

Deutsche Initiative für  
Netzwerkinformation (DINI)



## **DINI-Certificate Document and Publication Repositories**

Working Group "Electronic Publishing"

October 2003

## Contents

1	Aims	4
2	Criteria	5
2.1	Policy	5
2.2	Author Support	5
2.3	Legal Aspects	5
2.4	Authenticity and Data Integrity	6
2.5	Indexing	6
2.5.1	Subject indexing	6
2.5.2	Export of Metadata	6
2.5.3	Interfaces	7
2.6	Logs and Statistics	7
2.7	Long-term Availability	7
3	Award and Evaluation	7
4	Explanations and Commentary	8
4.1	Policy	8
4.2	Author Support	10
4.3	Legal Aspects	10
4.4	Authenticity and Data Integrity	11
4.4.1	The Repository	12
4.4.2	The Documents	13
4.5	Indexing	14
4.6	Logs and Statistics	14
4.7	Long-Term Availability	15

## Summary

Installing and institutionalizing document and publication repositories at universities creates the opportunity to offer and archive scholarly publications that were produced at the respective universities for a worldwide audience. This new service by the infrastructure institutions library and computing center helps disseminate the idea of electronic publishing as a new tool for academic work.

DINI – the Deutsche Initiative für Netzwerkinformation (German Initiative for Networked Information) – supports this development to reach a higher level of scientific and scholarly communication nationwide and internationally. To reach this goal it is necessary to network document and publication repositories. A network of local publications would complement the dominant (commercial) publications through publishing houses. This network may therefore also function as a regulating body against the monopolistic tendencies in scholarly publishing.

The German Science Council and the Conference of University Rectors in Germany call for the installation of document and publication repositories, and funding institutions such as the Ministry for Science and Technology (Bundesministerium für Forschung und Bildung) and the German Research Foundation (Deutsche Forschungsgemeinschaft) support (and fund) this.

It is deemed important to develop this according to international standards and to use proven technology.

Taking the DINI publication “Electronic Publishing in Higher Education” (March 2002), that was published as a first guideline, as a starting point, this publications aims at three targets:

- give a detailed description of the requirements of a certified document and publication repository,
- highlight directions of development in the areas of configuration of repositories and of information exchange through these repositories,
- offer a certificate visible to both users and operators to certify the adherence to defined standards and recommendations by this repository.

By issuing this certificate, DINI for the first time introduces a quality control for document and publication repositories.

A set of minimum requirements of a repository and its operator mandatory for modern scholarly communication is formulated. At the same time DINI makes recommendations highlighting foreseeable developments that might turn into future requirements.

A working group within DINI audits the criteria for the DINI certificate against international standards and developments and updates them accordingly. The certificate is for this reason issued with a year-of-award stamp.

## 1 Aims

Utilizing the world wide web (www), science has created a new tool for communication, that is increasingly used commercially. Despite intensive use of the www by academics in their everyday work, the options to improve scholarly communication through the www are by no means exploited to their full extent.

Deficits in this regard are most prominent in the dissemination of high quality scholarly publications. Acceptance of electronic media has risen enormously, however, willingness to actively use the technology as a tool for publication (to publish electronically) is to a great extent limited to commercial publishers, especially in the STM areas. It is in these areas of science, that the quasi monopoly of the publishers has led to exorbitant prices, more and more often exceeding the abilities of libraries and universities to pay the demanded charges.

Building local publication repositories may work as a regulating factor, and should the great majority of scholarly publications be published originally on non-commercial repositories, a communication network will have been created, that will at least make it more difficult to continue unjustified raises in prices for the sole purpose of profit maximization.

Offering publications on local repositories allows for the improvement of publication services in science areas of less commercial interest. The publication of doctoral theses on the internet is a good example for this process. Major prerequisites for success were:

- A nationwide development of metadata standards,
- publication on local repositories through cooperation of libraries and computing centers
- a workflow for the upload to the German National Library (Die Deutsche Bibliothek), who in turn assumed responsibility for long-term archiving of the theses.

The German Science Council and the Conference of University Rectors in Germany call for the installation of document and publication repositories. The German Ministry for Science and Technology (Bundesministerium für Forschung und Bildung) and the German Research Foundation (Deutsche Forschungsgemeinschaft) support implementation of document and publication repositories at the German universities. National scientific institutions, such as the Max-Planck-Society, are preparing their extensive use. On an international level, organizations like SPARC are supporting this development. The MIT and other universities in the USA and the United Kingdom are offering open-source software to run institutional repositories. In the Netherlands, the DARE project is preparing introduction of services.

