

# ***DRIVER***

---

## **Long Term Preservation for Enhanced Publications in the DRIVER Infrastructure**

Dale Peters, Scientific Technical Manager  
Göttingen State and University Library



WePreserve Workshop, October 2008



# *DRIVER Vision*

---

- **Open Access**
  - European and worldwide research output (publications and data sets )  
openly accessible through institutional repositories
- **Interoperability**
  - ensuring automated aggregation of all scholarly research output into one  
virtual open knowledge base
- **Open Content**
  - enables service providers to provide a wide range of end-user services to  
researchers (search, browse, profiling, visualisation, citation, impact  
metrics...)



# Partners



...more global partners.....



## DIGITAL REPOSITORIES INFRASTRUCTURE VISION FOR EUROPEAN RESEARCH

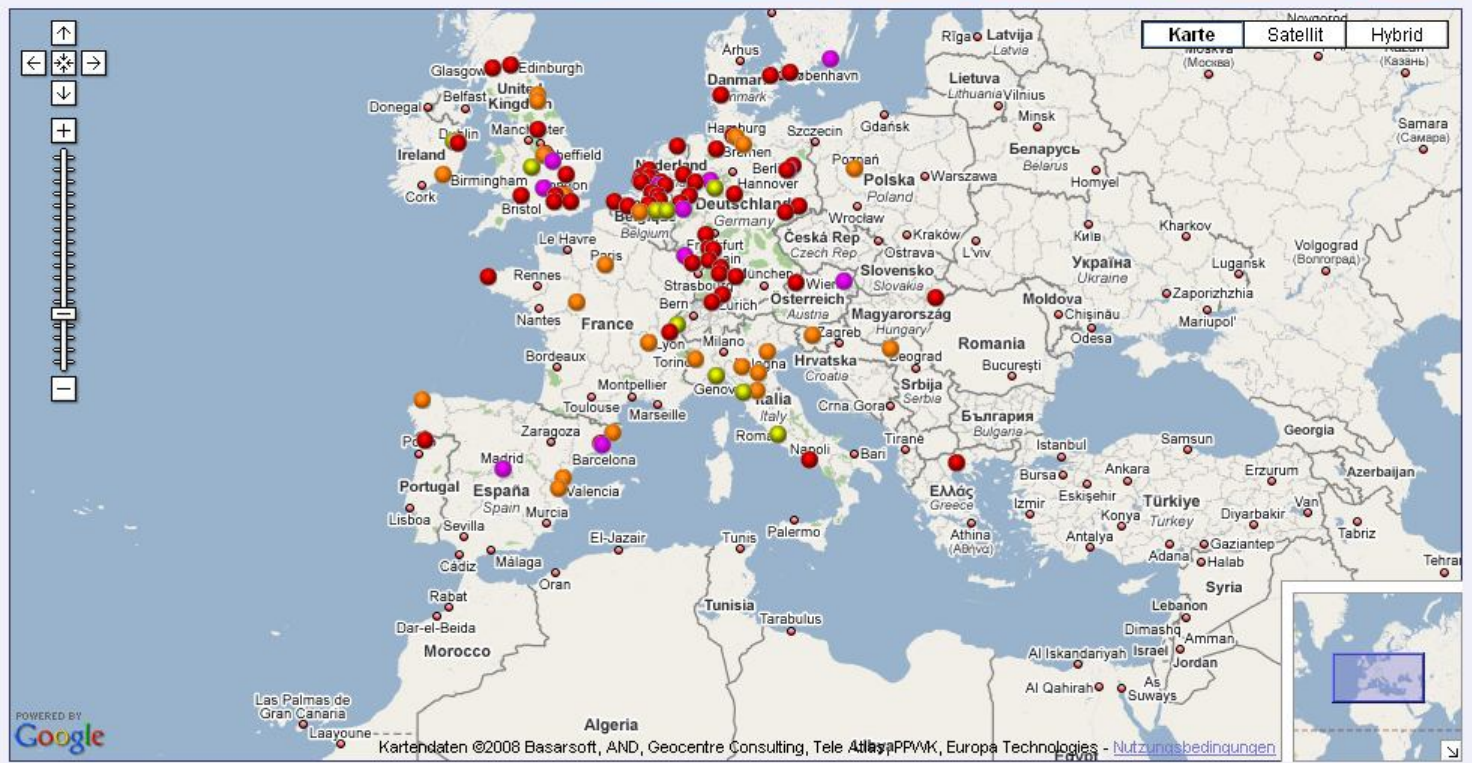
Home **Repositories** Network Collections



