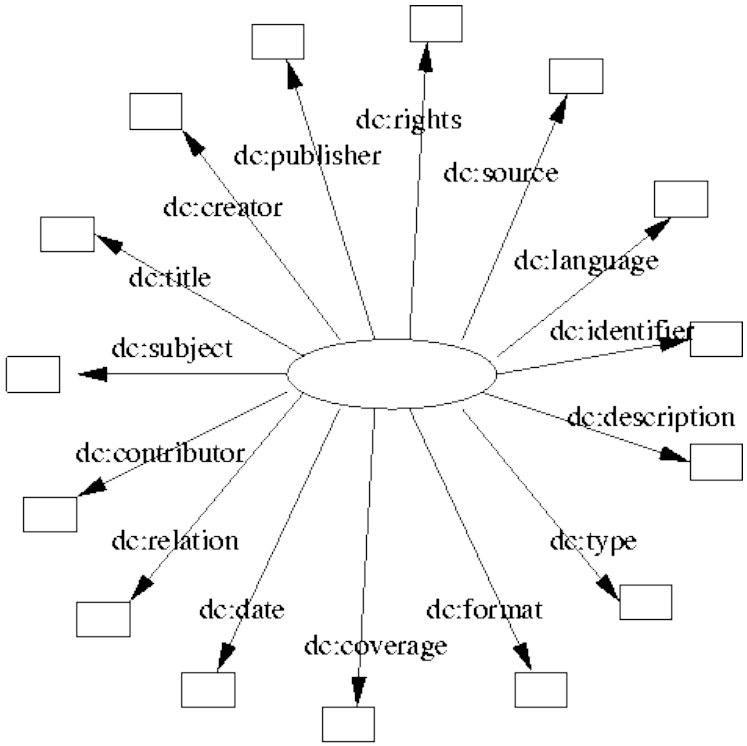
Metadata for Research and Education

Odile Bénassy
OFSET

Taiwan, Oct.2009

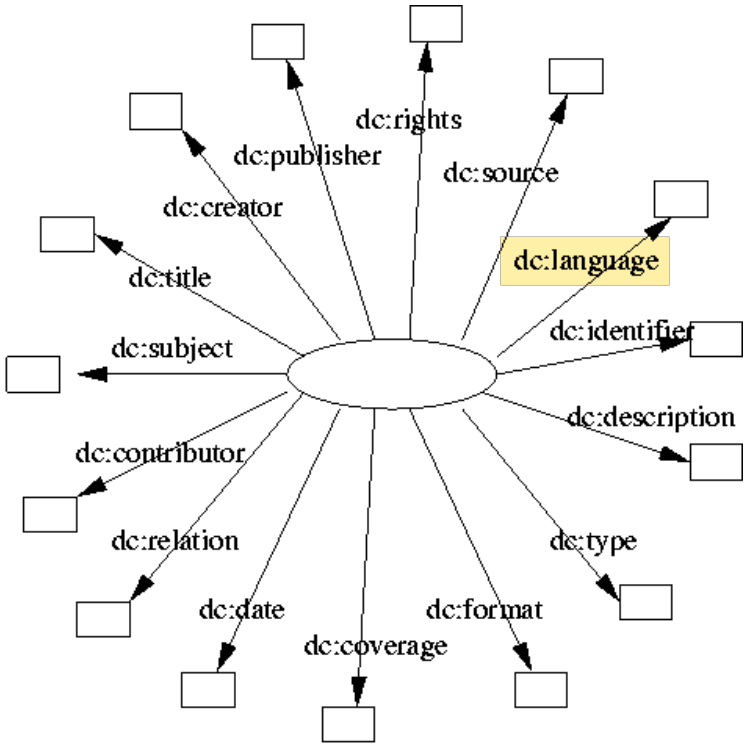
Outline

Dublin Core Hedgehog



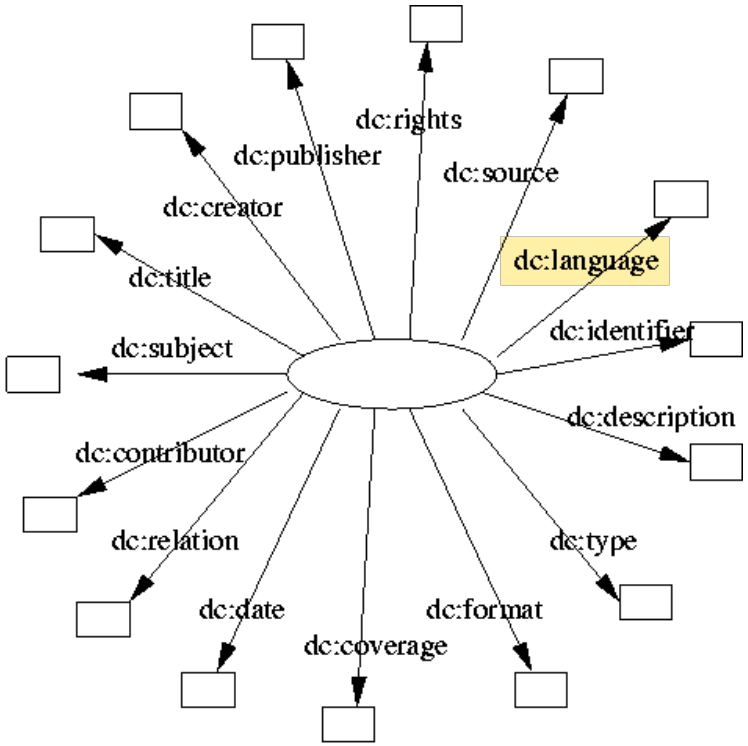
