GIVE ME A SIGN: WHAT DO THINGS MEAN ON THE SEMANTIC WEB?

Michael K. Bergman[1], Coralville, Iowa USA
January 24, 2012
AI3:::Adaptive Information blog

The crowning achievement of the semantc Web is the simple use of URIs to identify data. Further, if the URI identifier can resolve to a representation of that data, it now becomes an integral part of the HTTP access protocol of the Web while providing a unique identifier for the data. These innovations provide the basis for distributed data at global scale, all accessible via Web devices such as browsers and smartphones that are now a ubiquitous part of our daily lives.

Yet, despite these profound and simple innovations, the semantic Web's designers and early practitioners and advocates have been mired in a muddled, metaphysical argument of at least a decade over what these URIs mean, what they reference, and what their actual true identity is. These muddles about naming and identity, it might be argued, are due to computer scientists and programmers trying to grapple with issues more properly the domain of philosophers and linguists. But that would be unfair. For philosophers and linguists themselves have for centuries also grappled with these same conundrums [1].

As I argue in this piece, part of the muddle results from attempting to do too much with URIs while another part results from not doing enough. I am also not trying to directly enter the fray of current standards deliberations. (Despite a decade of controversy, I optimistically believe that the messy process of argument and consensus building will work itself out [2].) What I am trying to do in this piece, however, is to look to one of America's pre-eminent philosophers and logicians, Charles Sanders Peirce (pronounced "purse"), to inform how these controversies of naming, identity and meaning may be dissected and resolved.

'Identity Crisis', httpRange-14, and Issue 57

The Web began as a way to hyperlink between documents, generally Web pages expressed in the HTML markup language. These initial links were called URLs (uniform resource locators), and each pointed to various kinds of electronic resources (documents) that could be accessed and retrieved on the Web. These resources could be documents written in HTML or other encodings (PDFs, other electronic formats), images, streaming media like audio or videos, and the like [3].

All was well and good until the idea of the semantic Web, which postulated that information about the real world — concepts, people and things — could also be referenced and made available for reasoning and discussion on the Web. With this idea, the scope of the Web was massively expanded from electronic resources that could be downloaded and accessed via the Web to now include virtually any topic of human discourse. The rub, of course, was that ideas such as abstract concepts or people or things could not be "dereferenced" nor downloaded from the Web.

One of the first things that needed to change was to define a broader concept of a URI "identifier" above the more limited concept of a URL "locator", since many of these new things that could be referenced on the Web

---

[1]Email: mike@mkbergman.com

went beyond electronic resources that could be accessed and viewed [3]. But, since what the *referent* of the URI now actually might be became uncertain — was it a concept or a Web page that could be viewed or something else? — a number of commentators began to note this uncertainty as the "identity crisis" of the Web [4]. The topic took on much fervor and metaphysical argument, such that by 2003, Sandro Hawke, a staffer of the standards-setting W3C (World Wide Web Consortium), was able to say, "This is an old issue, and people are tired of it" [5].

Yet, for many of the reasons described more fully below, the issue refused to go away. The Technical Architecture Group (TAG) of the W3C took up the issue, under a rubric that came to be known as *httpRange-14* [6]. The issue was first raised in March 2002 by Tim Berners-Lee, accepted for TAG deliberations in February 2003, with then a resolution offered in June 2005 [7]. (Refer to the original resolution and other information [6] to understand the nuances of this resolution, since particular commentary on that approach is not the focus of this article.) Suffice it to say here, however, that this resolution posited an entirely new distinction of Web content into "information resources" and "non-information resources", and also recommended the use of the HTTP 303 redirect code for when agents requesting a URI should be directed to concepts versus viewable documents.

This "resolution" has been anything but. Not only can no one clearly distinguish these *de novo* classes of "information resources" [19], but the whole approach felt arbitrary and kludgy.

Meanwhile, the confusions caused by the "identity crisis" and httpRange-14 continued to perpetuate themselves. In 2006, a major workshop on "Identity, Reference and the Web" (IRW 2006) was held in conjunction with the Web's major WWW2006 conference in Edinburgh, Scotland, on May 23, 2006 [8]. The various presentations and its summary (by Harry Halpin) are very useful to understand these issues. What was starting to jell at this time was the understanding that the basis of identity and meaning on the Web posed new questions, and ones that philosophers, logicians and linguists needed to be consulted to help inform.

