Institutional Repositories, Long Term Preservation and the changing nature of Scholarly Publications

Paul Doorenbosch
KB, National Library of the Netherlands

Open Repositories 2010, Madrid, 2010-07-07
Background information

DRIVER II (2007-2009)

Current State of repository infrastructure in Europe

Roll out of Driver I repository infrastructure over the rest of Europe and make it more sustainable

Research into nature, technique and preservation of Enhanced Publications as a new expression of (scholarly) publishing

Project under FP7

Participants from: GB, DU, BE, NL, GR

Coordination for NL: SURF FOUNDATION
Overview

1. The Dutch LTP and Repository Landscape
2. Enhanced Publications
4. Challenges, Issues
Overview

1. The Dutch LTP and Repository Landscape
2. Enhanced Publications
4. Challenges, Issues
Dutch LTP landscape

National Coalition for Digital Preservation

National Library (KB)  Text (publications)

DANS  Primary scientific data

National Archive  Governmental data

Sound & Vision  Broadcast - AV

National Archival Coalition  local and regional archives

Cultural Heritage Coalition (prop)  museums, monuments, archeology

??????  Medical data

??????  Industrial data

??? data

Dutch research LTP stakeholders

- Primary research data (humanities, technology)
- Publications (text+)
- KB
- IR (Universities, Scientific Institutions)
- Primary data (technology)
- DANS
- 3TU/DELFT

[6]
Overview

1. The Dutch LTP and Repository Landscape
2. Enhanced Publications
4. Challenges, Issues
Enhanced Publications

Publication (e-print) combined with research data

- Improve interpretation and verification
- Promote available data
- Browsable network of related items
Demonstrator EP - goals

- Suggest the perfect solution
- Trigger researchers
- Trigger repository community
Collect publication, data and metadata

- define and identify aggregations of (web) resources using OAI-ORE
Demonstrator EP - approach

... in a durable way

- using persistent identifiers, digital author identifiers, ontologies, etc.
Demonstrator EP - approach

... in an exchangeable format

- describe aggregations as triple-relations using RDF

```xml
<!-- SIMPLIFIED ORE FOR DEMO ONLY -->
<rdf:Description rdf:about="urn:nbn:nl:ui:13-qrs-tuv">
  <rdf:type rdf:resource="http://www.openarchives.org/ore/terms/Aggregation"/>
  <dans:epType dans:resource="info:eu-repo/semantics/enhanced_publication"/>
  <ore:aggregates rdf:resource="urn:nbn:nl:ui:26/15677"/>
  <ore:aggregates rdf:resource="urn:nbn:nl:ui:26/15678"/>
  <ore:aggregates rdf:resource="urn:nbn:nl:ui:26/15680"/>
  <dcterms:creator rdf:parseType="Resource">
    <foaf:name>M. Versluis</foaf:name>
    <dai:identifier info:eu-repo/dai/nl/298870398</dai:identifier>
  </dcterms:creator>
</rdf:Description>
```
Demonstrator EP - approach

...visualise in browser

- translate RDF to HTML using XSLT
Visualisation of an EP

- Nodes can be clicked to center them.
- Nodes can be dragged to rearrange them.
- When moving over a node, its description will appear (if available).

Challenges

- Optimal generic presentation
- Feedback & annotations
- Identification
- Resource consistency
- Availability of resources
- Creation/Ownership/Responsibilities
- Long Term Preservation
Overview

1. The Dutch LTP and Repository Landscape
2. Enhanced Publications
4. Challenges, Issues
Long Term Preservation

- Demonstrator
- Visualisation
- Report

- Not the solution
- Simplified
- Research and Experiment
- Identifying Issues
- Dutch situation as a test case
Dutch research LTP stakeholders

- Publications (text+)
- KB
- IR (Universities, Scientific Institutions)
- Primary research data (humanities, technology)
- DANS
Relations between PID

For future use: Expansion of the national resolver?: Constructing a new resource map?
Archiving Agreement and EPs

- Known technical challenges: “business as usual”
- Organisational challenges need to be solved
Overview

1. The Dutch LTP and Repository Landscape
2. Enhanced Publications
Challenges, issues

- Organisation, preservation policies, archival agreement
- Ownership, responsibilities
- Rights (national, international)
- Authenticity and integrity; look and feel
- Versioning/updates
- Nature of consisting files
- Persistency (files, authors, identification, citation)
- Future use, usability and access
- Implementation
Developments

- Evolving forms of scholarly publication (EP, enhanced journals, VRE, e-books, databases, linked data, websites, twitter, triples)
- Growing diversity of file formats
- Changing ways of organising research
- ....

Nevertheless:

There are a lot of technical challenges in preserving and giving long term access to new kinds of scholarly ‘publications’, but still the organisational one looks the most challenging.
Thank you

Paul Doorenbosch
Eugène Dürr
Arjan Hogenaar
Maarten Hoogerwerff
Jens Ludwig
Birgit Schmidt
Barbara Sierman
Maurice Vanderfeesten

Paul Doorenbosch
Koninklijke Bibliotheek
paul.doorenbosch@kb.nl