



HCLS Tutorial: The W3C Health Care and Life Sciences Interest Group





Introduction to HCLS

M. Scott Marshall
co-chair HCLS IG
Leiden University Medical Center
University of Amsterdam



Background of the HCLS IG



- Originally chartered in 2005
 - Chairs: Eric Neumann and Tonya Hongsermeier
- Re-chartered in 2008
 - Chairs: Scott Marshall and Susie Stephens
 - Team contact: Eric Prud'hommeaux
- Broad industry participation
 - Over 100 members
 - Mailing list of over 600
- Background Information
 - <http://www.w3.org/blog/hcls>
 - <http://esw.w3.org/topic/HCLSIG>



Mission of HCLS IG



- The mission of HCLS is to develop, advocate for, and support the use of Semantic Web technologies for
 - Biological science
 - **Translational medicine**
 - Health care
- These domains stand to gain tremendous benefit by adoption of Semantic Web technologies, as they depend on the **interoperability of information** from many domains and processes for efficient decision support



Translating across domains



- Translational medicine – use cases that cross domains
- Link across domains and research:
 - What are the links?
 - gene – transcription factor – protein
 - pathway – molecular interaction – chemical compound
 - drug – drug side effect – chemical compound

But also:

- Link discourse to raw data
- 

Group Activities



- **Document use cases** to aid individuals in understanding the business and technical benefits of using Semantic Web technologies
- **Document guidelines** to accelerate the adoption of the technology
- **Implement** a selection of the **use cases** as proof-of-concept demonstrations
- **Develop** high-level **vocabularies**
- **Disseminate information** about the group's work at government, industry, and academic events



How does the HCLS IG work?




- Task forces
- Regular teleconferences using teleconference bridge (Zakim), IRC, minutes
- Face2Face (F2F) twice a year
- Procedures for publishing W3C notes



Current Task Forces



- **BioRDF** – federating (neuroscience) knowledge bases
 - Kei Cheung (Yale University)
 - **Clinical Observations Interoperability** – patient recruitment in trials
 - Vipul Kashyap (Cigna Healthcare)
 - **Linking Open Drug Data** – aggregation of Web-based drug data
 - Chris Bizer (Free University Berlin)
 - **Translational Medicine Ontology** – high level patient-centric ontology
 - Christi Denney (Eli Lilly)
 - **Scientific Discourse** – building communities through networking
 - Tim Clark (Harvard University)
 - **Terminology** – Semantic Web representation of existing resources
 - John Madden (Duke University)
- 

Accomplishments



Demonstrations:

- http://hcls.deri.org/hcls_demo.html
- Demonstrator of querying across heterogeneous EHR systems
 - <http://hcls.deri.org/coi/demo/>
- <http://www.w3.org/2009/08/7tmdemo>
- <http://ws.adaptivedisclosure.org/search>
- HCLS KB hosted at 2 institutes
- Linked Open Data contributions

Interest Group Notes:

- HCLS KB
- Integration of SWAN and SIOC ontologies for Scientific Discourse
 - SWAN
 - SIOC
 - SWAN-SIOC

Technologies: <http://sourceforge.net/projects/swobjects/>



Accomplishments II



- Conference Presentations:
 - Bio-IT World, WWW, ISMB, AMIA, etc.
- (Co)Organized Workshops:
 - C-SHALS, SWASD, SWAT4LS 2009, IEEE Workshop
- Publications:
 - Proceedings of LOD Workshop at WWW 2009: Enabling Tailored Therapeutics with Linked Data
 - Proceedings of the ICBO: Pharma Ontology: Creating a Patient-Centric Ontology for Translational Medicine
 - AMIA Spring Symposium: Clinical Observations Interoperability: A Semantic Web Approach
 - BMC Bioinformatics. A Journey to Semantic Web Query Federation in Life Sciences
 - Briefings in Bioinformatics. Life sciences on the Semantic Web: The Neurocommons and Beyond



We've come a long way



- Triplestores have gone from millions to billions
- Linked Open Data cloud
- <http://lod.openlinksw.com/>
- On demand Knowledge Bases: Amazon's EC2
- Terminologies: SNOMED-CT, MeSH, UMLS, ..
- Neurocommons, Flyweb, Biogateway, Bio2RDF, Linked Life Data
- Shared Names, NCBO's BioPortal, Concept Web Alliance (CWA)
- HCLS/CWA/SIB Uniprot project started



Penetrance of ontology in biomedicine



- OBO Foundry - <http://www.obofoundry.org>
- BioPortal - <http://bioportal.bioontology.org>
- National Centers for Biomedical Computing <http://www.ncbcs.org/>
- Shared Names <http://sharednames.org>
- Concept Web Alliance
<http://conceptweblog.wordpress.com/conferences/>
- Semantic Web Interest Group PRISM Forum
<http://www.prismforum.org/>
- Work packages in ELIXIR <http://www.elixir-europe.org/>



Provenance



- Represent knowledge so that others can discover where a fact (or triple) came from and evaluating how to use it – link facts to data as evidence
- Named graphs
- HCLS works together with the Provenance IG (Yolanda Gils)
- Has come up in work in BioRDF and SciDisc task forces
- Concept Web Alliance has adopted SWAN/SIOC as the starting point for modeling provenance



Shared Identifiers



- Must use common URI's in order to link data
- Provenance related:
 - Identifiers for people (researchers)
 - Terminology servers (identifiers for terms)
 - Identifiers for compounds
- Shared Names initiative



A common misconception




- Creating a Semantic Web application does not necessarily require converting and storing all data in RDF
- “Leave the data where it lives”
- Create *views* of the data accessible via RDF export or **BETTER** a SPARQL endpoint



Getting Involved



- Benefits to getting involved include:
 - Expertise - Early access to use cases and best practice
 - Influence standard recommendations
 - Cost effective exploration of new technology through collaboration and networking
 - Get involved
 - Speak to any of us after the session!
 - Email chairs and team contact
 - team-hcls-chairs@w3.org
 - Participate in the next F2F (last one was here):
 - http://esw.w3.org/topic/HCLSIG/Meetings/2009-11-30_F2F
- 

Tutorial Outline



- **Intro to W3C HCLS IG** - M. Scott Marshall
- **Health Care Use Case** - Vipul Kashyap
- **Pharma Use Case** - Elgar Pichler
- **Scientific Discourse Use Case** - Tim Clark
- **Terminology Considerations** - John Madden
- **Federated Query** - Kei Cheung

