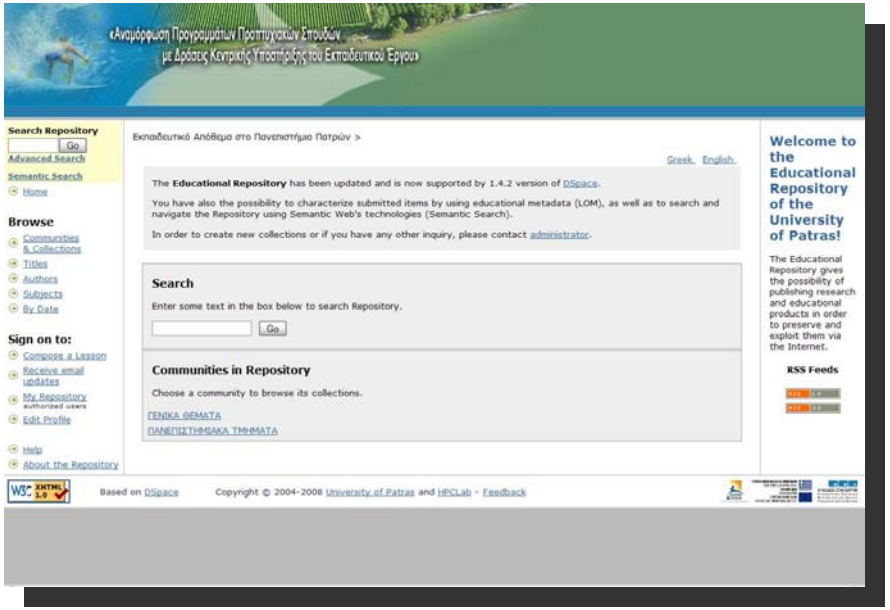
Digital Repositories and the Semantic Web: Semantic Search and Navigation for DSpace

Dimitrios Koutsomitropoulos, Georgia Solomou,
Andreas Alexopoulos and Theodore Papatheodorou

High Performance Information Systems Lab
Computer Engineering & Informatics Dpt.
School of Engineering, University of Patras
http://www.hpclab.ceid.upatras.gr

University of Patras Institutional Repository

A mechanism for the efficient **description**, **preservation, management**, **exploitation** and **distribution** of the University's educational and scientific material



http://repository.upatras.gr/dspace

- Built upon the open-source **DSpace** digital repository system
- Item description using the **Dublin Core** metadata schema

Articles, Books, Theses, Journal Papers, Images, Videos, Learning Objects, Data Sets, …

University of Patras Institutional Repository

Additional features

- Multilingual support
  - User Interface (Greek, English, …)
  - Metadata - Characterization of items in more than one language
- Advanced search service
  - Full text
  - Metadata
  - **Semantic Search**
- Advanced browsing
  - **Semantic navigation**

The Semantic Web

- New potential for the Web
  - Rich descriptions of resources + co-relations
  - Ability to reason about information
  - Knowledge acquisition and discovery (Inference based)
- Ontology Languages: OWL
  - Extensions to RDF(S)
  - Standard vocabulary for ontology representation
  - Decidable, sound & complete (! OWL *Full*)
- OWL 2: extension to OWL
  - More expressive constructs
    - Role-chains and characteristics
    - Negative assertions
    - Punning

Metadata & Applications

- Metadata Standards
  - Capture a *level of meaning* of (web) resources
  - Predate SW standards
- The Dublin Core
  - 15 main elements
  - Many other *qualifications*
    - Sub-elements – correspond to ***relations***
  - Popular in describing resources
    - In Digital Libraries / Repositories (like DSpace)
    - Supports interoperability
      - ***Structural*** consistency in information exchange

Metadata in DSpace

- Based on Dublin Core
  - Influenced by the Library Application Profile (DC-LAP)
  - A total of 66 elements (some invisible)
    - Including qualifications
- Includes non-standard elements
  - Cannot be mapped to DC
    - e.g. **“author”** and **“sponsorship”**
- Exportable through OAI-PMH
  - Provided a mapping is specified

The Problem

- Metadata in DSpace: Monolithic approach
  - Metadata flatly organized
  - Meaning lies *implicitly* in the **structure**
    - or in the (human-understandable) specifications!
    - Not machine "*understandable*"
  - Semantically **semi-structured** knowledge models
    - In contrast to fully-structured ontology models

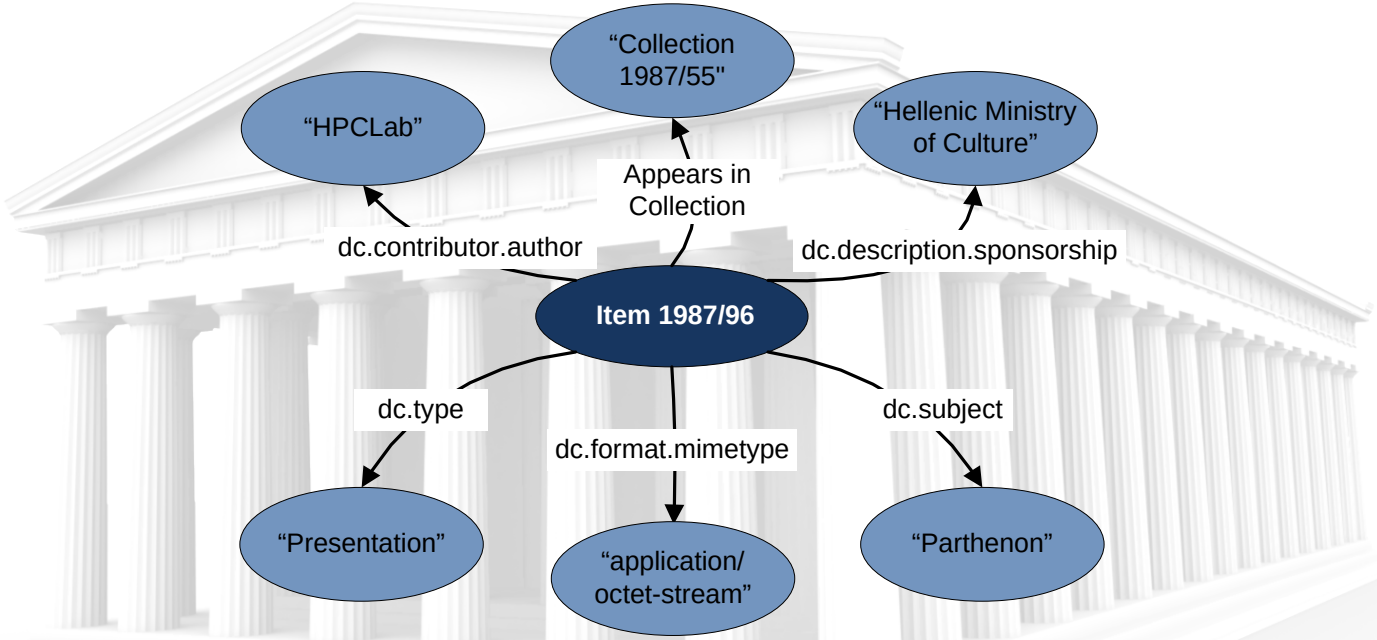
Item 1987/96: Presentation about the 3D reconstruction of the Parthenon of Athens

