

Bericht über aktuelle Entwicklungen

iPRES – ECDL



Bericht über aktuelle Entwicklungen

iPRES

International Conference on Preservation of Digital Objects, 15.-16. September 2005 in Göttingen, Deutschland

ECDL

9th European Conference on Research and Advanced Technology for Digital Libraries (ECDL 2005), 18.-23. September 2005 in Wien, Österreich

iPRES

International Conference on Preservation of
Digital Objects, 15.-16. September 2005 in
Göttingen, Deutschland

<http://rdd.sub.uni-goettingen.de/conferences/ipres/>

iPRES

Basis-Informationen

Schwerpunkte – Inhalte

Tendenzen – Einordnung – Kritik

Ausblick

Hinweise



iPRES

Basis-Informationen (I)

Organisation:

- SUB Göttingen, Deutschland
- nestor - Kompetenznetzwerk Langzeitarchivierung, Deutschland

Weitere Organisatoren:

- Österreichische Nationalbibliothek - ÖNB, Wien, Österreich
- California Digital Library - CDL, USA
- Cornell University Library - CUL, USA
- electronic Information for Libraries - eIFL, EU
- Chinese Academy of Sciences - CAS, China

iPRES

Basis-Informationen (II)

Teilnehmer: ca. 100 Teilnehmer aus 17 Ländern: u.a.
aus Australien, Belgien, China, Estland, Frankreich,
Großbritannien, Israel, den Niederlanden, Norwegen,
Österreich, Polen, der Schweiz, Slowenien, den USA;
etwa die Hälfte der Teilnehmer war aus Deutschland

Konferenzsprache: Englisch (Simultanübersetzung)

iPRES

Basis-Informationen (III)

Unterstützung – Finanzierung:

- Bundesministerium für Bildung und Forschung (BMBF)
- Deutsche Forschungsgemeinschaft (DFG)



iPRES

Schwerpunkte – Inhalte (I)

Keynote 1: Elisabeth Niggemann (DDB)

Topic I: Preservation Policies

Topic II: Technical Workflow

Topic III: Web Archiving

Keynote 2: Stephen Abrams (Harvard University Library)

Recent Developments

iPRES

Schwerpunkte – Inhalte (II)

Keynote 1: Preserving Our Digital Heritage - Weaving The Web Of Trust, Elisabeth Niggemann (Die Deutsche Bibliothek, DDB)

- Hoffnung auf innerdeutschen Kommunikations- und Koordinierungseffekt der iPRES
- Verantwortlichkeiten im Bereich der LZA sollten organisationsübergreifend - nicht zwingend orientiert an den gewachsenen Organisationsstrukturen – kooperativ und partnerschaftlich verteilt werden („web of trust“)
- die Finanzierung digitaler Langzeitarchivierungsaktivitäten ergibt sich durch ihren Charakter als nationale Aufgabe

iPRES

Schwerpunkte – Inhalte (III)

Topic I: Preservation Policies

- Developing A National Preservation Network For STM E-Journals, Zhixiong Zhang (Chinese Academy of Science, CAS)
- Developing A National Preservation Policy: Experiences In Germany, Stefan Strathmann (Göttingen State and University Library, SUB)
- A National Preservation Policy For The UK (?) - Thoughts And Challenges, Manjula Patel (UKOLN)

iPRES

Schwerpunkte – Inhalte (IV)

Topic II: Technical Workflow

- Information Life Cycle Management And Long Term Preservation - Technical Aspects Of Workflow Organization, Reinhard Altenhöner (Die Deutsche Bibliothek, DDB)
- Workflow Management In An Operational Archiving Environment, Johan Stapel (e-Depot - National Library of The Netherlands, KB)
- Safe Havens In A Choppy Sea: Digital Object Management Workflows At The National Library of Australia, Gerard Clifton (National Library of Australia, NLA)

iPRES

Schwerpunkte – Inhalte (V)

