

The eDoc-Server Project

– building an institutional repository for the Max Planck Society

Gerhard Beier, Theresa Velden

Abstract

With the eDoc-Server the Heinz Nixdorf Center for Information Management in the Max Planck Society (ZIM) provides the research institutes of the Max Planck Society (MPS) with a platform to disseminate, store, and manage their scientific output. Moreover, eDoc serves as a tool to facilitate and promote open access to scientific information and primary sources. Since its introduction in October 2002 eDoc has gained high visibility within the MPS. It has been backed by strong institutional commitment to open access as documented in the ‘Berlin Declaration on Open Access to the Data of the Sciences and Humanities’, which was initiated by the MPS and found large support by major research organizations in Europe.

This paper will outline the concept as well as the current status of the eDoc-Server, providing an example for the development and introduction of an institutional repository in a multi-disciplinary research organization.

Introduction

The eDoc-Server project is part of the strategy of a multi-disciplinary research organization to embrace the Internet as a powerful medium which will revolutionize scientific and scholarly communication and to increase access to all information resources relevant for its research. In this context the MPS has identified open access as a vital pre-condition. Therefore the Max Planck Society has initiated the “Berlin Declaration” which was signed on 22nd October 2003 at the end of a 3-day conference on “Open Access to Knowledge in the Sciences and Humanities” by all major German research organizations, as well as major French research organizations and further institutions also form the cultural heritage domain¹. This remarkable step of formal adoption of the principle of open access by major research organizations was the outcome of intensive internal discussions within the Max Planck Society (MPS) and with other key-players in Germany. The conference is seen only as starting point for the adoption and implementation of open access policies by research organizations and funding agencies in Europe. Further institutions have been invited to sign the ‘Berlin Declaration’ and a follow-up conference is planned in order to take stock of concrete measures taken and to coordinate political efforts to achieve open access.

The MPS strategy is to pursue two roads towards open access. Firstly, the MPS will encourage its scientists and scholars to publish in open access journals and will take these publications into account in tenure and evaluation. In this context the president of the MPS, Peter Gruss, has underlined his intention to strengthen the principle that the evaluation of research is to be based

on the intrinsic value and quality of a publication, not the reputation and formal impact factor of the journal it is published in. Secondly, the MPS will encourage its scholars and scientists to use the full potential of the internet for the dissemination of research results by self-archiving their work on institutional or disciplinary servers.

The Max Planck Society electronic Document (eDoc) Server represents a crucial element in this strategy. It has been set up as a prototype system in October 2002 to enable scientists and scholars in the MPS to archive, disseminate, manage, publish and share information and research results. Driven by the fast adoption of the Server by the Society during the last twelve months its functionalities have been extended step-by-step to fulfill the needs of a very heterogeneous user community.

Institutional Background and Aims

The Berlin Declaration demands open access both to primary research literature as well as to digital representations of objects of cultural heritage. The latter is a central part of the agenda of the EC funded project European Cultural Heritage Online (ECHO)² which brings together 17 institutions, among them three Max Planck Institutes. Both these aims are also part of the concepts of the eDoc-Server project. On the one hand, it aims to make the research results from the Max Planck Society worldwide openly accessible. On the other hand, it aims to support the management and storage of digital collections to make cultural heritage and primary sources publicly accessible. This latter usage scenario is currently explored working with a few selected digital collections from the domain of the history of art.

The Max Planck Society is currently organized in 80 institutes all dedicated to basic research in their field. The institutes are independent and highly autonomous both in their research and internal organization. They are organized in three sections, a Chemistry-Physics-Technology section, a Biology-Medicine Section and a Humanities Section.

Therefore, the concepts applied to eDoc have to take into account the different traditions, aims and habits of scientific and scholarly communication established in the various disciplines. The aim of the eDoc project is to provide a central service and software for all these different communities within the MPS. Bearing in mind that a central service for such a diverse environment might not meet all the needs of the institutes, the focus was placed equally on the development of a central system and services as on provision of interfaces and configuration possibilities to allow customization of the system to local and discipline specific needs.