Item 1987/96

Full Metadata Record

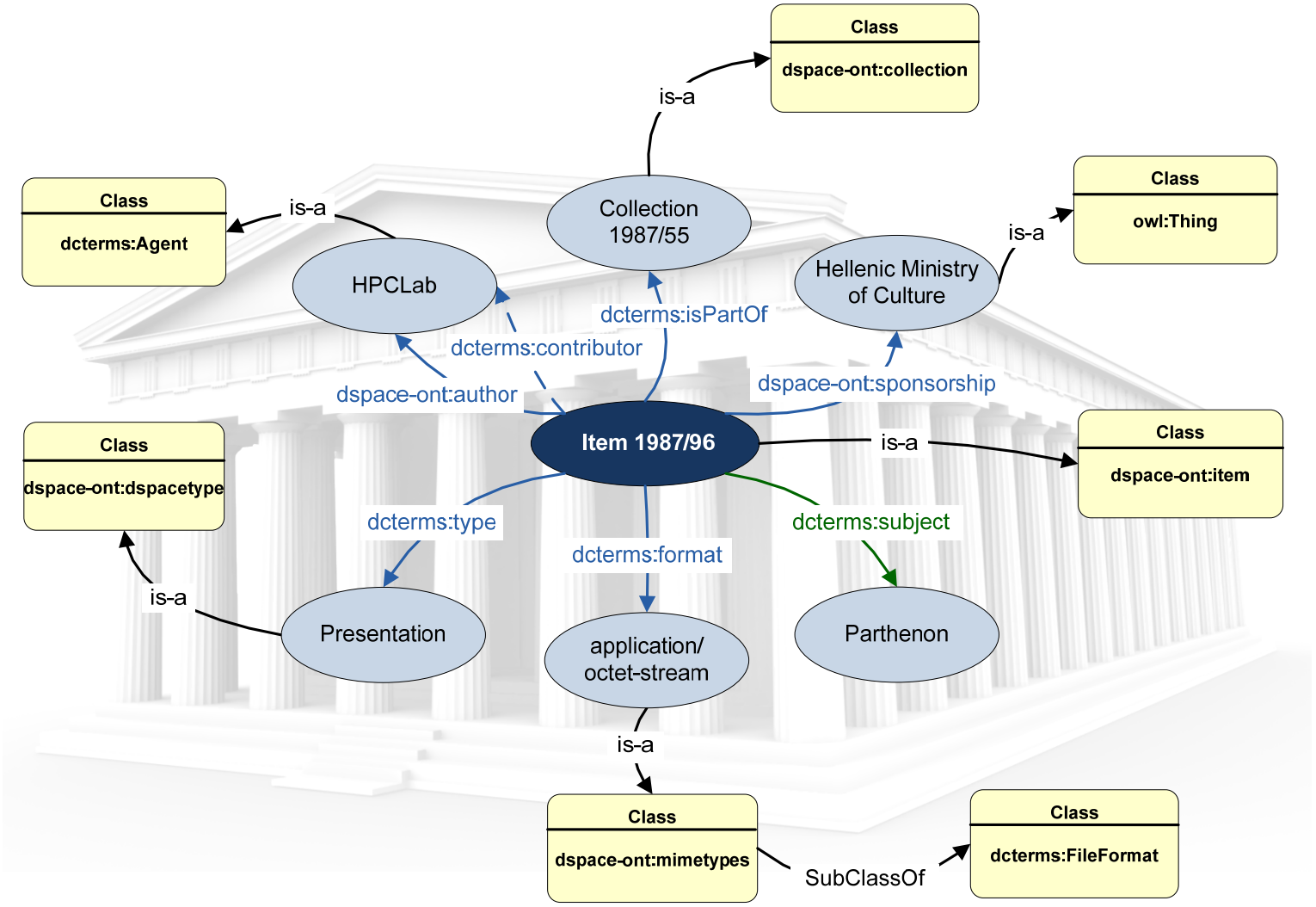
DC Field	Value	Language
dc.contributor.author	HPCLab	en
dc.date.accessioned	2006-12-11T11:41:05Z	-
dc.date.available	2006-12-11T11:41:05Z	-
dc.date.issued	2006-12-11T11:41:05Z	-
dc.identifier.uri	http://hdl.handle.net/1987/96	-
dc.description	This presentation was shown in Copenhagen 11/12/2006	en
dc.description	Η παρουσίαση αυτή προβλήθηκε στην Κοπενχάγη στις 12/11/2006	el
dc.description.abstract	The presentation that is attached describes the virtual reconstruction and 3D Real-time walkthrough inside the Parthenon temple.Issues of objectivity, accuracy and detail are also described and as a conclusion innovations and benefits are presented too.	en
dc.description.abstract	Η παρουσίαση που επισυνάπτεται είναι μια περιγραφή του αντίστοιχου έργου που σκοπό είχε να αποδώσει πραγματικού χρόνου περιπάτους σε μια εικονική αναπαράσταση του Παρθενώνα.Κατά την παρουσίαση αυτή γνωστοποιείται πως αντιμετωπίστηκαν τα θέματα της αντικειμενικότητας, της λεπτομέρειας και της ακρίβειας.Τέλος παρουσιάζονται οι καινοτομίες και τα πλεονεκτήματα του συγκεκριμένου εργου.	el
dc.description.sponsorship	Hellenic Ministry of Culture	en
dc.description.sponsorship	Υπουργείο Πολιτισμού	el
dc.format.extent	25177195 bytes	-
dc.format.mimetype	application/octet-stream	-
dc.language.iso	en	-
dc.subject	Parthenon	en
dc.subject	Παρθενώνας	el
dc.title	Virtual Reconstruction & 3D Real-time Walkthrough Environment	en
dc.type	Presentation	en
dc.type	Παρουσίαση	el
Appears in Collections:	Μνημεία Πολιτισμού	

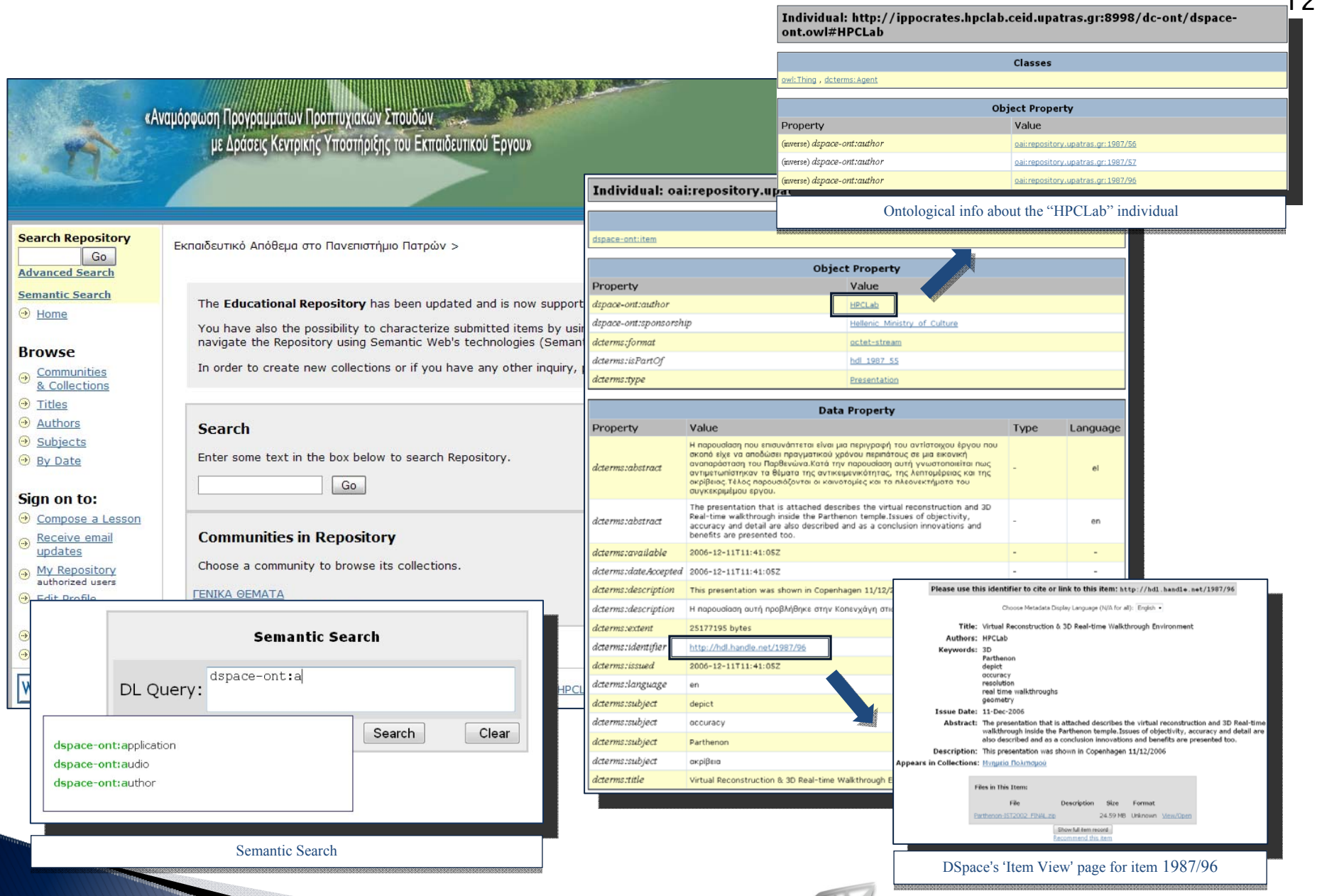
Metadata Relationships



Metadata are flatly organized in the DB. Even structure *is often unimplemented (exists only in the label)*

