Nordic Studies in Pragmatism
Helsinki | 2010

Catherine Legg

"Peirce, Meaning and the Semantic Web"

In: Bergman, M., Paavola, S., Pietarinen, A.-V., & Rydenfelt, H. (Eds.) (2010). *Ideas in Action: Proceedings of the* Applying Peirce *Conference* (pp. 173–188). Nordic Studies in Pragmatism 1. Helsinki: Nordic Pragmatism Network.

ISSN-L 1799-3954
ISSN 1799-3954
ISBN 978-952-67497-0-9

Nordic Pragmatism Network,
Helsinki 2010

www.nordprag.org

Pragmaticism on the Semantic Web

Catherine Legg
University of Waikato

1. Introduction

It was part of Tim Berners-Lee's original vision for the World Wide Web (Berners-Lee, 2002) that it would shortly evolve into a so-called 'Semantic Web', which would (famously) replace a "web of links" with a "web of meaning". The enormous work that has gone into trying to realize this vision raises (for the astute observer) fascinating philosophical questions, most notably: *What does it mean to 'give' Web pages meaning?* The question is philosophical, but the domain of Information Technology (IT) renders investigation of it fascinatingly concrete. It thus supplies an ideal opportunity to apply Peirce's "pragmatic maxim", which urges that to better understand abstract concepts (such as 'meaning') it is most helpful to think about their specific applications.

Many complex and technical discussions of meaning have taken place in philosophy over the past 400 years. However such debates have almost all shared a basic set of assumptions about meaning which is most unhelpful from an engineering perspective. We call this the "Cartesian Framework for Understanding Meaning". Terrain on the far side of these assumptions is only just being glimpsed (and understood as inviting) with the help of Peirce. This paper will outline the Cartesian framework for meaning (section 2), then the Peircean alternative (section 3), then, after a quick sketch of the semantic web project (section 4), trace some of the differing strategies and results which these two broad approaches may be perceived to bring about (sections 5 and 6).

2. The Cartesian framework for understanding meaning

Key idea: **The meaning of a sign is the intention of its producer**. This 'intention' has 2 key features, which form the basic assumptions of the Cartesian framework:

i) It is *private*. It has a location somehow 'in' a person's mind. The intention's physical location is not the key issue, though, it is that only the producer of the sign has knowledge of it. For Descartes, it was so inaccessible as to constitute a non-physical substance – hence the famous 'Cartesian dualism'.

ii) It is *incorrigible*. I am the ultimate authority on what the signs I produce mean. They mean what I intend them to mean. (This is sometimes referred to as a claim of 'first-person authority' with respect to meaning.)

Although Descartes doesn't discuss meaning explicitly in his *Meditations*, these views are extracted from what he says about *ideas*, which for him are the basic building blocks of thought and meaning. In *Meditation II* (Descartes, 1996), he claims that we only have direct access to the world of our ideas, that things in the world are quite separate from the ideas that accurately or falsely represent them. Thus for Descartes the mind is methodologically disconnected from the world so much so that he claimed to doubt whether the entire external world even exists and the ultimate authority on what its ideas mean. Error is possible, but not about what one's ideas *mean*, only about the way they are put together to form a representation of reality.

Later philosophers in the so-called 'early modern period', such as Locke and Hume, embraced a naturalistic *empiricism*, and gave up Descartes' dualistic understanding of mind as a separate substance from matter. However they retained his concept of the *idea* (private and incorrigible) as the basic unit of meaning. Thus Locke states:

> [W]ords, in their primary or immediate signification, stand for nothing but the ideas in the mind of him that uses them.
>
> Locke, 1994: 3, II, ii

In the 19th century Frege rejected the early modern understanding of meaning in terms of ideas. He pointed out that the any word, for instance 'dog', can be associated with many different, bizarre ideas in the minds of different people (disturbing ideas of being attacked, happy memories of working at the local pound, and so on). For the purposes of logic, Frege wanted a concept of meaning that could be definable more objectively, that could make a distinction between how people actually *do* understand the meaning of a sign and how they *should* understand it in order to grasp true propositions.

He therefore claimed that associated with every term was a "sense" (*Sinn*), which existed over and above its "reference" (*Bedeutung*). This 'Sinn', was an abstract object, common to everyone who grasps the meaning of a term. He sometimes referred to it as the 'mode of presentation' of the sign's reference. Thus Frege gave up the *privacy* of the Cartesian model of meaning. However he seems to have kept the *incorrigibility*. For how can I be wrong about the 'mode of presentation' which I associate with a given term?

