

Appendix 2: A Glossary for Semantic Technology (with thanks to the Ensemble technology developers for their help).

Appendix 2	GLOSSARY FOR SEMANTIC TECHNOLOGY
Abbreviation	Meaning
API	An application programming interface (API) is an interface implemented by a software program that enables it to interact with other software. It facilitates interaction between different software programs similar to the way the user interface facilitates interaction between humans and computers. (Wikipedia, accessed 6 th Oct 2010)
Babel	See ‘conversion tools’
Conversion tools	Tools that ‘translate’ digital (human readable) data into machine readable formats (into RDF - triplets). Various types, like Babel, or RDF123
Data object	digital object with resources associated that normally belongs to a particular collection; object that presents the actual contents
DDI	DDI (Data Documentation Initiative, an extension of Dublin Core oriented to Social Sciences research data
Digital object	An item as stored in a digital library, consisting of data, metadata, and an identifier. www.cs.cornell.edu/wya/DigLib/MS1999/Glossary.html Fedora digital objects can be related to other Fedora objects in many ways. For example there may be a Fedora object that represents a collection and other objects that are members of that collection. Also, it may be the case that one object is considered a part of another object, a derivation of another object, a description of another object, or even equivalent to another object. http://fedora-commons.org/download/2.1.1/userdocs/digitalobjects/introRelsExt.html
Digital repository	See Fedora.
Exhibit	A semantic web application that can ‘read’ RDF, and aggregate data from diverse sources; allows visualization of data or querying data across vast data sets. Exhibit only one type of semantic web application. SIMILE toolkit–developed by MIT.
External reference	External referenced content is similar to the first one with the difference that the contents of the data stream are not stored internally in the Fedora repository but come from a different server or web site.
Facet	A search facility in Exhibit, allows sorting data according to a selected aspect or feature of the phenomenon.
Faceted search	See ‘facet’
Fedora	Digital repository that can contain vast varieties of different types of digital data, as well as metadata in RDF.
Inline XML data	The last type of datastream is inline xml data. This type of datastream is really useful because it enables to store different metadata schemas within the same digital object and the different datastreams included in the object can be described using those schemas. Therefore we can store contents stored in RDF format that can be inserted into the triplestore. Examples of this kind of usage are provided with the plant distribution demonstrator where one collection of digital objects is stored into Fedora repository. The contents are spreadsheets with the original data and RDF versions of them are provided as

Appendix 2: A Glossary for Semantic Technology (with thanks to the Ensemble technology developers for their help).

	well. This data can be now stored in the triplestore and presented easily using visualization tools.
JSON	<p>JavaScript Object Notation, is a lightweight text-based open standard designed for human-readable data interchange. It is derived from the JavaScript scripting language for representing simple data structures and associative arrays, called objects.</p> <p>The JSON format is often used for serializing and transmitting structured data over a network connection. It is used primarily to transmit data between a server and web application, serving as an alternative to XML.</p> <p>(Wikipedia accessed 5th October 2011)</p>
MYSQL	A popular open source data base.
Mulgara Triplestore	Metadata manager – data imported and exported in RDF triplets (i.e. in Machine readable formats). Communicates with Semantic web applications.
Ontology	Formal representation of knowledge as a set of concepts within a domain and the relationship between these concepts; provides a shared vocabulary, which can be used to model a domain.
RDF	Resource definition framework. Part of the W3 consortium metadata data model. RDF is based on the idea of making statements about resources in the form of subject-predicate-object (e.g. the car is red), also known as <i>triples</i> . The subject denotes the resource, and the predicate denotes traits or aspects of the resource and expresses a relationship between the subject and the object. (http://en.wikipedia.org/wiki/Resource_Description_Framework accessed 25.10.11)
Repository	A digital storage for data or meta data.
The Semantic Web	The Semantic web is about common formats of data drawn from diverse sources; about encoding different types of data into common, machine readable formats; the internet would become one vast data base.
SIMILE	MIT built toolkit for building Exhibits. See ‘exhibit’.
SPARQL	RDF query language; can be used to query e.g. Mulgara Triplestore; Standardized by W3C; considered key semantic technology; allows for the query to consist of triple patterns, conjunctions; disjunctions and optional patterns – this allows writing of globally unambiguous queries.
Triple Store	See Mulgara Triplestore.
Triplets	Data expressed in subject-predicate-object format; unique formats that are understandable by ‘machines’; define relationship between data (see RDF).
URI	Unique Resource Identifier (given to data in the conversion process into RDF language or html, XML).
URL	Universal resource locator.
XML	Extensible Mark-up language, used with semantic technologies.