



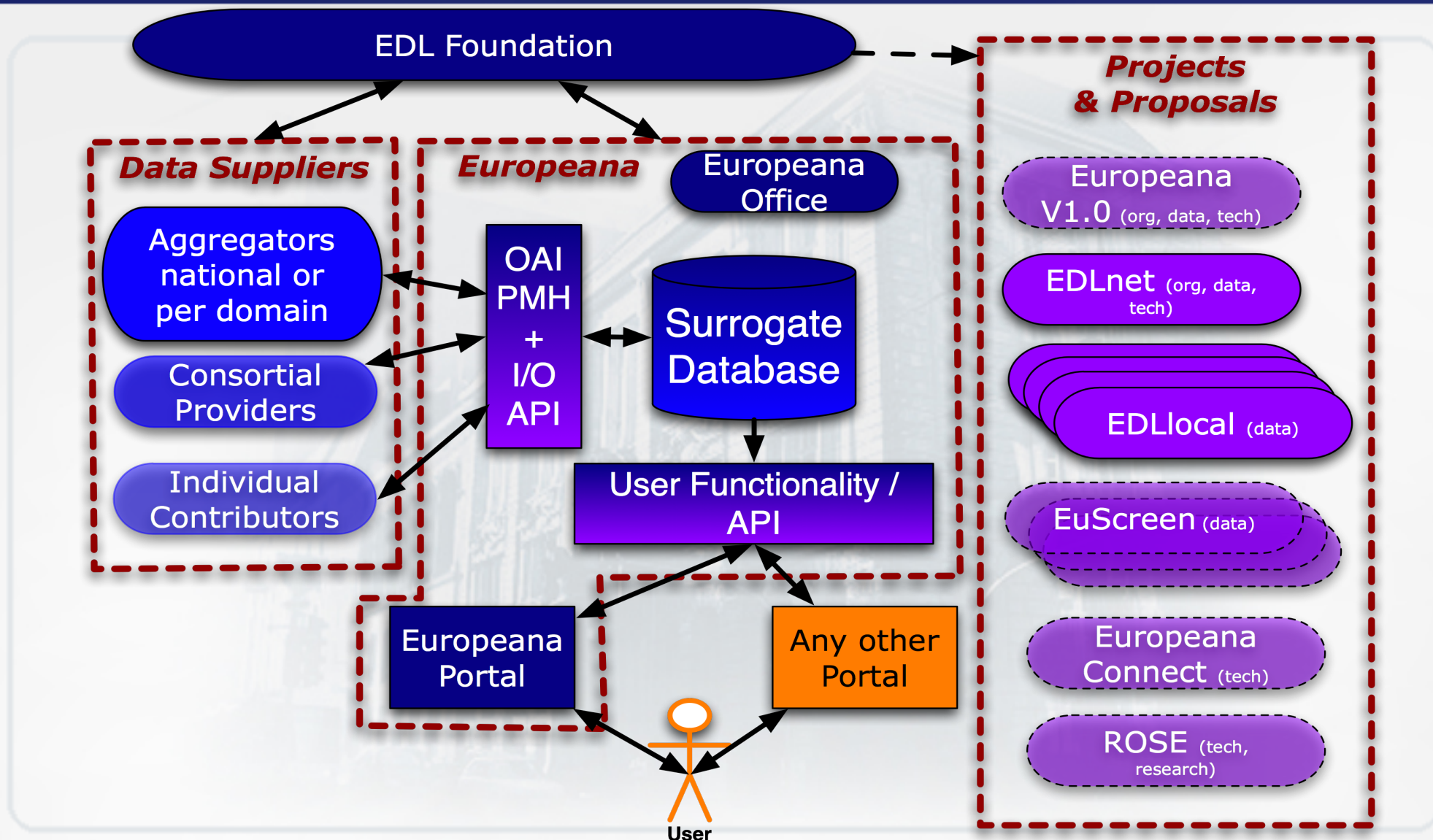
Making Europeana Interoperable (the) Six Most Challenging Issues

