Catalogues, Dictionaries, Libraries, Data and Metadata

Gerald Radack

2009-03-19

Kinds of IT Standards

- Exchange standards
  - define a document form for conveying info
  - languages define data elements and structures
- Interface standards
  - define an interaction for providing a service
  - simple interactions defined by messages or invocations (request/response messages)
  - complex interactions defined by choreography of message exchanges

- Ed Barkmeyer

- Language standards
- Vocabulary standards
  - define concepts within a field using terms, definitions and abbreviations

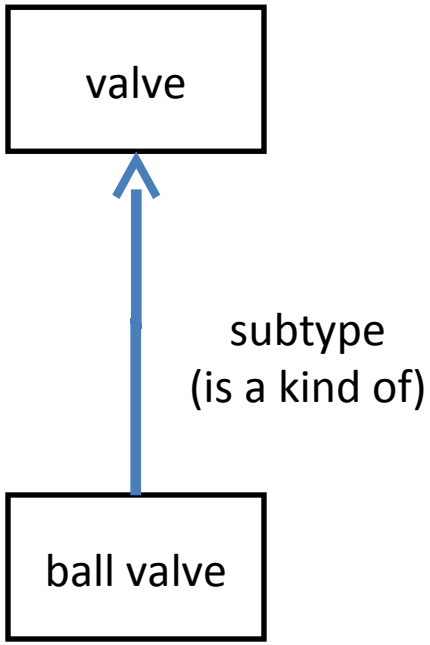
Extending Data Models

- Assumption: standardizing data models is expensive and slow, standardizing reference data is less expensive and faster