(source: Dublin Core Metadata Initiative)

Controlled Vocabulary



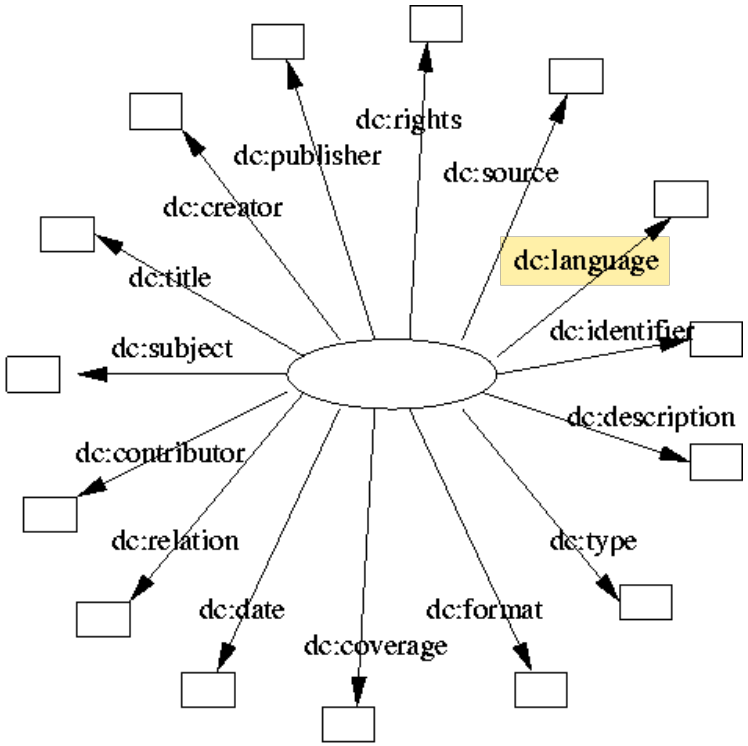
- attribute = Language

Controlled Vocabulary



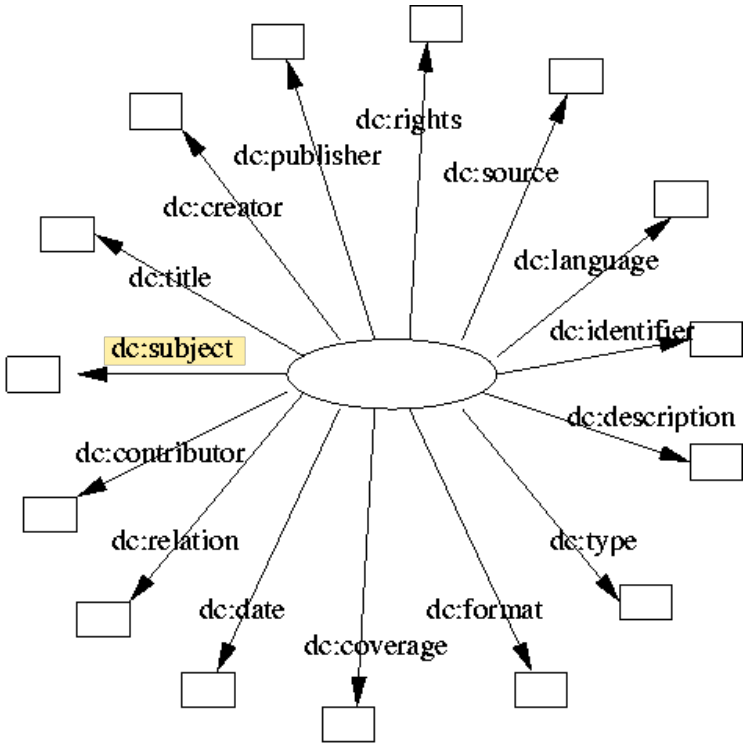
- attribute = Language
- standard classification = iso 639-2 – iso 639-3

