Interoperability is not just about Content & Functionality

Yannis Ioannidis   **University of Athens, Hellas**

**DLF3 @ ECDL 2008, Aarhus, Denmark, 18 September**

Information Space

- ~~*Digital library*~~ : A (potentially virtual) organization that comprehensively collects, manages, and preserves for the long term rich **digital content** and offers to its **user** communities specialized **functionality** on that content, of measurable **quality**, and according to prescribed **policies**.

Information Mgmt

- ~~*Digital Library System*~~: A software system that is based on a (potentially distributed) **architecture** and provides all functionality that is required by a particular Digital Library.

"Artists"'s Rendition of DL (Sys)

DELOS
AN ASSOCIATION FOR
DIGITAL
LIBRARIES

Policy

Quality

User

Functionality

Content

Architecture
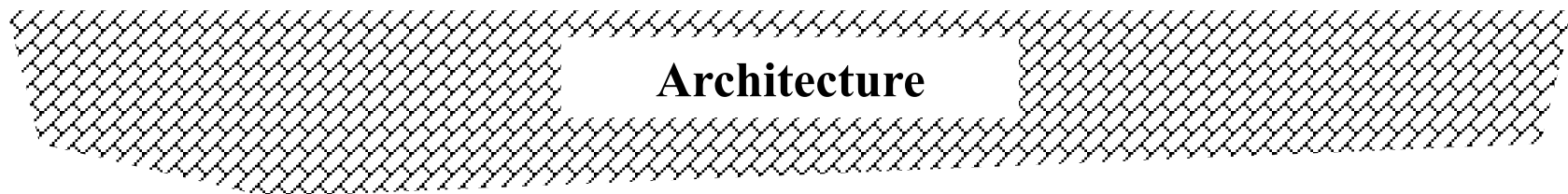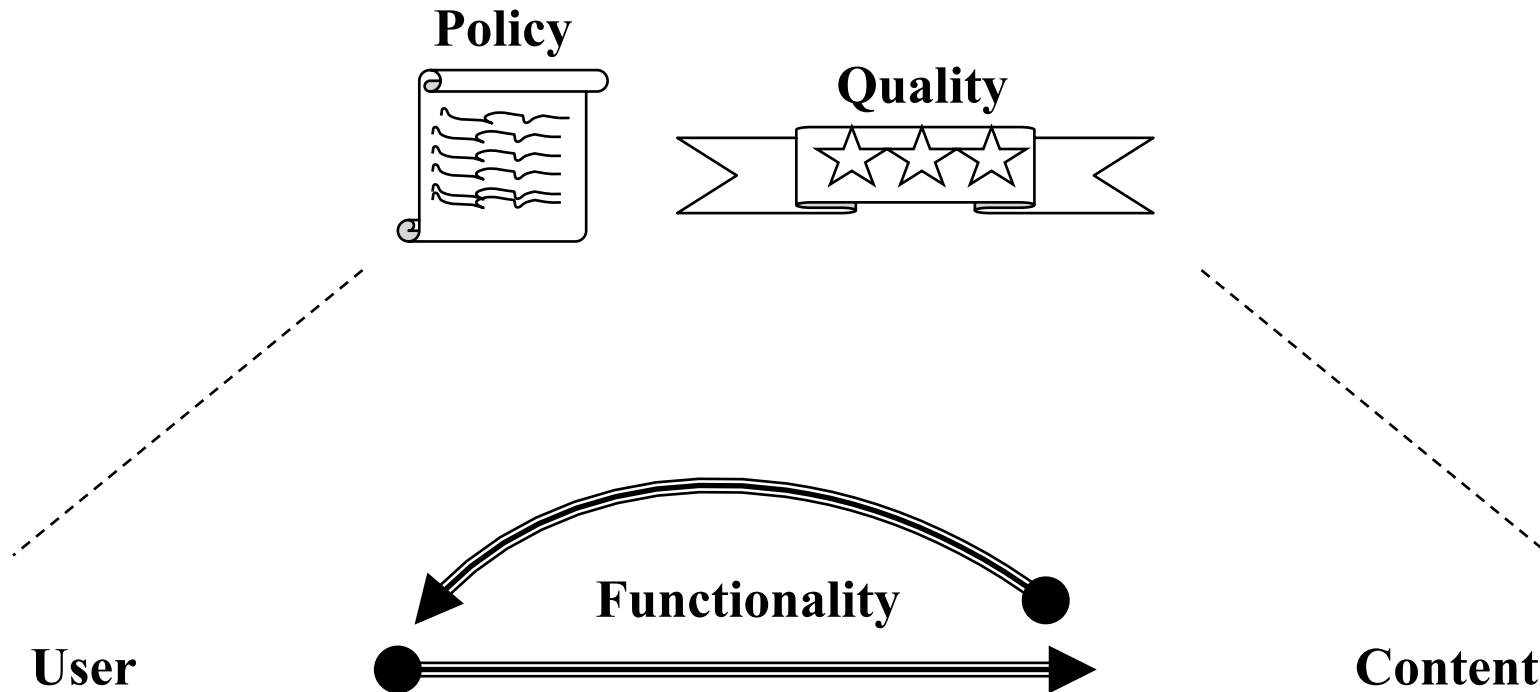
DL Interoperability Background