Frege dreamed that with his new 'concept-script' he might enable a newly clear and objective understanding of the meaning of all our signs. He hoped it would then be possible to build all knowledge into an integrated taxonomic system which was *deductively complete*. (This dream was of course shattered by Russell's Paradox.) Frege's insights helped to shape twentieth century philosophy's so-called "linguistic turn", which shifted from seeing meaning as an 'idea-world relationship' to seeing it as a 'word-world relationship' (Hacking, 1975). Such theories were played out with many variations: for instance, Quine tried to do away with the concept of meaning altogether for behaviorist reasons, without success, Davidson developed an account of the meaning of propositions in terms of their 'truth-conditions', a theory which was then vastly complicated and sophisticated *via* the technical concept of possible worlds. But the one aspect of the Cartesian picture that still went unchallenged was its *incorrigibility*. For, it was thought, surely I know what the signs I use mean?[1]

3. A Peircean alternative framework for understanding meaning

Key Idea: **The meaning of a sign is the process of interpretation which occurs as the sign is used**. Peirce denied both the privacy and the incorrigibility of the Cartesian framework. In its relationship between the sign (idea or word) and the thing in the world, the Cartesian framework possessed an essentially *dyadic* structure. (Frege *nearly* escapes this dyadicity by postulating a sense as well as a reference for every sign. However given that

[1] To be strictly accurate, this assumption was finally challenged in the 1970s in the discovery of so-called 'a posteriori necessities' – for example 'water', it is claimed, 'means' H_2O whether its users know that water is H_2O or not. However this erudite debate is of limited application to the Semantic Web and will be ignored (for further details, however, see Legg, 2005).

sense for him is an abstract object, logically speaking he has arguably replaced a single dyadic relationship with two dyadic relationships.) Peirce's *triadic* model of the sign, by contrast, consists in an irreducible relation between three elements:

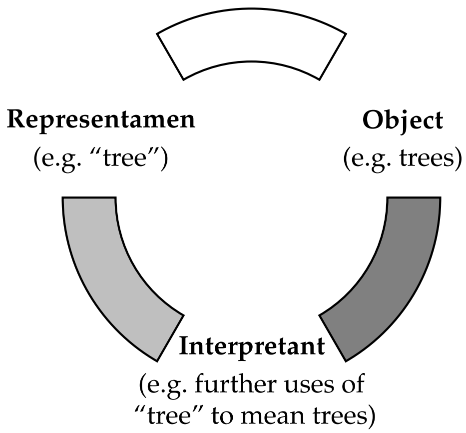

Fig. 1: Peirce: Structure of a sign.

The *representamen* is the actual signifying item. The *object* is what the sign refers to in the world. The *interpretant*, however, is Peirce's original addition to understanding meaning. It consists in further uses of the *same sign* to represent the *same object*. This is just to say that a sign must represent an object in such a way that it is understood and used again. For example, imagine that I decide to name a new star. This will not work unless other people learn the name and use it to pick out the same star. If I just stare at the night-sky, pick a name, and tell no-one about it, the process is literally meaningless, for Peirce, whatever my intentions.

Note that the interpretants, although they pick out the same object as the original sign, can 'interpret' that object in ways that differ to some degree from the ways it was interpreted originally. That is, they can not just continue but also add to, or even *shift* the meaning of the sign. One classic example is the word 'atom' as used by Democritus, and by us. Etymologically, in ancient Greek 'a-tom' meant something that cannot be broken up, but of course we have now 'split the atom'. Yet in *some* sense we are arguably still talking about the same things Democritus was, and the transition from the ancient to the present meaning was not clean or discrete. Thus by contrast to the Cartesian framework, we now have *corrigibility* with respect to meaning. The intention of the sign's producer is no longer the ultimate authority – when Democritus talked of 'atoms', he meant more than he knew.

