Open Access
knowledge for all

Technologies & Tools

for the Implementation of

Repositories and E-journals

Dr Nikos Houssos

Head of Software Development Unit, EKT / NHRF

www.ekt.gr

International Conference:
"Open Access Infrastructures: The Future of Scientific Communication"

15-16 December 2008, Athens-Greece (National Hellenic Research Foundation)

EKT

ΕΘΝΙΚΟ ΚΕΝΤΡΟ
ΤΕΚΜΗΡΙΩΣΗΣ
NATIONAL
DOCUMENTATION
CENTRE

Agenda

- Implementation options and challenges repositories and e-journals
- What is already available
- What needs to be improved

Repositories and E-Journals Implementation options

- Host in-house
  - Develop on open-source platform
  - Develop on commercial platform
  - Develop from scratch
- Outsource hosting
- Platform selection

Open source platforms

- Flexibility, unlimited customisation capabilities
- Tested solution - support already thousands of open access systems installations worldwide
- Suitable for a wide range of implementations
  - From the smallest to the largest
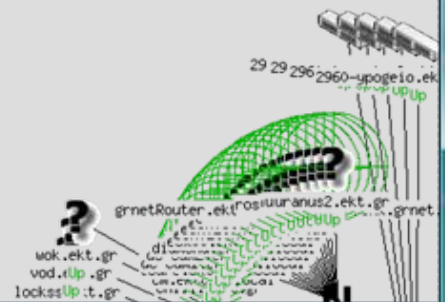 - Not only for high-end organisations at the edge of technology

Systems infrastructure requirements

- Reliability / availability / safety
- Rapid response to changing requirements
  - On-demand, "late" resource allocations
  - Staging environments - virtualisation
  - Particularly important for development on open source
- Logging / reporting

Systems infrastructure issues

- Virtualisation management
- Storage management
- Monitoring availability and performance
- Logging and statistics

Open
Access
knowledge for all

Software infrastructure issues

- Robust open source platforms are already in place for repositories and e-journals
- Basic functionality is well supported
- Some important advanced features are still not widely available

Mature functions

- Indexing of metadata and full-text
- Metadata search and browse
- Support for fundamental workflows (e.g., content submission, peer review)
- Protocols for remote metadata harvesting and searching
- Basic multi-lingual material support

Topics for further consideration (1 / 2)

- Advanced mechanisms for content submission
- Enhancing the end-user experience - web 2.0
- Sophisticated metadata representation standards (e.g., for compound objects)
- Semantic search / retrieval
- Applications of data matching

Topics for further consideration (2 / 2)

- Usability of full-text search (e.g., highlighting)
- Protocols and APIs for remote updates and harvesting full-text
- Distributed workflows
- Single-sign on
- Interoperability of repositories, e-journals and CRIS

How to attract researchers?

- Improve the user experience of researchers in the role of data providers
  - Help them with loading data to repositories
- Provide value-added services over repositories and e-journals
  - Reference management
  - Reporting
  - Usage statistics of researcher output

Researchers as data providers

- Combination of multiple mechanisms to make things easier for the researcher
  - Loading and pre-processing of researcher data from multiple bibliographic sources
  - Integration with workflows for production of scientific documents - "zero-click ingest"
  - Automatic metadata extraction
- Several production systems in use worldwide
- No open source solution yet

The end user experience

- Web 2.0 – like features
- Faceted browsing
- Maps, time lines, time plots
- Social networking features
  - Tagging, reviews
- Personalisation
- Make the above cross-repository?

