

THE URI NOTE

Requirements on the proposed URI Note

Please don't add items below "In scope" unless approved ahead of time by the BioRDF subgroup and the document editor.

For the editor's elaboration on these requirements and examples of solutions that have been suggested, please see [..../UriNoteRequirements](#). See also work plan, [..../RequirementsTalk](#).

In scope

Here are the areas the document needs to cover:

- For our own resources, what URIs to mint and what contracts to adhere to regarding well-definedness and documentation
- What particular URI's to use for resources related to public databases (esp. database records) (>4 proposals on table)
 - What entity is responsible for choosing and maintaining these URIs
- How to get stuff
 - How to use a URI to get metadata (RDF) about an identified resource
 - How to use a URI to retrieve the bits of an information resource

Here are some other important areas that the document needs to discuss (but not necessarily solve):

- How to deal with unusual situations while getting stuff (see above)
 - broken links
 - non-locator URIs (e.g. urn:, info:)
 - locally cached information
- Versioning

Basically what I get from your document is you are saying, "let's use defined terms, and let's be clear about saying what they mean, and then don't lose or change their definition."

Tony Loeser

[Cartoon from *The New Yorker*, May 28, 2007, page 61
by Frank Cotham

http://www.cartoonbank.com/product_details.asp?mscssid=LRH6CFRB36UX9KPHL09X5GKV62UUC3K6&site_type=1&did=4&sid=123995&pid=&keyword=simple+declarative§ion=all&title=undefined&whichpage=1&sortBy=popular]

Problem addressed here: How not to get confused.

(Especially, how not to confuse a computer.)

Outline: discourse; clear definition choice; hygiene; access; versioning

See spot run.

h:necessity24 a h:Necessity;

h:subject h:HCLS;

h:verb h:see;

h:object h:event23.

h:event23 a h:Event;

h:subject h:spot;

h:action h:run.

DISCOURSE.

RDF is a language of discourse, just as a natural language is.

It is not a data format, any more than Vietnamese is.

You can say it, hear it, believe it, be skeptical of it, argue with it.

We want RDF to be a language of scientific discourse.

THE LANCET Neurology

Volume 6, Issue 7, July 2007, Pages 611-619

Abstract

 Add to my Quick Links  Cited By  E-mail Article  Save as Citation Alert  Export Citation  Cite

doi:10.1016/S1474-4422(07)70170-9  Cite or Link Using DOI

Copyright © 2007 Elsevier Ltd All rights reserved.

Review

Silent brain infarcts: a systematic review

Dr Sarah E Vermeer MD^a, , , William T Longstreth Jr MD^b and Peter J Koudstaal PhD^a

^aDepartment of Neurology, Erasmus Medical Center, Rotterdam, The Netherlands

^bDepartments of Neurology and Epidemiology, University of Washington, Seattle, USA

Available online 18 June 2007.

Summary

As the availability and quality of imaging techniques improve, doctors are identifying more patients with no history of transient ischaemic attack or stroke in whom imaging shows brain infarcts. Until recently, little was known about the relevance of these lesions. In this systematic review, we give an overview of the frequency, causes, and consequences of MRI-defined silent brain infarcts, which are detected in 20% of healthy elderly people and up to 50% of patients in selected series. Most infarcts are lacunes, of which hypertensive small-vessel disease is thought to be the main cause. Although silent infarcts, by definition, lack clinically overt stroke-like symptoms, they are associated with subtle deficits in physical and cognitive function that commonly go unnoticed. Moreover, the presence of silent infarcts more than doubles the risk of subsequent stroke and dementia. Future studies will have to show whether screening and treating high-risk patients can effectively reduce the risk of further infarcts, stroke, and dementia.

example illustrating scientific discourse. what about it?

