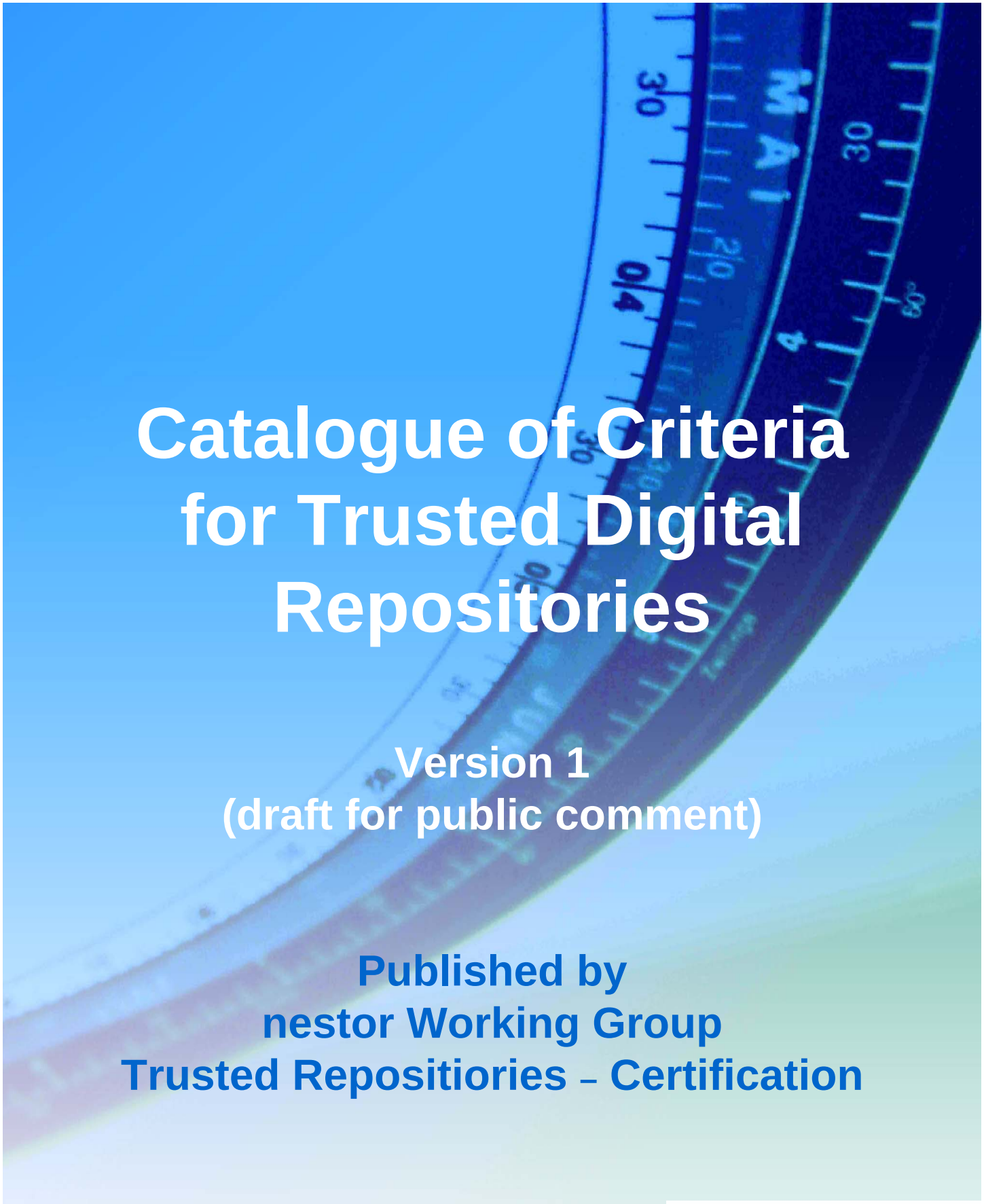

Catalogue of Criteria for Trusted Digital Repositories

Version 1
(draft for public comment)

Published by
nestor Working Group
Trusted Repositiories – Certification

nestor-materials 8

urn:nbn:de:0008-2006060703

nestor - studies - 8

Catalogue of Criteria for Trusted Digital Repositories

Version 1
(draft for public comment)

published by
nestor Working Group
Trusted Repositories - Certification

urn:nbn:de:0008-2006060703

Frankfurt am Main, December 2006

Publication details

nestor.-.studies 8: nestor – Network of Expertise in long-term STORage / nestor Working Group on Trusted Repositories Certification: Catalogue of Criteria for Trusted Digital Repositories, Version 1 (draft for public comment), June 2006, Frankfurt am Main: nestor c/o Deutsche Nationalbibliothek, urn:nbn:de:0008-2006060710

nestor Working Group -Trusted Repositories Certification

c/o Humboldt-Universität zu Berlin, Universitätsbibliothek
Susanne Dobratz
Unter den Linden 6
D-10099 Berlin
Tel.: +49-30-2093-7070
Fax.: +49-30-2093-2959
E-Mail: dobratz@cms.hu-berlin.de

or

c/o Bayerische Staatsbibliothek, Digitale Bibliothek
Dr. Astrid Schoger
80328 München
Tel.: +49-89-28638-2600
Fax.: +49-89-28638-2672
E-Mail: astrid.schoger@bsb-muenchen.de

nestor – Network of Expertise for Long-Term Storage and Long-Term Availability of Digital Resources in Germany

c/o Deutsche Nationalbibliothek
Adickesallee 1
D-60322 Frankfurt am Main
Email Address: lza-info@langzeitarchivierung.de
Web: http://www.langzeitarchivierung.de

Authors of Criteria Catalogue:

Dobratz, Susanne: Humboldt-Universität zu Berlin, University Library
Dr. Hänger, Andrea: Bundesarchiv Koblenz
Huth, Karsten: Bundesarchiv Koblenz
Kaiser, Max: Österreichische Nationalbibliothek Wien
Dr. Keitel, Christian: Landesarchiv Baden-Württemberg
Dr. Klump, Jens: Geoforschungszentrum Potsdam
Rödig, Peter: Institut für Softwaretechnologie, Universität der Bundeswehr München
Dr. Rohde-Enslin, Stefan: Institut für Museumskunde Berlin
Dr. Schoger, Astrid: Bayerische Staatsbibliothek München
Schröder, Kathrin: Deutsche Nationalbibliothek / Bundesarchiv Koblenz
Strathmann, Stefan: Staats- und Universitätsbibliothek Göttingen
Wiesenmüller, Heidrun: Württembergische Landesbibliothek Stuttgart

For their suggestions we are also grateful to:

Dr. Beckschulte, Klaus: Börsenverein des deutschen Buchhandels, Landesverband Bayern
Dr. Korb, Nikola: Deutsche Nationalbibliothek
Dr. Lupprian, Karl-Ernst: Generaldirektion der Staatlichen Archive Bayerns
Dr. Schomburg, Silke: Hochschulbibliothekszentrum Köln
Steinke, Tobias: Deutsche Nationalbibliothek

We would also like to thank the participants of the workshop held on 21 June 2005 in the Bavarian State Library in Munich and those who took part in the specialist conference held on 29 March 2006 at the German National Library in Frankfurt/Main.

Contents

SUMMARY 1
I. Introduction 2
Long-term preservation of digital objects - Basic concepts 2
Threats to the preservation of information, trustworthiness 2
Digital objects 2
Metadata 2
Digital repository 2
Use by designated community 2
Trustworthiness 3
The road to creating a trusted digital repository 3
The nestor criteria catalogue 4
Basic principles for the derivation of criteria 4
Abstraction 4
Conformity with OAIS terminology 4
Basic principles for application of criteria 4
Documentation 4
Transparency 4
Adequacy 4
Measurability 5
The nestor Working Group on Trusted Repositories Certification 5
II. Criteria catalogue 7
A. Organisational framework 7
B. Object management 15
C. Infrastructure and Security 27
III. Checklist 29
IV. Glossary and abbreviations 33
V. Bibliography 35

SUMMARY

Digital information has become an indispensable part of our cultural and scientific heritage. Scientific findings, historical documents and cultural achievements are increasingly being presented in electronic form, and in many cases exclusively so. However, despite the irrefutable benefits offered by digital content, there are a number of associated disadvantages. Users must invest a great deal of technical effort in order to access such information. Underlying technology continues to undergo development at an exceptionally fast pace and the rapid obsolescence of access technologies combined with at times imperceptible physical decay of storage media themselves represents a serious threat to preservation of the information content, both contemporaneously and in the long term.

These circumstances have provoked questions of information trustworthiness. Information producers and consumers wish to identify the memory organisations that are capable of ensuring the authenticity, integrity, confidentiality and availability of digital information. Confronted with the inexorable flood of digital objects, those responsible within the institutions are similarly motivated to establish and communicate their trustworthiness whether it is to fulfil a legal requirement or to simply survive within the market.

This is the main focus of the work of the nestor Working Group on Trusted Repositories Certification. It identifies criteria which facilitate the evaluation of digital repository trustworthiness, both at organisational and technical levels. The criteria are defined in close collaboration with a wide range of different memory organisations, information producers, experts and other interested parties. This open approach is the basis for achieving a high degree of universal validity and practical applicability and facilitates broad-based acceptance of the results of any evaluations conducted on the basis of these criteria. The present criteria catalogue for public comment represents an important milestone on the road towards achieving the working group's goals. The memory organisations should be given a well-constructed, coordinated and practical tool for achieving and demonstrating their trustworthiness. However, the intention is also to present the opportunity for repository certification within a standardised national or international process as a formal endorsement of an organisation’s trustworthiness. The document’s current draft also supports active participation in existing international standardisation efforts.

This document begins by offering a brief introduction into the problems surrounding the long-term preservation of digital objects. A description of key concepts and principles underpinning the criteria catalogue follows, ensuring understanding and limiting ambiguity. The aims and methods of the working group are then briefly outlined. The criteria catalogue itself follows this introduction, in its full, unabridged form. The document concludes with a compact overview of the catalogue in checklist format, and a glossary.

I. Introduction

Long-term preservation of digital objects - Basic concepts

Threats to the preservation of information, trustworthiness

Information in the form of digital objects faces numerous threats to its integrity, authenticity and security. In the worst cases this can result in total lossaccessibility and usability. Physical aging of storage media, separation of information from its original data carriers, and rapid changes in the technical infrastructure required to interpret digital objects represent key challenges for long-term preservation.
The purpose of many digital repositories is to preserve information over long periods of time. Both organisational and technical measures must be taken in order to counter these threats. Trusted digital repositories will have their own targets and specifications. Their trustworthiness can be tested and assessed on the basis of a criteria catalogue.

Digital objects

Within the context of this criteria catalogue, a digital object is a logically discrete unit of information in the form of digital data. Data is a machine-readable and processable representation of information in digital form (a sequence of bits, that is, zeros and ones). In order to use the information this digital data must be interpreted (decoded).
Within this context 'information' covers all types of communicable knowledge content, including works of intellectual creativity, results of research and development, and documentation of political, social and economic events.
Digital objects are frequently organised into files. A digital object can be a single file (such as a digital photo saved as a TIFF file) or several different, but related files (often described as a complex object, for instance an electronic journal consisting of individual articles saved as PDF files). In addition to the content data, a digital object may also contain metadata. Furthermore, a file may incorporate a number of digital objects (for example. a database file).
The concept of the digital object presented here is based on the information model described in the ReferenceModel for an Open Archival Information System (ISO 14721:2003, OAIS).

