

Metadata in the wild

– the laborious creation of a new international standard for describing learning resources

Tore Hoel

Oslo University College

vice chair CEN WS-LT

co-editor ISO MLR Part 5 Educational

Vienna 2009-11-12



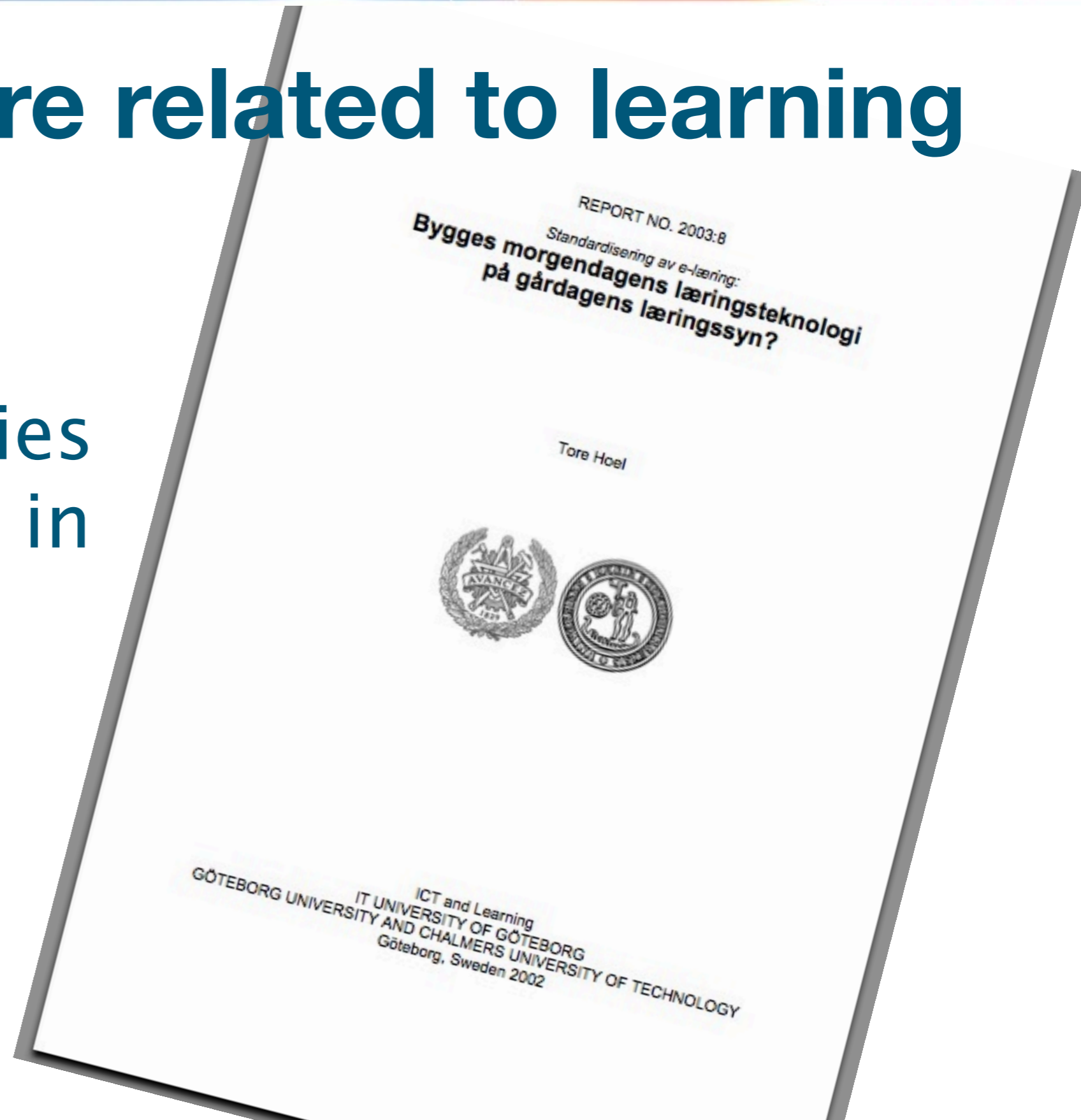
www.hio.no

Me (and where I am, related to metadata)

- Vice chair CEN WS-LT
- Work Package co-leader Dissemination & Roadmapping ICOPER
(EU funded project on competency-based learning)
- Co-editor MLR Part 5 Educational of ISO/IEC JTC1 SC36 WG4
- Worked on metadata since 2002 (same year WG4 was established)

Standards are related to learning theories!

- 2002: What learning theories are embedded in IEEE LOM and SCORM?



Me, You, Us – in context

Brian Lamb at WCET



Brian Lamb
Emerging Technologies Discoordinator, OLT UBC <http://olt.ubc.ca/>
Vancouver Canada

Me, You, Us – in context

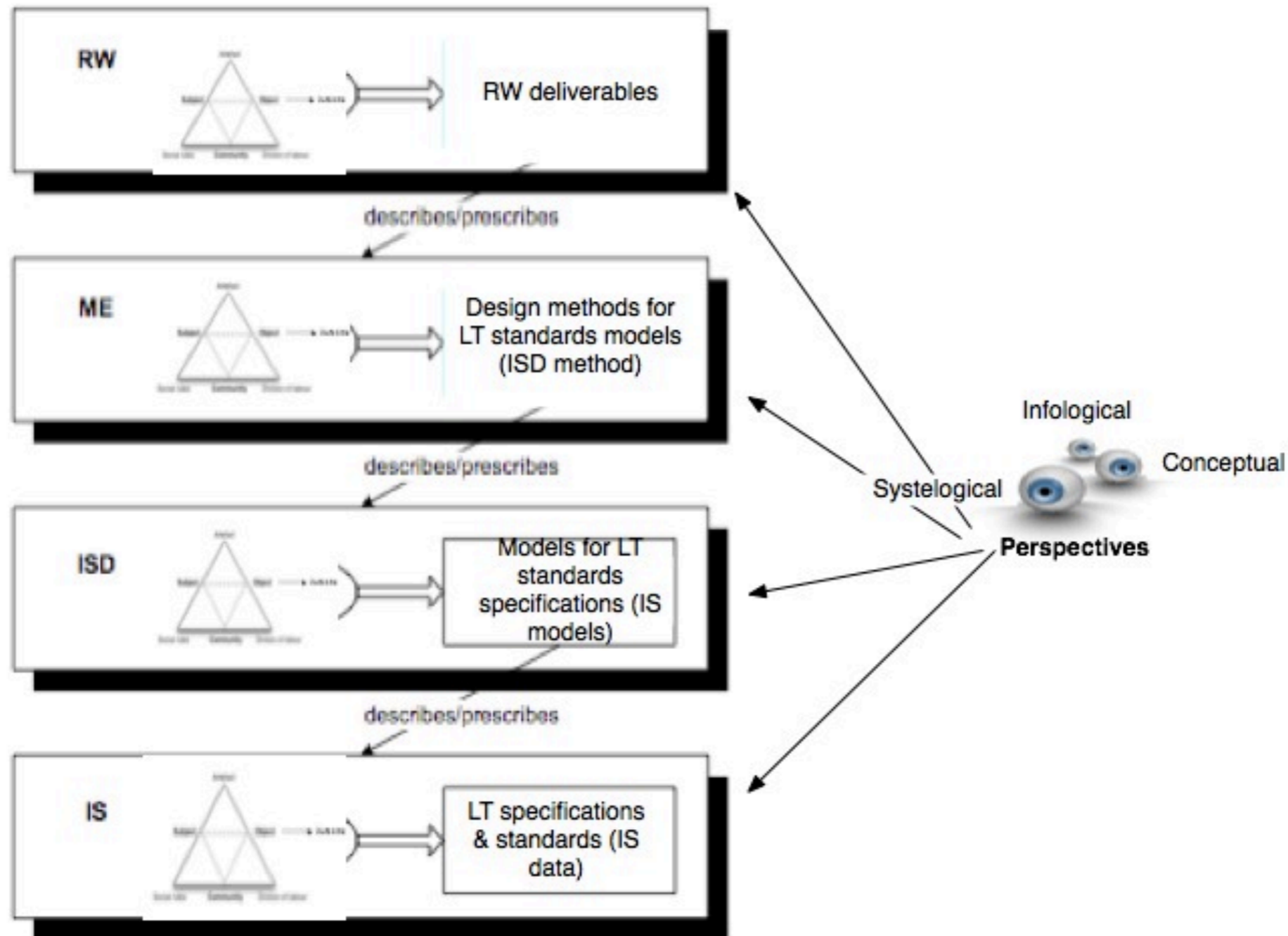
- Extinct concepts:
 - Learning object
 - Repositories
 - Metadata records
- Now: We just search for it

Brian Lamb at WCET



Brian Lamb
Emerging Technologies Discoordinator, OLT UBC <http://olt.ubc.ca/>
Vancouver Canada

Perspectives on metadata



The MLR story: International standardisation gone wild



- MLR = Metadata for Learning Resources
- What is an international standard?
- The process
 - The personalities
 - Formalisms, often for the sake of formalism
- What problems are the standard supposed to solve?

A yellow sticky note with a close button in the top right corner and a double-slash icon in the bottom right corner. The text on the note is 'Polish - repeated'.

Polish - repeated

Lessons from another SC36 project

What we did ...

- Take some proven specifications
 - Convert them to an ISO/IEC standard
 - Produce common vocabularies
- and now we will
- Produce more parts
 - Provide guidance on how to use them
 - Disseminate the results ...

Source: Presentation Australian expert SC36 WG7 document 2008

Lessons from another SC36 project

really

What we [^] did ...

- Take some proven ~~specifications~~ **applications**
 - Convert them to ISO/IEC ~~standard~~ **speak**
 - Produce common ~~vocabularies~~ **words**
- and now we will
- Produce more parts
 - Provide guidance on how to use them
 - Disseminate the results ...

Source: Presentation Australian expert SC36 WG7 document 2008

ISO Standardisations should be based on

- Proven practice from the field
- Tested specifications
- Stakeholder needs and commitment

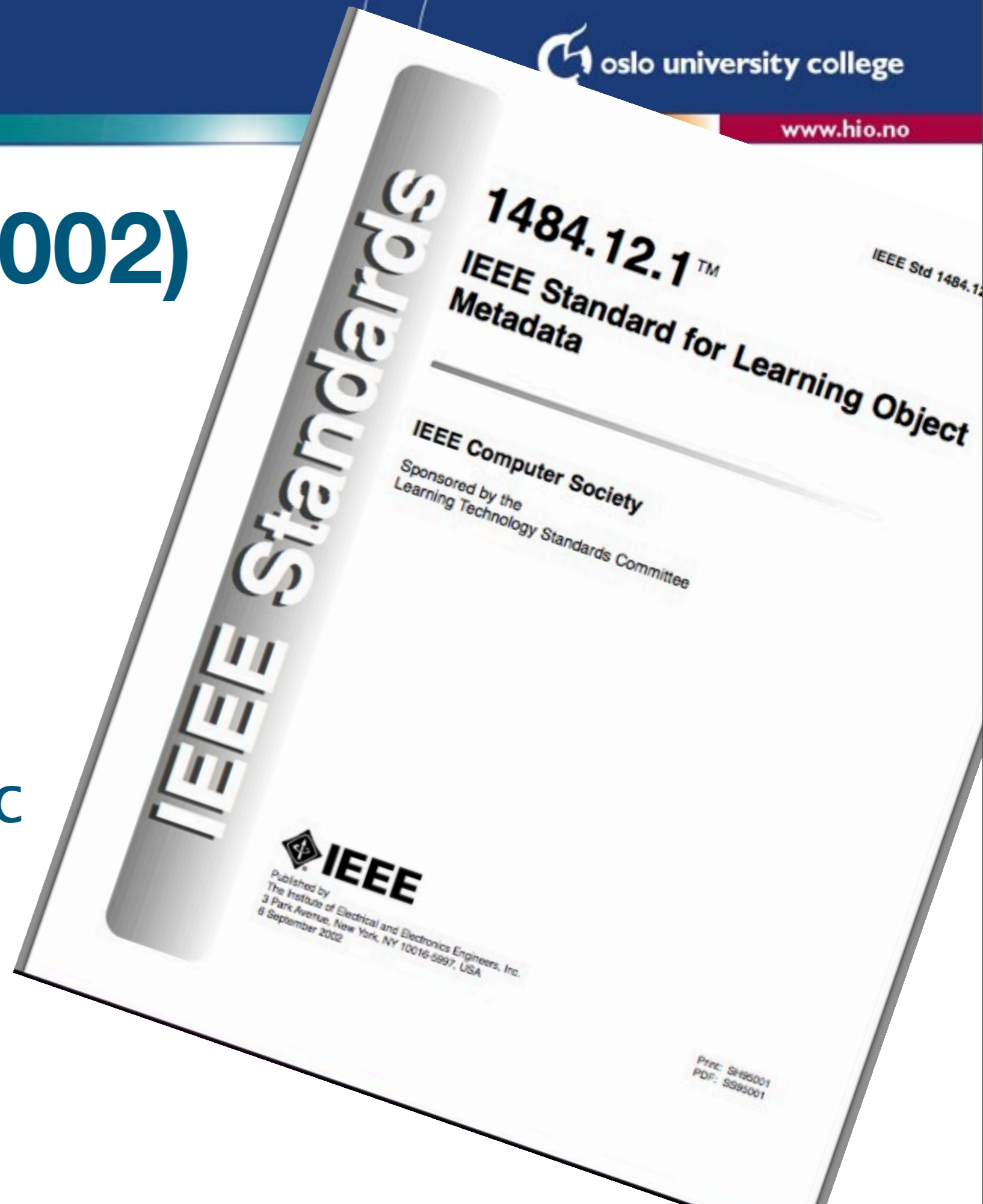
ISO Standardisations should be based on

- Proven practice from the field
- Tested specifications
- Stakeholder needs and commitment

What was the case for learning technologies?

IEEE LOM (2002)

- Our first e-learning standard!
- Superseded IMS LR metadata spec
- Should it be fast-tracked as an ISO standard?



Dublin Core

- Developed since mid 1990 – describing all kinds of web resources
- 2009 a ISO standard (15836)
 - (But only the Simple Dublin Core – 15 elements)



ISO/IEC JTC1 SC36 WG4 on Management and Delivery for Learning, Education, and Training

- Sept 2002: New WG4 established
- 2003: New Work Item on Metadata for Learning Resources (MLR)
 - 13 approve; 1 no (Sweden)
 - “Still, Sweden is of the opinion that the SC36 version of LOM is a compromise that **hardly fits as a standard to build on for future needs.**
 - Sweden believes future work within SC36 should be based on a **fully new approach on metadata modeling.**”

2003: What motivated Sweden's no vote?

