

Tools for semantic interoperability : hsubjects

*The hub does not know
what the wheel is about*

This version : 2005-06-29

Latest version : <http://www.mondeca.com/lab/bernard/hsubjects.pdf>

Editor : [Bernard Vatant](#)

Background image : <http://www.bigfoto.com/miscellaneous/photos-11/cart-wheel-photo.jpg>

Of hubs and other hollow things

Thirty spokes share the wheel's hub;
It is the center hole that makes it useful.

Shape clay into a vessel;
It is the space within that makes it useful.

Cut doors and windows for a room;
It is the holes which make it useful.

Therefore profit comes from what is there;
Usefulness from what is not there.

[Lao Tsu, Tao-te-King, Chapter 11](#)

Of subjects - some possible definitions

- Wordnet 2.0
 - The subject matter of a conversation or discussion.

- Dictionary.reference.com
 - Logic: The term of a proposition about which something is affirmed or denied.
 - Philosophy: The essential nature or substance of something as distinguished from its attributes.

- Topic Maps Reference Model
 - Anything whatsoever, regardless of whether it exists or has any other specific characteristics, about which anything whatsoever may be indicated by any means whatsoever.



Terminology disclaimer

- Q: *Why use « subject » rather than any other term?*
 - Why not « resource », « entity », « concept » ?
- A: *This is an arbitrary choice indeed*
 - But you need a word anyway
 - Any other choice would have been the same
- Q: *Is not « subject » too Topic Map biased?*
- A: *Yes, but it was the better choice I had*
 - I 'm used to it in Topic Map land
 - I like the notion of conversation
 - The pun « hsubject » just works well with it



Representations are defined in contexts

- Context provides elements needed to handle representations following their specific type and logical structure
- Context generally includes
 - Framework
 - Language, logical rules, protocols, specific tools ...
 - Example : RDF, Topic Maps, Conceptual Graphs, UML ...
 - Scheme
 - Structured and hopefully consistent set of representations
 - Generally defined in a single framework
 - Example : Subject Headings, Thesaurus, Ontology ...
- Context can also include a variety of specific elements
 - Scope, time and space, version ...



Identification of subjects

- Identity is conferred using representations ...
 - Assignment of identifying properties
 - Identification protocols
- Therefore identification is context-specific ...
 - Identification uses context rules
 - « Same-ness » rules are context-specific
- Therefore subjects have no absolute identity ...
- So ... how can one provide semantic interoperability between different contexts?

Is semantic interoperability possible ?

- Inside the same context : YES
 - Identification, mapping, merging, consistency checking ... are possible inside the same context thanks to ...
 - Shared notion of true/false
 - Shared axioms and inference rules
 - Shared identification process

- Across different contexts : Not exactly ...
 - Using different languages, different rules and tools
 - Identification, mapping, merging, consistency checking is a real challenge
 - Unless a meta-context is provided
 - Or specific rules of mapping are defined



So why not define a single meta-context ?

- Because it 's a silly, endless and childish game
 - A: « My context is more meta than yours.»
 - B: « No, MINE is more meta than yours ! »
 - C: « MINE is more meta than both of you !! »
 - D: « Prove it, but use MY context, please !!! »
- People have played it for centuries
 - With no visible results except religion wars



Semantic diversity is Life

- No representation is exhaustive
 - A representation only provides a limited set of properties which are useful and possible to express inside its context
- All contexts have limited expressivity
 - Bound to their specific language and logical rules
 - Relative to their specific purposes
- Anybody can represent anything anyway ...
 - So we 're bound to live with various representations of the « same » subject
 - But we would like to be able to assess this « same-ness » and use it efficiently - like in natural language and thought



Shall we forget about semantic interoperability?

- In a sense, yes ...
 - Forget about any « meta-framework »
 - The « more meta than you » game is both arrogant and useless ...
 - Leads at best to infinite recursion, at worst to unique thought
 - Forget about one-to-one mediators
 - Their sheer number raises serious scalability issues

- In a sense, not yet ...
 - Subjects are wheels yet to be built
 - But we have a lot of spokes : the representations
 - We just need hubs able to provide simple and robust connections between representations, across contexts



Re-inventing the Wheel

■ Hub

- The central piece of a wheel, binding spokes together
- A center of activity or interest or commerce or transportation; a focal point around which events revolve
- To hub : to connect through a hub



Requirements for a-semantic hubs

- Hubs should be shared and neutral

Thirty spokes share the wheel's hub;

- Hubs should be hollow

It is the center hole that makes it useful.

- Hubs should be agnostic about content and nature of things they connect, how they are produced, managed and used.

