

Acronyms

Readers are referred to the glossary for definitions.

- ABARES Australian Bureau of Agricultural and Resource Economics and Sciences
- ABS Australian Bureau of Statistics
- ACT Australian Capital Territory
- ACVC Australian Climate Variability & Change
- AEMET Agencia Estatal de Meteorologica (Spanish Meteorological Office)
- AIHW Australian Institute of Health and Welfare
- ANDS Australian National Data Service
- API Application Programmer Interface
- BOM (Australian) Bureau of Meteorology
- CSH Canadian Subject Headings
- CSIRO (Australian) Commonwealth Scientific and Industrial Research Organization
- CST Government of Canada Core Subject Thesaurus
- CSV Comma-Separated Value data format
- D2RQ Database to RDF Queueing
- DAG Directed Acyclic Graph
- DC Dublin Core Element Set
- DCMI Dublin Core Metadata Initiative
- DDC Dewey Decimal Classification
- DFID Department for International Development of the United Kingdom
- DNS Domain Name System
- DOAP Description of a Project vocabulary
- DOI Digital Object Identifier
- DSEWPC Australian Department of Sustainability, Environment, Water, Population and Communities
- DTD Document Type Definition
- EAD Encoded Archival Description
- e-Gov Electronic Government
- ETL Extract, Transform and Load
- EPA U.S. Environmental Protection Agency
- FDsys Federal Digital System
- FAR U.S. Federal Acquisition Regulations
- FLOSS Free/Libre/Open Source Software
- FOAF Friend of a Friend ontology
- FOIA Freedom of Information Act of the United States
- GC Government of Canada

GIS Geographic Information Systems
 GML Geographic Markup Language
 GORs Government Office Regions of England
 GPO U.S. Government Printing Office
 GSA General Services Administration (U.S.), Geoscience Australia (Australia)
 GSP Gross State Product
 GRDDL Gleaning Resource Descriptions from Dialects of Languages
 HTML Hypertext Markup Language
 HTTP Hypertext Transfer Protocol
 ICT Information and Communications Technology
 IETF Internet Engineering Task Force
 IGN National Geographic Institute of Spain
 IMOS Integrated Marine Observing System of Australia
 INE National Statistic Institute of Spain
 ISO International Standards Organization
 IT Information Technology
 KOS Knowledge Organization System
 LCC Library of Congress Classification
 LCSH Library of Congress Subject Headings
 LED Linking Enterprise Data
 LOD Linked Open Data
 LOGD Linking Open Government Data portal at TWC
 MARC MACHine-Readable Cataloging system
 MIME Multipurpose Internet Mail Extensions
 MODS Metadata Object Description Schema
 MRRS Management, Resources and Results Structure
 N3 Notation 3
 OGD Open Government Data
 OWL Web Ontology Language
 PAA Program Activity Architecture
 POTUS President of the United States
 PSI Public Sector Information
 R2RML Relational to RDF Markup Language
 RDB Relational Database
 RDF Resource Description Framework
 RDFa Resource Description Framework Attributes
 RDFS Resource Description Framework Schema
 RDF/XML Resource Description Framework eXtensible Markup Language serialization format
 RMMS Government of Canada Records Management Metadata Standard
 RVM Répertoire de vedettes-matière
 REST Representational State Transfer
 RFC Request for Comments
 RPI Rensselaer Polytechnic Institute
 SCOVO Statistical Core Vocabulary
 SKOS Simple Knowledge Organisation System
 SOA Service Oriented Architecture
 SOAP Simple Object Access Protocol
 SPARQL SPARQL Protocol and RDF Query Language
 SQL Structured Query Language
 SSN Semantic Sensor Network Ontology
 TSV Tab-Separated Value data format
 TWC Tetherless World Constellation at RPI
 URI Uniform Resource Indicator
 URL Uniform Resource Locator

USDA United States Department of Agriculture
USAID United States Agency for International Development
UUID Universally Unique Identifier
VoID Vocabulary of Interlinked Datasets
W3C World Wide Web Consortium
WCED World Commission on Environment and Development
WEHAB Water, Energy, Health, Agriculture and Biodiversity framework of the United Nations
XHTML Extensible Hypertext Markup Language
XML Extensible Markup Language
XSD XML Schema
XSLT Extensible Stylesheet Language Transformations

Glossary

Readers are encouraged to also see the acronym list.

Abox One of two types of statements in an ontology (the other being Tbox). Abox statements represent facts (or “assertions”, hence the “A”), e.g. John is a Person (where Person is a defined class).

Application Programmer Interface An abstraction implemented in software that defines how others should make use of a software package such as a library or other reusable program.

Backward chaining A method of reasoning that begins with a conclusion being sought and works backward to determine if any data supports that conclusion. Backward chaining is the reverse of forward chaining.