1. author loses interest quickly (moves to next project/job/career)
2. [publishers have priced themselves out of the market for sub-journal work?]
3. community has interest in preservation (e.g. libraries)
4. work is speculative: value to society is usually only *potential*
5. distance (time and space) between writer and reader -- little shared context

Concept: Transient Ischemic Attack

CUI: [C0007787](#)

Semantic Type: [Disease or Syndrome](#)

Definition:

Brief reversible episodes of focal, nonconvulsive ischemic dysfunction of the brain having a duration of less than 24 hours, and usually less than one hour, caused by transient thrombotic or embolic blood vessel occlusion or stenosis. Events may be classified by arterial distribution, temporal pattern, or etiology (e.g., embolic vs. thrombotic). (From Adams et al., Principles of Neurology, 6th ed, pp814-6) ([MeSH](#))

recurring, transient episodes of neurologic dysfunction caused by cerebral ischemia; onset is usually sudden, often when the patient is active; the attack may last a few seconds to several hours; neurologic symptoms depend on the artery involved. ([CRISP Thesaurus](#))

A brief attack (from a few minutes to an hour) of cerebral dysfunction of vascular origin, with no persistent neurological deficit. ([NCI Thesaurus](#))

A brief attack (from a few minutes to an hour) of cerebral dysfunction of vascular origin, with no persistent neurological deficit. ([NCI Thesaurus](#))

Korta, övergående episoder av lokal, icke-konvulsiv, ischemisk funktionsstörning i hjärnan med en varaktighet av mindre än 24 timmar, och ofta mindre än en timme, orsakade av kortvarig trombos eller emboli. Ischemin kan klassificeras utifrån distribution i kärlen, tidsmönster eller etiologi (t ex embolisk eller trombotisk art). ([MeSH Swedish](#))

(this is from UMLS KS)

CLARITY. 1. DIFFERENT DEFINITION, DIFFERENT URI

Language rests on words.

Words must be understood; truth = $M(v)(M(s),M(o))$

Common understanding is a PRECONDITION for useful discussion

1: [Med Sci Monit.](#) 2007 Mar;13(3):RA50-3.

A reappraisal of the definition and pathophysiology of the transient ischemic attack.

[**Hadjiev DI, Mineva PP.**](#)

Medical University, Sofia, Bulgaria. dimiter_hadjiev@hotmail.com

Recent data on the pathophysiology of brain ischemia obtained by neuroimaging methods and the new concept of transient ischemic attack (TIA) emergency have called for a redefinition of TIA. According to the new definition proposed by the TIA Working Group, TIA is a brief episode of neurological dysfunction caused by focal brain or retinal ischemia with clinical symptoms typically lasting less than one hour and without evidence of acute brain infarction. This new definition leads to a discussion on the duration of the neurological dysfunction and the availability of appropriate neuroimaging for all patients. It has been reported that the diffusion-weighted imaging abnormalities could be seen in TIA patients with durations of the neurological symptoms of less than 30 minutes, but they were not detected in 29% of patients with transient deficit lasting as much as 6 to 24 hours. Persisting perfusion abnormalities in TIA patients are also observed. Therefore, a cutoff period of any duration of TIA is inaccurate. From the pathophysiological viewpoint, TIA may be considered an ischemic penumbra of varied duration, which could proceed to cerebral infarction or reduce to benign oligemia. TIA, characterized as an ischemic penumbra, presents an ideal target for rapid reperfusion and neuroprotection. Follow-up perfusion imaging can guide and individualize its treatment.

PMID: 17325647 [PubMed - indexed for MEDLINE]

PEOPLE LIKE TO CHANGE DEFINITIONS.

two definitions of TIA

1. before this article
2. after this article

What if NCI changed its mind around definition of nci:TIA ? CHAOS.

Better to do nci:TIA and nci:TIA.2



TIA (in the sense of Hadjiev 2007)
<http://purl.org/NET/hadjiev2007#TIA>

TIA (in the sense of 2007 MeSH)
<http://purl.org/commons/mesh/2007//TIA>

Thanks to victorianweb.org for scanning this public domain image.

Thursday, November 15, 2007

9

BUT WHEN THEY USE A LABILE TERM, THEY CITE.

URI = bundle {term + reference to source possessing definition}

Different things

- The temperature of a patient (not an information resource)
- A instrument that measures and reports temperature (not an information resource)
- The record retrieved when you query the instrument (an information resource)
- The record that you retrieved at a certain time and you copied and saved (an information resource)

but related



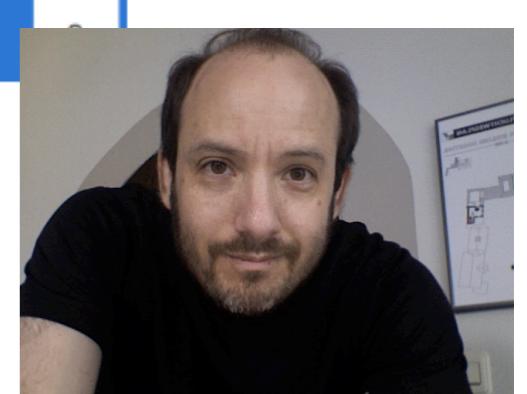
x has_atomic_mass NNN.
x last_modified_on DDD.
x a ???.