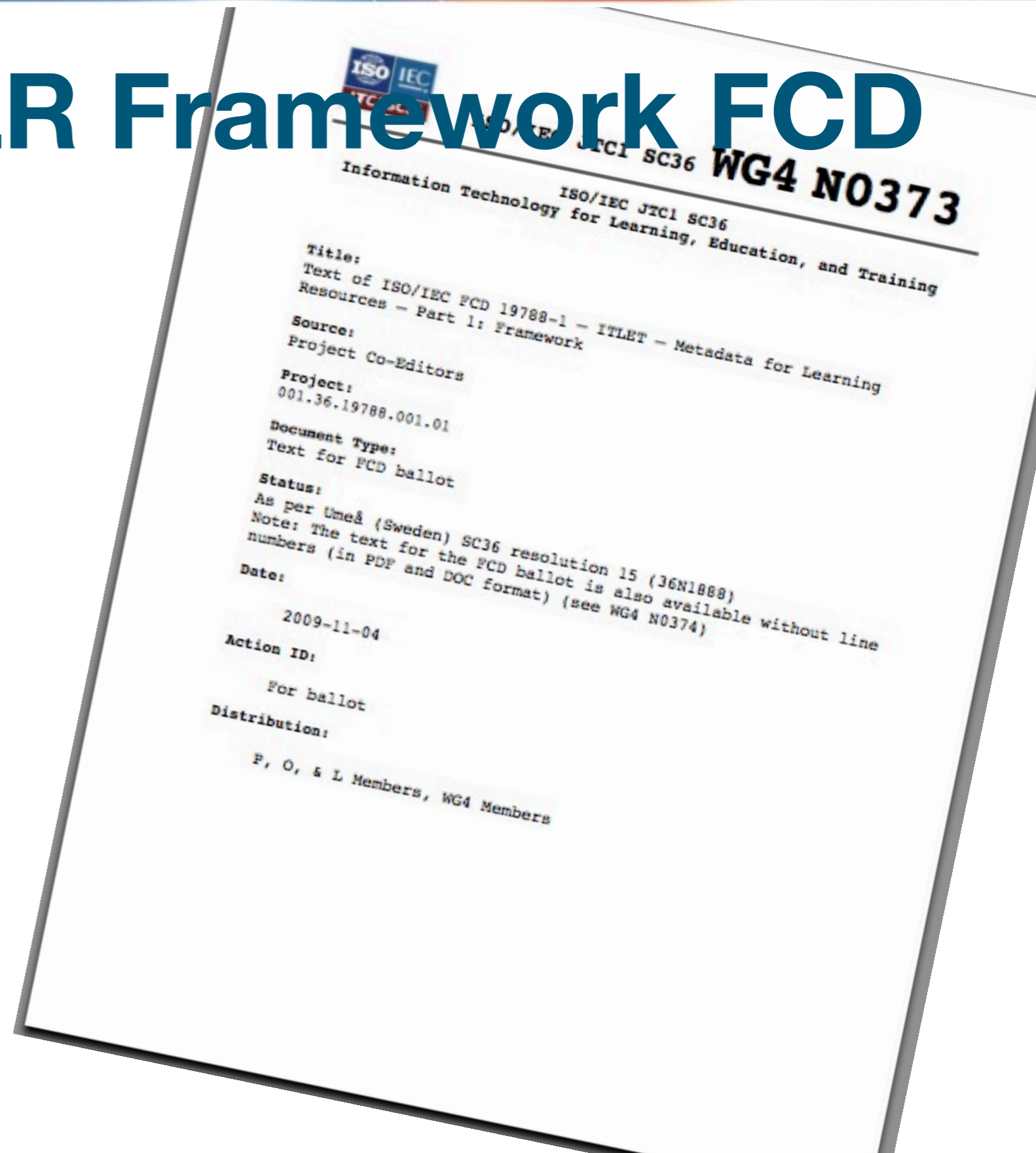
SC36_WG4_N0007.pdf (page 22 of 22)

The report has introduced a number of angles on metadata and metadata management. It is perfectly clear that, from a metadata perspective, data management and conceptual modelling appears to be new, at the same time as it is fairly conventional from a data perspective. All current trends indicate that a description of the resources that can be found and managed over the Internet will play a key role in the growth of the next web generation, one that presently goes under the name of the "Semantic Web". To this end established and generally accepted conceptual modelling languages will be needed by which presumably an almost endless number of conceptual models may be created.

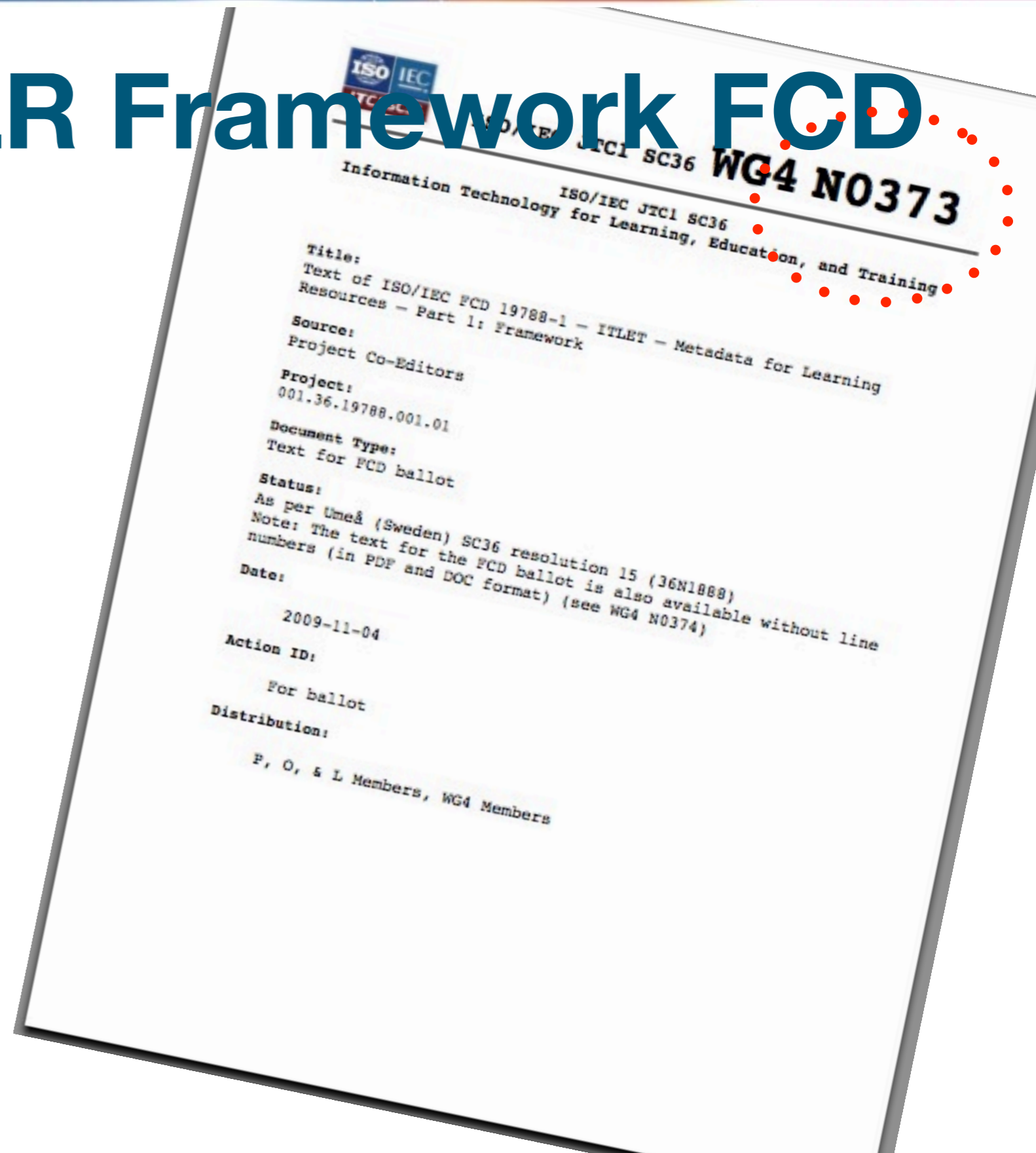
WG4 Working document No 7 of 2003-01-16

Nov 2009: MLR Framework FCD

Nov 2009: MLR Framework FCD



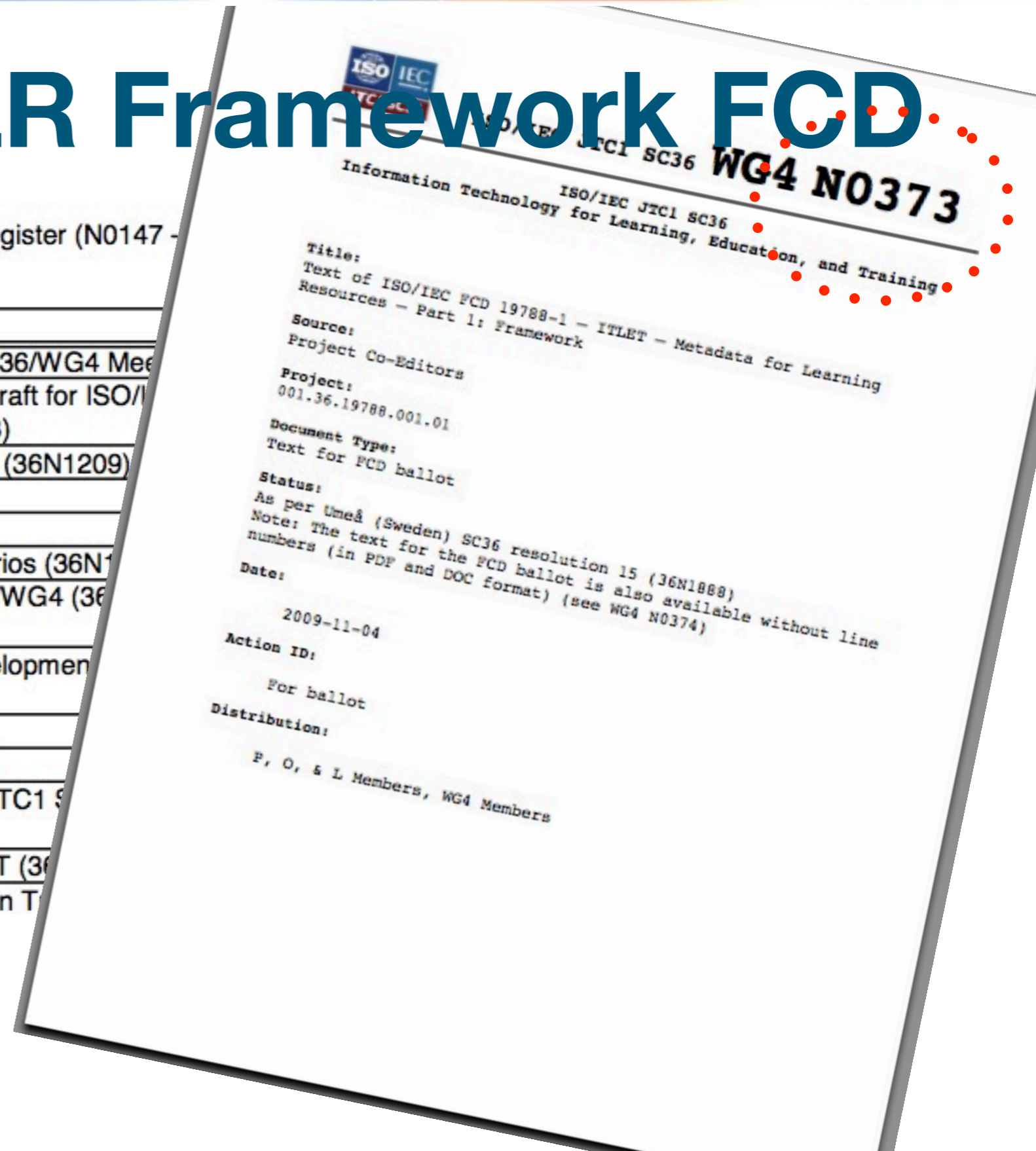
Nov 2009: MLR Framework FCD



Nov 2009: MLR Framework FCD

Document_Register (N0147 -

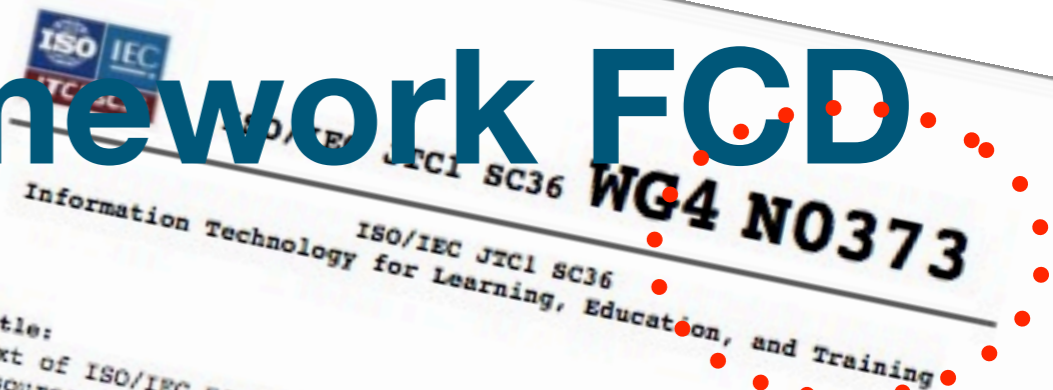
WG4 #	Title
N0147	Resolutions of the 2006-03 ISO/IEC JTC1 SC36/WG4 Meeting
N0148	DCMI Comments on WG4 N0145: Working Draft for ISO/IEC 19788-1:2006 Information Technology for Learning, Education, and Training Resources – Part 2: Data Elements (36N1203)
N0149	Draft - Aviation Industry Metadata Description (36N1209)
N0150	Metadata for Revisions (36N1226)
N0151	Metadata for Collections (36N1227)
N0152	Metadata for Collections and Example Scenarios (36N1228)
N0153	Canadian response to draft Agenda for SC36/WG4 (36N1229)
N0154	Contribution to "Principles governing the Development of Metadata (36N1233)
N0155	Flexible Hierarchy For Metadata (36N1235)
N0156	Hierarchy Attributes For Metadata (36N1236)
N0157	Final Draft Agenda for the 2006-03 ISO/IEC JTC1 SC36/WG4 Meeting (36N1253)
N0158	Aviation Industry Metadata Description DRAFT (36N1209)
N0159	SC36 Plenary Report of SC36/WG4 Meeting in Trondheim



Nov 2009: MLR Framework FCD

Document_Register (N0147 -

WG4 #	Title
N0147	Resolutions of the 2006-03 ISO/IEC JTC1 SC36/WG4 Meeting
N0148	DCMI Comments on WG4 N0145: Working Draft for ISO/IEC 19788-1:2006 Information Technology for Learning, Education, and Training Resources – Part 2: Data Elements (36N1203)
N0149	Draft - Aviation Industry Metadata Description (36N1209)
N0150	Metadata for Revisions (36N1226)
N0151	Metadata for Collections (36N1227)
N0152	Metadata for Collections and Example Scenarios (36N1228)
N0153	Canadian response to draft Agenda for SC36/WG4 (36N1229)
N0154	Contribution to "Principles governing the Development of the MLR Framework" (36N1233)
N0155	Flexible Hierarchy For Metadata (36N1234)
N0156	...



