



Preservation of Scientific Publications

“Tackling the Preservation Challenge”

Barbara Sierman, 12-12-2008



Overview

- KB and the e-Depot
- The DARE project
- Reflections
- Driver2



KB, National Library of the Netherlands

- Founded in 1798
- 275 fte
- 3.5 million paper based publications , 49 km of books, 18 km periodicals
- (Voluntary) deposit since 1974
- 12 million digital publications
- Funded by the Ministry of Education, Culture and Science
- Ambition to be one of the Safe Places for scientific publications
- Mission (part of): *we give researches and students access to research information*

KB and the e-Depot

- 1994 e-publications part of deposit
- Several pilot projects with publishers
- E-Depot started in 2003, based on DIAS system (IBM)

Currently (12 TB):

- Articles from 14 (main) STM publishers
- National deposit collection
- Research results of the Dutch universities
- Planned (500 TB in 2013):
 - “masters” of digitized collections
 - Selection of Dutch websites

DARE (Digital Academic Repositories)



- 2003 Digital Academic Repositories (DARE)
- Goal: to make the research results digitally accessible and available for different services
- Via portal: DARE-net; network of university repositories in the Netherlands (now part of NARCIS)
- Long term preservation at KB / e-Depot
- Partners:
 - 14 Universities and Organisations of Higher Education
 - Netherlands Organisation for Scientific Research
 - Royal Netherlands Academy of Arts and Sciences
 - KB, National Library of the Netherlands
 - Coordination: SURF Foundation

DARE (Digital Academic Repositories)



- Agreement on repository principles:
 - Repository is OAI compatible (Open Archiving Initiative), OAI-PMH
 - Descriptive metadata in Dublin Core Simple
 - DARE Guidelines for DC
 - Structural metadata in MPEG21-DIDL
 - Metadata in XML
 - ...

DARE harvesting

- Scientific publications, how to define?

“Those objects that are made available on the public Internet through Institutional Repositories”

- DARE partners decide individually what they'll deliver
 - No restricted publications
 - Not under embargo etc.
 - No data sets
 - Thesis: KB only stores digital publication



Agreement (practical)

- Periodical harvest by KB (OAI-PMH protocol)
- Metadata and/or objects (“modification date”)
- Limited set of file formats/version
- No possibility to delete objects in LTP
- Evaluation of agreement in 2009

Agreement on file formats

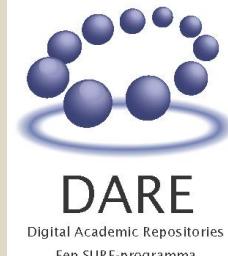
- Mostly PDF files (90%)
- Around 20 other file formats
- List of file formats in agreement
 - Zip files allowed uncompressed
 - Only those 20 file formats in Zip
 - Dynamic webpages received as HTML pages
- Negotiation about new file formats
- KB process: Normalisation (Word → PDF/A)
Archive original as well

Agreement on access

- No restricted access!
- No dark archive!
- Publications in KB Catalogue
- Access preferable from repository

The screenshot shows a Microsoft Internet Explorer window displaying search results for "waaijers" in the KB e-Depot. The results page is titled "Resultaten" and shows 2 of 12 titles. The first result is for "Waaijers, Leo" with a link to a detailed record. The record includes fields such as In: KB e-Depot, Titel: Wetenschappelijk publiceren: diefje-niet-verlos, Auteur: Waaijers, Leo, Onderwerp: lezing opening 0 Maastricht 0, Annotatie: DARLUTer gelegenheid van de opening van de Universiteitsbibliotheek Maastricht, op 21 mei 2003, hield Leo Waaijers een toespraak waarin hij de wereld van het wetenschappelijk publiceren analyseerde. Hier de vrijwel integrale weergave daarvan., Uitgever: Cramwinkel, Type: Conference lecture, Link: (with a right-pointing arrow), Bron: Raadpleeg publicatie in broncollectie / View publication in source collection, Formaat: Ophalen depotexemplaar / Download deposit copy, Informe Professional; [7] 7/8, P. 30-35 (2003), Formaat: image/jpg, Copyright: (c) Waaijers, Leo, 2003.

New developments e-Depot / DARE



- Improvement of communication and reporting
 - Central place with info
 - Contact persons
 - Harvest URL's
 - News on updates, changes
- Improvement of the harvester and ingest process at KB
 - MODS
 - More checks (DROID, JHOVE)
- Update archiving agreements

New developments

- Balance between dp requirements and material offered
Limited influence of repository manager
- Important to know scholarly communication cycle
 - Participation of KB in Planets(preservation planning), Driver1/2, PARSE Insight, PEER etc
- New kind of objects
 - Compound, with additional objects, data sets etc.
 - Dutch Preservation landscape: 3 organizations involved!



DriverII and Repositories

- Goal:
 - Repository Infrastructure for Enhanced Publications
- KB participates in:
 - Research into Enhanced Publications
 - Consequences for LTP Archiving of EP's
 - Create a demonstrator for harvesting these EP's
 - LTP Archiving agreements repositories

DriverII and Repositories

- Inventory study amongst European repositories
 - Is long term availability secured?
 - 2006 : 72.8 % yes
 - 2008: 52.2 % yes Decrease!
 - More aware of implications? More knowledge?
 - 37.6 % task National Library
 - 76 % internal procedure (projects like SHERPA2, PRESERV2, ABES etc.)

A photograph of a modern building facade made of light-colored rectangular panels. A large, stylized yellow logo consisting of the letters 'KB' is prominently displayed on the side of the building.

Thank you for your
attention!