The fiat of the TAG's 2005 resolution has failed to take hold. Over the ensuing years, various eruptions have occurred on mailing lists and within the TAG itself (now expressed as Issue 57) to revisit these questions and bring the steps moving forward into some coherent new understanding. Though linked data has been premised on best-practice implementation of these resolutions [9], and has been a qualified success, many (myself included) would claim that the extra steps and inefficiencies required from the TAG's httpRange-14 guidance have been hindrances, not facilitators, of the uptake of linked data (or the semantic Web).

Today, despite the efforts of some to claim the issue closed, it is not. Issue 57 and the periodic bursts from notable semantic Web advocates such as Ian Davis [10], Pat Hayes and Harry Halpin [11], Ed Summers [12], Xiaoshu Wang [13], David Booth [14] and TAG members themselves, such as Larry Masinter [15] and Jonathan Rees [16], point to continued irresolution and discontent within the advocate community. Issue 57 currently remains open. Meanwhile, I think, all of us interested in such matters can express concern that linked data, the semantic Web and interoperable structured data have seen less uptake than any of us had hoped or wanted over the past decade. As I have stated elsewhere, unclear semantics and muddled guidelines help to undercut potential use.

As each of the eruptions over these identity issues has occurred, the competing camps have often been characterized as "talking past one another"; that is, not communicating in such a way as to help resolve to consensus. While it is hardly my position to do so, I try to encapsulate below the various positions and prejudices as I see them in this decades-long debate. I also try to share my own learning that may help inform some common ground. Forgive me if I overly simplify these vexing issues by returning to what I see as some first principles . . . .

What's in a Name?

One legacy of the initial document Web is the perception that Web addresses have meaning. We have all heard of the multi-million dollar purchasing of domains [17] and the adjudication that may occur when domains are hijacked from their known brands or trademark owners. This legacy has tended to imbue URIs with a perceived value. It is not by accident, I believe, that many within the semantic Web and linked data communities still refer to "minting" URIs. Some believe that ownership and control over URIs may be equivalent to grabbing up valuable real estate. It is also the case that many believe the "name" given to a URI acts to *name* the referent to which it refers.

This perception is partially true, partially false, but moreover incomplete in all cases. We can illustrate these points with the global icon, "Coca-Cola".

As for the naming aspects, let's dissect what we mean when we use the label "Coca-Cola" (in a URI or otherwise). Perhaps the first thing that comes to mind is "Coca-Cola," the beverage (which has a description on Wikipedia, among other references). Because of its ubiquity, we may also recognize the image of the Coca-Cola bottle to the left as a symbol for this same beverage. (Though, in the hilarious movie, *The Gods, They Must be Crazy*, Kalahari Bushmen, who had no prior experience of Coca-Cola, took the bottle to be magical with evil powers [18].) Yet even as reference to the beverage, the naming aspects are a bit cloudy since we could also use the fully qualified synonyms of "Coke", "Coca-cola" (small C), "Classic Coke" and the hundreds of language variants worldwide.

On the other hand, the label "Coca-Cola" could just as easily conjure The Coca-Cola Company itself. Indeed, the company web site is the location pointed to by the URI of http://www.thecoca-colacompany.com/. But, even that URI, which points to the home Web page of the company, does not do justice to conveying an understanding or description of the company. For that, additional URIs may need to be invoked, such as the description at Wikipedia, the company's own company description page, plus perhaps the company's similar heritage page.

Of course, even these links and references only begin to scratch the surface of what the company Coca-Cola actually is: headquarters, manufacturing facilities, 140,000 employees, shareholders, management, legal entities, patents and Coke recipe, and the like. Whether in human languages or URIs, in any attempt to signify something via symbols or words (themselves another form of symbol), we risk ambiguity and incompleteness.

URI shorteners also undercut the idea that a URI necessarily "names" something. Using the service bitly, we can shorten the link to the Wikipedia description of the Coke beverage to http://bit.ly/xnbA6 and we can shorten the link to The Coca-Cola Company Web site to http://bit.ly/9ojUpL. I think we can fairly say that neither of these shortened links "name" their referents. The most we can say about a URI is that it *points to* something. With the vagaries of meaning in human languages, we might also say that URIs *refer to* something, *denote* something or *identify* (but not in the sense of completely define) something.

From this discussion, we can assert with respect to the use of URIs as "names" that:

1. In all cases, URIs are pointers to a particular *referent*
2. In some cases, URIs do act to "name" some things
3. Yet, even when used as "names," there can be ambiguity as to what exactly the referent is that is denoted by the name
4. Resolving what such "names" mean is a matter of context and reference to further information or links, and
5. Because URIs may act as "names", it is appropriate to consider social conventions and contracts (e.g., trademarks, brands, legal status) in adjudicating who can own the URI.