ISO/IEC JTC1 SC36
 Information Technology for Learning, Education, and Training
WG4 N0373

Title:
 Text of ISO/IEC FCD 19788-1 – ITLET – Metadata for Learning Resources – Part 1: Framework

Source:
 Project Co-Editors

Project:
 001.36.19788.001.01

Document Type:
 Text for FCD ballot

Status:
 As per Umeå (Sweden) SC36 resolution 15 (36N1888)

Note: The text for the FCD ballot is also available without line numbers (in PDF and DOC format) (see WG4 N0374)

Date:
 2009-11-04

Action ID:

Request for added Agenda Item for SC36/WG4 Turku, Finland March, 2006 Meetings

Canada notes that its Dublin, September, 2004 Meeting, that SC36/WG4 adopted Resolution 2 (see SC36/WG4 N0134) instructing the Project Editor to develop MLR 1 based on agreed upon principles which are listed in that Resolution along with SC36/WG4 document references.

Canada notes that this has not been done in development of the Part 1 MLR work. This SC36/WG4 resolution also applies to all other present and future Parts of MLR work.

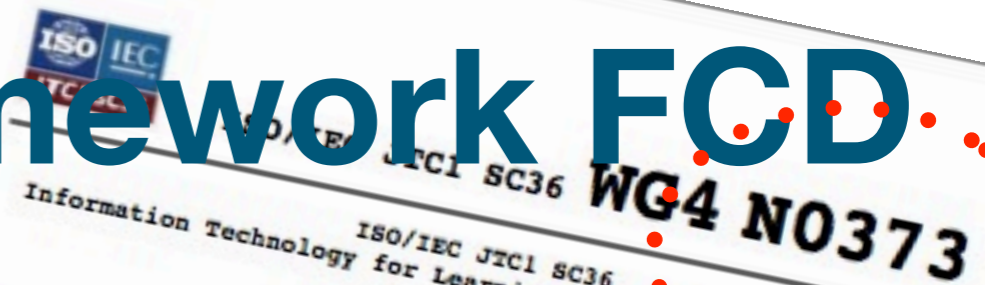
Canada requests that apart from resolution of ballot comments on the CD for Part 1, that SC36/WG4 schedule as an Agenda item a (1hr+) discussion on "Principles governing the Development of the MLR multipart standard".

bers, WG4 Members

Nov 2009: MLR Framework FCD

Document_Register (N0147 -

WG4 #	Title
N0147	Resolutions of the 2006-03 ISO/IEC JTC1 SC36/WG4 Meeting
N0148	DCMI Comments on WG4 N0145: Working Draft for ISO/IEC 19788-1:2006 Information Technology for Learning, Education, and Training Resources – Part 2: Data Elements (36N1203)
N0149	Draft - Aviation Industry Metadata Description (36N1209)
N0150	Metadata for Revisions (36N1226)
N0151	Metadata for Collections (36N1227)
N0152	Metadata for Collections and Example Scenarios (36N1228)
N0153	Canadian response to draft Agenda for SC36/WG4 (36N1229)
N0154	Contribution to "Principles governing the Development of the MLR Framework" (36N1233)
N0155	Flexible Hierarchy For Metadata (36N1234)
N0156	Request for added Agenda Item for SC36/WG4



ISO/IEC JTC1 SC36
 Information Technology for Learning, Education, and Training
WG4 N0373

Title:
 Text of ISO/IEC FCD 19788-1 – ITLET – Metadata for Learning Resources – Part 1: Framework

Source:
 Project Co-Editors

Project:
 001.36.19788.001.01

Document Type:
 Text for FCD ballot

Status:
 As per Umeå (Sweden) SC36 resolution 15 (36N1888)

Note: The text for the FCD ballot is also available without line numbers (in PDF and DOC format) (see WG4 N0374)

Date:
 2009-11-04

Action ID:

Request for added Agenda Item for SC36/WG4 Turku, Finland March, 2006 Meetings

Canada notes that its Dublin, September, 2004 Meeting, that SC36/WG4 adopted Resolution 2 (see SC36/WG4 N0134) instructing the Project Editor to develop MLR 1 based on agreed upon principles which are listed in that Resolution along with SC36/WG4 document references.

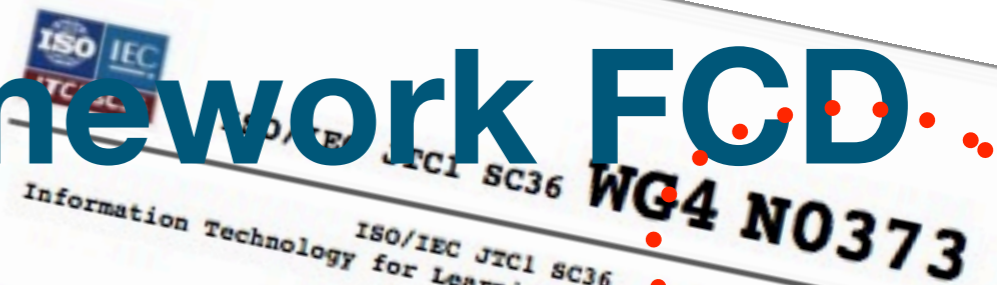
Canada notes that this has not been done in development of the Part 1 MLR work. This SC36/WG4 resolution also applies to all other present and future Parts of MLR work.

Canada requests that apart from resolution of ballot comments on the CD for Part 1, that SC36/WG4 schedule as an Agenda item a (1hr+) discussion on "Principles governing the Development of the MLR multipart standard".

Nov 2009: MLR Framework FCD

Document_Register (N0147 -

WG4 #	Title
N0147	Resolutions of the 2006-03 ISO/IEC JTC1 SC36/WG4 Meeting
N0148	DCMI Comments on WG4 N0145: Working Draft for ISO/IEC 19788-1:2006 Information Technology for Learning, Education, and Training Resources – Part 2: Data Elements (36N1203)
N0149	Draft - Aviation Industry Metadata Description (36N1209)
N0150	Metadata for Revisions (36N1226)
N0151	Metadata for Collections (36N1227)
N0152	Metadata for Collections and Example Scenarios (36N1228)
N0153	Canadian response to draft Agenda for SC36/WG4 (36N1229)
N0154	Contribution to "Principles governing the Development of the MLR Framework" (36N1233)
N0155	Flexible Hierarchy For Metadata (36N1234)
N0156	...



Information Technology for Learning, Education, and Training
 ISO/IEC JTC1 SC36
WG4 N0373

Title:
 Text of ISO/IEC FCD 19788-1 – ITLET – Metadata for Learning Resources – Part 1: Framework

Source:
 Project Co-Editors

Project:
 001.36.19788.001.01

Document Type:
 Text for FCD ballot

Status:
 As per Umeå (Sweden) SC36 resolution 15 (36N1888)

Note: The text for the FCD ballot is also available without line numbers (in PDF and DOC format) (see WG4 N0374)

Date:
 2009-11-04

Action ID:

Request for added Agenda Item for SC36/WG4 Turku, Finland March, 2006 Meetings

Canada notes that its Dublin, September, 2004 Meeting, that SC36/WG4 adopted Resolution 2 (see SC36/WG4 N0134) instructing the Project Editor to develop MLR 1 based on agreed upon principles which are listed in that Resolution along with SC36/WG4 document references.

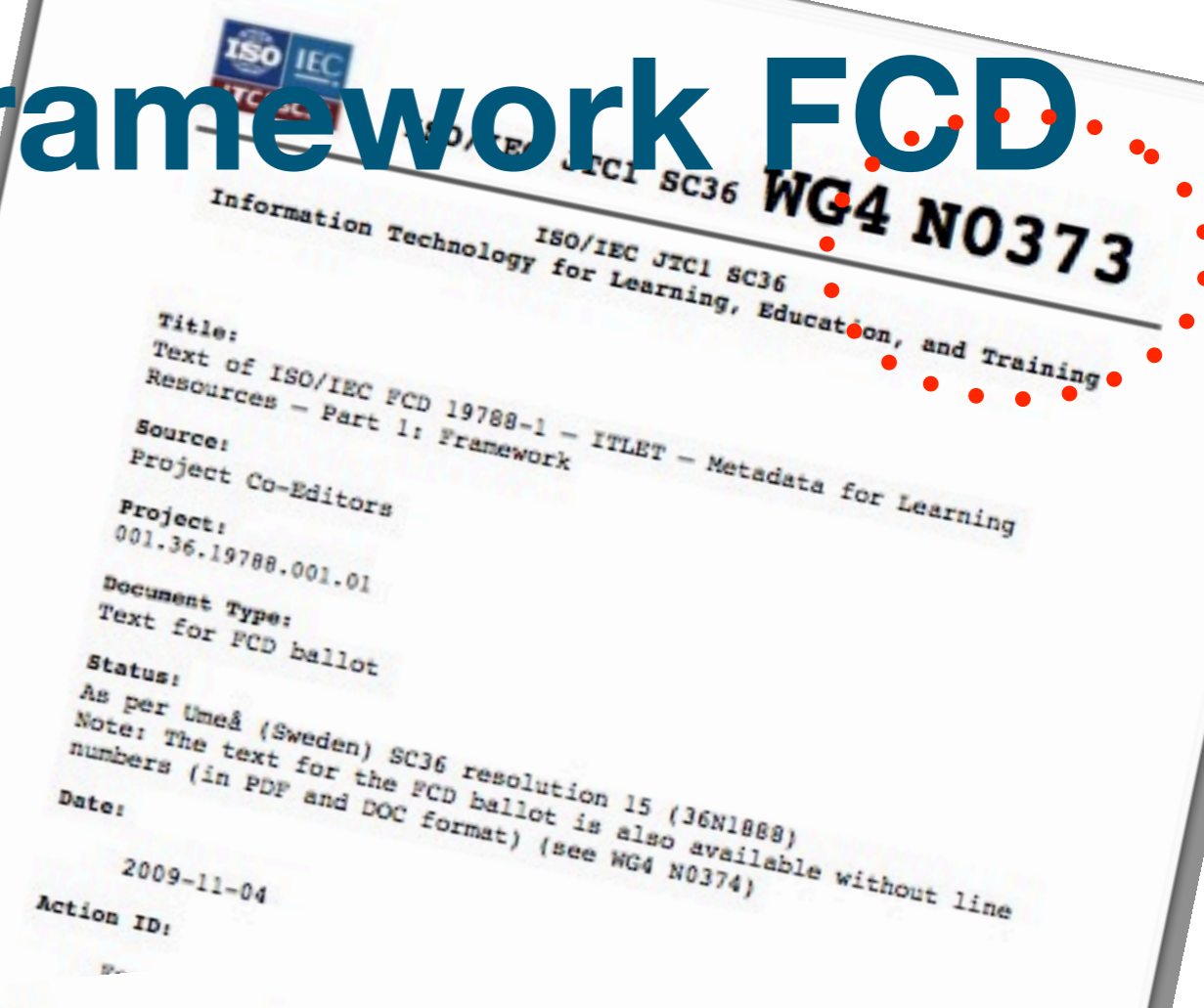
Canada notes that this has not been done in development of the Part 1 MLR work. This SC36/WG4 resolution also applies to all other present and future Parts of MLR work.

Canada requests that apart from resolution of ballot comments on the CD for Part 1, that SC36/WG4 schedule as an Agenda item a (1hr+) discussion on "Principles governing the Development of the MLR multipart standard".

Nov 2009: MLR Framework FCD

Document_Register (N0147 -

WG4 #	Title
N0147	Resolutions of the 2006-03 ISO/IEC JTC1 SC36/WG4 Meeting
N0148	DCMI Comments on WG4 N0145: Working Draft for ISO/IEC 19788-1:2006 Information Technology for Learning, Education, and Training Resources – Part 2: Data Elements (36N1203)
N0149	Draft - Aviation Industry Metadata Description (36N1209)
N0150	Metadata for Revisions (36N1226)
N0151	Metadata for Collections (36N1227)
N0152	Metadata for Collections and Example Scenarios (36N1228)
N0153	Canadian response to draft Agenda for SC36/WG4 (36N1229)
N0154	Contribution to "Principles governing the Development of the MLR Framework" (36N1233)
N0155	Flexible Hierarchy For Metadata (36N1234)
N0156	...



Request for added Agenda Item for SC36/WG4 Turku, Finland March, 2006 Meetings

Canada notes that its Dublin, September, 2004 Meeting, that SC36/WG4 adopted Resolution 2 (see SC36/WG4 N0134) instructing the Project Editor to develop MLR 1 based on agreed upon principles which are listed in that Resolution along with SC36/WG4 document references.