Prof. Dr. Stefan Gradmann

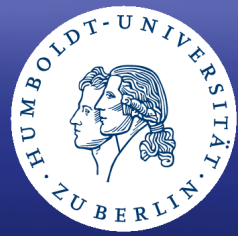
Humboldt-Universität zu Berlin / School of Library and Information Science

stefan.gradmann@ibi.hu-berlin.de

Europeana 1.0 Big Picture



Overview

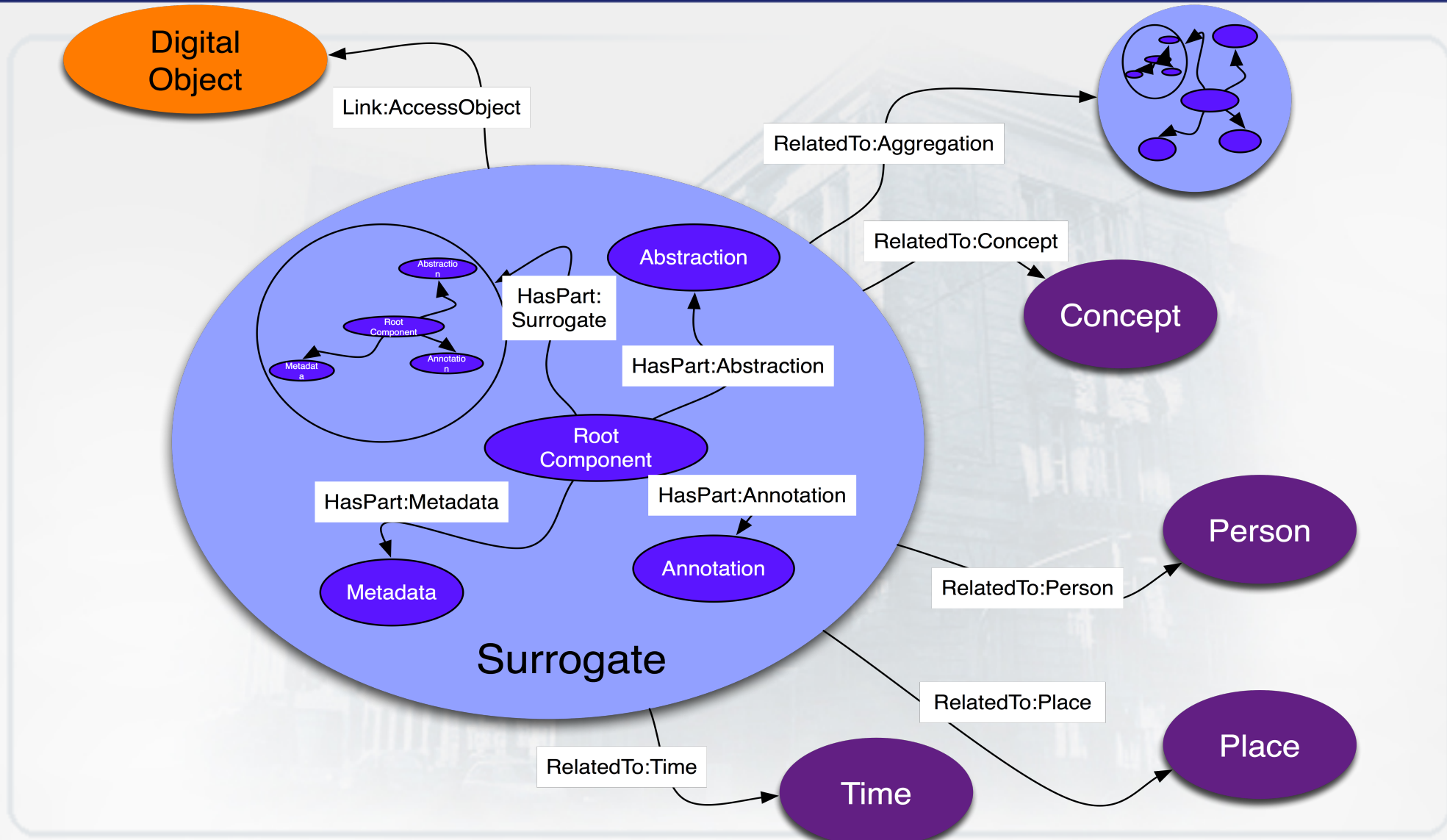


- Six Issues:
 - Object Model
 - Functional Primitives
 - Authentication based services
 - Semantic Mapping and Alignment
 - Multilingualism
 - Data quality
- Finding the appropriate abstraction level
- ... to what end?

