Linked Data Infrastructures and HE Challenges
The Outcomes of the SemTech Project

Thanassis Tiropanis

tt2@ecs.soton.ac.uk
the talk

• semantic technologies, semantic Web and linked data

• the case for linked data in HE (SemTech project)
  - the survey
  - the value of linked data
  - the roadmap

• progress so far
semantic technologies
semantic Web
linked data
Ontologies and Knowledge Modelling

• Concepts
  - Student
  - Class

• Relationships
  - attends
  - is_a_classmate_of

• Instances
  - Student: Alice, Bob, John
  - Class: French
Ontologies and Knowledge Modelling

• Concepts
  - Student
  - Class

• Relationships
  - attends
  - is_a_classmate_of

• Instances
  - Student: Alice, Bob, John
  - Class: French

is_a_classmate_of

Wednesday, 16 February 2011
RDF and SPARQL

http://id.ecs.soton.ac.uk/interest/linked_data

http://rdf.ecs.soton.ac.uk/ontology/ecs#hasInterest

http://id.ecs.soton.ac.uk/person/11208

linked data

SPARQL Query

SPARQL Endpoint

Thanassis
Linked Data Cloud

SOURCE: http://linkeddata.org/
the case for linked data in HE
semantic tech in edu scenario?

- Agreed Ontologies
- Learning Content
- Metadata

- learning content discovery
- personalisation & adaptation
it was time to re-think the value of semantic technologies in HE?

• Web 2.0 promise for content generation annotation
• Value of lightweight knowledge modelling and reasoning
• The HE institutional challenges
• The learning and teaching challenges
• Linked data movement and the Web of data
HE institutional perspective

- Visibility of degree programmes and research output of HE institutions
- Curriculum design
- Recruitment and retention of students
- Efficiency of accreditation
- Collaboration across departments and institutions through workflows
- Integration of knowledge capital, cross-curricular initiatives
- Transparency of data held by educational institutions
learning and teaching perspective

- Course creation and delivery workflows
- Group formation for learning and teaching activities
- Critical thinking and argumentation support
- Personal and group knowledge space construction
- Assessment, certification and addressing of plagiarism
semantic tech in a web 2.0 world

• Soft semantics
  ‣ Meaning in formats that humans can process
  ‣ Lightweight knowledge modelling in Web 2.0 applications

• Hard semantics
  ‣ Meaning in formats that machines can process
  ‣ Processing is independent of domain specific schemas

Wednesday, 16 February 2011
the survey
surveyed semantic technologies

semtech-survey.ecs.soton.ac.uk

- Collaborative Authoring and Annotation
- Searching and Matching
- Repositories, VLEs and Authoring tools
- Infrastructural Technologies for Linked Data and Semantic Enrichment
Collaborative Authoring and Annotation Tools

**Mymory**
Unobtrusive user observation
Meaning co-ordination
Annotation of resource sections

**Compendium**
Visualisation of arguments
Collaborative domain modelling
Real time meeting capture
Searching and Matching tools

**Arnetminer**
- Find experts
- Associations between experts
- Mining RDF from existing repositories

**LUISA**
- Discovery, selection, negotiation and composition of LOs
- Annotation techniques
- Use of Semantic Web Services
Freebase
Collaboratively authored, open repository of structured topics
Topics mined from other repositories
Accessible via open APIs

SKUA
Distributed network of semantically aware shared annotation services in the form of RDF stores
Support for user-facing applications
the value
surveyed semantic technology value

• Technologies:
  - Wiki
  - Tagging
  - Blog/Electronic Journal
  - Shared Bookmarking
  - RDF
  - OWL
  - FOAF
  - SKOS
  - Triple Store
  - Ontology/Taxonomy
  - Archive/Repository

• Value:
  - Well-formed Metadata
  - Interoperability/Data Integration
  - Improved Data Analysis/Reasoning
insights

- Most of the identified HE challenges can be addressed by querying across institutional repositories (databases, web pages, VLEs)
- Significant learning and teaching challenges can be addressed by accessing resources across departments, schools, institutions
- Argumentation and critical thinking could benefit from advanced reasoning over a large scale of resources
- Could we adopt a bottom-up approach starting from linked data which can be related to (layers of) ontologies later in the context of specific applications?