Metadata

Additional data may be supplemented with the content information in order to help identify, search for, reconstruct, interpret or document the integrity and authenticity of the content and manage its usage rights. Such metadata can be created at various times within the lifecycle of digital objects (e.g. during production, archiving or provision for use). Metadata are interpreted as being part of the logical "digital object" unit and can be physically linked to the content data, or recorded separately

Digital repository

For the purposes of this criteria catalogue, a digital repository is defined as an organisation (consisting of both people and technical systems) that has assumed responsibility for the long-term preservation and long-term accessibility of digital objects, ensuring their usability by a specified target group, or 'designated community'. "Long-term" in this context means beyond technological changes (to hardware and software) and also any changes to this designated community. Once more, this definition of digital repository is based on that introduced within the OAIS Reference Model.

Use by designated community

Future use is contingent not only upon the integrity, authenticity, confidentiality and availability of the digital objects being preserved, but also on the designated community being able to continue to understand and use the digital objects. Legal or organisational changes and technical developments can result in changes within the designated community and intheir needs and expectations. A digital repository must therefore monitor these changes and react accordingly.

Trustworthiness

Trustworthiness is the capacity of a system to operate in accordance with its objectives and specifications (that is, to do exactly what it claims to do). From an IT security perspective, the fundamental considerationsare integrity, authenticity, confidentiality and availability. IT security is therefore an important prerequisite for trusted digital repositories.

There is great diversity within existing and emerging digital repositories, as can bedemonstrated with the following typical examples:

Example 1: A large academic library with responsibility for continually growing collections of digital publications from publishers and official sources, for scientifically relevant Internet resources and for the results of digitisation projects. The designated community for this digital repository is the general public. There are many different producers including publishers, digitisation centres, and private individuals etc. This repository might also carry out long-term preservation services for smaller institutions. It may also be part of a network which permits cooperation with other libraries and grants users uniform access to cooperatively organised resources.

Example 2: A university library that, in addition to commercial scientific literature, also maintains eLearning modules, university publications, and publications by university staff members. Within this example the users are the students and the university employees. The producers are mostly university staff members.

Example 3: A research institution that generates and archives large quantities of specialist data. Its designated community consists of scientists with the necessary specialist knowledge for interpreting this data.

Example 4: An archive that stores electronic documents from administrative organisations on the basis of legal archive requirements. In addition to the general public, its main designated community is the producers. Use may be prohibited over longer periods by means of protective rights.

Example 5: A museum that manages the digitisation of museum objects and also original digital art. Users are the general public, art experts, and artists.

Example 6: A service provider that carries out long-term preservation contract work for other institutions and their collections. The institutions themselves are responsible for building up the collections; the service provider offers reliable preservation of the digital objects, ensuring their ongoing availability and usability.

The road to creating a trusted digital repository

A long-term digital repository is a complex interrelated system. Implementation of the individual criteria must always be undertaken in the light of the objectives of the overall system. Both realisation of the long-term digital repository as a whole and the fulfilment of individual criteria are multi-stage processes:
1. Conception
2. Planning and Specification
3. Realisation and Implementation
4. Evaluation

These steps should not be regarded as a rigid phase model. Rather they must be repeated at regular intervals as the result of continuous improvements. Quality management is deployed to monitor this development process.

The nestor criteria catalogue

Users of the Criteria Catalogue
The present criteria catalogue is principally aimed at memory organisations (archives, libraries, museums) and serves as a manual for devising, planning and implementing a trusted digital long-term repository. It can also be used at all stages of development for self-checking.

In addition, this catalogue is intended to provide guidance to all institutions currently administering archives, commercial and non-commercial service providers, and third party service providers.

Basic principles for the derivation of criteria

Abstraction

The aim of this catalogue is to formulate criteria that can be used for a broad spectrum of digital long-term repositories and that will retain their validity over a longer period. The assumption is that a selection of relatively abstract criteria is appropriate. The criteria are each accompanied by extensive explanations and concrete examples from different fields. The examples are state-of-the-art in terms of technology and organisation, although in some cases they may only make sense within the context of a particular archiving task. They make no claim to being exhaustive.

Conformity with OAIS terminology

The OAIS reference model [CCSDS: Producer-Archive Interface Methodology - Abstract Standard, Blue Book, 2004] together with its functional entities and information model serves - where possible - as the basis for providing common terms and for structuring the criteria catalogue. The OAIS is used to describe the core processes from ingest of the digital objects into the digital repository, via archival storage through to usage; on the other hand it is also used to describe the life cycle of digital objects from the producer via the digital long-term repository through to the user. For this the following information units have been considered: Submission Information Package (SIP) for ingest, Archival Information Package (AIP) for archival storage and Dissemination Information Package (DIP) for access.

Basic principles for application of criteria

Documentation

The objectives, basic concept, specifications and implementation of the digital long-term repository should be documented. The documentation can be used to evaluate the status of development both internally and externally. Early evaluation can serve to avoid errors caused by inappropriate implementation. Correct documentation of workflow also allows` verification of any evaluatory conclusions. All quality and security standards must also be suitably documented.

Transparency

Transparency is achieved by publishing appropriate parts of the documentation. External transparency for users and partners enables these stakeholders to themselves gauge the degree of trustworthiness. Transparency afforded to producers and suppliers enables these groups to determine to whom they wish to entrust their digital objects.
Internal transparency facilitates reflective self-assessment by the operators, backers, management and also employees With respect to sensitive or confidential documentation (e.g. company secrets, security-related information), transparency can be restricted to a specified group, such as certifying auditors
The principle of transparency relates closely to trust as it permits interested parties to make a direct assessment of the quality of a digital repository.

Adequacy

The principle of adequacy derives from the fact that the conception of absolute standards is somewhat unfeasible; rather that evaluation is always based on the objectives and tasks of the individual digital repository concerned. The criteria have to be related to the context of each individual archiving task. Individual criteria may therefore prove irrelevant. Depending on the

objectives and tasks of the digital repository, the required degree of compliance for a particular criterion may differ.

Measurability

In some cases - especially with regard to long-term issues - there are no objectively measurable characteristics. In such cases we must instead rely on indicators that demonstrate the degree of trustworthiness. Again, transparency makes the indicators accessible for evaluation.

The nestor Working Group on Trusted Repositories Certification

The nestor Working Group on Trusted Repositories Certification has been established within the BMBF (Federal Ministry of Education and Research) sponsored nestor project in order to define a first catalogue of criteria for trustworthiness and to prepare for the certification of digital repositories in accordance with nationally and internationally coordinated procedures. The members of the working group represent a range of communities including libraries, archives, museums, research institutions, publishing houses, software developers and certifying agencies. The current status of long-term preservation provides the basis for developing evaluation criteria which are realistic within the context of contemporary organisations and technology. In order to gain an overview the working group surveyed a representative selection of institutions such as libraries, archives, museums, research institutions, publishing houses, companies, broadcasting corporations and weather forecasting services on the status of their long-term preservation activities. The questionnaires used can be downloaded from the Working Group on Trusted repositories Certification - section of the nestor portal1. The results of the survey demonstrated that the procedures and the organisational systems used are highly heterogeneous and that, in many cases, no standards exist as yet.

A related workshop was held by the working group on 21 June 2005, and was attended by roughly 70 representatives from a variety of areas. It provided further evidence that in Germany, scarcely any guidelines, methods and tools are available which are suitable for daily use and which support the systematic construction and operation of digital repositories. This was particularly apparent from the audience's demands for a criteria catalogue to serve as an orientation and self-check tool in the design, planning and implementation of digital repositories.

A first draft of the nestor criteria catalogue was presented and discussed on 29 March 2006 in an expert round table meeting involving roughly 50 participants. The overall aims of the criteria catalogue, the principles on which it is based and also the catalogue itself met with broad-based acceptance. Suggestions made during the expert round table were then fed into the current version. Participants welcomed each of the standardisation and certification outcomes identifiedby the working group.

The nestor catalogue has been compiled mainly for application in Germany, however it is also being discussed and standardised within the international context. It is crucial to identify generally valid criteria amongst the specific, national conditions. These lie, among other areas, within the legal framework, the provision of public institutions with adequate financial and human resources, the national organisational structure and the status of national development in the field of digital long-term preservation.

1 http://www.langzeitarchivierung.de/ag-repositories

http://nestor.cms.hu-berlin.de/tiki/tiki-index.php?page=wg-repositories

The nestor criteria catalogue takes into consideration national and international approaches and findings such as the DINI Certificate for document and publication servers [Dokumenten- und Publikationsserver der Humboldt-Universität zu Berlin: Ziele und inhaltliche Kriterien, 2006], the RLG-OCLC report "Trusted Digital Repositories: Attributes and Responsibilities" (May 2002) [RLG Working Group on Digital Archive Attributes (2002): Trusted Digital Repositories: Attributes and Responsibilities] and the draft "Audit Checklist for Certifying Digital Repositories" (2006) [RLG NARA Task Force on Digital Repository Certification (2005): Audit Checklist for Certifying Digital Repositories] published by the RLG/NARA Task Force. Audit Checklist for Certifying Digital Repositories]. The working group is also in contact with the RLG/NARA Digital Repository Certification Task Force2, the Digital Curation Centre3, the EU project "Digital Preservation Europe"4 and the DELOS Digital Preservation Cluster5

In order to develop a broadly accepted criteria catalogue, nestor needs input and comment from the institutions that are affected or have an interest. For this reason the working group adopted an open procedure from the outset to tackle the problems jointly and to involve all interest groups from an early stage.

The aim of the public invitation to comment upon this criteria catalogue, in which a large number of suggestions have already been incorporated, is to create a solid and practical foundation for developing an evaluation and certification procedure. This task is to be continued in the follow-on project "nestor II" which will include national and international standardisation activities.

For reasons of brevity the term "digital repository" is abbreviated to "DR" in the catalogue below.

The following overview shows the structure of the criteria catalogue:

Criterion
General explanation of the criterion
`Examples, comments, notes from different application areas, with no claim to exhaustiveness`
Literature related to this criterion

2 http://www.rlg.org/en/page.php?Page_ID=367.

3 http://www.dcc.ac.uk/.

4 http://www.digitalpreservationeurope.eu/.

5 http://www.dpc.delos.info/.

II. Criteria catalogue

A. Organisational framework

The digital repository acts within an organisational framework that is determined by the definition of its goals, the legal context and the staffing and financial resources available.

1 The digital repository has defined its goals.

The DR should have a clear conception of its objectives. It has determined which tasks it fulfils, and which principles it observes in doing so. This is crucial, as trustworthiness is not an absolute term, rather it depends on the goals of the particular DR. Following the principle of adequacy, evaluation of the individual criteria is always based on the specific goals. The DR ensures that its objectives are transparent so that others - most notably users and producers - can themselves gauge the repository's trustworthiness. (The goals are often published in the form of a "policy".)

[PANDORA: The purpose of the PANDORA Archive, 2006]

[Oxford Digital Library: Background, Services, Principles and Guidelines, 2006]

[Dokumenten- und Publikationsserver der Humboldt-Universität zu Berlin: Ziele und inhaltliche Kriterien, 2006]

[National Archives: Custodial policy for digital records, 2006]

[Erpanet: Erpanet-Tagung "Policies for Digital Preservation", 2003]

1.1 The digital repository has developed criteria for the selection of its digital objects.

The DR should have laid down which digital objects fall within its scope. This is often determined by the institution's overall task area, or stipulated by laws. The DR has developed collection guidelines, selection criteria, evaluation criteria or heritage generation criteria. The criteria may be content-based, formal or qualitative in nature.

In the case of both state-owned and non-state-owned archives, the formal responsibility is generally derived from the relevant laws or the entity behind the archive (a state-owned archive accepts the documents of the state government, a corporate archive the documents of the company, a university archive, the documents of the university).

German National Library law - draft law approved by Bundesrat, Article 2 Tasks and authorisation:

The Library is tasked with:
1. collecting, making an inventory of, analysing and bibliographically recording a) originals of all media works published since 1913 and b) originals of all foreign media works published in German since 1913, and ensuring the long-term preservation of these works, rendering them accessible to the general public, and providing central library and national library services.

