
a note on ontologies

giovanni bono – università degli studi di milano – august 20, 2004, INEM, amsterdam

overview (1)

- ontologies are systems of definitions
- providing a context of meaning for statements in their domain
- practical ontological knowledge resides in institutions or is encoded in artifacts
- warning: work in progress

object

- on the encoding of the economists' knowledge – *study of a disciplinary community by looking at its tools*
- in the context of an information processing network – *humans backed by software agents*
- focusing on a special class of textual artifacts – *instances of a generalized markup language; markup is the notation used in (virtually all) written languages for attaching information to a subset of a text*
- without loss of generality – *machine-readable, human-readable, device-independent language*
- or: on encoding as it may be

overview (2)

ontologies

- one subclass of such textual artifacts – *specifying a conceptualization*
- semantic coordination technologies – *analogous to disciplines in a way*
- essential infrastructure of the Semantic Web

diversion: working papers

- produced by another subclass of such textual artifacts – *specifying a document's structure and format*
- typographic/composition technologies – *encoding as it is*
- well integrated in the WWW (World Wide Web)
- printed documents contra conceptual relations as the currency of scientific communication
- noteworthy: debate in the digital typography community on the separation of presentational and structural markup – *in particular the LaTeX and SGML communities*

web (1)

- from the WWW towards the Semantic Web
- original proposal by Tim Berners-Lee in 1989 – *least recently modified web-page dated 1991*

WWW – World Wide Web

- network–spanning information space of interrelated resources – *Hyper Text (in the open)*
- simple architecture based on URIs – *Universal/Uniform Resource Identifier*
- every resource addressable – *agents may use a URI to access the referenced resource; this is called dereferencing the URI*
- highly---yet partially ---decentralized system – *connective tissue is HTML (Hyper Text Markup Language)*

The Web is by design and philosophy a decentralized system, and its vulnerabilities lie wherever a central facility exists. ... The value of a common document language has been so enormous that HTML has gained a dominance on the Web, but it does not play a fundamental key role. Web applications are required to be able to process HTML, as it is the connective tissue of the Web, but it has no special place architecturally.

TBL, Web Architecture from 50,000 feet, Oct 1998

<http://www.w3.org/DesignIssues/Architecture.html>

web (2)

Semantic Web

- information space of self–describing resources – *Knowledge Representation (in the open)*
- simple architecture based on metadata – *the connective tissue would be RDF (Resource Description Framework) and XML (eXtensible Markup Language) described below*
- everything addressable – *URIs can meaningfully reference people, concepts, whatever*
- taking off?

The original idea of the Web being a creative space for people to work together in ("intercreative") seems to be making very slow progress. TBL, Web Architecture from 50,000 feet, Oct 1998

<http://www.w3.org/DesignIssues/Architecture.html>

The Semantic Web ... is an extension of the current web in which information is given well-defined meaning, better enabling computers and people to work in cooperation. Tim Berners-Lee, James Hendler, Ora Lassila, The Semantic Web, Scientific American, May 2001

- look at the W3C ubiquitous 'layers' picture

languages (1)

markup languages

- as noted every modular language is a markup language – or: markup is a pattern in a language
- generalized markup and richness of a language

- scribal and procedural markup

categories of markup

- paralinguistic – e.g. *in speech*
- punctuational – e.g. *.,,: ...!* (mostly scribal)
- presentational – e.g. *page numbering* (mostly procedural)
- descriptive – e.g. *TeX, wikis*
- referential – e.g. *SGML (Standard Generalized Markup Language)*
- meta – e.g. *below*

languages (2)

XML – eXtensible Markup Language

- universal format for structured documents and data – *XML documents are trees*
- XML vocabularies contains names for the markup of elements – *XML namespaces are URIs identifying those names uniquely*

RDF – Resource Description Framework

- express subject–predicate–object statements – *triples*
- a resource (URI) has property (URI) whose value is such and such (URI – literal)

OWL – Ontology Web Language

- family of languages, extensions of RDF – nested expressiveness
- formalize a domain by defining individuals, classes, and properties of those classes – *intended for distributed formalization*
- allow automated reasoning – *first order logic entailment*

languages (3)

N3 serialization

- quick notation for triples – *could go hand in hand with collaborative narrative tools, e.g. wikis*
- see an example

```
@prefix <#>
@prefix method: <http://ontologia.esempio/economicmethodology>
@prefix rdfs: <http://www.w3.org/2000/01/rdf-schema>

:notationOfTechnology rdfs:subClassOf method:ontologicNotion
```