Semantic Relationships




Individual: http://ippocrates.hpclab.ceid.upatras.gr:8998/dc-ont/dspace-ont.owl#HPCLab
Classes
owl:Thing , dcterms:Agent
Object Property
Property
Value
(inverse) dspace-ont:author
oai:repository.upatras.gr:1987/56
(inverse) dspace-ont:author
oai:repository.upatras.gr:1987/57
(inverse) dspace-ont:author
oai:repository.upatras.gr:1987/96
Ontological info about the "HPCLab" individual
«Αναμόρφωση Προγραμμάτων Προπτυχιακών Σπουδών με Δράσεις Κεντρικής Υποστήριξης του Εκπαιδευτικού Έργου»
Search Repository
Go
Advanced Search
Semantic Search
Home
Browse
Communities & Collections
Titles
Authors
Subjects
By Date
Sign on to:
Compose a Lesson
Receive email updates
My Repository
authorized users
Edit Profile
Εκπαιδευτικό Απόθεμα στο Πανεπιστήμιο Πατρών >
The Educational Repository has been updated and is now support
You have also the possibility to characterize submitted items by usi
navigate the Repository using Semantic Web's technologies (Seman
In order to create new collections or if you have any other inquiry,
Search
Enter some text in the box below to search Repository.
Go
Communities in Repository
Choose a community to browse its collections.
ΓΕΝΙΚΑ ΘΕΜΑΤΑ
Individual: oai:repository.upa
dspace-ont:item
Object Property
Property
Value
dspace-ont:author
HPCLab
dspace-ont:sponsorship
Hellenic Ministry of Culture
dcterms:format
octet-stream
dcterms:isPartOf
hdl 1987 55
dcterms:type
Presentation
Data Property
Property
Value
Type
Language
dcterms:abstract
Η παρουσίαση που επισυνάπτεται είναι μια περιγραφή του αντίστοιχου έργου που σκοπό είχε να αποδώσει πραγματικού χρόνου περιπάτους σε μια εικονική αναπαράσταση του Παρθενώνα.Κατά την παρουσίαση αυτή γνωστοποιείται πως αντιμετωπίστηκαν τα θέματα της αντικειμενικότητας, της λεπτομέρειας και της ακρίβειας.Τέλος παρουσιάζονται οι καινοτομίες και το πλεονεκτήματα του συγκεκριμένου εργου.
el
dcterms:abstract
The presentation that is attached describes the virtual reconstruction and 3D Real-time walkthrough inside the Parthenon temple.Issues of objectivity, accuracy and detail are also described and as a conclusion innovations and benefits are presented too.
en
dcterms:available
2006-12-11T11:41:05Z
dcterms:dateAccepted
2006-12-11T11:41:05Z
dcterms:description
This presentation was shown in Copenhagen 11/12/2
dcterms:description
Η παρουσίαση αυτή προβλήθηκε στην Κοπενχάγη στι
dcterms:extent
25177195 bytes
dcterms:identifier
http://hdl.handle.net/1987/96
dcterms:issued
2006-12-11T11:41:05Z
dcterms:language
en
dcterms:subject
depict
dcterms:subject
accuracy
dcterms:subject
Parthenon
dcterms:subject
ακρίβεια
dcterms:title
Virtual Reconstruction & 3D Real-time Walkthrough E
Please use this identifier to cite or link to this item: http://hdl.handle.net/1987/96
Choose Metadata Display Language (N/A for all): English
Title: Virtual Reconstruction & 3D Real-time Walkthrough Environment
Authors: HPCLab
Keywords: 3D
Parthenon
depict
accuracy
resolution
real time walkthroughs
geometry
Issue Date: 11-Dec-2006
Abstract: The presentation that is attached describes the virtual reconstruction and 3D Real-time walkthrough inside the Parthenon temple.Issues of objectivity, accuracy and detail are also described and as a conclusion innovations and benefits are presented too.
Description: This presentation was shown in Copenhagen 11/12/2006
Appears in Collections: Μνημεία Πολιτισμού
Files in This Item:
File
Description
Size
Format
Parthenon IST2002_FINAL.zip
24.59 MB
Unknown
View/Open
Show full item record
Recommend this item
DSpace's 'Item View' page for item 1987/96
Semantic Search
DL Query:
dspace-ont:a
Search
Clear
dspace-ont:application
dspace-ont:audio
dspace-ont:author
Semantic Search

Approach

- **A.** Create an ***Ontology*** for Dublin Core and DSpace
  - up to OWL 2 level (non-standard inferences)
  - Produce ***meaning*** out of ***structure***
    - e.g. implement qualifiers as *sub-properties*
  - Make **explicit** the spec and common-sense constraints
    - e.g. inverse relation between `dc:hasPart` and `dc:isPartOf`
- **B.** Populate the ontology
  - Transform and map ***existing*** DC metadata to a new ontological model
- **C.** Semantics-aware services for DSpace
  - Semantic Search
  - Semantic Navigation

An Ontological Model for DC & DSpace

- Based on existing DC implementation in RDF
- Create incremental semantic ***profiles*** of DC
  - *...by applying the* ***semantic profiling***[1] *technique*
- Gradually, tailor to the specific domain
  - University of Patras DSpace Installation
  - Based on DSpace
- Preserve the original DC model
  - Physically separate profiles
  - One `owl:imports` the other
  - Smoothly refine the original model

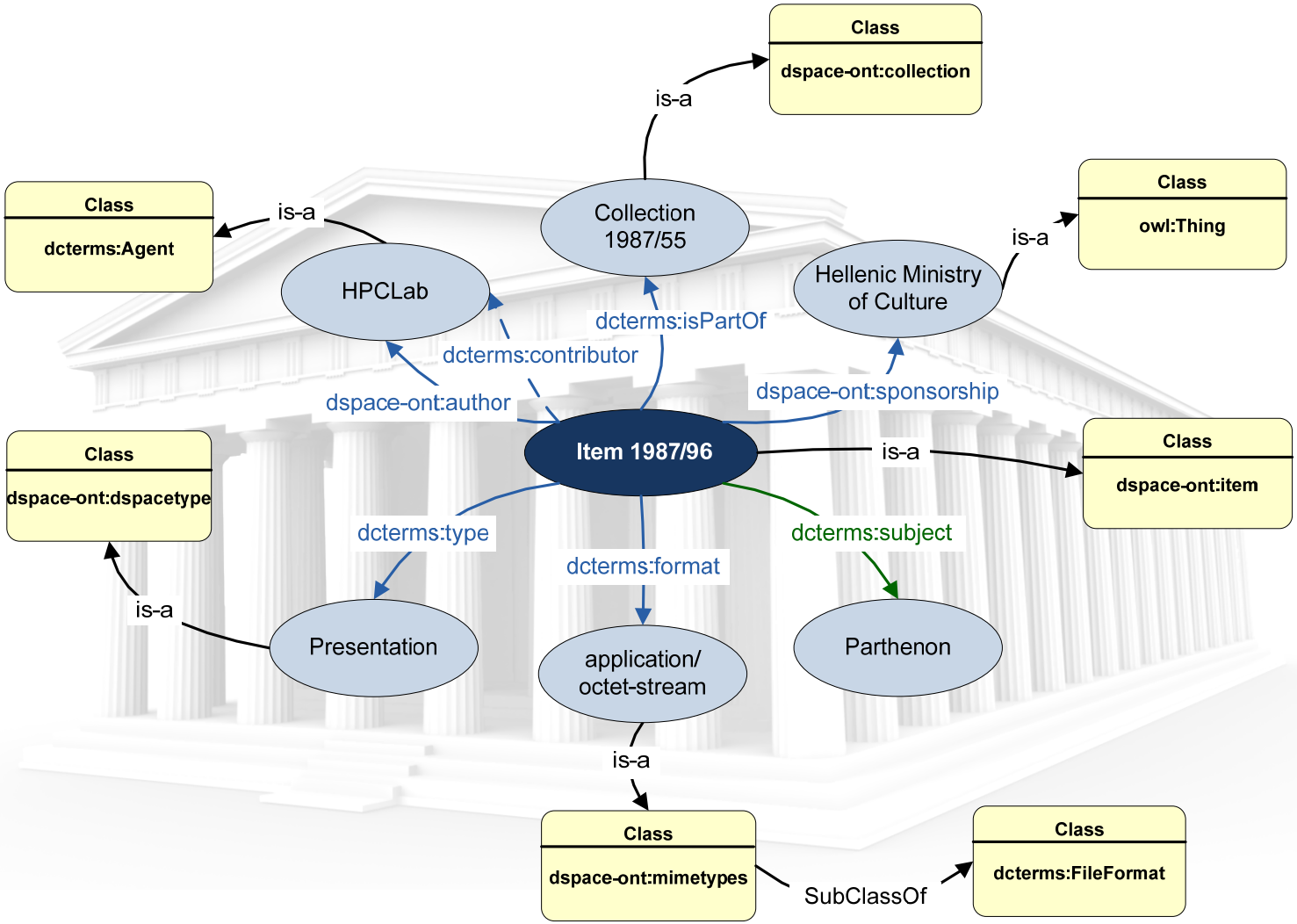
[1] Semantic Interoperability of Dublin Core Metadata in Digital Repositories. In Proc. of 5th International Conference on Innovations in Information Technology (Innovations 2008), pp. 233–237, 2008.