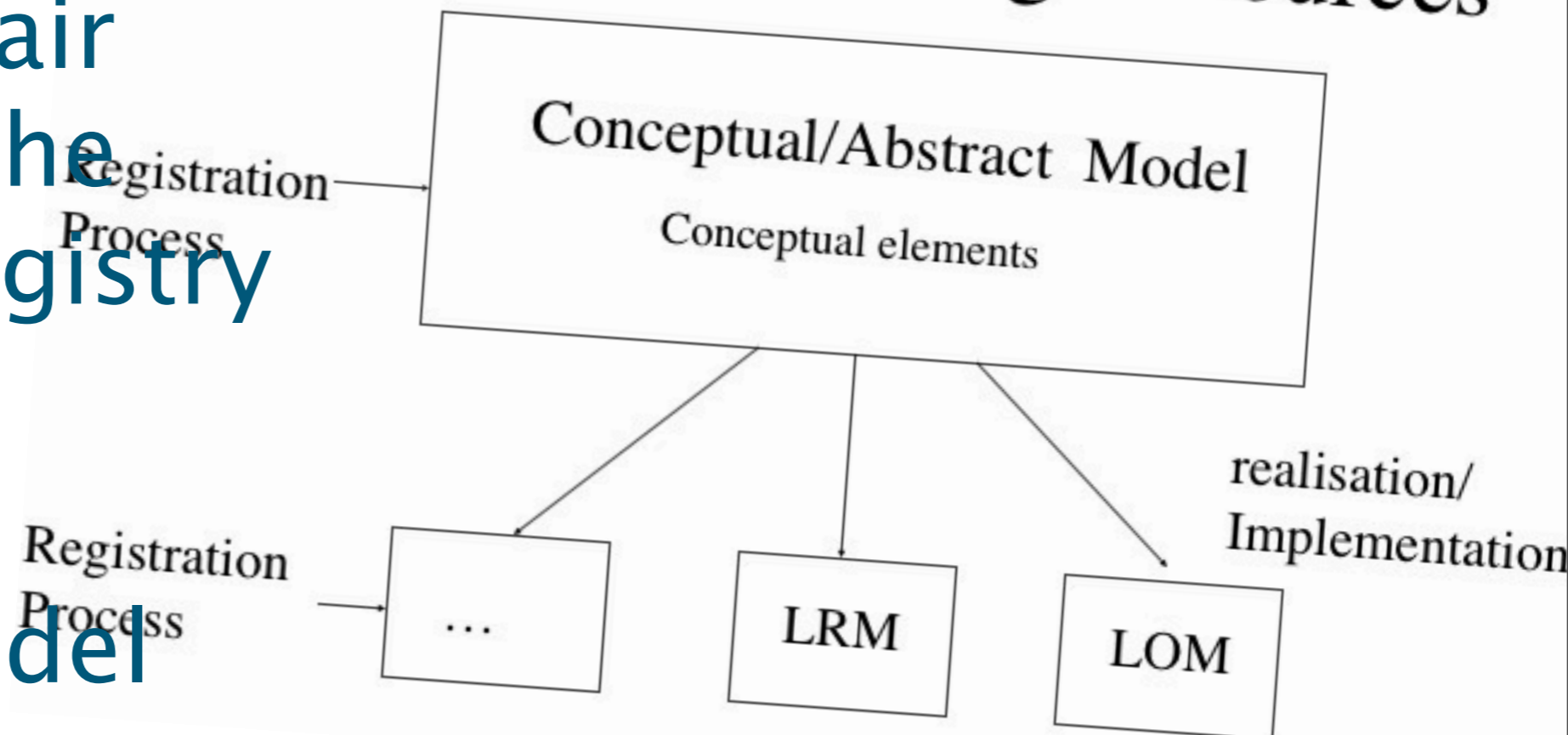
Canada notes that this has not been done in development of the Part 1 MLR work. This SC36/WG4 resolution also applies to all other present and future Parts of MLR work.

Canada requests that apart from resolution of ballot comments on the CD for Part 1, that SC36/WG4 schedule as an Agenda item a (1hr+) discussion on "Principles governing the Development of the MLR multipart standard".

MLR first years

- The WG4 chair champions the Metadata Registry (ISO 11179) approach
- Abstract model necessary

Metadata for Learning Resources



LOM survey 2003

Conclusion

The general, preliminary findings of this survey --themselves a product of interpretation-- invite further interpretation and speculation about the LOM and the way it tends to be implemented. However, it seems clear that a number of these findings support the conclusion that fewer and better defined elements might be more effective than the range of choice and interpretive possibilities currently allowed by the LOM. This seems to be especially the case regarding educational elements, which are surprisingly underutilized for metadata that it ostensibly and primarily educational. The need for a smaller number of elements is further supported by the common identification of Dublin Core element equivalents in the application profiles surveyed. There are also a number of findings supporting the conclusion that clear and easily-supported means of working with local, customized vocabularies would also be very valuable --and that the means of retaining a minimum of interoperability between these variant vocabularies would also be important. Finally, it also seems useful to ensure that structures are provided to accommodate complex but more conventional aspects of resource description. These would include multiple title versions, as well as multilingual descriptions and values.

WG4 document N0057 2003-09-03

The document mill

- 2003: Work started
- During 2004: 135 documents
- End of 2005: 145
- End of 2006: 180
- End of 2007: 233
- End of 2008: 305
- And now 2009: 373 WG4 documents + all the other SC36 documents

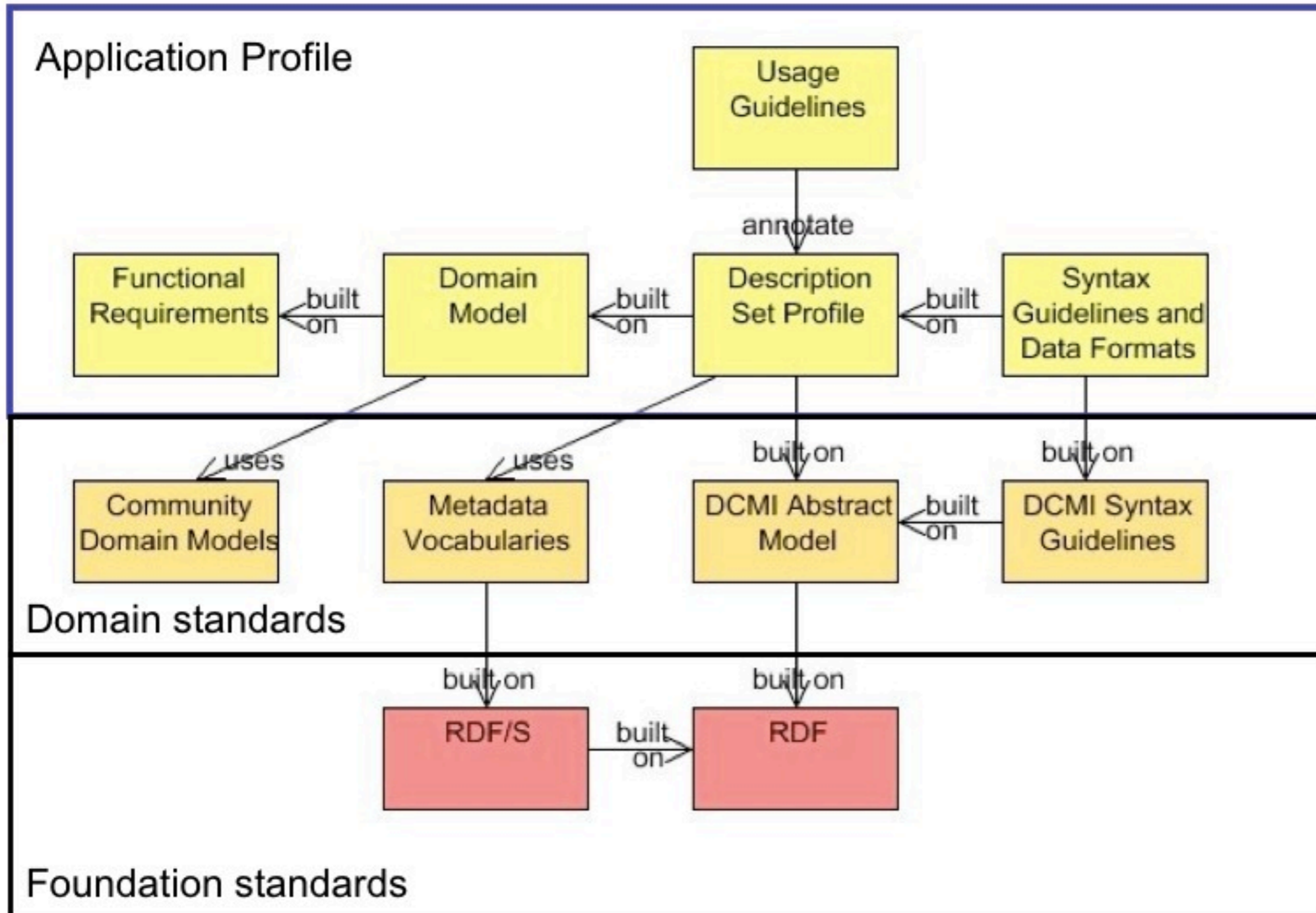
The exhausting MLR process – what is it about?

- IEEE LOM / DC compatibility
 - Interoperability – how is it understood?
- A global metadata standard – what is it?
 - Application profiles – which role do they play?
- What is the metamodel (Abstract model) of the standard?
- And then again, the personalities' role in standardisation

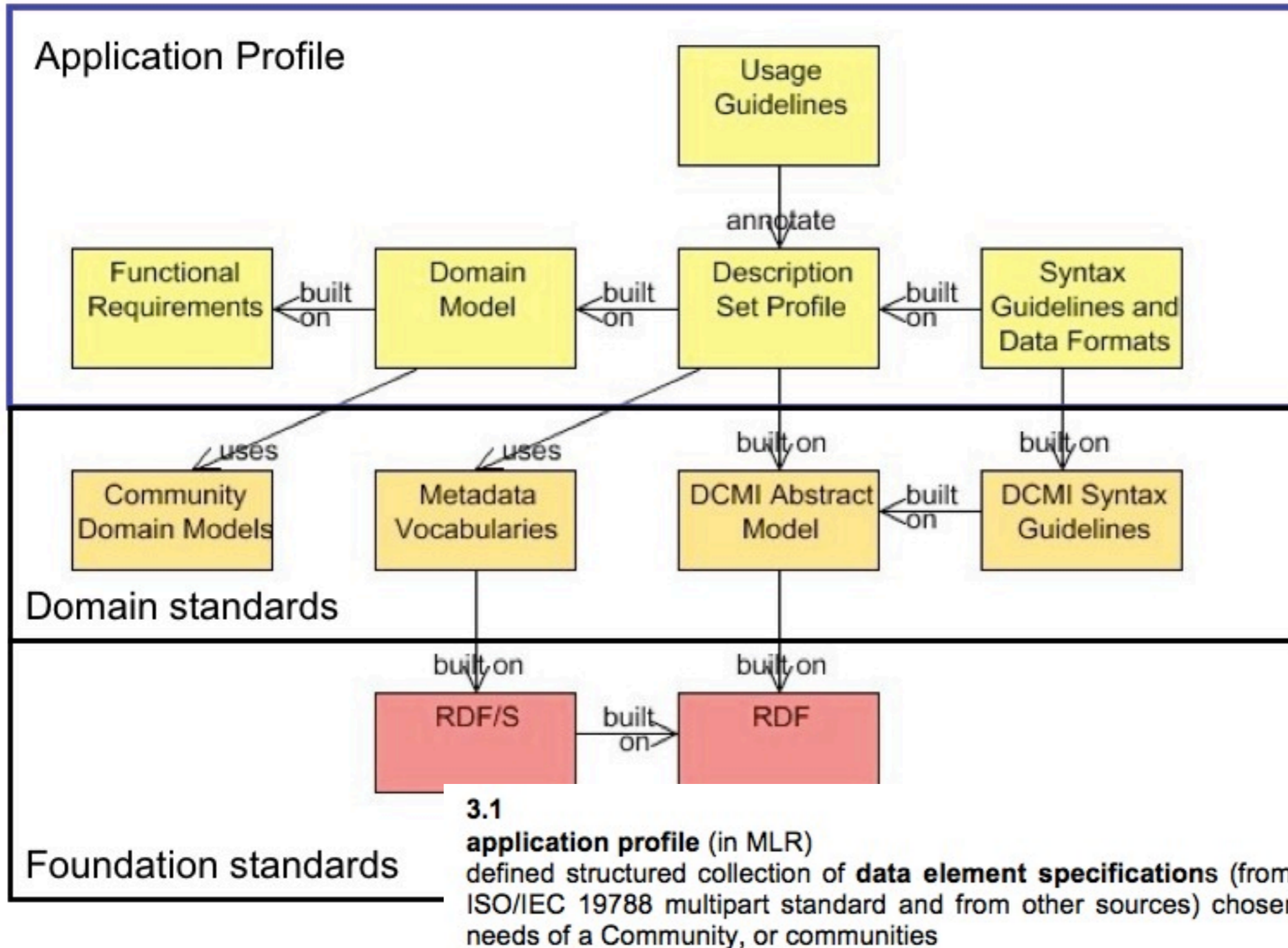
Interoperability

- Full compatibility with both DC and LOM is not possible (ref the strive for a RDF binding of LOM)
- LOM elements not usable in combination with DCMI elements (e.g. Dublin Core APs)
 - The concept of “element” differ substantially between the two standards
 - Surface interoperability:
 - XML namespaces
 - RDF
 - ...but the interpretation of these expressions differ

