MINT – METADATA INTEROPERABILITY SERVICES

Nikolaos Simou – National Technical University of Athens

What is MINT?

- Mint is a herb having hundreds of varieties that is mostly known as a mouth and breath fresher

Basic Concepts

- Metadata
  - Data about data
- Interoperability of metadata
  - Transformation of heterogeneous metadata into one inter-operable metadata standard
- Europeana
  - www.europeana.eu is an internet portal that acts as an interface to millions of books, paintings, films, museum objects and archival records that have been digitized throughout Europe

What is MINT?

- MINT is an open source, web based platform for Metadata INTeroperability
  - http://mint.image.ece.ntua.gr
  - It has been successfully used in more than 15 Europeana feeder projects
    - 300 cultural organizations
    - 500 users
  - More than 6.000.000 metadata records have been produced by it and published on Europeana

Scope

- Present MINT and its functionalities
- Familiarize the audience with a typical metadata aggregation workflow
- Highlight the importance of high-quality metadata

MINT – The early days

- MINT development started by the NTUA team during the Athena project back in 2008
- Athena's main objectives were
  - The delivery of 4.200.000 metadata records coming from 23 different museums and organisations to Europeana
  - The development of a metadata standard for the museum sector that would act as an intermediate for the project
  - The development of technical tools for assisting content providers to meet project objectives

MINT – Initial Workflow

MINT

Harvesting → XML → Mapping to Intermediate → XML → Aut. Mapping to ESE → XML → **Europeana**
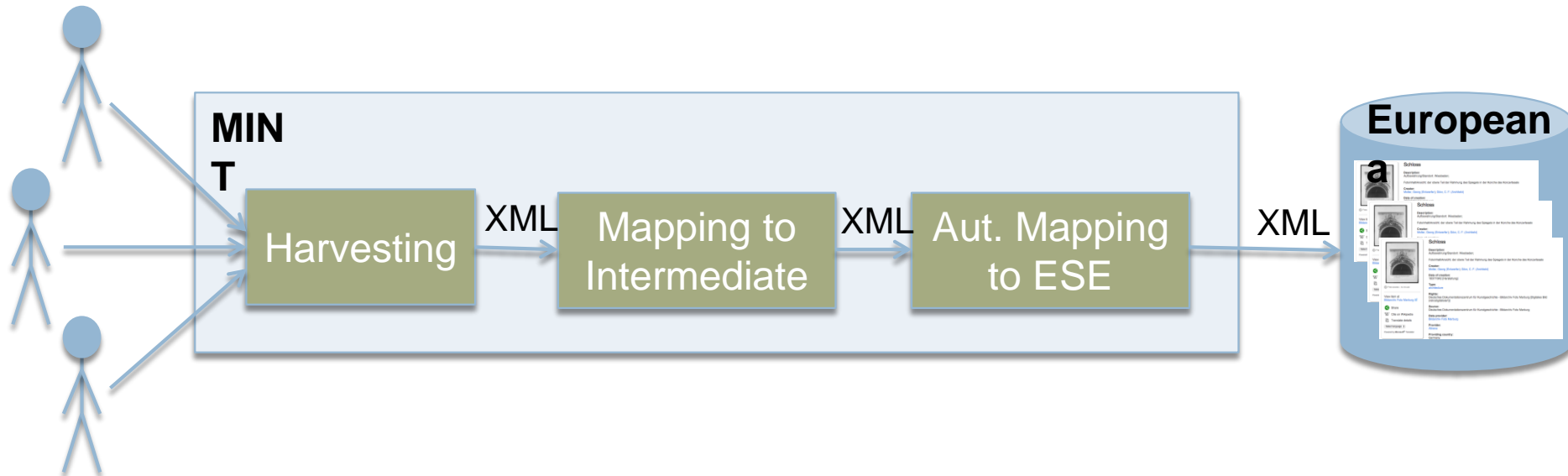
Harvesting requirements

- Different metadata models
  - Standards e.g. Dublin Core
  - In house metadata models
- Different File formats
  - XML
  - CSV
- Different delivery protocols
  - File upload (i.e. HTTP)
  - Open Archives Initiative - Protocol for Metadata Harvesting (OAI-PMH)

Mapping & Transformation req.