*Therefore profit comes from what is there;
Usefulness from what is not there.*



Introducing Hsubjects

■ Hsubject

- A **hub** connecting different representations of a **subject** inside the same or across different contexts
- Think about the subject as a **wheel** of which various representations are **spokes**
 - To complete the metaphor, think of context as the direction of the spoke



Subjects are Hubs

- Subjects are central
 - They bind representations as so many spokes of the subject
- Subjects are hollow
 - They do not provide any more semantics than the representations they are binding
 - No specific type, attribute, property or semantic of any kind

*Their profit comes from what is there
(the representations are there)*

*Usefulness from what is not there
(nothing else is specified)*

Looking for hsubjects

■ Reference.com

<http://www.reference.com/>

- Hubs a Dictionary, a Thesaurus, and Web resources
 - <http://dictionary.reference.com/search?q=hub>
 - <http://thesaurus.reference.com/search?q=hub>
 - <http://reference.com/search?q=hub>
- ... but not exactly hollow
 - The hub is identified by the value of the query string

■ ISBN.nu

<http://isbn.nu>

- Hubs offers for a given book in different data bases
- .. but not exactly hollow
 - The hub is identified by the value of the ISBN number



Not yet exactly hsubjects, but almost ...

■ Google News

<http://news.google.com/>

- Hubs the « same » news from different sources
 - Or more exactly « related news ». Google does not claim sameness, just similarity, with good recall, and low noise. Much redundancy. But that 's OK.
- Clustering rules are not declared.
 - Who cares? Users don't as soon as the results make sense.

■ So, are « Google News » hsubjects? Almost ...

- The news itself is never identified nor defined
 - That 's good.
- But news clusters are categorized ...
 - So much for semantic emptines ... but that 's a less semantic evil
 - It figures : categorization comes before identification



Existing examples are cool, but:

- They handle a specific kind of subjects
 - News, concepts, books ...
- They provide no explicitation of the context
 - Proprietary and ad hoc rules ...
- They are mostly designed for human consumption
 - Aka smart search engines ...
- They provide no formal representation of hsubjects
 - *Wait ... formal representation of hsubjects?*
Is not that the more-meta-than-you game again?



Hubjects are not yet another meta-username

- Hubjects bear no proper semantics
- Hubjects provide neither semantic interpretation of the representations they connect nor absolute indication of the subject

The hub does not know what the wheel is about

- Hubjects have a structure as simple as possible
 - Easy to express in a large variety of frameworks