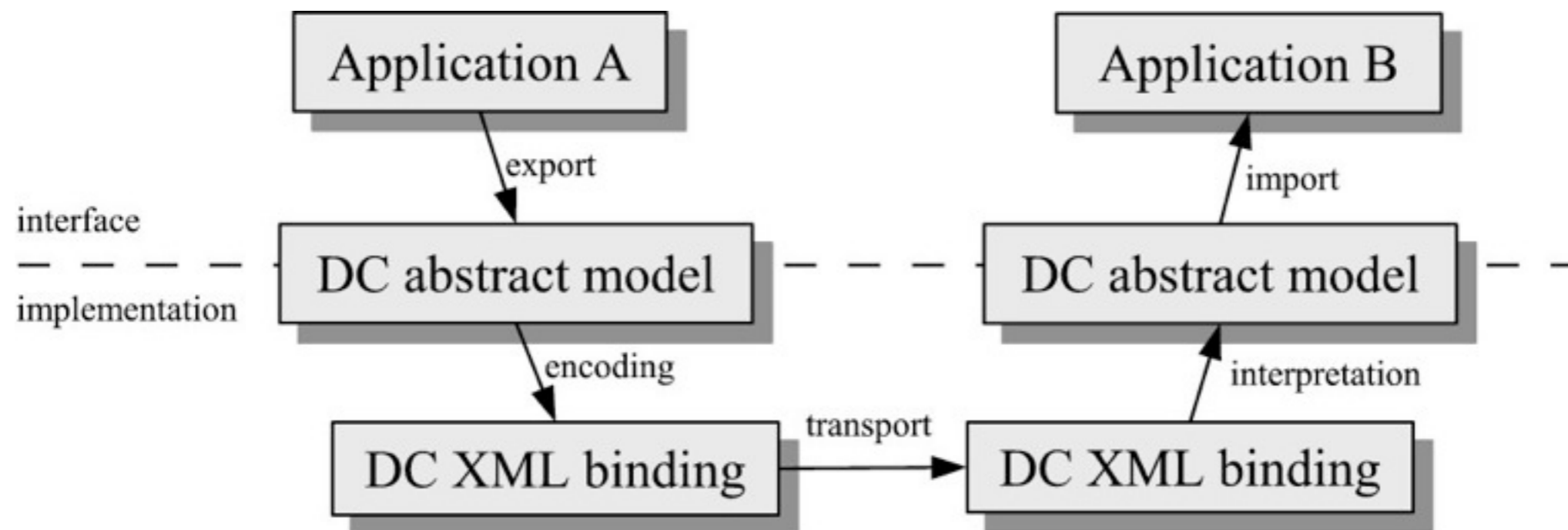
Application profiles – DC view



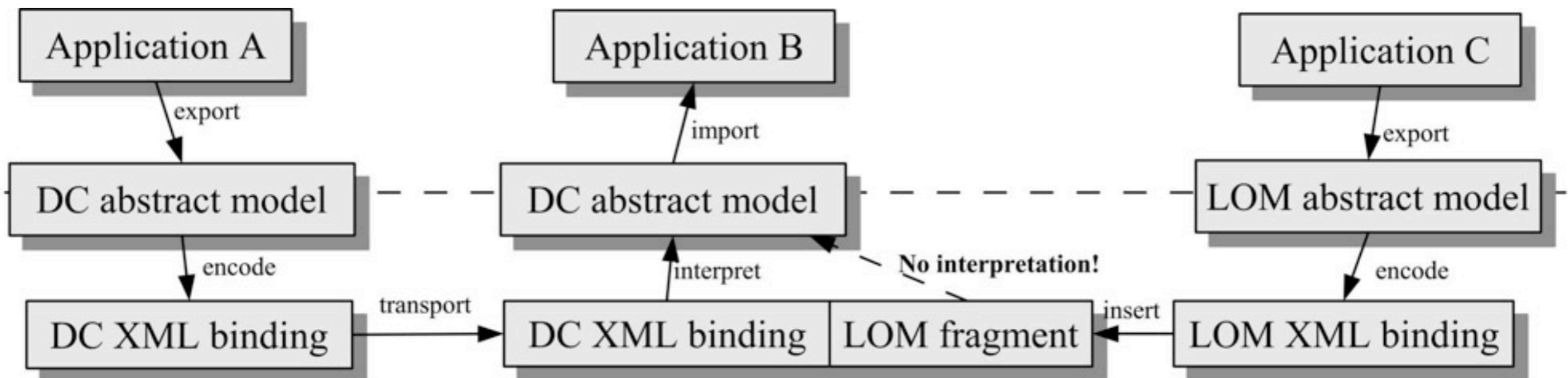
Application profiles – DC view



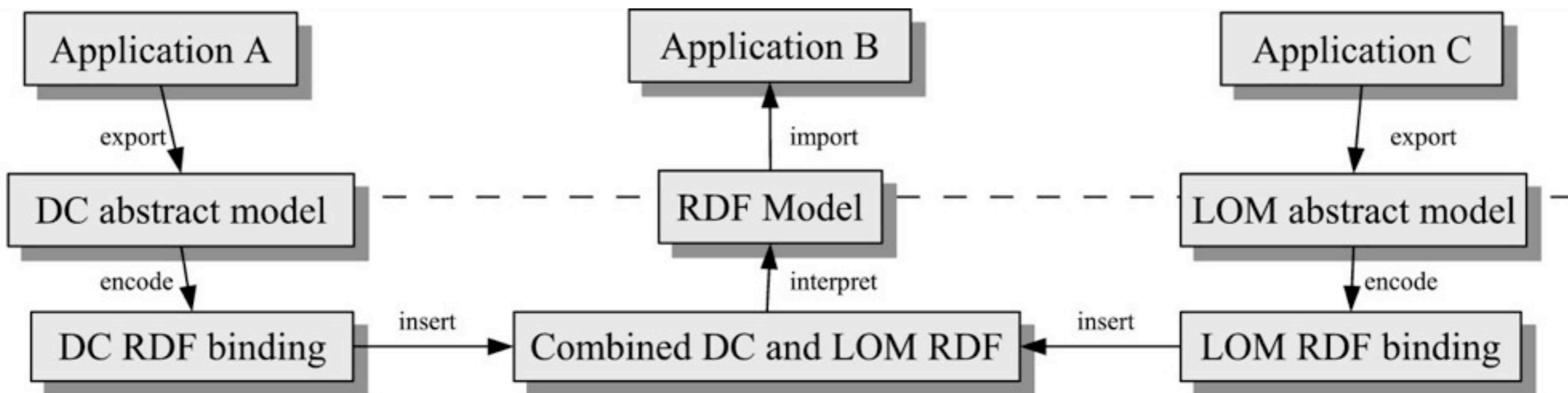
Interpreting metadata between two DC applications



Combining XML fragments from DC and LOM



Combining RDF fragments



What is wrong with LOM?

- Too complex
 - (the “semantic density” is too high?)
- Identification of data elements done by referring to names (labels)
- Lots of “hidden” semantics in the structure of the elements
- Use of compounds – (boxes within boxes)

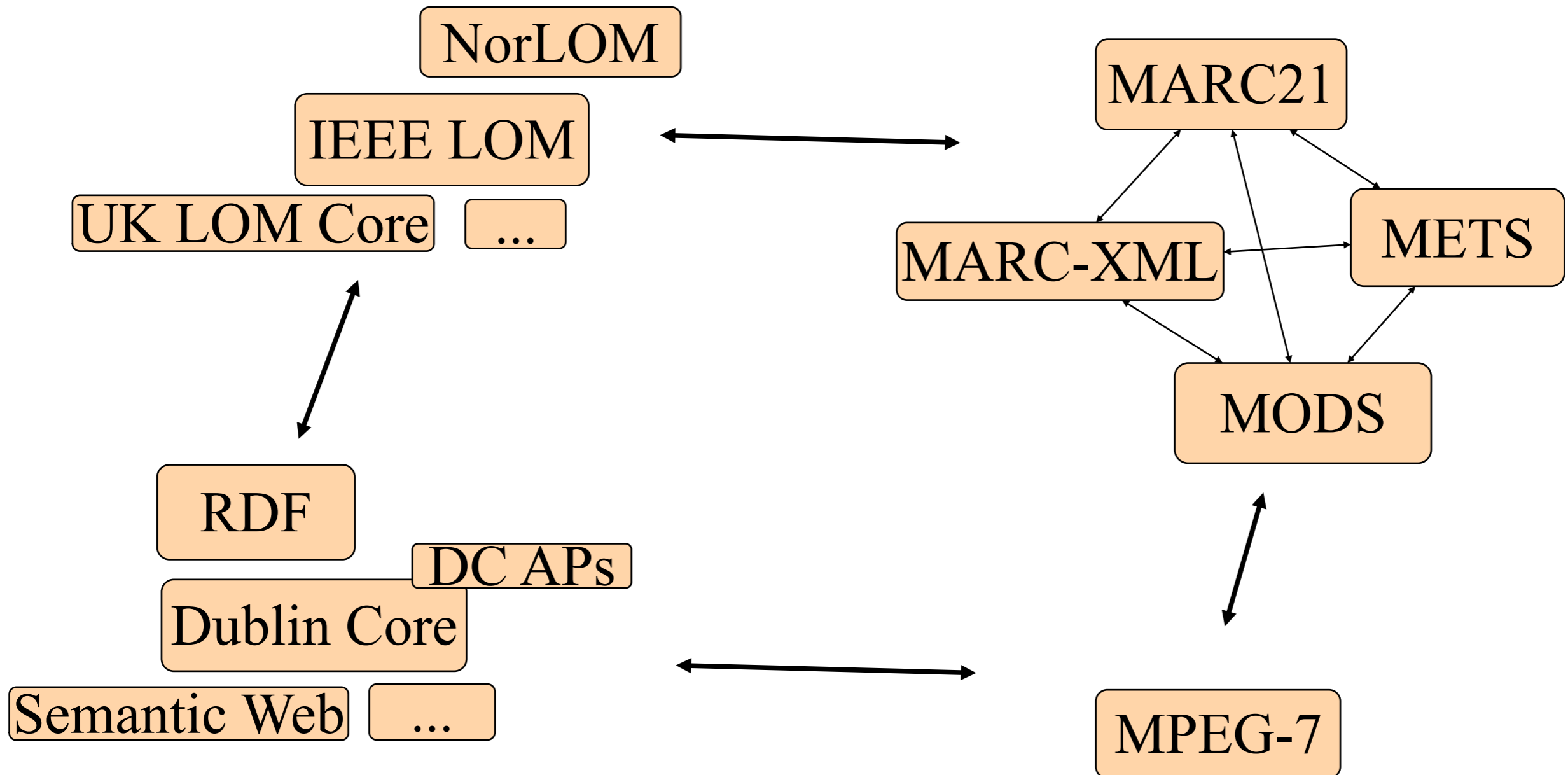
The MLR U-turn – the “Semantization project”

- Document WG4_N0278
- Late 2007 – 2008
 - Co-ordinated efforts
 - The importance of written expert contributions (if views not in a N document, do not count)
- Btw, where was Google at that time?

Introductory slide of Mikael Nilsson Stuttgart Sept 2008

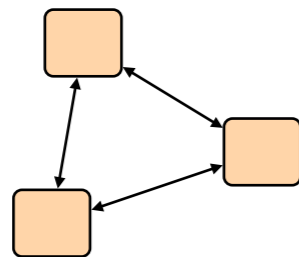
- Islands of metadata interoperability, for example
 - The “LOM island” – IEEE LOM and LOM-based profiles
 - The “MODS island”
 - The “MPEG-7 island”
 - The “Dublin Core & RDF island”
- Two approaches to Application Profiles
 - Base standard – profiles customise the base
 - (LOM, MODS, MPEG-7)
 - Framework only, profiles combine terms arbitrarily (Dublin Core, RDF) <-- MLR wants to be here (?)

Metadata interoperability now

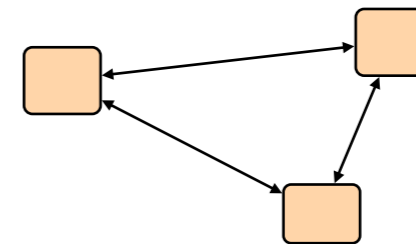


Metadata interoperability vision

Education

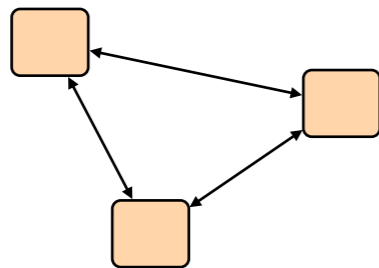


Libraries

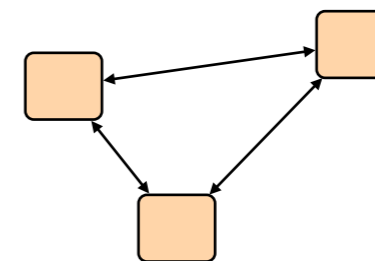


Semantic Web

Government



Multimedia



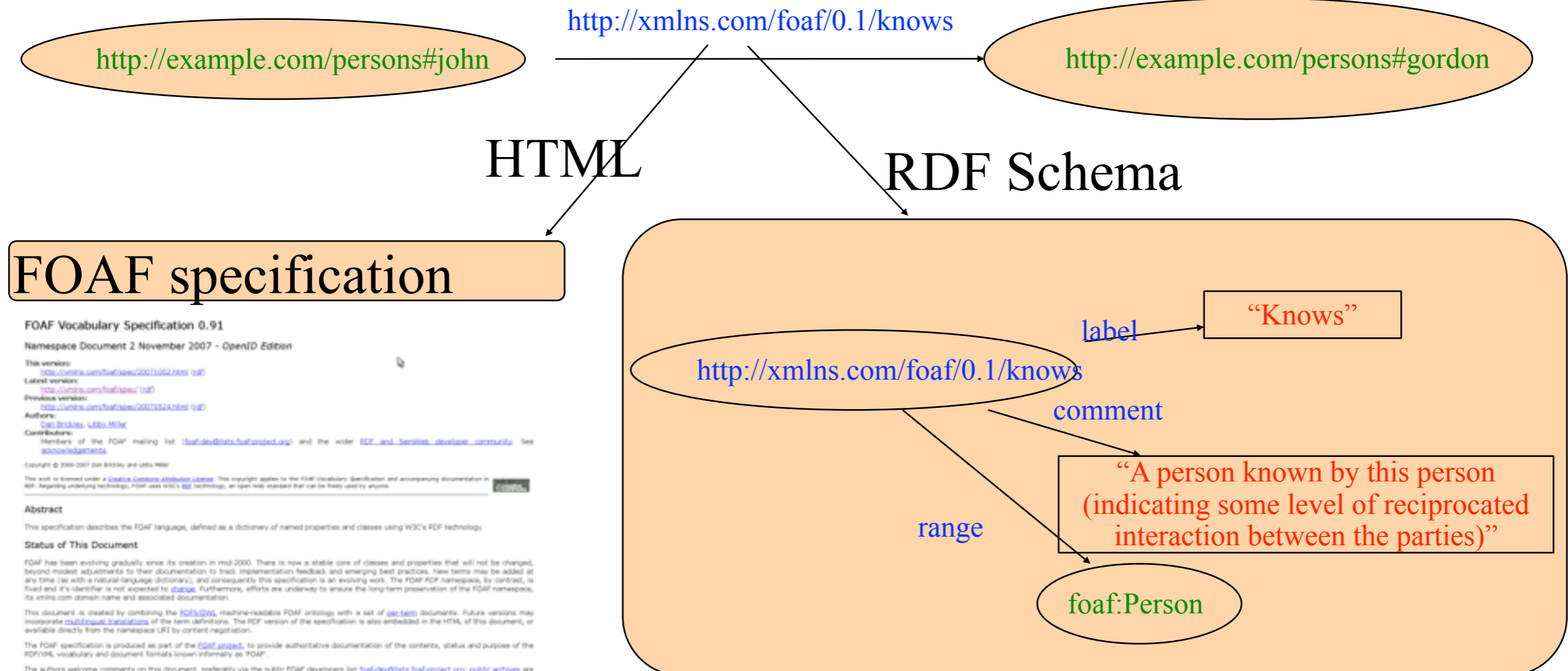
Levels of interoperability

- **Human interoperability – words**
 - Use the same definition of words, regardless of technical framework
- **Semantic interoperability – the cloud**
 - Machines apply the same processing to terms wherever they appear
 - This is the purpose of RDF
- **Profile interoperability – the domain**
 - Domain-specific interoperability based on shared profiles, vocabularies, etc.
 - Quality control, syntax validation etc.