- □ Agnostic to metadata input
- □ Target schema based on a metadata model
  - ◻ XSD support
- □ Crosswalks between known schemas
  - ◻ Project's intermediate to ESE
- □ User friendly interface
  - ◻ Most of the content providers did not have any technical background, they were (in most cases) well aware of their metadata

Mapping Editor

Further requirements…

- Better control of metadata and its quality
  - Metadata cleaning services
    - Many of the providers have realised that they also hold a lot of rubbish they wanted to get rid before publishing to Europeana
  - Validation according to target schema
    - Mapping to a common target schema itself did not provide high quality metadata
  - Preview interfaces
    - Providers wanted to preview their metadata for controlling its quality before the actual publication

Mappings – Concatenate mapping

□ Allows users to combine more than one mappings for producing a new mapping

Mappings – Conditional Mapping

- Allows users to set conditions on their mappings
  - If the value of xpath tns:Organisation **is equal to** "IVML" and the value of xpath tns:PhotoURI starts with "http://www.image.ntua.gr" and either the value of xpath tns:PhotoURI contains "nsimou" of EuPhoto then the value of the xpath tns:Organisation will be mapped to the xpath lido:appellationValue of the target schema
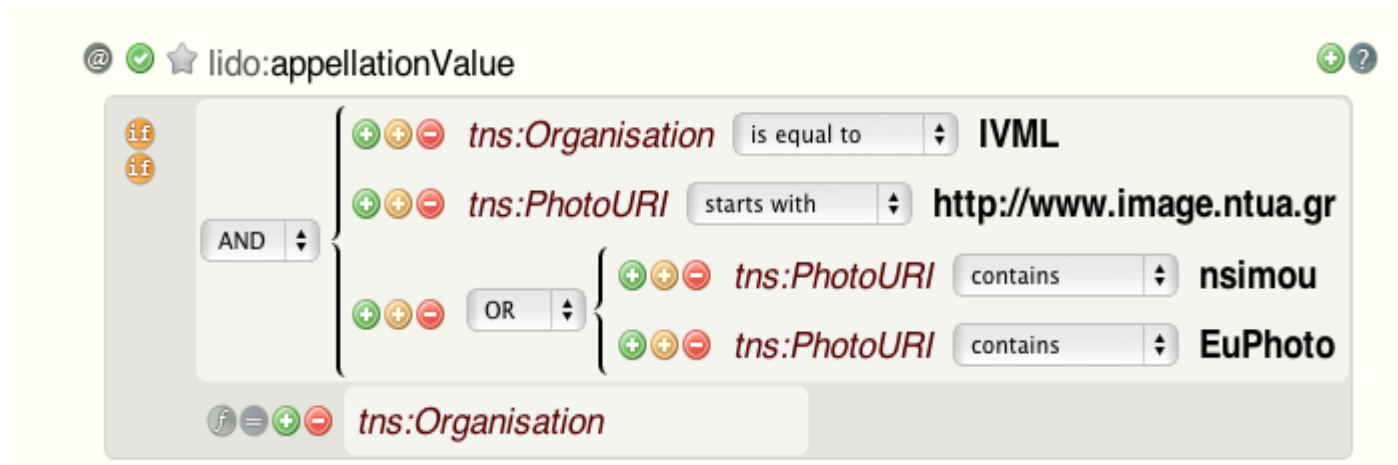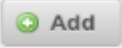
Mappings – Value Mapping