The MPS pursues four strategic goals with the eDoc-Server:

1. **Distributed, global knowledge space:** The MPS wants to position itself in the emerging worldwide network of scholarly communication and provide open access to its publications. In this aspect the MPS joins efforts undertaken also by other large institutions like CERN, MIT, Lund University³ etc.
2. **Innovative Models of Scholarly Communication:** Moreover, eDoc should serve as a backbone for initiatives by scientists and scholars in the MPS who want to make use of the full potential of the internet and establish new models for scholarly communication.

3. **Institutional Memory and Showcase:** Furthermore, eDoc shall serve as a comprehensive resource of scientific information produced by the Max Planck Institutes, providing a stable location for its preservation and dissemination. Thus, it functions as an institutional memory and allows Institutes to create showcases to demonstrate research productivity and excellence, supporting alternative approaches to research evaluation as well as immediate access to research results.
4. **Open Access:** eDoc should strengthen the Society and the scientific community in negotiations with publishers about ownership of and access to scientific research documents at a time where sky-rocketing journal prices and restrictive publishers' policies undermine their wide dissemination and persistent accessibility. eDoc is introduced and will be further promoted as the self-archiving tool for publications.

It follows that the eDoc-Server cannot be seen primarily as a software project, but rather as a project which is developing and introducing a new software and at the same time is promoting the idea for open access and working on a paradigm shift regarding the dissemination and evaluation of scientific information.

In addition to these strategic aims, a very pragmatic need arose during the first phase of the project, which accelerated the introduction of the eDoc-Server to virtually all Institutes of the MPS. The trigger was the modernization of the production process of the annual report and other formal reports and handbooks of the MPS, which was moved to a fully digital workflow, requiring the Institutes no longer to provide their annual publication lists in written documents, but in a structured format. Here the eDoc-Server offered itself as common web-based database system, which would transfer the bibliographic data intended for the formal reports to the content management system supporting the work flow for the production of the official web pages as well as formal reports of the MPS. This decision led to the fast adoption and integration of the eDoc-Server into the official reporting channel between Institutes and Society headquarters and made it possible that the server is used by literally all institutes for the management of their bibliographical records for reports. It required to focus on features supporting specifically the handling of bibliographic data and production of publication lists to allow easy re-use of data once entered into the system. On the other hand, the wide adoption of the system provides now an excellent ground work for not only introducing the software, but also the idea of open access to scientific information to the institutes.

Main Concepts

From the beginning the peculiarities of the MPS had to be taken into account for the development of an appropriate strategy for building an institutional repository. As the institutes are acting very autonomously only the infrastructure for a document repository can be provided centrally, while collection and user management as well as quality control have to be done locally. The strong autonomy of the institutes was also the reason why the Heinz Nixdorf Center for Information Management (ZIM) which is in charge of developing the software and establishing this server as a central service in the MPS, involved the institutes from the very beginning in shaping the project and formulating requirements. Hence a small group of "pilot Institutes" acted as consultants and early adopters of the eDoc-Server. One of these institutes, the Fritz-Haber-Institute in Berlin, also contributed the initial development of the eDoc-Server

software⁴.

In order to gain acceptance within these different communities – ranging from plasma physics to history of art and from astronomy to zoology – the ZIM tries to accommodate as best as possible all these different needs, never losing sight of a standardization on a central level. This can be illustrated by the variety of document types (genre types) which can be entered on eDoc. They range from widely accepted types such as article or book to more community specific types like expert opinion, book review etc. (all in the same system which is configured and administered locally).

Right from the beginning the aim was to provide open access to scientific results – if not restricted by copyright agreements with third parties – and to closely integrate the eDoc-Server with other systems, such as OAI-Service Providers or discipline-specific archives to make it attractive to deposit material on eDoc and to get maximal impact for the research output of the MPS.