One might ask at this point: So what is the *real* meaning of the sign? The original or the ultimate interpretation? However, do we have to choose? Peirce's theory raises the possibility that we do not. Arguably now it is more helpful to understand meaning not as an *object* (whose 'properties' can be argued over and had better not be contradictory), but as a *process*. In some real sense the meaning of a sign is what that sign *does* – how it spreads and grows (if, indeed, it does spread and grow). Thus, Peirce wrote (in a striking anticipation of contemporary use of the word 'virtual'):

> no present actual thought (which is a mere feeling) has any meaning, any intellectual value; for this lies not in what is actually thought, but in what this thought may be connected with in representation by subsequent thoughts; so that the meaning of a thought is altogether something virtual. EP 1:42

Note how this account renders meaning *public*. In the Cartesian framework, to really know what a sign means, you would need to get into the head of its producer (which alas is not possible). In the Peircean framework, to know what a sign means, look at what people are *doing* with it. Thus the responsibility for the meaning of a sign resides in a whole community. Relatedly, Peirce derived his account of *truth* by idealizing over this process of developing and using signs in a 'community of inquiry', a view which has been widely criticized as insufficiently objective. Yet Peirce denied this, claiming that over the long-run, within a broad enough community, sign-use was intrinsically self-correcting. It is also important to note that what the community is 'doing with' a given sign is not just what they are 'doing with it in their heads' by thinking about it, but what kinds of *practical activities* they are scaffolding with its help. Consider for instance, the term 'potting mix'. For Peirce it is part of its very meaning that people actually buy a certain brown stuff, put it in pots and insert plants in it.

As co-founder of the new predicate logic, Peirce pursued a vision interestingly different to Frege's regarding how it should advance human knowledge. As a pragmatist Peirce thought Frege's attempt to explicitly formalize the entire meaning of signs impossible, for an irreducible dimension of the meaning of any sign, such as 'tree', is the effects which an agent situated in the world would experience in relevant situations, such as tree-climbing, botanical investigations of new tree species, and so on, and not all of these can be anticipated in advance. In short, then, Peirce replaces a static model of meaning-as-object with a new dynamic model of meaning-as-process, where what a sign means is open to view (public), able to shift and change over time (corrigible), and inextricably entwined with actual tasks and projects.

4. The Semantic Web: an overview

4.1 Goals and challenges

Semantic Web developers embrace a wide variety of goals, including (from lesser to greater ambitiousness):

- disambiguating 'merely syntactic' Web searches, for instance distinguishing "Turkey" the country from "Turkey" the bird
- finding 'semantic joins' in databases
- indexing text and semantic markup together in order to improve Web retrieval performance (to turn the entire Web into one enormous 'distributed database')
- enabling software agents to interpret the meaning of websites in order to solve a wide range of arbitrarily complex tasks (from document-search to scheduling doctor's appointments)

Challenges for implementing it may be divided into 'technical' and 'human'. Technical challenges include inferential tractability, logical consistency, and the rapid changeability of information on the Web. The human challenges are equally problematic, and include: "Who will mark up Web pages with the required semantic metadata?", and "Who gets to say what that metadata means"?

4.2 Basic technologies

Semantic web development so far has centered around two new markup languages, which however by themselves are not sufficient to create a 'Web semantics'.

1) XML. XML was initially conceived of as a simple way to send documents across the Web, allowing authors to define their own tags, and thus document formats, subject to a simple syntax. Each new tag is linked to some unique 'namespace'. Though the term 'namespace' might suggest some further document which includes definitions for the tags, in practice

it is often just a naked URI, essentially only a way of indexing different tags uniquely *via* prefixes.

Anyone can define an XML namespace. So how do they relate to each other, semantically-speaking? Do two tags from different namespaces have the same meaning if they consist of the same character-string? No, for I could define a `<pine>` tag in my namespace to 'mean' pine *trees*, while a `<pine>` tag in another namespace is designed to apply to pine *wood* and anything made from it. Thus each namespace's tags are assumed to be distinct in meaning, and translating between them is a further problem. Thus XML arguably only provides 'syntactic' interoperability at best. This should not be too surprising since XML was not designed to share *meaning* so much as 'document format', a concept which includes any kind of structure within data (e.g. that a document contains just four elements). Semantic web developers' desire to represent meaning more purely and explicitly led to the development of RDF.