- same example in RDF – *warning: unchecked for conformance*

```
<rdf:RDF
```

```

xmlns      = "http://ontologia.esempio/economics#"
xmlns:econ = "http://ontologia.esempio/economics#"
xml:base   = "http://ontologia.esempio/economics#"
xmlns:method="http://ontologia.esempio/economicmethodology#"
xmlns:owl  = "http://www.w3.org/2002/07/owl#"
xmlns:rdf  = "http://www.w3.org/1999/02/22-rdf-syntax-ns#"
xmlns:rdfs = "http://www.w3.org/2000/01/rdf-schema#"
xmlns:xsd  = "http://www.w3.org/2001/XMLSchema#">
<owl:Ontology rdf:about="">
  <rdfs:comment>An example OWL ontology adapted from http://www.w3.org/TR/owl-guide/<
  <owl:imports rdf:resource="http://ontologia.esempio/economics"/>
  <rdfs:label>Ontology of Economics</rdfs:label>
</owl:Ontology>
</rdf:RDF>
<owl:Class rdf:ID="notionOfTechnology">
  <rdfs:subClassOf rdf:resource="&method;ontologicNotion" />
</owl:Class>

```

exempla

- open archives – *in economics Netec and cognates*
- bibliographic data – *bibster is a p2p ontology-based system*
- digital commons/private right management – *Creative Commons* – <http://creativecommons.org/> - *DRM (Digital Rights Management) languages*
- the Debian GNU operating system – *large distributed maintenance*
- etc.

institutions

- W3C – <http://www.w3.org/> – *leading the Web to its full potential*
- FSF – Free Software Foundation – <http://www.fsf.org/fsf/> – *promoting computer users' right to use, study, copy, modify, and redistribute computer programs*
- Dublin Core Metadata Initiative – <http://dublincore.org/> – *open forum engaged in the development of interoperable online metadata standards that support a broad range of purposes and business models*
- TEI – Text Encoding Initiative – <http://www.tei-c.org/> – *international and interdisciplinary standard that helps libraries, museums, publishers, and individual scholars represent all kinds of literary and linguistic texts for online research and teaching, using an encoding scheme that is maximally expressive and minimally obsolescent*
- OASIS – Organization for the Advancement of Structured Information Standards – <http://www.oasis-open.org/> – *is a not-for-profit, international consortium that drives the development, convergence, and adoption of e-business standards*
- DAML – DARPA (Defense Advanced Research Projects Agency) Agent Mark Up Language - <http://www.daml.org/> – *a language built upon XML that allows users to provide machine-readable semantic annotations for specific communities of interest*
- etc.

intermediate summary and questions

- the WWW had a major impact on the working of disciplines – *in just about 10 years*
- the Semantic Web is expected to double the shock – *something is already at hand*
- Semantic Web technologies are coming of age – *large body of viable standards, early implementation experiments*
- how should the working habits of economists change to leverage the potential of Semantic Web technologies?
- could economics provide positive and/or normative insight on the shape of the Semantic Web technologies to come?

disciplines and ontologies

- the exchange of conceptual links expose in public a larger fraction of the research activity – *to maximize social welfare*
- this requires documentation work – *self-reflective work*
- standardization of conceptual links involve richer interpretation – *mapping between contexts*

... to the best of his ability the author means to put his cards on the table, which is by no means the same as playing the game. TW Adorno, NEGATIVE DIALECTICS, 1966

semantic coordination

- knowledge base of economist queried with natural language sentences – *creative commons approach*
- example – *data base interoperability between interpretation contexts*
- the query would add to the knowledge base – *as in marketing*
- mechanism design problems address semantic coordination – *as in market design*
- reflexivity at the level of the theory is an example of a normative requirement – *if an optimal semantic coordination mechanism is known it is used in the production of economic ideas*
- property rights over digital objects is an example of a field addressed by the literature
- the web is evolving around the confrontation between licensing schemes – *copyleft licences in place before the web*

W3C patent policy

goal of this policy is to assure that Recommendations produced can be implemented on a Royalty-Free (RF) basis. compliance is a condition of participating in a Working Group and of submission, not a condition of W3C membership. if she receives a disclosure request, a W3C member has a good faith disclosure obligation. adapted from: 5 February 2004 (stable, normative version)

<http://www.w3.org/Consortium/Patent-Policy-20040205/>

FSF on the W3C patent policy

The problem comes from the "field of use" restrictions that patent holders are allowed to put in their royalty-free patent licenses. Such restrictions say that you are allowed to practice the patented idea, but only for implementing the standard precisely as specified--not in any other way. Thus, if you change the code to depart from the spec even slightly, the patent license no longer protects you from against being sued for infringing the patent.

<http://www.gnu.org/philosophy/w3c-patent.html>