Six Challenges (DFL3@ECDL2008) / Århus 18/09/2008



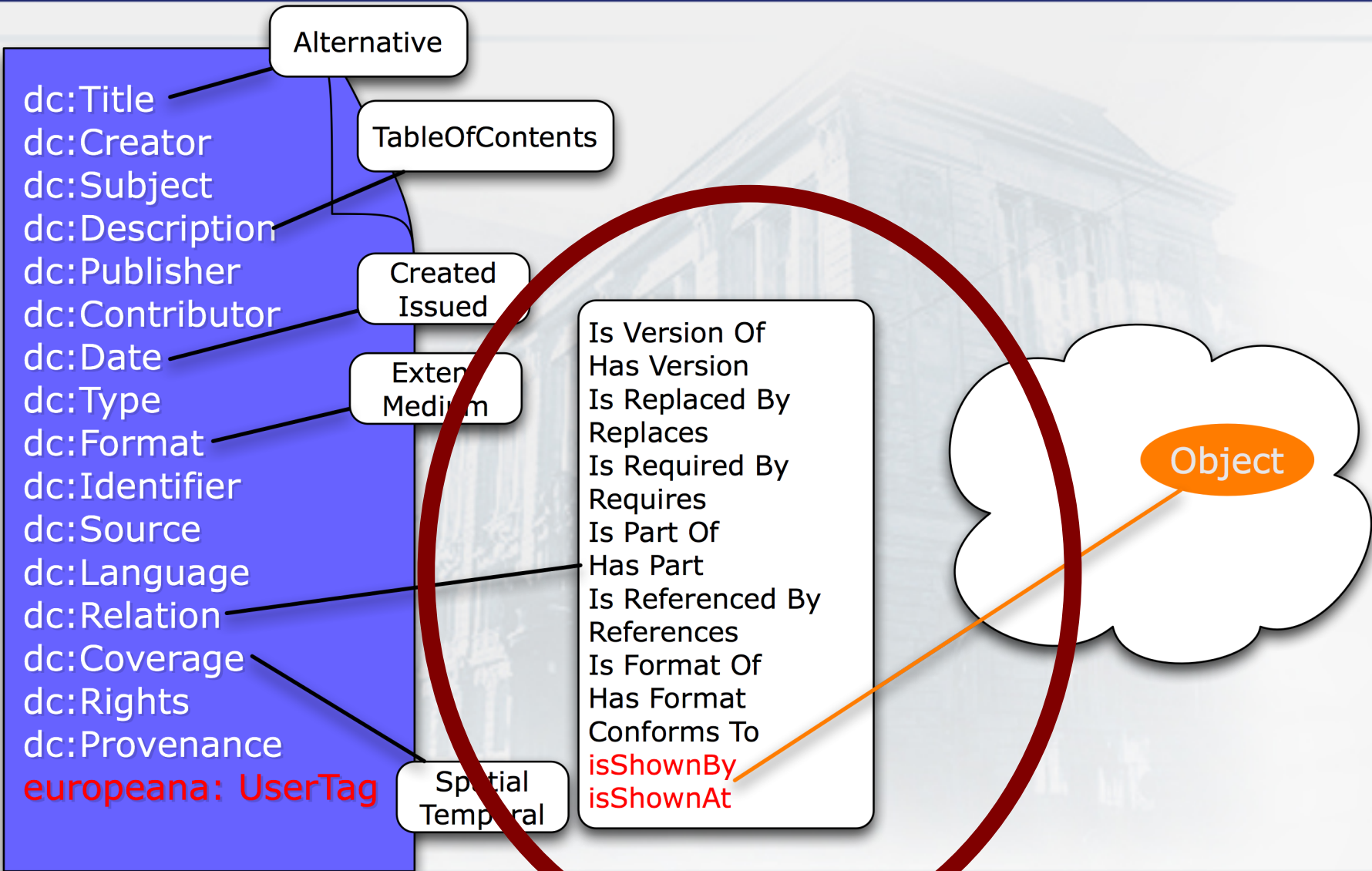
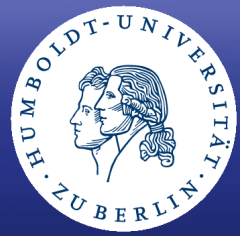
Objects: Europeana Perspective