□ Allows users to normalize their metadata by mapping values of their input to specific values they set
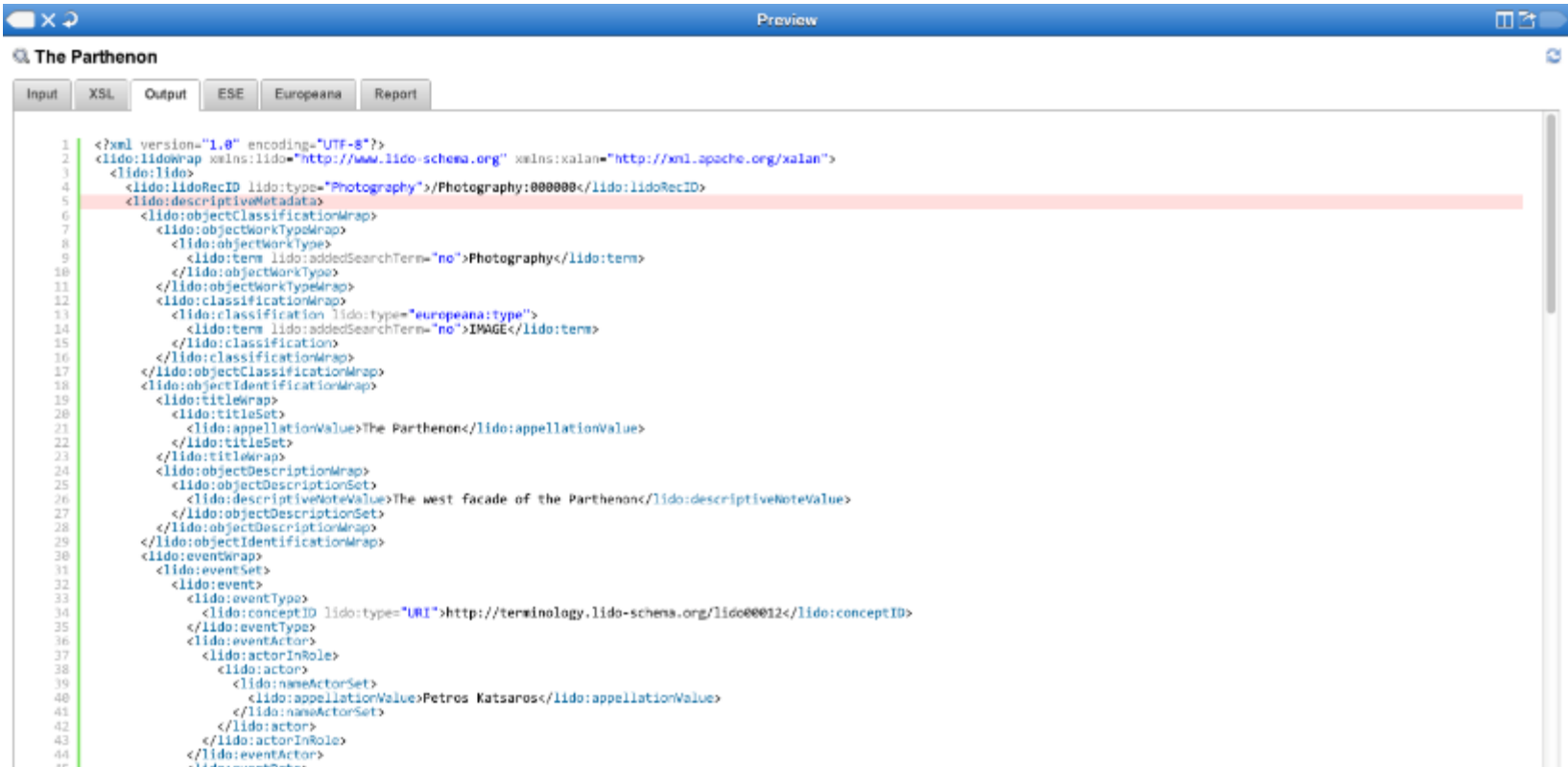
Mappings – Functional Mapping

- **Substring**: User sets the start and the end index.
- **Substring after**: User sets the substring of the original string after which the value is taken.
- **Substring before:** User sets the substring of the original string before which the value is taken.
- **Substring between:** User sets the substrings of the original string after and before which the value is taken.
- **Split**: User sets the delimiter for tokenization and the start index of the original string.
- **Tokenize content and generate an element per content**: User sets the delimiter for tokenization.

