



Semantic Web for Healthcare and Life Sciences Interest Group: A Vision for Advancing Research Communities

> Eric Neumann Clinical Semantic Group W3C HCLS chair, MIT Fellow



- Today's Informatics Problems
- Semantic Web Key-Points
- HCLSIG: The Vision
- The group's Scope and Participants
- Current Activities and Demonstrations:
 - Data Aggregation,
 - Science Collaborations
 - Knowledge Mining
- => http://www.w3.org/2001/sw/hcls/

A web of information



R. Stevens

HCLS Challenges ⇒Bench-to-Bedside Vision

- No common way to specifically link data and docs together
 HTML links carries no meaning with them
- Today's integration approaches prevent data re-use
- No global way to annotate our experiments
 - Most annotations cannot be found by context
 - Need a "sci-blog" for data interpretation
- Enterprise Information access and discoverability are weak
 - Making timely discoveries!
 - Why we all like Google
- Cutting and pasting between docs promotes fact mutation
 and loss of provenance
 - Address business tracking operations, reduce static data copying

Distributed Nature of R&D



The Big Picture In R&D: "Omics"







Hard to understand from just a few isolated Points of View



Complete view tells a very different Story

What is the Semantic Web?

Myths

Semantic Web...

- is Top-Down
- needs ontologies at the beginning
- requires all information to be converted to RDF
- must be centralized
- handles only binary relations
- requires the entire graph to exist on one memory store



- A means of describing (a Web of) Data
- A system defining and incorporating semantics
- A mechanism for making statements on things
- A format for associating metadata
- A strategy for federating data systems (with or without triplestore)



The Technologies: RDF

- Resource Description Framework (RDF)
- W3C standard for making statements or hypotheses about data and concepts
- Descriptive statements are expressed as triples: (Subject, Verb, Object)



The URI - global identification

URI serves as a universal and uniform identifier for all web based resources.

Used to map web documents AND data together!

Can be a URL or URN*

* urn:lsid:biopathways.org:biopax:hs_kreb_cycle



XML or RDF?

XML

- Known technology
- Existing Developer base
- Web Standard
- Parsers must work with DTDs

RDF

- Emerging technology
- Limited developer base
- Web Standard
- Normative with OWL
- No parser dev needed (1 Syntax for RDF-XML)
- How fast is change occurring in data semantics?
- *Will it need to integrate/reference outside data?*
- How well are DTDs and parsers being written and maintained?

The Data Space

Separate RDF documents are merged automatically into one aggregate graph.



Recombination between homologous chromosomes:

Preserves coding integrity



16

Recombinant Data

Graphs can be filtered and pivoted, without losing meaning



Recombinant Data

- Dynamic mixing of data
- Mash-ups that don't lose semantics
- Provide Different Views for Different Roles and Functions
 - Dashboards
 - Reports
- Works Directly with SPARQL output

Extending Computational Biology



Where does SW Fit?



Key Functionality offered by Semantic Web

Descriptiveness

- Data, Metadata, and Statements (Hypotheses)

- Ubiquity
 - Same identifiers for anything from anywhere
- Discoverability
 - Global search over the Web
- Application Independent => "Recombinant Data"
 - Interoperability becomes a consequence
 - No new parsers needed!

Emerging Standards Supporting SW

- SPARQL (CR) Query
- GRDDL (PR) XML to RDF
- **RDFA -** Embed RDF into HTML
- RIF (soon CR) Using Rules on RDF-OWL

HCLSIG Vision: Scope and Participants

HCLSIG Vision: Scope and Participants



W3C Semantic Web Health Care and Life Sciences Interest Group

The Semantic Web Health Care and Life Sciences Interest Group is designed to improve collaboration, research and development, and innovation adoption in the health care and life science industries. Aiding decision-making in clinical research, Semantic Web technologies will bridge many forms of biological and medical information across institutions.

Contents: Mission and Scope | Membership and Joining | Charter /History | Resources | Presentations | Articles | New and Events | Conferences | Task Forces

Nearby: Discussion archive | HCLS WIKI | Applications and Demonstrations | OWL | RDF Data Access | Rules | Semantic Web Best Practices and Deployment

Introduction

Both Life Science Research and Health Care are areas undergoing phenomenal growth, holding much promise for our future as long as we can manage and apply the new knowledge gained without driving up costs. Key to their success is the implementation of new informatics models that will unite many forms of biological and medical information across all institutions, through the encoding of meaning into the data and their interpretations. By focusing on the semantics of information, researchers will have more access to the knowledge required to effectively find cures to diseases, while doctors will have better tools for individualized clinical management of patients.