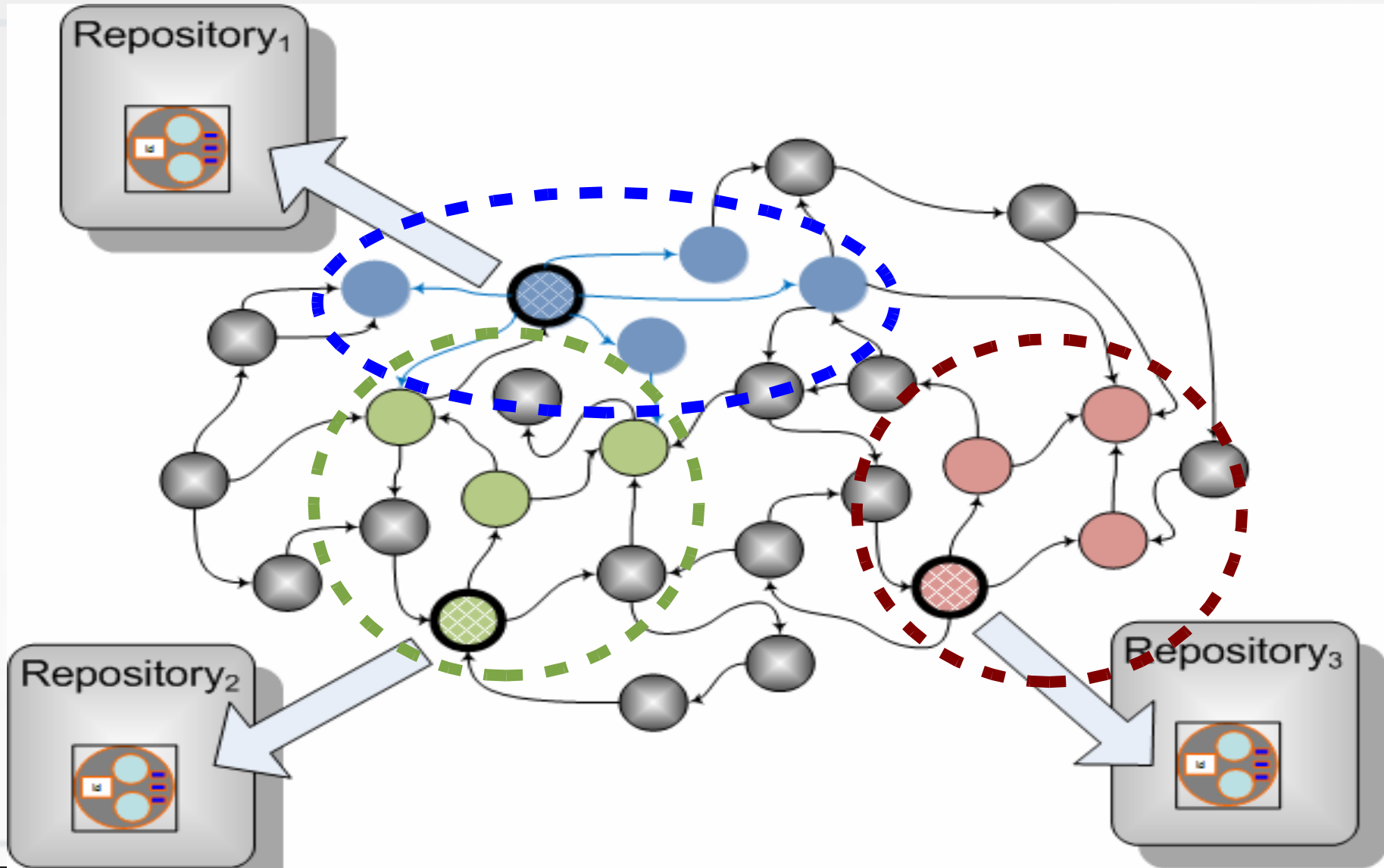
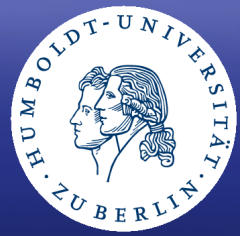
Six Challenges (DFL3@ECDL2008) / Århus 18/09/2008



Objects and Relations

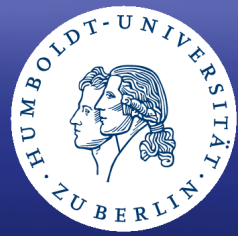


Objects: OAI-ORE Perspective



SIX CHALLENGES (DI LO@LODL2000) / MITUS 10/09/2000

Document / Object Model: Beyond Metadata

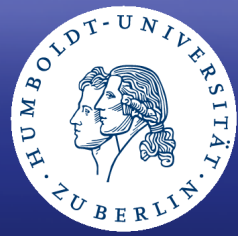


- Things are relatively simple with digitised objects - but what about born digital ones?
- What constitutes the boundaries of a complex digital entity?
- What are its basic constituents?
- What are typical relations between components within an entity?
- What are typical relations between autonomous entities?
- Framework candidates include
 - ORE
 - CIDOC-CRM
 - DCMI Abstract

Six Challenges (DFL3@ECDL2008) / Århus 18/09/2008



Modelling of Non-textual Media



- Media types
 - still images
 - audios
 - videos
 - 3D objects
- Too many standards: Plethora of content models and formats!