Preview & Validation

Europeana Preview

Metadata Quality

☐ [Link to record](#)

The shift from ESE to EDM

□ Metadata Quality

    ▫ Metadata reconciliation

        ■ Connecting provider's collection-specific vocabularies to controlled vocabularies on the Web

    ▫ Metadata Annotation

        ■ Some of the providers had metadata of limited expressivity for publishing on Europeana

Mappings – Thesaurus mapping

□ Target schema elements can take values from SKOS vocabularies
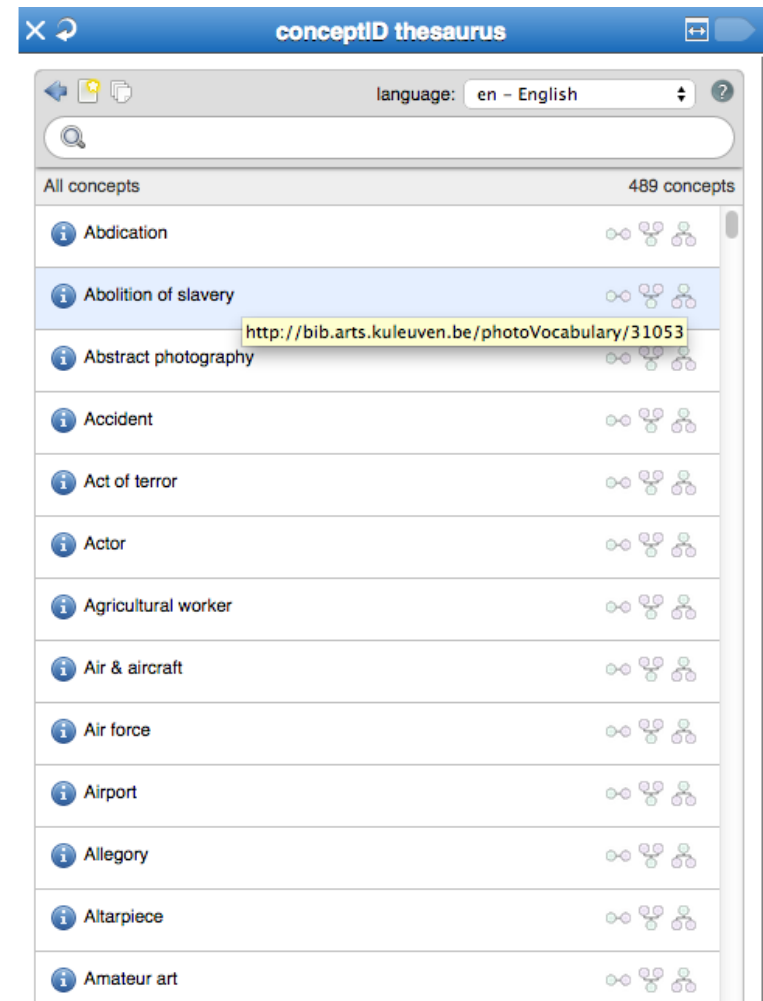
Mappings – Thesaurus mapping

- The user can
  - select a term from the thesaurus that will be applied to all items

Mappings – Thesaurus mapping

□ The user can
  ◻ align its in-house values to terms selected from the thesaurus

Metadata Annotation

- Annotator
  - Create, delete and modify records

Metadata Annotation

□ Group Edit

◻ Tag records using elements of the schema

Metadata Quality

□ [Link to record](#)