Extending Data Models

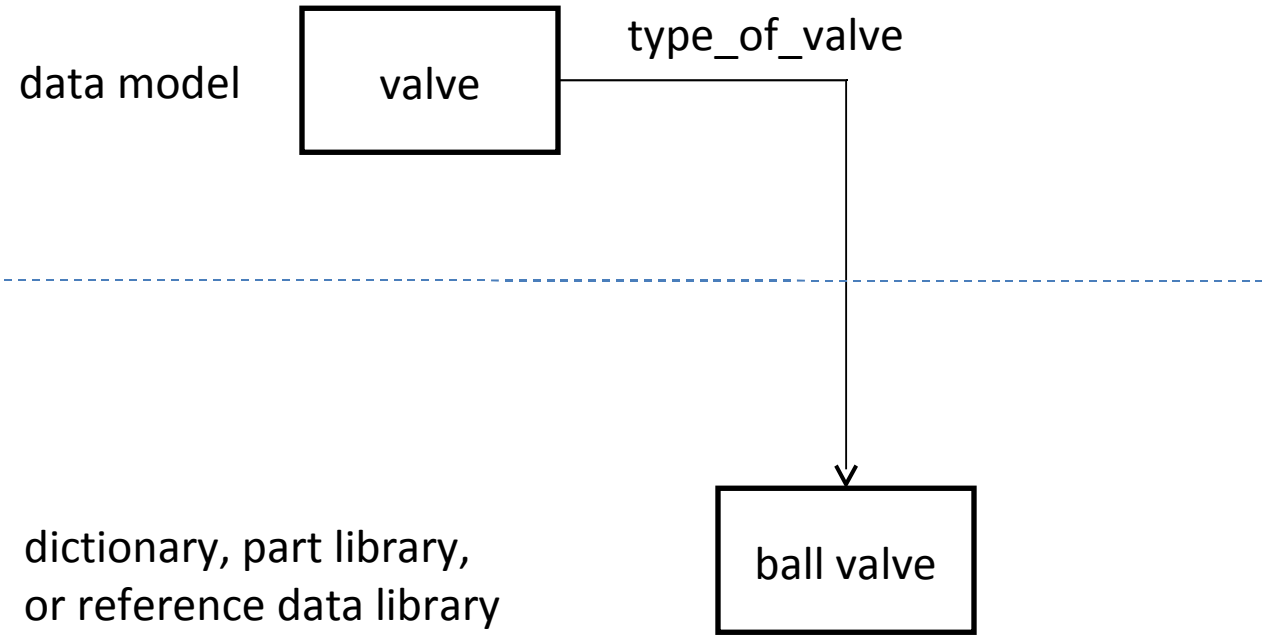
- Method 1

data model

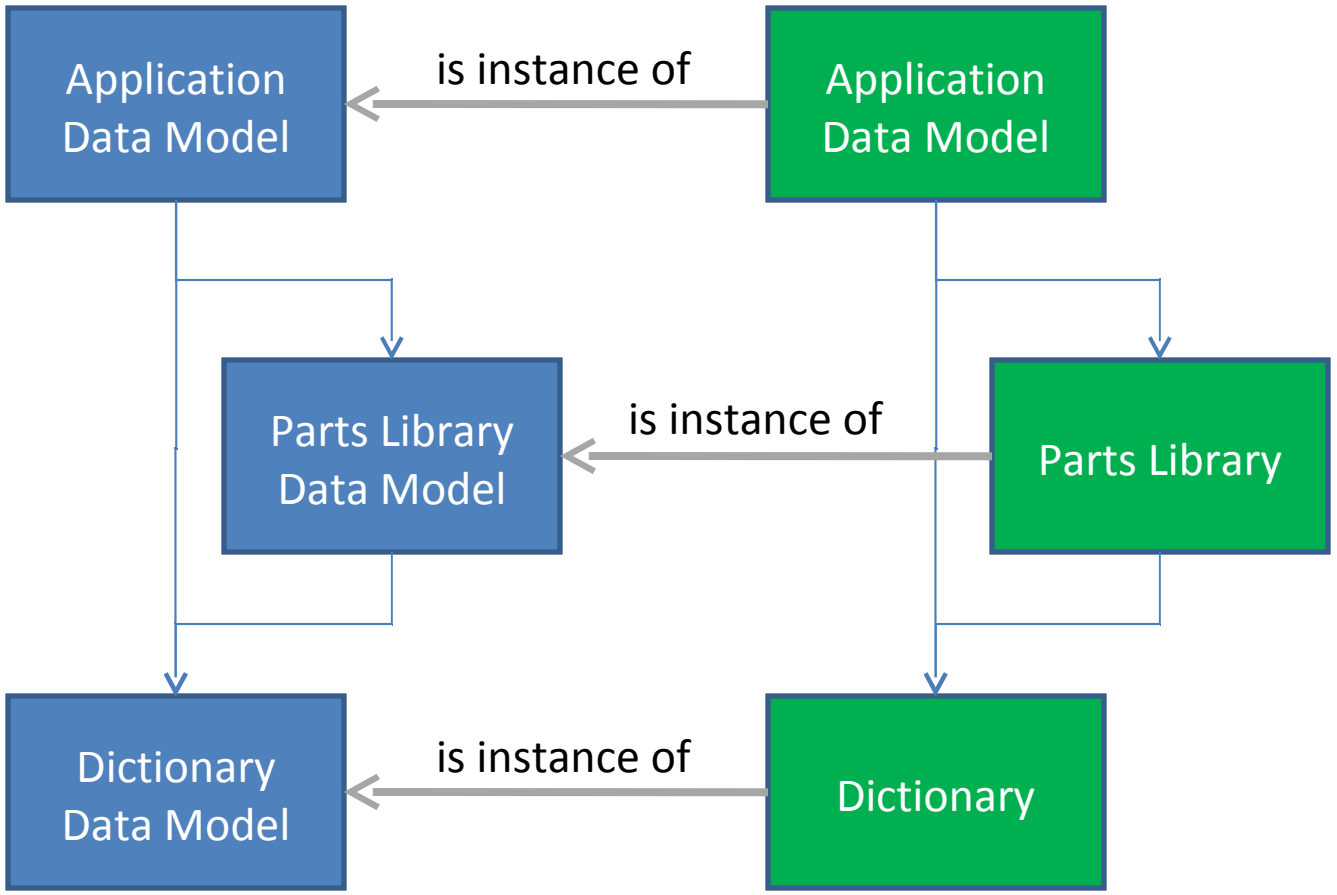


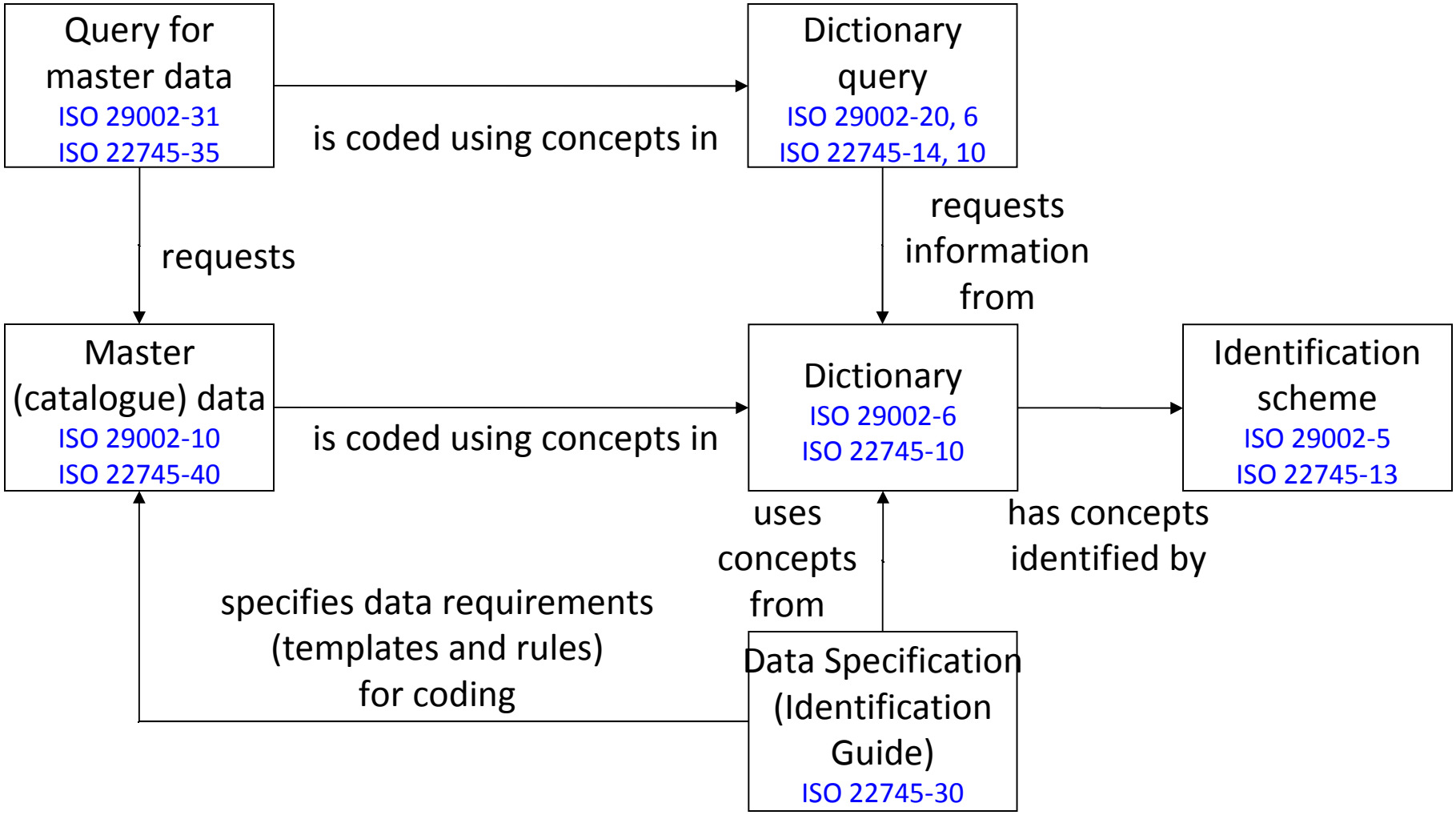
Extending Data Models

- Method 2



Typical Architecture




Query for master data
ISO 29002-31
ISO 22745-35
is coded using concepts in
Dictionary query
ISO 29002-20, 6
ISO 22745-14, 10
requests
requests information from
Master (catalogue) data
ISO 29002-10
ISO 22745-40
is coded using concepts in
Dictionary
ISO 29002-6
ISO 22745-10
Identification scheme
ISO 29002-5
ISO 22745-13
uses concepts from
has concepts identified by
specifies data requirements (templates and rules) for coding
Data Specification (Identification Guide)
ISO 22745-30