Proposal for MLR

- DON'T create a new metadata island
 - DON'T create a need for more crosswalks
- DO use a framework-based approach
 - DO allow for application profiles combining terms from other sources
- DON'T reinvent the framework
 - DON'T require others to redefine their terms for use in MLR
- DO base the framework on the RDF model

“Follow your nose”



FOAF specification

FOAF Vocabulary Specification 0.91
 Namespace Document 2 November 2007 - OpenID Edition

This version: <http://xmlns.com/foaf/spec/20071002.html> (rdf)

Latest version: <http://xmlns.com/foaf/spec/> (rdf)

Previous version: <http://xmlns.com/foaf/spec/20070524.html> (rdf)

Authors: [Dan Brickley](#), [Liz Miller](#)

Contributors: Members of the FOAF mailing list (foaf-dev@lists.foaf-project.org) and the wider [FOAF and Semantic developer community](#). See [acknowledgements](#).

Copyright © 2000-2007 Dan Brickley and Liz Miller

This work is licensed under a [Creative Commons Attribution License](#). The copyright applies to the FOAF vocabulary specification and accompanying documentation in RDF. Regarding underlying technology, FOAF uses W3C's [RDF](#) technology, an open web standard that can be freely used by anyone.

Abstract

This specification describes the FOAF language, defined as a dictionary of named properties and classes using W3C's RDF technology.

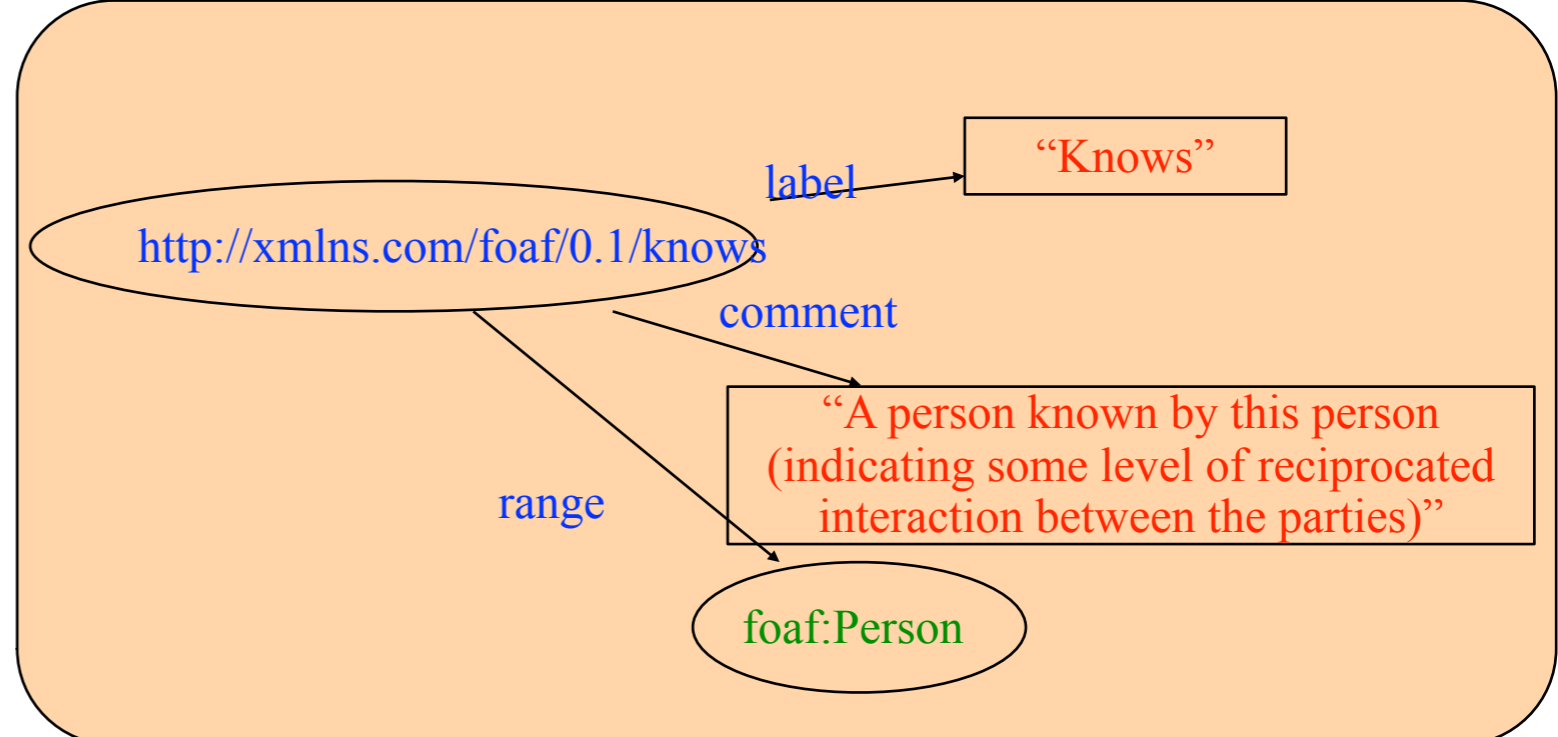
Status of This Document

FOAF has been evolving gradually since its creation in mid-2000. There is now a stable core of classes and properties that will not be changed, beyond modest adjustments to their documentation to track implementation feedback and emerging best practices. New terms may be added at any time (as with a natural language dictionary), and consequently this specification is an evolving work. The FOAF RDF namespace, by contrast, is fixed and its identifier is not expected to [change](#). Furthermore, efforts are underway to ensure the long-term preservation of the FOAF namespace, its `xmlns.com` domain name and associated documentation.

This document is created by combining the [RDFS/OWL](#), machine-readable FOAF ontology with a set of [web-based](#) documents. Future versions may incorporate [multilingual translations](#) of the term definitions. The RDF version of the specification is also embedded in the HTML of this document, or available directly from the namespace URI by content negotiation.

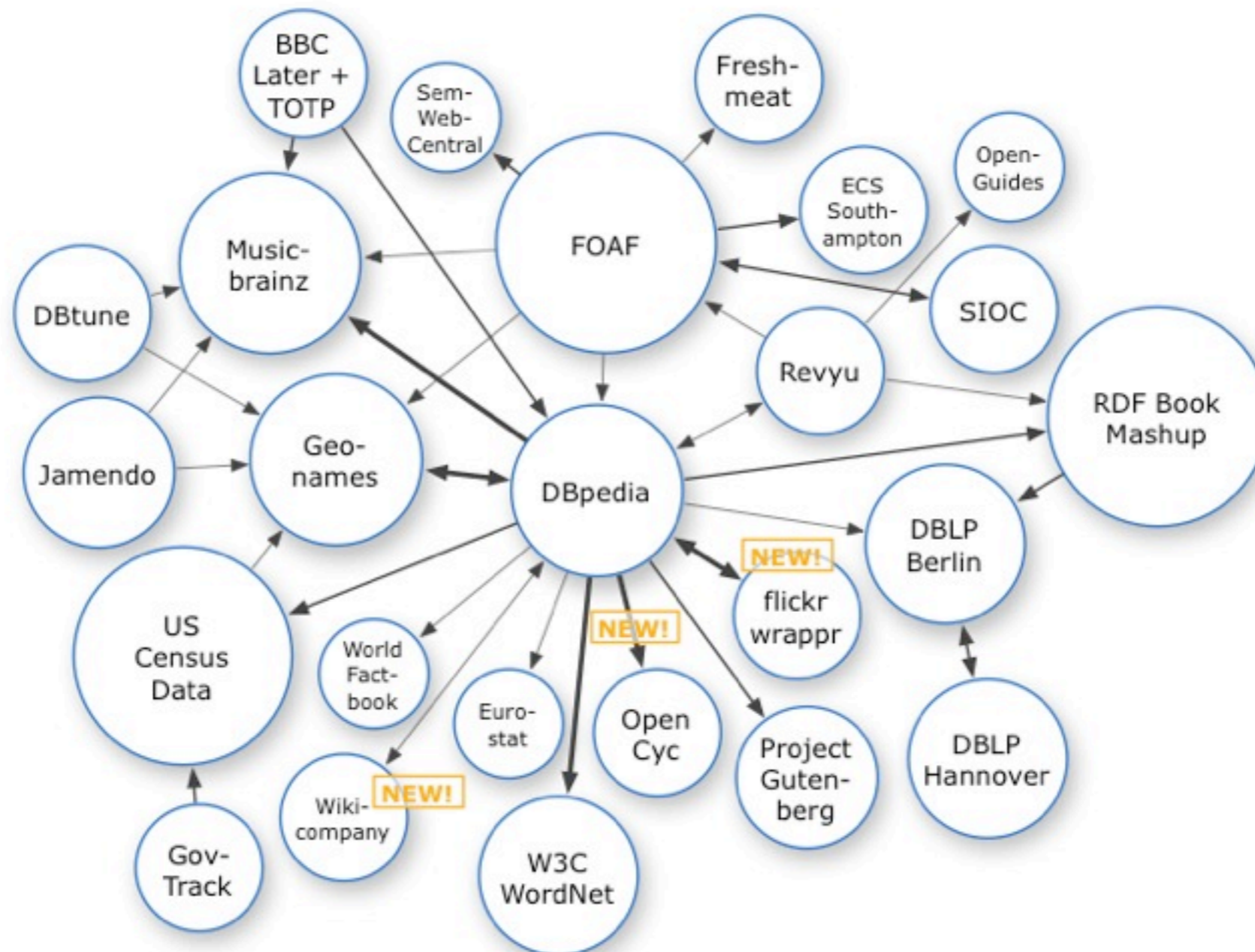
The FOAF specification is produced as part of the [FOAF project](#), to provide authoritative documentation of the contents, status and purpose of the RDFS/OWL, vocabulary and document formats known informally as 'FOAF'.

The authors welcome comments on this document, preferably via the public FOAF developers list foaf-dev@lists.foaf-project.org, [public archives](#) are available. A historical backlog of known technical issues is acknowledged, and available for discussion in the [FOAF wiki](#). Proposals for resolving these issues are welcomed, either on [foaf-dev](#) or via the wiki. Further work is also needed on the exploratory text in this specification and on the [FOAF website](#); progress towards this will be measured in the version number of future revisions to the FOAF specification.

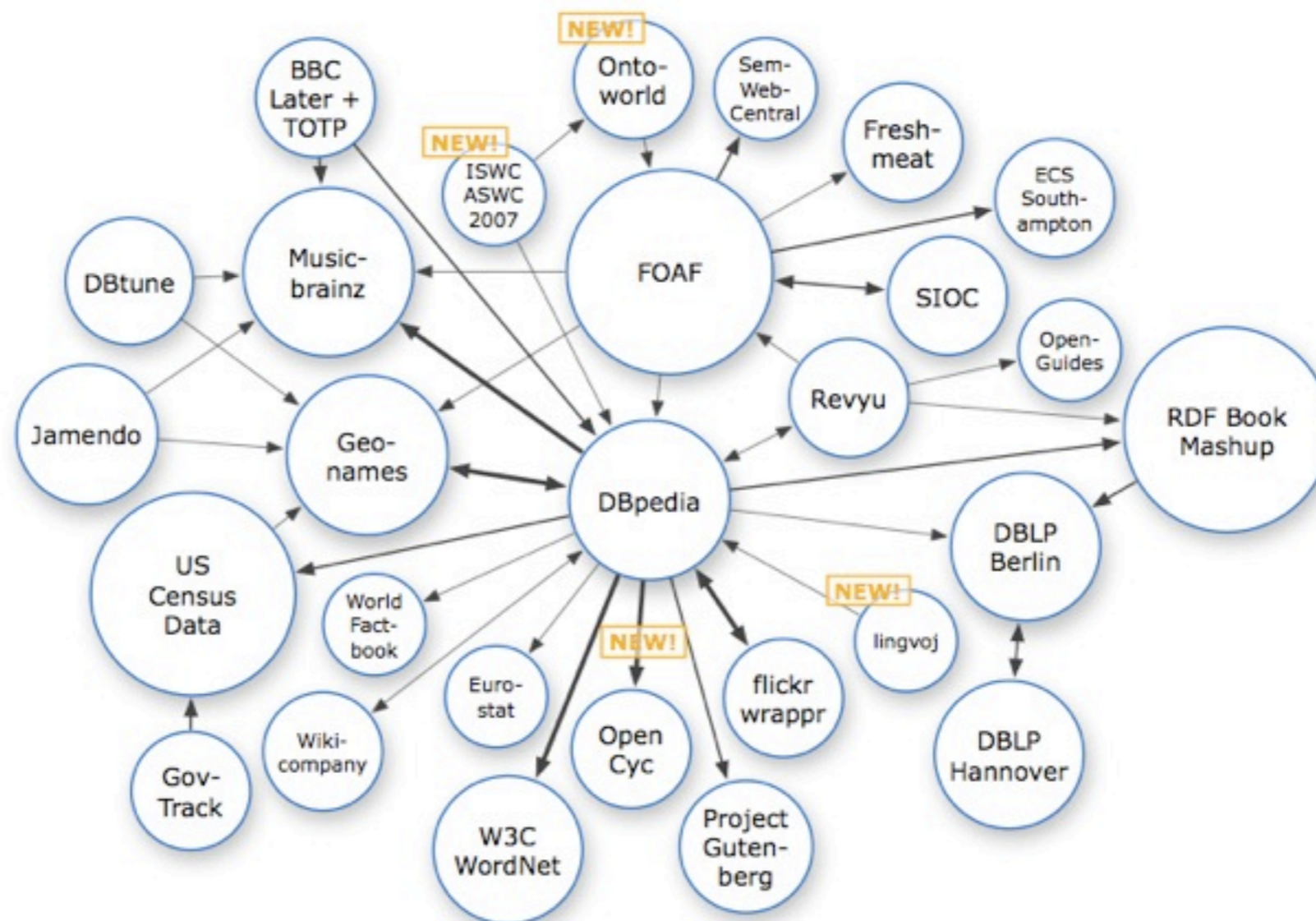


```
<rdf:Property rdf:about="http://xmlns.com/foaf/0.1/knows"
vs:term_status="testing" rdfs:label="knows" rdfs:comment="A person known by this
person (indicating some level of reciprocated interaction between the parties).">
  <rdf:type rdf:resource="http://www.w3.org/2002/07/owl#ObjectProperty" />
  <rdfs:domain rdf:resource="http://xmlns.com/foaf/0.1/Person" />
  <rdfs:range rdf:resource="http://xmlns.com/foaf/0.1/Person" />
  <rdfs:isDefinedBy rdf:resource="http://xmlns.com/foaf/0.1/" />
</rdf:Property>
```

