a standardized secondary homepage has been designed that summarizes important institutional information on each of the participating organizations. Institutes at universities and research institutions are cooperating with the project now to build up relevant data resources. (Other German projects will make available the printed journal literature in digital form to complete the first stage of the future social science portal. These projects are united in the *infoconnex* cluster [15], of which SozioNet is also a member.)

To summarize this work for a special social science portal, at the beginning of 2003 the working group *telos* started a project called *Generic and Component-Based Scientific Portals – A Framework for the Building of User-Oriented Services* [16]. This project aims at creating modules for the standardized exchange of structured information on the basis of Web services standards. Information resources that are specified as services and retrievable via registries can be integrated in manifold ways and offer new possibilities of usage. The framework is meant to facilitate the building of portals (that can use each other as a service, too) and the construction of scientific networks on the basis of an international standard that is gaining increasingly in usage and support. We plan to realize this framework beyond sociology in cooperation with mathematicians and chemists in a paradig-

matic way. One of the tasks of this project is the coordination of the national network of competence mentioned above.

To conclude this short description of the activities of *telos* in Darmstadt, it should be mentioned that Rudi Schmiede served on the advisory board of the European DELOS network of excellence, and Stephan Koernig was coordinating one of the DELOS working groups. This NoE is now slated to obtain funding for the years to come within the new European funding program. This work on the European level and on the level of global cooperation – and last but not least via the Journal on Digital Libraries – will improve the work of *telos*, and, it is hoped, the working group will be able to make one contribution or another to the desired international advances in digital libraries.

References

1. Schmiede R, Wenzel H (1996) Darmstädter Bibliotheken im World Wide Web. Ein Projektbericht (zusammen mit H. Wenzel). Bibliotheksdienst 30(1):75–90
2. Schmiede R, Wenzel H (1997) Der Darmstädter Virtuelle Gesamtkatalog. Ausbau und Erweiterungen 1996. Bibliotheksdienst 31(1):88–97
3. http://www.iuk-initiative.org
4. http://www.global-info.org
5. http://www.soziologie.de
6. Schmiede R (1998) GLOBAL INFO. The German Digital Library Project – development and perspectives. In: ABI-Technik, Jg. 18(2):147–153. Also at: http://www.global-info.org/doc/980407-schmiede.html
7. Schmiede R (1999) Digital library activities in Germany. The German Digital Library Program GLOBAL INFO. In: Proceedings of the IEEE forum on research and technology advances in digital libraries, Baltimore, MD, pp 73–83. http://www.global-info.org/doc/990413-schmiede.pdf
8. Koch W, Koernig S, Schmiede R (2000) Informationen über Informationen – Perspektiven einer besseren Versorgung mit Fachinformation in Global Info. ABI-Technik 20(1):62–63
9. Oberhage R, Drewitz A, Winkler S (1999) InterDoc Vorabevaluationsbericht (2000): http://interdoc.theo-physik.uni-essen.de/vorab/
10. Meier W, Müller MNO, Winkler S (2000) Die virtuelle Fachbibliothek Sozialwissenschaften. Problembereich und Konzeption. Bibliotheksdienst 34(7/8):1236–1244
11. Meier W, Müller MNO, Winkler S (2000) Die virtuelle Fachbibliothek Sozialwissenschaften. In: Knorz G, Kuhlen R (eds) Informationskompetenz – Basiskompetenz in der Informationsgesellschaft. In: Proceedings des 7. Internationalen Symposiums für Informationswissenschaft (ISI 2000) Konstanz, Universitätsverl. (Schriften zur Informationswissenschaft) 38:179–194, http://www.vibsoz.de
12. http://www.sozionet.org
13. Meier W, Schumann N, Heise S, Schmiede R (2003) SozioNet: networking social science resources. In: Koch T, Solvberg IT (eds) Research and advanced technology for digital libraries, Proceedings of the 7th European conference (ECDL 2003), Trondheim, Norway. Lecture notes in computer science, vol 2769. Springer, Berlin Heidelberg New York
14. Meier W (2003) eXist Native XML database. In: Chaudri A, Rashid A, Zicari R (eds) XML data management. Reading, MA, 2003, http://www.exist-db.org
15. http://www.infoconnex.de
16. Schmiede R, Körnig S (2003) Infrastrukturen für innovative Digital Library-Dienste. Perspektiven des Kompetenznetzwerks "Neue Dienste, Standardisierung, Metadaten". In: Schwänzl R (Hg) Sharing knowledge: scientific communications. Beiträge zur Jahrestagung der IuK-Initiative der wissenschaftlichen Fachgesellschaften 10.–13. März 2003 in Osnabrück (in press)