Supported by the state libraries, the Baden-Württemberg online archive (BOA - http://www.boa-bw.de/) collects net publications …"which originate in Baden-Württemberg, or the content of which is related to the state, its towns and villages or inhabitants."

The Oxford Text Archive http://ota.ahds.ac.uk/ collects "high-quality scholarly electronic texts and linguistic corpora (and any related resources) of long-term interest and use across the range of humanities disciplines". The website contains a detailed "collections policy".

The document and publication server of the Humboldt University in Berlin collects "electronic academic documents published by employees of the

Humboldt University" http://edoc.hu-berlin.de/e_info/leitlinien.php.

[Erpanet: Erpanet "Appraisal of Scientific Data" conference, 2003]

[Interpares Appraisal Task Force: Appraisal of Electronic Records: A Review of the Literature in English, 2006]

[Wiesenmüller, Heidrun et al.: Auswahlkriterien für das Sammeln von Netzpublikationen im Rahmen des elektronischen Pflichtexemplars : Empfehlungen der Arbeitsgemeinschaft der Regionalbibliotheken, 2004]

1.2 The digital repository assumes responsibility for long-term preservation of the information represented by the digital objects.

The DR explicitly declares its responsibility for the long-term preservation of the digital objects collected as described under 1.1. Long-term preservation here means permanent retention of the usability of the information represented by the digital objects (cf. the OAIS information model).

Formulation on the website of the Internet Archive http://www.archive.org/about/about.php: "The Internet Archive is working to prevent the Internet (...) and other "born-digital" materials from disappearing into the past. Collaborating with institutions including the Library of Congress and the Smithsonian, we are working to preserve a record for generations to come."

Formulation on the website of the Oxford Digital Library http://www.odl.ox.ac.uk/principles.htm: "Like traditional collection development, long-term sustainability and permanent availability are major goals for the Oxford Digital Library."

1.3 The digital repository has defined its designated community(ies).

The general definition of the framework for a DR involves defining the designated community(ies)/designated community. This includes knowledge of the specific requirements of the designated community(ies) influencing the selection of the services to be provided.
If the designated community or its requirements change over time, the DR should respond by adapting its services.

Possible designated communities include:

- Employees of an official body, a research institute etc.
- Scientists working in a particular discipline
- The general public

2 The digital repository grants its designated community adequate usage of the information represented by the digital objects.

The DR should regard its primary task as ensuring the current and future use of the information represented by the digital objects on the part of its designated community. Use of the digital objects relies on their preservation, their accessibility and their understandability. Use may be adequate despite legal restrictions(c.f. 3.3.) loss of some characteristics of the original (c.f. 9.2.).

So-called "dark archives" are established with no external access; they will only be used if the primary archive is rendered non-functional for whatever reason. In the event of such a crisis, use must then be guaranteed.

2.1 The digital repository ensures its designated community can access the digital objects.

The DR should ensure that authorised users have access to the digital objects. This includes the provision of adequate research opportunities. When determining its service portfolio, the DR takes considers the needs of its designated community into account. The DR announces in advance its conditions of use and any costs whichcosts that may arise, listing documentedthese in a transparent manner.

Access can mean:

- Accessing the digital objects
- Creating or supplying an analogue copy (e.g., as print-out by the user or in the form of a print-on-demand service)
- Creating or supplying a digital copy (e.g. download to a storage medium by the user, email delivery)
- Creating interfaces to permit access via other systems to the digital objects.

2.2 The digital repository ensures that the designated community can interpret the digital objects.

The DR should take appropriate measures to ensure that the digital objects can be interpreted on a long-term basis, thereby creating the basic requisites for adequate usage. This includes the ability to interpret both content and metadata. In ensuring this, the DR should consider the needs of its designated. The more specialised the designated community, the more know-how and technical equipment (such as specialist software) is required, or the repository must demonstrate greater willingness to provide additional equipment (for example, the installation of plug-ins).
Changes to the technical environment or the designated community can influence the interpretability of objects. The DR should therefore check at regular intervals, using appropriate procedures, to determine whether the objects are still interpretable by the designated community.

Possible measures include:

- Conversion into a current standard format
- Provision of emulation software (e.g. open source DOS emulator "DOSBox")
- Provision of representation information: e.g. documentation of data structures and field content to ensure that users can transfer data from specialist applications (databases) into the respective current database applications
- Provision of instructions for use, installation instructions, help texts
- DR carries out research or evaluation work as (charged) service
- Checking of interpretability on the basis of regular spot checks
- Provision of a feedback form with which users can register interpretation problems

3 Legal and contractual rules are observed.

The DR's actions should reflect legal regulations. These may cover the acquisition of the digital objects and also their archiving and use. The DR should strike a balance between the legitimate interests of the producers and those of the users and also, where applicable, the individuals concerned (in the case of person-related data).

[Solicitors Goebel and Scheller (Bad Homburg v.d.H.): nestor - materialien 1: Digitale Langzeitarchivierung und Recht, 2004]

3.1 Legal contracts exist between producers and the digital repository.

In order to ensure planning and legal security the DR, where possible, should conclude formal agreements with the producers or suppliers. The nature and scope of the delivery, the DR's archival obligations, the conditions of use and, where applicable, the costs should be legally defined. The legal agreements should be supplemented with concrete implementation provisions. If it is not possible to conclude a formal agreement, the grounds for this should be given.

Possible agreements include:

1. Laws, ordinances: law of obligation, archive laws: law governing the German National Library - draft law approved by Bundesrat, https://www.umwelt-online.de/PDFBR/2005/0396_2D05.pdf

2. Contracts, agreements:

Licence agreements (cf. archiving clause in JISC model contract for electronic journals: http://www.nesli2.ac.uk/NESLi2_licence_journals_final011003.htm)

Framework contracts (cf. DDB framework contract with Börsenverein des deutschen Buchhandels), http://deposit.ddb.de/netzpub/web_rahmenvereinbarung.htm

contracts of custody, archiving agreements, archiving and usage permits (cf. Austrian National Library(), http://www.onb.ac.at/about/lza/

Such an agreement, or its implementation clauses, define e.g.:

a) in which form the cooperation between producer / supplier and the DR should take; how feedback is organised.

b) Type and scope of supply:
Scope, schedules, acquisition procedure (data carrier, file transfer via networks, upload, download), file formats, other file properties (e.g. without active elements), additional information (e.g. the content and structure of descriptive metadata, SXL schema etc).

c) DR obligation:
Time of assumption of legal responsibility, duration of archiving, use of preservation measures (multiple copies, change-causing actions e.g. during migrations), significant characteristics.

d) Conditions of use:
designated communities, services offered, usage rights, costs.

There is no possibility of a formal agreement regarding the archiving of STASI documents as neither the rights holder nor a legal successor exists.

3.2 In carrying out its archiving tasks, the digital repository acts on the basis of legal rulings.

The DR should take legal requirements and contractual obligations into consideration regarding its archival storage and the use of preservation measures.

Restrictions imposed on archival storage by copyright can e.g. be countered by explicit agreements on the right to multiple storage, file-altering actions etc.

3.3 *With regard to use, the digital repository acts on the basis of legal requirements.*

The DR should take legal requirements and contractual obligations into consideration regarding the use of digital objects. If this results in restrictions to their use, the reason(s) for the restrictions should be documented.

Legal requirements which can influence use include copyright, data protection, other legal regulations (e.g. periods of copyright for archives), contractual obligations or the contractual or legal purpose-tying of use.

Restrictions on use can be countered in some cases by controlled access to the digital objects. For the observance of copyright restrictions this could involve registration, exclusive use on the premises or on the Intranet or through charged usage/billing models. Separate declarations of commitment or the issue of anonymised user copies are possible options for compliance with data protection and archive regulations.

4 The organisational form is adequate for the digital repository.

The digital repository should be organised in such a way that it can fulfil its short, medium and long-term goals. Its effectiveness and sustainability should facilitate evaluation by users and producers. This evaluation is based on the following points.

[Erpanet: Erpanet-Tagung BusinessModelsrelatedtoDigitalPreservation, 2004]

4.1 Adequate financing of the digital repository is secured.

The digital repository should be able to demonstrate that the proposed services can be financed, both in the short and long term.

The financing of the digital repository should have a legally secured basis.
In the case of state-financed digital repositories, the financing should be included in the formal planning documents (at least medium-term).
A private digital repository should be able to guarantee its financial sustainability on the basis of charged use of its services and on a long-term business plan.

[Digitaleduurzaamheid: Kostenmodell für die Langzeitarchivierung siehe: Vers van de pers..... 'Kostenmodel digitale bewaring', 2006]

[Palm,Jonas: The Digital Black Hole, 2006]

[Oltmans,Erik and Kol,Nanda: A Comparison Between Migration and Emulation in Terms of Costs, 2006]

4.2 Sufficient numbers of appropriately qualified staff are available

The qualifications and training of the staff should be adequate for the goals, tasks and processes of the DR. Suitable schemes should be in place to ensure adequate training and further training in the long term. Staff numbers should be sufficient to allow all necessary processes to be fully completed. The long-term planning of the DR should consider staffing resources.

Staff development includes task-based initial and further training, e.g. through courses and the provision of appropriate literature.

An aspect of training is active participation in relevant national and international conferences and working groups plus work on standardisation bodies. Such active participation makes the training levels of the staff externally visible.

Any shortfall in internal capacity can be compensated by external capacity.

4.3 Appropriate organisational structures exist for the digital repository.

The organisational structure should be adequate for the targets, tasks and processes of the DR. The processes and the allocation of staff and other resources should be structured in such a way that the defined goals can be met.

4.4 The digital repository engages in long-term planning.

The DR should engage in pre-emptive planning covering imminent or expected tasks, plus the deadlines by which they are to be completed. The management should have suitable structures and procedures for strategic planning. The basis for long-term planning is the adequate monitoring of legal and social changes, the demands and expectations of the designated communities (in OAIS: "Monitor Designated Community") and all technical developments (in OAIS: "Monitor Technology")that are relevant for the sustained preservation and appropriate use of the information represented by the digital objects. The planning also includes securing the necessary resources.

Relevant legislative processes should be monitored right from the early phases (e.g. law on basic conditions for electronic signatures).

Strategic planning requires access to reliable current data. Process cost accounting, for instance, helps long-term planning of the required resources.

4.5 Continuation of the preservation tasks is ensured even beyond the existence of the digital repository.

The DR should have made emergency plans. These should describe processes to enable the preservation work to continue within an alternative organisational framework, thereby ensuring that the requirements can continue to be completed. Where this is not possible, any restrictions should be documented. The DR should take precautions to ensure that any transition process can be defined, planned and implemented in good time. Suitable documentation is the basis for the success of such transition.

This includes exportability of all the archive objects (including metadata) in a form which can be interpreted by the successor as a means of guaranteeing the interpretability and authenticity of the data.

An external or higher body should guarantee continuation of the defined tasks.

Continuation should be governed by an agreement with a comparable organisation.

5 Adequate quality management is conducted

Quality management should ensure that the DR's goals are reached. The general targets should be broken down into specific aims and objectives. Suitable quality assurance structures should be established and monitored by the quality management system.
The quality management should be a cross-sectional process covering all parts of the DR.

[Erpanet: Erpanet-Tagung AuditandCertificationinDigitalPreservation, 2004]

[ISO 9000:2005 Quality management systems -- Fundamentals and vocabulary, 2005]

[Liggesmeyer,Peter: Softwarequalität, 2002]

[Kneuper,Ralf: Verbesserung von Softwareprozessen mit Capability Maturity Model Integration, 2006]

5.1 All processes and responsibilities have been defined.

The quality management system should ensure that all processes and their interactions are defined, in particular that individuals are assigned responsibility for each process. This also applies to external (outsourced) processes.

It is easier to determine the completeness of the processes and their interactions if a suitable reference model is available. The OAIS "functional entities" of Ingest, Archival Storage and Access can be used as the basis for defining the core processes. Support and management processes (data management, quality management, etc.) can then be defined on the basis of these core processes.