Mission and Scope

The Semantic Web for Health Care and Life Sciences Interest Group (HCLSIG) is chartered to develop and support the use of Semantic Web technologies and practices to improve collaboration, research and development, and innovation adoption in the of Health Care and Life Science domains. Success in these domains depends on a foundation of semantically rich system, process and information interoperability. (more).

The scope of HCLSIG will include:

- · Core vocabularies and ontologies to support cross-community data integration and collaborative efforts
- · Guidelines and Best Practices for Resource Identification to support integrity and version control
- · Better integration of Scientific Publication with people, data, software, publications, and clinical trials

News and Events

- Last Call: SPARQL Query Language for RDF 2007-03-27: Comments are due by 18 April.(<u>Permalink</u>)
- <u>HCLS demo</u>, planned for <u>WWW2007 in Banff</u>. To help participate in the demo, please contact <u>Alan</u> <u>Ruttenberg</u>.
- FIRST INTERNATIONAL WORKSHOP ON HEALTH CARE AND LIFE SCIENCES DATA INTEGRATION FOR THE SEMANTIC WEB, May 8, WWW2007 in Banff.
- Eric Prud'hommeaux, new W3C staff contact for HCLS.
- <u>GRDDL</u> links Microformats and Semantic Web: Working Draft <u>xmlns="http://www.w3.org/2000/svg"(Permalink)</u>
- Upcoming <u>HCLS Workshop at ISWC2006</u>, November 6th, Athens, Georgia - <u>Conference Registration</u>
- Next F2F HCLS Meeting Scheduled for Oct 3-4, 2006, in Amsterdam, NL <u>Announcement</u> and <u>Registration</u>

Vision

- Aggregating *data* and *statements* using the Web
 - Aggregation specified by need and role
 - "Recombinant Data"
- Common system of referencing things (no copying)
 - even is they exist in one of many databases
 - All resources can be indexed
- URI resources can have types and be tagged
 - Common and ad hoc vocabularies
- Supporting the collective knowledge of an R&D Community
 - A Wiki that is type aware and has semantic relations
 - Substrate for Next Generation Discovery Tools

Vision

- Aggregating Life Sciences *data* and their *statements* on the Web
- Common system for referencing things across all research domains
 - even is they exist in one of many databases
- URI resources can be typed and associated with cross-community vocabularies
- Support the collective knowledge of an R&D Community
 - A Wiki that is type aware and has semantic relations
 - Substrate for Next Generation Discovery Tools

(Neumann and Prusak Briefings in Bioinformatics, 2007.)

HCLSIG Scope

- Advance Data Semantics and Interoperability in support of Scientific Research
- Promote good usage of vocabularies
- Foster good practices for resource identification (URI)

Community

- Academic Researchers
- Drug R&D
- Healthcare Providers
- Governmental Agencies
- Publishers

HCLS Ecosystem



HCLSIG Task Forces

- BioRDF
- BioONT
- AHPP Adaptive Healthcare Protocol & Pathways
- DSE Drug Safety and Efficacy
- See http://esw.w3.org/topic/HCLSIG

HCLS Activities

HCLS Conferences and Workshops

- HCLSIG F2F's (W3C members and invited experts)
 - Jan 2006
 - Oct 2006
 - http://www.w3.org/2001/sw/hcls
- ISWC 2006 http://esw.w3.org/topic/HCLS/ISWC/Workshop
- WWW2007 http://www2007.org/workshop-W2.php

HCLS Neuro Demo: Harnessing the Semantic Web to Answer Scientific Questions

http://esw.w3.org/topic/HCLS/Banff2007Demo







A simple query: Biological processes in dendrites?

Alzheimer's disease is characterized by neural degeneration. Among other things, there is damage to dendrites and axons, parts of nerve cells.

What resources do we have available to learn more about biological processes in dendrites?

Integration and analysis of heterogeneous data sets

· Hypothesis, Genome, Pathways, Molecular Properties, Disease, etc.