Local publication systems are common in Germany. To bring them to greater success it is necessary to standardize further developments. DINI with its publication "Electronic Publishing in Higher Education" laid a foundation for a widespread introduction of general regulations and standards. "DINI-Certificate – Document and Publication Repository" takes this a step further by clearly describing criteria, that will guarantee repositories to be set up and operated according to national standards and international developments. Repositories fulfilling these criteria may be awarded a certificate, testifying to their quality.

The certified repositories will be links in a chain of content providers, who offer their metadata to specialized service providers in a standardized format and procedure in accordance with the OAI model. The service providers in turn may aggregate these metadata and among other possible functions serve as a connector to commercial content and service providers.

If the worldwide movement to install and operate local (institutional) repositories could be based on the solid foundations of universities' infrastructure institutions like libraries and computing centers and their respective international networks, a new basis may be created for a communication infrastructure in the sciences. Additionally, money invested in information infrastructure would be used much more efficiently without much extra effort.

Certification of repositories to state their international compatibility within an arising communication network of the sciences for high-quality publications does not have to be

restricted to university repositories. It is desirable that commercial publishers integrate their services in a standardized scholarly communication.

## **2 Criteria**

The following criteria are split into two sections. In the first section *minimum standards and requirements* are specified, that must be met by the document and publication repositories or their operators to be awarded the certificate.

As DINI plans a continuous adaptation of the certificate to retain its validity in an ever changing environment, the second section lists *recommendations*, that, from today's point of view, are likely to be future requirements for the certificate.

### **2.1 Policy**

The operator of the document or publication repository formulates a policy regarding content and operation of the repository and makes this policy publicly available.

This policy must state rights and obligations of operators of a repository and of authors and/or editors.

The policy must contain:

- Provision of standards for the publications in the repository with regard to content and functional and technical quality
- A guarantee to archive for defined time spans depending on content and functional and technical quality of the publication
- Procedures for the operation of the repository; definition of services that the operator of the document repository offers to authors and editors

### **2.2 Author Support**

Minimum standards

- Offer consultancy services and support via web pages and/or e-mail and/or telephone as well as person-to-person support
- Support of the entire publication process (including technical and legal areas)

Recommendations

- Curriculum of courses on electronic publishing offered at least once per semester
- Specialized courses on "structured writing" for authors
- Utilization of (electronic) help-desk systems

### **2.3 Legal Aspects**

Minimum standards

- The operator of the repository must be permitted to publish the uploaded document onto the repository (right to store electronically, especially in databases, and make available to the public for individual retrieval, rendition on a monitor and print (online use) in part and/or as a whole)
- The operator of the repository must be permitted to forward the document to an archiving institution
- The operator of the repository must be permitted to alter the documents technically to secure long-term availability
- Exemption from liability must be formulated in a disclaimer

Recommendation

- List rights and privileges in metadata

## 2.4 Authenticity and Data Integrity

### Minimum standards for repositories

- Documentation of the technical system (It is expected of the repository to be available on a 24/7 basis)
- Back-up system to secure repository, metadata and documents
- Technically controlled and verifiable acceptance of documents

### Recommendation for repositories

- SSL certification

### Minimum standards for documents

- Persistent identifier must be assigned
- A document with altered content must be treated as a new document (new persistent identifier)
- Archiving of authors' uploaded files in their original format

### Recommendations for documents

- Application or procedure to account for the integrity of documents (e. g. a hash number)
- Advanced digital signature (§ 2 Abs. 2 SigG 2001<sup>1</sup>)
- Rendering of archiving file formats with reference to the export of documents into long-term archiving facilities or institutions (cf. 2.7 and 4.7)

## 2.5 Indexing

### 2.5.1 Subject indexing

#### Minimum standards

- Availability of a defined policy for indexing
- Verbal indexing with keywords or classificatory indexing

#### Recommendations

- Dewey Decimal Classification (DDC) according to usage in the German National Bibliography as the general classificatory indexing system for documents (v. DINI-OAI recommendations)
- At least one additional standardized system of verbal or classificatory indexing (general or subject specific, e. g. Schlagwortnormdatei, LoC, CCS, MSC, PACS etc.)
- Keywords in English
- Abstracts in German and English

