

Annotation of Heterogeneous Database Content for the Semantic Web

Eero Hyvönen, Mirva Salminen, Miikka Junnila

Mirva Salminen

mirva.salminen@cs.helsinki.fi

Semantic Computing Research Group

Helsinki Institute for Information Technology (HIIT) & University
of Helsinki, Dept. of Computer Science

<http://www.cs.helsinki.fi/group/seco/>



Outline

- Specification for content and annotation process need
- Three step annotation process
 - Syntactic homogenization
 - Terminology creation
 - Annotation creation
- Conclusion

Specification for content need

- Museum items are distributed to different museums
- Museum web sites are built for each museum separately
 - The user is interested in the historical artifacts, not in museum organizations
- MuseumFinland is a shared semantic portal for searching and browsing museum items
 - Global view to distributed collections
 - One seamless national collection (virtually)



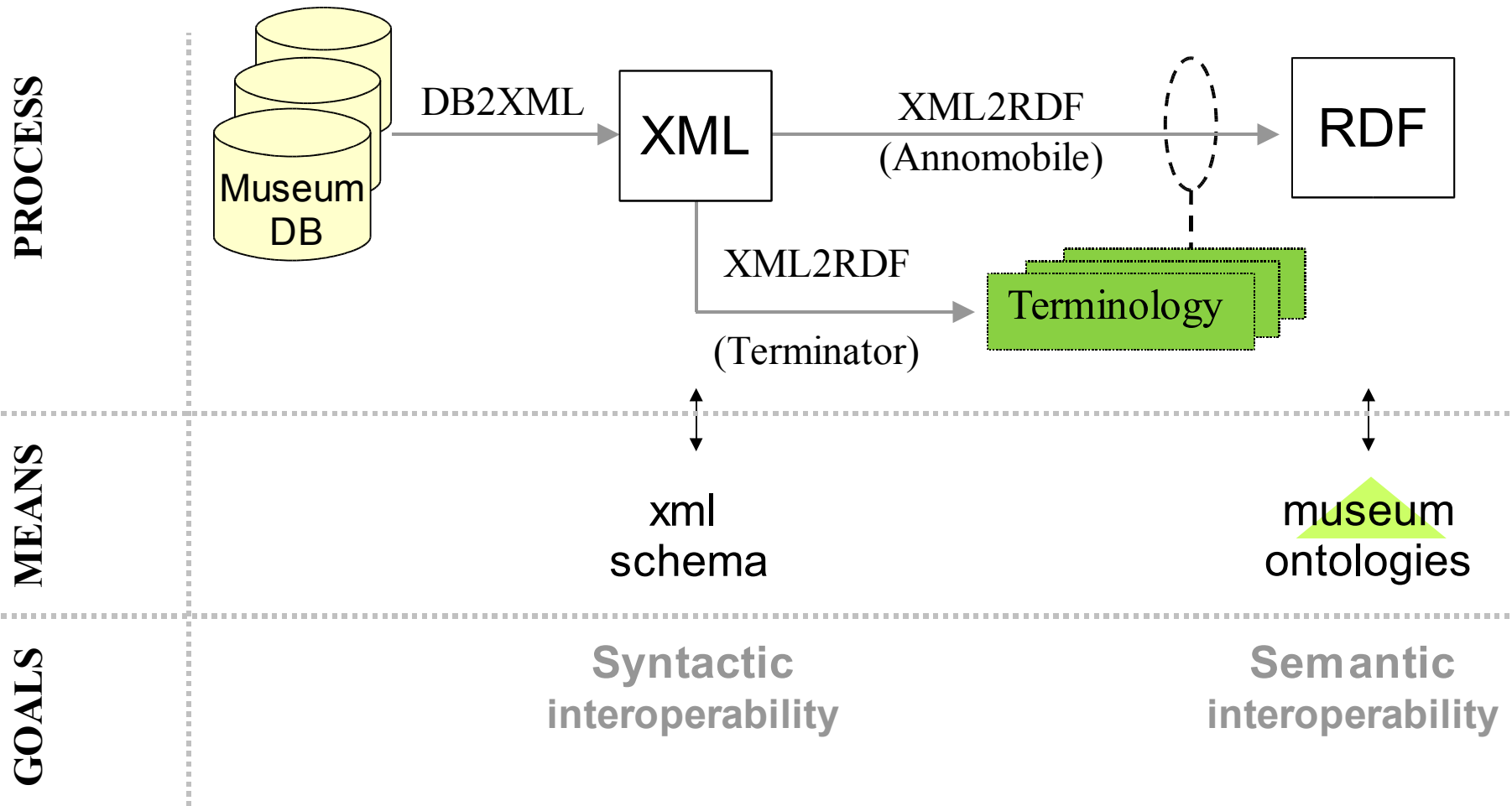
Specification for annotation process need

