



The Hellenic Aggregator

Ioannis Troxopoulos

Veria Central Public Library

WHERE WE ARE



ATLA 2010

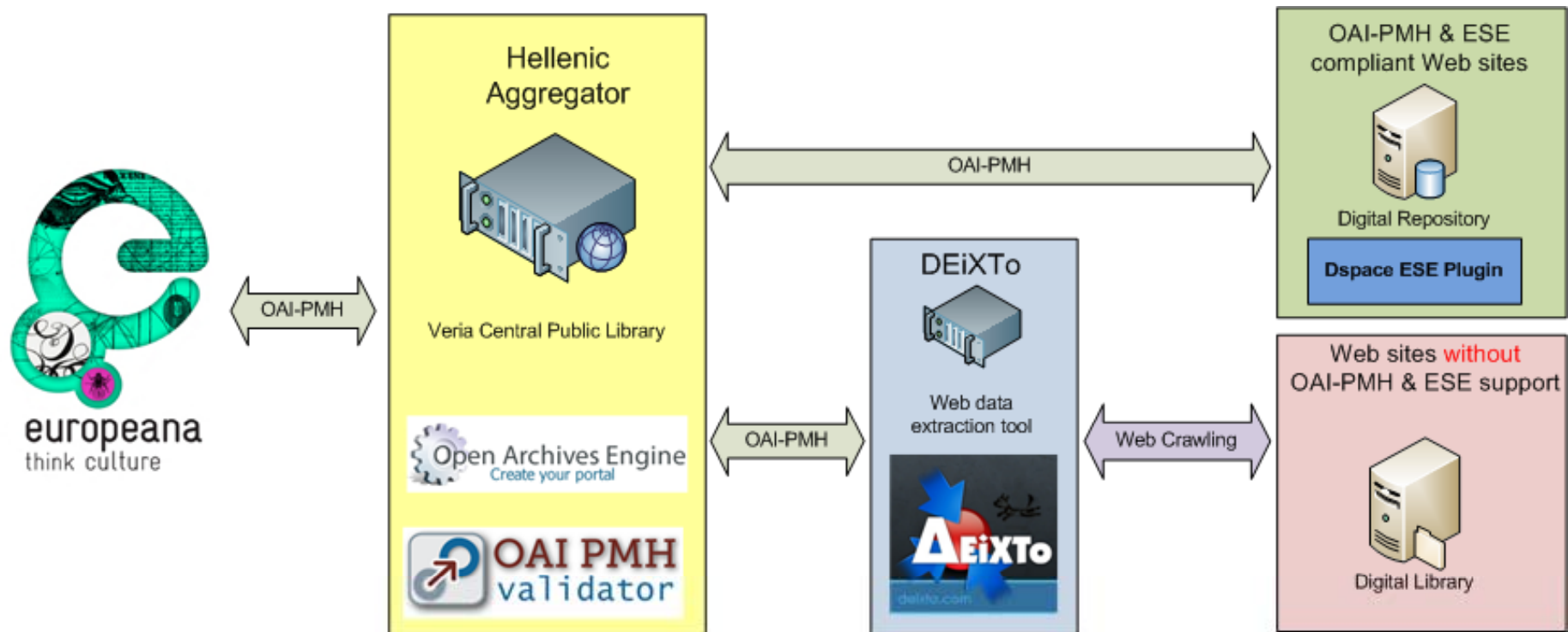
*“...a small library
creates big impact”*



The Hellenic Aggregator

- The contributing parties to this project are currently:
 1. Pandektis - National Documentation Center of Greece
 2. Medusa - Veria Central Public Library
 3. The Historical Archives of the American Farm School of Thessaloniki
 4. Technical Chamber of Greece Regional Department of Corfu
 5. Central Library of NTUA
 6. Music Library - Lilian Voudouri
 7. Corgialenios Digital Library
 8. University of Athens - Pergamos
 9. Hellenic Ministry of Education - Educational Television
- The total number of records available at 1/2/2011 is **128.439**
- Visit: <http://aggregator.libver.gr>

The Hellenic Aggregator Architecture



DeiXTo – Web Data Extraction Tool

- **DEiXTo** (or **ΔEiXTo**) is a powerful web data extraction tool that is based on the W3C Document Object Model (DOM).



- It allows users to create highly accurate "extraction rules" (wrappers) that describe what pieces of data to scrape from a website.
- DEiXTo can contend with a wide range of digital libraries with high precision and recall.
- **DEiXTo** is able to transform the contents of **your** digital library into **OAI-PMH ESE** or another suitable format.
- Visit: <http://www.deixto.com>