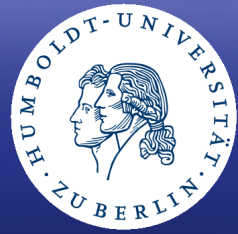
Six Challenges (DFL3@ECDL2008) / Århus 18/09/2008



Functional Primitives

- What functional primitives can we distinguish beyond 'read' and 'write'?
- On which level of abstraction?
 - Bytestream – as in iRods microservices?
 - Content management – as in JCR 283?
 - Higher level function blocks – as in DRIVER?
 - Web Service Level – as in JISC Information Environment?

Authentication based Operations



- Authorisation
 - closely related to functional primitives!
 - what kinds of operations do we want to distinguish in terms of authorisation?
- Access control / licensing
- Personalisation

Six Challenges (DFL3@ECDL2008) / Århus 18/09/2008

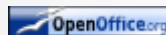


Semantic Alignment & Mapping



- Terminology alignment and identification of preferred concepts
- Ontology matching
- Ontology mapping
- Mapping evolution
- ... the underlying issue is conceptual interoperability => next & last challenge

Six Challenges (DFL3@ECDL2008) / Århus 18/09/2008



Multilingualism

- 'European' essential: **linguistic diversity is vital for Europe!**
- Four levels of multilinguality:
 - **Interface** - *elementary and mostly available*
 - **Browsing via a common multilingual ontology** mapping onto versions for each language. Invest in SKOSification!
 - **Search on a monolingual baseline**
 - Monolingual search for multiple languages (L1 => L1)
Invest in tokenizers (essential), stopword lists (essential), stemmers or morphological analysers (essential), decomposers (optional), phrase recognizers (optional), named entity recognizers (optional)!
 - Simple cross language search (L1 => Ln)
Invest in the above plus automated query and document translation!
 - **Full multilingual search** (Ln => Ln)
Holy grail and still in a distant future only!
- Multilingualism beyond linguistic mimicry requires conceptual interoperability!

