# Representation and resolution of digital resources on the semantic web – an ontology-based approach

Matthias Samwald, Alan Ruttenberg

The use of Unique Resource Identifiers (URIs) to identify resources is fundamental to the architecture of the Semantic Web. However, when using URIs in ontologies the community has not yet decided if all or some URIs should be resolvable, how such resolution mechanisms should work, what kind of data should be retrieved, and what the semantic relations between the resource identified by the URI and the retrieved information should be.

Within the life science community, there are two solutions proposed: 1) All URLs, when resolvable, are retrieved using HTTP, possibly using content negotiation and 2) Use Life Science Identifiers (LSID) instead of HTTP URIs and process the retrieval of information by consulting an LSID resolver, which informs an agent of appropriate methods of retrieving data and metadata associated with the URI.

Neither of these solutions address the question of which URIs should be resolvable, and each suffers from technical problems impeding adoption – HTTP servers frequently become unavailable, and there is no defined mechanism for retrieving the information from an alternate resource. LSID has infrastructure requirements that are not widely deployed, namely the resolvers and the client software needed to interact with them.

We propose instead that the information needed to handle this problem be directly embedded in the ontology. First, an explicit distinction needs to be made instances that represent information resources and those that do not, and suggest that this distinction be made by making them instances of two disjoint classes. Only for information resources do we expect to be able to go out to the web and retrieve further information. Second, we proposed that the type and method of resolution for information resources be encoded as property values of those instances. Thus we make the distinction between:

- the thing itself (e.g. *the patient John Doe*)
- the digital representation of a thing – an information resource (e.g. *a radiograph of John Doe*)
- A method for taking the URI representing the information resource, *(e.g. contact an HTTP based LSID resolver)*

We will present examples of such an ontology, demonstrating suggested patterns for handling some common cases that have been raised in the W3C Health Care and Life Sciences interest group: Digital resources that can be retrieved from more than one server, retrieved by more than one protocol, that are versioned and those that contain additional triples to augment a graph.