2) RDF. RDF stands for 'Resource Description Framework'. Strictly speaking, RDF is not a language but a data-model. In a key advance on XML, RDF introduces propositional structure. Each RDF 'proposition' has three parts, sometimes referred to as 'subject', 'predicate' and 'object' (e.g. Beckett 2004; Swartz, 2002), and sometimes as 'object', 'attribute', and 'value' (e.g. Decker et al, 2000). As an example, take the proposition: `The Kauri is a kind of pine tree.` Here the subject/object would be 'Kauri', the predicate/attribute would be 'a kind of', and the object/value would be 'pine tree'.

Does marking up web pages with RDF propositions make the Web 'semantic'? In the example above, a propositional structure exists, with all three components envisaged to be assigned URIs. However we have seen that URIs are merely indices. Once again, RDF does not determine what they are indexed to. As Sowa has written:

> By standardizing the notations, XML and RDF take an important first step, but that step is insufficient for data sharing without some way of comparing, relating, and translating the vocabularies. Sowa, 2000

5. Cartesian approaches to web semantics

Key idea: **Try to define an authoritative sign-producer's intention for what each sign should mean**. If one believes that the meaning of a sign resides in what the user of a sign *intends* it to mean, it would appear that the way to give the Web meaning is to try to define that intention, in machine-readable fashion, as fully and determinately as possible. This thinking has resulted in many attempts to set up *silos of meaning*, also known as 'formal ontologies'. Some key examples will now be discussed.

1) RDFS. RDF Schema, an extension of RDF, allowed one to declare classes, and properties, populate classes with instances, and organize them into a subsumption hierarchy. It also allowed range and domain constraints to be added to properties, and properties to be ascribed to individuals. It was initially envisaged that web-semantics-defining ontologies would be stored in this language. However, RDFS turned out to be too logically simple to express a great deal of what one might wish to say to authoritatively define the meanings of terms. Though one can declare new classes and populate them with instances, one cannot say anything further about these classes and instances (Delteil et al, 2001). For instance, one cannot state that two names denote the same person. At the end of the day RDFS is still just a set of terms indexed *via* namespaces whose further meaning is opaque. RDFS was never widely used and its main components are now folded into the more expressive OWL (see next).

2) OWL. OWL ('Web Ontology Language') was a renaming and reworking of DAML+OIL. It became a W3C Recommendation in February 2004 and is currently the flagship ontology of the W3C group. OWL goes beyond RDFS by providing additional vocabulary and a formal semantics. The additional vocabulary includes the ability to define classes in terms of logical relationships between other classes, the ability to state class cardinality, equality (for both classes and individuals), and logical characteristics of properties. It was hoped that this greater expressivity would enable it to outdo RDFS in capturing all information needed to define the semantics of terms on the web. Greater expressivity has costs in inferential tractability, however, so OWL has three versions, each an extension of the previous: OWL Lite, OWL DL and OWL Full.

The W3C envisaged that once they provided the OWL language, the world would respond by defining and contributing ontologies, and a number of ontology libraries/clearing-houses have been set up for this purpose (for instance: the DAML ontology library `http://www.daml.org/ontologies/`, and the Protégé ontology library `http://protege.cim3.net/cgi-bin/wiki.pl?ProtegeOntologiesLibrary`). However at present coverage is patchy at best. For instance regarding our test-concept, 'tree', a search on Swoogle, UMBC's ontology search engine (`http://swoogle.`

umbc.edu), produces just a few very scattered assertions.[2] It would thus appear that OWL is not currently widely used outside the academic research context (though OWL DL is used more than the other two).

There are a number of reasons for this. First of all, it is clumsy and verbose: the OWL translation of, "A student is a person who is enrolled in at least one thing" runs to 10 complex nested lines. Secondly, it is complained that its graph/tree data-structure does not scale for real-world applications foundering for example, when dealing with the information in a typical business spreadsheet (Bergman, 2006), nor does it allow user-defined datatypes. Finally, the exact formal relationship between OWL and RDF is a delicate matter. While OWL Full can be viewed as an extension of RDF, OWL Lite and OWL DL can only be viewed as extensions of a restricted view of RDF. (McGuinness and van Harmelen, 2004). This creates a problem for layering OWL over RDF. From the Peircean perspective, however, OWL's most fundamental issue is the 'human' one: its current lack of *use*.