Topic III: Web Archiving

- Future-Proofing The Web: What We Can Do Today, John Kunze (California Digital Library, CDL)
- Digital Preservation Activities In The National Library Of China, Chunming Li (National Library of China, NLC)
- Web Archiving: A Collaborative Effort In Progress, Julien Masanès (European Digital Archives)

iPRES

Schwerpunkte – Inhalte (VI)

Keynote 2: Digital Formats And Preservation, Stephen Abrams (Harvard University Library, HUL)

- Überblick über die im Archivierungsablauf gegebenen Formatabhängigkeiten
- Darstellung neuer Entwicklungen im Bereich der Werkzeuge und Verfahrensweisen, die die Formate berücksichtigen
- Beschreibung von Format-Migrationen zur Archivierung (Fallstudien)
- Betonung der Bedeutung der Einrichtung öffentlicher Format Registry-Mechanismen

iPRES

Schwerpunkte – Inhalte (VII)

Recent Developments

- PREMIS: Implementation & Preservation Metadata, Robin Dale (Research Library Group, RLG)
- DELOS - New Developments, Michael Day (Network of Excellence on Digital Libraries, UKOLN)
- Cornell University Library Open Archival Information System, Marcy Rosenkrantz (Cornell University Library, CUL)
- METS: From Presentation To Preservation, Markus Enders (Göttingen State and University Library, SUB)

iPRES

Tendenzen – Einordnung – Kritik (I)

- ein erstes Resümee zur iPRES fällt sehr positiv aus
- die Konferenz ist bei den Teilnehmern – eine Evaluation wurde mittels einer Fragebogenaktion durchgeführt – sehr gut angekommen
- die thematische Ausrichtung, die Inhalte und Qualität der Vorträge, die Referenten sowie die Organisation der Tagung wurden als sehr positiv bewertet
- eine Pressekonferenz mit ca. 10 Journalisten fand am ersten Konferenztag statt
- die (Fach-)Presse hat umfangreich über die Konferenz berichtet; auch der Hörfunk nahm sich der Thematik an

iPRES

Tendenzen - Einordnung – Kritik (II)

- die Beteiligung aus der deutschen Archiv- und Museumsszene war enttäuschend; dies wurde sehr bedauert und muss Anlass zu weiteren Überlegungen und Aktivitäten durch nestor sein
- es wurde angeregt, Einführungen auf Anfängerniveau in Deutschland durchzuführen, um die internationalen Diskussionen und Erfahrungen auf deutsche Verhältnisse herunterzubrechen und so Kompetenz im Bereich der Langzeitarchivierung digitaler Objekte auf allen Ebenen, in Museen, Archiven und Bibliotheken zu entwickeln
- ein erster Schritt in diese Richtung wird am 29. November 2005 an der SUB Göttingen getan - in einem „nestor-Seminar“ wird in die Langzeitarchivierung digitaler Objekte eingeführt

iPRES

Ausblick

- während der Abschlussdiskussion regte der Direktor der SUB Göttingen, Prof. Mittler, an, dass die iPRES-Folgekonferenz im Jahre 2006 in den USA stattfinden solle (eine iPRES-Vorläuferkonferenz hatte Mitte 2004 in China stattgefunden)
- dieser Anstoß wurde sowohl von der California Digital Library als auch von der Cornell University Library positiv aufgenommen
- eine definitive Entscheidung gibt es noch nicht, beide amerikanischen Bibliotheken haben ihre endgültige Bereitschaft signalisiert!

iPRES

Hinweise (I)

Die gesamte Konferenz kann im Nachhinein noch einmal „besucht“
und damit nachvollzogen werden unter:

<http://rdd.sub.uni-goettingen.de/conferences/ipres/>

- hier finden sich im Bereich „Media Files“ alle Präsentationen und Dokumente der Konferenz: u.a. die Abstracts der Vorträge, die Biographien der Referenten
- im Bereich „Programm“ können zudem alle Vorträge als **Video Stream** abgerufen werden

iPRES

Hinweise (II)

- in Kürze wird eine DVD erscheinen, die alle Vorträge enthält – sie kann im Universitätsverlag Göttingen (<http://www.univerlag.uni-goettingen.de/>) käuflich erworben werden
- die Webseite bietet darüber hinaus noch viele weitere Informationen: so etwa ein thematisches Glossar, die Teilnehmerliste, Fotos von der Konferenz etc.