Biological processes naming dendrites

PREFIX owl: <http://www.w3.org/2002/07/owl#>

PREFIX go: <http://purl.org/obo/owl/GO#>

PREFIX obo: <http://www.geneontology.org/formats/oboInOwl#>

PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>

select ?name ?class ?definition from http://purl.org/commons/hcls/20070416

where

{ graph <http://purl.org/commons/hcls/20070416/classrelations>
 {?class rdfs:subClassOf go:GO_0008150}
 ?class rdfs:label ?name.
 ?class obo:hasDefinition ?def.
 ?def rdfs:label ?definition
 filter(regex(?name,"[Dd]endrite"))
A SPARQL query spanning 4 sources

prefix go: <http://purl.org/obo/owl/GO#> prefix rdfs: <http://www.w3.org/2000/01/rdf-schema#> prefix owl: <http://www.w3.org/2002/07/owl#> prefix mesh: <http://purl.org/commons/record/mesh/> prefix sc: <http://purl.org/science/owl/sciencecommons/> Mesh: Pyramidal Neurons prefix ro: <http://www.obofoundry.org/ro/ro.owl#> select ?genename ?processname where { graph <http://purl.org/commons/hcis/pubmesh> { ?paper ?p mesh:D017966 . ?article sc:identified_by_pmid ?paper. Pubmed: Journal Articles ?gene sc:describes gene or gene product mentioned by ?article. graph <http://purl.org/commons/hcls/goa> { ?protein rdfs:subClassOf ?res. ?res owl:onProperty ro:has function. ?res owl:someValuesFrom ?res2. ?res2 owl:onProperty ro:realized_as. Entrez Gene: Genes ?res2 owl:someValuesFrom ?process. graph <http://purl.org/commons/hcls/20070416/classrelations> {{?process <http://purl.org/obo/owl/obo#part of> go:GO 0007166} union {?process rdfs:subClassOf go:GO_0007166 }} ?protein rdfs:subClassOf ?parent. ?parent owl:equivalentClass ?res3. ?res3 owl:hasValue ?gene. **GO: Signal Transduction** graph <http://purl.org/commons/hcls/gene> { ?gene rdfs:label ?genename } graph <http://purl.org/commons/hcls/20070416> { ?process rdfs:label ?processname} Inference required

The Allen Brain Atlas Project

20,000 genes, 400000 images



HCLS <--> Allen Brain Atlas

How it works (standing on the shoulders of giants)



...but you can also use Exhibit to visualize



Acknowledgements

HCLS Demo Contributors

- John Barkley (NIST)
- Olivier Bodenreider (NLM, NIH)
- Bill Bug (Drexel University College of Medicine)
- Huajun Chen (Zhejiang University)
- · Paolo Ciccarese (SWAN)
- · Kei Cheung (SenseLab, Yale)
- Tim Clark (SWAN)
- · Don Doherty (Brainstage Research Inc.)
- Kerstin Forsberg (AstraZeneca)
- · Ray Hookaway (HP)
- Vipul Kashyap (Partners Healthcare)
- June Kinoshita (AlzForum)
- · Joanne Luciano (Harvard Medical School)
- · Scott Marshall (University of Amsterdam)
- Chris Mungall (NCBO)
- Eric Neumann (Teranode)
- Eric Prud'hommeaux (W3C)
- Jonathan Rees (Science Commons)
- Alan Ruttenberg (Science Commons)
- · Matthias Samwald (Medical University of Vienna)

HCLS Demo Contributors

- Susie Stephens (Eli Lilly)
- Mike Travers
- Gwen Wong (SWAN)
- Elizabeth Wu (SWAN)

Data Providers

- Judith Blake (MGD)
- Mikail Bota (BAMS)
- David Hill (MGD)
- Oliver Hoffman (CL)
- Minna Lehvaslaiho (CL)
- Colin Knep (Alzforum)
- Maryanne Martone (CCDB)
- Susan McClatchy (MGD)
- Simon Twigger (RGD)
- Allen Brain Institute

Vendor Support

- OpenLink Kingsley Idehen, Ivan Mikhailov, Orri Erling, Mitko Iliev, Patrick van Kleef
- · HP Ray Hookaway, Jeannine Crockford

Drug Safety and Efficacy Horizontal Connectivity for Translational Medicine

Translational Medicine in Drug R&D



Across Clinical Studies

- Multi-Study Tox analysis
- Surrogate POC markers
- Strategy comparisons



 Sren
 Celores
 Energy
 Prob

 Weight
 (kod)
 (kJ)
 (g)