Αρχείο Επεξεργασία Προ
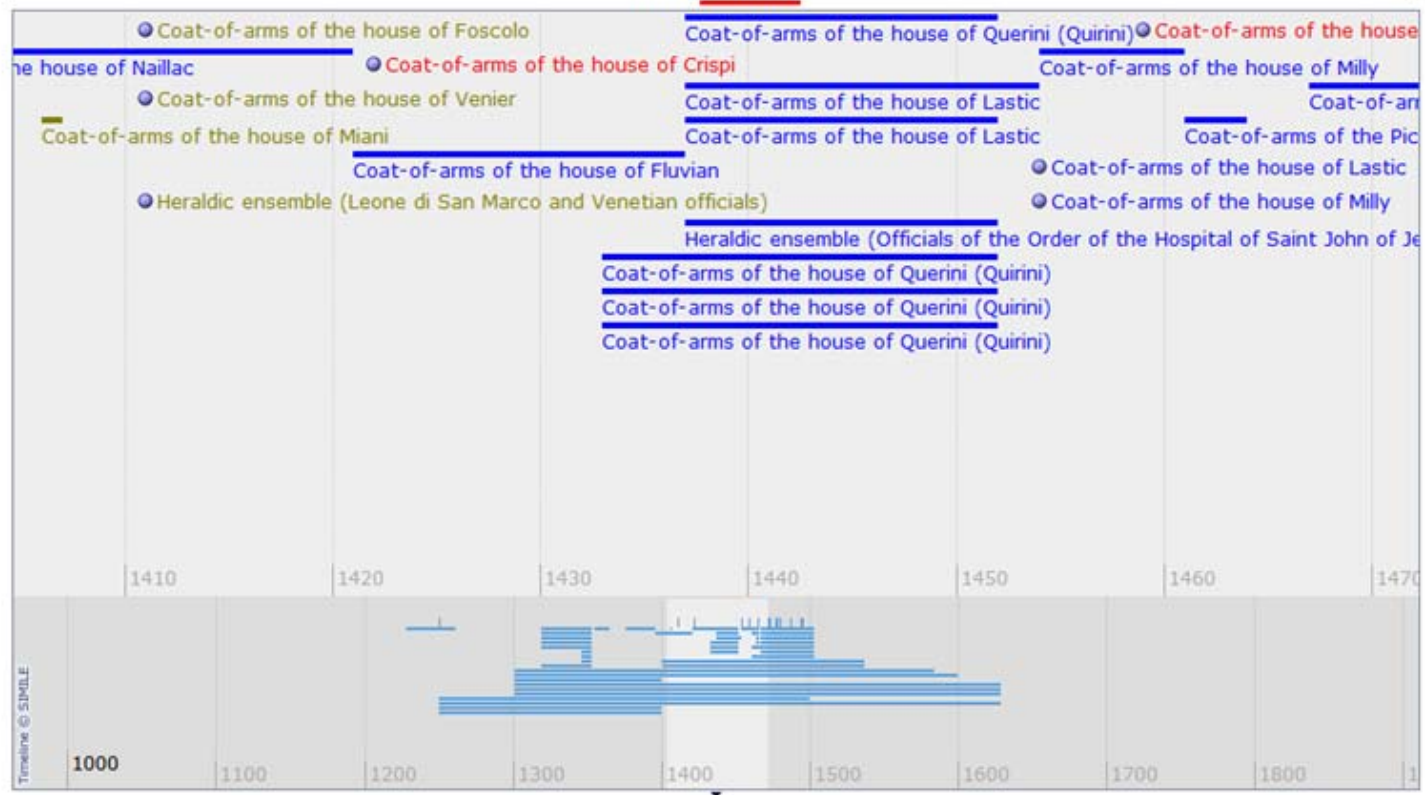
Αρχείο Επεξεργασία Προβολή Ιστορικό Delicious Σελιδοδείκτες Εργαλεία Βοήθεια

http://pandektis.ekt.gr/dspace/handle/123456789/20221

Google

Heraldic data ov

Heraldic data overview (faceted browsing)

MAP • **TIMELINE**

- Coat-of-arms of the house of Foscolo
- Coat-of-arms of the house of Querini (Quirini) • Coat-of-arms of the house
he house of Naillac
- Coat-of-arms of the house of Crispi
- Coat-of-arms of the house of Milly
- Coat-of-arms of the house of Venier
- Coat-of-arms of the house of Lastic
- Coat-of-arms of the house of Miani
- Coat-of-arms of the house of Lastic
- Coat-of-arms of the Pic
- Coat-of-arms of the house of Fluvian
- • Coat-of-arms of the house of Lastic
- • Heraldic ensemble (Leone di San Marco and Venetian officials)
- • Coat-of-arms of the house of Milly
- Heraldic ensemble (Officials of the Order of the Hospital of Saint John of Je
- Coat-of-arms of the house of Querini (Quirini)
- Coat-of-arms of the house of Querini (Quirini)
- Coat-of-arms of the house of Querini (Quirini)

| 1410 | 1420 | 1430 | 1440 | 1450 | 1460 | 1470 |

Timeline © SIMILE

| 1000 | 1100 | 1200 | 1300 | 1400 | 1500 | 1600 | 1700 | 1800 | 1 |

🟩 **Crete**   🟦 **Dodecanese islands**   🟥 **Kyklades islands**   🟨 **Peloponnese**