…and more requirements…

□ Publication

    ▫ Support of multiple publications

       ■ Europeana

       ■ Thematic portals

       ■ Linked Open Data

MINT – Current Workflow
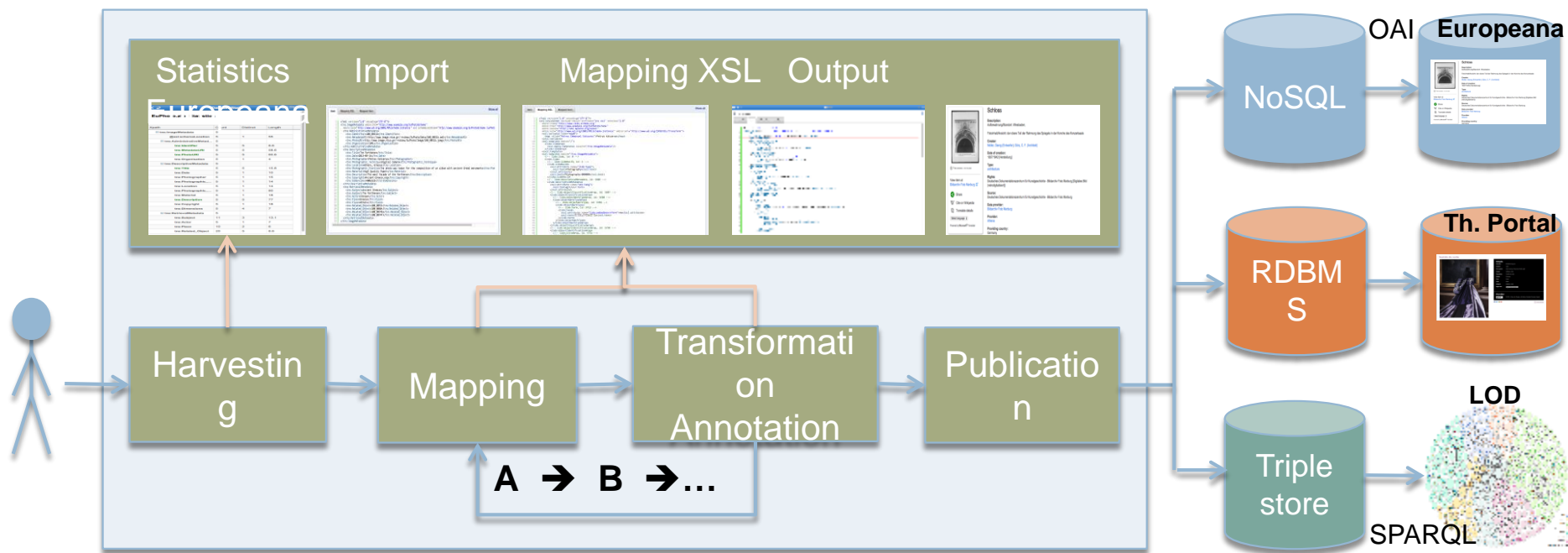
Lesson learnt

- Metadata
  - The production of high quality metadata will always depend on the user. A tool can only provide useful services
- Never-ending requirements
  - Be one of them, find out what they'll need before they do
  - Be flexible by modularizing everything as much as possible

Future developments

- Cleaning and Linking
  - Developing functionalities based on semantic technologies for cleaning and linking metadata to existing Linked Open Data sources
- APIs for creative applications
  - Simplify access to MINT manipulated metadata for building applications, websites and mash-ups

Useful links

- MINT User Documentation
  - http://mint-wordpress.image.ntua.gr/mint-end-user-documentation/
- MINT on github
  - https://github.com/mint-ntua
  - http://mint.image.ntua.gr/redmine/projects/mint/wiki/Mint_Installation_instructions
- MINT Instance for AthenaPlus Museum
  - http://mint-projects.image.ntua.gr/athenaplusmuseum
- Europeana Fashion Thematic Portal
  - http://www.europeanafashion.eu/

Contact

Nikolaos Simou

National Technical University of Athens

Email: nsimou@image.ntua.gr

LinkedIn: http://www.linkedin.com/pub/nikolaos-simou/68/31a/9aa