Ontological Data

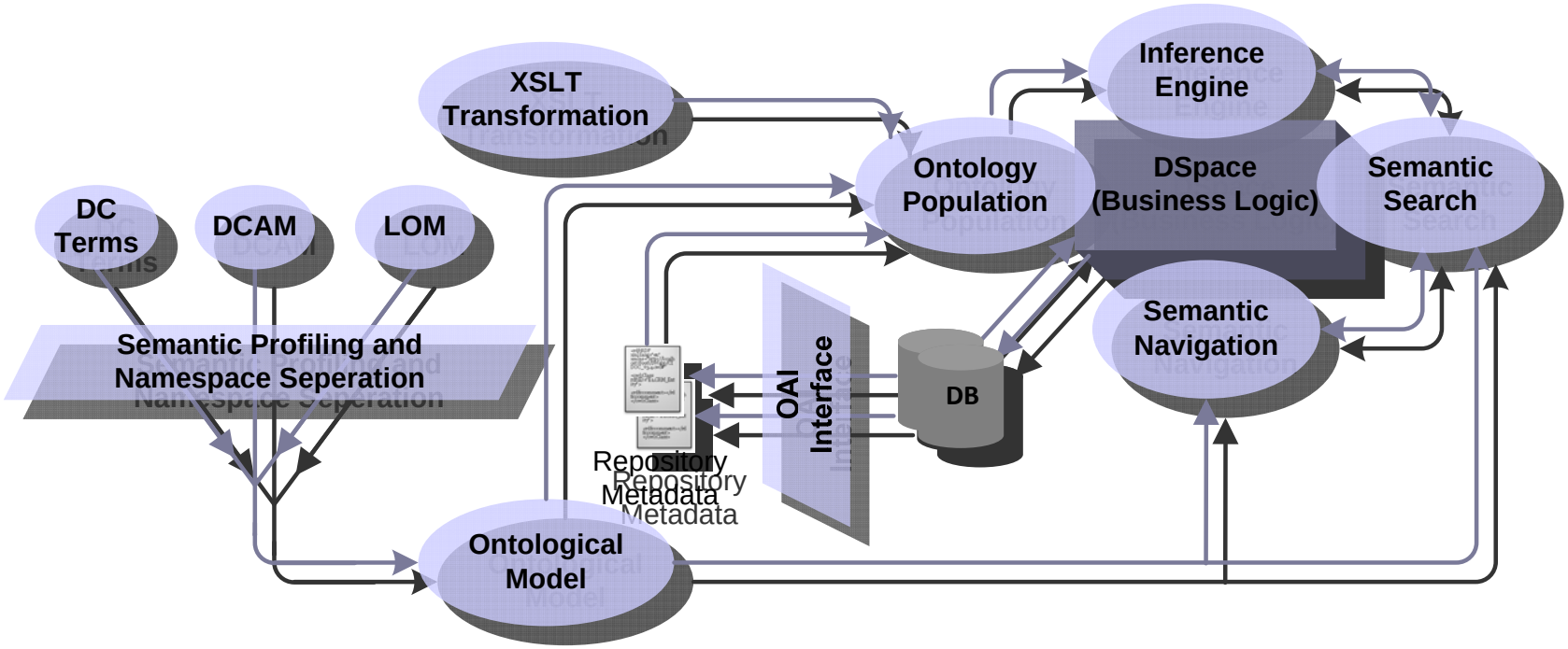
- How to populate the ontology?
  - Harvest and map repository's metadata
  - Through the standard OAI-PMH interoperability interface
- Minimum intervention
  - Not ***altering*** the database
  - Not ***accessing*** the database
- ***Automated*** population
  - Using standard XML-based technologies (XSLT)