Big O notation A definition of the worst case performance of a mathematical function, often a computer algorithm.

Canadian Subject Headings An English language thesaurus on Canada and Canadian topics for use in bibliographic records, maintained by Library and Archives Canada. Compare to RVM and LCSH.

Closed world The presumption that what is not known to be true must be false. The assumption underlying relational databases, most forms of logical programming, OWL DL and OWL Lite.

Controlled vocabularies Carefully selected sets of terms that are used to describe units of information; used to create thesauri, taxonomies and ontologies.

Database to RDF Queueing A mechanism to query information in traditional management systems such as relational databases via the SPARQL query language. D2RQ may refer to the language definition or the Open Source Software project.

Data warehouse A storage and retrieval system for enterprise information designed to centralize information from other stores to facilitate cross-system querying and reporting.

DBPedia An RDF representation of the metadata held in Wikipedia and made available for SPARQL query on the World Wide Web.

Dewey Decimal Classification A commonly used proprietary system of library classification, currently maintained by The Online Computer Library Center (OCLC).

Directed Acyclic Graph A directed graph (like RDF) with the additional restriction that no loops or cycles are permitted. A cycle is a path from a given node that would allow one to find their way back to the starting node.

Directed graph A graph in which the links between nodes are directional (they only go from one node to another). RDF represents things (nouns) and the relationships between them (verbs) in a directed graph. In RDF, the links are differentiated by being assigned URIs.

Document Type Definition A type of schema for defining a markup language, such as in XML or HTML (or their predecessor SGML).

Domain Name System The Internet's mechanism for mapping between a human-readable host name (e.g. www.example.com) and an Internet Protocol (IP) Address (e.g. 203.20.51.10).

Dublin Core Element Set A vocabulary of fifteen properties for use in resource descriptions, such as may be found in a library card catalog (author, publisher, etc). The most commonly used vocabulary for Semantic Web applications.

Dublin Core Metadata Initiative An open international organization engaged in the development of interoperable metadata standards, including the Dublin Core Element Set.

Encoded Archival Description A standard for encoding archival finding aids using Extensible Markup Language (see XML), maintained by a partnership of the U.S. Library of Congress and the Society of American Archivists.

Enterprise For the purposes of this book, any human organization that uses computer systems to store, retrieve and analyze information.

Federal Digital System An electronic information management system to authenticate, preserve, version, and provide permanent public access to U.S. federal government information operated by the U.S. GPO (see GPO).

Forward chaining A method of reasoning that begins with statements of all the relevant facts and infers new facts based on a set of rules. Equivalent to the logical operation *modus ponens*. The reverse of forward chaining is backward chaining.

Free/Libre/Open Source Software A generic and internationalized term for software released under an Open Source license.

Friend of a Friend A Semantic Web vocabulary describing people and their relationships for use in resource descriptions.

Government of Canada Records Management Metadata Standard A metadata element set for records management systems used in the Government of Canada.

Government of Canada Core Subject Thesaurus A bilingual (French and English) thesaurus of terms used in information resources of the Government of Canada.

Graph A collection of objects (represented by "nodes") any of which may be connected by links between them. See directed graph.

Hypertext Markup Language The predominant markup language for hypertext pages on the Web. HTML defines the structure of Web pages. A family of W3C standards.

Hypertext Transfer Protocol The standard transmission protocol used on the World Wide Web to transfer hypertext requests and information between Web servers and Web clients (such as browsers). An IETF standard.

International Standards Organization A network of the national standards institutes of 162 countries that cooperate to define international standards. Defines many standards including in the context of this book formats for dates and currency.

Internet Engineering Task Force An open international community concerned with the evolution of Internet architecture and the operation of the Internet. Defines standards such as HTTP and DNS.

Jena An Open Source Software implementation of a Semantic Web development framework. Supports the storage, retrieval and analysis of RDF information. See <http://openjena.org> and compare to Mulgara and Sesame.

Knowledge Organization System A tool or system that provides an organized interpretation of knowledge structures.

Library of Congress Classification A system of library classification developed and maintained by the U.S. Library of Congress.

Library of Congress Subject Headings A thesaurus of subject headings for use in bibliographic records, maintained by the U.S. Library of Congress.

Linked data A pattern for hyperlinking machine-readable data sets to each other using Semantic Web techniques, especially via the use of RDF and URIs. Enables distributed SPARQL queries of the data sets and a “browsing” or “discovery” approach to finding information (as compared to a search strategy).

Linked Open Data An open community project to interlink data on the Semantic Web using URIs and RDF.

Linking Government Data The use of tools and techniques of the Semantic Web to connect, expose and use data from government systems.