[Need example of ontological non-hygiene – one URI used in 2 ways]

Talk about Uniprot experience. Settled on x = record.
has_atomic_mass says that *the subject protein of the record* has that property.

What goes wrong with URLs (courtesy public-semweb-lifesci)

- 1 • The server disappears
- 2 • The content disappears - 404
- 3 • The content might change and you want to know and communicate what it used to be
- 4 • Access to the content is too slow
- 5 • Access to the content is too public
- 6 • The content is very big **denotes**
- 7 • You don't know if a URI ~~is~~ an information resource or not
- 8 • You want to record and access metadata - information about some information resource - and you don't know where to get it.
- 9 • You don't know what format an information resource is encoded in.

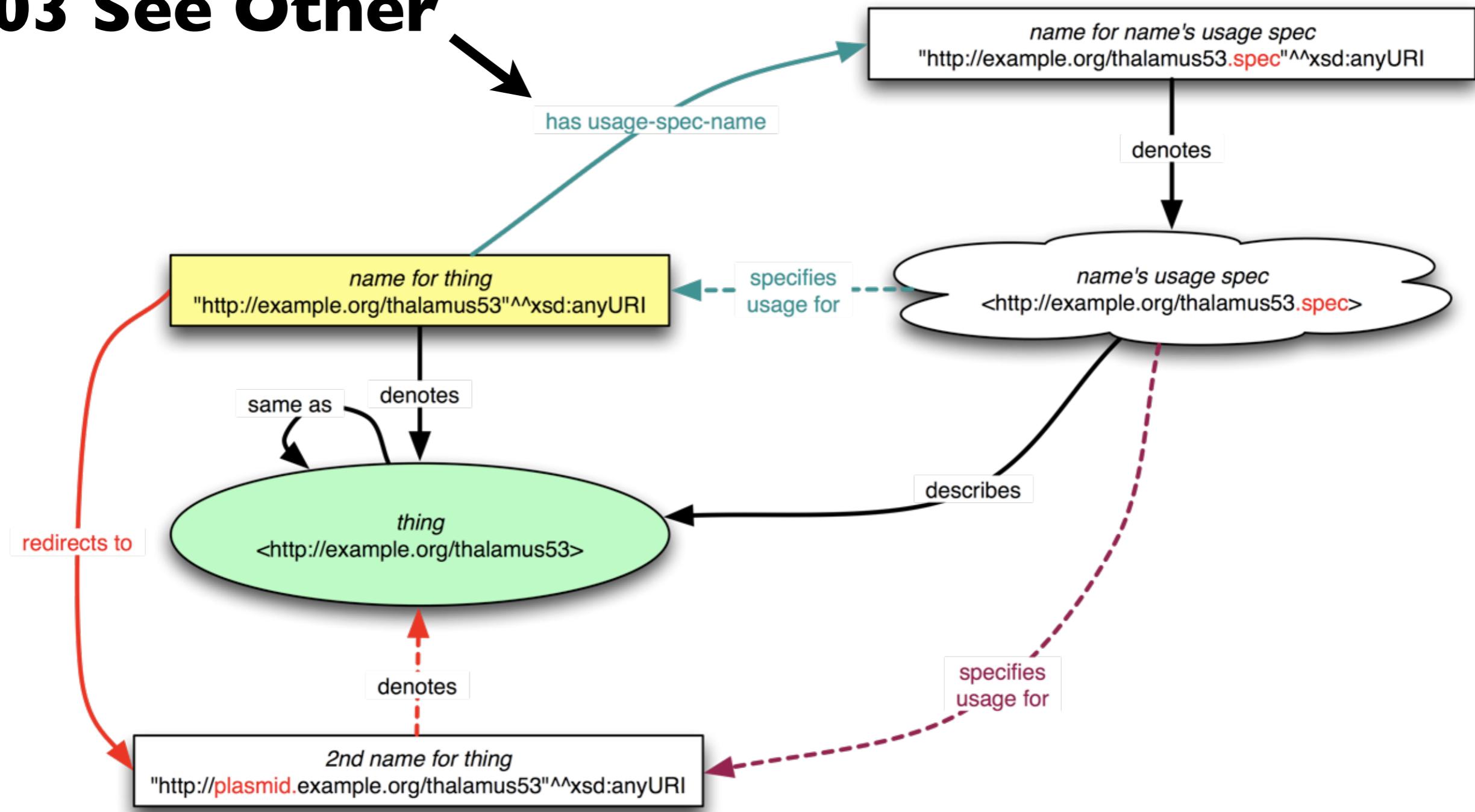


3. ACCESS TO STUFF.

(Access to clarity.)
(Clarity about access.)