Region	4☑
22 **Crete**	☑
69 **Dodecanese islands**	☑
6 **Kyklades islands**	☑
9 **Northern Aegean islands**	☐
8 **Peloponnese**	☑

Settlement

2	**Geraki (Peloponnisos)**
4	**Kalymnos**
4	**Kamariotis (Malevizi-Iraklio, island of Crete)**
4	**Kastri of Agios Mamantas (Mylopotamos-Rethymno, island of Crete)**
4	**Kefali (Kissamos-Hania, island of Crete)**
1	**Kefalos (island of Paros)**
29	**Kos (city)**

Historical Period	1☑
105 **13th-15th century**	☑
111 **16th-17th century**	☐
42 **18th-19th century**	☐
2 **20th century**	☐

Region

27	**Attiki**
20	**Ionian islan**
42	**Kyklades is**
2	**Peloponnes**

POWERED BY Google

100 μίλ
200 χλ

Ολοκληρώθηκε

Ολοκληρώθηκε

Metadata representation

- Current systems support mostly flat metadata standards

- Limited support for:
  - Heterogeneity in metadata standards – model-driven development
  - Compound objects
  - Linking

The significance of data matching

- Authority files / unique identifiers for names (authors, research organisations, journals, publishers, research areas)
- Benefits:
  - Enhanced search / browse functionality
  - Reliability in statistics and reporting
  - Easier data integration and consolidation

MADS Collection Editor

Νέο Αρχείο   Άνοιγμα   Αποθήκευση   Έξοδος

Λίστα Συγγραφέων

Agriantoni, Christina
Alexis, Michael N.
Anagnostakis, Ilias
Andrianou, Dimitra
Andriotis, A. N.
Angelidi, Christine
Apostolopoulos, Dimitris
Argyropoulos, Roxane
Asdrachas, Spuros
Aslanis, Ioannis
Aspromallis, G.
Athanasiou, K.
Bacalis, Naoum C.
Balta, Evangelia
Bonas, G.
Calogeropoulou, Theodora
Cazianis, C.T.
Cefalas, Alciviadis Constantinos
Chatziioannou, Aristotelis

Αλλαγή Στοιχείων

Ονόματα Προβολής

Γλώσσα	Όνομα
en	Chryssikos, Georgios D.
gr	Χρυσικός, Γεώργιος Δ.

Εναλλακτικά Ονόματα

Γλώσσα	Όνομα
en	Chrysikos, G.D.
en	Chryssikos, D. G.
en	Chryssikos, G.
en	Chryssikos, G. D.
en	Chryssikos, G.D.
gr	Χρυσικός, Γ.
gr	Χρυσικός, Γ. Δ.
gr	Χρυσικός, Γ.Δ.

Υπό Σκέψη

Adamopoulos, G.
Aggeli, A.
Agiostratidou, G.
Athanasiou, C.
Benten, W. P. M.
Berberian, J.
Beretta, Marco
Bergman, T.
Bernardi, P.
Bernet, B.
Besnardramage, M. J.
Besra, G. S.
Bethell, D.
Bex, V.
Bezrukov, V.
Bhat, A. G.
Bia, M.
Bianchini, F.
Bibby, M. C.
Biberacher, W.
Bichard, C. J. F.

Example: data matching

data matching list washing object identity resolution entity disambiguation coreference resolution reference reconciliation duplicate detection merge/purge processing data deduplication instance identification record matching database hardening name matching identity uncertainty entity resolution entity matching record linkage

Interoperability standards

- Interoperability guidelines for repositories (DRIVER II, DINI, ...)

- Harvesting standard (OAI-PMH)

- Search standards (Z39.50, SRW/U)

- Structure (METS, OAI-ORE, MPEG-DIDL)

- Batch remote submission (SWORD)

- Need for standard update APIs

Conclusions

- Open access installations have demanding requirements on the systems infrastructure

- Robust open source software platforms are available for both repositories and e-journals

- Fundamental functionality is well supported

- Further work needed for several important advanced functions / features

Back up slides

Technical enablers

- Systems
  - Virtualisation
  - Single sign-on

- Software
  - Web services / Service-oriented architecture
  - Model-driven development
  - User interface technologies

Multi-lingual information management

- Multi-lingual information management
  - Encoding
  - User interface localisation
  - Indexing, searching and browsing