MACHINE-Readable Cataloging system A family of formats for the representation of bibliographic information in libraries (ISO 2709, ANSI/NISO Z39.2).

Management, Resources and Results Structure Government of Canada policy linking the management of government resources and program results to their organizational structures. It provides a representation of how a department is managed through a) strategic outcomes; b) program activity architecture (PAA); and c) governance structure. See also PAA.

Metadata Information used to administer, describe, preserve, present, use or link other information held in resources, especially knowledge resources, be they physical or virtual.

Metadata Object Description Schema A bibliographic description system intended to be a compromise between MARC and DC metadata. Implemented in XML Schema (see DC, MARC, XSD).

Mulgara An Open Source Software implementation of an RDF database. Supports the storage, retrieval and analysis of RDF information. See <http://mulgara.org> and compare to Jena and Sesame.

Multipurpose Internet Mail Extensions A means of representing binary content in textual messages, such as in electronic mail and HTTP. An IETF standard.

Notation 3 An RDF syntax intended to be readable by humans. See also Turtle.

Ontology A formal representation of relationships between items in a directed graph structure. See taxonomy.

Open world The presumption that what is not known to be true may yet be true if additional information is later obtained. The assumption underlying RDF and OWL Full.

Pattern A general reusable approach to solving a commonly occurring type of problem.

Program Activity Architecture A component of a MRRS. It provides a hierarchical linkage for all departmental programs and activities to a department’s strategic outcomes. It documents how Government of Canada departments allocates and manages its resources to achieve their intended results. See also MRRS.

Protocol A set of instructions for transferring data from one computer to another over a network. A protocol standard defines both message formats and the rules for sending and receiving those messages.

Public Sector Information Information created by a government in the course of governing.

Quad store A colloquial phrase for an RDF database that stores RDF triples plus an additional element of information, often used to collect statements into groups.

RDF database A type of database designed specifically to store and retrieve RDF information.

Répertoire de vedettes-matière A partially bilingual (English and French) thesaurus of terms in French hosted and developed by Université Laval. It is complemented by the Canadian Subject Headings (CSH) developed by Library and Archives Canada.

Representational State Transfer An architectural style for information systems used to greater or lesser degree on the Web and explains some of the Web's key features, such as extreme scalability and robustness to change.

Request for Comments A document submitted to the IETF. Internet standards started as RFCs and are often referenced by their RFC numbers.

Resource Description Framework RDF: An international standard for data interchange on the Web. A W3C standard.

Resource Description Framework Attributes An RDF syntax encoded in HTML documents. A W3C standard.

Resource Description Framework Schema The simplest RDF vocabulary description language that provides much less descriptive capability than SKOS or OWL. A W3C standard.

Resource Description Framework eXtensible Markup Language serialization format An RDF syntax encoded in XML. A W3C standard.

Schema A data model that represents the relationships between a set of concepts. Some types of schemas include relational database schemas (which define how data is stored and retrieved), taxonomies and ontologies.

Semantic technologies The broad set of technologies that relate to the extraction, representation, storage, retrieval and analysis of machine-readable information. The Semantic Web standards are a subset of semantic technologies and techniques.

Semantic Web An evolution or part of the World Wide Web that consists of machine-readable data in RDF and an ability to query that information in standard ways (e.g. via SPARQL).

Semantic Web standards Standards of the World Wide Web Consortium (W3C) relating to the Semantic Web, including RDF, RDFa, SKOS and OWL.

Service Oriented Architecture A set of architectural design guidelines used to expose services, often as Web Services.

Sesame An Open Source Software implementation of a Semantic Web development framework. Supports the storage, retrieval and analysis of RDF information. See <http://www.openrdf.org> and compare to Jena and Mulgara.

Simple Knowledge Organisation System A vocabulary description language for RDF designed for representing traditional knowledge organization systems such as enterprise taxonomies in RDF. A W3C standard.

Simple Object Access Protocol A protocol over HTTP for exchanging structured information in XML to and from Web Services.

SPARQL Protocol and RDF Query Language A query language standard for RDF data on the Semantic Web; analogous to the Structured Query Language (SQL) for relational databases. A W3C standard.

Structured Query Language A query language standard for relational databases.

Taxonomy A formal representation of relationships between items in a hierarchical structure. See ontology.

Tbox One of two types of statements in an ontology (the other being Abox). Tbox statements describe a knowledge system in terms of controlled vocabularies (or “terminology”, hence the “T”), e.g. A Person is a Mammal.

Term For the purposes of this book, an entry in a controlled vocabulary, schema, taxonomy or ontology.

Triple An RDF statement, consisting of two things (a “subject” and an “object”) and a relationship between them (a verb, or “predicate”). This subject-predicate-object triple forms the smallest possible RDF graph (although most RDF graphs consist of many statements).