…an* interoperable *approach for ontology construction and population

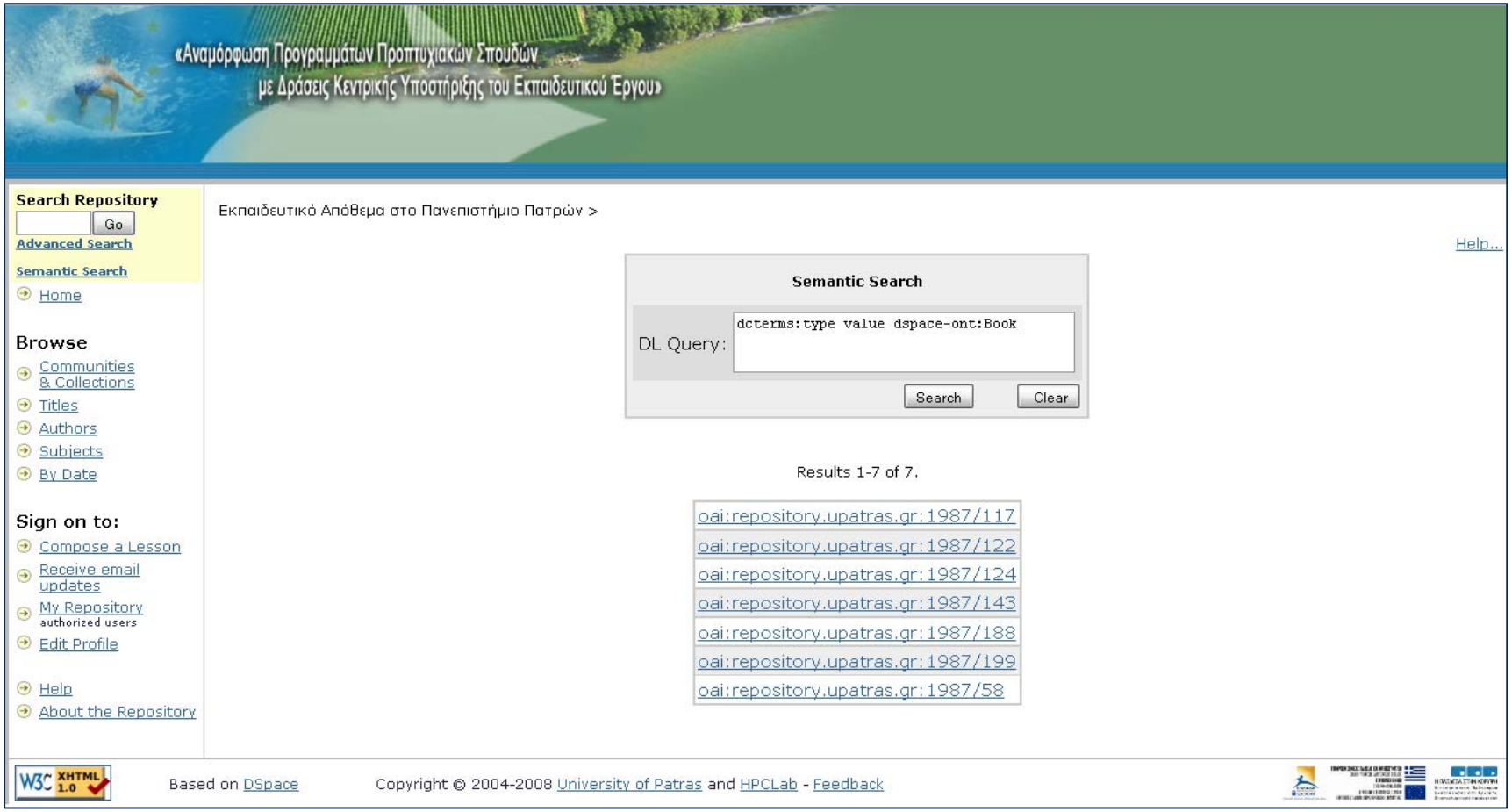
Semantic Relationships



Semantic Services: Architecture

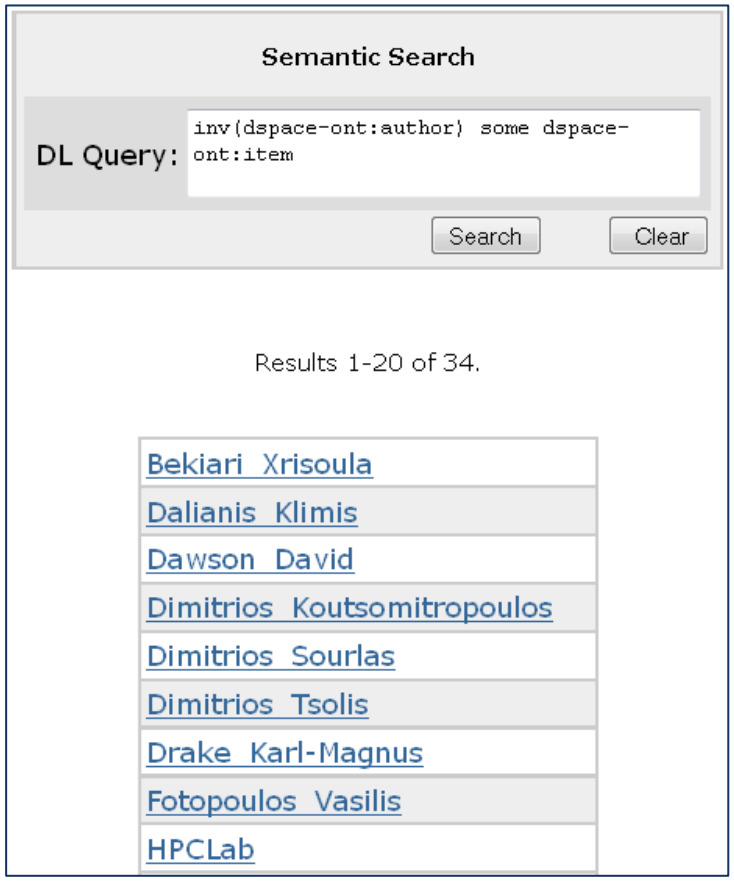


The Semantic Search Interface



Semantic Search

- Queries are typed as simple text using the Manchester OWL syntax
- Type of accepted queries:
  - Valid ontological class names
  - Class expressions (existential qualifications , cardinality restrictions, …)
  - Boolean combinations of class expressions



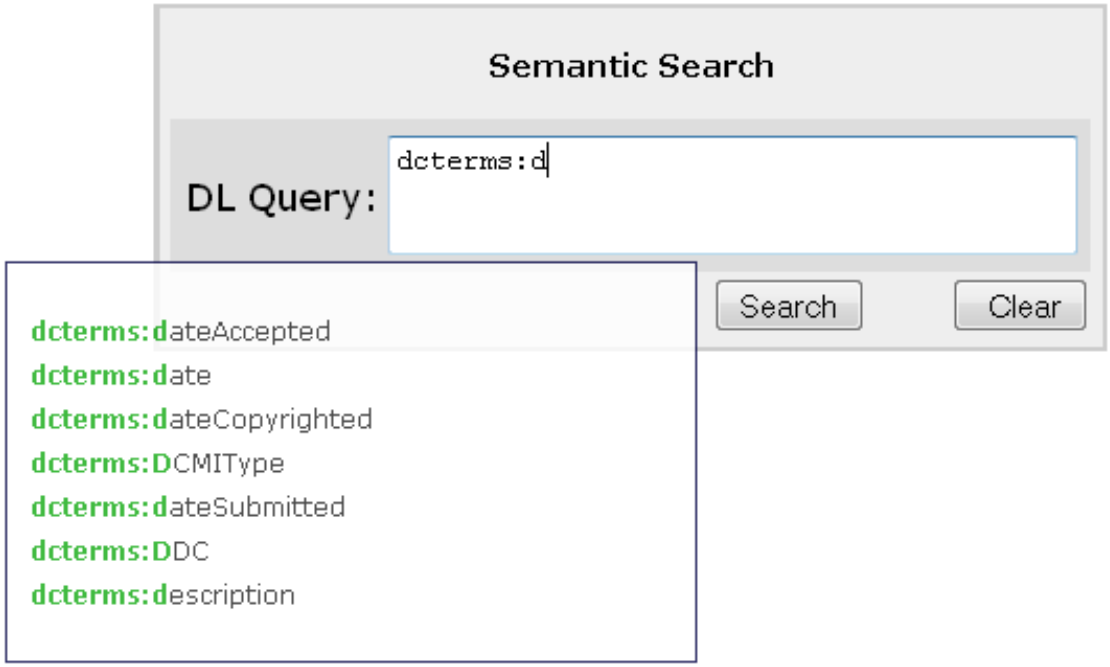
The Manchester OWL Syntax

- A user-friendly syntax for OWL
  - Maps Description Logics symbols to English words and phrases
- Designed for writing OWL class expressions or even complete OWL ontologies

OWL Expression	Description Logics Symbol	Manchester Syntax
someValuesFrom	∃	some
allValuesFrom	∀	only
hasValue	∋	value
minCardinality	≥	min
cardinality	=	exactly
maxCardinality	≤	max
intersectionOf	⊓	and
unionOf	⊔	or
complementOf	¬	not
SubPropertyChain	∘	o