In summary, I think we can say that URIs may act as names, but not in all or most cases, and when used as such are often ambiguous. Absolutely associating URIs as names is way too heavy a burden, and incorrect in most cases.

What is a Resource?

The "name" discussion above masks that in some cases we are talking about a readable Web document or image (such as the Wikipedia description of the Coke beverage or its image) versus the "actual" thing in the real world (the Coke beverage itself or even the company). This distinction is what led to the so-called "identity crisis", for which Ian Davis has used a toucan as his illustrative thing [10].

As I note in the conclusion, I like Davis' approach to the identity conundrum insofar as Web architecture and linked data guidance are concerned. But here my purpose is more subtle: I want to tease apart still further the apparent distinction between an electronic description of something on the Web and the "actual" something. Like Davis, let's use the toucan.

In our strawman case, we too use a description of the toucan (on Wikipedia) to represent our "information resource" (the accessible, downloadable electronic document). We contrast to that a URI that we mean to convey the actual physical bird (a "non-information resource" in the jumbled jargon of httpRange-14), which we will designate via the URI of http://example.com/toucan.

Despite the tortured (and newly conjured) distinction between "information resource" and "non-information resource", the first blush reaction is that, sure, there is a difference between an electronic representation that can be accessed and viewed on the Web and its true, "actual" thing. Of course people can not actually be rendered and downloaded on the Web, but their bios and descriptions and portrait images may. While in the abstract such distinctions appear true and obvious, in the specifics that get presented to experts, there is surprising disagreement as to what is actually an "information resource" *v.* a "non-information resource" [19]. Moreover, as we inspect the real toucan further, even that distinction is quite ambiguous.

When we inspect what might be a definitive description of "toucan" on Wikipedia, we see that the term more broadly represents the family of *Ramphastidae*, which contains five genera and forty different species. The picture we are showing to the right is but of one of those forty species, that of the keel-billed toucan (*Ramphastos sulfuratus*). Viewing the images of the full list of toucan species shows just how divergent these various "physical birds" are from one another. Across all species, average sizes vary by more than a factor of three with great variation in bill sizes, coloration and range. Further, if I assert that the picture to the right is actually that of my pet keel-billed toucan, *Pretty Bird*, then we can also understand that this representation is for a specific individual bird, and not the physical keel-billed toucan species as a whole.

The point of this diversion is not a lecture on toucans, but an affirmation that distinctions between "resources" occur at multiple levels and dimensions. Just as there is no self-evident criteria as to what constitutes an "information resource", there is also not a self-evident and fully defining set of criteria as to what is the physical "toucan" bird. The meaning of what we call a "toucan" bird is not embodied in its label or even its name, but in the context and accompanying referential information that place the given referent into a context that can be communicated and understood. A URI *points to* ("refers to") something that causes us to conjure up an understanding of that thing, be it a general description of a toucan, a picture of a toucan, an understanding of a species of toucan, or a specific toucan bird. Our understanding or interpretation results from the context and surrounding information accompanying the reference.

In other words, a "resource" may be anything, which is just the way the W3C has defined it. There is not a single dimension which, magically, like "information" and "non-information," can cleanly and definitely place a

referent into some state of absolute understanding. To assert that such magic distinctions exist is a flaw of Cartesian logic, which can only be reconciled by looking to more defensible bases in logic [20].

Peirce and the Logic of Signs

The logic behind these distinctions and nuances leads us to Charles Sanders Peirce *(1839 – 1914)*. Peirce (pronounced "purse") was an American logician, philosopher and polymath of the first rank. Along with Frege, he is acknowledged as the father of predicate calculus and the notation system that formed the basis of first-order logic. His symbology and approach arguably provide the logical basis for description logics and other aspects underlying the semantic Web building blocks of the RDF data model and, eventually, the OWL language. Peirce is the acknowledged founder of pragmatism, the philosophy of linking practice and theory in a process akin to the scientific method. He was also the first formulator of existential graphs, an essential basis to the whole field now known as model theory. Though often overlooked in the 20th century, Peirce has lately been enjoying a renaissance with his voluminous writings still being deciphered and published.

The core of Peirce's world view is based in semiotics, the study and logic of signs. In his seminal writing on this, "What is in a Sign?" [21], he wrote that "every intellectual operation involves a triad of symbols" and "all reasoning is an interpretation of signs of some kind". Peirce had a predilection for expressing his ideas in "threes" throughout his writings.