While the strategic focus is on providing open access to all MPS research output, strong emphasis is put also on facilitating comprehensive access to information internally, for members of the MPS. Hence, whenever an open access version of a work registered on eDoc is not available, a fine-grained hierarchy of access levels on eDoc allows to provide a version of the document at least internally. For the same reason eDoc is integrated with the MPS Virtual Library, which is a central portal for MPS researchers for their information searches, and which provides dynamic links to articles published in electronic journals.

Usage Scenarios

The early introduction of the eDoc prototype within the MPS has led to an extensive discussion among the institutes about the possible uses of the eDoc system. From responses and discussions within and outside the MPS the ZIM identified a number of desired usage scenarios for an electronic document server and was able to support one usage scenario – eDoc Archival – extensively and to experiment with a second usage scenario – eDoc Primary Sources – in order to evaluate requirements and the feasibility of supporting them by a central infrastructure. The two scenarios are characterized as follows:

eDoc Archival

serves to document, disseminate and archive results of scientific research of the Max Planck Institutes and to make them openly accessible. Ideally, not only bibliographic metadata are stored here, but also the full text is archived or a reference is given to a server where open access is guaranteed. The eDoc-Server provides a stable location for all sorts of publications including works not formally published such as presentations, talks, posters, PhD thesis, interactive resources etc. The metadata of all records on eDoc are made publicly available and access to the full text is restricted only if strong reasons (e.g. copyright regulations) prohibit its public dissemination. eDoc Archival can also serve as a showcase of the scientific productivity of the MPS and its individual institutes.

The eDoc Archival scenario is currently widely applied by all institutes of the Max Planck Society, in particular because of its function to collect the data for the official Max Planck Annual Report. Apart from that usage which is mainly focused on bibliographical metadata a

number of institutes are archiving and disseminating full texts via eDoc. They submit posters, talks at events, conference papers, PhD thesis, articles, inBooks etc. and most of them are publicly available. The most common file format is PDF, but also Microsoft Powerpoint or RTF, LaTeX etc. had been stored, especially if the focus was on sharing the uploaded document with some colleagues.

➤ **eDoc Primary Sources**

aims to make primary sources openly accessible and to care for their long term preservation. It supports ongoing research projects in the Institutes and captures collections of primary source material such as images, scans of texts, datasets etc. which are referenced in publications or constitute the basis for further research. Objects of primary sources collections generally have research-specific metadata sets which are not covered by the general metadata model for genre types defined for eDoc Archival. Objects are likely to consist of several files which are jointly up- and downloaded with the tools developed by the specific research community.

The implementation of the eDoc Primary Source scenario is still in an early, explorative phase. Work so far has concentrated on high-resolution scans. First achievements are the development of interfaces for the automated up- and download of complex objects and the prototypical integration of an image collection (Photothek) of the Kunsthistorische Institut in Florence. Subsequently, 10,000 digitized images of cultural heritage will be uploaded onto eDoc where they are archived and made accessible to a wider audience. The discipline-specific metadata for the objects are administered in a local database and mapped to the general model in order to provide basic views on the data. Moreover, a special image viewer, Digilib⁵, is integrated on eDoc. DigiLib is an open source tool, which is collaboratively developed by a number of Institutes to satisfy research-specific needs. It supports zooming, changing of contrast, measurements, and annotations which can be shared among colleagues. The integration of Digilib enables scholars to really use the eDoc environment as a kind of workbench for their research.

Structure of the System

Collections & User Management

All documents and material on eDoc are organized in collections. Documents will be submitted to collections and administered there. Responsibilities and user roles are defined on collection level. For every collection at least a moderator and an authority have to be assigned, who take charge of the quality control process (see below in section workflow). Typically the collections will correspond to organizational units (e.g. departments of an Institute or research schools), and the head of this unit defines the policy for the material accepted, e.g. articles + posters, talks, un-refereed articles and what the rules for the quality assessment shall be for the collection in question.