External processes require an internal process which contractually defines the services and a process which checks these services. Responsibility should be assigned for these internal processes.

[CCSDS: Producer-Archive Interface Methodology - Abstract Standard, Blue Book, 2004]

[Erpanet: Workshop on Workflow, 2004]

5.2 The digital repository documents all its elements based on a defined process.

The quality management system should provide a suitable procedure for documentation, that is, a system to manage all necessary documents. The DR should lay down rules regarding the completeness, correctness, validity, comprehensibility and accessibility of the documentation, which are implemented and monitored.

Standardised terminology, for instance, which is adapted to the needs of the documentation users, helps improve comprehensibility. Accordingly the documentation can be formal (e.g. for description of critical software processes), semi-formal (for conceptual description of processes and IT

infrastructure) or natural (e.g. for external description of archive's objectives).

- Software documentation
- Process documentation
- Documentation of object formats

5.3 The digital repository reacts to substantial changes

Substantial alterations are those threaten the continued fulfilment of goals or introduce additional risks. Substantial alterations can be technical, organisational or community-based.

For this the management should incorporate a process that monitors changes, recognises the likelihood of their occurrence, evaluates their possible effects on task fulfilment and plans, and implements and monitors any necessary alterations.

The monitoring of technical developments includes e.g. the development and standardisation of new file formats and new storage techniques and, accordingly, any obsolescence of existing technologies arising as a result.

A substantial technical change could be a fundamental change in human-machine-communication.

A substantial change affecting the organisation as a whole could be the loss of a backer and therefore the financial base.

B. Object management

The digital repository should analyse its goals and strategies, and specify all object-related requirements for digital object management during the lifecycle of the objects in the DR. The main phases correspond in the OAIS reference model to the processes ("functional entities") of submission (ingest), storage (archival storage, including implementation of long-term preservation measures), and usage (access). Additions to these functions may become necessary depending on the goals of the digital repository. Object management is based on the OAIS information model that defines Submission Information Packages (SIPs), Archive Information Packages (AIPs) and Dissemination Information Packages (DIPS). Appropriate object-related planning of the long-term archiving measures ("Preservation Planning" in OAIS) should be undertaken (cf. 8). The DR should ensure appropriate data management (cf. 12) in its handling of digital objects. Object management requirements are the prerequisites for planning and operating the technical infrastructure and security system (cf. 13).

Object management generally covers the following aspects:

- Object integrity (including metadata)
- Object authenticity (including metadata)
- Object availability (including metadata)
- Object confidentiality (including metadata)

These aspects can be assigned to the area of IT security.

The sustainability of these and further aspects:

- long-term tracing and referencing capacity of the objects (including metadata) and
- long-term interpretability of the objects (including metadata)

requires monitoring beyond what is commonly understood by IT security.

6 The digital repository ensures the integrity of the digital objects during all processing stages.

Integrity refers to the completeness of the digital objects and to the exclusion of unintended modifications as defined in the preservation rules. Integrity is measured in terms of the characteristics of the digital object being preserved (cf. 9.2.).
Innappropriate modifications may be causedby human error (deliberate or accidental), technical imperfections or damage to/theft of the technical infrastructure.

The DR should take both organisational and technical precautions to secure the integrity of objects within their custody.

The DR should operate a data management system suitable for preserving integrity for the processes of ingestion, archival storage and access. The DR should also take precautions regarding the integrity of the data management system itself.

An example of deliberate or accidental modification is the input of virus-infected objects, the execution of which can result in the changing or modification of objects or other system elements (e.g. database scripts which delete objects or metadata).

Examples of technical imperfections are: faulty or incomplete software, especially that used for complex transformations (migrations), and storage media which is obsolete or has not been stored in conformity with the specifications. Generally, however, technical imperfections which are foreseeable should be remedied or flagged by means of appropriate error correction or error identification procedures. In some cases the user can select higher level error correction procedures for certain system components (e.g. through a higher degree of redundancy). This should be made full use of, where possible.

[ISO 15489-1, Information and documentation, Records Management, 2001]

[Network Working Group,Shirey R: Internet Security Glossary, 2000]

[ISO/IEC 15408-x:2005 Information technology -- Security techniques -- Evaluation criteria for IT security, 2005]

[Interpares: Ergebnisse des InterPares-Projekts, 2006]

6.1 Ingest: the digital repository ensures the integrity of the digital objects.

The DR should define agreements with its producers and/or suppliers regarding the technical aspects of the submission (ingest) transactions. In particular, there should be an agreement governing the transfer of responsibility for the maintenance of object integrity.

The DR should agree with the producer/supplier any digital objectcharacteristics that are required to mitigate integrity risks.. Examples include the prior removal of executable code in digital documents.

The DR should ensure secure transfer channels from the producer / supplier to the DR.

The DR should conduct checks to ensure the completeness and quality of the deliveries.

[CCSDS: Producer-Archive Interface Methodology - Abstract Standard, Blue Book, 2004]

[Littman,Justin : A Technical Approach and Distributed Model for Validation of Digital Objects, Volume 12 Number 5, 2006]

6.2 Archival Storage: the digital repository ensures the integrity of the digital objects.

The DR should have transparent procedures for determining the required degree of physical redundancy and suitable locations for storage media and related subsystems.

The DR should stipulate the required specification for storage media (e.g. the use of standardised and certified storage media).

The DR should define a policy regarding logical access to the archive store; this should include internal DR users such as system administrators.

The DR should strictly regulate physical access to the IT systems.

6.3 Access: the digital repository ensures the integrity of the digital objects.

The DR should define a clear interface for the user. It should permit the user to check the integrity of the digital objects.

The DR should ensure that no unauthorised user can obtain rights over digital objects, metadata or other system elements.

The DR should define how far its responsibility for the integrity of the digital objects extends in the delivery process.

The DR should analyse the quality of the representation information it provides and make the results available to the users.

7 The digital repository ensures the authenticity of the digital objects during all stages of processing.

Authenticity here means that the object actually contains what it claims to contain. The DR should document where authenticity cannot be demonstrated for a particular object.
The DR should operate a data management system suitable for preserving authenticity within the ingest, archival storage and access processes. This is achieved to some extent by documenting all changes to the objects (including metadata) (see 12.4).

Authenticity means that the producer or sender and the given production or transmission time correspond to the facts.For example, that an email supposedly generated and transmitted by a particular person at a particular time is actually from this person and was sent at the given time.

[ISO 15489-1, Information and documentation, Records Management, 2001]

[Network Working Group,Shirey R: Internet Security Glossary, 2000]

[ISO/IEC 15408-x:2005 Information technology - Security techniques - Evaluation criteria for IT security, 2005]

[PREMIS Working Group: Data Dictionary for Preservation Metadata: Final Report of the PREMIS Working Group (with a glossary), 2005]

7.1 Ingest: the digital repository ensures the authenticity of the digital objects.

The DR should demand the formal registration of producers/suppliers with an authorised body.

In certain contexts the use of digital signatures can ensure the authenticity of the objects being transferred.

The DR should require the producer/supplier to define and undertake procedures to assess the authenticity of the digital objects, for example on the basis of metadata describing the origin.

CCSDS, Producer-Archive Interface Methodology – Abstract Standard, Blue Book, 2004

7.2 Archival Storage: the digital repository ensures the authenticity of the digital objects.

The DR should keep full documentation of all transforming (that is, altering or deleting) operations on the digital objects (including metadata).

7.3 Access: the digital repository ensures the authenticity of the digital objects.

The digital repository should be capable of authenticating itself to the user as the supplier of usage objects. It should provide documentation to the user in cases where the authenticity of the digital objects is not clear.

The DR should provide the user with metadata that documents the origin and all changes in the archiving process, thereby permitting evaluation of authenticity.

The DR should register with an authorised body, for example the regulator for postal and telecommunications affairs, from which it should receive a digital signature key certificate to be used to generate digital signatures.

The DR should use digital signatures for the delivery of usage objects.

The method deployed in the ArchiSafe project by the Physikalisch-Technische Bundesanstalt Braunschweig for the use of digital signatures, cf. http://www.archisafe.de/s/archisafe/index

8 The digital repository has a strategic plan for its technical preservation measures.

In order to fulfil its responsibility for preserving information, the DR should have a strategic plan covering all outstanding or expected tasks, and the timetable fortheir completion. This strategic planning (cf. 4.4) should be specified at the object level. Such measures should keep pace with ongoing technical developments (such as changes to data carriers, data formats, and user demands).

Measures for physical data preservation (integrity, authenticity), its accessibility and the preservation of its interpretability should be conceived to provide long-term preservation functionality. Long-term preservation measures cover both content and metadata.

See 10.4 regarding implementation of the long-term preservation measures.

Output onto analogue media (e.g. microfilm) and redigitisation may be appropriate for certain digital objects.

The following are the main methods used to preserve interpretability:

Conversion to a current format or a current format version (migration)

Recreation of the old application environment within a new technical infrastructure (emulation).

Long-term planning of the tasks arising from the formats can be based e.g. on a format register. Format registers are currently being developed by e.g. Harvard (Global Digital Format Registry: http://hul.harvard.edu/gdfr/) and the National Archives, Kew (PRONOM: http://www.nationalarchives.gov.uk/pronom/).

[DigiCULT: Technology Watch Reports, 2006]

[Rauch,Carl und Rauber,Andreas: Anwendung der Nutzwertanalyse zur Bewertung von Strategien zur langfristigen Erhaltung digitale Objekte, 2006]

9 The digital repository accepts digital objects from the producers based on defined criteria.

The general collection guidelines, selection criteria, evaluation criteria or criteria for heritage generation (cf. 1.1) and the general aims of long-term preservation should be specified at the object level.
Acquisition can be performed by submission of the objects to the DR by the issuing party or through manual or automatic collection on the part of the DR.

[DOMEA: Aussonderung und Archivierung elektronischer Akten, Erweiterungsmodul zum DOMEA-Organisationskonzept 2.0, 2005]

[The U.S. National Archives & Records Administration: Disposition of Federal Records. Transfer of Records to the National Archives of the United States, 2006]

[National Digital Archive of Datasets (NDAD): Transfer Procedures (Overview), 2005]

9.1 The digital repository specifies its transfer objects (Submission Information Packages, SIPs).

Either by agreement or the expression of explicit conditions the DR should communicate to producers or suppliers the types of digital objects (transfer objects)that it will accept. These agreements should allow the transfer or the collection to be automated, and for workflows for submission to the DR to be implemented.
These specifications are the basis for quality checking of the transfer objects.

Transfer objects may contain content data and also metadata, e.g. to establish their authenticity.

In the case of harvesting based on offline browsers, only text content, but not audio, video and other multimedia content is collected (through the selection or exclusion of specific file formats).

The file formats of the transfer objects can be validated using JHOVE (cf. http://hul.harvard.edu/jhove/) as a quality check.

The DR should recommend file formats for the transfer objects, e.g. GeoTif for remote reconnaissance data, or Seed/MiniSeed as the format for geodata, as used in GeoFon (http://www.gfz-potsdam.de/geofon/).

9.2 The digital repository identifies which characteristics of the digital objects are significant for information preservation.

In determining the scope of the characteristics to be preserved, a balance should be struck, between the technical possibilities and the costs of long-term preservation on the one hand and the needs of the designated community on the other hand.

It may be necessary to maintain the digital objects in a number of different forms to preserve an optimal number of characteristics.

Regarding information from databases it may be sufficient to archive the data as so-called "flat files" (including a precise description of the data structure).

With regard to electronic files, the individual documents should be saved as image files, following the DOMEA specifications. This excludes the possibility of full-text searches and the executability of some documents (Excel tables or PowerPoint presentations).

Regarding web pages containing text-image information, one archive can store only text information, one, only images, and the third the entire interrelation. The different objectives lead to correspondingly different archiving strategies.