Semiotics is often split into three branches: 1) syntactics – relations among signs in formal structures; 2) semantics – relations between signs and the things to which they refer; and 3) pragmatics – relations between signs and the effects they have on the people or agents who use them.

Peirce's logic of signs in fact is a taxonomy of sign relations, in which signs get reified and expanded via still further signs, ultimately leading to communication, understanding and an approximation of "canonical" truth. Peirce saw the scientific method as itself an example of this process.

A given sign is a representation amongst the triad of the *sign* itself (which Peirce called a *representamen*, the actual signifying item that stands in a well-defined kind of relation to the two other things), its *object* and its *interpretant*. The *object* is the actual thing itself. The *interpretant* is how the agent or the perceiver of the sign understands and interprets the sign. Depending on the context and use, a *sign* (or *representamen*) may be either an *icon* (a likeness), an indicator or *index* (a pointer or physical linkage to the object) or a *symbol* (understood convention that represents the object, such as a word or other meaningful signifier).

An *interpretant* in its barest form is a sign's meaning, implication, or ramification. For a *sign* to be effective, it must represent an object in such a way that it is understood and used again. This makes the assignment and use of signs a community process of understanding and acceptance [20], as well as a truth-verifying exercise of testing and confirming accepted associations.

John Sowa has done much to help make some of Peirce's obscure language and terminology more accessible to lay readers [22]. He has expressed Peirce's basic triad of sign relations as follows, based around the Yojo animist cat figure used by the character Queequeg in Herman Melville's *Moby-Dick*:
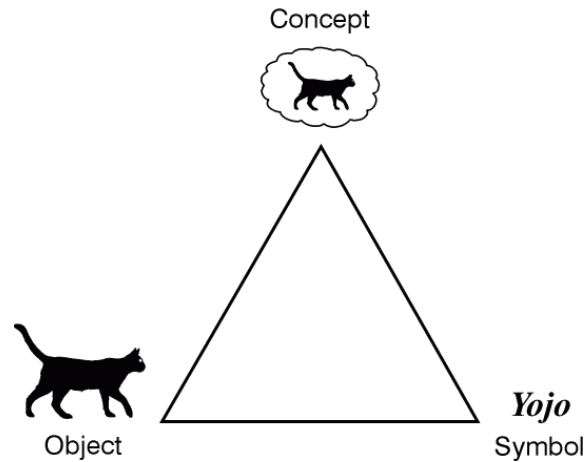
In this figure, *object* and *symbol* are the same as the Peirce triad; *concept* is the interpretant in this case. The use of the word 'Yojo' conjures the concept of cat.

This basic triad representation has been used in many contexts, with various replacements or terms at the nodes. Its basic form is known as the Meaning Triangle, as was popularized by Ogden and Richards in 1923 [23].