### 2.5.2 Export of Metadata

#### Minimum standards

- Metadata are available for free
- Unqualified Dublin Core (ANSI/NISO Z39.85–2001<sup>2</sup>)

#### Recommendations

- Qualified Dublin Core
- ONIX<sup>3</sup>
- Technical and/or archival metadata; print-on-demand data (e. g. METS<sup>4</sup>)
- Export to long-term archiving institutions or repositories
- Export into bibliographical databases (e. g. bibliographical management system, networks, OPACs)

---

<sup>1</sup> SigG –Signaturgesetz – German Signature Law

<sup>2</sup> [http://www.niso.org/standards/standard\\_detail.cfm?std\\_id=725](http://www.niso.org/standards/standard_detail.cfm?std_id=725)

<sup>3</sup> <http://www.editeur.org/onix.html>

<sup>4</sup> <http://www.loc.gov/standards/mets/>

### 2.5.3 Interfaces

Minimum standards

- User interface to the web repository
- OAI PMH 2.0<sup>5</sup>

Recommendations

- Expansion of the OAI protocol to allow for exchange of complex metadata schemata
- Web-service interface (e. g. SOAP<sup>6</sup>)
- Z 39.50<sup>7</sup>

### 2.6 Logs and Statistics

Minimum standards

- Every individual repository must (within legal boundaries) log statistical data on access to the repository and to individual publications

Recommendation

- The log containing data on access to a publication should be added to the publication as dynamic metadata

### 2.7 Long-term Availability

Minimum standards

- Persistent linking of metadata and documents (e. g. via a persistent identifier or storage of metadata and document in one single container)
- Definition within the policy (v. 2.1) of a minimum time span of availability of a document of no less than five years)

Recommendation

- Securing long-term availability where necessary through cooperation with a archiving institution

## 3 Award and Evaluation

The DINI office or an authorized working group is responsible for awarding the DINI Certificate – Document and Publication Repository. The seal of the certificate carries the year when it is awarded. The certification document acknowledges that the certified repository meets the minimum standards of a DINI-certified Document and Publication Repository.

A fee is charged for the issuing of the DINI certificate.

#### Non-profit organizations

DINI members	50.00 €
Other	100.00 €

#### Profit organizations

DINI members	150.00 €
Other	250.00 €



<sup>5</sup> <http://www.openarchives.org/OAI/openarchivesprotocol.html>

<sup>6</sup> <http://www.w3.org/TR/SOAP/>

<sup>7</sup> [http://www.niso.org/standards/resources/Z3950\\_Resources.html](http://www.niso.org/standards/resources/Z3950_Resources.html)

The operator of a document or publication repository applies at DINI for certification using a web form on the DINI website. This form contains the minimum standards as listed in the criteria (v. 2). The applicant fills out the form therewith stating to what extent the criteria for certification are met by the repository to be certified. Further information about the repository may be added in designated fields in the form (text, URL et al.).

An authorized DINI working-group verifies the submitted data. Access to the repository must be permitted to this working group. The operator must be prepared to answer questions by the working group. A visit at the actual location of the repository will be the exception. Additional costs that emerge during the certification process must be paid by the operator of the repository. The operator of the repository will be informed about possible additional costs beforehand. The certification process should generally be completed within two months.

There is no time limit on the certificate. As the certificate shows the year it was awarded, re-verification of criteria can be done at any time.

The operator of a certified repository is entitled to call the certified repository "DINI-certified document and publication repository" and to show the DINI certificate's seal on a web page or in other applicable forms.

Any misuse of the seal or certificate will be prosecuted according to applicable laws.

## **4 Explanations and Commentary**

A local document repository is to ensure storage, management, provision, archiving, search and retrieval of electronic resources of an institution.

### **4.1 Policy**

As an example for a repository policy see below the Policy for the Operation of the Document and Publication Repository of the:

Humboldt University, Berlin.

#### **1 Objectives and Criteria for Content of the Document and Publication Server of Humboldt University**

The Document and Publication Server offers the organisational and technical framework to all members of Humboldt University for publishing scientific documents digitally. In course of this joint services offered by the Computer and Media Services and the University Library scientific documents of high importance are being published on the internet under strict quality control.

The digital documents are provided with persistent identifiers and addresses, and are indexed within national and international library catalogues, search engines, and other reference tools. The document and publication server provides protection against distortion by using digital signatures and digital time stamps. Furthermore, long-time preservation of digital publications is guaranteed.