On institute level the local eDoc Manager is responsible for the local customization of the system and administration of users, i.e. he can register users and assign specific rights to them. The local eDoc Manager can choose from the following pre-defined roles:

- Depositor: may submit documents to a collection on eDoc

- Collection Moderator: checks the formal quality of the submission
- Collection Authority: checks the scientific quality of the paper and ensures that it follows the policy standards for this collection
- Metadata Editor: may correct and add metadata
- Publication Editor: can select and release items for annual report
- Local eDoc Manager: can add users, assign rights, customize

Workflow (eDoc Archival)

A strong emphasis was put on the workflow for making a document publicly available. It is based on the assumption that the self-archiving of documents by individual researchers is desired and supported. A typical usage scenario will look as follows:

The scientist enters the document and all the necessary metadata and uploads the full text. He/she can recommend the appropriate access level for the full text of the resource, i.e. whether it should be accessible only internally, to a group of privileged users, for the whole Max-Planck-Society or worldwide⁶. After the submission the document will undergo a quality control procedure performed by two additional people. The moderator, who is in charge of organizational aspects of the collection, will check the metadata and corrects them if necessary. The authority guarantees the scientific quality and decides whether this document is accepted for this collection. This scientific quality control process is no peer-review process in the “traditional” sense known from journals, but guarantees that the document follows the rules and policies set up for this particular collection. With the authorization of the document the authority acknowledges that it is worthy to be displayed as part of the scientific output of the institute. If a document has already passed a “traditional” peer-review this can be further indicated in the document’s metadata.

After moderator and authority have approved and accepted the document the metadata will be publicly released, and the full text made available according to the specified access level as outlined above. The released metadata record will get an additional tag containing the name of the authority which has authorized the document. This is displayed as “communicated by <name of authority>” as part of the records’ metadata.

The scholars and scientists of the MPG strongly demanded such a quality review process is implemented, even though it will not replace or question the traditional peer-review system. It will guarantee a minimum quality standard for openly accessible scientific information on the web and thus every outside user can trust that the material found on eDoc follows the quality standards of the issuing Max Planck Institute.⁷

Version Control

The versioning system of eDoc is guided by the ambition to offer reliable and transparent information on eDoc. Once a document is submitted by an individual scientist or scholar it

cannot be altered anymore. Any modification of metadata or upload of full text by the depositor of the document will create a new version which is indicated with a number after the eDoc ID (e.g. 2356.2 indicates version 2 of the document with ID 2356). This guarantees that changes to documents can be traced and the appropriate copy cited by referencing the full document ID.

Genre Types and File Formats

For eDoc Archival a metadata model was developed which allows to capture all kinds of publications used in the scientific and scholarly communication process. It is based on international standards such as vCard, DC, OpenURL, AMF, LOM, Ariadne, ODRL, OAI, CLD⁸ and is extensible for community specific requirements and covers genre types such as articles, inBooks, books, conference papers, proceedings, poster, talks etc. Currently, 17 different genre types are supported and more will be added on request. The concept of genre types is based on the separation of the intellectual concept and the medium of publication. An example might illustrate this: An article will always be submitted as an article irrespectively whether it is published in a traditional print journal, on a CD-ROM or in an online-journal.⁹ A similar logic applies to the status an article might have in an external publication process. This status can be expressed by adding additional metadata like “submitted”, “accepted / in press” or “published”. This avoids the creation of a preprint genre where no consensus among all disciplines can be reached how the term is used. The creation of new genres is handled centrally by the ZIM to ensure a common understanding of the terms applied and to guarantee consistency throughout the system. As long as discipline-specific genre types do not overlap with general types they will be created as well, e.g. expert opinion is a valid genre type for law institutes, but not in use in the science section and does not cause any misinterpretations in this field either. Moreover, the institutes can decide on local level which genre types they want to offer on the web form for document submissions, this way simplifying the process.

The main aim of the eDoc server is to capture all types of documents and resources which are considered part of the scientific output of the MPS. For this reason the eDoc server does not place any restriction on the file formats which can be uploaded to the system. Disciplines should use the formats which are most common in their fields – for physicists this might be LaTeX or dvi, for historians DOC, RTF or PDF. Nevertheless, users are advised to provide at least one version of a document which can be easily read in a web environment, independently of the platform used (we currently recommend PDF) and to give wherever possible open non-proprietary file formats such as RTF, TIFF, JPG etc. At the moment eDoc can only ensure that files can be accessed and downloaded, but does not provide any additional functionality to read these files (except Digilib for images) as it relies on the availability of browser plug-ins and appropriate programs in the respective communities. Files which require a special program have to be downloaded first and then opened with the local program.

