



# The „many ways“ to interoperability

Outcomes and Challenges within  
DRIVER & OpenAIRE



Wolfram Horstmann  
Universität Bielefeld

# This Talk

- Discusses „Interoperability“ of DLs in the OA-domain (repositories)
- Exemplified by DRIVER- & OpenAIRE-initiatives
- „**Already history**“, story often told: but what can we actually **learn** WRT interoperability?
- Proposes a **multi-faceted** perspective on interoperability
- Suggests accomodating interoperability for **autonomy** in DLs by strict semantic focus

# **TERMS USED**

# „Interoperability“: Definitions

- “the ability of two or more systems or components to exchange information and to use the information that has been exchanged”

IEEE Standard Computer Dictionary. A Compilation of IEEE Standard Computer Glossaries : 610. IEEE, New York (1991)

- “the capability to communicate, execute programs, or transfer data among various functional units in a manner that requires minimal knowledge of the unique characteristics of those units”

ISO/IEC 2382:2001. Information Technology Vocabulary – Fundamental Terms

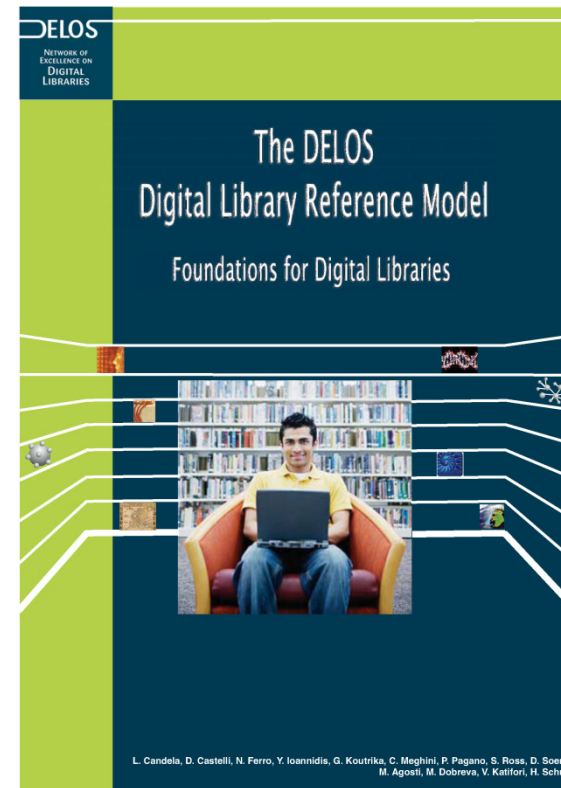
# „Interoperability“: Levels

- **Technical** agreements cover formats, protocols, security systems, so that messages can be exchanged.
- **Semantic**: Content agreements cover the data and metadata, and include semantic agreements on the interpretation of the information.
- **Organisational** agreements cover the ground rules for access, preservation of collections and services, payments, authentication, etc.

Also adopted by *IDABC: European Interoperability Framework for pan-European eGovernment Services European Commission, Luxembourg (2004)*, now introducing **political** and **legal** level (here subsumed in organisational).

# Digital Library

- DL.org Reference Model
  - Digital Library
  - Digital Library System
  - Digital Library Management System
  - Users
    - DL End users
    - DL Designer
    - DL System Administrator
    - DL Application Developer
  - Content
  - Functionality
  - Policy
  - Architecture
  - Quality



# „Digital Libraries“ & „Interoperability“

- Digital libraries are complex systems, intrinsically interdisciplinary [and heterogeneous]. They involve collaboration support, digital preservation, digital rights management, distributed data management, hypertext, information retrieval, human-computer interaction, library automation, publishing.
- **Autonomy challenge** -- degrees of freedom in DLs needed for representing heterogeneous requirements; standards for interoperation

# „Open Access“

- *“By open access, we mean its immediate, free availability on the public internet, permitting any users to read, download, copy, distribute, [export], search or link to the [materials], crawl them for indexing, pass them as data to software or use them for any other lawful purpose.”*

Squared brackets indicate changes of the original wording of the **Budapest Declaration**



# „Repositories“

- *„Vessels with tentacles and glands“*
  - *Here: for resources of scholarly communication*

# **OA-INTEROPERABILITY IN DRIVER & OpenAIRE**

# DRIVER & OpenAIRE

- DRIVER
  - „Digital Repository Vision for European Research“
  - Construction plans for a distributed DL backbone
  - Generic infrastructure development & production
- OpenAIRE
  - „Open Access Infrastructure for Research in Europe“
  - Support for European Open Access Policy (pilot phase)
  - Multi-faceted application of DRIVER-results
- Ultimate goal: Multi-sited OA-DLS system
- > Perfect example for OA interoperability challenges with complete walkthrough coverage of all levels