 28U
 6UU
 251.04
 5.0

10 21 1 25

Descriptor usond Hela", ckead

'Almond Pella', chedida

OFU, CREDOAR

TOFUL CHEDDAR

SDTM* ala RDF

* Study Data Tabulation Model

```
<http://clinic.com/study/T2271/subject/4183542663506>
```

```
a cdisc:Subject ;
nci:sex_code nci:Female ;
cdisc:treatment
<http://clinic.com/study/T2271/subject/4183542663506/observation/02241> ;
cdisc:vitalSigns
<http://clinic.com/study/T2271/subject/4183542663506/observation/06561> ;
cdisc:adverseEvent
<http://clinic.com/study/T2271/subject/4183542663506/observation/06622> ;
```

// ROUTE DRGGROUP DOSE pid treatment tpfday tptday
// IV B 7 MG 4183542663506 7mg then 14mg SEMWEB 6/11/84 7/11/84
<http://clinic.com/study/T2271/subject/S83221/observation/O2241 >

```
a cdisc:Treatment ; // cdisc:Treatment is a subclass of cdisc:Observation
cdisc:design_arm <http://clinic.com/study/T2271/treated_B/double_dose> ;
cdisc:route cdisc:IV_route ;
cdisc:drug_group "B";
cdisc:dose "7" ;
cdisc:dose_units nist:mg ;
cdisc:treatment "7mg then 14mg SEMWEB" ;
cdisc:first_date "6/11/84" ;
cdisc:term_date "7/11/84" .
```

SDTM + Biomarkers

<http://clinic.com/study/T2271/subject/S83221/observation/02241 >

a dse:Biomarker_Measure ; // a subclass of cdisc:Observation

dse:biomarker_proc <http://clinic.com/study/T2271/treated_B/biomarker_sample> ;

dse:mol_analyses nci:gene_expression_analysis ;

dse:biomarker_set <http://nci.nih.gov/biomarkers/colon_cancer/B324> ;

dse:biomarker_values [2.343, 1.211, 0531, 23.34, 83.12, 4.323, 9.543];

nist:units nist:norm_ratio ;

dc:date "6/11/84" ;

http://simile.mit.edu/exhibit

covt

Study Demography View

Treatment View

Adverse Event View

DEMOGRAPHY • RACE AGE-WEIGHT • AE AGE-WEIGHT • SMOKER AGE-WEIGHT • DRUG TREATMENT AGE-WEIGHT

SMOKED

300 Person total.								146 FEMALE	
Subject	Age	Sex	Weight	Height	Race	Smoker	Wine	Adverse Events	154 MALE
3793476724374	35.0	FEMALE	62.0	161.0	HISPANIC	SMOKER	0.0		
2076247169253	39.0	FEMALE	69.0	175.0	WHITE	SMOKER		http://127.0.0.1/ae_315, http://127.0.0.1/ae_288, http://127.0.0.1/ae_92, http://127.0.0.1/ae_75, and http://127.0.0.1/ae_228	
4441828514263	41.0	MALE	91.0	180.0	WHITE	EX	1.0		smkclasst
						SMOKER			85 EX SMOKER
3419674301026	41.0	MALE	82.0	171.0	HISPANIC	NEVER SMOKED	1.0		72 SMOKER
1173462489794	42.0	MALE	105.0	176.0	BLACK	NEVER SMOKED			
1184610454148	43.0	MALE	73.0	168.0	ASIAN	EX SMOKER		http://127.0.0.1/ae_151, http://127.0.0.1/ae_382, http://127.0.0.1/ae_156, and http://127.0.0.1/ae_50	RACE
4668933019258	43.0	MALE	85.0	159.0	HISPANIC	NEVER SMOKED		http://127.0.0.1/ae_322	52 ASIAN 49 BLACK 102 HISPANIC
1070626710873	43.0	MALE	99.0	180.0	ASIAN	NEVER SMOKED	1.0	http://127.0.0.1/ae_202 and http://127.0.0.1/ae_54	97 WHITE
164836978298	44.0	MALE	78.0	166.0	WHITE	NEVER			
				1 and 1 a					treatment
316/83136/99/	44.0	MALE	93.0	188.0	WHITE	SMOKER	0.0		39 Smg HCLSIGDSE
4694858090847	45.0	FEMALE	84.0	170.0	ASIAN	EX SMOKER	0.0	http://127.0.0.1/ae_323 and http://127.0.0.1/ae_401	76 Smg then 10mg HC 77 7mg SEMWEB
1492306644693	45.0	FEMALE	76.0	173.0	HISPANIC	EX SMOKER		http://127.0.0.1/ae_100, http://127.0.0.1/ae_44, http://127.0.0.1/ae_117, and http://127.0.0.1/ae_157	32 PLACEBO
2771536224259	45.0	FEMALE	101.0	182.0	HISPANIC	NEVER SMOKED	2.0		AFEDIT
	10000		Lage and	1000000	10000000	Catalogue	-		119 MILD
332173461604	45.0	FEMALE	67.0	162.0	WHITE	NEVER SMOKED		http://127.0.0.1/ae_340, http://127.0.0.1/ae_339, http://127.0.0.1/ae_209, and http://127.0.0.1/ae_245	101 MODERATE 48 SEVERE
2182615806968	45.0	MALE	78.0	169.0	HISPANIC	NEVER	0.0	http://127.0.0.1/ae 130, http://127.0.0.1/ae 293, http://127.0.0.1/ae 116, http://127.0.0.1/ae 347, and	