Alternative Resolution

303 See Other



ACCESS:

If you can say where something is, you are enabled to control access method.
server gone, 404, private mirror, non-HTTP

Note: the HCLS/NC KB doesn't depend on resolution.

1

n:hadjiev2007 a foaf:Document.

2

n:hadjiev2007 a hcls:StableDocument.

n:hadjiev2007 a hcls:FixedDocument.

3

n:hadjiev2007 **hcls:isDenotedBy**

“<http://hcls.org/local-copies/hadjiev2007>”^^xsd:anyURI.

4

<<http://hcls.org/usagespecs/hadjiev2007.usage>>

hcls:specifiesUsageFor

“<http://hcls.org/local-copies/hadjiev2007>”^^xsd:anyURI.

5

<<http://hcls.org/about/hadjiev2007.cliffnotes>>

hcls:describes n:hadjiev2007.

```
@prefix u: <urn:lsid:ubio.org:classificationbank:>
@prefix m: <http://lsid.tdwg.org/urn:lsid:ubio.org:classificationbank:>
```

1 u: 1164063 a foaf:Document.

2 u: 1164063 **hcls:isDenotedBy**

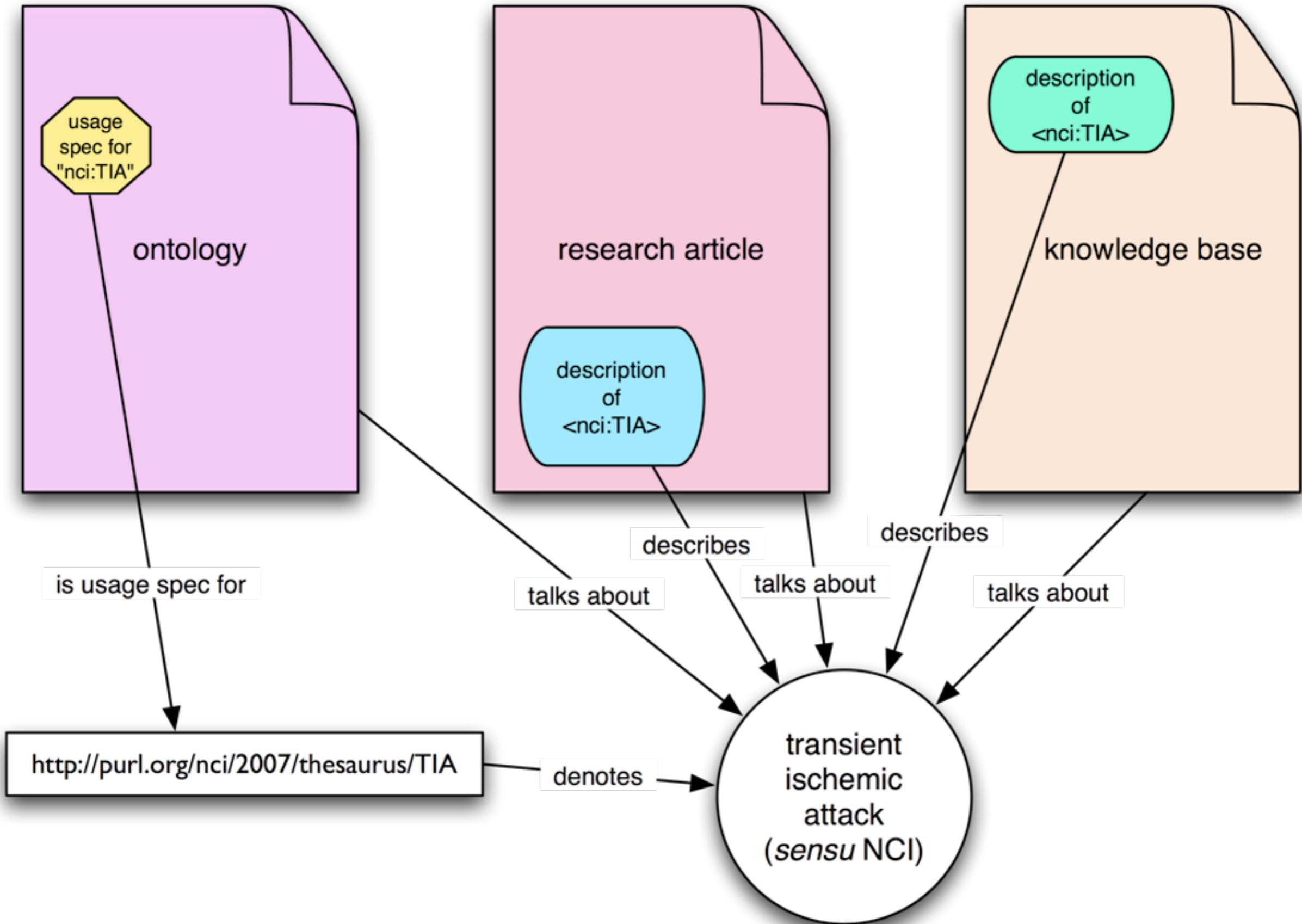
“http://lsid.tdwg.org/data/urn:lsid:ubio.org:classificationbank: 1164063”^^xsd:anyURI.