3) CYC. The original (yet continuing) most ambitious formal ontology project of all is the Cyc project (http://www.cyc.com). It has deep roots in classical AI. It is most ambitious in terms of size (over 600 000 categories), depth of knowledge (over 2 million axioms), and time devoted to it (over 700 person-years) (Sowa, 2004). It has its own purpose-built inference engine, and natural language interface. The Cyc project is the most systematic, unified attempt to not just index terms but to *describe* their meanings in machine-readable terms. Thus its representation of a tree, #$Tree-ThePlant, is distinguished from #$Tree-PathSystem. It comes with *axiomatic assertions* (for instance, "A tree is largely made of wood") and *rules* (for instance, "If a tree is cut down, then it will be destroyed"), from which further facts can be deduced (for instance, "If the pine tree in my backyard is cut down, then it will be destroyed."). It manages to bypass the W3C's problems with layering OWL on RDF, by using its own in-house language, the purpose-built CycL (which has the expressivity of higher-order logic).

The company has made strenuous efforts to position itself for the Semantic Web, by for instance mapping in databases such as FIPS (Federal In-

[2] For example, "A Tree is a kind of LandscapeProduct" in http://individual.utoronto.ca/hesham/Ontology/IPDLite.owl, and nothing else, "A TreeRing is a kind of Vegetation" in http://sweet.jpl.nasa.gov/ontology/biosphere.owl, and nothing else. These assertions are mixed with many others concerning trees as mathematical structures, with no obvious way of telling that this is a different concept.

formation Processing Standards), the CIA WorldFactbook (Reed & Lenat, 2002), and WordNet, and producing automated OWL annotation of text documents (Witbrock et al, 2004). Nevertheless, once again, unfortunately Semantic Web developers outside the company have so far made little *use* of this ontology. Its system of categories is extremely complex, requiring philosophical training to understand. Inferential tractability is a particular worry given the expressivity of the CycL language, and the monolithic holism of such a giant ontology unfortunately leads to brittleness.

In conclusion, then, the attempts of these formal ontology projects to 'create *ex nihilo*' the meaning of signs on the semantic web *via* a set of antecedent definitions misunderstand what it is for something to have meaning. From a Peircean perspective the mere fact that these projects are not widely used *is* the key argument against their having real 'significance'.

6. Peircean approaches to web semantics

Key idea: **Build applications which allow interpretants to freely *grow*, within whatever communities choose to use them**. What *is* growing right now on the Web? Some developments manifestly are. (These are sometimes referred to as the 'lower-case semantic web', by contrast to the W3C's official efforts):

1) Tagging. Tags are labels added to the Web voluntarily by users. Ontologically speaking, the practice is entirely uncontrolled – no categories are prepared or agreed upon in advance. (Thus a given CD might be labelled "`boring`", "`Mike_likes_this`", "`driving_music`", and "`songs_about_fish`"). Tagging began as a way of labelling web-pages with words or phrases meaningful to oneself in order to rediscover them quickly, but has spread to embrace a number of other much more public uses, as a variety of websites has emerged to serve as tag clearing-houses. Examples of such sites include del.icio.us. for tagged bookmarks (`http://del.icio.us/`) and Flickr for tagged photographs (`http://www.flickr.com/`).

Tagging is said to produce not a *taxonomy* (in the sense of a mark-up according to a pre-given ontology) but a '*folksonomy*' (Weinberger, 2005). Despite the 'feral' source of tags, it has been argued that at the level of the entire Web the impact of individual idiosyncrasy lessens, and that, "[b]y forgoing formal classification, tags enable a huge amount of user-produced organizational value, at vanishingly small cost" (Shirky, 2005).

2) RSS *autodiscovery*. This technology 'syndicates' websites (frequently weblogs) by providing summaries of their content, links to the full version,

and other meta-data, in an XML file called an RSS feed. Content is filtered for individual users using keywords (the choice of which once again is wholly personal and idiosyncratic).

3) *Collaborative websites*. These provide a medium in which speakers of any language define, describe and discuss topics of contemporary relevance. The resulting information is freely available, electronically encoded and conveniently presented. Such websites are quickly springing up on every conceivable subject, for instance: music (`http://musicbrainz.org/`), exercise (`http://www.favoriterun.com/`) and biosecurity (`http://paipm.cas.psu.edu/biosecurity.html`), to give just a few examples. One of the original and most impressive websites, however, and by far the most comprehensive, is the online encyclopaedia Wikipedia. This project is a remarkable and unanticipated realisation of Peirce's 'community of inquiry', its ever-increasing level of accuracy causing considerable surprise in those who do not hold to Peirce's theory of truth (but a sense of vindication in those who do).