The Nasty Bit: Data Quality

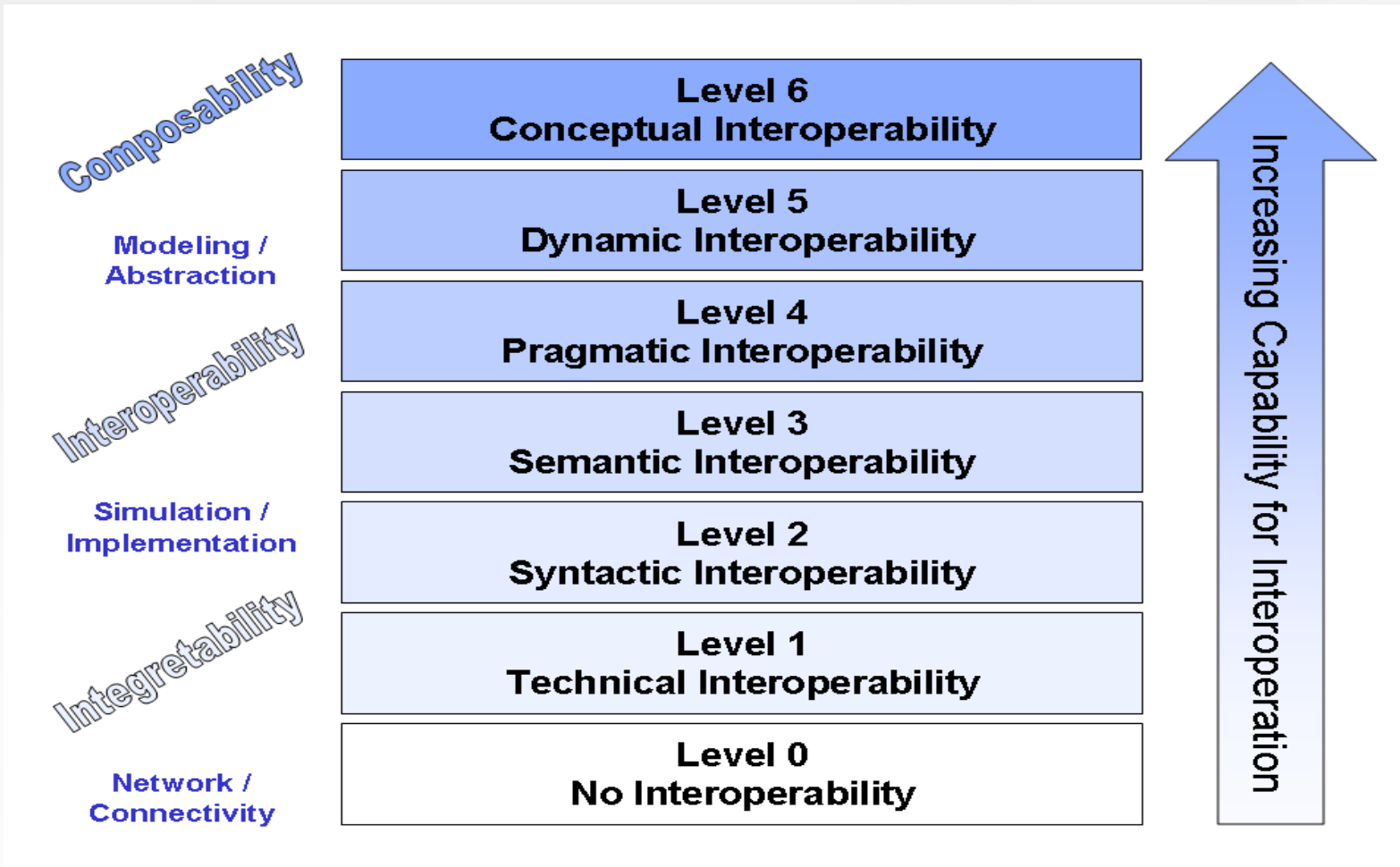
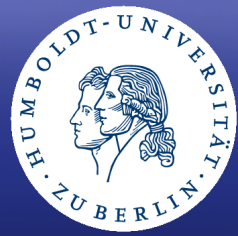


- A perfect framework combining
 - object modeling
 - well understood functional primitives
 - including authorisation methods
 - as well as using aligned semantic elements
 - and fully multilingual

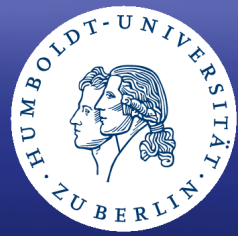
may still result in a dramatic lack of interoperability:

- When operating on 'dirty', heterogeneous data!
- This is a truth both trivial and critical

