Federating Distributed Social Data to Build an Interlinked Online Information Society

Alexandre Passant\textsuperscript{1}, Matthias Samwald\textsuperscript{1,2},
John Breslin\textsuperscript{1,3}, Stefan Decker\textsuperscript{1}

\textsuperscript{1} Digital Enterprise Research Institute, NUI Galway
\textsuperscript{2} Konrad Lorenz Institute for Evolution and Cognition Research, Austria
\textsuperscript{3} Department of Electronic Engineering, NUI Galway

WebSci09, Athens
19th March 2009
Agenda

- **Social Semantic Information Spaces**
  - Bridging the gap between Web 2.0 and the Semantic Web

- **Implementing SSIS in Enterprise 2.0**
  - Improving search in collaborative environments

- **Integrating, interlinking and re-using SSIS**
  - Connecting data and people on a global scale

- **Relationships with the Web Science agenda**
  - Relevance of SSIS in this context
Social Semantic Information Spaces

Collaboration and Communication Tools
- Blogs, Forums, OSNs, Wikis

World Wide Web
- URIs, HTML, HTTP

Social Semantic Information Spaces
- Web, Desktop

Semantic Web
- RDFS, OWL, SPARQL

Syntax

Semantic
Requirements for SSIS

- Communities sharing and building data collectively
  - Blogs, wikis, social networking tools ...

- An additional layer of semantics
  - To model activities of online communities and related user-generated data in order to provide interoperable applications
Requirements for SSIS

- Communities sharing and building data collectively
  - Blogs, wikis, social networking tools ...

- An additional layer of semantics
  - To model activities of online communities and related user-generated data in order to provide interoperable applications

- Web 2.0 and the Semantic Web are not disjoint
  - But should be combined to realize the vision of “social machines” [1]
“A little semantics goes a long way”

- Semantics regarding the structure of the communities and their social interactions
  - FOAF - Friend Of A Friend - people and social networks
  - SIOC - Semantically-Interlinked Online Communities - content and interactions within the community
“A little semantics goes a long way”

- **Semantics regarding the structure of the communities and their social interactions**
  - FOAF - Friend Of A Friend - people and social networks
  - SIOC - Semantically-Interlinked Online Communities - content and interactions within the community

- **Semantics regarding the content itself**
  - Domain ontologies (RDF(S)/OWL) and taxonomies (SKOS)
  - Reusing existing data - The Linking Open Data project
“A little semantics goes a long way”

- Semantics regarding the structure of the communities and their social interactions
  - FOAF - Friend Of A Friend - people and social networks
  - SIOC - Semantically-Interlinked Online Communities - content and interactions within the community

- Semantics regarding the content itself
  - Domain ontologies (RDF(S)/OWL) and taxonomies (SKOS)
  - Reusing existing data - The Linking Open Data project

- A need to combine the two levels
  - Thanks to Semantic tagging principles
Semantically-Interlinked Online Communities [5]

- A lightweight (but powerful enough) model to represent online communities and user-generated content using Semantic Web technologies
Semantically-Interlinked Online Communities [5]

- A lightweight (but powerful enough) model to represent online communities and user-generated content using Semantic Web technologies.
Adoption?
Adoption?
Adoption?
Adoption?
Enterprise 2.0 [9]

“the use of emergent social software platforms within companies, or between companies and their partners or customers”
SSIS for Enterprise 2.0

- **Enterprise 2.0 [9]**
  - "the use of emergent social software platforms within companies, or between companies and their partners or customers"

- **Information fragmentation and reusability**
  - How to find all information about a particular object, spread between wikis, blogs, RSS feeds, etc.
  - Most of this information is plain-text only
SSIS for Enterprise 2.0

- Enterprise 2.0 [9]
  - “the use of emergent social software platforms within companies, or between companies and their partners or customers”

- Information fragmentation and reusability
  - How to find all information about a particular object, spread between wikis, blogs, RSS feeds, etc.
  - Most of this information is plain-text only

- Tagging
  - Ambiguity, heterogeneity, lack of organization
  - Expertise-related gap in tagging behaviors, depending on the basic levels of knowledge of each user [14]
A first layer of common semantics

- Lightweight add-ons, completely automated exports
SIOC for Enterprise 2.0

- A first layer of common semantics
  - Lightweight add-ons, completely automated exports
- Enable integration of user-generated content
  - From various distributed applications
- **A first layer of common semantics**
  - Lightweight add-ons, completely automated exports
- **Enable integration of user-generated content**
  - From various distributed applications
Semantic wikis in use

- Ontology population is a complex task
  - Knowledge engineering skills, RDF(S)/OWL modeling
  - Semantic wikis can ease the process
Semantic wikis in use

- **Ontology population is a complex task**
  - Knowledge engineering skills, RDF(S)/OWL modeling
  - Semantic wikis can ease the process

- **UfoWiki**
  - Combining forms and lightweight ontologies
  - Community-oriented ontology population and evolution
Semantic wikis in use

- Ontology population is a complex task
  - Knowledge engineering skills, RDF(S)/OWL modeling
  - Semantic wikis can ease the process

- UfoWiki
  - Combining forms and lightweight ontologies
  - Community-oriented ontology population and evolution
MOAT - Meaning Of A Tag