Controlled Vocabulary



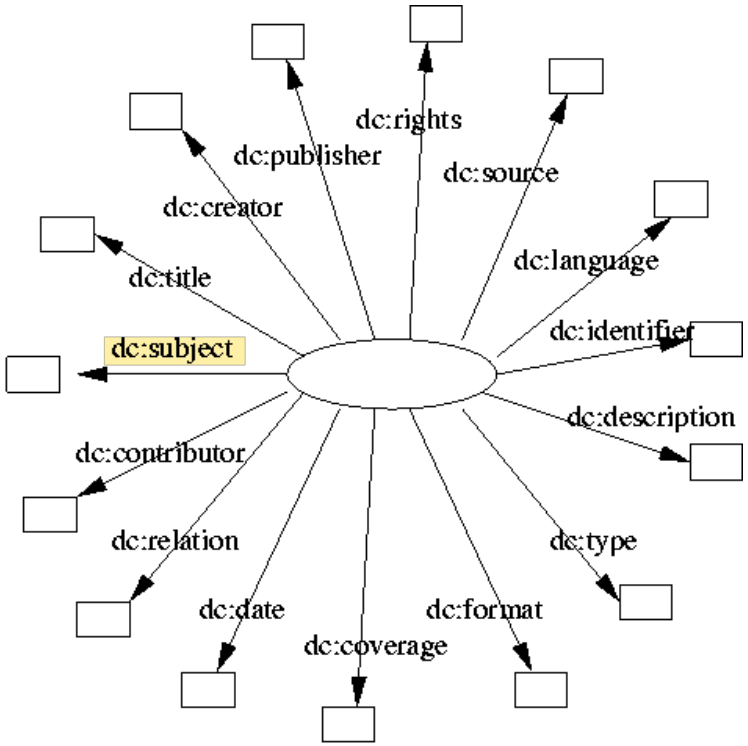
- attribute = Language
- standard classification = iso 639-2 – iso 639-3
- issues = Chinese: written or spoken?