- Bringing new museum collections into MuseumFinland → by annotating the museum items.
- Requirements for the annotation process:
 4. Transforming museum item descriptions into MuseumFinland format should be
 - easy and
 - need only little manual work and technical expertise.
 5. No need to change old conventions of creating museum item descriptions.

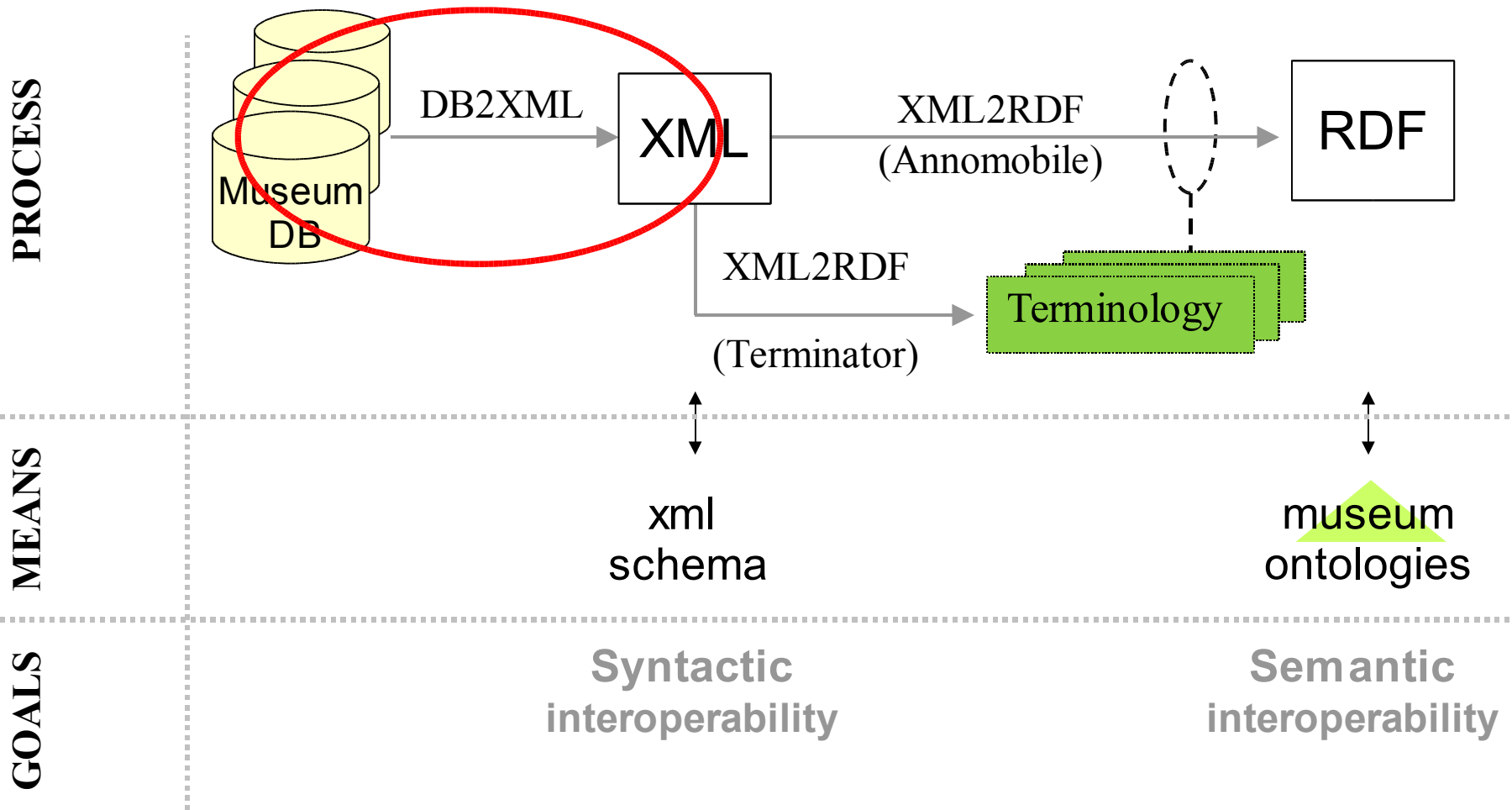
Outline

- Specification for content and annotation process need
- Three step annotation process
 - Syntactic homogenization
 - Terminology creation
 - Annotation creation
- Conclusion