- A lightweight model and framework to bridge the gap between tagging and semantic indexing [12]
- *In this blog post, I’m using the tag “apple” as in* [http://dbpedia.org/Apple_Inc.](http://dbpedia.org/Apple_Inc.) *(the computer brand), not as* [http://dbpedia.org/Apple](http://dbpedia.org/Apple) *(the fruit)*
Semantic tagging

- MOAT - Meaning Of A Tag
  - A lightweight model and framework to bridge the gap between tagging and semantic indexing [12]
  - *In this blog post, I’m using the tag “apple” as in* <http://dbpedia.org/Apple_Inc.> *(the computer brand), not as* <http://dbpedia.org/Apple> *(the fruit)*

- Linking tags to ontology instances from the wikis
  - And allow to create new instances
Semantic tagging

- **MOAT - Meaning Of A Tag**
  - A lightweight model and framework to bridge the gap between tagging and semantic indexing [12]
  - *In this blog post, I’m using the tag “apple” as in* <http://dbpedia.org/Apple_Inc.> *(the computer brand), not as* <http://dbpedia.org/Apple> *(the fruit)*

- **Linking tags to ontology instances from the wikis**
  - And allow to create new instances
A complete interlinked graph
Improving search by using semantic annotations

- But hide the RDF(S)/OWL and SPARQL complexity to end-users
- From plain-text to concept search
Semantic search for Enterprise 2.0

- Improving search by using semantic annotations
  - But hide the RDF(S)/OWL and SPARQL complexity to end-users
  - From plain-text to concept search

- Tags related to that concept (MOAT + Domain ontologies)
- Main wiki page (SIOC + Semantic Wikis + Domain ontologies)
- Related wiki pages from several wikis
- Latest blog posts tagged with that URI (SIOC + MOAT)
Providing and interlinking SSIS

- The previous approach can be re-used in other domains
  - As soon as people share and build information in a collaborative way regarding common topics
  - Interlinking various SSIS provides an interlinked on-line information society rather than isolated islands
The previous approach can be re-used in other domains
- As soon as people share and build information in a collaborative way regarding common topics
- Interlinking various SSIS provides an interlinked on-line information society rather than isolated islands

Health-Care and Life Sciences
- SWANSIOC - http://esw.w3.org/topic/HCLSIG/SWANSIOC
- Integrating SWAN and SIOC for discourse representation
- Combined with biomedical knowledge bases
Interlinking via vocabularies

- Agreement on some core vocabularies for publishing data using lightweight semantics
  - Unified querying capabilities w/ SPARQL
Interlinking via people

- Distributed an open social-networking
  - Integrating networks from disconnected data silos
  - Can be achieved thanks to FOAF exporters
Interlinking via common topics

- Re-use existing URIs for common topics
  - Linking Open Data as a reference

http://richard.cyganiak.de/2007/10/iod/
Interlinking via common topics

- Re-use existing URIs for common topics
- Linking Open Data as a reference

[http://richard.cyganiak.de/2007/10/lod/]
... and increasing information value

- Any data independently published can be reused globally
  - Distributed Collective Intelligence with semantics
  - Low-cost Semantic Mash-ups - reusing data across SSIS
... and increasing information value

- Any data independently published can be reused globally
  - Distributed Collective Intelligence with semantics
  - Low-cost Semantic Mash-ups - reusing data across SSIS
Web Science is "a science that seeks to develop, deploy, and understand distributed information systems, systems of humans and machines, operating on a global scale" [2]
Web Science is "a science that seeks to develop, deploy, and understand distributed information systems, systems of humans and machines, operating on a global scale" [2]

SSIS (and the Semantic Web in general) can help by providing:

- Ecosystems where machines help humans to build knowledge collaboratively
- Ways to make the process of the study and understanding of these systems easier thanks to standard representation formats
“How can we extend the current Web infrastructure to provide mechanisms that make the social properties of information-sharing explicit and guarantee that the use of this information conforms to relevant social-policy expectations?” [7]
“How can we extend the current Web infrastructure to provide mechanisms that make the social properties of information-sharing explicit and guarantee that the use of this information conforms to relevant social-policy expectations?” [7]

SSIS explicit these social properties in a machine-readable form

- Without extending the Web architecture [15]
“How can we extend the current Web infrastructure to provide mechanisms that make the social properties of information-sharing explicit and guarantee that the use of this information conforms to relevant social-policy expectations?” [7]

- SSIS explicit these social properties in a machine-readable form
  - Without extending the Web architecture [15]
- More open and interlinked data != less privacy
  - But advanced policies for access control [11]
Summary

- **Web 2.0 and the Semantic Web**
  - A complementary approach to build *social machines*

- **SSIS can be applied to any socially-aware ecosystem**
  - Using lightweight semantics and simple add-ons

- **Connecting SSIS**
  - Provides a complete network of people, machines and data

- **The Social Semantic Web**
  - Can make the process of studying the Web easier
Thank you for your attention!

- Any questions?

- Contact
  - <alexandre.passant@deri.org>
  - http://apassant.net

- Paper available at:
  - http://journal.webscience.org/179/