6.1 A case-study in semantics extraction from user-supplied web content

Wikipedia's immense potential as an automated, just-in-time source of semantic knowledge, by contrast to manually encoded, 'frozen' silos of meaning, is just beginning to be explored scientifically. Each web page/ article in Wikipedia defines a specific concept and is inter-linked with other articles in the encyclopaedia. Milne et al. (2006) extract a thesaurus by treating article names as terms and hyperlinks as semantic relations between them. By looking at different types of links, they are able to identify three types of semantic relations that are commonly used in manually crafted thesauri:

- *Synonymy/Polysemy* – Redirect pages in Wikipedia link synonymous phrases to the same article (e.g. 'Pine tree' is linked to 'Pine'). Disambiguation pages help to identify ambiguous terms (e.g. 'Tree' as 'woody plant' and 'Tree' as 'data structure', along with 14 other possible senses)

- *Hierarchical relations* – Wikipedia's category structure defines relations between broader and narrower concepts (e.g. 'Pine' belongs to the category 'Pinaceae', which is in turn a part of the category 'Plant families')

- *Associative relations* – Any other hyperlinks connecting article pages are association between the concepts of different strength. (e.g. On the page 'Pine' there are links to articles 'pine nuts', 'evergreen', 'christmas trees' and 'parks')

In this way they concretely demonstrate how a semantic knowledge base can be created on-the-fly, tailored to any document collection. Figure 2 demonstrates an example, where a mini-version of a thesaurus was extracted given merely Wikipedia and the following short document:

> *"Tane Mahuta is New Zealand's tallest Kauri Tree, growing in Waipoua Forest. Its massive smooth, grey-white trunk rises 59 feet before a branch appears."*

Thick lines represent hierarchical relations, thin lines are association relations; dotted lines reflect polysemy relations to homonyms. Note the detail of this result by contrast to the random and patchy coverage of current OWL ontologies. Furthermore, this semantic structure reflects *public* opinion on the relatedness between document terms, it reflects an up-to-the-minute version of it, and the restriction to a particular document guarantees that all included terms are relevant for this particular knowledge domain.

Given a large agricultural document collection and a thesaurus Agrovoc, manually created to cover the same domain, Milne et al. report that Wikipedia covers more than twice as many document concepts as Agrovoc.

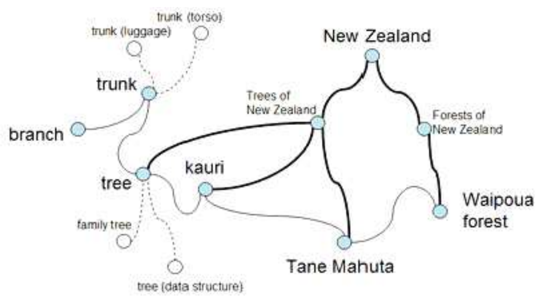

Figure 2. Thesaurus extracted from Wikipedia for a sample document.

7. Conclusion

Cartesians assume that in order to make the Semantic Web happen it is necessary to make a huge defining effort, to somehow encode for the computer the private intentions we have when we produce signs. The Peircean approach by contrast involves realizing that vast quantities of semantic data *already exists* on the Web, our job is to work out how to leverage it. It might be objected that it is difficult to envisage how any kind of coherent inferencing might be built on such a turbulent and amateur base as (is most of) this user-supplied metadata. Still, these criticisms might be made of Google's deployment of its spectacularly successful page-rank algorithm across the turbulent and amateur World Wide Web. Google's genius was to realize that it did not need to pay people to inspect and rate websites, as such data already existed in the form of hyperlinks. In Peircean terms, we can say that Google realized that hyperlinks constitute *interpretants* of the web-pages they link to. For in most cases such links indicate that the creator of the linking page thought that their page was in some sense relevant to, and thus *about the same thing* (object) as the linked-to page. In a similar way, then, tags can be considered as interpretants of the web-pages they describe, blog syndications as interpretants of the blogs syndicated, Wikipedia entries as interpretants of the terms defined, and so on. The kinds of inferencing that will trace such interpretants and transform them into semantic data is not the neat, deductivist rule-based reasoning of 'good old fashioned AI'. We need new models.