Three step process

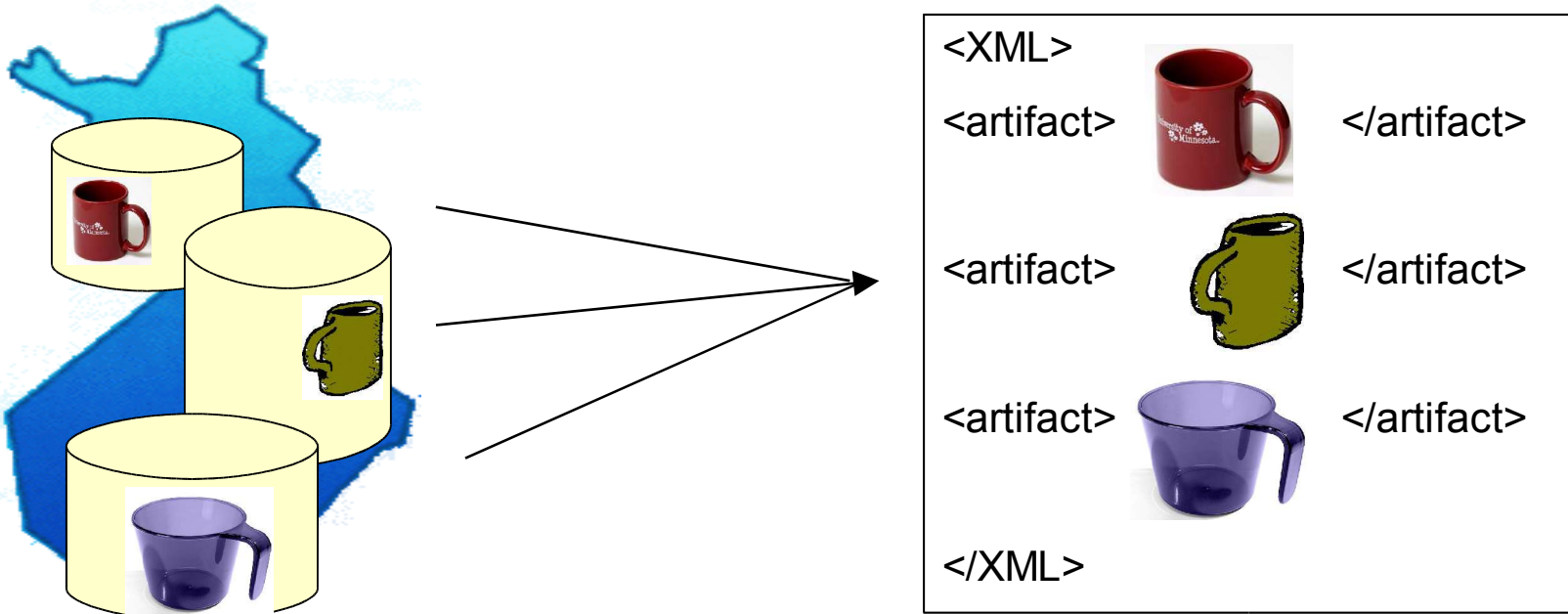


Three step process: step 1

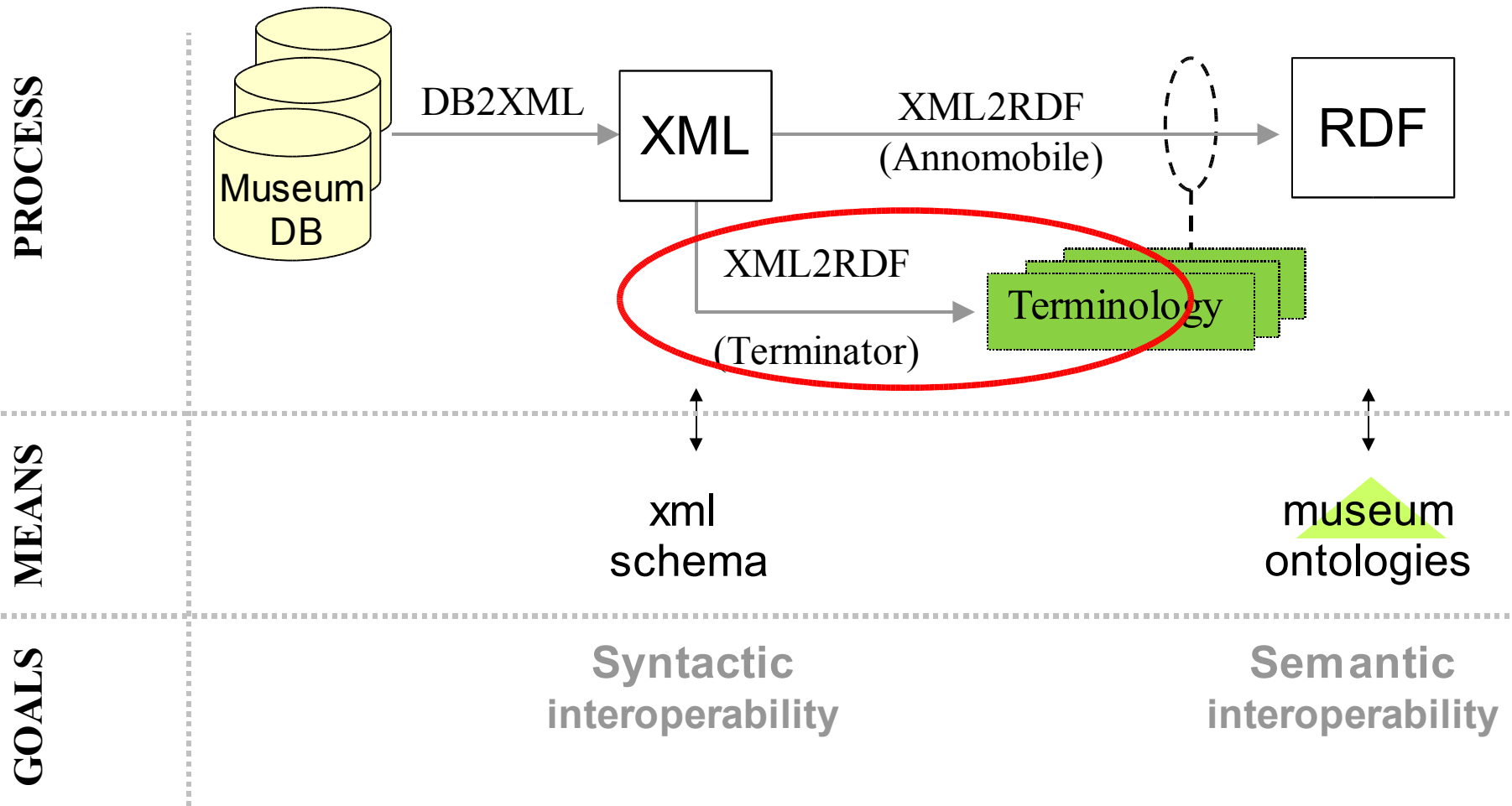


Reaching syntactic interoperability

- Syntactic heterogeneity
 - The data located in physically distributed databases
 - Databases are made by different manufacturers
 - Logical structure of DB , eg. schemas, tables, fields, differ
- The wanted information combined through shared presentation language, like XML