Screenshots from a standard browser are taken of web pages, but the text information is also stored for ease of research.

[Kunze, John: Future-Proofing The Web: What We Can Do Today, 2005]

9.3 The digital repository has technical control of the digital objects in order to carry out long-term preservation measures.

Many digital objects contain technical features that restrict their use, either for commercial or legal reasons. For the long-term preservation of digital objects it is crucial that the digital repository is capable of opening and processing the objects with no restrictions. All technical restrictions on use must therefore be removed before submission to the DR.

Internal settings may prevent e.g. copying, printing or saving of objects; other objects are encrypted and require the input of codewords or cannot be opened after expiry of a certain period or after a specified number of sessions.

"Music and publishing industry agree duplication of copy-protected works with German National Library", joint press information released by Deutsche Nationalbibliothek, Börsenverein des Deutschen Buchhandels and Bundesverbands der Phonographischen Wirtschaft, dated 18 January 2005, http://www.ddb.de/aktuell/presse/pressemitt_vervielfaeltigung.htm

10 Archival storage of the digital objects is undertaken to defined specifications.

At the heart of a digital repository is an implementation of an archival process. This encompasses the definition of the archive objects, storage of the digital objects and implementation of the long-term preservation measures.

10.1 The digital repository defines its archival objects (Archival Information Packages, AIPs).

Archive objects consist of the content data in a suitable archive format and all the relevant metadata for long-term preservation. This should be stored within a defined structure.
Archive objects definitions should describe the object structures and archive formats used, and the metadata necessary for long-term preservation (cf. 12). Selection of the archive objects should depend on the object types (for example, digital script or 3D animated clip) and the characteristics of the objects to be preserved.

Open, disclosed and widely proliferated formats are preferred as archive formats, the assumptions being that these will have a longer life, and that there are likely to be more tools and techniques available to support their conversion or emulation, given that they are used by a wide circle of users.

KOPAL stores its objects in a universal document format: UOF, see examples:

http://kopal.langzeitarchivierung.de/downloads/kopal_UOF_DDB_mets.xml

http://kopal.langzeitarchivierung.de/downloads/kopal_UOF_SUB_mets.xml

Examples of currently used archive formats:

- for unformatted text: ASCII/Unicode
- for structured text: XML
- for formatted text: PDF/A
- for raster graphics: TIFF_6
- for audio formats: WAVE
- for video files: MPEG 4 File Format Version 2
- for executable programs: source text and documentation of the programming language

When making a decision regarding (lossless) compression of data, a balance should be struck between optimising storage on the one hand and subsequent dependency on the compression technologies on the other. Open or disclosed techniques (e.g. TIFF-LZW) are preferable to proprietary or strictly regulated technologies, which should be avoided on account of the associated licensing issues.

For the structural description of the archive objects, XML is currently favoured, especially the METS schema, which allows metadata and references to be managed within the individual files of an object.

[Coy,Prof.Dr.Wolfgang: nestor - materialien 5: Perspektiven der Langzeitarchivierung multimedialer Objekte, 2006]

[Witthaut,Dirk.Unter Mitarbeit von Andrea Zierer,Arno Dettmers,Stefan Rohde-Enslin: nestor - materialien 2: Digitalisierung und Erhalt von Digitalisaten in deutschen Museen, 2005]

[Erpanet : Erpanet-Tagung FileFormatsforPreservation, 2004]

[Abrams,Stephen: Digital Formats And Preservation, 2005]

[Library of Congress: Sustainability of Digital Formats: Planning for Library of Congress Collections, 2006]

[W3C: Extensible Markup Language (XML), 2004]

[ISO 19005-1. Document management - Electronic document file format for long-term preservation, 2006]

[Adobe: TIFF, Revision 6.0, 1992]

[Microsoft: Multimedia Data Standards Update, 1994]

[ISO/IEC 14496-14:2003. Information technology.Coding of audio-visual objects.MP4 File Format, 2003]

10.2 The digital repository takes care of transforming the transfer objects (SIPs) into archival objects (AIPs).

As part of the ingest process, the SIPs should be transferred into AIPs, with the addition of specific long-term preservation metadata. This might involve conversion of the format.

In KOPAL, a PDF document object is converted into a METS object with appropriate XMetaDiss and LMER metadata.

DOC files are converted into PDF/A files.

10.3 The digital repository guarantees the storage and readability of the AIPs.

The digital repository should use appropriate methods to ensure that the archival objects are correctly stored and can be read, using means available within the system. Readability here refers to the capacity to read the storage media and the appropriate bit sequence.

See 6.2 on ensuring the integrity of archival objects.

Possibilities for storing and ensuring readability include:

- Use of RAID systems
- Persistent storage on suitable media such as tapes, records, CDs, DVDs

10.4 The digital repository implements strategies for the long-term preservation of the AIPs.

The long-term preservation measures specified in point 8 should be implemented. A process should be defined to determine for each archive object whether a long-term preservation measure such as migration or the provision of emulation software - must be undertaken. If necessary, the corresponding measure should be carried out and documented (cf. 12.4).

For example, this strategy could involve ensuring that a decision is made in 2007 whether or not to migrate documents stored in PDF1.1 PDF/A format.

11 The digital repository permits usage of the digital objects based on defined criteria.

The usage purposes described under point 2 must be specified at the object level. The objects may be usable by individuals and/or client systems. The search and access possibilities regarding the usage objects should be defined. Each search should result in a clear response from the system. Usage objects (DIPs) are the information units which users receive as a response to inquiries to the DR.

11.1 The digital repository defines its usage objects (Dissemination Information Packages, DIPs).

The DR should define its DIPs within the context of both its designated community(s) and the archival objects (AIPs). A precondition for this is that the application environment for use has been determined.. An archival object may be offered as different usage objects, depending on the particular usage context. Use of the information represented by the digital objects in most cases does not mean access to the archival objects themselves, rather the use of copies or derivatives (possibly in combination with other information) which aid interpretability. This could be a technical description, additional application software or emulation software.

To exchange data with other digital repositories, or to migrate the data to a different technical infrastructure it is necessary to transform parts of, or the entire content of, the DR into a documented, standardised export format. The information can thus be preserved beyond the life of the DR itself (cf. 4.5).

```
Image archive: for use in the Web, low resolution files are generated
from the master images which can be displayed by today's browsers. High-
resolution files can be supplied electronically for reproduction
purposes.
cf. http://www.bsb-muenchen.de/karten/bilddatenb.htm
```

11.2 The digital repository ensures transformation of AIPs into DIPs.

The usage objects should be derived from the archival objects according to a defined procedure. Usage objects can be held in the digital repository and, in the event of changed conditions be regenerated, or can be created directly from the archival object when requested.

```
Information should be stored on the conversion process (conversion
software, date, participants etc.) for the conversion of high resolution
master images into low-resolution usage versions which can be displayed
by standard browsers.
```

12 The data management system is capable of providing the necessary digital repository functions.

Data management is an all-encompassingprocess which supports the core processes of a DR - ingest, archival storage and access - and also the planning and implementation of the preservation measures, while ensuring integrity and authenticity at all stages of processing. The scope of data management is dictated by the goals of the DR.
A number of aspects of data management are integral:

- identification of the digital objects and their relationships is essential for administration of the objects
- formal description of digital objects' content and structure is a precondition for their discovery
- ensuring interpretability and integrity, and planning and implementing preservation measures presumes technical

description of the objects
- documentation of all changes to digital objects is necessary to ensure authenticity of the data
- recording of all legal restrictions and their basis (laws, ordinances, contracts, agreements) is necessary to ensure that legal requirements are observed throughout the preservation process

The generation and storage of metadata currently fulfill these tasks. Metadata can be recorded in a structured manner in a metadata plan. Various metadata schema have become established for a range of purposes (e.g. descriptive, structural, technical, administrative, legal metadata) in a range of disciplines(e.g. archives, libraries, museums). Conformance to a national or international standard or deployment of a widespread metadata schema is often possible and beneficial in terms of data sustainability, and also for cooperation and data exchange between producers / suppliers, the DR and users. A metadata schema contains defined fields (data elements) within which the respective content can be recorded. The result is a data structure that can ought to be usable by both humans and machines.

The DR should establish rules for populating the fields (for example, the use of controlled terminology). Different tools permit the automatic generation or extraction of metadata , such as JHOVE for technical metadata.

In this criteria catalogue, metadata is treated as part of the logical information units: submission object, archival object and usage object. These can be managed for example in databases and/or XML structures.

[Bischoff,Frank M.: Metadata in preservation : selected papers from an ERPANET seminar at the Archives School Marburg, 2004]

[METS: Überblick und Anleitung, 2006]

[PREMIS Working Group: Preservation Metadata: Implementation Strategies, 2005]

[PREMIS Working Group: Data Dictionary for Preservation Metadata: Final Report of the PREMIS Working Group, 2005]

[LMER: 2006]

12.1 The digital repository uniquely and permanently identifies its objects and their relationships.

A DR should use internal identifiers to manage the objects and their parts and relationships (part/totality, different manifestations, versions for instance), especially for unique assignment of the content data to the metadata (cf. 12.7).

The use of externally visible, standardised persistent identifiers should ensure objects' reliable referencing and citability.

Transferred from the world of printed materials to electronic media are:

Signatures

- ISBN (International Standard Book Number) for monographs
- ISBN (International Standard Book Number) for periodicals
- ISBN (http://www.ietf.org/rfc/rfc3187.txt) and ISSN (http://www.ietf.org/rfc/rfc3044.txt) are registered as URN namespaces.

For electronic media other systems are used, e.g.:

- Uniform Resource Names (URN):
  An international Internet standard for unique, permanent identification of objects. In libraries National Bibliography Numbers (NBN) are used, a sub-namespace of the URNs, e.g.
  - URN: urn:nbn:de:0008-20050117016

URL: http://nbn-resolving.de/urn:nbn:de:0008-20050117016

- the handle system (HDL):
  a persistent identifier representing a sub-area of URNs, e.g.
  - HDL: 1721.1/30592
    URL: http://hdl.handle.net/1721.1/30592
- Digital Object Identifier (DOI)
  DOIs are used by publishers but also increasingly for specialist and primary data. Their technical basis is the handle system, e.g.
  - DOI: 10.1045/april2004-dobratz
    URL: http://dx.doi.org/10.1045/april2004-dobratz
- SRef – the scientific reference linking system (http://www.sref.org/site/index.php)

Use of a resolving service allows persistent identifiers to be incorporated in the URL address, thereby ensuring permanent access. This requires continuous data maintenance by the resolving service to which a DR commits itself.

URN-Service Der Deutschen Bibliothek: http://www.persistent-identifier.de/

Allgemeine Anforderungen an URNs: ftp://ftp.rfc-editor.org/in-notes/rfc1737.txt

Registrierung von URN-Unternamesräumen: IANA Registry, http://www.iana.org/assignments/urn-namespaces

Digital Object Identifier homepage: http://www.doi.org/

Handle-System homepage: http://www.handle.net/

URI: http://info-uri.info/

PURL: http://purl.oclc.org/

ARK: http://www.ietf.org/internet-drafts/draft-kunze-ark-10.txt

PADI – Preserving Access to Digital Information, Topic: Persistent Identifiers: URL: http://www.nla.gov.au/padi/topics/36.html

Nestor-Informationsdatenbank, Themenschwerpunkt: Persistente Identifikatoren: URL: http://nestor.sub.uni-goettingen.de/nestor_on/browse.php?show=21

ERPANET Workshop „Persistent Identifier", 2004:

URL: http://www.erpanet.org/events/2004/cork/index.php

12.2 The digital repository acquires adequate metadata for formal and content-based description and identification of the digital objects.

The scope, structure and content of the descriptive metadata should depend on the goals of the DR, its designated community and the object types. Formal and content-based description of the objects in the form of metadata makes it possible to find objects; this is essential in terms of the research options which are offered to users.