Falls Sie die Konferenz also auf einem der genannten Wege nachträglich „besuchen“ wollen:

Viel Spaß dabei!

Im unmittelbaren Anschluss an die iPRES-Konferenz begann dann in Wien die ...

ECDL

9th European Conference on Research and
Advanced Technology for Digital Libraries
(ECDL 2005), 18.-23. September 2005 in
Wien, Österreich

<http://www.ecdl2005.org/>

ECDL

Basis-Informationen
Schwerpunkte – Inhalte
Resümee



ECDL

Basis-Informationen

Organisation:

- Technische Universität Wien
- Österreichische Computer Gesellschaft, Wien
- Österreichische Nationalbibliothek, Wien

Teilnehmer:

- ca. 400 Teilnehmer (inklusive der Workshops) aus nahezu allen europäischen Ländern (Schwerpunkt West- und Mitteleuropa) sowie viele Besucher aus den USA

ECDL

Schwerpunkte – Inhalte (I)

Struktur der ECDL:

- 6 Tutorials (vor der Konferenz)
- 2 Keynotes
- 12 Sessions
- 2 Panels
- Poster and Demo Session
- 5 Workshops (im Anschluss an die ECDL)



ECDL

Schwerpunkte – Inhalte (II)

Keynote 1: Plenty of Room at the Bottom? Personal Digital Libraries and Collections, Neil Beagrie (British Library - JISC)

- die Bedeutung persönlicher Sammlungen, speziell solche von Wissenschaftlern, wird zukünftig immer mehr zunehmen
- “Personal Digital Collections” und “Personal Curation” als Trends
- diese Sammlungen werden zukünftig oftmals Basis für Sammlungen in Archiven, Bibliotheken und Museen sein, was diese Einrichtungen berücksichtigen sollten

ECDL

Schwerpunkte – Inhalte (III)

Beispiele für „Collective Services“:

- www.memoriesforlife.org (Forschung zu Erinnerungen)
- www.webarchive.org.uk (Archivierung von Personal Blogs)
- www.paradigm.ac.uk (Projekt zu digit. Papieren von Politikern)
- www.ourmedia.org (freie Speicherung für Individuen)
- www.flickr.com (Speicherung von Photos)
- www.connotea.org (Weitergabe von Referenzen / Bookmarks)

Der Beitrag Beagries ist auch nachlesbar in der Juni-Ausgabe des D-Lib-Magazins:

<http://www.dlib.org/dlib/june05/beagrie/06beagrie.html>

ECDL

Schwerpunkte – Inhalte (IV)

Keynote 2: From Centralized Systems to Flexible Content and Service Federations: Promises & Challenges of Next Generation Digital Library Architectures, Erich Neuhold (TU Wien)

- Beschreibung des Kontextes in dem sich DL bewegen (Umfeld: Technologien, Nutzer- und Marktanforderungen, Diensteanbieter)
- „Cyber Infrastructure“ (e-Science) und „Grid Computing“ als vielversprechende Technologien speziell für den Austausch großer Datenmengen und die Verarbeitung an verteilten Standorten

ECDL

Schwerpunkte – Inhalte (V)

Sessions der ECDL (insgesamt 12):

- Digital Library Models and Architectures
 - Multimedia Digital Libraries
 - Metadata
 - Text Classification in Digital Libraries
 - Building Digital Libraries
 - Digital Libraries and e-Learning
 - XML
 - Text Digital Libraries
 - Searching
 - Digital Preservation
 - User Studies 1 + 2
- Die Sessions bestanden jeweils aus drei bis vier Vorträgen mit anschließender Diskussion (Befragung des Referenten).
- Die Sessions wurden moderiert, Zusammenhänge zwischen den Vorträgen oder eine Abschlussdiskussion gab es nicht.

