Intro to

LoCloud

Holly Wright
Archaeology Data Service
University of York, UK

International TAG CLOUD Project Workshop
Digital Heritage, 29 September, 2015
Granada, Spain

LoCloud Video
• Best practice network funded by the EU under the CIP ICT PSP programme
• 1\textsuperscript{st} March 2013 to 29\textsuperscript{th} February 2016
• 33 partners in 26 countries
  • Scientific coordinator: National Archives, Norway
  • Project management: 2Culture Associates, UK
  • Strong group of technical partners
  • National and regional aggregation services, and individual content providers

http://www.locloud.eu

International TAG CLOUD Project
Main goals

- Was developed based on the findings of EuropeanaLocal and CARARE
- Supporting small and medium-sized institutions in making their content available to Europeana
- Creating technical solutions suitable for smaller institutions
- Exploring the potential of cloud-computing technologies
- Creating a variety of metadata enrichment services
- Providing support, guidance and training
Innovating the aggregation infrastructure and exploring the potential of the cloud:

- **MINT**
  - Mapping tool with support for multiple metadata schemas
- **MORe aggregator framework**
  - Storage layer with decentralized and scaleable components
  - Pluggable services for metadata validation and enrichment
  - User-friendly interface for data providers
- **Crawler ready tagging tools**

International TAG CLOUD Project
MINT Mapping Tool

• Allows users to map native XML or CSV metadata to one of the Intermediate schemas supported by LoCloud: LIDO, CARARE, EDM.

• Enables content providers to prepare their metadata for publication.

• Integrated with MORE International TAG CLOUD Project.
• Aggregation framework
  – Based on Fedora Commons and Apache Cassandra
  – Storage layer with decentralized and scaleable components
  – Pluggable components
  – Developers area provides REST API and virtual enrichment drivers
• Cloud-based IaaS

International TAG CLOUD Project

http://store.locloud.eu
• User-friendly interface allows data providers to select data sources, manage ingestion, validation, enrichment and publication of their metadata
• Automatic creation of data sources in MINT and LoCloud Collections
Cloud-based micro-services for metadata enrichment and capture:

- Geo-location API
- Vocabulary service
- Historic Place Names service
- Geo-coding application
- Vocabulary matching service
- Background linking service
- Wikimedia application

International TAG CLOUD Project
• Six development teams: AIT, Athena RC, AVINET, IPCHS, UPV/EHU, VUKFC
• All the services are implemented on virtual machines in a cloud testlab (using the OpenNebula cloud computing platform)
• API documentation is available for each service
• End-user interfaces are provided for three microservices
• Each service has been integrated in MORE
LoCloud

LoGeo API

- Uses NER-Name Entity Recognition
- Recognizes and returns one or more place names candidates accompanied by geographical coordinates.

International TAG CLOUD Project
• End-user oriented application
• allows institutions to add geographical locations to existing content in a controlled crowd-sourcing environment and then
• load the augmented data back into their original production databases.
LoCloud Historic Place Names Service

- A **semi-automatic** historical geo-information management function and web service.
- Enables local cultural institutions to **collaborate in developing** the Thesaurus of HPN and enriching ingested metadata.
- Enables **export** of selected LoCloud HPN Thesaurus data sets.

International TAG CLOUD Project
Vocabulary services

- Based in TemaTres
- Incorporates 29 standard thesauri
- Allows users to add and create vocabularies, and add terms or translations
- Add thesauri terms to all items of a metadata package
- Integrated in MORE
- Exploited by the LoCloud vocabulary matching microservice via its API

International TAG CLOUD Project
Vocabulary matching

- Builds on work in the PATHS project (FP7)
- Automatically assigns relevant concepts and terms to item records
- Retrieves terms from the Vocabulary service

International TAG CLOUD Project
• Based on work in PATHS
• Follows the Named Entity Disambiguation approach
• Enriches items with links to external contextual information in DBpedia and Wikipedia
A web service that uses a REST interface to communicate with Wikimedia commons.

The main functionalities are harvesting content from Wikimedia, parsing the harvested content, and identifying useful entities that can be mapped to the ESE or EDM metadata schemas.

The mapped ESE / EDM records are then sent to MORE to be delivered to Europeana.
<table>
<thead>
<tr>
<th>Service</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Language</td>
<td>EDM Language identification</td>
</tr>
<tr>
<td>GeNoa</td>
<td>EDM Geo Normalization</td>
</tr>
<tr>
<td>LoGeo</td>
<td>EDM GeoLocation Enrichment</td>
</tr>
<tr>
<td>CLoSer</td>
<td>EDM Geonames Geocoding</td>
</tr>
<tr>
<td>GeNeSi</td>
<td>EDM Geonames Rev. Geocoding</td>
</tr>
<tr>
<td>Vocabulary</td>
<td>EDM Vocabulary</td>
</tr>
<tr>
<td>Vocab. match</td>
<td>EDM Vocabulary Matching</td>
</tr>
<tr>
<td>Back Link</td>
<td>EDM Background Link</td>
</tr>
</tbody>
</table>
LoCloud Collections

https://locloudhosting.net/

- Developed by PSNC
- Based on the open source Omeka system
- Allows users to publish digital objects on-line in a few minutes
- Developed and tested in 2014, now available in production
- Scalable, user friendly, with OAI-PMH interface
- Storage of metadata and content
- Incorporates web traffic analytics
- Cloud-based SaaS

International TAG CLOUD Project
**Featured Item**

Święta Maria Magdalena cząstująca

**Recently Added Items**

- Portret Adama Mickiewicza
  - Koniec XIX w.
  - Styka Jan

- Swieta Maria Magdalena
  - Koniec XIX w.
  - Brugnol Bartholomaeus de

- Seksualny mąż
  - autor nieznany

- Portret miłości
  - autor nieznany

- Sekcja ks. Brunata w. Spira
  - autor nieznany

**Featured Collection**

Witajcie w Galerii im. Władysława

Instytucja Cultura Nauki Poznańskiego Towarzystwa Przyjaciół Nauk była prowadzona od 2 połowy XIX wieku. Pod koniec lat siedemnastych XIX w.
LoCloud support portal incorporates:
- user and technical documentation
- eLearning materials
- Question-and-Answer service
- Access to a ticket desk system

http://support.locloud.eu
LoCloud has developed services and aggregation infrastructure in the cloud

These new services are being used to publish content online from smaller data providers and to supply it to Europeana

Support, guidance and training has been put in place

Current focus is on operational and strategic planning
Thank you!
holly.wright@york.ac.uk

Please join us for
Sharing Cultural Heritage Online with LoCloud
Thursday, 1 October 10.00-13.30