Having mentioned AI, it's worth noting that here also philosophical theories of meaning are not mere abstract speculation but directly influence what we envision and attempt to build. This is not surprising since the Semantic Web at its most boosterish arguably consists in many old AI goals in 1990s dress (Halpin, 2004). The classic 1950s-era model of AI – something like a digital encyclopedia in the head of a robot – may now be seen as a poignant attempt to make concrete the Cartesian picture of meaning as idea *in* the head. By contrast, Peirce's account of meaning as interpretants led him to write, "just as we say that a body is in motion, and not that motion is in a body, we ought to say that we are in thought, and not that thoughts are in us." (EP 1:42 fn.). In this sense, perhaps as with 'Web semantics' also with 'Web intelligence' we already have more at our disposal than we realize.[3]

[3] A considerably revised and expanded version of this paper is forthcoming in *Semiotica*, under the title "Peirce, Meaning and the Semantic Web".

References

Beckett, D. (2004). "RDF/XML Syntax Specification (Revised)." Retrieved 10 August 2007 from
`http://www.w3.org/TR/2004/REC-rdf-syntax-grammar-20040210/`

Bergman, M. K. (2006). "OWL Ontologies: When Machine-Readable is not Good Enough." Retrieved 20 August 2007 from
`http://www.mkbergman.com/?p=194`

Berners-Lee, T. (2002). "Foreword." In D. Fensel (Ed.), *Spinning the Semantic Web: Bringing the World Wide Web to its Full Potential*. Cambridge: MIT Press.

Berners-Lee, T., Hendler, J, & Lassila, O. (2001). "The Semantic Web." *Scientific American*, 284(5), 34–43.

Brickley, D. & Guha, R. V. (2003). "RDF Vocabulary Description Language 1.0: RDF Schema." Retrieved 16 May 2007 from `http://www.w3.org/TR/rdf-schema/`

Decker, S., van Harmelen, F., Broekstra, J., Erdmann, M., Fensel, D., Horrocks, I., Klein, M., & Melnik, S. (2000). "The Semantic Web – On the Respective Roles of XML and RDF." IEEE Internet Computing, 4(5) 63–74.

Descartes, R. (1996). *Meditations on First Philosophy: With Selections from the Objections and Replies.* (Eds.) Cottingham, J., Williams, B. & Ameriks, K. Cambridge: Cambridge University Press.

Fikes, R., Hayes, P., & Horrocks, I., (Eds.) (2002). "DAML Query Language (DQL) Abstract Specification." Retrieved 20 August 2007 from
`http://www.daml.org/2002/08/dql/`

Frege, G. (1879). *Begriffsschrift: eine der arithmetischen nachgebildete Formelsprache des reinen Denkens*. Halle (Appears in English in *Conceptual Notation, and Related Articles.* Bynum, T. W. (Ed. & Trans.) Oxford: Clarendon Press, 1972.)

Gruber, T. R. (1995). "Toward Principles for the Design of Ontologies Used for Knowledge-Sharing." *International Journal of Human and Computer Studies*, 43 (5/6), 907–928.

Guha, R. (2003). "Semantic Issues in Web-Scale Knowledge Aggregation." *Knowledge Systems Laboratory*, 2003. Retrieved 16 May 2007 from
`http://ksl.stanford.edu/KSL_Abstracts/KSL-03-15.html`

Guha, R. V., Lenat, D.B., Pittman, K., Pratt, D., & Shepherd, M. (1990). "Cyc: A Midterm Report." *Communications of the ACM* 33 (8).

Guha, R. & McCool, R. (2003). "TAP: A Semantic Web Platform." *Journal of Network Computing* 42, 557–577.

Hacking, I. (1975). *Why Does Language Matter to Philosophy?* Cambridge: Cambridge University Press.

Halpin, H. (2004). "The Semantic Web: The Origins of AI Redux." *Proceedings of the Fourth International Workshop on the Philosophy and History of Mathematics and Computation*, San Sebastien, Spain, 2004.

Horrocks, I., Patel-Schneider, P., & Van Harmelen, F. (2002). "Reviewing the Design of DAML+OIL: An Ontology Language for the Semantic Web." *Proceedings* *of the 18th National Conference on Artificial Intelligence* AAAI2002. Edmonton, Canada, 792–797.

Legg, C. (2005). "The Meaning of Meaning-Fallibilism." *Axiomathes*, 15 (2), 293–318.