3 m: 1164063 **hcls:specifiesUsageFor**

“urn:lsid:ubio.org:classificationbank: 1164063”^^xsd:anyURI.

- ✗ a RESTEndpoint.
- ✗ a Document.
- ✗ a StableDocument.
- ✗ a FixedDocument.

x a Document.
x a VersionableDocument.
x.1 is_a_version_of x.
x.2 is_a_version_of x.
x.1 a StableDocument.
x.2 a StableDocument.



Thursday, November 15, 2007

18

OK, look at relation of usage specs / descriptions.

All the domain-level statements relating docs to TIA factor through the URI: the URI 'occurs in' the document, and 'denotes' TIA.

Dan C: Public Subject Indicator

Requirements on the proposed URI Note

Please don't add items below "In scope" unless approved ahead of time by the BioRDF subgroup and the document editor.

For the editor's elaboration on these requirements and examples of solutions that have been suggested, please see [..../UriNoteRequirements](#). See also work plan, [..../RequirementsTalk](#).

In scope

Here are the areas the document needs to cover:

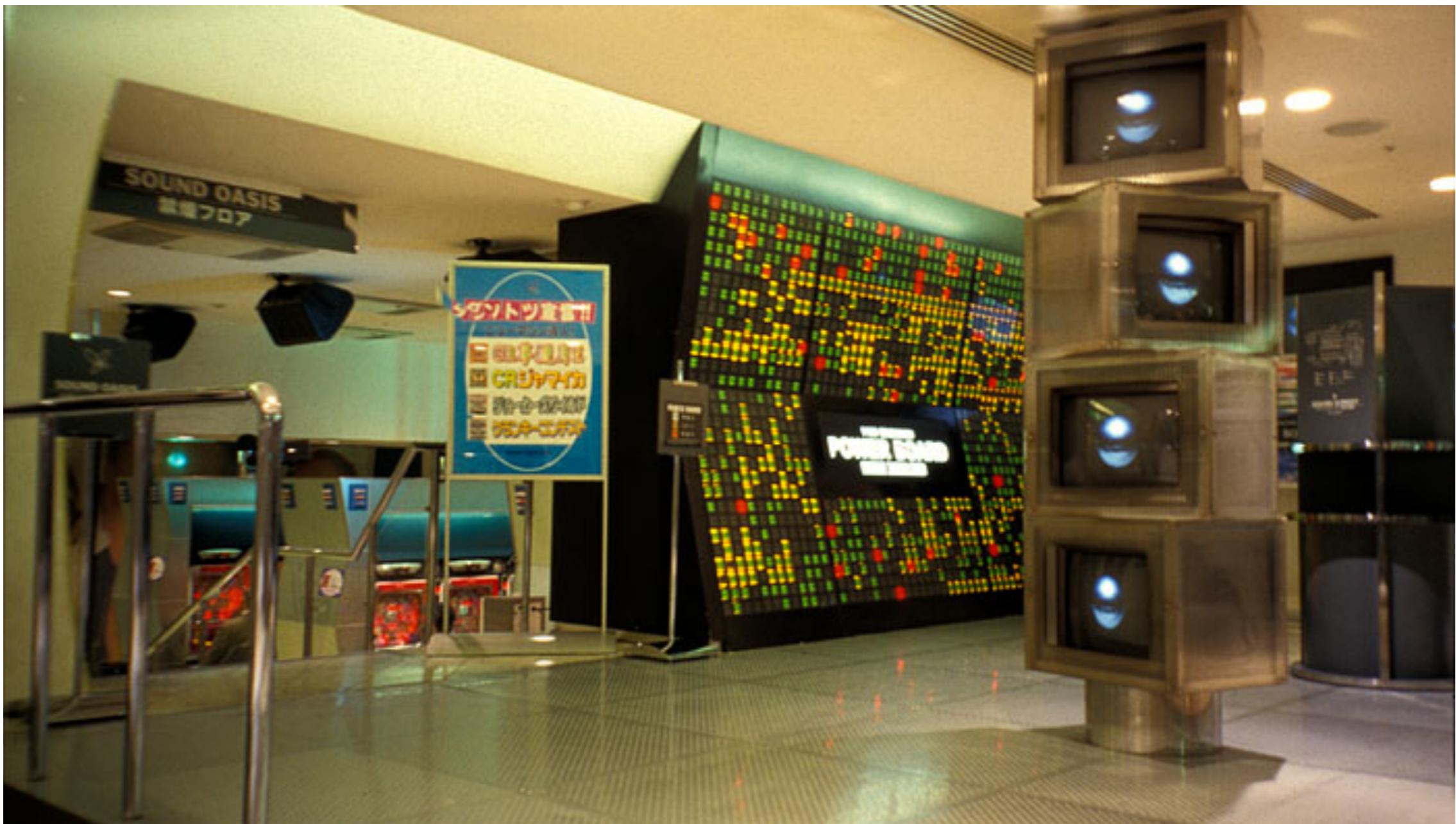
- 1 • For our own resources, what URIs to mint and what contracts to adhere to regarding well-definedness and documentation
- 2 • What particular URI's to use for resources related to public databases (esp. database records) (>4 proposals on table)
 - What entity is responsible for choosing and maintaining these URIs
- 3a • How to get stuff
 - How to use a URI to get metadata (RDF) about an identified resource
 - How to use a URI to retrieve the bits of an information resource
- 3b

Here are some other important areas that the document needs to discuss (but not necessarily solve):