Thesaurus



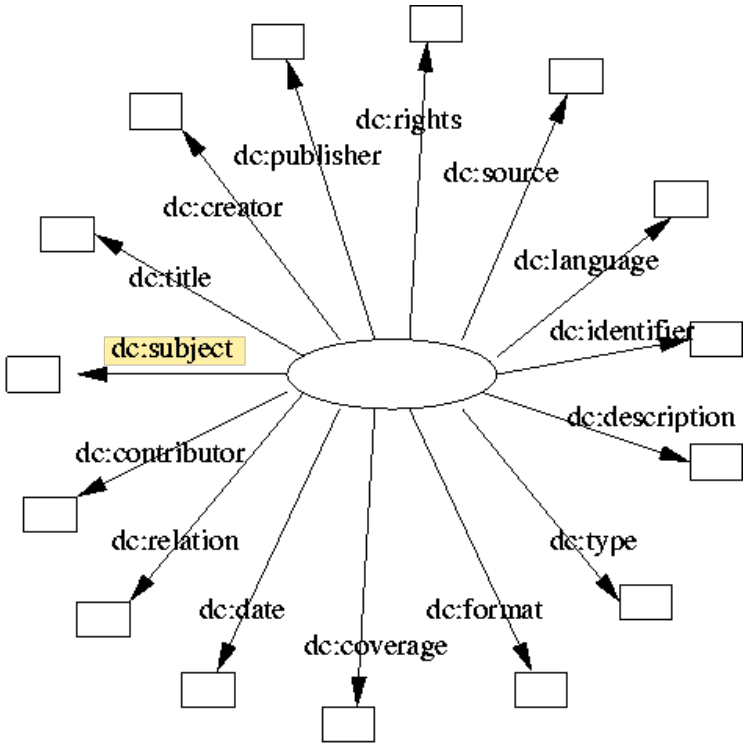
- attribute = Subject

Thesaurus



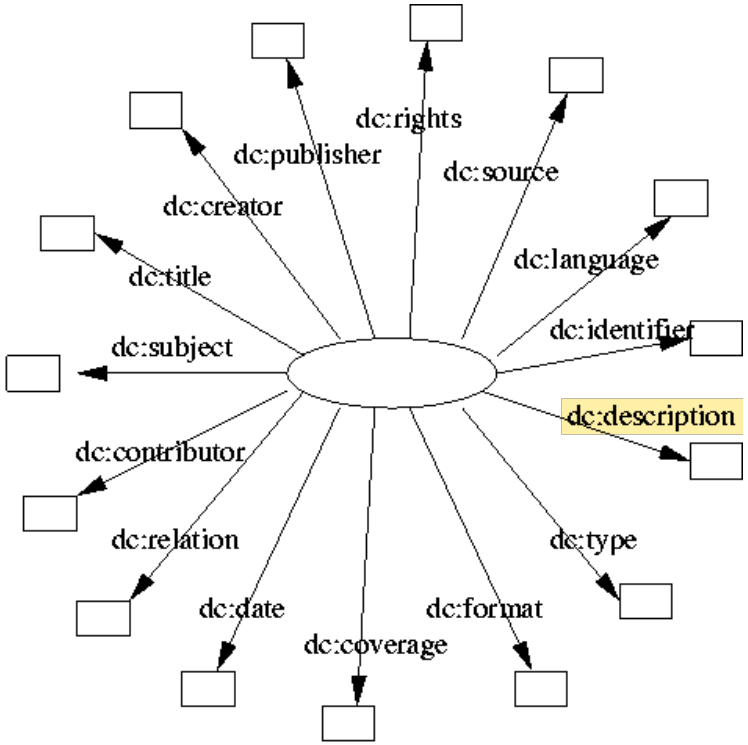
- attribute = Subject
- standard classification = Dewey/Library of Congress

Thesaurus



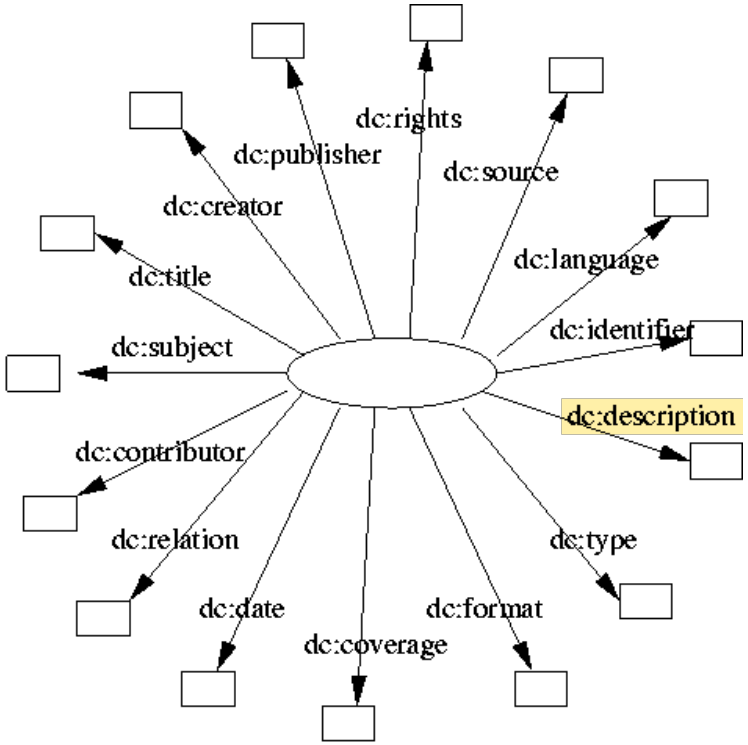
- attribute = Subject
- standard classification = Dewey/Library of Congress
- issues = educational subject depends on audience

Thesaurus (2)