- *Distributed Heterogeneous Digital Libraries*

- *Information in all forms*

- RM a future unifying factor, but *interoperability* crucial for
  - legacy systems
  - reconciling different future approaches

- Why/when: integration, composition, matching, mapping, deduction, and activation
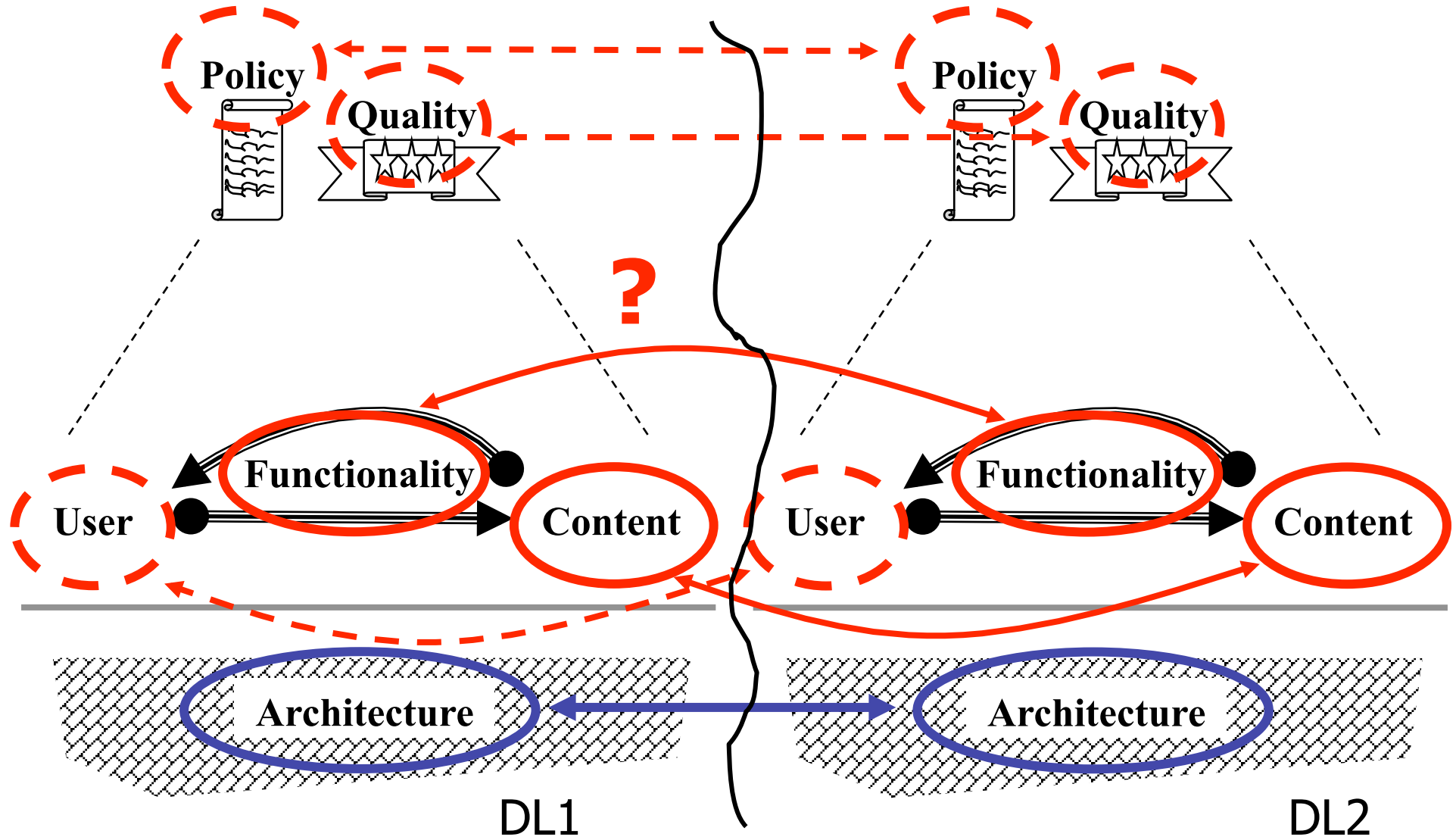
Biological Data Banks (DLs)
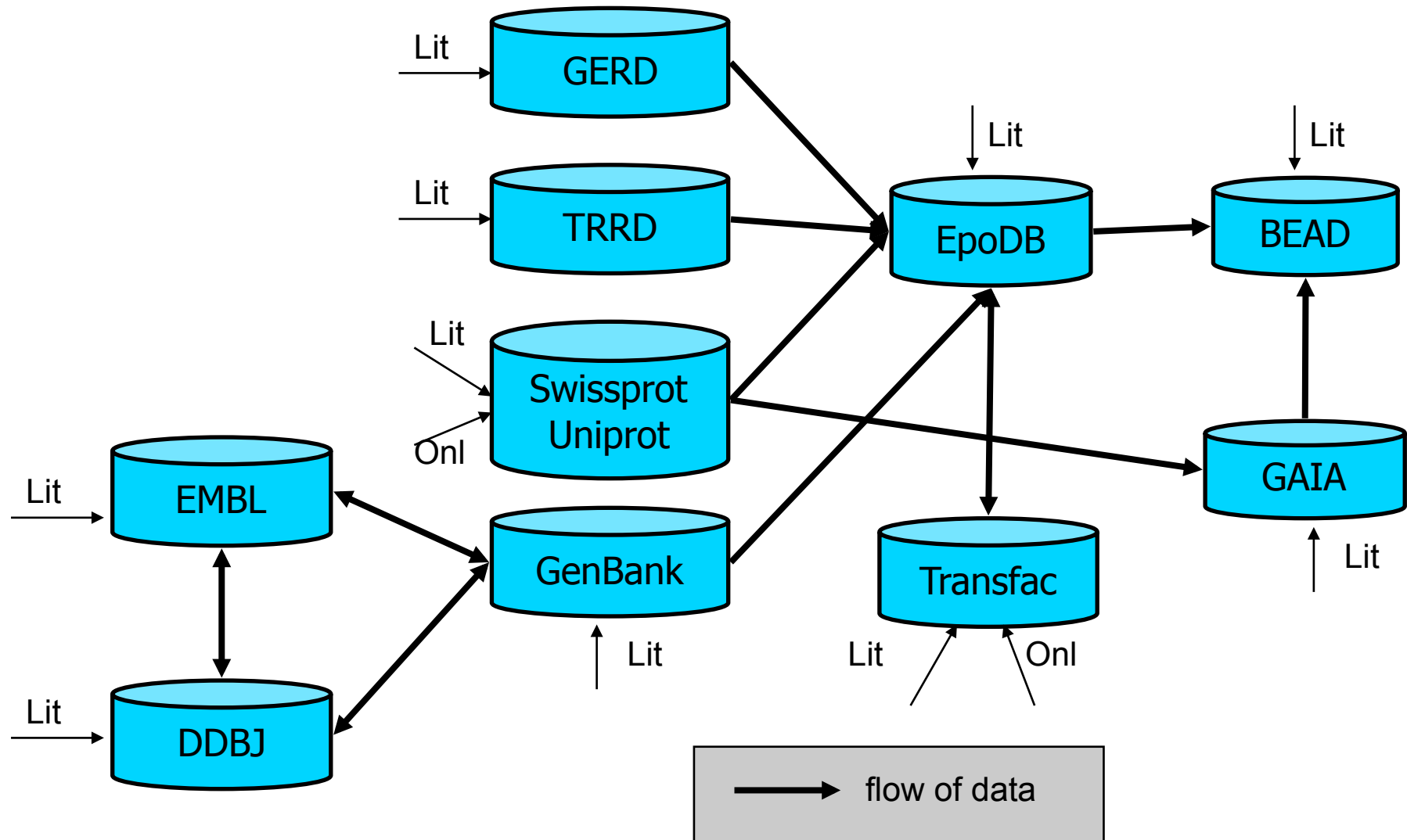
DL Interop/ility State of the Art