- 4 • How to deal with unusual situations while getting stuff (see above)
 - broken links
 - non-locator URIs (e.g. urn:, info:)
 - locally cached information
- 5 • Versioning

LET'S SEE HOW WELL THE URI NOTE MEETS THESE REQUIREMENTS:

- 1) http: URIs that have a fighting chance of persisting.
purl.org is best bet. Use NET or get your own domain.
- 2) successor to <http://purl.org/commons/> & bio2rdf.org -- needs an organization
- 3a) #/303 (US) + seeAlso (non-US) 3b) 200 or 303 to sameAs
- 4) resolution ontology. -- for purls, we'll need community administration
- 5) simple ontology.



(c) 2007 flickr.com user Wanderungen /
CC attribution-nc 2.0 license

Thursday, November 15, 2007

20

FUTURE WORK:

Ontologies: resolution, documents, versions
Bibliographic reference ontology
Web dereference (e.g. HTTP) semantics
Public resource URIs
Registry of thing -> set of documents about thing
URI resolver as web service (replicable)
(retro futurism, Tokyo department store)

Major remaining trouble spots

1. [/DenoteVsDereference](#) - how to explain the relationship between denotation and dereference
2. [/RacineSharing](#) - tolerate or discourage?
3. [/DeclarationDelineation](#) - how strongly to encourage?
4. [/AttitudeTowardMigration](#) - tolerate ephemeral locations (deal via resolution rules) or discourage (SHOULD use only highly stable URIs such as purls)?
5. [/Purls](#) - how strongly can we recommend purl.org?
6. [/PublicResources](#) - exactly what URIs to use for NCBI records, journal articles, etc?
7. [/AttitudeTowardNonlocators](#) - ok to use info:, urn:lsid:, tag: ?