ECDL

Schwerpunkte – Inhalte (VI)

Beispiel-Session: Digital Preservation

- The DSpace Open Source Digital Asset Management System: Challenges and Opportunities (Robert Tansley, Hewlett Packard)
- File-based storage of Digital Objects and constituent datastreams: XMLtapes and Internet Archive ARC files (Herbert Van de Sompel, Los Alamos National Laboratory)
- A No-Compromises Architecture for Digital Document Preservation (Thomas A. Phelps, University of Liverpool)
- A study into the effect of digitisation projects on the management and stability of historic photograph collections (Veronica Davis Perkins, Middlesex University)

ECDL

Schwerpunkte – Inhalte (VII)

Panels der ECDL:

- Digital Libraries over the GRID: Heaven or Hell?
- Does eScience need Digital Libraries?

Hier gab es Kurzreferate der Referenten („Panelists“), die dann durch eine Podiumsdiskussion und anschließend durch eine offene Diskussion mit dem Publikum ergänzt wurde.

ECDL

Schwerpunkte – Inhalte (VIII)

Poster and Demo Session:

- für die Poster und Demonstrationen war eine Anmeldung erforderlich, die einzelnen Beiträge wurden evaluiert
- in einer „Poster Spotlight Session“ konnten die Verantwortlichen kurz ihren Beitrag skizzieren
- anschließend wurden die Forschungsvorhaben, Projekte und sonstigen Beiträge mittels der Poster oder Demonstrationen präsentiert
- insgesamt waren ca. 60 Poster und Demonstrationen vertreten

ECDL

Schwerpunkte – Inhalte (IX)

Workshops (vom 21.-23. September, im Anschluss an die ECDL):

- HDL (3rd Workshop on Digital Libraries in Health Care)
- NKOS (4th Workshop on Mapping Knowledge Organisation Systems)
- KED (1st Workshop on Knowledge Extraction and Deployment for Digital Libraries and Repositories)
- CLEF (Workshop on the Cross-Language Evaluation Forum)
- IAWAW (5th International Workshop on Web Archiving and Digital Preservation) [... hierzu später mehr ...]

ECDL

Resümee (I)

- ein allgemeingültiges, wertendes Resümee einer solch großen Konferenz ist schlecht möglich
- mein persönlicher Eindruck: bezogen auf die einzelnen Veranstaltungen, aber auch innerhalb einzelner Sessions war das Niveau sehr unterschiedlich
- die Themen „e-Science“ und „Grid“ werden zukünftig größere Bedeutung haben
- Suchmaschinen-Technologie wurde auch in Bezug auf Google-Aktivitäten stark diskutiert

ECDL

Resümee (II)

- beeindruckend fand ich den Vortrag von Neil Beagrie (Keynote)
- allgemein war es für mich eine gute Möglichkeit, Gespräche zu führen, Kontakte zu knüpfen und Erfahrungen auszutauschen

Zu einer vertiefenden Beschäftigung mit den Themen der ECDL 2005 empfiehlt sich der Tagungsband:

Lecture Notes in Computer Science. Publisher: Springer-Verlag GmbH, ISSN: 0302-9743, Volume 3652 / 2005. Title: Research and Advanced Technology for Digital Libraries: 9th European Conference, ECDL 2005, Vienna, Austria, September 18-23, 2005. Proceedings Editors: Andreas Rauber, Stavros Christodoulakis, A Min Tjoa; ISBN: 3-540-28767-1, DOI: 10.1007/11551362.

ECDL

... nun zum Schluss noch einen Blick in die Workshops,
die direkt im Anschluss an die ECDL stattfanden.

Einige Kolleginnen und Kollegen der SUB Göttingen
haben an dem IWAW-Workshop teilgenommen.