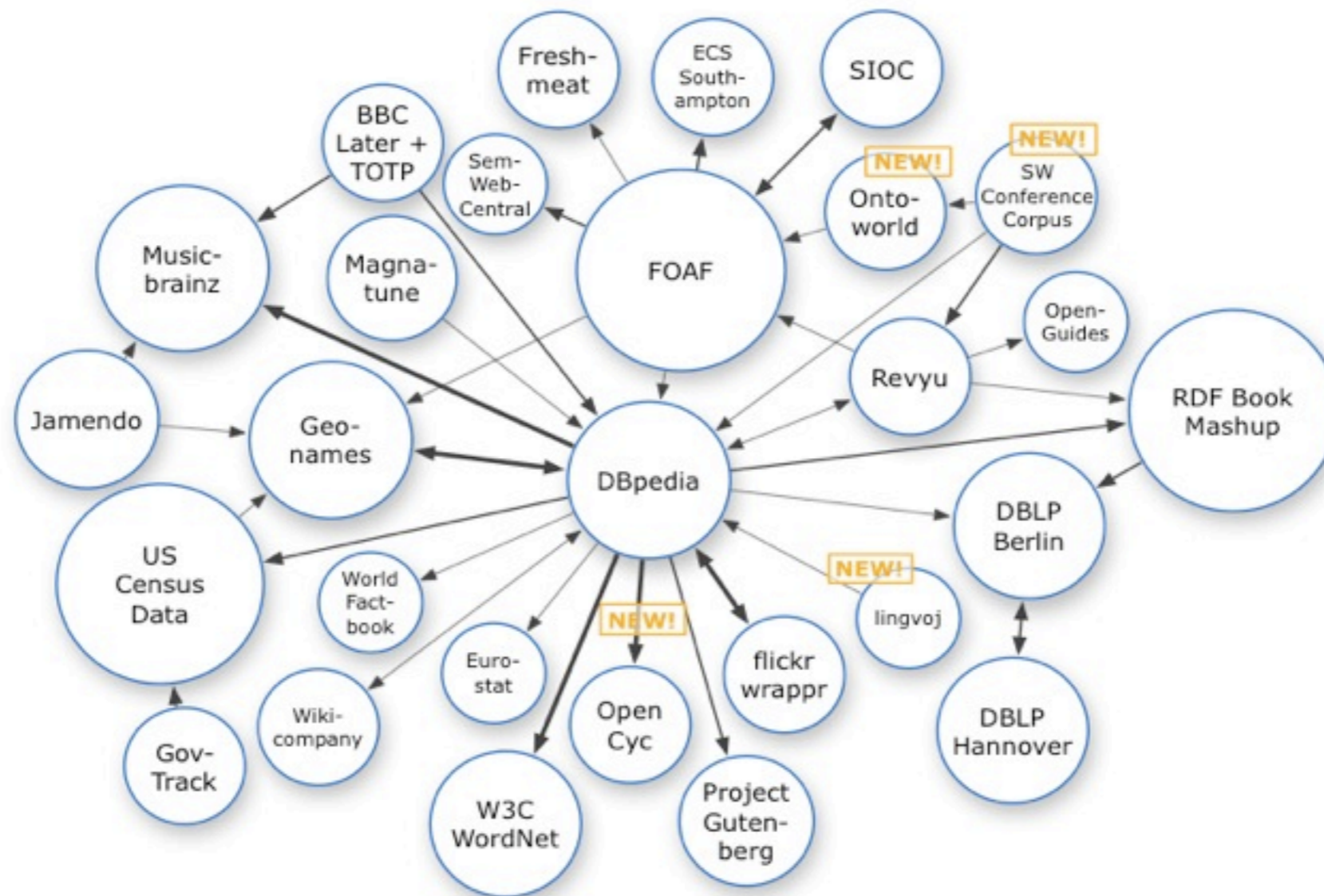
Linked data, October 2007



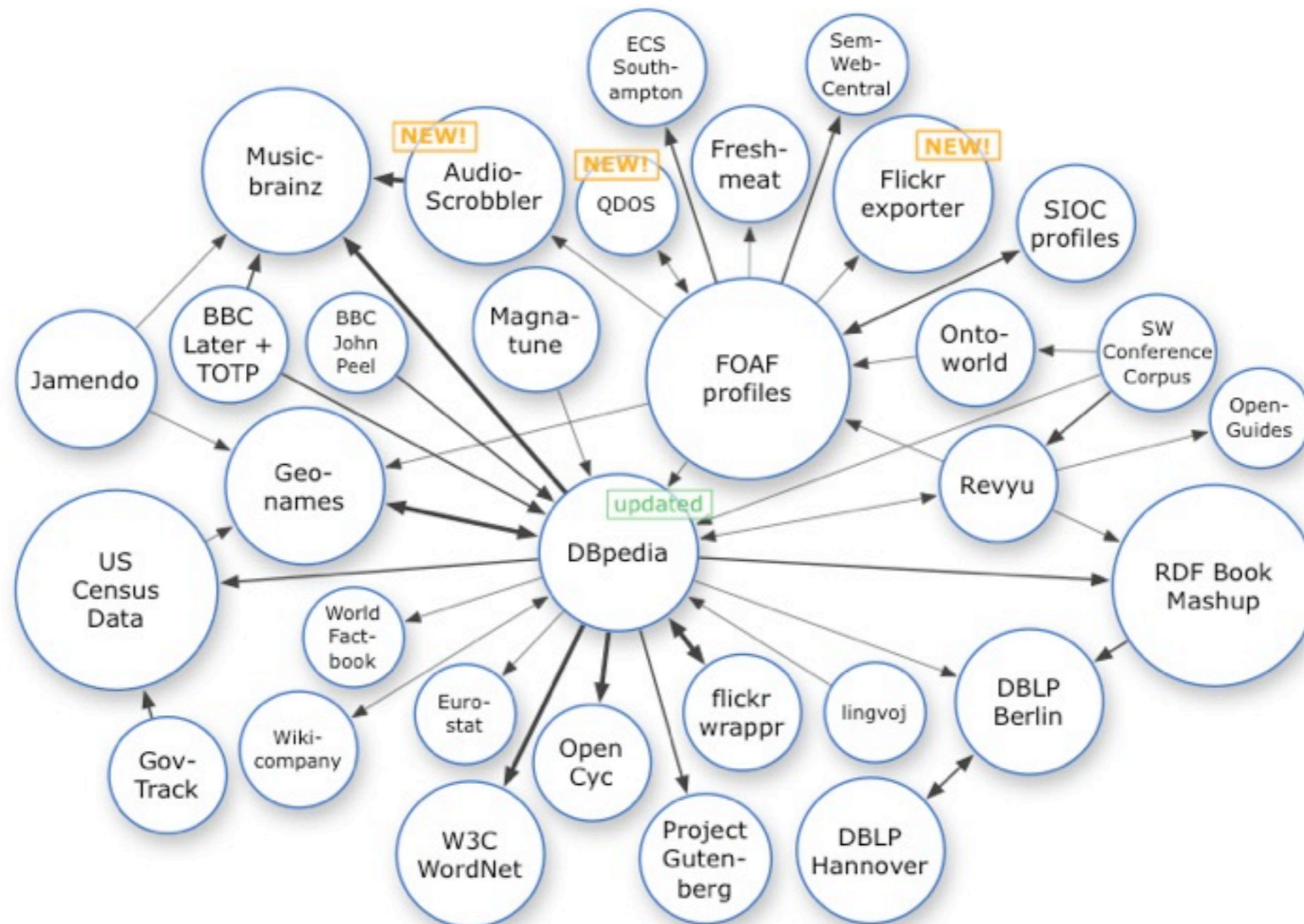
Linked data, November 2007



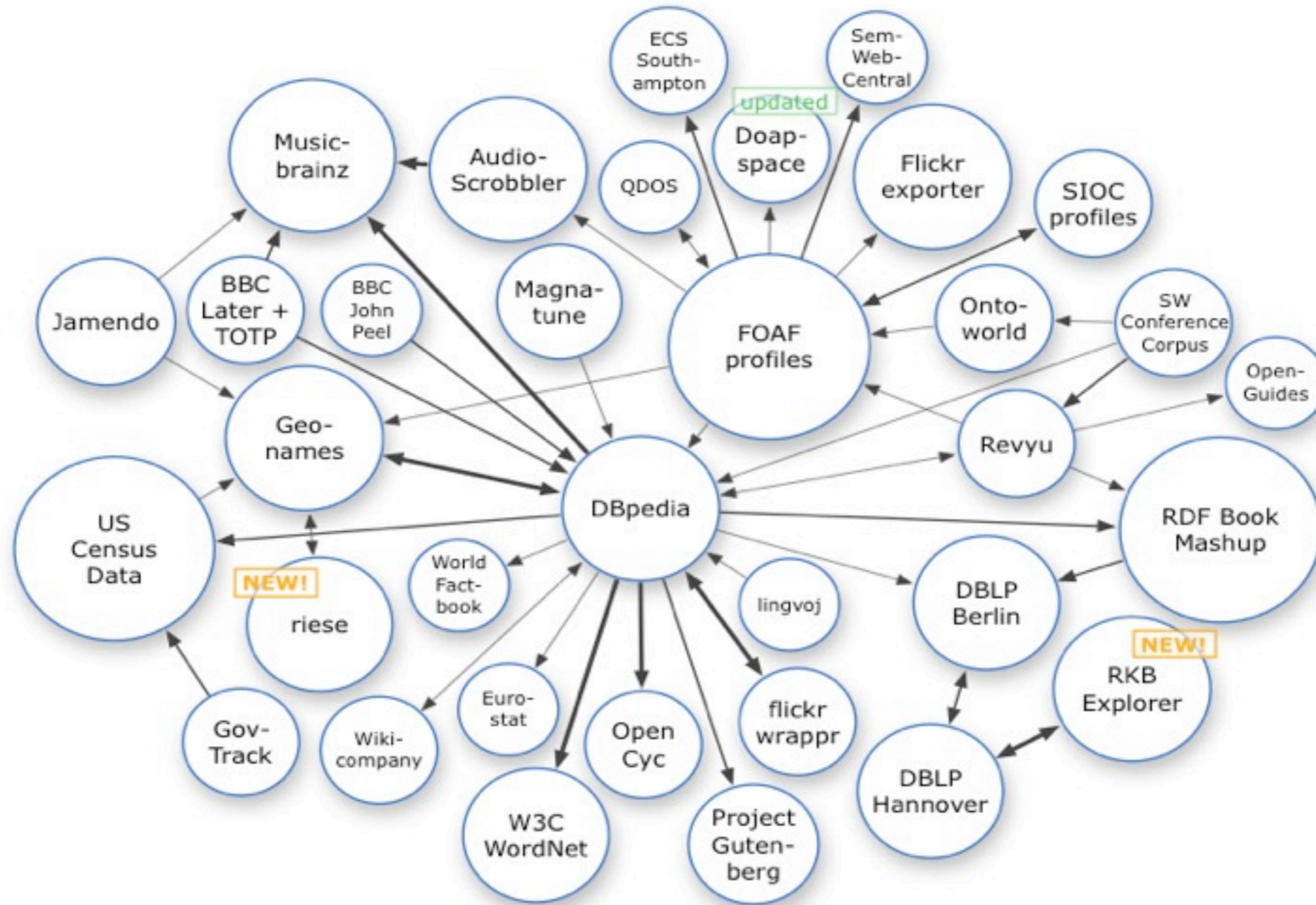
Linked data, December 2007

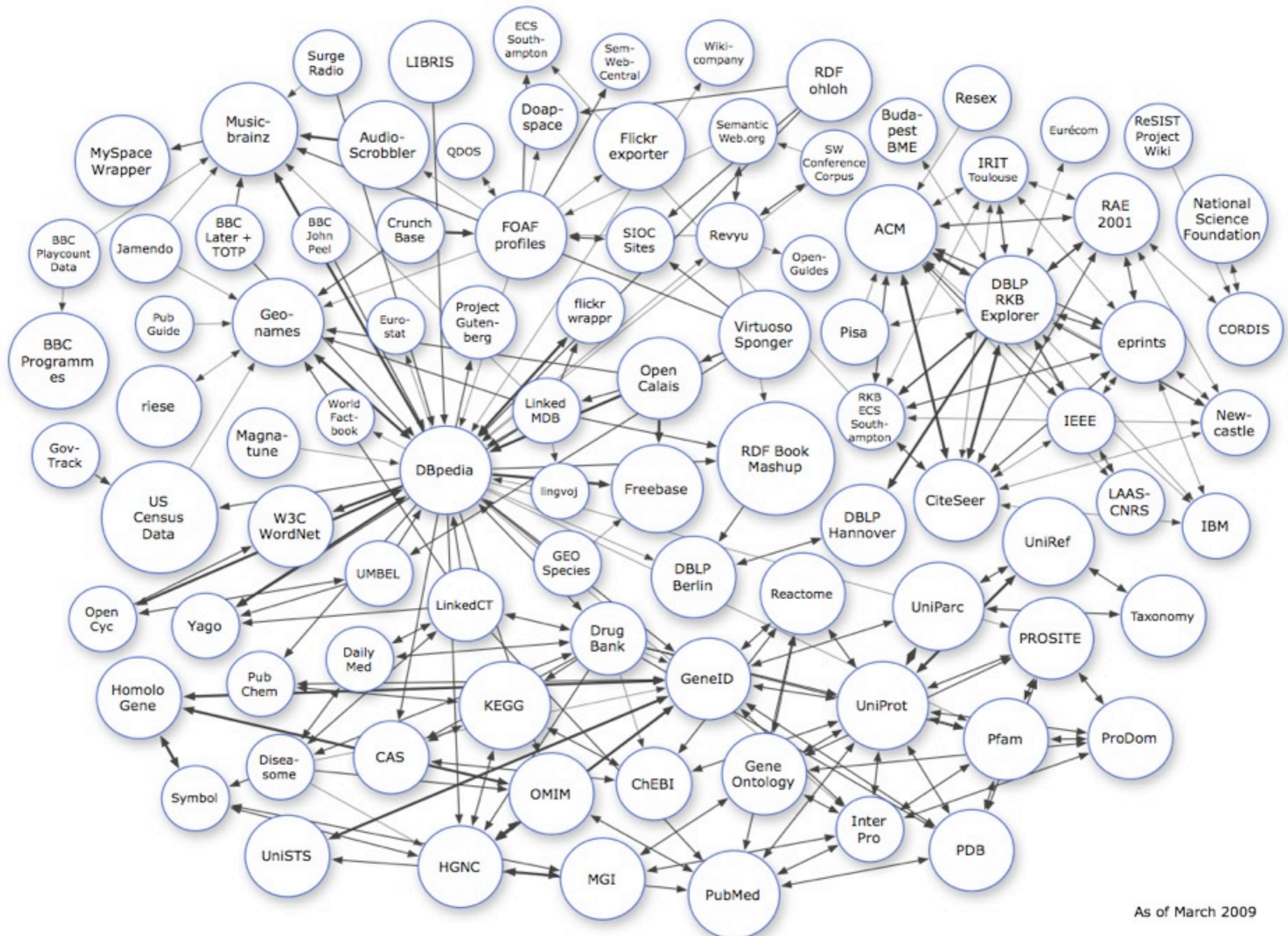


Linked data, February 2008



Linked data, March 2008





As of March 2009

Linking Open Data

- Global set of connected open databases
- ca 5 billion RDF triples
- W3C project



BBC been there a while – and now the New York Times....