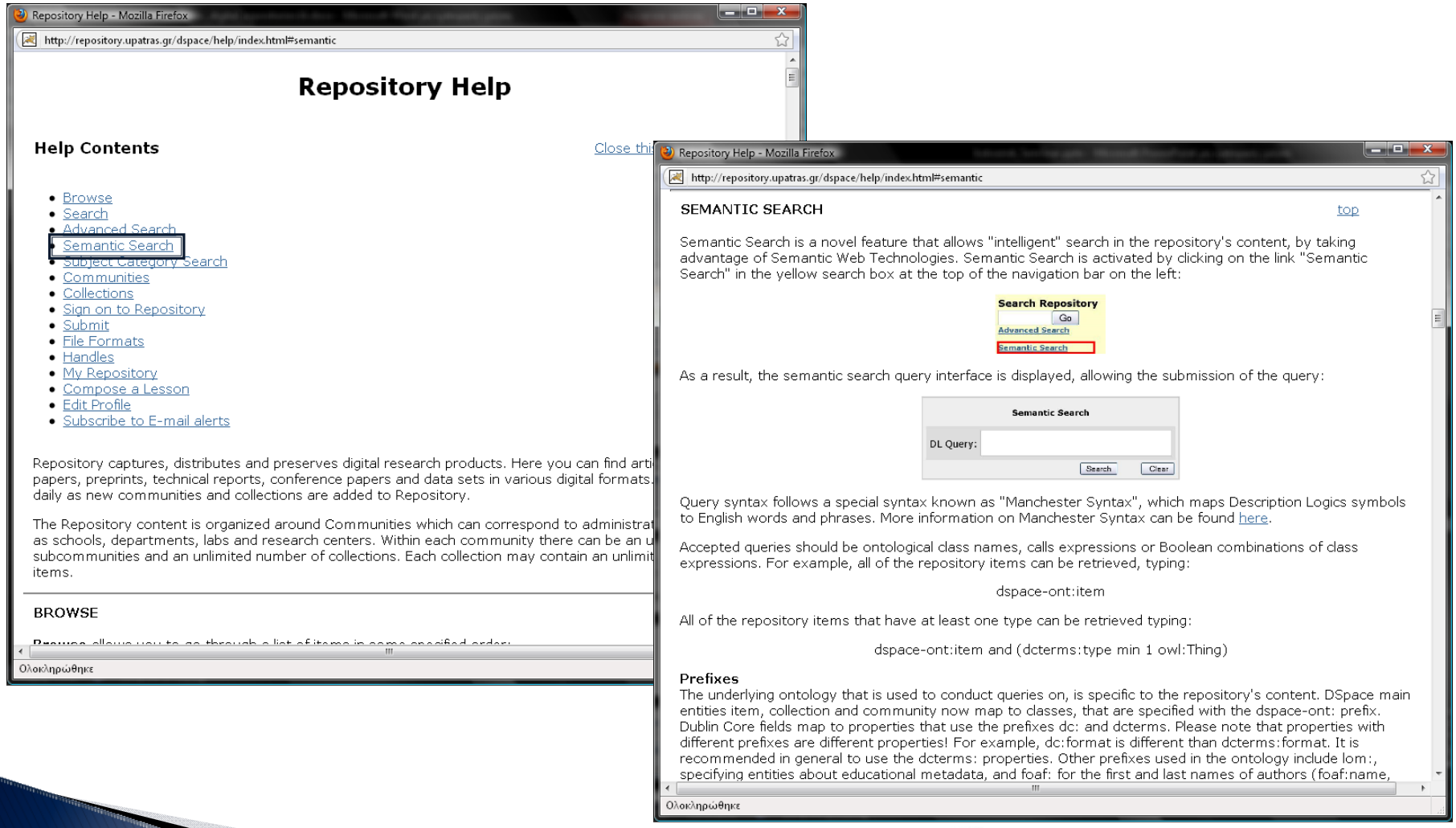
⇨Offers a simpler way to end users for formulating their queries

Auto-complete Facility

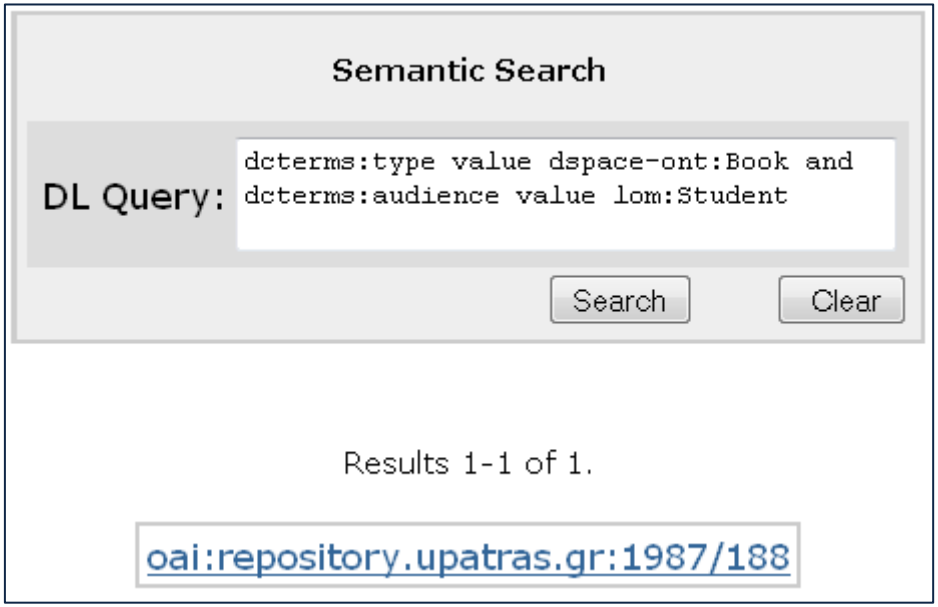


Suggestion of a list of entities that belong to the knowledge base (classes, properties and individuals names)

Help for Semantic Search



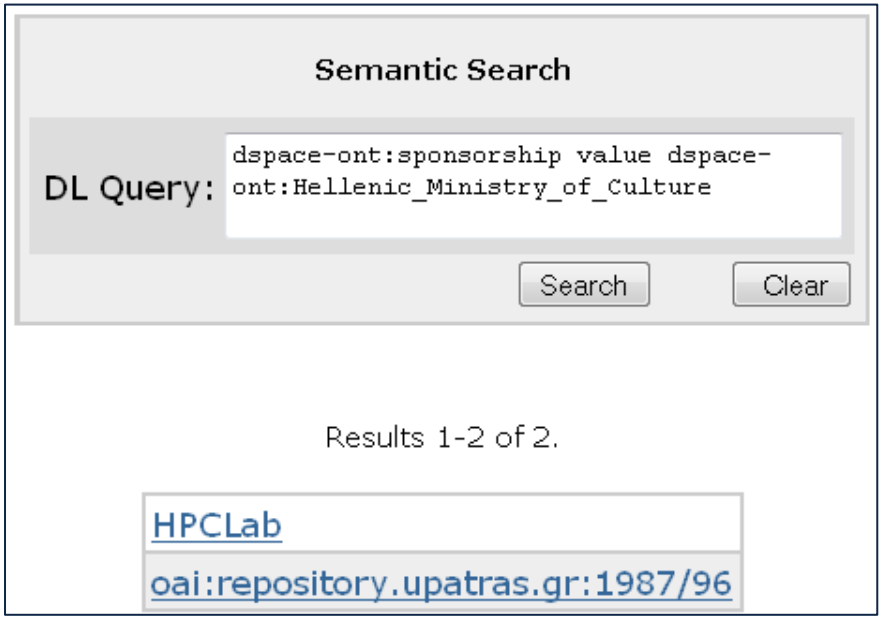
Example–*Conjunctive Queries*



Show all items of type "*Book*" that are mainly intended for Students (audience of type "*Student*")

Example – *Role Chains*

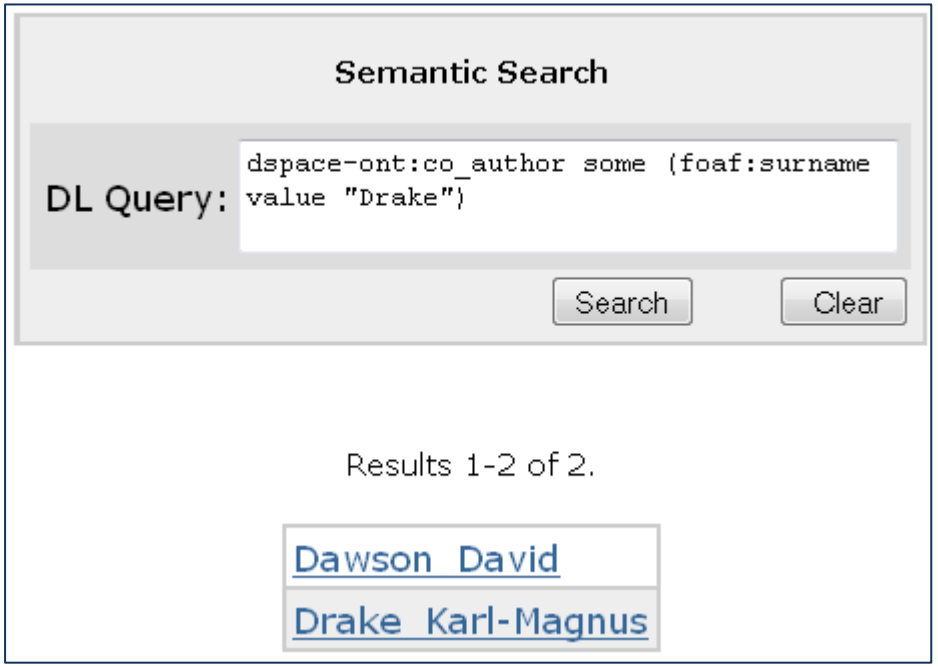
Show who draws sponsorship from the "*Hellenic Ministry of Culture*"