Lenat, D. B. (1995). "Cyc: A Large-Scale Investment in Knowledge Infrastructure." *Communications of the* ACM, 38(11), 33–38

Locke, J. (1690/1994). *An Essay Concerning Human Understanding*. Prometheus Books.

McCool, R., Fikes, R., & Guha, R. (2003). "Semantic Issues in Web-Scale Knowledge Aggregation." Retrieved 1 August 2007 from `http://www-ksl.stanford.edu/KSL_Abstracts/KSL-03-15.html`

McGuinness, D., & van Harmelen, F. (2004). "OWL Web Ontology Language: Overview." Retrieved 8 August 2007 from `http://www.w3.org/TR/owl-features/`

Maedche, A., Motik, B., & Stojanovic, L. (2003). "Managing Multiple and Distributed Ontologies on the Semantic Web." *The* VLDB *Journal*, 12, 286–302.

Melnik, S. (1999). "Bridging the Gap Between RDF and XML." Retrieved 6 August 2007 from `http://www-db.stanford.edu/ melnik/rdf/fusion.html`

Milne, D., Medelyan, O., & Witten, I. (2006). "Mining Domain-Specific Thesauri from Wikipedia: A Case Study." IEEE/WIC/ACM International *Conference on Web Intelligence (WI 2006 Main Conference Proceedings).* WI'06, 442–448.

Patel-Schneider, P.F. & Fensel, D. (2002). "Layering the Semantic Web: Problems and Directions." In I. Horrocks and J. Hendler (Eds.), ISWC *2002,* LNCS *2342* (pp. 16–29).

Patel-Schneider, P.F. & Siméon, J. (2002). "The Yin/Yang Web: XML Syntax and RDF Semantics." *WWW* 2002. Honolulu, Hawaii.

Phipps, S. (2000). "Meaning, not Markup." XML *Journal*, 1(1), 66.

Ramachandran, D., Reagan, P., & Goolsbey, K. (2005). "First-Orderized Research-Cyc: Expressivity and Efficiency in a Common-Sense Ontology." In P. Shvaiko et al (Eds.), *Papers from the 2005* AAAI *Workshop on Contexts and Ontologies: Theory, Practice and Applications* (pp. 33-40). Menlo Park: The AAAI Press.

Reed, S., & Lenat, D. (2002). "Mapping Ontologies into Cyc." AAAI *2002 Conference Workshop on Ontologies For The Semantic Web*, Edmonton, Canada.

Shah, U., Finin, T., Joshi, A., Cost, R.S., & Mayfield, J. (2002). "Information Retrieval on the Semantic Web." In D. Grossman (Ed.), *Proceedings of the 11th International Conference on Information and Knowledge Management:* CIKM'*02* (pp. 461-468). New York: ACM Press.

Shirky, C. (2005). "Ontology Is Over-Rated: Categories, Links and Tags." Retrieved 24 May 2007 from `http://www.shirky.com/writings/ontology_overrated.html`

Sowa, J. (2000). "Ontology, Metadata, and Semiotics." In B. Ganter & G. W. Mineau (Eds.), *Conceptual Structures: Logical, Linguistic, and Computational Issues, Lecture Notes in AI #1867* (pp. 55–81) Berlin: Springer-Verlag.

Sowa, J. (2001). "Signs, Processes, and Language Games." Retrieved 10 December 2002 from `http://www.jfsowa.com/pubs/signproc.htm`

Sowa, J. (2004). "The Challenge of Knowledge Soup." Retrieved 24 May 2007 from `http://www.jfsowa.com/pubs/challenge.pdf`

Swartz, A. (2002). "The Semantic Web in Breadth." Retrieved 20 August 2007 from `http://logicerror.com/semantic-Web-long`

Weinberger, D. (2005). "Taxonomies to Tags: From Trees to Piles of Leaves." Release 1.0, 23(2), 1–33.

Witbrock, M., Panton, K., Reed, S.L., Schneider, D., Aldag, B., Reimers, M., & Bertolo, S. (2004). "Automated OWL Annotation Assisted by a Large Knowledge Base." In *Workshop Notes of the 2004 Workshop on Knowledge Markup and Semantic Annotation at the 3rd International Semantic Web Conference* ISWC*2004* (pp. 71-80). Hiroshima, Japan.