- 2009–10–30: The New York Times has started to publish parts of its subject headings as Linked Data under a CC BY license.



The “semantic task force” Summary for MLR

- Semantic technologies allow for
 - Large-scale interoperability (triples, AP-independent, follow-your-nose, linked data)
 - Ontology support
 - Reuse of existing standards
 - Collaboration between standards bodies
 - Reuse of existing tools
 - Implementation in many environments
 - From mobile or AJAX applications
 - Through HTML (RDFa) and RSS
 - To multi-billion-triples RDF stores

Requirements for Reusability

- The components must be **unambiguously identified**
- The components must **adhere to compatible abstract models.**
- A metadata format must be used that allows for **consistent interpretation** of the components with respect to their respective abstract models.

Metadata – not just an index

- Metadata is not always objective information
 - It must allow for subjective expressions & opinions.
- Metadata is not produced “once and for all”
 - an eco-system of metadata.
- Metadata is not just a document
 - But a global network of information.
- Metadata is much more than just a digital catalog.
- Metadata is not just for machines
 - We need conceptual metadata for people!

Four rules for exposing information on the web

- Use URIs as names for things
- Use HTTP URIs so that people can look up those names.
- When someone looks up a URI, provide useful information.
- Include links to other URIs, so that they can discover more things.

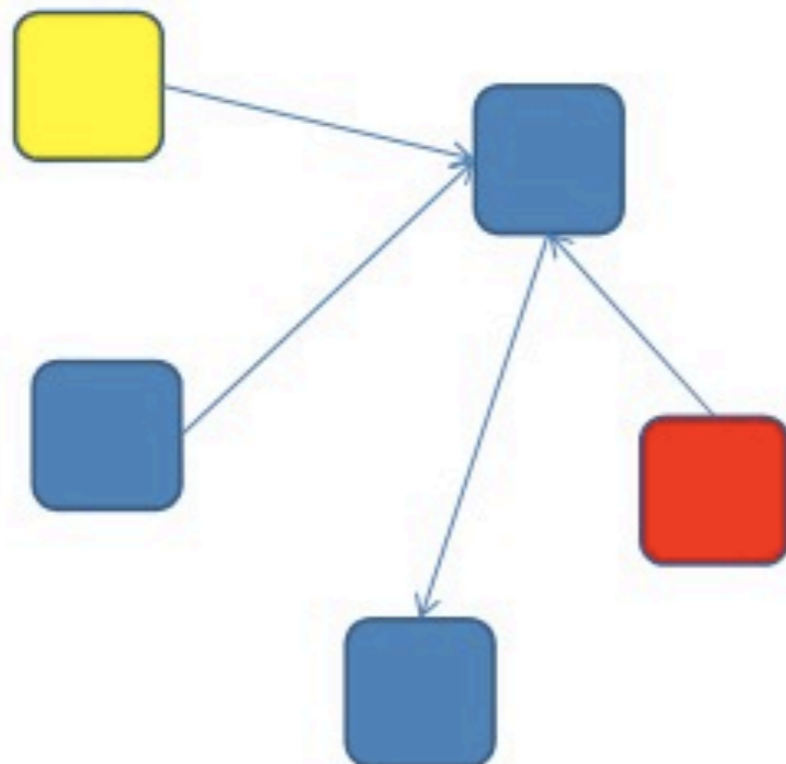
The “new” MLR

- Final Committee Draft still to be balloted (before March 2010)
- MLR-1 Framework showing how MLR elements are mapped to RDF
- Identifiers that enables URIs
- Easy to produce the new parts (Educational, Technical, Rights, etc.)



Where do metadata live?

In centralized or distributed Learning Object Repositories



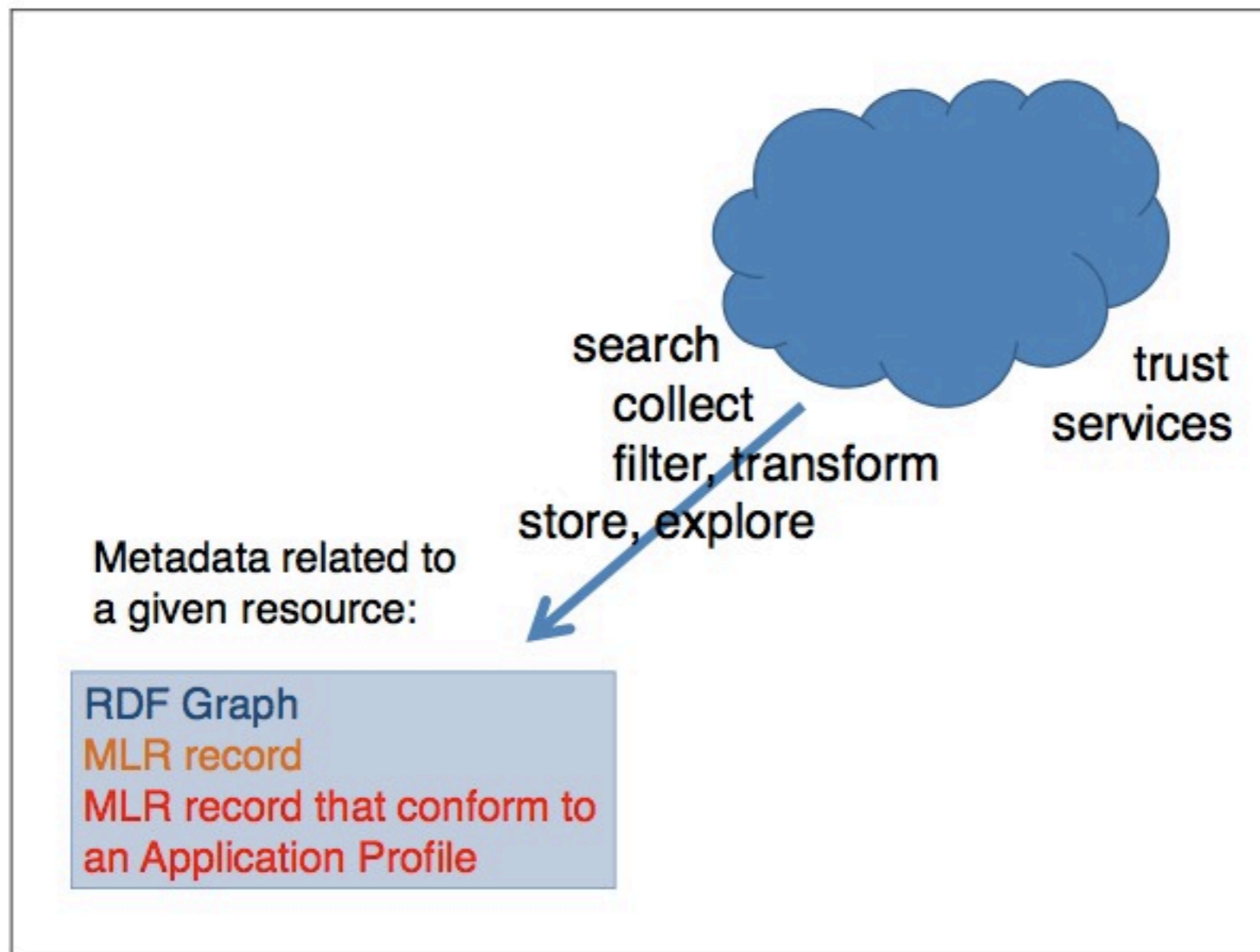
Free in the Web



Who contributes metadata for a learning resource?

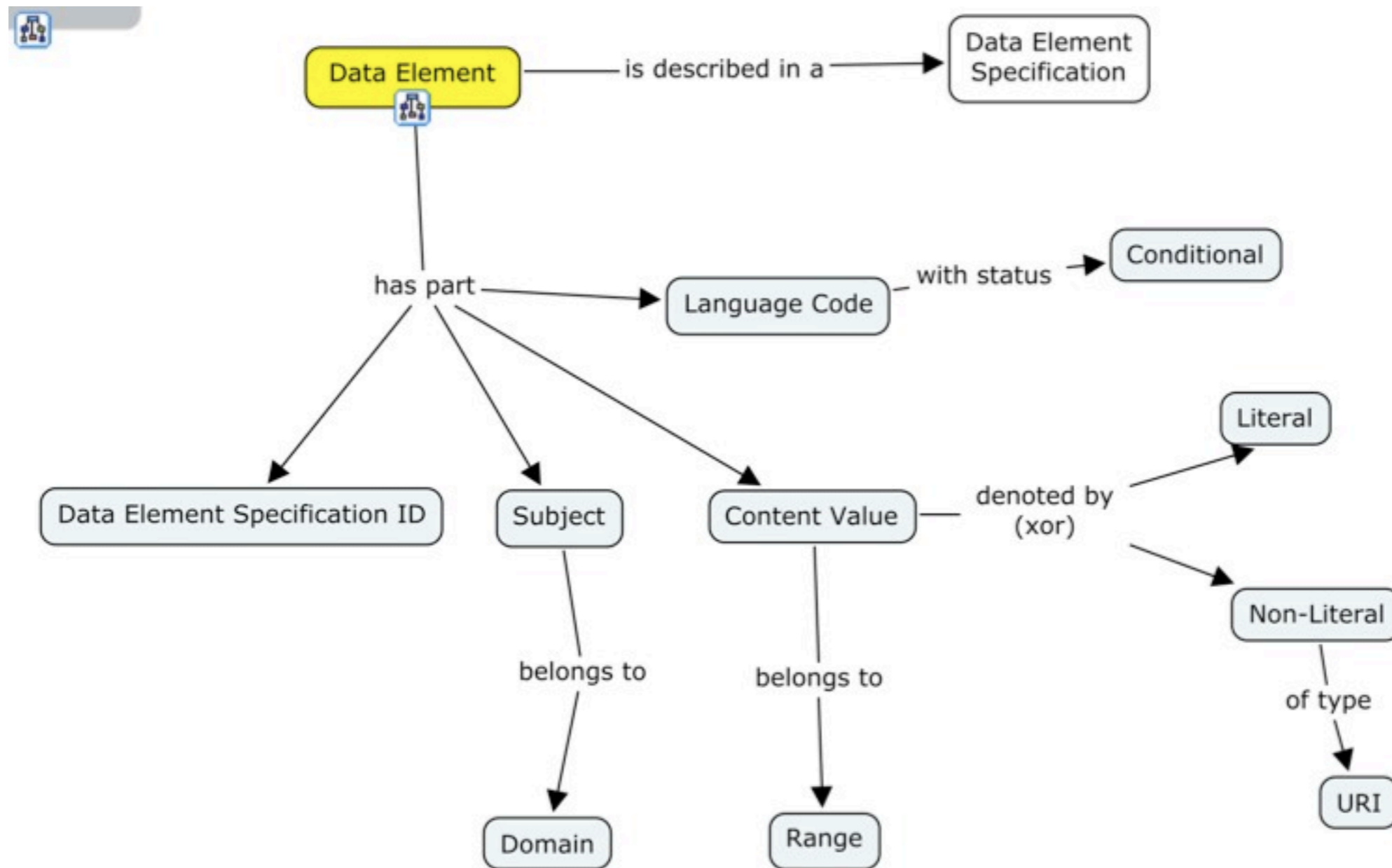
Source: ISO/IEC JTC1 SC36 WG4 N0340

The MLR response



Source: ISO/IEC JTC1 SC36 WG4 N0340

Conceptual modelling



Concept map for Data Elements from FCD of MLR-1

More transparent processes

- MLR Educational
 - Open online meetings
 - Concepts maps
 - Consultations with the DC community and stakeholder groups
- MLR Educational and Dublin Core Ed is going to exchange maps (as of this morning :-))

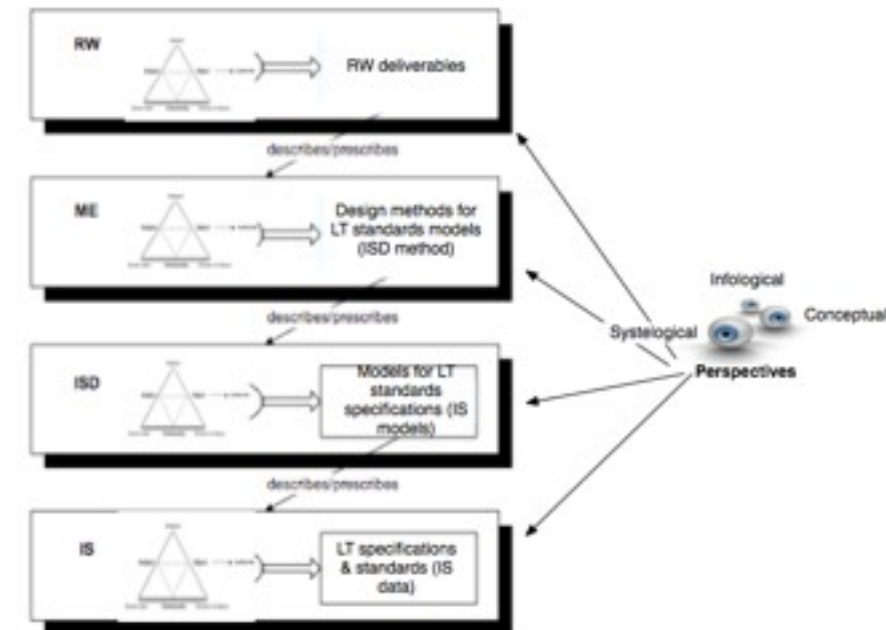
First lesson to be learnt: Understand your domain!

- Is Learning, Education and Training (LET) best supported by an eBusiness approach or a web architecture approach?
- Is LET about exactness or fuzziness?
- Is LET about truth or points of views?
- Is LET a well or a ill-defined domain?

Google for New York Times and see what you get

Lessons to be learnt?

- Don't embark on a standards project without a metamodel and methodology
- When you cannot do much about the personalities...
 - ... make sure you control the tools
 - ... learn the process and use it for your own purpose
 - ... foster transparency



Thanks & More Information

- Thanks to Mikael Nilsson for allowing reuse of his slides
- For more information:
 - tore.hoel@hio.no
 - www.icoper.org
 - www.hoel.nu/wordpress