SC 4 Products

- Data models
  - Multiple modeling languages in use: IDEF1X, EXPRESS, UML, etc.
  - Different modeling styles and patterns
- Dictionaries / part libraries / reference data libraries
  - Represented according to several data models
- Vocabulary
- EXPRESS and other foundational standards

Issues and Needs

- Seamless integration
  - Vertical (application → parts library → dictionary → EXPRESS meta-model)
  - Horizontal — as needed
  - Can ontology languages avoid the need for heterogeneous representation of concepts (data model, reference data dichotomy, etc.)
- Definition of “ontology”
- Migration path — providing existing models in multiple formats, including EXPRESS, Part 21 based, and OWL
- Management of identifiers/URIs
- Filling in gaps in application-level standards

The Integration of Taxonomies Using Ontology Structures

Presented by Gerry Radack, *CTC*

2009-03-18

The Integration of Taxonomies Using Ontology Structures

- Relate Government/Commercial Item Descriptions/Taxonomies to Supplier Capabilities
  - DLIS must be able to integrate DOD item classifications with commercial systems
  - Commercial cataloging systems have domain-specific classifications that are not interconnected
  - Information is stored in different, often proprietary, and incompatible formats

Approach: Guiding Principle

- Use the ECCMA Open Technical Dictionary (eOTD) as the basis for the integration of classifications
- The eOTD is the industry version of the FCS and seeks recognition as the international standard for e-catalogs via the ISO 22745 designation