Sponsorship refinement:
inv(dcterms:contributor) **o** sponsorship **SubPropertyOf** sponsorship

Example – *Role Chains* (2)

Find co-authors of author with surname *"Drake"*

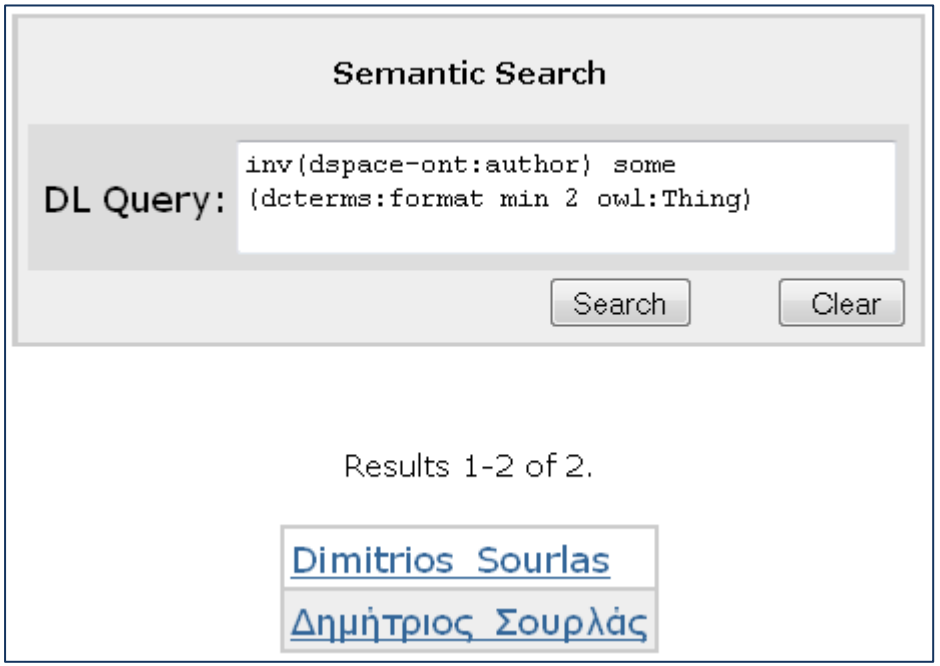


Co-author declaration:

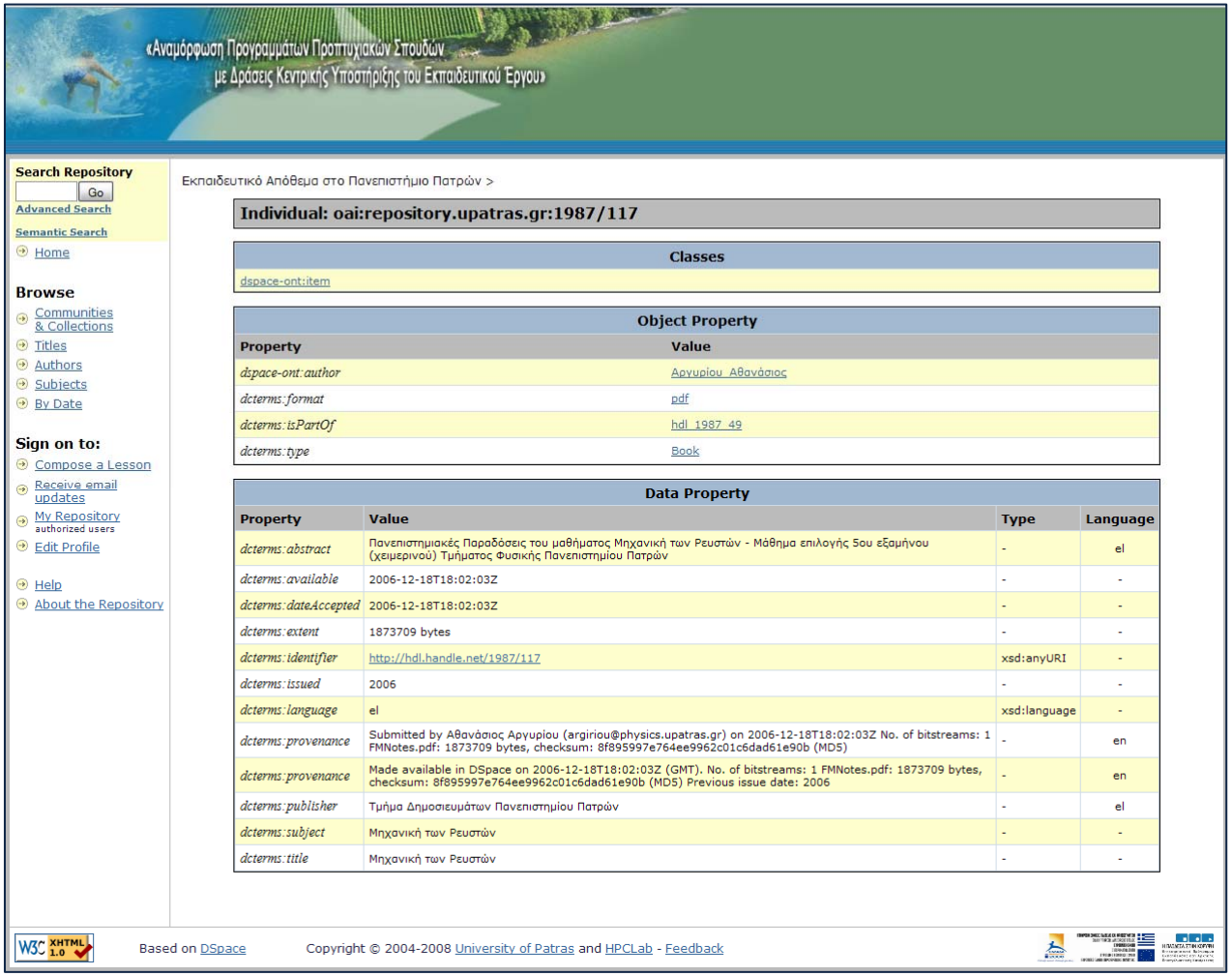
```
inv(author) o author SubPropertyOf co_author
```

Example – *Cardinality Restrictions*

Authors of items that have at least two different formats



Semantic Navigation Pane



Semantic Navigation – Example

Individual: oai:repository.upatras.gr:1987/117

Classes
dspace-ont:item

Object Property	
Property	Value
dspace-ont:author	Αργυρίου Αθανάσιος
dcterms:format	pdf
dcterms:isPartOf	hdl_1987_49
dcterms:type	Book

Data Property			
Property	Value		
dcterms:abstract	Πανεπιστημιακές Παραδόσεις του μαθήματος Μηχανική των Ρευστών - επιλογής 5ου εξαμήνου (χειμερινού) Τμήματος Φυσικής Πανεπιστημίου		
dcterms:available	2006-12-18T18:02:03Z		
dcterms:dateAccepted	2006-12-18T18:02:03Z		
dcterms:extent	1873709 bytes		
dcterms:identifier	http://hdl.handle.net/1987/117		
dcterms:issued	2006		
dcterms:language	el		
dcterms:provenance	Submitted by Αθανάσιος Αργυρίου (argiriou@physics.upatras.gr) on 2006-12-18T18:02:03Z No. of bitstreams: 1 FMNotes.pdf: 1873709 b checksum: 8f895997e764ee9962c01c6dad61e90b (MD5)		
dcterms:provenance	Made available in DSpace on 2006-12-18T18:02:03Z (GMT). No. of bi FMNotes.pdf: 1873709 bytes, checksum: 8f895997e764ee9962c01c6c (MD5) Previous issue date: 2006		
dcterms:publisher	Τμήμα Δημοσιευμάτων Πανεπιστημίου Πατρών	-	el
dcterms:subject	Μηχανική των Ρευστών	-	-
dcterms:title	Μηχανική των Ρευστών	-	-