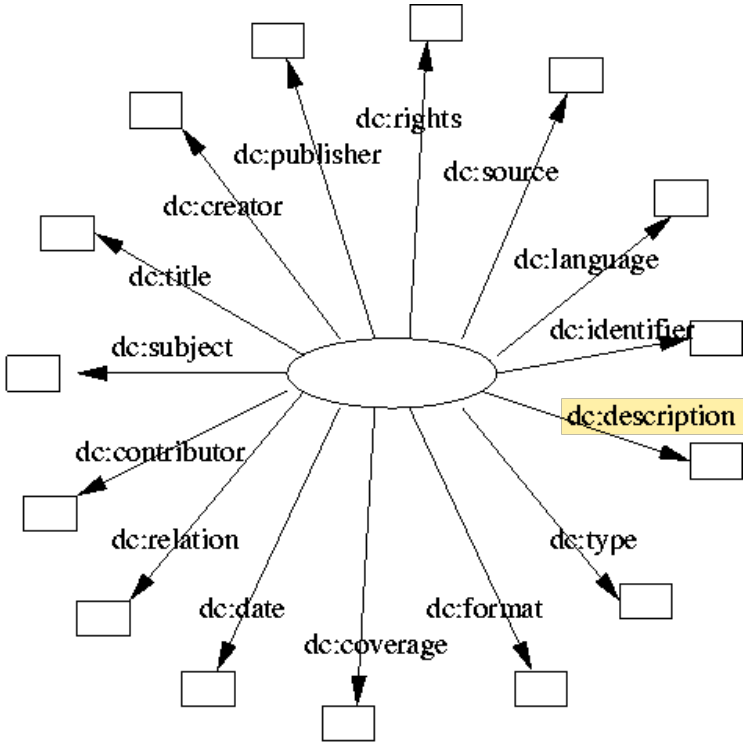
- attribute = Keyword

Thesaurus (2)



- attribute = Keyword
- standard classification? expl “Rameau” vocab. at the French National Library (BNF)

Thesaurus (2)



- attribute = Keyword
- standard classification? expl “Rameau” vocab. at the French National Library (BNF)
- issues = translation non univocal
  - “show” and “demonstrate” → same word in french

How do you export?

Interoperability: Exporting/Importing

- RSS, Atom = for Dublin Core only

How do you export?

Interoperability: Exporting/Importing

- RSS, Atom = for Dublin Core only
- OWL = rich but difficult

How do you export?

Interoperability: Exporting/Importing

- RSS, Atom = for Dublin Core only
- OWL = rich but difficult
- SKOS = easier
  - Library of Congress Subject Headings
  - Rameau

How do you export?

Interoperability: Exporting/Importing

- RSS, Atom = for Dublin Core only
- OWL = rich but difficult
- SKOS = easier
  - Library of Congress Subject Headings
  - Rameau
- OAI = output format + harvesting protocol

How do you export?

Interoperability: Exporting/Importing

- RSS, Atom = for Dublin Core only
- OWL = rich but difficult
- SKOS = easier
  - Library of Congress Subject Headings
  - Rameau
- OAI = output format + harvesting protocol
- JSON = object serialization, not standardized

How do you export?

Interoperability: Exporting/Importing

- RSS, Atom = for Dublin Core only
- OWL = rich but difficult
- SKOS = easier
  - Library of Congress Subject Headings
  - Rameau
- OAI = output format + harvesting protocol
- JSON = object serialization, not standardized
- REST = highly useful but not standardized in itself

A New Thinking

A New Thinking ref. newthinking.de

- Physical vs Electronical

A New Thinking

A New Thinking ref. newthinking.de

- Physical vs Electronical
- Text vs Hypertext

A New Thinking

A New Thinking ref. newthinking.de

- Physical vs Electronical
- Text vs Hypertext
- Hierarchical vs Semantic

A New Thinking

A New Thinking ref. newthinking.de

- Physical vs Electronical
- Text vs Hypertext
- Hierarchical vs Semantic
- Tree vs Graph

A New Thinking

A New Thinking ref. newthinking.de

- Physical vs Electronical
- Text vs Hypertext
- Hierarchical vs Semantic
- Tree vs Graph
- Ontology: grammar

A New Thinking

A New Thinking ref. newthinking.de

- Physical vs Electronical
- Text vs Hypertext
- Hierarchical vs Semantic
- Tree vs Graph
- Ontology: grammar
- Example: Harry Potter by Gautier Poupeau (French)

Outline

Physical Library Standards

Research Papers in Physical Libraries

- MARC, UNIMARC, Z39-50 (BiblioML?)