Preservation

So far the eDoc-Server can only ensure the “bit preservation” meaning that we do our best that data can be retrieved in the form in which they were created by using high quality storage systems and servers, backup systems and a mirror at a different location.

It would be highly desirable to provide also “functional preservation”, that is not only to allow the retrieval of the file, but to support proper interpretation of it. This would mean that either an

emulation or a migration would have to be carried through in the long run. Obviously giving a guarantee for 'functional preservation' is currently beyond the reach of any institution housing large, heterogeneous digital collections and costs implied are barely understood.

User Interfaces

The whole system can be administered via web-based interfaces. Documents are entered via a step-by-step or single form submission interface which guides depositors through the process and offers extensive documentation of metadata fields in help texts. Also the upload of files with structured metadata e.g. in XML or Endnote format can be initiated from a web page. The different user roles have different interfaces and visibility of administrative metadata depends on the level of rights in the system. There are three types of searches available, quick search, full text search and an advanced search which allows a combination of various different search terms to exploit the full potential of the underlying metadata model. The main entry point for getting an overview on records deposited on eDoc is by browsing through institutes and collections or by browsing through the alphabetic lists of authors.

Furthermore, some user interfaces were added to support the workflow of gathering, selecting and releasing bibliographic data intended for the annual report of the MPS. In this process also conformance of the metadata to the guidelines for the annual report is checked and a view for the editor of the report provided. User interfaces are also available to manage baskets (equivalent to shopping carts) where users can place items and reuse the baskets afterwards to reference a particular set of documents on the eDoc-Server or to export the selected record sets. These baskets can also be used to save search result sets and print them formatted in the citation style provided by eDoc or to download them e.g. in PDF.

A major strength of the system is the web-based administration of the metadata and the configuration of the display of metadata on the submission interfaces as well as in the browse views. With the so-called metadata modeler the ZIM can easily create new genres and assign genre-specific layouts drawing from the overall eDoc metadata set.

(Figure metadata_modeller.gif to be included)

Figure 1: View on element "publisher" in genre type "PhD thesis"

Technology Platform

After evaluating existing open source e-document/digital library software early in 2002, taking into account in particular CDS, Greenstone, EPrints.org, DSpace and MyCore¹⁰, the ZIM decided to rely in the first phase of the project on an adaptation of an in-house development, which was provided by one of the pilot Institutes, the Fritz-Haber-Institute.¹¹ This decision was based on the assumption that 1) the system should be available for loading with first pilot collections within three months, 2) the multidisciplinary environment and acceptance at the institutes would require a complex metadata model and workflows and interfaces which none of the available systems provided – hence system intervention would be required, 3) after this first explorative phase an advanced 2nd generation system would be introduced in due time to replace the earlier system.

This 1st generation of the eDoc-Server is based on an Apache web server, a PostgreSQL database and is programmed in Embedded Perl. The DigiLib image viewer is programmed in Java and

requires a TomCat server. The whole system runs on a Linux server, hosted at a central computer center (GWDG) of the MPS where it is regularly monitored and backed-up. Currently, the software is not available as an open source package as it was not – in contrast to EPrints or DSpace – within the scope of the project at this stage to create a distributable, open source document server. Instead the main focus of the project is on the adoption of the eDoc-Server within the MPS and the advocacy of open access in order to get a maximum of content into the system.

Interfaces to other Systems

For the architecture of a central system in a distributed environment interfaces to external systems are absolutely vital. This means in particular an integration of eDoc into systems used in the MPS and integration into the global scientific information space.