Show Map

Number of records: ● < 50 ● 50 - 500 ● 500 - 1000 ● > 1000

- Repositories
- IT Alma Mater Digital Library
  - IT AperTo
  - FR Archimer, Archive Institutionnelle de l'Ifremer
  - IT Archivio Eprints Universita di Firenze
  - ES Archivo Digital UPM
  - GR Aristotle University of Thessaloniki
  - DE ask23 repository (lkw, hfbk hamburg, germany)
  - NL AUP publications
  - CH Bibliotheque Numerique RERO DOC
  - DE BieSON - Bielefelder Server fuer Online-Publikationen (University of Bielefeld, GERMANY)
  - UK Birkbeck ePrints
  - UK Bristol Repository of Scholarly Eprints (ROSE)
  - UK British Library Research Archive
  - SE DALEA - Högskolan Dalarnas elektroniska arkiv
  - HU Debreceni Egyetem elektronikus Archivumának (DEA)
  - DE DESY Publication Database
  - DE DIPP NRW
  - NL Dissertations of the Universiteit van Amsterdam



# DIGITAL REPOSITORIES INFRASTRUCTURE VISION FOR EUROPEAN RESEARCH



Home | **Repositories** | Network | Collections

Show Map

Legend ● Host ● Node

Karte | Satellit | Hybrid

- Hosting Nodes
- node1.p.driver.research-infrastructures.eu (146.48.87.214)
    - 8002 (IS\_Store)
    - 8003 (ResultSet)
    - 8001 (IS\_Registry)
    - 8007 (IS\_LookUp)
    - 8000 (IS\_SN)
    - 8009 (IS\_UI, CS\_UI)
  - node2.p.driver.research-infrastructures.eu (146.48.87.215)
    - 8003 (ResultSet)
    - 8009 (IS\_UI, CS\_UI)
    - 8010 (CS\_Engine)
    - 8008 (MS\_RO)
    - 8180 (A2)
  - node3.p.driver.research-infrastructures.eu (146.48.87.216)
    - 8003 (ResultSet)
    - 8006 (Publisher)
    - 8180 (Search, Recommendation, UserInterface, Community, UserProfiling)
  - node9.p.driver.research-infrastructures.eu (212.87.15.95)
    - 8003 (ResultSet)
    - 8004 (TextEngine)
    - 8180 ( )



Copyright © 2007 Driver Project | All Rights Reserved

https://admin.driver.research-infrastructures.eu/IS/NetworkMap#network



# *DRIVER Portal*



Digital Repository Infrastructure Vision for European Research

Driver is a collaboration, co-funded by the European Commission, to build a network of freely accessible digital repositories with content across academic disciplines.



Search the repositories portal



Find support



Learn about DRIVER



For repository managers

# Enhanced Publications

---

**An Enhanced Publication is a compound digital object that may consist of various heterogeneous, but related web resources**

*independent parts of the compound object might be distributed over several repositories, and located in different countries*







## *Milestone : LTP Connector*

---

A demonstrator of data/publication transfer from DRIVER Information Space to a long term archiving facility will be tested by relevant research communities.