http://127.0.0.1/ae 16

http://simile.mit.edu/exhibit

Study Demography View

reatment View

dverse Event View



http://simile.mit.edu/exhibit

Adverse Events View

Demography View

Treatment View

407

ADVERSE EVENT TABLE • AE FIRST-TERM DAY • AE AGE-WEIGHT • AE TIMELINE

Subject .	AE_ID	Adverse Event	Sex	first day	last day	Symptom	System	Duration (days)
1009121340704	ae_274	MILD	FEMALE	11.0	93.0	NAUSEA DIGESTIVE		82
1009121340704	ae_186	MILD	FEMALE	4.0	94.0	GOUT	METABOLIC AND NUTRITIONAL	90
1044093510666	ae_298	MODERATE	FEMALE	34.0	46.0	PERIPHERAL EDEMA	METABOLIC AND NUTRITIONAL	12
1044093510666	ae_105	MILD	FEMALE	-13.0	1.0	CONSTIPATION	DIGESTIVE	14
1044093510666	ae_52	MILD	FEMALE	-5.0	-2.0	ARTHROSIS	MUSCULOSKELETAL	3
1050794548864	ae_56	SEVERE	MALE	-9.0	-6.0	ASTHENIA	BODY AS A WHOLE	3
1050794548864	ae_290	MODERATE	MALE	20.0	30.0	PARESTHESIA	NERVOUS	10
1050794548864	ae_90	MODERATE	MALE	87.0	87.0	CARCINOMA	BODY AS A WHOLE	0
1060417365103	ae_365	MODERATE	FEMALE	-15.0	-13.0	STUPOR	NERVOUS	2
1060417365103	ae_242	MODERATE	FEMALE	-16.0	-14.0	LAB TEST ABNORMAL	BODY AS A WHOLE	2
1060417365103	ae_330	SEVERE	FEMALE	-15.0	-13.0	RECTAL DISORDER	DIGESTIVE	2
10674902998869	ae_292	MILD	FEMALE	-14.0	68.0	PARESTHESIA NERVOUS		82
10674902998869	ae_377	MODERATE	FEMALE	17.0	70.0	TOOTH DISORDER DIGESTIVE		53
10674902998869	ae_134	MILD	FEMALE	3.0	70.0	DIARRHEA DIGESTIVE		67
1070626710873	ae_202	MILD	MALE	6.0	24.0	HERNIA BODY AS A WHOLE		18
1070626710873	ae_54	MILD	MALE	20.0	24.0	ARTHROSIS MUSCULOSKELETAL		4
1074594411070	ae_386	MILD	MALE	36.0	86.0	ULCERATIVE STOMATITIS DIGESTIVE		50
1075396035292	ae_383	MODERATE	MALE	3.0	3.0	ULCERATIVE STOMATITIS DIGESTIVE		0
1075396035292	ae 7	MODERATE	MALE	-4.0	-4.0	ALKALINE PHOSPHATASE INCREASED METABOLIC AND NUTRITIONAL		0

AESEVT

- 203 MILD 151 MODERATE
- 53 SEVERE

PREFTEXT

- 2 ABDOMEN ENLARGED 1 ABDOMINAL PAIN
- 1 ABSCESS
- 2 ACCIDENTAL INJURY
- 4 ALKALINE PHOSPHATASE INCREASED
- BODYTEXT
- 57 BODY AS A WHOLE
- 56 CARDIOVASCULAR
- 75 DIGESTIVE
- 18 HEMIC AND LYMPHATIC
- 59 METABOLIC AND NUTRITIONAL
- 18 MUSCULOSKELETAL