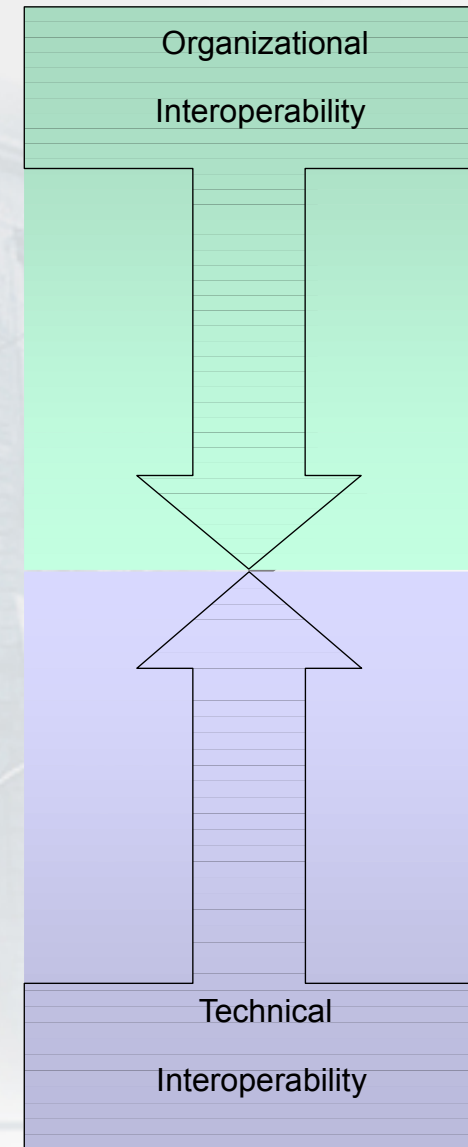
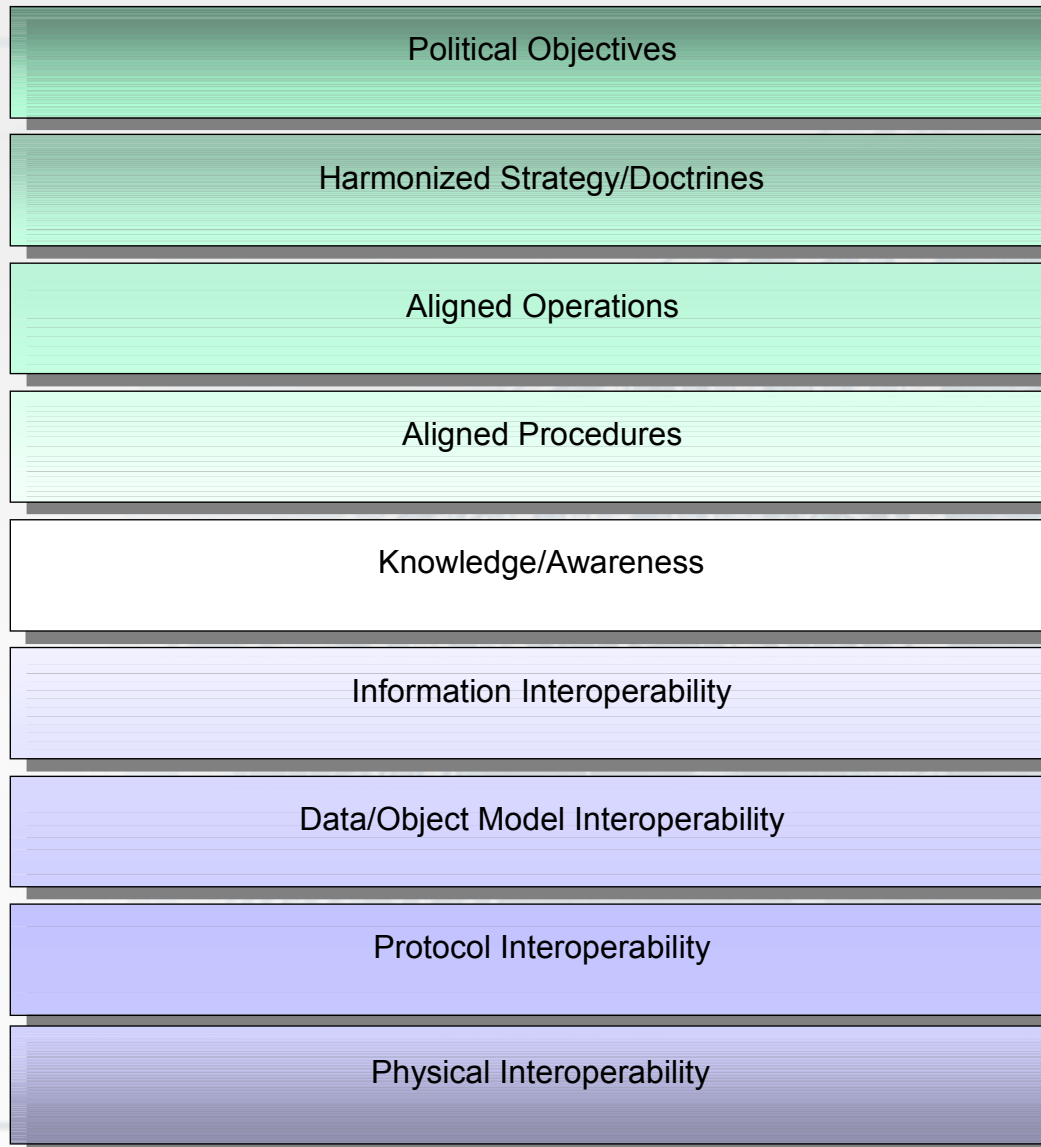
Abstraction Layers (Tolk2006)



Abstraction Layers (Legrand2006)