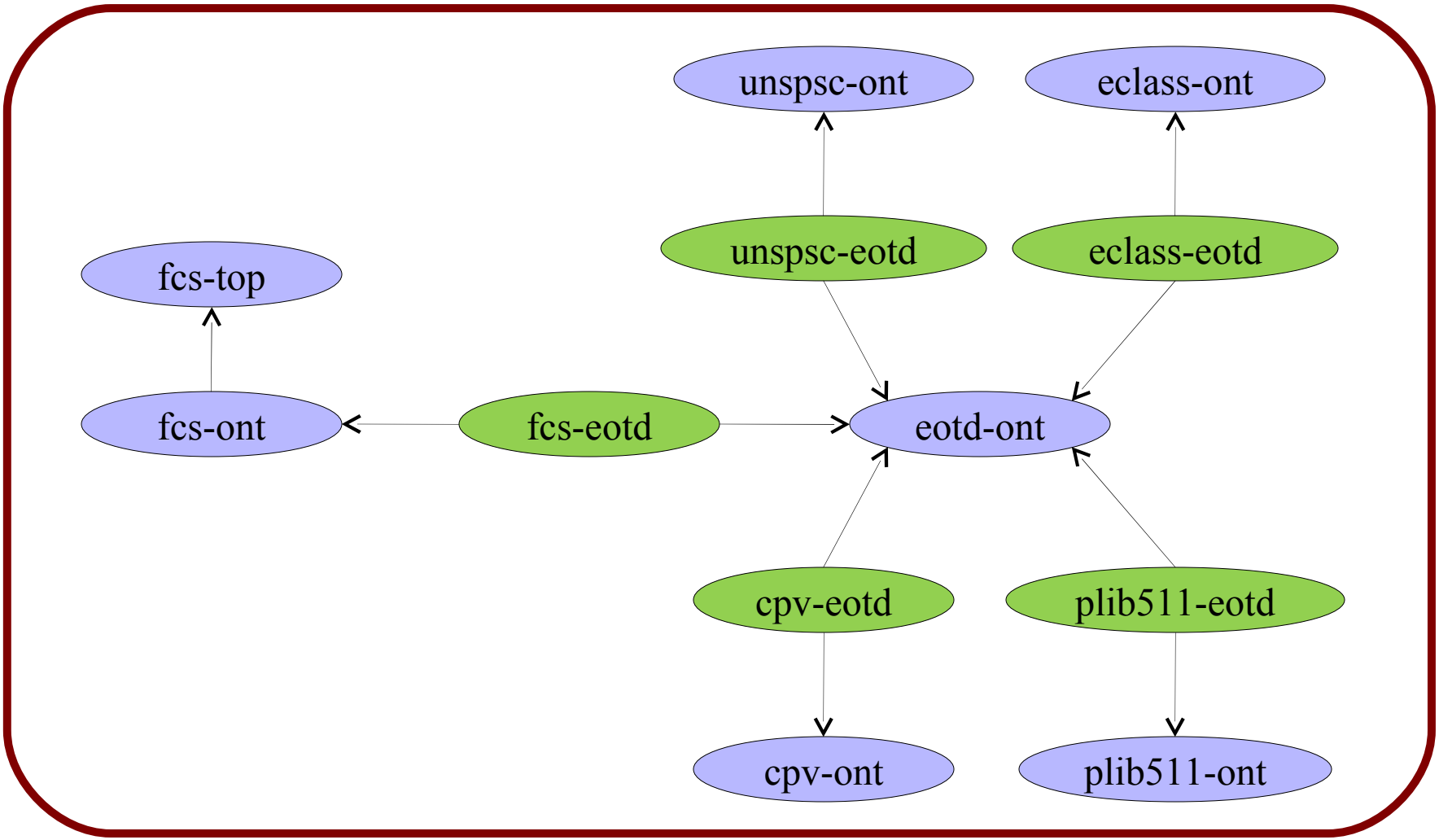
Mapping Between Taxonomies

- Mapping keywords is insufficient
  - one-to-one correspondences aren't always possible
  - overlapping classes
    - » functional vs. compositional classifications
    - » e.g. grinding machine:cutlery (0161-1#01-007071#1) vs. grinding machine:carbide tool bit (0161-1#01-007098#1)
- Need deeper analysis
- Focus on systemic classification of attributes that connects communities of information
- This connection can be established with the Semantic Web

Ontology Mapping

- OWL axioms are used to map each ontology to the eOTD
  - fsc:BearingsPlainUnmounted ≡ eotd:BearingPlain ∩ eotd:BearingUnmounted
  - fsc:BushingsRingsShimsAndSpacers ≡ eotd:Bushing ∪ eotd:Ring ∪ eotd:Shim ∪ eotd:Spacer

Ontologies Overview



Challenges

- Taxonomies differ in scope and purpose
- Naming conventions differ across classifications
  - e.g. “bearing, roller” versus “roller bearing”
- Target taxonomies have one or more deficiencies:
  - lack of definitions or inaccurate definitions
  - lack of freely available electronic version
  - lack of sample data
  - poor superclass/subclass structures
  - inconsistent modeling
  - failure to state/observe modeling conventions

Lack of Strict is-a Hierarchy

- OWL defines strict **is-a** hierarchies
  - *A* rdfs:subClassOf *B* is interpreted as "Every *A* is a *B*"
- Many product taxonomies are not is-a hierarchies
  - They were created to support purchasing
  - eCl@ss example:

27 Electric engineering, automation, process control engineering
- 27-05 Accumulator, battery
- 27-05-01 Station. batt., accum.
- 27-05-02 Traction battery, starter battery
- 27-05-04 Portable battery
- 27-05-06 Battery charger
- 27-05-90 Accumulator, battery (other)
- 27-05-91 Accumulator, battery (parts)
- 27-05-92 Accumulator, battery (accessories)
- 27-05-98 Accumulator, battery (maintenance, service)
- 27-05-99 Accumulator, battery (repair)

These are not batteries!

Integration Demonstration

- DL reasoner was used to integrate the FCS ontology with the target ontologies
  - Computed which target classes are implicit subclasses of FCS classes (subsumption)
  - Automatically "merged" two taxonomies

Project Team and Responsibilities

- Lehigh University
  - Project Management
  - Ontology Development
  - Taxonomy Integration
  - Translation Compiler Development
  - NSN Screening Tool Development
- ECCMA
  - Development of terms and definitions
  - Inclusion of taxonomy terminology in eOTD
  - Facilitation of OWL output from eOTD
- CTC
  - Technical guidance on data modeling, eOTD core model, and ISO 22745
  - ISO/IEC JTC1/SC32 metadata standards
  - ISO TC37 terminology standards