The key aspect of signs for Peirce, though, is the ongoing process of interpretation and reference to further signs, a process he called *semiosis*. A sign of an object leads to interpretants, which, as signs, then lead to further interpretants. In the Sowa example below, we show how meaning triangles can be linked to one another, in this case by abstracting that the triangles themselves are concepts of representation; we can abstract the ideas of both *concept* and *symbol*:
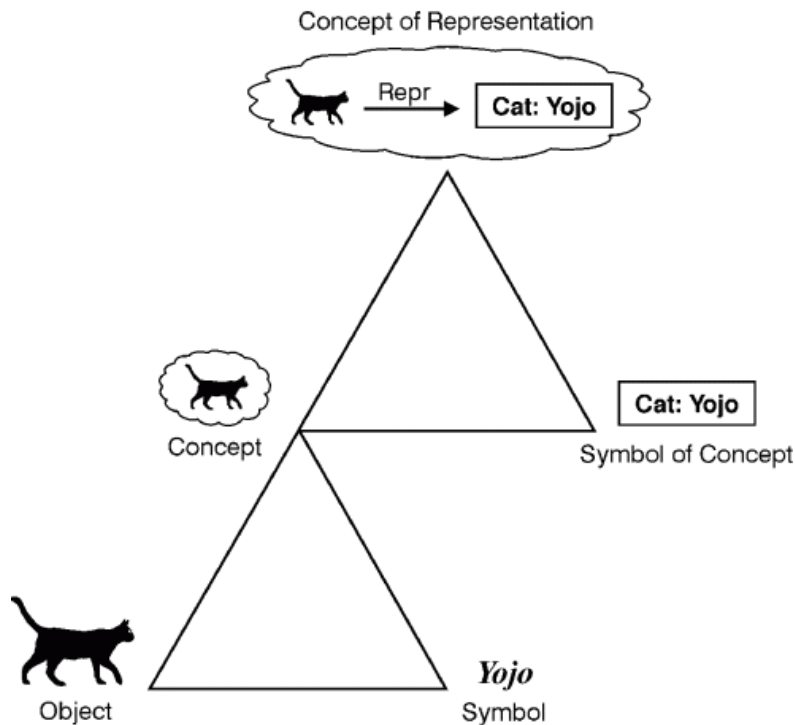
We can apply this same cascade of interpretation to the idea of the *sign* (or *representamen*), which in this case shows that a name can be related to a word symbol, which in itself is a combination of characters in a string called 'Yojo':
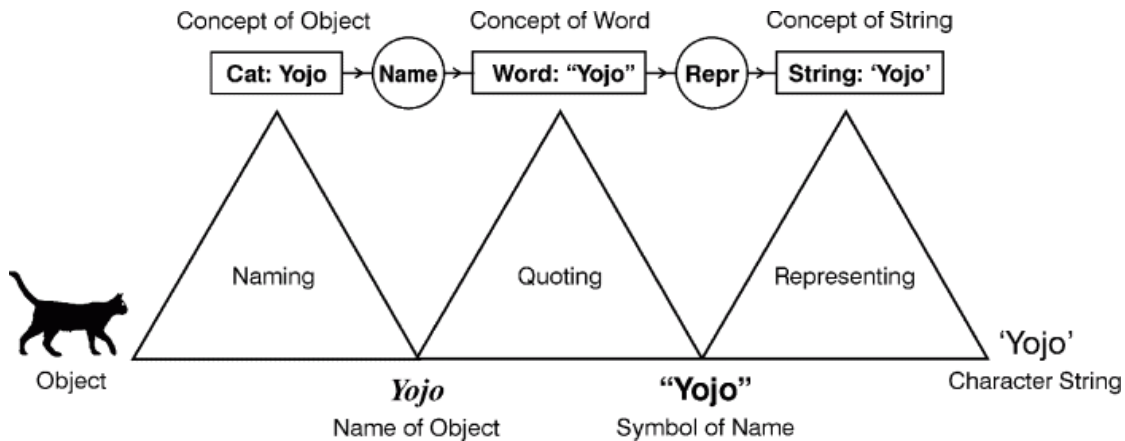
According to Sowa [22]:

"*What is revolutionary about Peirce's logic is the explicit recognition of multiple universes of discourse, contexts for enclosing statements about them, and metalanguage for talking about the contexts, how they relate to one another, and how they relate to the world and all its events, states, and inhabitants.*

"*The advantage of Peircean semiotics is that it firmly situates language and logic within the broader study of signs of all types. The highly disciplined patterns of mathematics and logic, important as they may be for science, lie on a continuum with the looser patterns of everyday speech and with the perceptual and motor patterns, which are organized on geometrical principles that are very different from the syntactic patterns of language or logic.*"

Catherine Legg [20] notes that the semiotic process is really one of community involvement and consensus. Each understanding of a sign and each subsequent interpretation helps come to a consensus of what a sign means. It is a way of building a shared understanding that aids communication and effective interpretation. In Peirce's own writings, the process of interpretation can lead to validation and an eventual "canonical" or normative interpretation. The scientific method itself is an extreme form of the semiotic process, leading ultimately to what might be called accepted "truths".

Peircean Semiotics of URIs

So, how do Peircean semiotics help inform us about the role and use of URIs? Does this logic help provide guidance on the "identity crisis"?

The Peircean taxonomy of signs has three levels with three possible sign roles at each level, leading to a possible 27 combinations of sign representations. However, because not all sign roles are applicable at all levels, Peirce actually postulated only ten distinct sign representations.

Common to all roles, the URI "sign" is best seen as an *index*: the URI is a pointer to a representation of some form, be it electronic or otherwise. This representation bears a relation to the actual thing that this referent represents, as is true for all triadic sign relationships. However, in some contexts, again in keeping with additional signs interpreting signs in other roles, the URI "sign" may also play the role of a symbolic "name" or even as a signal that the resource can be downloaded or accessed in electronic form. In other words, by virtue of

the conventions that we choose to assign to our signs, we can supply additional information that augments our understanding of what the URI is, what it means, and how it is accessed.

Of course, in these regards, a URI is no different than any other sign in the Peircean world view: it must reside in a triadic relationship to its actual object and an interpretation of that object, with further understanding only coming about by the addition of further signs and interpretations.