A number of different schemata have become established in the different fields:

Libraries: Dublin Core (DC); MAB and MARC, Metadata Objects Description Schema (MODS). Library codes can be used in combination with this, e.g. RAK or AACR2 for formal identification, and RSWK or a classification (e.g. DDC, RVK) for content identification.

Archives: General International Standard Archival Description (ISAD(G)), Encoded Archival Description (EAD), supplemented by Encoded Archival Context (EAC).

For space-related data: ISO Standard 19115.

NASA DIF (Data Interchange Format, see http://gcmd.nasa.gov/User/difguide/difman.html) as NASA descriptive format which has developed into a de-facto standard, and which is also used for the Global Change Master Directory (http://gcmd.nasa.gov/).

[DC: The Dublin Core Metadata Element Set, ISO 15836, http://dublincore.org/]

[MAB: http://www.ddb.de/standardisierung/formate/mab.htm]

[MARC: http://www.loc.gov/marc/]

[MODS: http://www.loc.gov/standards/mods/]

[EAD: http://www.loc.gov/ead/]

[EAC: http://jefferson.village.virginia.edu/eac/]

[ISAD(G): http://www.icacds.org.uk/eng/ISAD(G).pdf]

[Shepherd, Elizabeth and Smith, Charlotte: The Application of ISAD(G) to the Description of Archival Datasets, 2000]

[Domea: DOMEA-Konzept]

[Generaldirektion der Staatlichen Archive Bayerns (Hrsg.): Metadaten für die Aussonderung und Archivierung digitaler Sachakten, 2004]

12.3 The digital repository acquires adequate metadata for structural description of the digital objects.

The structure of complex objects must be adequately described so that they can be reconstructed and subsequently used as complete entities.

METS is appropriate for representing the structures of digital objects; however structure information can also be managed in the descriptive metadata and in the metadata for long-term preservation (e.g. PREMIS and LMER).

A digital record generally consists of procedures that in turn consist of documents to which further documents (appendices) may belong. This hierarchy is described by a file containing metadata about each level, (at the document level) metadata and references to the documents themselves (primary information).

The digitised version of a conventional book consists of 200 individual image files. The metadata should list the correct order of book pages and the corresponding image files.

An archived website consists of a number of HTML pages and JPEG image files which are bound to each other via links. These links should be recorded in the metadata.

12.4 The digital repository acquires adequate metadata to record the changes made by the digital repository to the digital objects.

The DR should document all changes made to the digital objects. This also includes recording the people,systems and rights involved (cf. 3.2). This documents authenticity (cf. 7) and also ensures technical preservation of the digital objects.
A side effect of migration as a preservation strategy is that digital objects exposed to process are changed to varying degrees. Additional changes may follow during the transformations that are carried out during submission to the DR and for delivery of usage objects.

This metadata (history, audit trail, provenance) can be managed: e.g. by METS (amdSec digiorivMD section), PREMIS (Events section), LMER (Processes section).

An archive should migrate objects stored in an obsolete data format to a current format using a conversion program. Metadata on the migration procedure, the technical protocol, the time of migration, the factors involved (staff and technical aids) and the result of the action should

be recorded and saved.

12.5 The digital repository acquires adequate metadata for technical description of the digital objects.

To ensure interpretability and integrity and to control the preservation measures, the objects and their associated files must be comprehensively described in technical terms.

The technical description should contain general information which can be used for all file formats, including:

- File name, storage location
- File size, different check sums
- Full description of file formats
- Hardware/software environment used for generation
- Hardware/software environment required for use
- Recording of all necessary additional objects (DTD, schema file, fonts etc.)

Also there should be specific information for the individual formats, e.g. resolution, colour space, compression etc. for TIFF files.

The general technical metadata is also managed by METS (amdSec, techMD sections), PREMIS or LMER.

Other standards have become established for format-specific metadata:

- Metadata for Images in XML schema (MIX) for images, based on NISO Technical Metadata for Digital Still Images, http://www.niso.org/committees/committee_au.html
- For text as extension of METS schema: textmd.xsd

File format registers can be referenced to describe file formats, e.g.: Global Digital Format Registry: http://hul.harvard.edu/gdfr/, PRONOM: http://www.nationalarchives.gov.uk/pronom/.

Tools are available for automatic extraction of the technical metadata, e.g. JHOVE.

Examples:

A DR stores files in version 1.4 PDF files. "Acrobat Reader 5.0" is required to view the files. This program runs on a Microsoft operating system - Windows 98 SE or later. The entire software, however, requires a PC with a processor of at least 350 MHz and 64 MB main memory. These technical details are part of the metadata that is recorded and stored by the DR.

A DR stores files in A-1 PDF files. This format is described in full in ISO standard 19005-1:2005. The DR appends the relevant ISO standard to the metadata or refers to it via a reference within the metadata.

A DR stores files in XML format. The relevant schema files are required to assess the validity of these files. The DR appends the relevant schema files to the metadata or refers to them via a reference within the metadata.

[Steinke,Tobias: Universelles Objektformat: Ein Archiv- und Austauschformat für digitale Objekte, 2006]

[National Library of New Zealand: Metadata Standards Framework – Preservation Metadata, 2002]

[JHOVE: Harvard Object Validation Environment, 2006]

[MIX, 2006]

[Textmd.xsd, 2006]

12.6 The digital repository acquires adequate metadata to record the corresponding usage rights and conditions.

Use of the digital objects may be restricted for legal or contractual reasons. These rights and conditions must be recorded in such as way as to facilitate use controls (such as controlled access and anonymising of user copies) and user feedback (cf. 3.3), to an extent determined by the nature of the conditions and the affected user groups.

A DR archives databases that are only released for use after a period of 60 years with an exception for scientific research use. The usage restriction is part of the DR's metadata and reference is also made to the relevant legal paragraphs (in this case Art. 2 paragraph 4 clause 2 and Article 5 paragraph 3 of the Federal Archive Act and Article 16 paragraphs 6-9 of the Federal Statistics Act).

This can be implemented for example in the METS RightsDeclationMD Extension Schema.

For issuing of author-based rights with markup see, for example, Creative Commons (http://www.creativecommons.org) or DPPL (Digital Peer Publishing License), (http://www.dipp.nrw.de/lizenzen)

[METS rights, 2006]

12.7 The assignment of metadata to the digital objects is guaranteed at all times.

The relationship between the metadata and the digital objects, both as a whole and individually (especially the content data), must be reliable and unambiguous. This can be achieved by:
a) Use of internally applied yet externally visible persistent identifiers for the digital objects and their constituent parts, especially content data and metadata (cf. 12.1)
b) Implementation of a defined object structure (SIP, AIP, DIP) and storage in the same location (encapsulation) to accommodate all object content and metadata .

The Open Archival Information System (OAIS) recommends storage of the necessary metadata together with the content data in an Archival Information Package (AIP).

The metadata schema METS offers the possibility of embedding a digital object which has been converted into a sequence of ASCII characters by means of a base 64 converter into the metadata.

C. Infrastructure and Security

Infrastructure and Security looks at the technical aspects of the overall system and aspects of IT security.

13 The IT infrastructure is adequate.

The IT infrastructure should implement all of the technical and security specifications for handling digital objects. This infrastructure is responsible for the full extent of all the objects within the repository.

13.1 The IT infrastructure implements the object management demands.

The digital object handling requirements specified by the DR should be implemented by the overall system at every stage of processing. This includes the main processes (in OAIS: "functional entities") of Ingest, Archival Storage and Access and the data management supporting process. Extension of these functions may become necessary as a consequence of the evolution of the DR's

goals.

Web-Ingest-Module, module for bulk ingest in batch operation

Storage module with access to a different, geographically separate storage system.

Usage module

If the DR policy includes registered users being able to feed their photo collections themselves into the DR, assuming these are available as JPEG files, the DR must then provide a suitable upload interface for users.

[Borghoff,Uwe M.u.Mitarb.Univ.d.Bundeswehr München,Fak.f.Informatik,Inst.f.Softwaretechnologie: nestor - materialien 3: Vergleich bestehender Archivierungssysteme, 2005]

13.2 The IT infrastructure implements the security demands of the IT security system.

Realisation should take the object management security requirements into consideration:
Ensuring the **integrity** of the objects (including metadata), i.e. protecting them from modifications arising from deliberate and unintentional human actions, and technical imperfection
Ensuring the **authenticity** of the objects (including metadata)
Ensuring the **confidentiality** of the objects (including metadata),that is, excluding the possibility of unauthorised access to information
Ensuring the **availability** of the objects (including metadata) through availability of the object management functions (including protection against sabotage and system failures for example)

Access to protected data (such archived STASI investigation committee documents) must be restricted to authorised users by means of appropriate technical security precautions (e.g. passwords or biometric access barriers).

The use of approved digital signatures as defined in the Digital Signature Act, and time stamps for the preservation of patent applications.

14 The infrastructure protects the digital repository and its digital objects.

The infrastructure should protect the digital objects from system-based and external hazards. System-based hazards may arise due to hardware problems or the failure of individual storage media for example. Externally, the DR's first priority must be to protect against natural threats (e.g. fire, water, seismic activity), and also against risks posed by humans. The objects may be harmed by direct employee interactions or through harmful programs that compromise the system (e.g. viruses). Protection of data also implies the prevention of unintentional forwarding of information by programs (trojans) or people (espionage).

In addition to the archive objects themselves, the DR must protect its facilities, its associated hardware and software, and, not least, its staff.

The various risks must be countered with a package of technical (such as virus protection programs) and organisational (such as access restrictions) measures.

A fire which breaks out in the main building of the institution housing the DR should not result in damage to the objects or data loss, as there should be a suitable backup system at a separate location which can assume operations in the event of an accident.

[IT-Grundschutzhandbuch, 2006]

III. Checklist

It is not possible to make an absolute evaluation of the measures for fulfilling the criteria. The evaluation is always based on the goals of the digital repository; however the adequacy of the measures should be checked.

Besides implementation of the criteria, publication of appropriate documentation helps increase the transparency of the archive, and helps to generate confidence in it. The checklist is therefore presented as a table including the 4 phases of completion and also publication.

		conceptual groundwork	planned /specified	implemented	evaluated	published
A	**Organisational framework**					
1	The DR has defined its goals.					
1.1	The DR has developed criteria for the selection of its digital objects.					
1.2	The DR assumes responsibility for long-term preservation of the information represented by the digital objects.					
1.3	The DR has defined its designated community/-ies.					
2	The DR grants its designated community adequate usage of the information represented by the digital objects.					
2.1	The DR grants its designated community access to the information represented by the digital objects.					
2.2	The DR ensures that the designated community can interpret the digital objects.					
3	Legal and contractual rules are observed.					
3.1	Legal contracts exist between producers and the digital repository.					
3.2	In carrying out its archiving tasks, the DR acts on the basis of legal rulings.					
3.3	With regard to use, the DR acts on the basis of legal requirements.					
4	The organisational form is adequate for the DR.					
4.1	Adequate financing of the digital repository is secured.					
4.2	Sufficient numbers of appropriately qualified staff are available					
4.3	Appropriate organisational structures exist for the DR.					

4.4	The DR engages in long-term planning.					
4.5	Continuation of the preservation tasks is ensured even after the existence of the digital repository.					
5	Adequate quality management is conducted					
5.1	All processes and responsibilities have been defined.					
5.2	The DR documents all its elements based on a defined process.					
5.3	The DR reacts to substantial changes					
B	**Object management**					
6	The DR ensures the integrity of the digital objects during all processing stages.					
6.1	Ingest: the DR ensures the integrity of the digital objects.					
6.2	Archival Storage: the DR ensures the integrity of the digital objects.					
6.3	Access: the DR ensures the integrity of the digital objects.					
7	The DR ensures the authenticity of the digital objects and metadata during all processing stages.					
7.1	Ingest: the DR ensures the authenticity of the digital objects.					
7.2	Archival Storage: the DR ensures the authenticity of the digital objects.					
7.3	Access: the DR ensures the authenticity of the digital objects.					
8	The DR has a strategic plan for its technical preservation measures.					
9	The DR accepts digital objects from the producers based on defined criteria.					
9.1	The DR specifies its transfer objects (Submission Information Packages, SIPs).					
9.2	The DR identifies which characteristics of the digital objects are significant for information preservation.					