Operation and development of the document and publication server are integrated into national and international initiatives and projects like the »Networked Digital Library of Theses and Dissertations« (NDLTD) or the »Open Archives Initiative« (OAI).

#### **2 The Collection Mandate of Humboldt University Library for Digital Documents**

The collection mandate of Humboldt University Library consists of collecting, cataloguing, and archiving all the scientific documents published by the members of Humboldt University. It refers to digitally born documents as well as digital versions of printed documents.

Also included are significant historic documents from the University Library and other institutions that are digitised due to terms of content, conservatory aspects, or the requirements of place-independent use.

Intellectual property rights will be preserved. Publishing the document with the publication and document server does not prevent publication elsewhere, like scientific journals or other document repositories. According to the recommendations of the German Science Council (Wissenschaftsrat) all scientists of Humboldt University are asked to secure further rights of use when negotiating with publishers. Later, maybe after a certain qualifying period, they will be asked to also publish their documents with the Humboldt University Document Server.

Observation of copyrights and rights of use for third parties are solely with the authors or the editors of the digital documents.

### **3 Digital Documents**

The term »digital document« as used here is defined as a document based on text and images, that is stored in digital form on a data medium, and that is distributed via computer networks. In the future this term will be extended toward multimedia documents that include audio and video sequences.

A document to be published with the document and publication server should meet the following requirements:

1. It is to be distributed for the public.
2. It is not a dynamic document. With any changes in the document a new version will be stored.
3. It fulfils the technical parameters given by the Computer and Media Services and the University Library of Humboldt University.

### **4 Digital Documents to be published and distributed by the Document and Publication Server**

1. The following categories of digital documents will be stored and distributed:
  - Single publications and publication series of Humboldt University with scientific content, as the public lectures of Humboldt University or the publication series of the University Library,
  - Single publications and publication series published by staff of Humboldt University, i.e. collections, conference papers, research reports, journals (e-journals),
  - Documents that are mandatory to be published according to examination rules and regulations (theses and dissertations),
  - Single publications and publication series of institutions and persons associated with Humboldt University,
  - Documents of Humboldt University students like master theses and seminar papers, should those be recommended by a lecturer.
2. Distribution of these documents can be restricted to local or temporary use only.

### **5 Technical Characteristics of the Document and Publication Server**

1. By providing and proving qualified digital signatures the digital documents receive a legally valid certificate of authenticity. The assignment of these qualified digital signature occur in full accordance with the German Digital Signature Law (Signaturgesetz - SigG).
2. The digital documents are provided with individual and persistent addresses. This enables direct access to the document.
3. Retrieving the digital documents is possible via library catalogues, queries of the bibliographical metadata, search terms within the structures of the digital documents, via alphabetical and indexing systems, and via dynamically produced lists and indexes.
4. For indexing, storage and archiving of the digital documents international standards like the guidelines of the Open Archives Initiative (OAI) are employed and developed.
5. When using SGML/XML an archiving time span of 50 years is guaranteed. Preserving other formats depends on the availability of these formats, their viewer software, and their conversion tools.

## **6 Organisational Regulations**

1. The document and publication server of Humboldt University is a joint operation of the Computer and Media Services and the University Library.
2. The electronic publication is free of charge for Humboldt University staff and members of associated institutions.
3. Submitting the digital documents for distribution via the Humboldt University document server takes place in the University Library.
4. Additional services that are necessary for publication, like processing digital documents or conversion into other formats, are conducted by staff of the Computer and Media Services or the University Library. Services that need significant additional work will be charged according to the scale of charges of the Computer and Media Services and the University Library.
5. Contact for all questions about the Document and Publication Server is the joint »Electronic Publishing Group« of the Computer and Media Services and the University Library

### **4.2 Author Support**

One of the goals of electronic publishing at universities is to register and archive all scholarly publications that were created at the universities and/or by scholars and scientists employed by the universities. Standards for storage and archiving with regard to file formats, digital signatures, metadata etc. are desirable and necessary for long-term archiving and possible later reuse of the electronic publications

The path to reaching this goal begins with advising and supporting authors and editors. The use of templates in word-processing applications is one basis for structural writing in the digital age.

It is therefore recommended to initiate course programs for electronic publishing at the universities. To raise the level of attraction to and acceptance of electronic publishing, strategic advantages such as support services et al. must be developed and offered to authors and editors. DINI encourages the use of Document Type Definitions (DTD) according to international standards for publications, or the support of templates and the enforcement of style sheets for word processors and/or TeX/LaTeX. This is the current practice at Humboldt University.