In shortened form, this means that a URI, acting alone, can at most play the role of a pointer between an object and its referent. A URI alone, without further signs (information), can not inform us well about names or even what type of resource may be at hand. For these interpretations to be reliable, more information must be layered on, either by accepted convention of the current signs or the addition of still further signs and their interpretations. Since the attempts to deal with the nature of a URI resource by fiat as stipulated by httpRange-14 neither meet the standards of consensus nor empirical validity, the attempt can not by definition become "canonical". This does not mean that httpRange-14 and its recommended practices can not help in providing more information and aiding interpretation for what the nature of a resource may be. But it does mean that httpRange-14 acting alone is insufficient to resolve ambiguity.

Moreover, what we see in the general nature of Peirce's logic of signs is the usefulness of adding more "triads" of representation as the process to increase understanding through further interpretation. Kind of sounds like adding on more RDF triples, does it not?

Global is Neither Indiscriminate Nor Unambiguous

Names, references, identity and meaning are not absolutes. They are not philosophically, and they are not in human language. To expect machine communications to hold to different standards and laws than human communications is naive. To effect machine communications our challenge is not to devise new rules, but to observe and apply the best rules and practices that human communications instruct.

There has been an unstated hope at the heart of the semantic Web enterprise that simply expressing statements in the right way (syntax) and in the right form (RDF) is sufficient to facilitate machine communications. But this hope, too, is naive and silly. Just as we do not accept all human utterances as truth, neither will we accept all machine transmissions as reliable. Some of the information will be posted in error; some will be wrong or ill-fitting to our world view; some will be malicious or intended to deceive. Spam and occasionally lousy search results on the Web tell us that Web documents are subject to these sources of unsuitability, why is not the same true of data?

Thus, global data access via the semantic Web is not — and can never be — indiscriminate nor unambiguous. We need to understand and come to trust sources and provenance; we need interpretation and context to decide appropriateness and validity; and we need testing and validation to ensure messages as received are indeed correct. Humans need to do these things in their normal courses of interaction and communication; our machine systems will need to do the same.

These confirmations and decisions as to whether the information we receive is actionable or not will come about via still more information. Some of this information may come about via shared conventions. But most will come about because we choose to provide more context and interpretation for the core messages we hope to communicate.

A Go-Forward Approach

Nearly five years ago Hayes and Halpin put forth a proposal to add `ex:refersTo` and `ex:describedBy` to the standard RDF vocabulary as a way for authors to provide context and explanation for what constituted a specific RDF resource [11]. In various ways, many of the other individuals cited in this article have come to

similar conclusions. The simple redirect suggestions of both Ian Davis [10] and Ed Summers [12] appear particularly helpful.

Over time, we will likely need further representations about resources regarding such things as source, provenance, context and other interpretations that would help remove ambiguities as to how the information provided by that resource should be consumed or used. These additional interpretations can mechanically be provided via referenced ontologies or embedded RDFa (or similar). These additional interpretations can also be aided by judicious, limited additions of new predicates to basic language specifications for RDF (such as the Hayes and Halpin suggestions).

In the end, of course, any frameworks that achieve consensus and become widely adopted will be simple to use, easy to understand, and straightforward to deploy. The beauty of best practices in predicates and annotations is that failures to provide are easy to test. Parties that wish to have their data consumed have incentive to provide sufficient information so as to enable interpretation.

There is absolutely no reason that these additions can not co-exist with the current httpRange-14 approach. By adding a few other options and making clear the *optional* use of httpRange-14, we would be very Peirce-like in our go-forward approach: We are being both pragmatic while we add more means to improve our interpretations for what a Web resource is and is meant to be.

Acknowledgements

---

[1] Throughout intellectual history, a number of prominent philosophers and logicians have attempted to describe naming, identity and reference of objects and entities. Here are a few that you may likely encounter in various discussions of these topics in reference to the semantic Web; many are noted philosophers of language:

- Aristotle *(384 BC – 322 BC)* – founder of formal logic; formulator and proponent of categorization; believed in the innate "universals" of various things in the natural world
- Rudolf Carnap *(1891 – 1970)* – proposed a logical syntax that provided a system of concepts, a language, to enable logical analysis via exactly formula; a basis for natural language processing;rejected the idea and use of metaphysics
- René Descartes *(1596 – 1650)* – posited a boundary between mind and the world; the meaning of a sign is the intension of its producer, and is private and incorrigible
- Friedrich Ludwig Gottlob Frege *(1848 – 1925)* – one of the formulators of first-order logic, though syntax not adopted; advocated shared senses, which can be objective and sharable
- Kurt Gödel *(1906 – 1978)* – his two incompleteness theorems are some of the most important logic contributions of all time; they establish inherent limitations of all but the most trivial axiomatic systems capable of doing arithmetic, as well as for computer programs
- David Hume *(1711 – 1776)* – embraced natural empiricism, but kept the Descartes concept of an "idea"
- Immanuel Kant *(1724 – 1804)* – one of the major philosophers in history, argued that experience is purely subjective without first being processed by pure reason; a major influence on Peirce
- Saul Kripke *(1940 –)* – proposed the causal theory of reference and what proper names mean via a "baptism" by the namer
- Gottfried Wilhelm Leibniz *(1646 – 1716)* – the classic definition of identity is Leibniz's Law, which states that if two objects have all of their properties in common, they are identical and so only one object
- Richard Montague *(1930 – 1971)* – wrote much on logic and set theory; student of Tarski; pioneered a logical approach to natural language semantics; associated with model theory, model-theoretic semantics
- Charles Sanders Peirce *(1839 – 1914)* – see main text

- [Willard Van Orman Quine](#) *(1908 – 2000)* – noted analytical philosopher, advocated the "radical indeterminancy of translation" (can never really know)
- [Bertrand Russell](#) *(1872 – 1970)* – proposed the direct theory of reference and what it means to "ground in references"; adopted many Peirce arguments without attribution
- [Ferdinand de Saussure](#) *(1857 – 1913)* – also proposed an alternative view to Peirce of semiotics, one grounded in sociology and linguistics
- [John Rogers Searle](#) *(1932 –)* – argues that consciousness is a real physical process in the brain and is subjective; has argued against strong AI (artificial intelligence)
- [Alfred Tarski](#) *(1901 – 1983)* – analytic philosopher focused on definitions of models and truth; great admirer of Peirce; associated with model theory, model-theoretic semantics
- [Ludwig Josef Johann Wittgenstein](#) *(1889 – 1951)* – he disavowed his earlier work, arguing that philosophy needed to be grounded in ordinary language, recognzing that the meaning of words is dependent on context, usage, and grammar.

Also, [Umberto Eco](#) has been a noted proponent and popularizer of [semiotics.](#)

[2] As any practitioner ultimately notes, standards development is a messy, lengthy and trying process. Not all individuals can handle the messiness and polemics involved. Personally, I prefer to try to write cogent articles on specific issues of interest, and then leave it to others to slug it out in the back rooms of standards making. Where the process works well, standards get created that are accepted and adopted. Where the process does not work well, the standards are not embraced as exhibited by real-world use.

[3] Tim Berners-Lee, 2007. [What Do HTTP URIs Identify?](#)

This article does not discuss the other sub-category of URIs, [URNs](#) (for names). URNs may refer to any standard naming scheme (such as ISBNs for books) and has no direct bearing on any network access protocol, as do URLs and URIs when they are referenceable. Further, URNs are little used in practice.

[4] Kendall Clark was one of the first to question "resource" and other identity ambiguities, noting the tautology between URI and resource as "anything that has identity." See Kendall Clark, 2002. "Identity Crisis," in XML.com, Sept 11 2002; see [http://www.xml.com/pub/a/2002/09/11/deviant.html](#). From the topic map community, one notable contribution was from Steve Pepper and Sylvia Schwab, 2003. "Curing the Web's Identity Crisis," found at : [http://www.ontopia.net/topicmaps/materials/identitycrisis.html](#).

[5] Sandro Hawke, 2003. *[Disambiguating RDF Identifiers.](#)* W3C, January 2003. See [http://www.w3.org/2002/12/rdf-identifiers/](#).

[6] The issue was framed as what is the proper "range" for HTTP referrals and was also the 14th major TAG issue recorded, hence the name. See further the [httpRange-14 Webography](#) .

[7] See W3C, "httpRange-14: What is the range of the HTTP dereference function?"; see [http://www.w3.org/2001/tag/issues.html#httpRange-14](#).

[8] See [http://www.ibiblio.org/hhalpin/irw2006/](#).

[9] Leo Sauermann and Richard Cyganiak, eds., 2008. *Cool URIs for the Semantic Web*, W3C Interest Group Note, December 3, 2008. See [http://www.w3.org/TR/cooluris/](#).