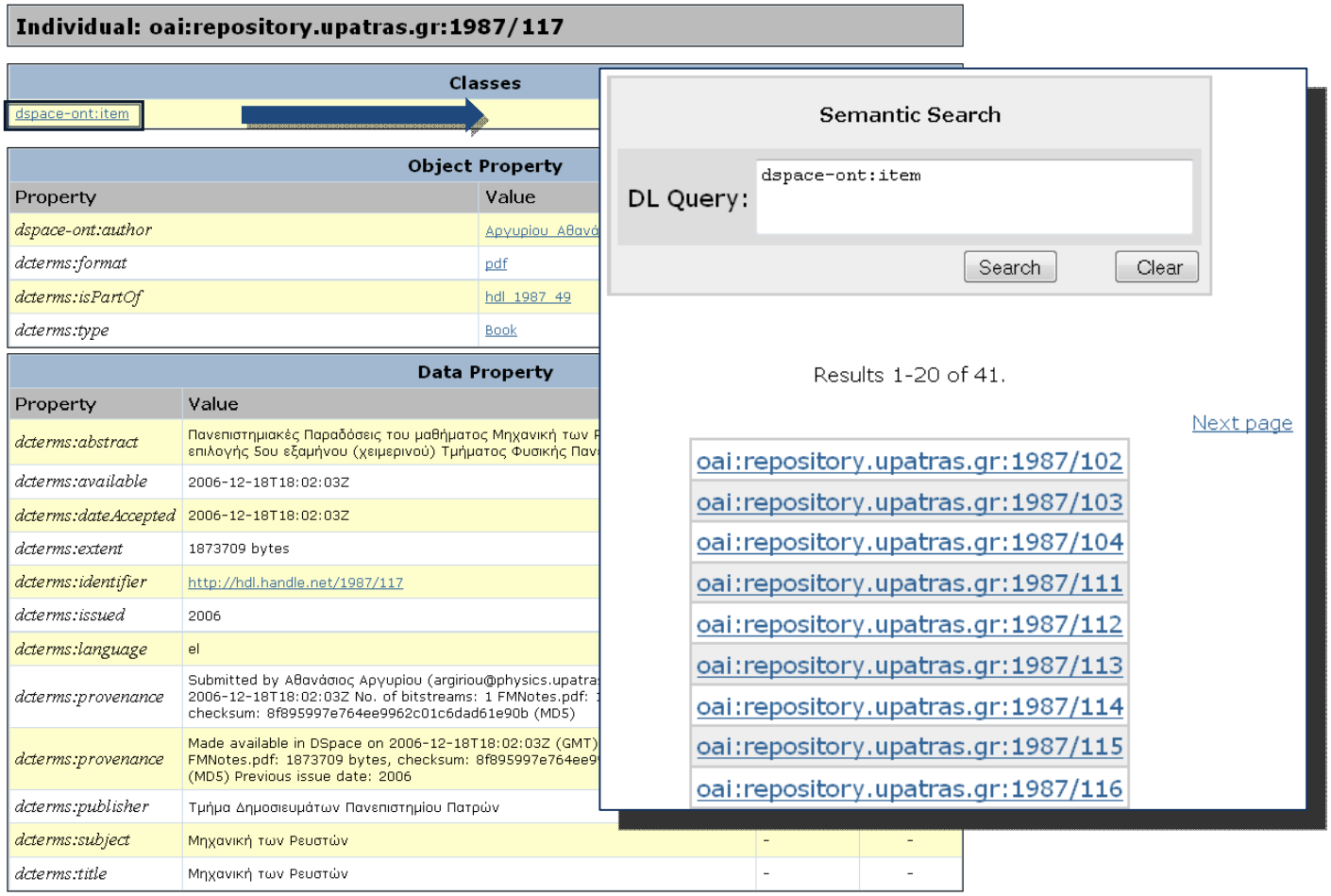
Individual: http://ippocrates.hpclab.ceid.upatras.gr:8998/dc-ont/dspace-ont.owl#Book

Classes
dspace-ont:dspacetype , rdfs:Class

Object Property	
Property	Value
(inverse) *dcterms:type*	oai:repository.upatras.gr:1987/117
(inverse) *dcterms:type*	oai:repository.upatras.gr:1987/122
(inverse) *dcterms:type*	oai:repository.upatras.gr:1987/124
(inverse) *dcterms:type*	oai:repository.upatras.gr:1987/143
(inverse) *dcterms:type*	oai:repository.upatras.gr:1987/188
(inverse) *dcterms:type*	oai:repository.upatras.gr:1987/199
(inverse) *dcterms:type*	oai:repository.upatras.gr:1987/58

Annotation			
Property	Value	Type	Language
label	Βιβλίο	-	el
label	Book	-	en

Semantic Navigation – Example



Semantic Navigation – Example

Individual: oai:repository.upatras.gr:1987/117

Classes
dspace-ont:item

Object Property	
Property	Value
dspace-ont:author	Αργυρίου Αθανάσιος
dcterms:format	pdf
dcterms:isPartOf	hdl 19
dcterms:type	Book

Data Prope			
Property	Value		
dcterms:abstract	Πανεπιστημιακές Παραδόσεις του μαθήματος Μηχ επιλογής 5ου εξαμήνου (χειμερινού) Τμήματος Φ		
dcterms:available	2006-12-18T18:02:03Z		
dcterms:dateAccepted	2006-12-18T18:02:03Z		
dcterms:extent	1873709 bytes		
dcterms:identifier	http://hdl.handle.net/1987/117		
dcterms:issued	2006		
dcterms:language	el		
dcterms:provenance	Submitted by Αθανάσιος Αργυρίου (argiriou@phys 2006-12-18T18:02:03Z No. of bitstreams: 1 FMN checksum: 8f895997e764ee9962c01c6dad61e90b		
dcterms:provenance	Made available in DSpace on 2006-12-18T18:02: FMNotes.pdf: 1873709 bytes, checksum: 8f895997e (MD5) Previous issue date: 2006		
dcterms:publisher	Τμήμα Δημοσιευμάτων Πανεπιστημίου Πατρών	-	el
dcterms:subject	Μηχανική των Ρευστών	-	-
dcterms:title	Μηχανική των Ρευστών	-	-

Please use this identifier to cite or link to this item: `http://hdl.handle.net/1987/117`

Choose Metadata Display Language (N/A for all): English

Title: Μηχανική των Ρευστών

Issue Date: 2006

Appears in Collections: Μετάδοση Θερμότητας - Μεταφορά Μάζας

Files in This Item:

File	Description	Size	Format	
FMNotes.pdf		1.83 MB	Adobe PDF	View/Open

Show full item record

Recommend this item

Possible Improvements

- A more user-friendly interface
  - Provision of a fixed list of common queries expressed in physical human language
  - Facility that will guide (non-familiar with OWL) users in creating queries in Manchester OWL syntax
- Integration with controlled vocabularies/thesauri
  - Expressed in SKOS (OWL)
  - Extend semantic search to include controlled vocabulary/thesaurus concepts
  - Augment subject search
- *Federated* Semantic Search

Conclusions

- Semantics for DSpace metadata
  - Knowledge discovery *(high, OWL 2 expressivity)*
  - Automatic model population
    - Alleviate the “bootstrapping” problem
- Novel Semantic Services
  - Augment traditional search and navigation
  - Intelligent retrieval and discovery
  - “Plug-in” philosophy
- Interoperable design
  - Easy to integrate in any digital repository (OAI-PMH)
  - Straightforward integration with other schemata (e.g. LOM)
  - Semantic interoperability

Questions? Comments?

Dr. Dimitrios Koutsomitropoulos
kotsomit@hpclab.ceid.upatras.gr
http://www.hpclab.ceid.upatras.gr

University of Patras Institutional Repository:
http://repository.upatras.gr/dspace

Semantic Search for DSpace:
http://wiki.dspace.org/index.php/User:Kotsomit
http://www.dspace.org/add-ons-and-extensions/addons/
(code available)