### **4.3 Legal Aspects**

#### **Policy and Contract**

To be awarded the DINI certificate, a written policy for the operation of the repository must exist. This policy must be accepted by authors and editors as well as the university. This policy must state rights and obligations of operators of a repository and of authors and/or editors. The operator of the repository must enter contractual relations with authors according to the policy.

The policy must contain items, where authors grant the following non-exclusive rights to the operator of the repository:

- The right to electronically store the publication, especially in databases, to make the publication available to the public for individual retrieval, rendition on a monitor and/or print (online use) in part and/or as a whole
- The right to forward the document to an archiving institution
- The right to technically convert the files for the purpose of long-term archiving

These rights are granted for the unaltered publication.

Granting the aforementioned non-exclusive rights to publish, to forward the publication, and to convert for archiving purposes serves not only to publish a document, but also to secure long-term availability of electronic documents on document and publication repositories.

Operators of a document and publication repository must explicitly grant the following

privilege to authors/editors:

- The author/editor may make the document available on personal repositories or other institutions' repositories in part and/or as a whole (right to self-archive)

This privilege is granted for the unaltered publication.

### **Rights of Third Parties**

- Authors/editors must assure the operator of a repository that no (copy)rights of third parties will be breached by publishing a document or parts thereof (e. g. pictures)
- Authors/editors must assure the operator of a repository of immediate notification in cases of doubt or in cases of alleged or actual legal hindrances to the publication of a document or parts thereof
- Questions of liability in cases of damages and/or legal prosecution must be clarified and regulated between authors/editors and the operator of a repository in a written contract

### **Print Publications**

The following items become relevant should universities or university presses offer the extra service of printing a publication.

Authors/editors must grant the university/university press the exclusive right to publish and distribute the work in print as a book. The option of printing an edition in a print-on-demand procedure must be stated in the contract.

### **Exploitation Rights**

In case of an additional print publication authors/editors must agree to not publish, transmit or reproduce the work in tangible or intangible form over a given period of time (abstention duty, duplicate registration with VG Wort).

In case of a publication with an ISBN assigned, a contract must be signed by authors/editors and universities. The contract must contain paragraphs regulating (copy)rights and exploitation rights. The importance of a legally binding contract increases, if a university press acts on behalf of the university.

### **Exemplary Contract**

The German version of this document contains in this place the text of a contract as signed by the Oldenburg University with authors publishing through the Oldenburg University Press.

As this contract is applicable only within the German legal system, it is not translated into English for this document. However, some of the important issues are listed below; they concur with the recommendations above.:

- Granting of non-exclusive rights to publisher to publish electronically and in print
- Granting of right to author to publish on personal server or institutional repositories
- Liabilities of authors and publisher (violation of third parties' rights)
- Exploitation rights in case of additional print publication

## **4.4 Authenticity and Data Integrity**

Storage of publications can be based on a content management system (featuring data replication, data backup) that organizes storage, administration and archiving of electronic media throughout their entire life cycles.

Access to the stored digital resources must be ensured for a long-term period. This is done through the use of unambiguous identification systems (persistent identifiers). Persistent identifiers (PI) can additionally be used to support harvesting. PIs can be used as workflow-control mechanisms, where unambiguous identification of objects is required, as in e. g. decentralized processes, checking for doubles, authentication mechanisms, alarm systems, or for multimedia or modular objects. Objects whose modules are stored

in different repositories can be merged effectively through the use of PIs. By selecting clearly identified modules, different versions can be generated without having to store each of these versions. Digital objects can be cited through their PIs in digital environments as well as in non-digital environments.<sup>8</sup>

#### 4.4.1 The Repository

##### Documentation

In the documentation for the technical system, the administrator(s) must be named and the system itself must be described.

Below see an excerpt from the documentation for the repository of the document management system of the university library Dortmund.<sup>9</sup>

##### Server:

- Data-Server dx1

##### Administrator:

- Michael Heyer, Tel. 2423, room 237

##### Services:

- The data-server dx1 serves as platform for ca. thirty web servers receiving roughly four million requests per month, and as FTP server handling ca. 100,000 FTP calls per month. Dx1 is used as a back-up server (TSM) for 350 clients on campus; an additional archive server (TSM) is installed. Additionally dx1 serves as mail server (POP, IMAP, SMTP) and as information system for the Cambridge Database (Chemistry). The DBMS Oracle is available. The application ArcView (classroom administration) runs on this server. Lastly, dx1 serves as spooler for a number of printers and plotters. Operating system is Solaris 2.7.