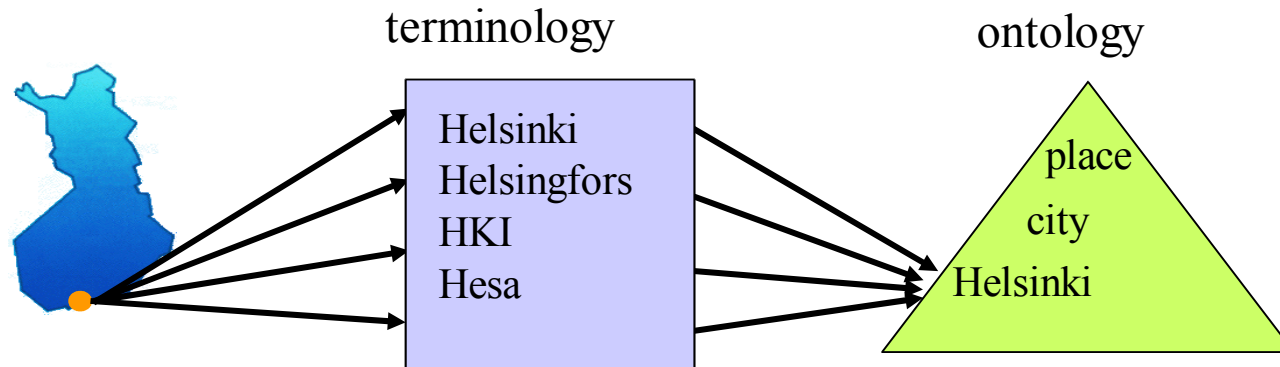


Three step process: step 2



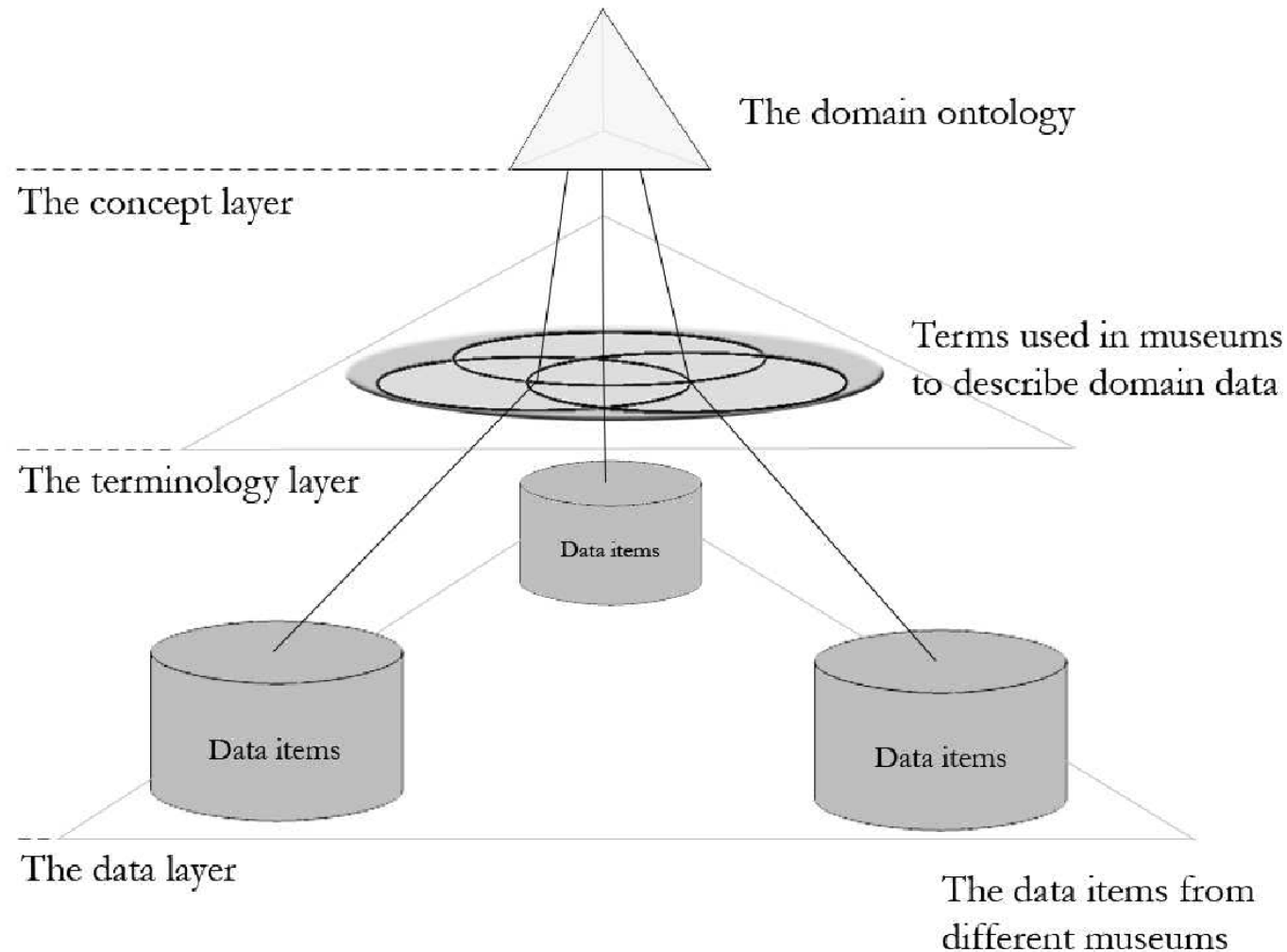
Terminology creation

- In MuseumFinland terms are associated with ontology concepts through term-ontology mappings.