sext 194 FEMALE

213 MALE

smkclasst 103 EX SMOKER

http://simile.mit.edu/exhibit

Adverse Events View

Demography View

Treatment View

ADVERSE EVENT TABLE • AE FIRST-TERM DAY • AE AGE-WEIGHT • AE TIMELINE

Subject .	AE_ID	Adverse Event	Sex	first day 2.0	last day 2.0	Symptom	System	Duration (days)
1121526482163	ae_196	MILD	FEMALE			HEMORRHAGE	CARDIOVASCULAR	
1137567663522	ae_85	MILD	FEMALE	-4.0	-1.0	BRUISE	HEMIC AND LYMPHATIC	3
1181140285973	ae_370	MODERATE	MALE	79.0	84.0	SYNCOPE	CARDIOVASCULAR	5
1261476779849	ae_20	MILD	MALE	31.0	87.0	ANEMIA	HEMIC AND LYMPHATIC	56
1310036402417	ae_34	MODERATE	MALE	30.0	37.0	AORTIC STENOSIS CARDIOVASCULAR		7
1396536741695	396536741695 ae_197 MILD		MALE	78.0	78.0	HEMORRHAGE CARDIOVASCULAR		0
1407862393203	07862393203 ae_195 MILD		FEMALE	1.0	8.0	HEMORRHAGE	CARDIOVASCULAR	7
1447855747874	ae_222	MODERATE	FEMALE	54.0	54.0	HYPERTENSION	CARDIOVASCULAR	0
145322219436	ae_321	MILD	FEMALE	3.0	51.0	POSTURAL HYPOTENSION	CARDIOVASCULAR	48
145322219436	ae_387	MODERATE	FEMALE	30.0	30.0	VASODILATATION	CARDIOVASCULAR	0
1563660755850	ae_60	MILD	FEMALE	5.0	52.0	ATRIAL FIBRILLATION	CARDIOVASCULAR	47
1666330047838	ae_269	MODERATE	MALE	77.0	83.0	MIGRAINE CARDIOVASCULA		6
1672182765314	ae_374	MILD	MALE	6.0	27.0	TACHYCARDIA	CARDIOVASCULAR	21
1677140190516	ae_193	SEVERE	FEMALE	43.0	43.0	HEART BLOCK	CARDIOVASCULAR	0
1791376707628	ae_223	MILD	FEMALE	42.0	47.0	HYPERTENSION	CARDIOVASCULAR	5
1807777224415	ae_289	MILD	FEMALE	48.0	50.0	PALPITATION	CARDIOVASCULAR	2
1855510751277	ae_372	MILD	FEMALE	83.0	83.0	TACHYCARDIA	CARDIOVASCULAR	0
1948387669736	ae_192	MODERATE	FEMALE	82.0	87.0	HEART BLOCK	CARDIOVASCULAR	5
2011401646386	ae_65	MODERATE	FEMALE	3.0	3.0	AV BLOCK FIRST DEGREE	CARDIOVASCULAR	0

37 MILD	
29 MODERATE	
8 SEVERE	
PREFTEXT	
6 ANEMIA	
3 ANGINA PECTORIS	
7 AORTIC STENOSIS	
2 ARRHYTHMIA	
ATRIAL FIBRILLATION	
S AV BLOCK FIRST DEGREE	
BODYTEXT	2
57 BODY AS A WHOLE	
56 CARDIOVASCULAR	\checkmark
75 DIGESTIVE	
18 HEMIC AND LYMPHATIC	\checkmark
59 METABOLIC AND NUTRITIONAL	
* sext	
38 FEMALE	
DC MALE	



Exhibit Source

http://simile.mit.edu/exhibit

<html> <head>

<title>Study Demography View</title>

```
<link href="demog_trt_ae.js" type="application/json" rel="exhibit/data" />
<script src="http://static.simile.mit.edu/exhibit/api-2.0/exhibit-api.js?views=timeline"</pre>
```

type="text/javascript"></script>

</head>

<body> <h1>Study Demography View</h1>

```
<h3><a href="treat_vu.html">Treatment View</a> </h3>
<h3><a href="ae_vu.html">Adverse Event View</a> </h3>
```

```
<div ex:role="collection" ex:itemTypes="Person"></div>
             <div ex:role="exhibit-viewPanel">
                 <div ex:role="exhibit-view"</pre>
                     ex:viewClass="Exhibit.TabularView"
                     ex:label="Demography"
                     ex:columns=".label, .AGE, .sext, .wt, .ht, .RACE, .smkclasst, .unitswine, !pid"
                     ex:columnLabels="Subject, Age, Sex, Weight, Height, Race, Smoker, Wine, Adverse Events"
                     ex:columnFormats="list, list, list, list, list, list, list, list, item
                     { title: expression(.uri) }"
                     ex:sortColumn="1"
                     ex:possibleOrders=".RACE"
                     ex:sortAscending="true"
                     ex:rowStyler="rowStyler" >
                 </div>
              <div ex:role="view"
                  ex:viewClass="Exhibit.ScatterPlotView"
                  ex:label="Race Age-Weight"
                  ex:x=".AGE"
                  ex:y=".wt"
                  ex:xLabel="Aae"
```