# e.g. semantic interoperability

- DCMES as content agreements
- OAI-PMH „standard“ as implementation
  - Addressing technical (transfer) as well as semantic (content) interoperability – with known problems

## e.g. Field Contents heterogeneity: <dc:type>

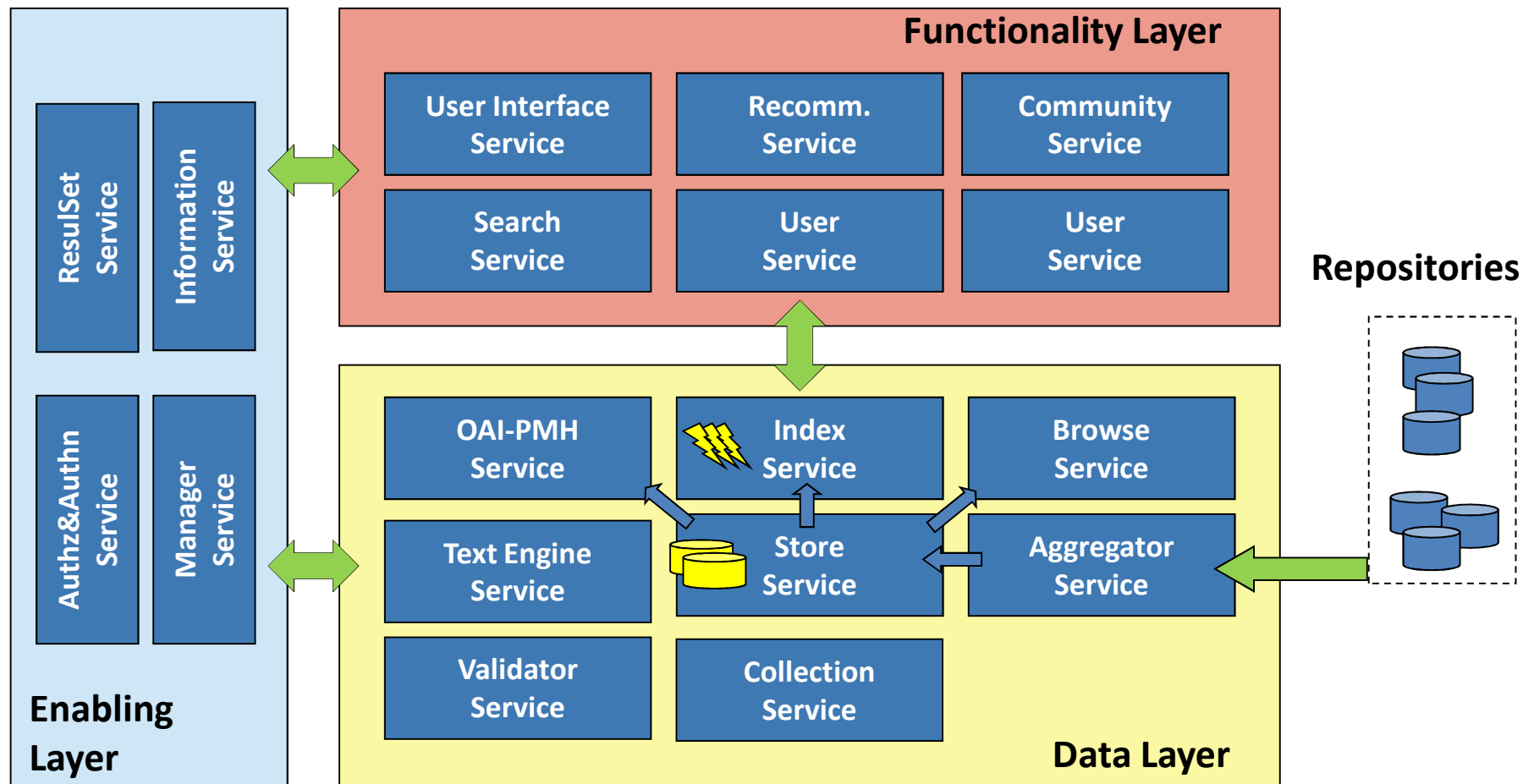
Text - 2597764	NULL - 8082
text - 2540080	Clippings Text  - 5004
TEXT - 1128327	This record contains content - 1473
Dataset - 580448	Dark Item - 1145
Image - 522643	Photographs; Mosaics; - 1
image - 514585	1 photographic print mounted on cardboard: b&w; 22 x 16 in. - 1
Electronic Thesis or Dissertation	80534 - 1
Tese ou Dissertacao Eletronica	Section V Flow Control - 1
article - 345093	Acrylic on board: 63 x 114 cm. - 1
Article - 337314	79508 - 1
Artikel - 310858	Colored graphite on paper: 47 x 63 cm. - 1
Other - 219331	texts, vocabulary ('stories, monologues, word list' - 1