# Archiving options

---

- Current Research Information Systems
  - Source university, institute
- Harvesting to aggregators
  - Disciplinary repositories
- Trusted digital repository
  - Designated responsibility



# *Challenges to archiving Enhanced Publications*

---

- Shared ownership
- Distributed deposit
- Scholarly perceptions on preservation and access
- Greater inherent risk
- Special handling requirements
- Organisational considerations



# Technology watch

---

- Connections with related infrastructure and digital library technologies
- GRID technology infrastructures
- DRIVER – CRIS interactions
- **Long-term preservation**
- Interoperability

**Acknowledgements:**  
**Barbara Sierman, KB**  
**Jens Ludwig, UGOE**



# Technology issues

## Enhanced Publications

- publications, datasets, websites, audio, video

## Characterised by

- file format, version, metadata, identifiers etc.

## Need

- structural metadata – with formal syntax  
declared semantics
- persistent identifiers

## *Status issues raised*

---

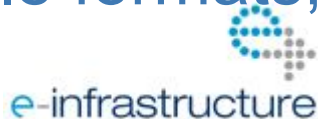
- When is an enhanced publication **ready** for long-term archiving?
  - Who is authorised to make that decision?
- When is it **complete**?
  - OAI-ORE Object model allows versioning
- **Legal** aspects
  - Root owner rights to allow harvesting and storage



# *SIP issues...*

---

- Representation Information
  - minimum set of metadata
  - significant properties
- Contextual information
  - designated community
  - future users
- Control over content
  - nestor, DRAMBORA, TRAC
  - file formats, software



## *Even more issues...*

---

- LTP archive rationalisation
  - Distributed functionality
- Digital curation contact
  - SIP definition
  - Preservation actions
- Persistent identifiers
  - Validity of ID method selected
  - Resolver maintenance





# Recommendations

***Consolidate root object and related objects in single LTP archive***

***Harvesting trigger mechanism to derive status of (all parts of) EP***

***Minimum set of structured metadata...***

***...including Persistent Identifier***

# *Recommended set of metadata for IR's*

---

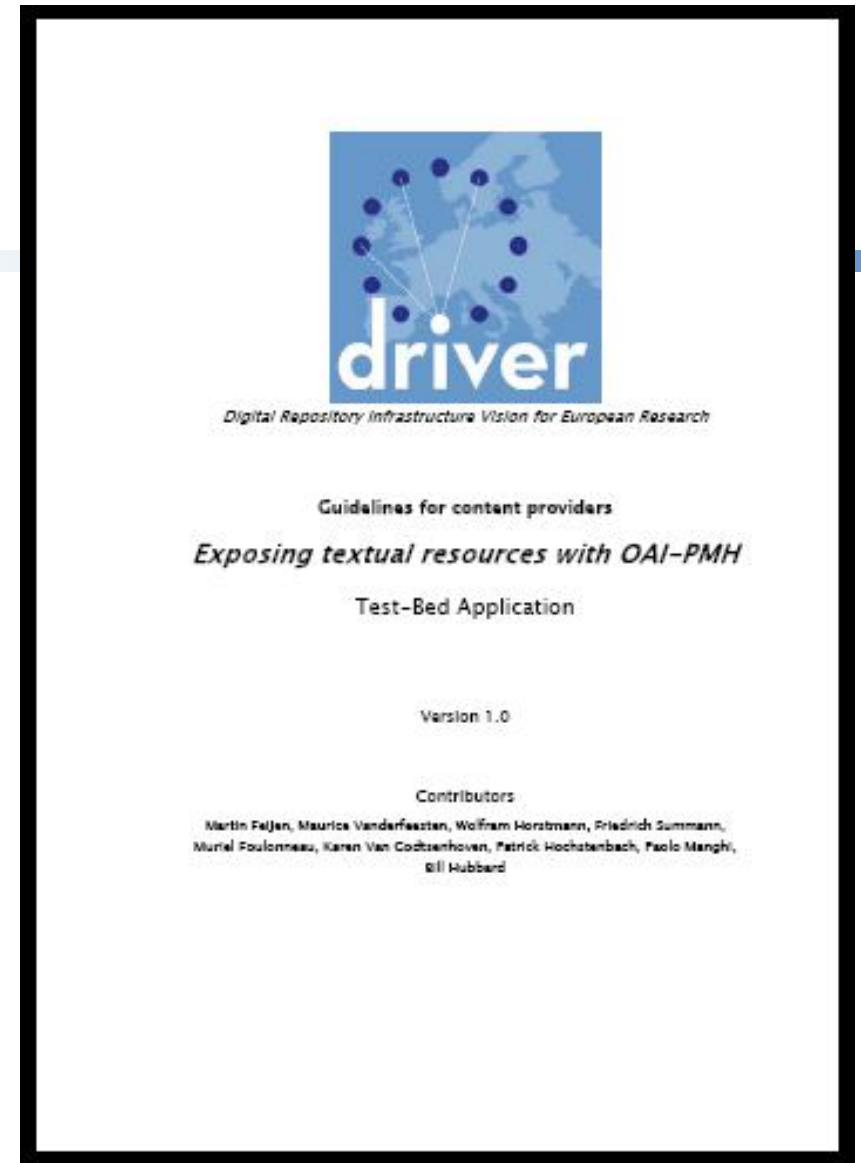
- File format
- File format version
- Date of creation
- Software name
- Software version