9.3	The DR has physical control of the digital objects in order to carry out long-term preservation measures.					
10	Archival storage of the digital objects is undertaken to defined specifications.					
10.1	The DR defines its archival objects (Archival Information Packages, AIPs).					
10.2	The DR takes care of transforming the transfer objects (SIPs) into archival objects (AIPs).					
10.3	The DR guarantees the storage and readability of the AIPs.					
10.4	The DR implements strategies for the long-term preservation of each AIP.					
11	The DR permits usage of the digital objects based on defined criteria.					
11.1	The DR defines its usage objects (Dissemination Information Packages, DIPs).					
11.2	The DR ensures transformation of AIPs into DIPs.					
12	The data management system is capable of providing the necessary digital repository functions.					
12.1	The DR uniquely and permanently identifies its objects and their relationships.					
12.2	The DR acquires adequate metadata for formal and content-based description and identification of the digital objects.					
12.3	The DR acquires adequate metadata for structural description of the digital objects.					
12.4	The DR acquires adequate metadata to record the changes made by the digital repository to the digital objects.					
12.5	The DR acquires adequate metadata for technical description of the digital objects.					

12.6	The DR acquires adequate metadata to record the corresponding usage rights and conditions.					
12.7	The assignment of metadata to the objects is guaranteed at all times.					
C.	**Infrastructure and Security**					
13	The IT infrastructure is adequate.					
13.1	The IT infrastructure implements the object management demands.					
13.2	The IT infrastructure implements the security demands of the IT security system.					
14	The infrastructure protects the digital repository and its digital objects.					

IV. Glossary and abbreviations

Archival Storage: An OAIS functional entity consisting of the functions and processes ensuring the storage and availability of the archival objects.

Archival object (Archival Information Package, AIP): An information unit stored in the DR, consisting of content data and metadata required for long-term preservation.

Ingest: An OAIS functional entity consisting of the functions and processes which receive the transfer objects from the producer/supplier, transform them into archival objects and incorporate them into the archive.

Authenticity: The object actually contains what it claims to contain.

Data: Formalised representation of information that enables it to be interpreted, processed and exchanged.

Digital repository (DR): An organisation (consisting of people and technical systems) that has assumed responsibility for the long-term preservation and availability of digital data and its provision for a specified designated community. "Long-term" here means lasting beyond technological changes (to hardware and software) and also any changes to the designated community (e.g. for future generations, indefinitely).

Digital object: Logical discrete unit of digital data. This could be a simple object consisting of a single file (e.g. a PDF document) or a complex object consisting of a number of different files (e.g. an electronic journal consisting of individual articles saved as files). Further data (metadata) may be added to the information that represents content (content data) in order to detail the formal and content description, the structural description, the preservation processes undertaken, or the means by which content should be interpreted (cf. transfer object, archival object, usage object).

Integrity: 1. The completeness of the digital objects or, 2. Exclusion of modifications that are prohibited within the preservation rules. Integrity is measured in terms of the characteristics of a digital object being preserved.

Preservation planning: A collective term describing the methods specifically used to archive digital objects indefinitely and to make them available for a sustained period. This includes methods for the physical preservation of the data and also the use of migration and emulation techniques to change the archived objects or their environments to guarantee their future use.

Metadata: Data representing information about other data. This may describedata's content, structure, composition, handling, or origin.
Metadata can be created at various times throughout the lifecycle of digital objects (during production, archiving or provision for use etc.).
The term is used primarily in the digital field (e.g. Dublin Core Metadata), although, for example, title listings in library catalogues, or archive catalogue entries may also be regarded as metadata. Metadata should be regarded as parts of the logical unit of "digital object".

Users: People or client systems that interact with the DR to discover and use the information represented by the digital objects.

Access: An OAIS functional entity consisting of the functions and processes that make the archived information accessible to the users.

Usage object (Dissemination Information Package, DIP): An information unit derived from one or more AIPs that a user receives from the DR in response to an inquiry. A usage object consists of the data representing the content and, where applicable, the information needed for interpretation (e.g. a csv format file and description of the data structure; a DOS program in source code and emulation software for the DOS operating system).

OAIS: Reference model (ISO 14721:2003) for DRs, which describes the core, processes of a DR (in terms of functional entities) and provides an information model.

Producer: People or client systems that transfer digital objects to the DR for long-term preservation. They are not necessarily the originators; they could also be the suppliers of the digital objects.

Quality: The quality of a DR is the extent to which its inherent characterising properties fulfil the specified requirements. Requirements are prerequisites or expectations which are expressed, and which are generally taken for granted or compulsory (following ISO 9000:2000).

Representation information: Information that is necessary to interpret digital data (for example, the file format of a file).

Transfer object (Submission Information Package, SIP): An information unit submitted by the producer to the DR. The content data may already be supplemented with metadata.

Availability: The extent to which data is available to the user at the required time.

Confidentiality: The extent to which unauthorised divulgence of the data is tolerable.

Trustworthiness: Trustworthiness is the capacity of a system to operate in accordance with its objectives and specifications (that is, it does exactly what it claims to do). The trustworthiness of a DR can be tested and assessed on the basis of a criteria catalogue.

Designated community/Target group: An identifiable group of potential users with specific interests and circumstances. It could be the general public or a group of specialist scientists, for instance. It can be heterogeneous and consist of different user groups.

V. Bibliography

[1] (2001): ISO 15489-1, Information and documentation, Records Management. URL: http://www.iso.org

[2] (2002): Rahmenvereinbarung zur freiwilligen Ablieferung von Netzpublikationen zum Zwecke der Verzeichnung und Archivierung. URL: http://deposit.ddb.de/netzpub/web_rahmenvereinbarung.htm

[3] (2005): ISO/IEC 15408-x:2005 Information technology -- Security techniques -- Evaluation criteria for IT security Part 1 - Part 3. URL: http://isotc.iso.org/livelink/livelink/fetch/2000/2489/Ittf_Home/PubliclyAvailableStandards.htm

[4] (2005): ISO 9000:2005 Quality management systems -- Fundamentals and vocabulary. URL: http://www.bsi-global.com/Quality_management/Management/bseniso9000.xalter

[5] (2006): LMER. URL: http://www.ddb.de/standards/lmer/lmer.htm

[6] (2006): IT-Grundschutzhandbuch. URL: http://www.bsi.de/gshb/

[7] (2006): MAB. URL: http://www.ddb.de/standardisierung/formate/mab.htm

[8] (2006): METS rights. URL: http://www.loc.gov/standards/rights/METSRights.xsd

[9] (2006): MARC. URL: http://www.loc.gov/marc/

[10] (2006): EAD. URL: http://www.loc.gov/ead/

[11] (2006): ISO 19005-1. Document management - Electronic document file format for long-term preservation - Part 1, Use of PDF (PDF/A).

[12] (2006): ISAD(G). URL: http://www.icacds.org.uk/eng/ISAD(G).pdf

[13] (2006): Handle System. URL: http://www.handle.net/

[14] (2006): ISO/IEC 14496-14:2003. Information technology - Coding of audio-visual objects - Part 14: MP4 File Format.

[15] (2006): GeoFon. URL: http://www.gfz-potsdam.de/geofon/

[16] (2006): METS: Überblick und Anleitung. URL: http://www.loc.gov/standards/mets/METSOverview.v2_de.html

[17] (2006): SREF. URL: http://www.sref.org/site/index.php

[18] (2006): Registrierung von Unternamesräumen: IANA Registry. URL: http://www.iana.org/assignments/urn-namespaces

[19] (2006): Qualitätsmanagement DIN EN ISO 9000ff.

[20] (2006): TEI (Text Encoding Initiative). URL: http://www.tei-c.org/

[21] (2006): URI. URL: http://info-uri.info/

[22] (2006): The Dublin Core Metadata Element Set, ISO 15836. URL: http://dublincore.org/documents/dces/

[23] (2006): Textmd.xsd. URL: http://dlib.nyu.edu/METS/textmd.xsd

[24] (2006): NBNs: Homepage des Projektes EPICUR. URL: http://www.persistent-identifier.de/

[25] (2006): MODS. URL: http://www.loc.gov/standards/mods/

[26] (2006): MIX. URL: http://www.loc.gov/standards/mix/

[27] (2006): NESLi2 licence journals. URL: http://www.nesli2.ac.uk/NESLi2_licence_journals_final011003.htm

[28] (2006): PURL. URL: http://purl.oclc.org/

[29] (2006): PADI – Preserving Access to Digital Information, Topic: Persistent Identifiers. URL: http://www.nla.gov.au/padi/topics/36.html

[30] (2006): Nestor-Informationsdatenbank, Themenschwerpunkt: Persistente Identifikatoren. URL: http://nestor.sub.uni-goettingen.de/nestor_on/browse.php?show=21

[31] (2006): Aussonderung und Archivierung elektronischer Akten, Erweiterungsmodul zum DOMEA-Organisationskonzept 2.0. URL: http://www.kbst.bund.de/cln_011/nn_836802/SharedDocs/Anlagen-kbst/Domea/erweiterungsmodul-aussonderung-und-archivierung-elektronischer-akten--pdf,templateId=raw,property=publicationFile.pdf/erweiterungsmodul-aussonderung-und-archivierung-elektronischer-akten--pdf.pdf

[32] (2006): ARK. URL: http://www.ietf.org/internet-drafts/draft-kunze-ark-10.txt

[33] (2006): ArchiSafe Projekt. URL: http://www.archisafe.de/s/archisafe/index

[34] (2006): Digital Object Identifier (DOI). URL: http://www.doi.org/

[35] (2006): Common Criteria. URL: http://www.commoncriteriaportal.org/

[36] (2006): Baden-Württembergisches Online-Archiv (BOA). URL: http://www.boa-bw.de/

[37] (2006): EAC. URL: http://jefferson.village.virginia.edu/eac/

[38] (2006): Allgemeine Anforderungen an URN. URL: ftp://ftp.rfc-editor.org/in-notes/rfc1737.txt

[39] Abrams, Stephen (2005): Digital Formats and Preservation, IPRES. URL: http://rdd.sub.uni-goettingen.de/conferences/ipres05/download/Digital%20Formats%20And%20Preservation%20-%20Stephen%20Abrams.pdf

[40] Adobe (1992): TIFF, Revision 6.0, Final. URL: http://partners.adobe.com/public/developer/en/tiff/TIFF6.pdf

[41] Baker, Thomas (2000): A Grammar of Dublin Core, Volume 6 Number 10, D-Lib Magazine.

[42] Bischoff, Frank M. (2004): Metadata in preservation : selected papers from an ERPANET seminar at the Archives School Marburg, Archivschule , Marburg, ISBN: 3-923833-77-6.

[43] Borghoff, Uwe M. u. Mitarb. Univ. d. Bundeswehr München Fak. f. Informatik Inst. f. Softwaretechnologie (2005): nestor - materialien 3: Vergleich bestehender Archivierungssysteme, graph. Darst., nestor c/o Die Deutsche Bibliothek, Frankfurt am Main.

[44] Borghoff; Uwe M.und Peter Rödig; Jan Scheffczyk und Lothar Schmitz (2003): Langzeitarchivierung - Methoden zur Erhaltung digitaler Dokumente, dpunkt.verlag, Heidelberg, ISBN: 3-89864-245-3.