# e.g. semantic interoperability

- DCMES Interoperability solutions
  - Simplicity and degrees of freedom make it applicable and widespread, at least a common denominator
  - OAI-PMH readily packs in into a technical solution
- DCMES problems
  - Low specificity causes semantic variation
  - Integration with OAI-PMH prevents healing directed only to one aspect
- DRIVER / OpenAIRE always in the dilemma to address a critical mass and to sort out the problems
  - Alternative not building on OAI-PMH endangers to loose standards at all and leave materials unused

# e.g. technical interoperability

- D-NET software for **multi-sited operation** at the level of **service interoperability**



# e.g. technical interoperability

- D-NET: Interoperability Solutions
  - Division of labour in operation and dev. between partners (Athens, Bielefeld, Pisa, Warsaw)
  - Functional responsibilities: e.g. aggregation, UI etc.
  - Versatile construction of various DLs w. service reuse
- D-NET: Interoperability Problems
  - Transitions between technical and organisational interoperability introduce overheads
    - MoUs, SLAs, Specialization, local interests
  - Threshold for uptake too high due to the multiple problem domains to be solved for service interoperability
- DRIVER/OpenAIRE always in the dilemma between controlling sites and expanding services to partners

# e.g. organisational interoperability

- Interoperability policy compliance requires organisational implementation and includes technical and semantic interoperability
  - DRIVER guidelines as OAI-PMH „interpretation“
    - „Become part of the network“ through validation
  - OpenAIRE guidelines as European OA-policy interpretation („SC39“)
  - OpenAIRE National Open Access Desks and DRIVER helpdesk as support structures est.



# e.g. organisational interoperability

- Guidelines: interoperability solutions
  - Enhances adherence to standards
    - Also implementation in platforms (DLMS)
  - Provides glue between all interoperability levels
- Guidelines: interoperability problems
  - Introduce compliance enforcement
  - Cause extra work at all levels (repositories, support etc.)

**PUTTING IT TOGETHER**

# Multi-Level Interoperability

- Interoperability levels not really separable
- Technical: Operation of distributed repositories and services
  - e.g. http, OAI-PMH (transfer), Web-Services, Platforms
- Semantic: Harmonization of distributed content
  - e.g. OAI-PMH (DCMES), Guidelines, Profiles/Rules
- Organisational: Guidelines compliance, policy implementation
  - Guidelines, Helpdesk/NOADs, Networking
  - Policy (EC-OA-policy), Legal (Copyright, SC39)

# Observations WRT Levels

- Ambiguity of interoperability levels
  - OAI-PMH integrates technical and semantic level
  - Guidelines span all interoperability levels
  - Policy and legal level indeed can be added but also introduce more complexity
- Interpretation: „Direction“ of interoperability unclear
  - Reducibility: Organisational wouldn't exist without technical and technical not without semantic but not vice versa > semantic is the core interoperability level > „Library-Content here to stay“
  - Interoperability in heterogeneous systems benefits from organisation/policy
- Suggestion
  - Interoperability measures should address semantic level

# Summary & Conclusion

- Interoperability is **multi-leveled** but rather a **network than** a layer-model is required
- **Simplicity** is the clue for uptake of **standards**
- **Semantics** are **core** DL interoperability **challenge**
- **Content** (semantic) and **functionality** (technical) to be **decoupled** as far as possible Look deeper into the DL.org-RM?
- Focus on semantic interoperability allow **DLs/repositories** the **autonomy** they need for meeting heterogeneous requirements

# Tentative Outlook

- Semantic interoperability can be enhanced...
  - (**not** immediately by **semantic web**; probably still too powerful and versatile for DLs to encode content)
  - through **rigid terminology** standards applied on the **local site** (DL/repository)
    - E.g. authority files for Journal Titles, Names, Institutions, ...
  - through **adaptive algorithms** (unsupervised, dynamic, fault-tolerant...) that tame semantic variability on the **central site** (specialized service providers)
- **Separation of „semanticists“ & technologists?**

Thanks.