As the management of publication lists is a very important usage scenario in the MPS eDoc is providing various importing and exporting mechanisms. Widely used bibliographical management systems such as Endnote, Reference Manager, BibTeX are supported. Moreover, to facilitate the administration and creation of publication lists the eDoc-Server provides interfaces to upload search results from the Web of Science according to the agreements with ISI. Furthermore, the generation of publication lists (via an XML query; XSLT; HTML) for integration into departmental or personal homepages which can be delivered in the customized personal layout is planned. Already, records can be exported in a general citation style to HTML, RTF or PDF by using the basket functionality. Bibliographic records selected for the annual report of the MPS are also integrated via a customized OAI-Interface into the Content Management System of the MPS and will so be included on the general MPS website.

In order to achieve global visibility for records and full texts on eDoc it is planned to extend the existing OAI-interface to provide a public access to data where no restrictions apply. Through an OAI-MHP interface the eDoc-Server will serve as a data provider for the larger scientific community and OAI enabled services. Besides the required Dublin core metadata set also an extended metadata set, based on the eDoc XML exchange metadata scheme can be exposed via this interface.

By Open-URL based integration with the MPS-SFX-Service, eDoc offers context-sensitive linking to external resources such as the Web of Science or journal web sites. This guarantees access for MPS scientists and scholars to the full text where the MPS holds licenses, in cases where the article is not stored on eDoc itself. The MPS-SFX-Service also provides interesting functionalities to non-MPS users, who in many cases are at least led to the abstract of the article in the journal.

A close integration with the institutes is especially necessary if the eDoc server is used as a storage backend and/or as a means for making data publicly available. For the flexible batch upload and download of objects from primary source collections the system provides an open interface which can be used by java-clients e.g. as developed by the MPI for the History of Science. Moreover, it is planned to use the OAI-protocol also for the integration of databases maintained locally by institutes using the eDoc-XML-Schema for data exchange¹².

Status and Challenges for the Future

The eDoc system was developed as a rapid prototype to be able to introduce an eDocument server for the MPS within a very short time frame in order to explore its uses with pilot institutes on a live-system. It was then extended and revised during the last twelve months to satisfy immediate needs and to take advantage of the opportunity to introduce it to all Institutes of the MPS as part of the reporting procedure for the annual report of the MPS.

To date the eDoc-Server has been adopted by the entire organization due to its capabilities to handle bibliographic data and support the workflow needed for the production of the publication lists of the official annual report for the Society. At present (early November 2003) the system holds about 15,000 records, 15% of them including a full text. The eDoc-Server is further backed by a growing awareness and support on the management level of the Society for the strategic importance of open access and the determination to build a comprehensive institutional repository. This currently evolving institutional policy in favor of open access and self-archiving is expected to boost the use of the eDoc-Server as open access platform for the entire research output of the MPS. The ZIM has reacted to that by providing guidelines for scholars and scientists on how to deal with contracts with publishers and exclusive licenses or restrictive copyright regulations.

As indicated above the speed of development and introduction of the eDoc Server came at the cost of system scalability and extensibility. As eDoc is being developed in close contact with the needs of the institutes it has triggered a lot of new ideas and requirements. To address these needs the ZIM plans to move to a new system (working title 'eDoc2') which will be more flexible and is intended to support more usage scenarios, workflows, metadata sets as currently possible. The concepts of this new system will build on experiences gained with eDoc1 as well as on experiences made by other institutional or disciplinary archives.

Conclusion

Within a year, with the eDoc-Server an institutional repository could be introduced to the 80 research Institutes of the Max Planck Society, which are geographically distributed and cover a wide range of disciplines in sciences and humanities. In close interaction with a group of pilot Institutes the discipline specific needs, both for managing and disseminating publications as well as digital primary source collections were explored. The development of the system was feature driven, that is focused on satisfying immediate and integration in existing work flows to facilitate acceptance of the system, rather than the creation of a scalable, distributable software.

Given that many institutional repositories struggle to get content into their systems the decision to focus on the introduction of the system rather than on provision of software seems justified and the close integration into the MPS will pave the way for further usage. The eDoc system is heavily used as a bibliographic management software and as institutes are getting more and more acquainted with the system and the idea of open access, the amount of free content available in eDoc is growing day by day.