Tox Commons

- Proposed Open Re-Use of Failed Compounds
- Common Effort by Pharmaceuticals
- Putting Together Pieces of the *Puzzle*
- No real IP in Failed Clinical Data
- Part of Science Commons Initiative
- Is a Drug Safety Commons Possible?, Bio-ITWorld

Other Activities...





A Genome – Phenome Integrated Approach for Mining Disease-Causal Genes using Semantic Web

Gudivada Ranga Chandra Email : gudx6u@cchmc.org Department of Biomedical Engineering/University of Cincinnati Division of Biomedical Informatics/ Cincinnati Children's Hospital Medical Center

Data Aggregation



Phenomic Data



Mapped each entity to UMLS CONCEPT using SKR - metamap

INTEGRATION

Integration has been done using Semantic Web Standards & Technologies such as RDF and OWL



Prioritize Resources On the network

57





Reference : Kleinberg, J. M. 1999. Authoritative sources in a hyperlinked environment. J. ACM 46, 5 (Sep. 1999)

Bhuvan Bamba, Sougata Mukherjea: Utilizing Resource Importance for Ranking Semantic Web Query Results. SWDB 2004: 185-198

Resource Importance



Total Aggregate is Complex...



ADVANTAGES OF USING SEMANTIC WEB TECHNOLOGIES

Flexible Integration of Genome to Phenome Networks



Adding Context through Semantic Weights

By incorporating context specific subjectivity (SW) and objectivity weights (OW), we were able to improve ranking of certain genes



SPARQL Querying capabilities

Provides flexible querying capabilities to query the RDF graphs in assigning and finding importance

'CONSTRUCT' Querying

>"CONSTRUCT" type of query is used in constructing sub - graph associated with matching graph template

>Used in calculating the importance and rank each query result

'DESCRIBE' Querying

> 'DESCRIBE' type returns a single large RDF graph pooling all the information associated with the a resource

>Can be used to investigate 'why Gene A is ranked higher' or 'why Gene B is ranked lower?'



Data Aggregation from Sources



Ranked Results

PREFIX CCHMC:<http://www.cchmc.com/Bio_RDF.owl#> PREFIX rdf:<http://www.w3.org/1999/02/22-rdf-syntax-ns#> SELECT DISTINCT ?pathway where { ?pathway rdf:type CCHMC:KEGG_Pathway . ?Resource ?PROPERTY ?pathway . }

Rank	Pathw ay	Score
1	Agrin in Postsynaptic Differentiation	0.35737
2	Actions of Nitric Oxide in the Heart	0.27969
3	Stress Induction of HSP Regulation	0.18511
4	Integrin Signaling Pathway	0.185
5	uCalpain and friends in Cell spread	0.185
6	How Progesterone Initiates the Oocyte Maturation	0.1844
7	Signaling of Hepatocyte Grow th Factor Receptor	0.15668
8	Y branching of actin filaments	0.15668
9	How does salmonella hijack a cell	0.15668
10	NFAT and Hypertrophy of the heart	0.15668

the neurocommons project



"open source knowledge management": text-mined annotations on literature data integration via RDF and OWL open source analytics platform open curation system for annotations

Neuro Commons

neurocommons text mining pilot: PubMed abstracts @ 16,000,000 > CNS classified abstracts @ 874,727 > text mining recognized @ 368,688 > text mining processed @ 94,381

extracted graph of 30,000+ relationships and 5,500 genes and proteins

Science Commons

- Publishing and Protecting Knowledge
- Copyright Jujitsu

Allow commercial uses of your work? (more info ⊐)
() Yes
(•) No
Allow modifications of your work? (more info 💷)
○ Yes
🕑 Yes, as long as others share alike (more info 🛄)
O No
Jurisdiction of your license (more info 🛄)
United States
Tell us the format of your work:
Audio
Select a License



This work is licensed under a Creative Commons Attribution-Noncommercial-Share Alike 3.0 United States License.