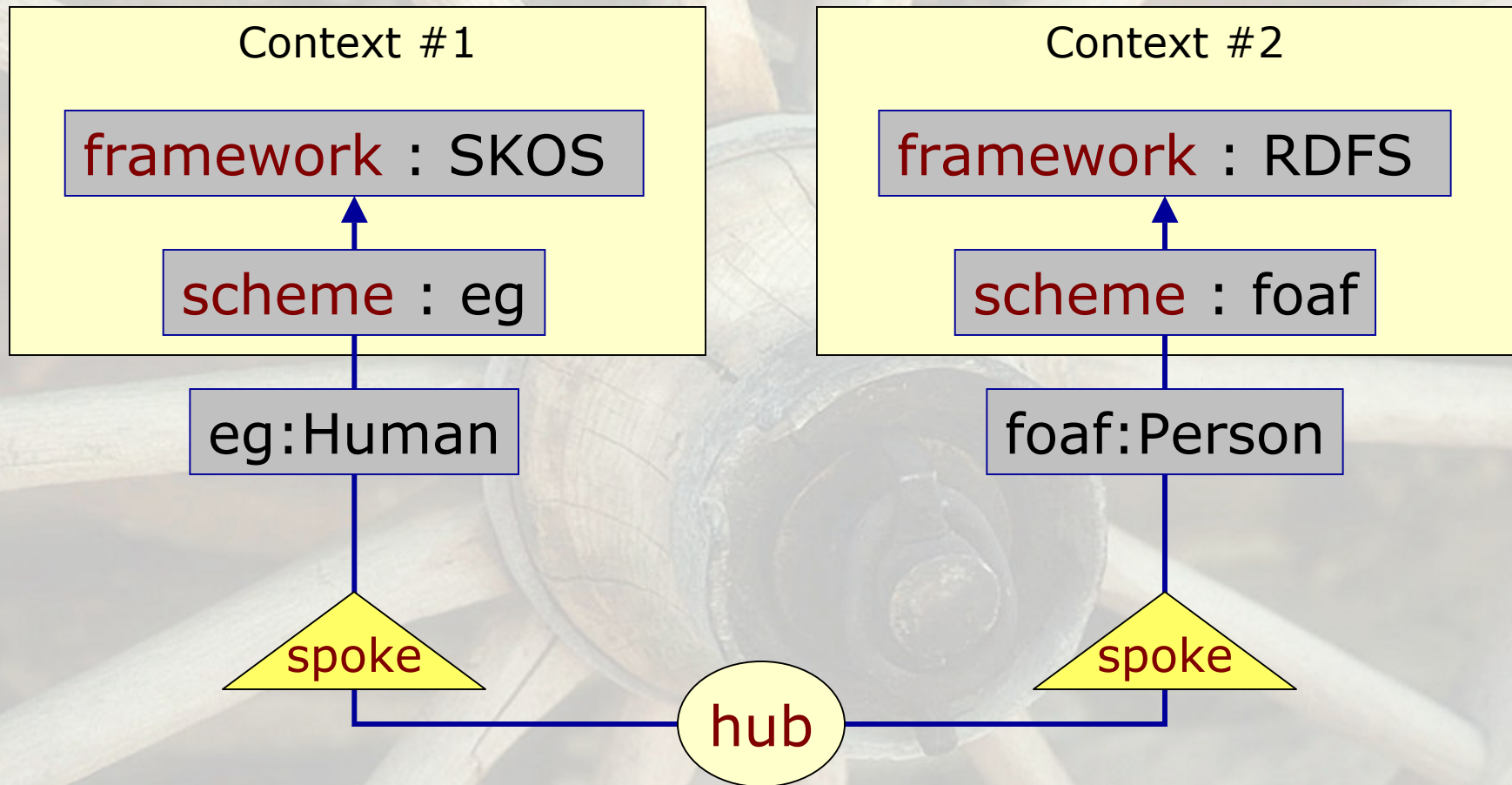


A subject use case from RDF land

- As set by Dan Brickley
 - How do I connect the following ?
 - a Concept defined in a SKOS concept Scheme (Thesaurus)
 - eg:Human
 - a Class in a RDFS or OWL ontology
 - foaf:Person
- An obvious solution seems to use some direct assertion
 - eg:Human a:foo foaf:Person
- But such a solution brings about tricky issues
 - In which context is this assertion made?
 - In the Thesaurus? the ontology? In SKOS? In RDFS? ...
 - In which framework is this assertion to be used?
 - IOW, what is the operational semantics of a:foo?
 - RDFS and SKOS have separately robust semantics
 - But how do they fit together in a more generic RDF framework?



Hubject structure



An XML micro-format

```
<hub>
<spoke>
  <context>
    <framework>http://www.w3.org/2004/02/skos/core</framework>
    <scheme>http://www.example.com/concepts</scheme>
  </context>
  <resource>http://www.example.com/concepts#Human</resource>
</spoke>
<spoke>
  <context>
    <framework>http://www.w3.org/2000/01/rdf-schema</framework>
    <scheme>http://xmlns.com/foaf/0.1/</scheme>
  </context>
  <resource>http://xmlns.com/foaf/0.1/Person</resource>
</spoke>
</hub>
```



An RDF variation



_:x	rdf:type	hub:Hub
_:x	hub:spoke	ex:Human
ex:Human	hub:context	_:a
_:a	hub:framework	"http://www.w3.org/2004/02/skos/core"
_:a	hub:scheme	"http://www.example.com/concepts"
_:x	hub:spoke	foaf:Person
foaf:Person	hub:context	_:b
_:b	hub:framework	"http://www.w3.org/2000/01/rdf-schema"
_:b	hub:scheme	"http://xmlns.com/foaf/0.1"



Q & A

- *Are hubjects resources in the RDF sense?*
 - They should not be, to avoid recursivity trap ...
 - But they can easily be represented as such (see previous slide)

- *Can I assign URIs to hubjects?*
 - You can, if hubs are represented in RDF
 - But you certainly should not (see above)

- *Can hubjects be used as subject indicators?*
 - Not by themselves, since they should not be addressable.
But they can be used to bind together several subject indicators.

- *Can hubjects have different types?*
 - No. A type would be a specific property.
It goes against the no-semantic requirement.

To do list

- Definition of an abstract micro-format for hubjects
 - As framework-independent as possible
 - Common Logic expression?
 - As simple as possible
 - Framework, context and resource seem enough to begin with...
- Expression in various frameworks
 - RDF, XTM, UML, XHTML ...
- Rules of utilisation
 - Publication and Management
- Documenting use cases and scenarios

References

- Topic Maps standards
 - Topic Maps Reference Model
 - <http://www.isotopicmaps.org/tmrm/>

- W3C Semantic Web Best Practices Group
 - Software Engineering Task Force
 - <http://www.w3.org/2001/sw/BestPractices/SE/>

- universimmedia
 - Where subject identification issues are tracked
 - <http://universimmedia.blogspot.com/>



Acknowledgments

- An open list of folks who help this to happen
 - In Semantic Web and Topic Maps community
 - **Dan Brickley**, who provided the example issue
 - **Steve Newcomb**, who pushed me to show we were not in disagreement
 - **Alistair Miles**, who pointed terminology pitfalls
 - **Jack Park**, who is still not convinced all this is making sense
 - **Patrick Durusau, Jan Algermissen**, for the TMRM perspective
 - **Michel Biezunski**, for his faithful interest in whatever I publish
 - **Mikael Nilsson**, who pushed me into more accuracy
 - **Phil Tetlow**, for his input on composite identification
 - In Mondeca
 - **Jean Delahousse**, who opens the bandwidth to follow this track
 - **Florence Amardeilh, Thomas Francart**, for their kind support