##### Hardware:

- UltraSPARC Enterprise 4000 Server
- 6\* UltraSPARC, 167 Mhz processors
- 3 GB main storage; external storages SPARC Storage Array 112 (63 GB), SPARC Storage Array 114 (126 GB), Sun StorEdge D1000 (218 GB)

##### Software:

- OS Solaris 2.7
- System management: Solstice Admin Tools
- Solstice DiskSuite
- File system NFS
- Perl
- Conversion Tools
- [etc.]

Contact: Michael Heyer, phone 2423

Last update: 2003-07-09

The controlled adding of documents to the repository ensures that only documents adhering to the defined standards and criteria of the repository will be published on the repository. Developing and implementing appropriate workflows is the operators' responsibility. Workflows must allow for technical control of the upload of documents.

##### SSL Certification

SSL Certification is a prerequisite to guarantee authenticity of the documents for the readers. Certification must be issued by an authorized CA. A digital signature ensures integrity of the documents.

---

<sup>8</sup> <http://www.persistent-identifier.de>

<sup>9</sup> <http://www.hrz.uni-dortmund.de/S2/Server/dataserver.html>

## 4.4.2 The Documents

### Persistent Identifier

A persistent identifier must be assigned to every document on the repository. A possible solution is the assignment of URNs.

To guarantee the integrity of a document for the reader, a hash must exist for every document in the repository. Changing demands on security foster technical developments in cryptography; algorithms for hash numbers have to be adapted according to these new developments. Current recommendations include the Message Digest Algorithm MD5 or the Secure Hash Algorithm SHA-1. Application of cryptography algorithms should be publicized for the readers' information.

The following is an excerpt of the University of Chemnitz' document repository's information page.<sup>10</sup>

#### The MONARCH Security System

The MONARCH system applies various hash algorithms and a digital signature to the documents published on the system.

#### Basics

- A hash file exists for every publication in MONARCH
- Three different hash numbers exist for each file of a publication; these numbers are listed in the hash file of the publication
- The hash file is digitally signed by the archivist

Using the hash numbers, it can be checked at any time whether the publication has been altered. The digital signature of the hash file itself allows for checking the hash numbers for changes or alterations since the creation of the hash file.

#### Applied algorithms:

- Secure Hash Algorithm (SHA)
- Message Digest Algorithm MD5
- RIPE-MD algorithm

#### Digital signature of the hash file

- The archivist digitally signs the hash file
- For the digital signature PGP 2.6.3 is used

Every alteration of a document creates a new document. A persistent identifier is assigned to this new document. Versioning allows for identification of previous versions of a document. Metadata work as links between generations of a document.

### Digital Signature

The German Digital Signature Law<sup>11</sup> differentiates various levels of signatures:

#### *Electronic Signature*

Electronic data added or logically linked to other electronic data used for authentication

#### *Advanced Electronic Signature*

Electronic signature as above, that

- is assigned exclusively to the holder of the signature key,
- allows identification of signature-key holder,
- is created through processes under the exclusive control of the signature-key holder, and
- is linked with the data it is referring to in such a manner as to allow identification of any alteration of said data after application of the advanced electronic signature.

<sup>10</sup> <http://archiv.tu-chemnitz.de/integrity.html>

<sup>11</sup> [http://bundesrecht.juris.de/bundesrecht/sigg\\_2001/index.html](http://bundesrecht.juris.de/bundesrecht/sigg_2001/index.html); legislation of May 16, 2001 published in BGBl I 2001, 876.

### *Qualified Electronic Signature*

Advanced electronic signature as above, that

- is based on an at the time of its creation valid and qualified certificate,
- is created utilizing a secure signature-creation unit.

It is recommended to at least apply the Advanced Electronic Signature.

To facilitate long-term archiving, it is advised that the institution accepting the original files of a publication creates from those new files in an archiving format. The operator of the document and publication repository should sign an agreement with the operator of an archiving repository. No archiving file-format is recommended at the moment.

Documents on the document and publication repository should be published in an established presentation file-format. Currently this is the Portable Document Format (PDF). Should the presentation of the documents not be in this format, either the viewing software should be offered for download with the publication, or a URL be given, where it is available.