[10] Ian Davis, 2010. *[Is 303](#)* R*[eally Necessary?](#)* Blog post, November 2010, accessed 20 January 2012. (See http://blog.iandavis.com/2010/11/04/is-303-really-necessary/.) A considerable thread resulted from this post; see [http://markmail.org/thread/mkoc5kxll6bbjbxk](#).

[11] See first Harry Halpin, 2006. "Identity, Reference and Meaning on the Web," presented at WWW 2006, May 23, 2006. See [http://www.ibiblio.org/hhalpin/irw2006/hhalpin.pdf](#). This was then followed up with greater elaboration by Patrick J. Hayes and Harry Halpin, 2007. "In Defense of Amibiguity," [http://www.ibiblio.org/hhalpin/homepage/publications/indefenseofambiguity.html](#).

[12] Ed Summers, 2010. Linking Things and Common Sense, blog post of July 7, 2010. See [http://inkdroid.org/journal/2010/07/07/linking-things-and-common-sense/](#).

[13] Xiaoshu Wang, 2007. [URI Identity and Web Architecture Revisited](#), Word document posted on posterous.com, November 2007. (Former Web documents have been removed.)

[14] David Booth, 2006. "URIs and the Myth of Resource Identity," see [http://dbooth.org/2006/identity/](#).

[15] See Larry Masinter, 2012. "The 'tdb' and 'duri' URI Schemes, Based on Dated URIs," 10th version, *IETF Network Working Group Internet-Draft*,January 12, 2012. See [http://tools.ietf.org/html/draft-masinter-dated-uri-10](#).

[16] Jonathan Rees has been the scribe and author for many of the background documents related to Issue 57. A recent mailing list entry provides pointers to four relevant documents in this entire discussion. See Jonathan A Rees, 2012. [Guide to ISSUE-57 (httpRange-14) document suite](#)January, 21, 2012.

[17] At least twenty domain names, led by insure.com, have sold for more the $2 million each; see this [Wikipedia listing](#).

[18] In the wonderful movie, *The Gods, They Must be Crazy*, Bushmen in the Kalahari Desert one day find an unbroken glass Coke bottle that had been thrown out of an airplane. Initially, this strange artifact seems to be another boon from the gods, and the Bushmen find many uses for it. But unlike anything that they have had before, there is only one bottle to go around. This creates jealousy, envy, anger, hatred, even violence. The protagonist, Xi, decides that the bottle is an evil thing and must be thrown off of the edge of the world. The hilarity of the movie comes from that premise and Xi's encounters with the modern world as he pursues his quest with the magic bottle.

[19] Wang [13] rhetorically asked which of the following things would be categorized as an "information resource":

1. A book
2. A clock
3. The clock on the wall of my bedroom
4. A gene
5. The sequence of a gene
6. A software
7. A service
8. A namespace
9. An ontology
10. A language
11. A number
12. A concept, such as Dublin Core's creator.

See the 2007 thread on this issue, mostly by Sean Palmer and Noah Mendelsohn, the latter aknowledging that various experts may only agree on 85% of the items.

[20] See further Catherine Legg, 2010. "Pragmaticsm on the Semantic Web," in Bergman, M., Paavola, S., Pietarinen, A.-V., & Rydenfelt, H. eds., *Ideas in Action: Proceedings of the Applying Peirce Conference*, pp. 173–188. Nordic Studies in Pragmatism 1. Helsinki: Nordic Pragmatism Network. See http://www.nordprag.org/nsp/1/Legg.pdf.

[21] Charles Sanders Peirce, 1894. "What is in a Sign?", see http://www.iupui.edu/~peirce/ep/ep2/ep2book/ch02/ep2ch2.htm.

[22] The figures in particular are from John F. Sowa, 2000. "Ontology, Metadata, and Semiotics," presented at ICCS 2000 in Darmstadt, Germany, on August 14, 2000; published in B. Ganter & G. W. Mineau, eds., *Conceptual Structures: Logical, Linguistic, and Computational Issues, Lecture Notes in AI #1867*, Springer-Verlag, Berlin, 2000, pp. 55-81. May be found at http://www.jfsowa.com/ontology/ontometa.htm. Also see John F. Sowa, 2006. "Peirce's Contributions to the 21st Century," presented at *International Conference on Conceptual Structures*, Aalborg, Denmark, July 17, 2006. See http://www.jfsowa.com/pubs/csp21st.pdf.

[23] C.K. Ogden and I. A. Richards, 1923. *The Meaning of Meaning*, Harcourt, Brace, and World, New York, 8th edition 1946.