This shift from pure management of bibliographic data towards an open access platform will further be supported and encouraged by the Heinz Nixdorf Center for Information Management

in the Max Planck Society. eDoc will be promoted as the self-archiving tool for scientists and scholars in the MPS. The focus of further activities will be to move to a more flexible and scalable system, to increase the integration with global and local systems, to facilitate alternative approaches to research evaluation in an open access environment, and to realize services that facilitate open access also to primary source collections and supplementary material.

Reader Response

If you have any comments on this article, please contact the editors

Author Details

Gerhard Beier

Heinz-Nixdorf-Center for Information Management in the Max-Planck-Society (ZIM)

E-Mail: g.beier@zim.mpg.de

Gerhard Beier is working as project manager of the eDoc-Project in the ZIM.

Theresa Velden

Heinz Nixdorf Center for Information Management in the Max-Planck-Society (ZIM)

E-Mail: velden@zim.mpg.de

Theresa Velden is working as Executive Director for the ZIM.

Heinz Nixdorf Center for Information Management in the Max Planck Society:

<http://www.zim.mpg.de>

Tel: +49 89 3299 1551

Address: Heinz Nixdorf Zentrum für Informationsmanagement in der Max-Planck-Gesellschaft,
Boltzmannstraße 2, c/o IPP, ITER-Gebäude
85478 Garching, Germany

Notes

¹ See for further details the website of the associated conference on ‘Open access to the Data of the Sciences and Humanities’, held from 20th-22nd October in Berlin: <http://www.zim.mpg.de/openaccess-berlin/>

² For further information this project see the ECHO website at <http://echo.mpiwg-berlin.mpg.de/>

³ All of these institutions are running institutional repositories; CERN: <http://cds.cern.ch>; MIT: DSpace <http://dspace.mit.edu>; Lund University <http://lu-research.lub.lu.se/>

⁴ The initial development was undertaken by Mike Wesemann and Heinz Junkes from PP&B department, Fritz-Haber-Institut. A variant of the initial code is available as ‘eDoc Advanced’ from their website: <http://w3.rz-berlin.mpg.de/eda/>

⁵ The Digilib source code is available from <http://developer.berlios.de/projects/digilib/>

⁶ Restricting access to a full text to the level of an institute will be made possible in the near future.

⁷ Also Jean-Claude Guédon pointed to the relevance of quality control in open access environments: Guédon, Jean-Claude (2002) Open Access Archives: from Scientific Plutocracy to the Republic of Science. In *Proceedings 68th IFLA General Conference and Council: Libraries for Life: Democracy, Diversity, Delivery*, Glasgow, Scotland. [deposited: 20 November 2002] <http://papyrus.bib.umontreal.ca/archive/00000065/>

⁸ vCard <http://www.w3.org/2001/vcard-rdf/3.0#>; DC <http://purl.org/dc/elements/1.1/>; OpenURL <http://www.sfxit.com/openurl/openurl.html>; AMF <http://amf.openlib.org/>; LOM <http://ltsc.ieee.org/wg12/>; Ariadne <http://ariadne.unil.ch/Metadata/>; ODRL <http://www.w3.org/TR/odrl/>; OAI <http://www.openarchives.org/OAI/2.0>; CLD <http://www.ukoln.ac.uk/metadata/cld/>

⁹ See also Functional Requirements for Bibliographical Records, Final Report, IFLA Study Group on the Functional Requirements for Bibliographical Records, K.G.Saur, München 1998, p.7-16.

¹⁰ At the time neither Dspace, MyCore or CDS were publicly released, and only a pre-version of CDS was obtained for evaluation. Intensive discussions and visits were conducted to meet developers or user communities of all these projects (with the exception of Greenstone, which by its characteristics was excluded early on from further evaluation).

¹¹ The software which is currently running on <http://edoc.mpg.de> is not available as open source due to a number of pragmatic reasons which inhibit this.

¹² The edoc-transfer XML schema is documented at http://edoc.mpg.de/doc/schema/zim_transfer.xsd