[45] Bundesamt für Sicherheit in der Informationstechnik (2004): Leitfaden IT-Sicherheit IT-Grundschutz kompakt.

[46] Bundesrat (2005): Entwurf eines Gesetzes über die Deutsche Nationalbibliothek. URL: https://www.umwelt-online.de/PDFBR/2005/0396_2D05.pdf

[47] Caplan, Priscilla and Guenther Rebecca (2005): Practical Preservation: The PREMIS Experience In: Library Trends, Vol. 54, No. 1, ("Digital Preservation: Finding Balance,"edited by Deborah Woodyard-Robinson), S. 111–124. URL: http://www.loc.gov/standards/premis/caplan_guenther-librarytrends.pdf

[48] CCSDS (2004): Producer-Archive Interface Methodology - Abstract Standard, Blue Book. URL: http://public.ccsds.org/publications/archive/651x0b1.pdf

[49] CCSDS (Consultative Committee for Space Data Systems (2002): Reference Model for an Open Archival Information System (OAIS). Blue Book. URL: http://www.ccsds.org/docu/dscgi/ds.py/Get/File-143/650x0b1.pdf

[50] Coy, Prof. Dr. Wolfgang Humboldt-Universität zu Berlin Institut für Informatik (2006): nestor - materialien 5: Perspektiven der Langzeitarchivierung multimedialer Objekte, nestor c/o Die Deutsche Bibliothek, Frankfurt am Main.

[51] Deutsche Initiative für Netzwerkinformation (DINI) AG Elektronisches Publizieren (2003): DINI-Zertifikat für Dokumenten- und Publikationsserver / Engl. Version, DINI-Certificate Document and Publication Repositories. URL: urn:nbn:de:kobv:11-10046073

[52] Die Deutsche Bibliothek, Börsenverein des Deutschen Buchhandels Bundesverband der Phonographischen Wirtschaft (2005): Gemeinsame Presseinformation über Vervielfältigung kopiergeschützter Werke. URL: http://www.ddb.de/aktuell/presse/pressemitt_vervielfaeltigung.htm

[53] DigiCULT (2006): Technology Watch Reports. URL: http://www.digicult.info/pages/techwatch.php

[54] Digitaleduurzaamheid (2006): Kostenmodell für die Langzeitarchivierung siehe: Vers van de pers..... 'Kostenmodel digitale bewaring'. URL: http://www.digitaleduurzaamheid.nl/detail.cfm?id=106&sub=nieuws&categorie=0

[55] Dobratz, S. and Schoger A. (2005): Digital Repository Certification: A Report from Germany, Vol.9, No.5, RLG DigiNews.

[56] Dokumenten- und Publikationsserver der Humboldt-Universität zu Berlin (2006): „Ziele und inhaltliche Kriterien". URL: http://edoc.hu-berlin.de/e_info/leitlinien.php

[57] DOMEA (2005): DOMEA-Konzept. URL: http://www.kbst.bund.de/cln_006/nn_836960/Content/Standards/Domea__Konzept/domea__node.html__nnn=true

[58] Erpanet (2003): Erpanet-Tagung "Appraisal of Scientific Data". URL: http://www.erpanet.org/events/2003/lisbon/index.php

[59] Erpanet (2003): Erpanet-Tagung "Policies for Digital Preservation". URL: http://www.erpanet.org/events/2003/paris/index.php

[60] Erpanet (2004): Erpanet-Tagung „File Formats for Preservation". URL: http://www.erpanet.org/events/2004/vienna/index.php

[61] Erpanet (2004): Erpanet-Tagung „Audit and Certification in Digital Preservation". URL: http://www.erpanet.org/events/2004/antwerpen/index.php

[62] Erpanet (2004): Erpanet-Tagung "Business Models related to Digital Preservation". URL: http://www.erpanet.org/events/2004/amsterdam/index.php

[63] Erpanet (2004): Erpanet Workshop „Persistent Identifier". URL: http://www.erpanet.org/events/2004/cork/index.php

[64] Erpanet (2004): Workshop on Workflow. URL: http://www.erpanet.org/events/2004/budapest/index.php

[65] Erpanet, Frank M. Bischoff (2004): Selected papers from an ERPANET seminar at the Archives School Marburg: Metadata in preservation, Archivschule Marburg - Institut für Archivwissenschaft, Marburg, ISBN: 3-923833-77-6.

[66] Generaldirektion der Staatlichen Archive Bayerns (Hrsg.) (2004): Metadaten für die Aussonderung und Archivierung digitaler Sachakten, München. URL: http://www.gda.bayern.de/digpub.htm

[67] Harvard College (2006): Global Digital Format Registry. URL: http://hul.harvard.edu/gdfr/

[68] Internet Archive (2006). URL: http://www.archive.org/about/about.php

[69] Interpares (2006): Ergebnisse des InterPares-Projekts. URL: http://www.interpares.org/reports.htm

[70] Interpares Appraisal Task Force (2006): Appraisal of Electronic Records: A Review of the Literature in English. URL: http://www.interpares.org/documents/interpares_ERAppraisalLiteratureReview.pdf

[71] JHOVE (2006): JSTOR/Harvard Object Validation Environment. URL: http://hul.harvard.edu/jhove/

[72] JHOVE - JSTOR (2005). URL: http://hul.harvard.edu/jhove/

[73] Jones, M. and Beagrie N. (2002): Preservation Management of Digital Materials: A Handbook, The British Library, London.

[74] Kneuper, Ralf (2006): Verbesserung von Softwareprozessen mit Capability Maturity Model Integration, dpunkt.verlag.

[75] KOPAL (2006): Universal Object Format - An archiving and exchange format for digital objects (English Version). URL: http://kopal.langzeitarchivierung.de/downloads/kopal_Universal_Object_Format.pdf http://kopal.langzeitarchivierung.de/downloads/kopal_UOF_DDB_mets.xml http://kopal.langzeitarchivierung.de/downloads/kopal_UOF_SUB_mets.xml

[76] KOPAL (2006): Universelles Objektformat - Ein Archiv- und Austauschformat für digitale Objekte (deutsche Version). URL: http://kopal.langzeitarchivierung.de/downloads/kopal_Universelles_Objektformat.pdf http://kopal.langzeitarchivierung.de/downloads/kopal_UOF_DDB_mets.xml http://kopal.langzeitarchivierung.de/downloads/kopal_UOF_SUB_mets.xml

[77] Kunze, John (2005): Future-Proofing The Web: What We Can Do Today, IPRES. URL: http://rdd.sub.uni-goettingen.de/conferences/ipres05/download/Future-Proofing%20The%20Web%20What%20We%20Can%20Do%20Today%20-%20John%20Kunze.pdf

[78] Library of Congress (2006): Sustainability of Digital Formats: Planning for Library of Congress Collections. URL: http://www.digitalpreservation.gov/formats/index.shtml

[79] Liggesmeyer, Peter (2002): Softwarequalität, Spektrum Akademischer Verlag.

[80] Littman, Justin (2006): A Technical Approach and Distributed Model for Validation of Digital Objects, Volume 12 Number 5, D-Lib-Magazin. URL: http://www.dlib.org/dlib/may06/littman/05littman.html

[81] Microsoft (1994): Multimedia Data Standards Update. URL: http://www-mmsp.ece.mcgill.ca/Documents/AudioFormats/WAVE/Docs/RIFFNEW.pdf

[82] National Archives (2006): Custodial policy for digital records. URL: http://www.nationalarchives.gov.uk/recordsmanagement/custody/

[83] National Archives (2006): PRONOM, Kew. URL: http://www.nationalarchives.gov.uk/pronom/

[84] National Digital Archive of Datasets (NDAD) (2005): Transfer Procedures (Overview). URL: http://www.ndad.nationalarchives.gov.uk/resources/pdf/xfer_notes_overview.pdf

[85] National Library of New Zealand (2002): Metadata Standards Framework – Preservation Metadata. URL: http://www.natlib.govt.nz/files/4initiatives_metaschema.pdf

[86] Network Working Group, Shirey R. (2000): Internet Security Glossary.

[87] Oltmans, Erik and Kol Nanda (2006): A Comparison Between Migration and Emulation in Terms of Costs. Volume 9, Number , In RLG DigiNews. URL: http://www.rlg.org/en/page.php?Page_ID=20571#article0

[88] Oxford Digital Library (2006): „Background" (u. a. „scope and objectives"), „Services", „Principles and Guidelines" etc. URL: http://www.odl.ox.ac.uk/about.htm

[89] Oxford Text Archive (2006). URL: http://ota.ahds.ac.uk/

[90] Österreichische Nationalbibliothek (2003): Archivierungs- und Nutzungsgenehmigungen. URL: http://www.onb.ac.at/about/lza/

[91] Palm, Jonas (2006): The Digital Black Hole. URL: http://www.tape-online.net/docs/Palm_Black_Hole.pdf

[92] PANDORA (2006): „The purpose of the PANDORA Archive", „Collecting responsibility" etc. URL: http://pandora.nla.gov.au/overview.html

[93] PREMIS Working Group (2005): Data Dictionary for Preservation Metadata: Final Report of the PREMIS Working Group (with a glossary). URL: http://www.oclc.org/research/projects/pmwg/premis-final.pdf

[94] PREMIS Working Group (2005): PREMIS (Preservation Metadata: Implementation Strategies). URL: http://www.oclc.org/research/projects/pmwg/

[95] Rauch, Carl und Rauber Andreas (2006): Anwendung der Nutzwertanalyse zur Bewertung von Strategien zur langfristigen Erhaltung digitale Objekte. URL: http://www.ifs.tuwien.ac.at/~andi/publications/pdf/rau_zfbb05.pdf

[96] Rechtsanwälte Goebel und Scheller (Bad Homburg v.d.H.) (2004): nestor - materialien 1: Digitale Langzeitarchivierung und Recht, nestor c/o Die Deutsche Bibliothek, Frankfurt am Main.

[97] RLG NARA Task Force on Digital Repository Certification (2005): Audit Checklist for Certifying Digital Repositories, RLG, Mountain View, CA.

[98] RLG Working Group on Digital Archive Attributes (2002): Trusted Digital Repositories: Attributes and Responsibilities. An RLG-OCLC Report, RLG, Mountain View CA, California. URL: http://www.rlg.org/longterm/repositories.pdf

[99] Shepherd, Elizabeth and Smith Charlotte (2000): The Application of ISAD(G) to the Description of Archival Datasets, Journal for the Society of Archivists.

[100] Steinke, Tobias (2006): Universelles Objektformat: Ein Archiv- und Austauschformat für digitale Objekte, Frankfurt am Main. URL: http://kopal.langzeitarchivierung.de/downloads/kopal_Universelles_Objektformat.pdf

[101] The Dublin Core (2006): DC: The Dublin Core Metadata Element Set, ISO 15836. URL: http://dublincore.org/

[102] The U.S.National Archives & Records Administration (2006): Disposition of Federal Records. Subpart L -- Transfer of Records to the National Archives of the United States, Part 1228, § 1228.270, Electronic records. URL: http://www.archives.gov/about/regulations/part-1228/l.html?template=print

[103] W3C (2006): Extensible Markup Language (XML) 1.1, Recommendation 04 February 2004. URL: http://www.w3.org/TR/xml11/

[104] Wiesenmüller, Heidrun et al. (2004): Auswahlkriterien für das Sammeln von Netzpublikationen im Rahmen des elektronischen Pflichtexemplars : Empfehlungen der Arbeitsgemeinschaft der Regionalbibliotheken, S. 1423-1444, Bibliotheksdienst 38.

[105] Witthaut, Dirk. Unter Mitarbeit von Andrea Zierer Arno Dettmers Stefan Rohde-Enslin (2005): nestor - materialien 2: Digitalisierung und Erhalt von Digitalisaten in deutschen Museen, graph. Darst., nestor c/o Die Deutsche Bibliothek, Frankfurt am Main.