Layers of Interoperability

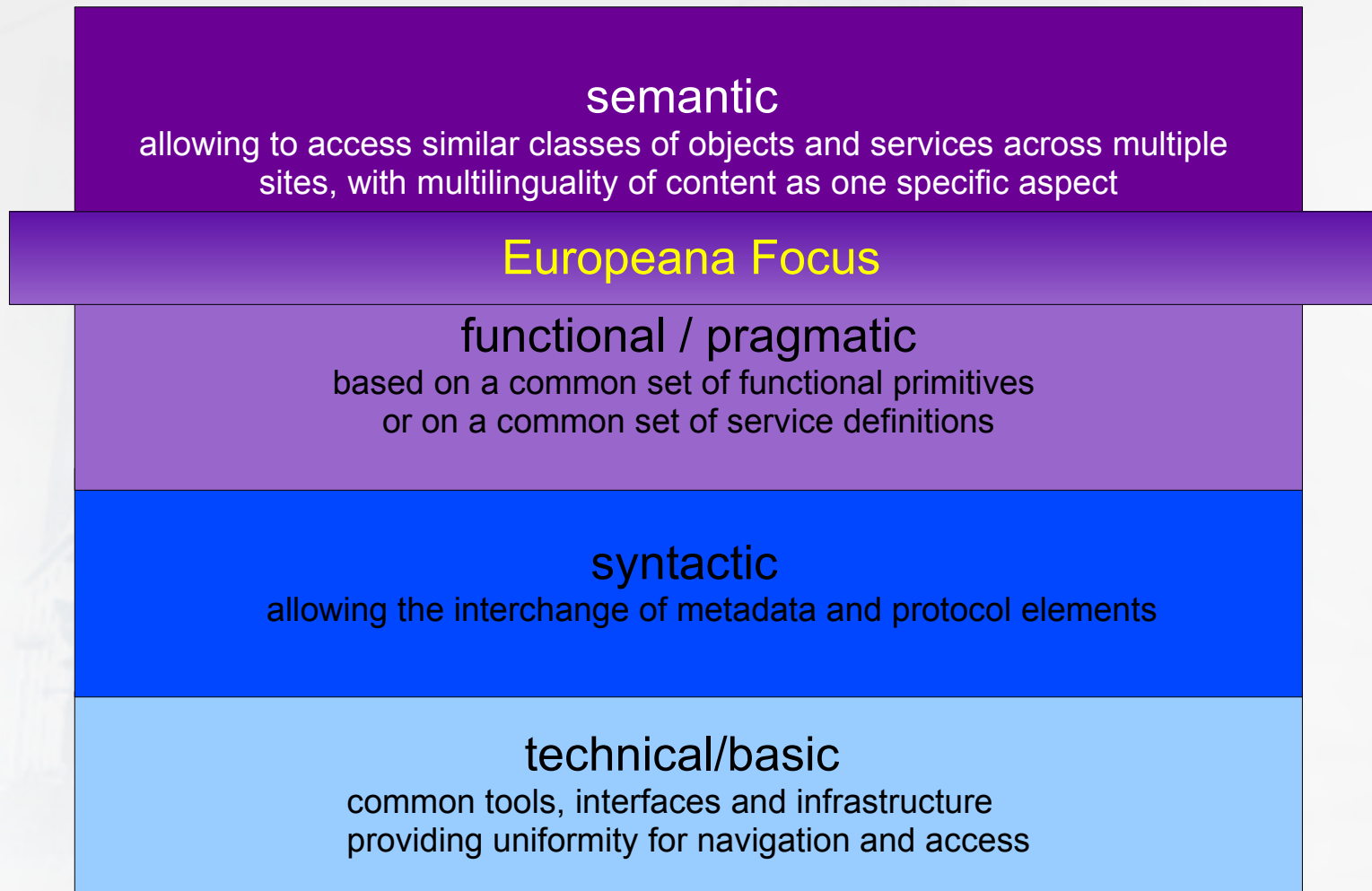


Six Challenges (DFL3@ECDL2008) / Århus 18/09/2008



Abstraction Layers

Abstract



Concrete

Six Challenges (DFL3@ECDL2008) / Århus 18/09/2008

... to what end?

- “Interoperability is the capability to **communicate, execute programs, or transfer data** among various **functional units** in a manner that requires the user to have **little or no knowledge** of the **unique characteristics** of those units.”
(ISO/IEC 2382 Information Technology Vocabulary)
- This is required both
 - for interoperation
 - and for making knowledge access platforms sustainable/persistent
- And in this sense Europeana is sharing the six challenges touched upon with all communities present in this workshop!
- And this requires continued work on the
... Reference Architecture