- Research has focused on common tools, enabling technologies, and standards
- Interoperability continues to be a difficult problem
- Understanding interoperability issues requires operational experience from large-scale deployment of DL systems

Interop/ility Abstraction Levels

- ## Superficial
  - Common tools and interfaces for navigation & access
  - Human intelligence for content coherence

- ## Syntactic
  - Common metadata models and object transmission protocols and formats for limited coherence
  - Supplementary human interpretation

- ## Semantic
  - Consistent and semantically coherent access to all digital objects and services
  - Federating/mediating software for site-by-site variations
  - "No" human involvement

Interoperability

DELOS
AN ASSOCIATION FOR DIGITAL LIBRARIES

Policy

Quality

?

User

Functionality

Content

User

Functionality

Content

Architecture

Architecture

DL1

DL2

Interop/ility is a Broad Concept

- Not just coherence among passive object repositories

- All services exposed in an interoperable fashion

- Not only data heterogeneity or software mismatch

- *Heterogeneity* the norm in all six fundamental RM concepts → All raise interoperation issues

- Nonexistent state of the art for many concepts

- Rudimentary state of the art for the rest

Content Interoperability

- Most common form of interoperability
- Almost exclusively assumed by most
- Reconciliation of primary data, schemas, metadata, and all other superimposed information, e.g., annotations
- Significant work for 30 years, mostly for structured and/or semi-structured alphanumeric data
- Key methodologies and tools (e.g., wrappers and mediators, use of ontologies)
- Semantic interoperability for unstructured and cross-medium content?

Functionality Interoperability

- Selection of services based on their semantic signatures

- Service composition for creation of complex workflows

- Service equivalence identification for optimisation

- Good work on service interoperability in recent years

- Several standards (e.g., WSRF for web services)

- Mostly syntactic, failing to capture internal functionality semantics realized by these services

User Interoperability

- Collect, exchange, and integrate information on users: profiles, preferences, and access rights

- User migration across systems, local or distributed operation

- Same services and system behaviour

- Limited work on access rights treats them as content and reuses content interoperability techniques

- Profile and preference fusion for personalization and recommendation services?

User Interoperability Example

- Donatella@DL1:
  - "Research Infrastructures" → 0.9
  - "Swimming" → 0.3

- Donatella@DL2:
  - "Research Infrastructures" ≤ "Swimming"

- Contradicting or Incomparable?
- Context dependent?
- Reconciliation approach?
  - E.g., More info and stronger statement in DL1

Quality Interoperability

- Quality mostly used for external system evaluation and internal optimisation
- Reconciliation of different metrics of quality for globally optimal behaviour
- Multi-objective optimisation and pareto optimality?

Quality Interoperability Example

- Query: Documents on folk jewelry with bird decorations or pink stones
- DL1:
  - 300 objects (200 bird, 100 pink) in 30 seconds
- DL2:
  - 400 (100 bird, 300 pink) objects in 60 seconds
  - If >60 seconds Then 600 objects (SLA)

- DL1, DL2 better at different things
- DL2 has guarantees
- Reconciliation approach?
- How are SLA terms incorporated?

Policy Interoperability

- Policies expressed as rules and regulations on content, functionality, and users

- Policy interoperability is rule integration, composition, deduction, and activation

- Rules in different languages, based on different logic

Policy Interoperability Example

- ## DL1:
  - No more than 2 videos streaming simultaneously per user
  - Only open access documents

- ## DL2:
  - Local users have priority over remote users
  - Only OAI-PMH documents (subset of OA)

- Dominating or integrated (how?) rule?

If you think …

If you think interoperability is difficult

You are WRONG!

It is VERY DIFFICULT!
It is (ALMOST) IMPOSSIBLE!

If you think interoperability is about content/functionality

You are WRONG! Again!

It is about ALL 6 CONCEPTS of the DELOS RM!
It is about (ALMOST) EVERYTHING!

So, …

Interoperability is a dirty job

Interoperability is a broad job

Interoperability is a complex job

Interoperability will never be solved completely

Interoperability is a critical job; someone has to do it

Interoperability is a partitionable job

Interoperability is a fun job

Interoperability must be solved even approximately

Conclusion

AN ASSOCIATION FOR DIGITAL LIBRARIES

The DELOS Reference Model
an excellent foundation
for interoperability work