## **4.5 Indexing**

### **Subject Indexing**

It is required that a written strategy for subject indexing of the published documents exist within the document and publication repositories' policies. In the long run, indexing not based on a controlled vocabulary will not be sufficient to guarantee good recall and precision in searches for publications.

To permit quality searches across multiple document repositories, international indexing systems should be used (v. DINI OAI-recommendations<sup>12</sup>).

### **Metadata Export and Interfaces**

DINI is not recommending any specific data structure for the publications on document and publication repositories. However, standardized metadata should be created and provided for exchange and harvesting services.

Administrative and technical metadata (long-term availability) of the electronic resources must be provided for archival purposes. Recommended standards such as METS or OAIS aim at this. It is recommended that national and international developments to further enhance long-term availability are taken into consideration. As these developments are currently not finalized, no specific interfaces to long-term-archiving systems can be named (v. 4.7).

## **4.6 Logs and Statistics**

Document-related access statistics, integrated into the document and publication repository, contain indications of the relevance of a publication in relation to other publications.

Access statistics of a document and publication repository are the qualitative, quantitative and technological basis for the evaluation of the repository.

Chronological progression within document-related access statistics can contain indications about the "up-to-dateness" of a publication throughout its life cycle.

---

<sup>12</sup> [http://www.dini.de/documents/dini\\_oai\\_Empfehlungenen\\_07-2002.pdf](http://www.dini.de/documents/dini_oai_Empfehlungenen_07-2002.pdf)

## Comparability

Currently, no standards exist for access statistics. Different protocols and harvesting technologies add the problem of uniform logging and evaluation of data.. Currently no mutual basis exists for a comparison of statistics of document and publication repositories nor of the published documents.

DINI will closely watch international developments like e. g. the standardization project “Counting Online Usage of Networked Electronic Resources”<sup>13</sup> and publish new recommendations when and where appropriate.

It is recommended to store the access statistics for no less than five years.

## 4.7 Long-Term Availability

During a period of five years following the publication date of a document, it is the operators’ responsibility to ensure access to publications in accordance with the requirements of 2.4. This period must be defined in the document and publication repository’s policy.

At the same time, the long-term availability of the electronic resources has to be ensured.

Recommendations to achieve this are provided by the “Network of Competence for Long-Term Archiving and Long-Term Availability”<sup>14</sup>; these recommendations should be adhered to. The operator of the document and publication repository may ensure this independently (depot system according to OAIS), financial and technical means permitting. Alternatively, administrative and organizational infrastructures would have to be implemented to archive in cooperation with an archiving institution<sup>15</sup>. Requirements on the transfer of the electronic resource and the metadata necessary for archiving and usage, as listed in 4.5, must fulfilled.

Currently, the Network of Competence has not published concrete recommendations on an archiving file-format.

---

<sup>13</sup> <http://www.projectcounter.org>

<sup>14</sup> <http://www.dl-forum.de/Foren/Langzeitverfuegbarkeit/>

<sup>15</sup> <http://www.rlg.org/longterm/repositories.pdf>

## DINI Working-Group “Electronic Publishing”

Name	Institution	Email
Kim Braun	Carl-von-Ossietzky Universität Oldenburg, Bibliotheks- und Informationssystem	braun@bis.uni-oldenburg.de
Susanne Dobratz	Humboldt-Universität zu Berlin, AG Elektronisches Publizieren, Universitätsbibliothek	dobratz@cms.hu-berlin.de
Prof. Dr. Eberhard Hilf	Universität Oldenburg, Institute for Science and Networking	hilf@physnet.physik.uni-oldenburg.de
Dr. Bruno Klotz- Berendes	Universität Dortmund, Universitätsbibliothek	klotz-berendes@fh-muenster.de
Dr. Nikola Korb	Die Deutsche Bibliothek	korb@dbf.ddb.de
Prof. Dr. Elmar Mittler	Staats- und Universitätsbibliothek Göttingen	mittler@mail.sub.uni-goettingen.de
Dr. Peter Schirnbacher	Humboldt-Universität zu Berlin, Computer- und Medienservice	schirnbacher@cms.hu-berlin.de
Frank Scholze	Universität Stuttgart Universitätsbibliothek	scholze@ub.uni-stuttgart.de
Thomas Severiens	Universität Oldenburg, Institute for Science Networking	severiens@physnet.physik.uni-oldenburg.de