

- a. Title
- b. Author(s)
- c. Affiliation
- d. City and country
- e. Email address
- f. Abstract
- g. Article text

a. Title

Long-term Accessibility and Preservation of Scientific Publications and Data

b. Author

Heike Neuroth (Abstract)

Stefan Strathmann (Article text)

c. Affiliation

Niedersächsische Staats- und Universitätsbibliothek Göttingen (SUB) / Göttingen State and University Library

nestor - Network of Expertise in Long-term STOrage of Digital Resources

d. City and country

Göttingen, Germany

e. Email address

strathmann@sub.uni-goettingen.de

f. Abstract

Scientific data are the basis for each scientific publication, e.g. in the natural sciences (space observation data, x-ray data, climate data etc.) or in the social science (e.g. statistical data). They are the key indicator to ensure the scientific relevance and authenticity of interpretations in the scientific community. Furthermore these interpretations, conclusions etc. may influence future research and projects. To ensure long-term access to scientific published information both the published articles, research reports etc. and the scientific data have to be saved/stored in a trusted way, optimal linked to each other. Trusted way means in this context digital long-term preservation, which includes long-term access and long-term preservation of all relevant digital information. This guarantees that other scientists and researchers are able to prove results of scientific work.

The presentation will give an overview of relevant aspects of long-term accessibility and preservation in the scientific field (heterogeneity of data, organizational models etc.).

g. Article text

0.)

This talk presents an overview of some aspects of digital preservation from the point of view of the Göttingen State and University Library (SUB)¹. It describes some activities or projects in this field.

1.)

Scientific data are the basis for all scientific research. They are the key indicator to ensure the scientific relevance and authenticity of interpretation in the scientific community. Furthermore these data may allow future research.

To ensure long-term access to scientific information both the published articles, research reports etc. and the scientific data have to be archived in a trusted way, optimally linked to each other. Trusted way in this context means digital long-term preservation, which includes long-term access and long-term preservation of all relevant digital information.

Over the past years, there have been discussions about strategies and techniques for long-term preservation of digital information, particularly within the distributed structure of Germany's library and archival institutions. It is obvious that the approach to a successful solution of these issues must be a cooperative one. It was decided to establish a permanent distributed infrastructure for long-term preservation and long-term accessibility of digital resources in Germany comparable to the Digital Preservation Coalition (DPC²) in the UK.

The initial phase of building this infrastructure is funded by the German Ministry of Education and Research and is called nector - Network of Expertise in long-term STORage of digital Resources³.

The nector partners are:

- Die Deutsche Bibliothek⁴ (German National Library) as the leading institution for the project
- Niedersächsische Staats- und Universitätsbibliothek Göttingen (Goettingen State and University Library)
- Computer and Media Service of Humboldt-University⁵, Berlin
- Bayerische Staatsbibliothek⁶ (Bavarian State Library)
- Institut für Museumskunde⁷ (Institute for Museum Information)
- Generaldirektion der Staatlichen Archive Bayerns⁸ (Bavarian State Archive – Head Office)

¹ <http://www.sub.uni-goettingen.de>

² <http://www.dpconline.org>

³ <http://www.langzeitarchivierung.de>

⁴ <http://www.ddb.de>

⁵ <http://www.cms.hu-berlin.de>

⁶ <http://www.bsb-muenchen.de>

⁷ <http://www.smb.spk-berlin.de/ifm>

⁸ <http://www.gda.bayern.de>

nestor will not archive anything – it will provide support in questions of long time preservation of digital objects.

Some of the aims of the nestor project are:

- to heighten awareness of the problem
- to produce and distribute expertise and information
- to facilitate cooperation
- to establish a durable organisational form
- ...

On the one hand nestor tries to establish a communication and information infrastructure (Subject Gateway⁹, Calendar, Mailing list, Newsletter, Online Forum ...) on the other hand nestor commissions and develops information materials. At the end of the project nestor will publish concrete national recommendations, guidelines and best practices for topics like:

- criteria for trustworthy digital archives
- certification of archives
- collection development policies and selection criteria for archiving digital sources
- ...

Expert reports have been commissioned dealing with topics like:

- Electronic Journals
- Perspectives of long-term preservation of multimedia objects
- A comparison of existing archiving systems
- ...

One expert report on scientific raw data will examine - among other things - questions like: Which scientific raw data are archived in Germany until now? How are they indexed? Which data are relevant for archiving? How should they be indexed? What is the expected volume of data? What data formats are in use? How can the data be accessed? Which business models exist in the field of scientific raw data? And more... The report will be published in the next year.

2.)

The KOPAL project (Cooperative development of a long-term digital information archive) can be seen as a technical addition to the nestor project.

The project is developed jointly by the following partners:

- Die Deutsche Bibliothek Frankfurt (DDB),
- Niedersächsische Staats- und Universitätsbibliothek Göttingen (SUB),
- Gesellschaft für wissenschaftliche Datenverarbeitung mbH Göttingen¹⁰ (GWDG)

⁹ <http://nestor.sub.uni-goettingen.de>

¹⁰ <http://www.gwdg.de>

- and IBM Deutschland GmbH¹¹.

KOPAL is supported financially by the German Federal Ministry of Education and Research.

The goal of the KOPAL project is to develop a technical and organisational infrastructure for the long term preservation of digital data and publications which can be used by different organisations in the academic, cultural heritage and business fields. The KOPAL project will cooperatively develop and operate a reliable long-term archive for digital data. This archive is based on DIAS (Digital Information and Archiving System), developed by IBM Netherlands and the National Library of The Netherlands¹². DIAS is compliant to the OAIS framework (Open Archival Information System). The KOPAL project will enhance the DIAS with monitoring functions, with preservation management functions, with flexible data import and export functions based on the object description scheme METS (Metadata Encoding and Transmission Standard) and other functions.

During the initial three years of the project the partners DDB and SUB will file digital materials into the long-term archive. The partners will start with relatively homogenous material like images of the Center for Retrospective Digitization (Göttinger Digitalisierungszentrum GDZ¹³) or TEX documents of the TeX-Doc Center¹⁴ (TeXDocC). Later more heterogeneous and complex material like web sites or digital video shall be archived in the system.

In the future, the long-term archive for digital information will provide other institutions with the opportunity to keep data available on a long term basis.

3.)

Beside national activities, the SUB Goettingen is involved in international initiatives to develop solutions in the field of digital preservation and accessibility of digital objects, for example in the OCLC/RLG working group on preservation metadata (PREMIS - PREServation Metadata: Implementation Strategies¹⁵) or in the DC Preservation working group¹⁶.

All these and other activities complement one another and can be seen as steps on the way to solving the problems of preservation and accessibility of scientific data and publications.

4.)

References:

Dobratz, Susanne; Neuroth, Heike: nestor. Network of Expertise in Long-term STORAGE of Digital Resources - A Digital Preservation Initiative for Germany. In: D-Lib Magazine, April 2004, Volume 10 Number 4, ISSN 1082-9873, DOI: 10.1045/april2004-dobratz
URL: <http://www.dlib.org/dlib/april04/dobratz/04dobratz.html>

¹¹ <http://www.ibm.com/de>

¹² <http://www.kb.nl/>

¹³ <http://gdz.sub.uni-goettingen.de>

¹⁴ <http://www.sub.uni-goettingen.de/projekte/tex-doc-center.html>

¹⁵ <http://www.oclc.org/research/projects/pmwg/default.htm>

¹⁶ <http://dublincore.org/groups/preservation/>