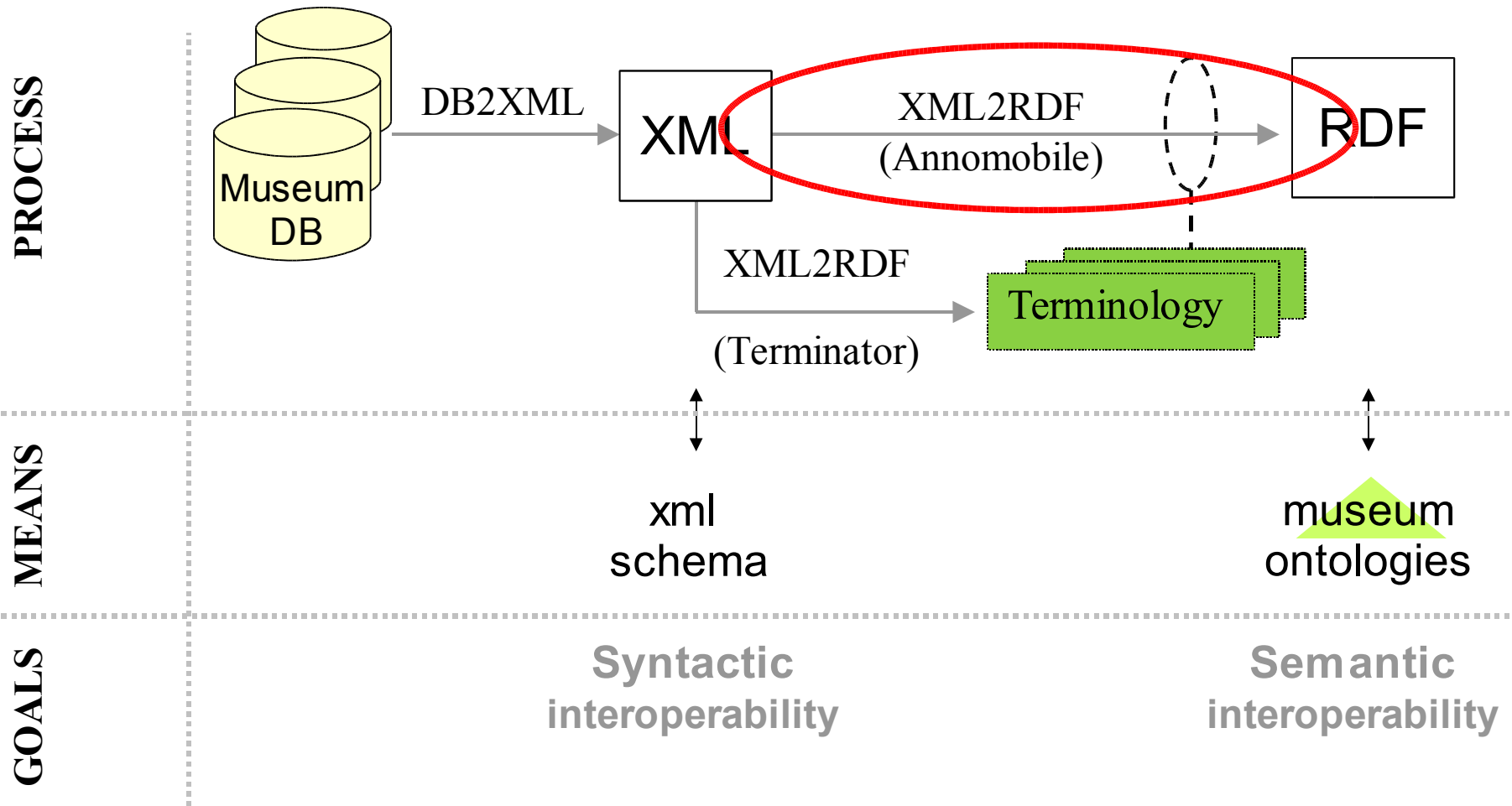


- Term-ontology mappings are created from
 - Ontology concepts
 - Museum item descriptions

Terminology creation: three layers of data

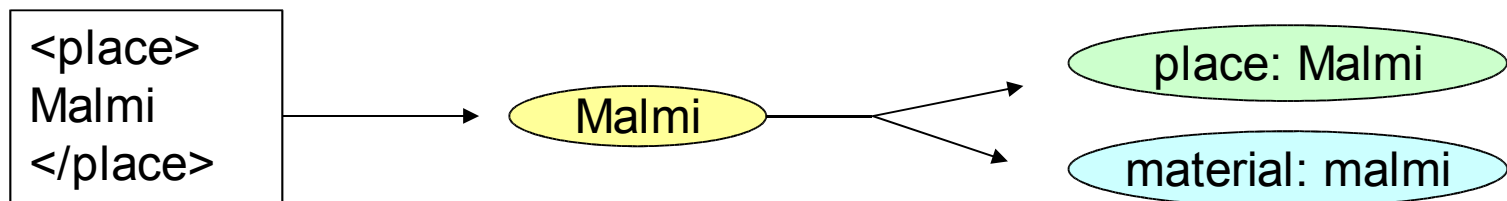


Three step process: step 3



Annotation creation

- Linking the artifact information to ontologies through term-ontology mappings
- Term-ontology mappings provide candidates for ontology links
- Properties are mapped to domain ontologies
 - disambiguation of homonymous terms referring to resources in different ontologies.
 - homonymous terms from the same domain need to be checked





XML - format

```
<artifact>
  <id>NBA:H26069:467</id>
  <target>cup and plate</target>
  <material>porcelain</material>
  <creationLocation>Germany
    </creationLocation>
  <creator>Meissen</creator>
</artifact>
```

RDF - format

```
<artifact:artifact
  rdf:about="&artifact;NBA_H26069_467"
  artifact:target_literal="cup and plate"
  artifact:material_literal="porcelain"
  artifact:creator_literal="Meissen"
  artifact:creationLocation_literal="Germany">
  <artifact:target rdf:resource="&artifacts;cups"/>
  <artifact:target rdf:resource="&artifacts;plates"/>
  <artifact:material
    rdf:resource="&materials;porcelain"/>
  <artifact:creationLocation
    rdf:resource="&locations;Germany"/>
  <artifact:creator rdf:resource="&actors;Meissen"/>
</artifact:artifact>
```

Outline

- Specification for content and annotation process need
- Three step annotation process
 - Syntactic homogenization
 - Terminology creation
 - Annotation creation
- Conclusion

Conclusion

- Pilot version opened on the web
 - Museums and public have been pleased with the portal
- Terminological interoperability reached
 - Ability to annotate heterogenous museum collections with shared ontologies
- Homonymy not a considerable problem
 - Disambiguation based on data fields and concept domains
- The current system cannot reach complete automation in annotation
 - museum cataloging systems should be enhanced with support for ontology-based indexing

Thank you!