# DRIVER Guidelines

## Objectives

- Policy development
- Improved services
- Additional functionality
- Check conformance
- Validation



# Organisational challenges

---

- Cost estimates (Beagrie, Chruszcz & Lavoie, 2008)
- Rights management
- Digital curation
  - management policies
  - preservation procedures
- Co-ordination of activities
- Emerging memory centres

Current initiatives...





## e-Depot and digital preservation

Home

About the KB

### e-Depot and digital preservation

Information for international publishers

The e-Depot system (DIAS)

Digital preservation research

Contact

Publications and links

Bibliotheek



Exposities en collecties



### 'To ensure permanent access to digital objects'

The e-Depot is a digital archiving environment that ensures long-term access to digital objects which would otherwise be threatened by rapidly evolving software and hardware platforms as well as media decay. It is the dedicated archiving environment for the KB's national electronic deposit collection. In addition, it will include the Dutch web archive and digitised master images. In line with the international nature of information provision, the KB has extended its e-Depot services to publishers worldwide. The e-Depot is supported by sustained research and development efforts geared towards maintaining the integrity of stored digital objects.

- [Information for international publishers](#)
- [Information on the e-Depot system \(DIAS\)](#)
- [Information on Digital preservation research](#)
- [Contact](#)
- [Publications and links](#)



### New

Report: [Alternative File Formats for Storing Master Images of Digitisation Projects](#)

[Information for Dutch publishers \(Dutch only\)](#)



Portico offers a not-for-profit service that provides a permanent archive of scholarly literature published in electronic form, beginning with electronic journals.