DeiXTO: Corgialenios Library use case

From HTML unstructured data

ΤΙΤΛΟΣ ΑΝΤΙΚΕΙΜΕΝΟΥ			
[Καρποκράτης]			
ΓΕΝΙΚΗ ΚΑΡΤΕΛΑ			
ΣΕΛΙΔΕΣ	ΣΥΓΓΡΑΦΕΑΣ	ΗΜ/ΝΙΑ ΕΚΔΟΣΗΣ	ΣΤΟΙΧΕΙΑ ΚΥΡΙΟΤΗΤΑΣ
8	Άννινος, Χαραλάμπης	N/A	ΠΛΗΡΗΣ ΚΥΡΙΟΤΗΤΑ
ΑΝΑΛΥΤΙΚΗ ΚΑΡΤΕΛΑ			
ΚΩΔΙΚΟΣ	ANN10.22		
ΗΜ/ΝΙΑ ΑΠΟΚΤΗΣΗΣ	1974		
ΓΛΩΣΣΑ	GRE		
ΧΩΡΑ	GR		
ΤΙΤΛΟΣ ΚΑΙ ΔΗΛΩΣΗ ΠΝΕΥΜΑΤΙΚΗΣ ΥΠΕΥΘΥΝΟΤΗΤΑΣ	[Καρποκράτης]		
ΦΥΣΙΚΗ ΠΕΡΙΓΡΑΦΗ	7 φ.		
ΘΕΜΑΤΑ (ΛΕΞΕΙΣ ΚΛΕΙΔΙΑ)	Χειρόγραφα Κεφαλονιά Λογοτέχνες Αρχεία		
ΟΝΟΜΑ ΠΡΟΣΩΠΟΥ (ΚΥΡΙΑ ΠΝΕΥΜΑΤΙΚΑ ΥΠΕΥΘΥΝΟΣ)	Άννινος, Χαραλάμπης		
ΠΡΟΕΛΕΥΣΗ	ΔΩΡΕΑ ΑΝΝΕΤΑΣ ΓΕΩΡΓΙΑΔΟΥ, ΚΛΕΙΟΥΣ, ΜΑΡΙΑΝΝΑΣ Χ. ΑΝΝΙΝΟΥ		



To ESE format!

```
-<europeana:record>
- <dc:identifier>
  http://www.corgialenios.gr/library/media.asp?aid=240
</dc:identifier>
<dc:identifier>ANN10.22</dc:identifier>
<dc:title>[Καρποκράτης]</dc:title>
<dc:creator>Άννινος, Χαραλάμπης</dc:creator>
<dc:creator>Άννινος, Χαραλάμπης</dc:creator>
<dc:date>N/A</dc:date>
<dc:format>8</dc:format>
<dc:rights>[Καρποκράτης]</dc:rights>
<dc:coverage>GR</dc:coverage>
<dc:language>GRE</dc:language>
- <dc:source>
  ΔΩΡΕΑ ΑΝΝΕΤΑΣ ΓΕΩΡΓΙΑΔΟΥ, ΚΛΕΙΟΥΣ, ΜΑΡΙΑΝΝΑΣ Χ. ΑΝΝΙΝΟΥ
</dc:source>
<dc:publisher>Χειρόγραφα Κεφαλονιά Λογοτέχνες Αρχεία</dc:publisher>
- <europeana:object>
  http://www.corgialenios.gr/files/low/2/ann10.22_1_22_000.jpg
</europeana:object>
<europeana:provider>Corgialenios Digital Library</europeana:provider>
<europeana:type>TEXT</europeana:type>
- <europeana:isShownAt>
  http://www.corgialenios.gr/library/media.asp?aid=240
</europeana:isShownAt>
</europeana:record>
```

Open Archives Engine

- Open Archives Engine (OAE) is a server-side software capable of creating a digital library metadata aggregator and search portal using OAI-PMH enabled, web accessible digital repositories 
- OAE utilizes the OAI-PMH protocol in order to retrieve metadata from multiple digital libraries and create an index which then can be used not only to search and filter information but also to export information in a variety of formats such as OAI-PMH Dublin Core (DC) and Europeana Semantic Elements (ESE).
- OAE leverages the technology of DeiXTO. The combination of OAE and DeiXTO can extract metadata from any web accessible digital library regardless of the technology, it uses and make it available through the web interface or the web service.
- Visit: <http://openarchivesengine.com>

OAIPMH.com – OAI-PMH validation & data extraction tool

- OAIPMH.com is a **free web application** capable of performing all the necessary checks required to ensure that an OAI-PMH enabled digital library is ready on a technical level to be part of Europeana. The logo for the OAIPMH validator. It features a square icon on the left with a blue circular arrow and a red arrow pointing upwards and to the right. To the right of the icon, the text "OAI PMH" is written in a dark red, serif font, and "validator" is written below it in a blue, sans-serif font.
- **Validation:** The validation of an OAI-PMH enabled digital library requires only the submission of the OAI-PMH web service URL. After this process is completed, the user is presented with a checklist of validation checks which have been performed and their results **in real time**.
- **Metadata extraction:** Users can provide the system with a list of OAI-PMH URLs and retrieve all the metadata records which are available from them in parallel. Using this feature, **users can retrieve a large number of metadata records from multiple libraries** rapidly and easily, thus enabling them to inspect them and evaluate them.
- Visit: <http://www.oaipmh.com>

Dspace ESE Plugins

- We have developed 2 Dspace plugins
- DSpace plugin for ESE
 - Add ESE support to an existing Dspace installation
 - <http://el.vbanos.gr/blog/2010/02/02/dspace-plugin-for-europeana-semantic-elements-ese/>
- Automated ESE schema & fields addition plugin
 - Batch insert ESE elements to an existing Dspace installation
 - <http://helios-eie.ekt.gr/EIE/handle/10442/8592>

Thank you

- Questions?