IWAW

IWAW –

5th International Workshop on Web Archiving
and Digital Preservation

22.-23. September 2005



IWAW

Basis-Informationen
Schwerpunkte – Inhalte
Resümee
Links



IWAW

Basis-Informationen (I)

- Teilnehmerzahl: ca. 60 Teilnehmer - vornehmlich aus Europa und den USA - aber auch aus Australien und Japan
- der Workshop wurde in diesem Jahr erstmals auf zwei Tage ausgedehnt
- Webarchivierung war weiterhin das Hauptthema des ersten Tages
- Ausdehnung auf Langzeitarchivierung und Projektvorstellungen aus dem Umfeld von Webarchivierung am zweiten Tag

IWAW

Basis-Informationen (II)

- der Workshop gliederte sich in 5 Sessions
- an beiden Tagen wurden Papiere zu Webarchivierung und Langzeitarchivierung präsentiert
- einige Präsentationen versuchten beide Themen zu verbinden
- nach den Sessions war jeweils Zeit für Fragen und Diskussionen eingeplant

IWAW

Schwerpunkte – Inhalte (I)

- **Session 1 - IIPC-Results:** Vorstellung der Entwicklungen der International Internet Preservation Consortium (IIPC), die gemeinsame Tools und gemeinsame Standards für die Webarchivierung zu entwickeln versucht
- **Session 2 - Audio und Video Web Archiving:** Präsentation von Thomas Drugeon vom Institut National de l'Audiovisuel (INA), zu einem staatlichen Sammelauftrags-Projekt für ca. 10.000 bis 15.000 Websites (u.a. spezifische Domains, wie das Medienportal TF1)
- **Session 3 - Time Dimension:** „Linkrotting“ und Vorstellung eines formalen Modells zur Feststellung von zeitlichen Relationen

IWAW

Schwerpunkte – Inhalte (II)

- **Session 4 - Digital Preservation:**
 - Strategien in der Langzeitarchivierung (Emulation, Entscheidungen)
 - Tools im Umfeld von Langzeitarchivierung (Grace, PANIC)
 - Modell zum Austausch von LZA-Objekten unter Verwendung von DSpace
- **Session 5 – Current Projects and Issues:**
 - Webarchivierung in Dänemark
 - Gründung des European Digital Archive (Webarchivierung)
 - das deutsche Projekt kopal

IWAW

Resümee (I)

- die Webarchivierungstools werden ebenso brauchbar für große Mengen an Daten
- ausgefeiltere Webharvestingstrategien werden gerade entwickelt und automatisiert
- Die Webarchivierungsszene ist gespalten. Selektive Ansätze, unterhalb von TLDs waren im Programm kaum vertreten.

IWAW

Resümee (II)

- es wurden viele interessante Tools für die Langzeitarchivierung vorgestellt, eines ist als Webservice verfügbar (PANIC)
- die Themen Webarchivierung und Langzeitverfügbarkeit werden weiterhin separat betrachtet: so wurde zum Beispiel die Integration von Webarchivgut und -metadaten in Container- und Metadatenformaten zur Langzeitarchivierung erst in der Diskussion thematisiert

IWAW - Links

- International Web Archiving Workshop (IWAW), <http://www.iwaw.net/05/index.html>.
- International Internet Preservation Consortium (IIPC), <http://www.netpreserve.org/>.
- Internet Archive, <http://www.archive.org/>.
- Heritrix: Internet Archive Web Crawler, <http://sourceforge.net/projects/archive-crawler>.
- NutchWAX - Nutch + Web Archive eXtensions, <http://sourceforge.net/projects/archive-access/>.
- Apache Lucene, <http://lucene.apache.org/>.
- Institut National de l'Audiovisuel (INA), <http://www.ina.fr/>.
- PANIC - Preservation webservices Architecture for Newmedia and Interactive Collections, <http://metadata.net/panic/>.
- Modular emulation as a long-term preservation strategy for digital objects, <http://www.iwaw.net/05/index.html>.
- Kopal - Cooperative development of a long-term digital information archive, <http://kopal.langzeitarchivierung.de/>.

iPRES – ECDL – IWAW

Vielen Dank für Ihre Aufmerksamkeit!

Haben Sie Fragen, Kommentare oder Anregungen?

Referenten:

Matthias Schulze (schulze@sub.uni-goettingen.de)

Olaf Brandt (brandt@sub.uni-goettingen.de)