VALUE IN A LINKED DATA FIELD ACROSS HE
the roadmap
breads vs. depth

• The value of semantic technologies on a large scale needs to be investigated

  - In addition to the value of reasoning using ontologies

• Mapping a critical volume of linked data to expressive ontologies can be promising

• Encouragement for community-agreed ontologies can be more effective and flexible

• Expressive semantics to enable pedagogy-aware applications over a large volume of linked data can be meaningful in a Web 2.0+ world
a roadmap of sem tech adoption

**Stage 1:** Linked Data Field
(Triple stores, SPARQL endpoints, RDF)
RDFisers, TALIS, Virtuoso, Collibra, dbpedia.org, freebase.com

**Stage 2:** Ontology-based applications
(Ontology building, mapping linked data, applications)
ArnetMiner

**Stage 3:** Pedagogy-aware reasoning
(Collaborative ontology building, pedagogy in reasoning)
Compendium, Debategraph
a roadmap of sem tech adoption

• Stage 1
  ‣ Exposing internal repositories as linked data, performance optimised triple stores
  ‣ Searching across repositories, matching students, teachers, curricula, research interests

• Stage 2
  ‣ Advanced searching and matching, argumentation and critical thinking applications
  ‣ Mapping linked data to application-wide or community-wide agreed ontologies

• Stage 3
  ‣ Collaborative semantic enrichment of linked data by communities
  ‣ Pedagogy-aware applications and services with reasoning
the network effect

• HE institutions exposing relational databases, VLE material, Web pages as linked data
  ‣ Relevant technologies: RDF, RDFa, VLE plugins
  ‣ Starting from information already available in (X)HTML!

• Applications that use exposed linked data across institutions
  ‣ Curriculum design or alignment
  ‣ Inline recommendation of resources or people
  ‣ Addressing HE challenges such as curriculum alignment, student retention and others using linked data across institutions
progress so far
Where are we now?

- It seems that the first step of the roadmap is well under way.
- Now is the time to discuss issues related to:
  - Exposure of linked data
  - Use of linked data across HE repositories
  - Privacy and confidentiality
  - Searching and matching applications to address HE challenges
Challenges

• Are we ready for the next steps in terms of:
  • Standards
  • Tools
  • Practices

Wednesday, 16 February 2011
Related PhD Research

• Farhana Sarker

Linked Data Technologies to Support Higher Education Challenges: Student Retention, Progression and Completion

• Xin Wang

Query optimisation over Linked Data Repositories in Specific Domains
Related Projects

• SLE (Southampton Learning Environment)
  - looking to enhance the student and staff experience around learning and living at UoS
  - Closely aligned with the Southampton Research Environment (SRE) and the Southampton Business Environment (SBE)
Related Reports

- Paul Miller, Linked Data Horizon Scan http://linkeddata.jiscpress.org/ JISC, 2010

• JISC-funded project working with CETIS

• Objectives:

  * Survey of semantic tools and services
  * Current adoption of semantic technologies in the UK higher education
  * Roadmap of semantic technology adoption in the next 5 years

• www.semtech.ecs.soton.ac.uk
Acknowledgements

The SemTech team:
Hugh Davis
Faith Lawrens
David Millard
Asma Ounnas
Heather S. Packer
Marcus Ramsden
Daniel A. Smith
Thanassis Tiropanis
Mark Weal
Su White
Gary Wills
Learning Societies Lab
(ECS-University of Southampton)

The JISC CETIS Semantic Technology Working Group:
Sheila MacNeill (CETIS)
Lorna Campbell (CETIS)
Phil Barker (CETIS)
Helen Beetham
Simon Buckingham-Shum (Open University, UK)
David Davies (University of Warwick)
Michael Gardner (University Essex)
Tony Linde (University of Leicester)
Wilbert Kraan (CETIS)
Sue Manuel (University of Loughborough)
Lou McGill
Graham Wilson (LT Scotland)
Robin Wylie (LT Scotland)
David Kernohan (JISC)

The SemTech workshop participants:
Colin Allison (University of St. Andrews)
Chris Bailey (University of Bristol)
Liliana Cabral (Knowledge Media Institute, Open University)
Patrick Carmichael (University of Cambridge)
Tom Franklin (Franklin Consulting)
David Kay (Sero Consulting)
George Magoulas (London Knowledge Lab, Birkbeck College)
Uma Patel (City University)
Alex Poulavassilis (London Knowledge Lab, Birkbeck College)
John Scott (University of Essex)
SemTech activities


• 1st International SemTech workshop: “A roadmap for semantic technology adoption in UK higher education” accepted to the ALT-C conference, Manchester, 8-10 September 2009.


• Publications
Publications


Thank you!

Thanassis Tiropanis - tt2@ecs.soton.ac.uk

The special issue (IEEE TLT): http://www.computer.org/

The project: www.semtech.ecs.soton.ac.uk

The survey: semtech-survey.ecs.soton.ac.uk

The workshop: www.semhe.org