# PORTICO

8,333,905  
preserved articles

66  
participating publishers

8,085  
committed titles

476  
participating libraries

5,603  
preserved titles

facts and figures

## Contact Us

## About Portico

## For Publishers

How to Participate

Archive License

## For Libraries

How to Participate

Archive License

## News and Resources

ALPSP Digital Preservation Survey

New Preservation Activities

Trigger Events

Holdings Comparison

Other Announcements

Portico Logo

## Audit the Archive

for designated librarians and publishers

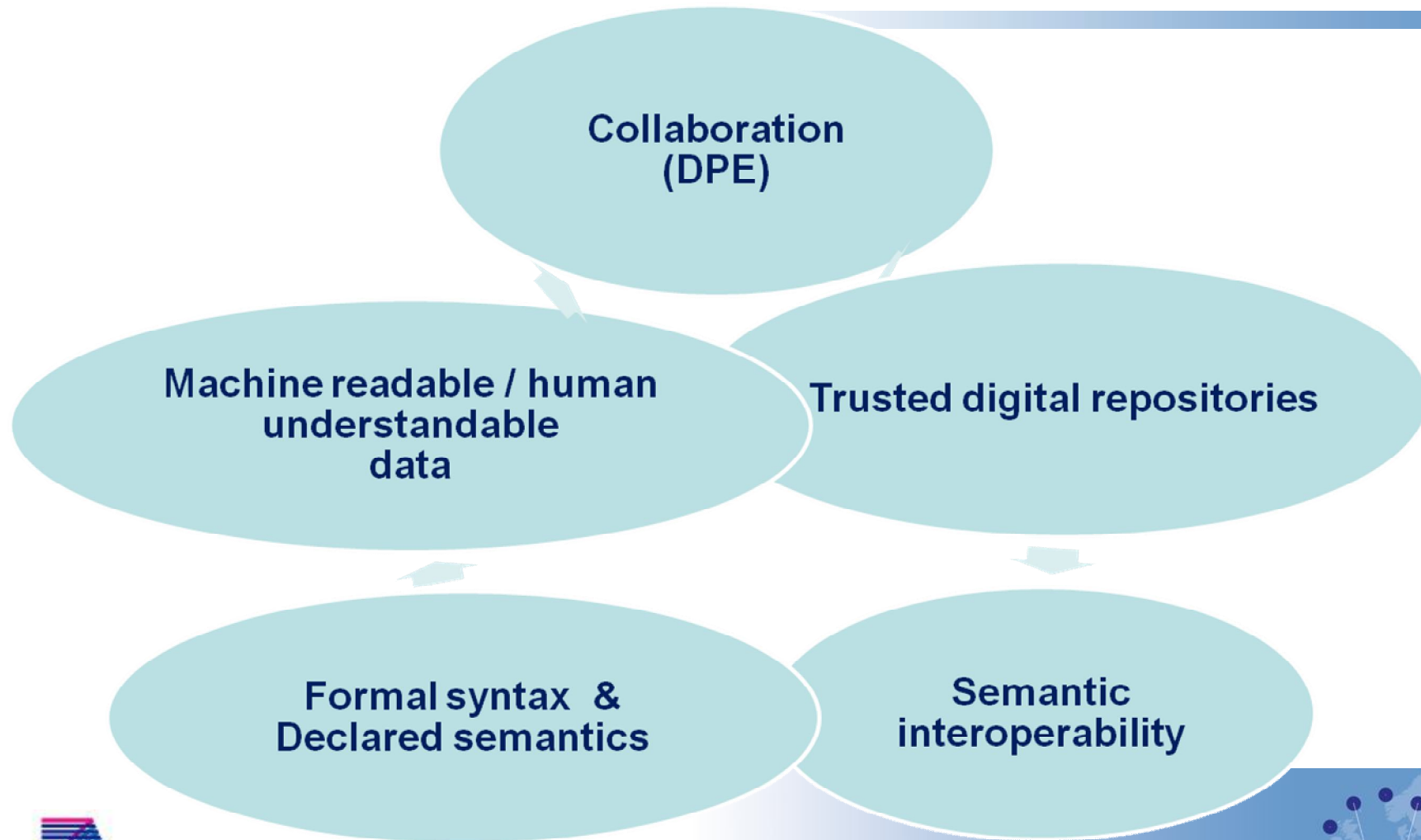
## Access the Archived Content

for researchers at participating institutions

Portico is committed to the digital preservation of scholarly literature published in electronic form. Initial support for Portico is provided by The Andrew W. Mellon Foundation, Ithaka, The Library of Congress, and JSTOR.

Copyright 2005-2008. Ithaka Harbors, Inc. | Privacy Policy | Terms of Use

# Conclusions



## Attribution-Noncommercial-Share Alike 3.0 Unported

### You are free:



to **Share** — to copy, distribute and transmit the work.



to **Remix** — to adapt the work.

### Under the following conditions:



**Attribution.** You must attribute the work in the manner specified by the author or licensor (but not in any way that suggests that they endorse you or your use of the work).



**Noncommercial.** You may not use this work for commercial purposes.



**Share Alike.** If you alter, transform, or build upon this work, you may distribute the resulting work only under the same or similar license to this one.

- For any reuse or distribution, you must make clear to others the license terms of this work. The best way to do this is with a link to this web page.
- Any of the above conditions can be waived if you get permission from the copyright holder.
- Nothing in this license impairs or restricts the author's moral rights.