Questions? Comments?

MuseoSuomi - Mozilla

File Edit View Go Bookmarks Tools Window Help

Back Forward Reload Stop

http://museosuomi.cs.helsinki.fi/?l=f&m=0&n=%2500%250B&g=c%2500%250B

Search Print

Home Bookmarks The Mozilla Organiza... Latest Builds

MuseoSuomi
- Suomen museot semanttisessa webissä -

HELSINKI INSTITUTE FOR INFORMATION TECHNOLOGY

UNIVERSITY OF HELSINKI

Uusi haku | Ohjeet | Näytä kaikki kategoriat | Tietoa ohjelmasta | MuseoSuomi-palautteet | In English: About Museum Finland

Sanahaku: Hae tarkenna hakua

Esinetyyppi: [kaikki](#) > [pukineet ja tekstiilit](#) (koko luokittelu)

[nahkatuotteet](#) (25), [pukineet](#) (1505), [tekstiilit](#) (338)

Materiaali (koko luokittelu) (ryhmittele kohteet)

[materiaalit](#) (1700)

Valmistaja (koko luokittelu) (ryhmittele kohteet)

[henkilöt](#) (406), [kaupungit](#) (1), [laitokset](#) (7), [tuotemerkit](#) (108), [yhdistykset](#) (2), [yhteisöt](#) (1), [yritykset](#) (695)

Valmistuspaikka (koko luokittelu) (ryhmittele kohteet)

[Aasia](#) (7), [Afrikka](#) (25), [Eurooppa](#) (1404), [Pohjois-Amerikka](#) (2)

Valmistusaika (koko luokittelu) (ryhmittele kohteet)

[aikakaudet](#) (1479), [vuosisadat](#) (1479)

Käyttäjä (koko luokittelu) (ryhmittele kohteet)

[henkilöt](#) (1204), [henkilöryhmät](#) (1), [laitokset](#) (15), [yritykset](#) (27)

Käyttöpaikka (koko luokittelu) (ryhmittele kohteet)

[Afrikka](#) (2), [Eurooppa](#) (1298), [Pohjois-Amerikka](#) (1)

Käyttötilanne (koko luokittelu) (ryhmittele kohteet)

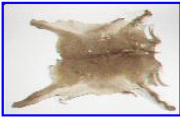
[kulttuuritapahtumat](#) (5), [kansalais-, harrastus- ja vapaa-ajantoiminta](#) (35), [institutionaalinen toiminta](#) (11), [juhlat ja seremoniat](#) (25), [kohteelle tehtävät toimenpiteet](#) (180), [kohteessa tapahtuvat muutokset](#) (9), [muut tapahtumat](#) (3), [maatalous ja karjanhoito](#) (1), [muut](#) (17)


Hakuehdot


Kategoria: Esinetyyppi > [pukineet ja tekstiilit](#) (ryhmittele kohteet) (poista)


Kohteet ryhmiteltyinä kategorian [pukineet ja tekstiilit](#) mukaisesti (näytä ilman ryhmittelyä)

[nahkatuotteet](#), kohteet 1-4/25 (ryhmittele kohteet)



harjalampaannahka, tisirit nwudad (NBA VK4835 96)



gasellinnahka (NBA VK4835 97)

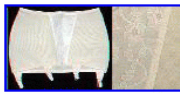

peski, turkki (NBA SU1826 1)

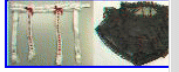

peski, porkipeska, purkka turkki (NBA SU4135 1) (edellinen)

[pukineet](#), kohteet 1-4/1505 (ryhmittele kohteet)






korkolappu, kantalappu (ECM 3594 298)


Kaularöyhelö, vauvan (L KM LHM LHM ES 72045 18)


Tukiliivit, naisen; lantioliivit, sukkanauhalliset (L KM LHM LHM ES 94084 287)


Sukkanauhaliivit ja alushousut: 1960-l. (L KM LHM LHM ES 98021) (edellinen)

[tekstiilit](#), kohteet 1-4/338 (ryhmittele kohteet)

Document: Done (0.735 secs)