1. don't
2. discourage
3. very strongly ?
4. insurance policy idea (Alan)
5. TBD
6. public - establish new entity
7. avoid non-http, use sameAs or redirection rules

Your trail:

[View](#) [Attach](#) [Info](#)

Bioinformatics existing creeps

Here is a sample of all kind of different URI syntax used in the bioinformatics web.

- Table of Contents

[Bioinformatics existing creeps](#)

[Existing syntax for URL](#)

[All different URI for uniprot:p26838](#)

[Official namespace lists](#)

[Categories](#)

Existing syntax for URL

- PDB use DOI
 - <http://www.rcsb.org/pdb/explore.do?structureId=2b3y>
 - <http://dx.doi.org/10.2210/pdb2b3y/pdb>
- UniProt at SwissProt, Pir and EBI
 - <http://www.expasy.org/enzyme/4.2.1.11>
 - <http://www.expasy.org/uniprot/Q8NKC2>
 - http://www.pir.uniprot.org/cgi-bin/upEntry?id=ENO12_SCHPO
 - http://www.ebi.uniprot.org/uniprot-srv/uniProtView.do?proteinId=ENO12_SCHPO
- NCBI
 - http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=pubmed&dopt=full_report&list_uids=1234567890

Id Reinvention

- Global Identity naming mechanism for data objects in the Life Sciences

`urn:lsid:uniprot.org:{db}:{id}`
`http://purl.uniprot.org/{db}/{id}`

- LSIDs and URIs and PURLs. WS-Naming and all its friends
- Half the debaters haven't actually read the LSID or URL or PURL specs. Or provided use cases.
- Web Pages are not Data Assets.
- "you could do this with HTTP based identifiers given <insert hack>".
- The debate rages! 124 messages in the last week.
- W3C Semantic Web Health Care and Life Sciences Interest Group public-semweb-lifesci@w3.org

--- Carole Goble, The Seven Deadly Sins of Bioinformatics

not http	info:/pubmed/12345
no RDF yet	http://view.ncbi.nlm.nih.gov/pubmed/12345
provisional	http://purl.org/commons/record/pmid/12345
ontology?	http://bio2rdf.org/pmid:12345
new org	http://hclshcls.org/pmid/record/12345
new org	http://uriresolver.org/refer?uri=info:/pubmed/12345

Identifying pubmed records (NOT articles).
 Entrez won't serve RDF until after we do, won't do 303s, etc

PMID = a LSRN

From: timbl@w3.org
Subject: Subgroup to handle semantics of HTTP etc?
Date: October 15, 2007 4:07:21 PM EDT
To: www-tag@w3.org
Cc: alanruttenberg@gmail.com, jar@mumble.net, connolly@w3.org
Archived-At: <http://www.w3.org/2007/10/15/semantics-of-http-etc.html>

At the Cambridge Semantic Web Gathering a few days ago I was was chatting with Jonathan Rees and Alan Ruttenberg from Science commons about basically web architecture from the semantic web point of view. Alan felt an urgent need for much more concrete basis for this than he could get by trying to red the current AWWW with semantic web-colored glasses. He, as , would really like to have an ontology for the things the AWWW document talks about, and a formal definition of the semantics of things like HTTP fetches, hash, etc.

At the same time, Anne van Kesteren has been suggesting that the HTTP spec doesn't have a very clear semantics. He asks, for example, what happens if a server sends two different content-type headers, for example? There are no HTTP validators, ad the significance of it is not obvious. This problem could also be helped partially by some semantics expressed more explicitly.

Two question Alan asked recently (on the list and offline) were

- "How can one ever show that a web site is behaving contrary to the web architecture?" and
- "How do i know what triples an RDF system is able to draw from an HTTP interaction?"

Both good questions.

The answer to the first question could be to draw all the triples from the HTTP transactions and the documents published, and then check for OWL inconsistencies. Which begs the second question.

This is relevant to the Tabulator project, as Tabulator does this, and uses the conclusions from HTTP transactions to (for example) select user interface operations to offer the user, and to generate warning messages about inappropriate behavior.

We wondered whether it would be good idea to put together some kind of a task force under the TAG to propose set of these axioms and an ontology.

Tim



(c) 2007 flickr.com user Mimir.nm /
CC attribution-nc 2.0 license

Thursday, November 15, 2007

26

How can we make the semantic web something one can bank on?
What URI scheme will be seen as a reliable foundation for content production?