BioSurveillance using RSS/RDF



EPISPIDER HOME PROMED MAPPER SITE STATS BLOG





CHOLERA DIARRHEA AND DYSEN

This map shows a summary of the las ProMED Mail and related sources. View filtered maps in <u>PROMED MAPPER</u>. Not overlap for a few countries.

EpiSPIDER News

EpiSPIDER New Layout, New Feature Linking ProMED Reports to News Sou More AJAX in EpiSPIDER New Year Greetings to ALL! Treemaps for ProMED Mail Topics Timeline browser in testing TreeMap Visualization in EpiSPIDER! Bayesian classifier operational IP-to-Location Mapping Overhauled EpiSPIDER Implements XML-RPC EpiSPIDER Blog Launched EpiSPIDER gets poster award EpiSPIDER Developer List Bayesian Classifier in EpiSPIDER Google Earth Feeds!

SWAN

Semantic Web Applications in Neuromedicine



70

MagLev Explosing Current DBs onto SW

900	C	Dashboard: p	project					
 + One http://localhost:3000/dashboard/project, 	/4?class=Compound&	focus_id=5			🜍 • 🔍 Google			
Signed in as Eric HOME SIGN-OU	T ADMIN H	IELP		Search	Annotate			
my Projects	Project: GSK	3b Diabet	tes			0 E		
D30-DA	Overvie	w	Timeline	Compounds	Pathways	Datasets		
P 30-RDA	Compounds	mol weight	formula	toxicity				
005.5D	Kenpaullone	328.163	C15H10ON3Br					
PDE-ED	SB216763	371.217	C19H12O2N2Cl2					
Diabetes-PPAR	Kenpaullo Tags: Kinase anta	ne 🝕						
	e e e http://localhost:3000/urispace/target/6							
Started:Wed Mar 07 10:36:00 EST 2007 Lead:Eric Neumann Connections for Kenpaullone + Hypotheses (1)	<pre><?xml version <rdf:RDF xmln www.myOrg.org</td><th>="1.0" end s:rdf="htt /2006/drug t rdf:abou ame>GSK3be escription icated as , and poss ssible dru ntrez_gene md=Retriev asAnnotati asAnnotati et></th><td>coding="iso-88 tp://www.w3.or g_semantics#"> ut="http://loc eta h>Glycogen syn a key player sibly depressi ug target is o s>http://www.n ye&dopt=fu ion rdf:resour ion rdf:resour</td><td><pre>59-1"?> g/1999/02/22-rdf- alhost:3000/urisp thase kinase 3 be in several diseas on. Based on thes bviously quite si cbi.nlm.nih.gov/e cbi.report&lis cc="http://localh cc="http://localh cc="http://localh</pre></td><td><pre>syntax-ns#" xmlns ace/target/6"> ta (GSK3Ã es: diabetes type e relations, its) gnificant.ntrez/query.fcgi? t_uids=1432ost:3000/urispace ost:3000/urispace ost:3000/urispace</pre></td><td><pre>:dd="http:// 81;Â² 2, Alzheimer\'s potential market scription> ntrez_gene> /annotation/28", /annotation/58", /annotation/60",</pre></td></pre>	="1.0" end s:rdf="htt /2006/drug t rdf:abou ame>GSK3be escription icated as , and poss ssible dru ntrez_gene md=Retriev asAnnotati asAnnotati et>	coding="iso-88 tp://www.w3.or g_semantics#"> ut="http://loc eta h>Glycogen syn a key player sibly depressi ug target is o s>http://www.n ye&dopt=fu ion rdf:resour ion rdf:resour	<pre>59-1"?> g/1999/02/22-rdf- alhost:3000/urisp thase kinase 3 be in several diseas on. Based on thes bviously quite si cbi.nlm.nih.gov/e cbi.report&lis cc="http://localh cc="http://localh cc="http://localh</pre>	<pre>syntax-ns#" xmlns ace/target/6"> ta (GSK3Ã es: diabetes type e relations, its) gnificant.ntrez/query.fcgi? t_uids=1432ost:3000/urispace ost:3000/urispace ost:3000/urispace</pre>	<pre>:dd="http:// 81;Â² 2, Alzheimer\'s potential market scription> ntrez_gene> /annotation/28", /annotation/58", /annotation/60",</pre>		

How to interact with the Semantic Web


Thank You

Joining HCLSIG-W3Chttp://www.w3.org/2001/sw/hcls/