Research Papers in Physical Libraries

- MARC, UNIMARC, Z39-50 (BiblioML?)
- Flat Records

Research Papers in Physical Libraries

- MARC, UNIMARC, Z39-50 (BiblioML?)
- Flat Records
- Physical Limitations: only one shelf

Open Access Initiative

- BOAI (B for Budapest)

Open Access Initiative

- BOAI (B for Budapest)
- OAI format and query protocol

Open Access Initiative

- BOAI (B for Budapest)
- OAI format and query protocol
- Example: HAL (Hyper Articles en Ligne) ≈130 000 full-text documents

Open Access Initiative

- BOAI (B for Budapest)
- OAI format and query protocol
- Example: HAL (Hyper Articles en Ligne) ≈130 000 full-text documents
- Example: IFREMER, Archimer

Open Access Initiative

- BOAI (B for Budapest)
- OAI format and query protocol
- Example: HAL (Hyper Articles en Ligne) ≈130 000 full-text documents
- Example: IFREMER, Archimer
- Find repositories: OAIster, OpenDOAR

OAI Software

- GNU eprints (Perl)

OAI Software

- GNU eprints (Perl)
- Fedora Commons, SDX (Java)

OAI Software

- GNU eprints (Perl)
- Fedora Commons, SDX (Java)
- Library Software like Koha, Greenstone ..

OAI Software

- GNU eprints (Perl)
- Fedora Commons, SDX (Java)
- Library Software like Koha, Greenstone ..
- General use CMSs like Drupal

OAI Software

- GNU eprints (Perl)
- Fedora Commons, SDX (Java)
- Library Software like Koha, Greenstone ..
- General use CMSs like Drupal
- Specialized CMSs like Open Journal System, Lodel (PHP), Silva and MOAI (Python)

Outline

Problem Statement

What difficulties?

- “Thales”

Problem Statement

What difficulties?

- “Thales”
- Theory of Probabilities

Problem Statement

What difficulties?

- “Thales”
- Theory of Probabilities
- Yet: Ease of translation in a few cases (Science, Foreign language ..)

Problem Statement

What difficulties?

- “Thales”
- Theory of Probabilities
- Yet: Ease of translation in a few cases (Science, Foreign language ..)
- Finally: education is a vocation and has its own needs

The Learning Object Model (LOM)

Note: In French: LOM-fr, AFNOR

- General: title, language, description, subject, document type, aggregation level, keywords, publisher ..

The Learning Object Model (LOM)

Note: In French: LOM-fr, AFNOR

- General: title, language, description, subject, document type, aggregation level, keywords, publisher ..
- Technical: format, file size, required software, duration (video) ..

The Learning Object Model (LOM)

Note: In French: LOM-fr, AFNOR

- General: title, language, description, subject, document type, aggregation level, keywords, publisher ..
- Technical: format, file size, required software, duration (video) ..
- Legal: author(s), rights, version, status (draft ..)

The Learning Object Model (LOM)

Note: In French: LOM-fr, AFNOR

- General: title, language, description, subject, document type, aggregation level, keywords, publisher ..
- Technical: format, file size, required software, duration (video) ..
- Legal: author(s), rights, version, status (draft ..)
- Cataloging: URL, identifier, metadata on metadata

The Learning Object Model (LOM)

Note: In French: LOM-fr, AFNOR

- General: title, language, description, subject, document type, aggregation level, keywords, publisher ..
- Technical: format, file size, required software, duration (video) ..
- Legal: author(s), rights, version, status (draft ..)
- Cataloging: URL, identifier, metadata on metadata
- Educational: subject (more precise), level, audience, country (in some cases), learning resource type, difficulty, curriculum, age, skill/competency landmarking ..

LOM & Vocabularies

Vocabularies?

- Not always defined by the standard

LOM & Vocabularies

Vocabularies?

- Not always defined by the standard
- Multiple possible sources and crafts (technical, legal, educational)

LOM & Vocabularies

Vocabularies?

- Not always defined by the standard
- Multiple possible sources and crafts (technical, legal, educational)
- Politics, nations

LOM & Vocabularies

Vocabularies?

- Not always defined by the standard
- Multiple possible sources and crafts (technical, legal, educational)
- Politics, nations
- Education systems being so diverse

Interactive Geometry Interoperability

- Cabri, Geoplan, ..

Interactive Geometry Interoperability

- Cabri, Geoplan, ..
- Drawing, Scripting

Interactive Geometry Interoperability

- Cabri, Geoplan, ..
- Drawing, Scripting
- Cataloging, Indexing ..

Interactive Geometry Interoperability

- Cabri, Geoplan, ..
- Drawing, Scripting
- Cataloging, Indexing ..
- An OWL Ontology called GeoSkills cf GeoSkills Project Page and GNU Edu/Intergeo SKOS export page