Triple store A colloquial phrase for an RDF database that stores RDF triples.

Tuple An ordered list of elements. RDF statements are 3-tuples; an ordered list of three elements.

Turtle An RDF serialization format, designed to be easier to read than others such as RDF/XML. A subset of N3.

Uniform Resource Indicator A global identifier for the Web standardized by joint action of the W3C and IETF. A URI may or may not be resolvable on the Web (see URL).

Uniform Resource Locator A global identifier for Web resources standardized by joint action of the W3C and IETF. A URL is resolvable on the Web and is commonly called a “Web address”.

Universally Unique Identifier A large hexadecimal number that may be calculated by anyone without significant central coordination and used to uniquely identify a resource. A standard of the Open Software Foundation.

U.S. Government Printing Office An agency of the U.S. Congress tasked with printing and making available documents of the U.S. Government.

Web 2.0 A colloquial description of the part of the World Wide Web that implements social networking, blogs, user comments and ratings and related human-centered activities.

Web 3.0 A colloquial description of the part of the World Wide Web that implements machine-readable data and the ability to perform distributed queries and analysis on that data. Considered synonymous with the phrases “Semantic Web” and “The Web of Data”.

Web Ontology Language A family of knowledge representation and vocabulary description languages for authoring ontologies, based on RDF and standardized by the W3C. Standardized variants include OWL Full, OWL DL (for “description logic”) and OWL Lite.

World Wide Web Consortium An international community that develops standards for the World Wide Web. Defines standards such as HTML, XML and RDF.

eXtensible Hypertext Markup Language A family of versions of HTML based on XML and standardized by the W3C.

eXtensible Markup Language A specification for creating structured textual computer documents. Many thousands of XML formats exist, including XHTML. A family of standards from the W3C.

XML Schema Limitations on the content of an XML document that defines what structural elements are allowed.

eXtensible Stylesheet Language Transformations Declarative programs to transform one XML document into another XML document.

Index

- 4Store, 37, 215
- 5-star, 7, 19, 29, 138–140, 146

- ABox, 32, 41, 45
- Altova Semantic Works, 34
- Australia, ix, 4, 51, 156, 162, 163, 181–183, 185–199, 201–203
- Azure, 69

- Baidu, viii
- Barak Obama, 4
- Basic Geo, 15, 118
- Beth Simone Noveck, 4
- Bibliographic Ontology, 15, 118
- Bolivia, 32
- Brazil, ix, 129, 131

- Callimachus, 10, 109
- Canada, ix, 4, 155–157, 160–168, 170, 176
- CKAN, 22, 36, 38, 44, 47, 69, 96, 97, 99, 108
- cookbook, viii, 3, 5, 6, 24
- Creative Commons, 10, 15, 101, 118
- Creative Commons License, 33
- Crown Copyright, 23
- CSIRO, 188, 203
- CSV, 7, 11, 35, 44, 46, 66, 96, 131, 139, 146, 147, 175, 188, 194–197, 202, 211, 213, 216

- David Cameron, 4, 185
- DBpedia, 10, 43, 67, 68, 88, 100, 140, 213, 214
- DCMI, 16, 117
- DFID, 215, 216
- Digital Enterprise Research Institute, ix
- Disco, 39, 148, 201
- DOAP, 15, 118

- Drupal, 10, 54, 115, 218
- Dublin Core, 10, 14, 58, 117, 160, 212

- Federal Acquisition Regulations, 23
- Finland, ix, 4, 97, 106, 110
- First International Open Government Conference, 7
- FOAF, 10, 14–16, 58, 66, 117, 124, 127

- Geonames, 43, 88
- Geospatial data, 7
- Good Relations Ontology, 16
- Google, viii, 38, 44, 54, 96, 102, 140, 193, 207, 218
- Google Refine, 35, 146, 147
- Gordon Brown, 4, 7
- GovTrack, 67, 69
- GRDDL, 35
- Greece, 4, 51

- HTTP, 9, 10, 13, 20, 32, 36, 65, 96, 114, 115, 147, 148

- Ireland, ix, 75, 141, 143

- Jay Myers, 4
- Jena, 37, 66, 104, 115
- Joseki, 37, 104
- Julia Child, 9, 10

- Library and Archives Canada, ix
- LIMES, 36
- Linked Open Data, 3, 6
- Linked Open Data principles, 6
- Linked Open Data Project, vii
- LinkSailor, 39
- LOV, 33

- MIME, 20

- Nicholas Negroponte, 9
 Norway, 4, 51
- OAI, 16, 118, 163
 Open Data Commons Attribution License, 33
 Open Database License, 33
 Open Government Data, 13, 16, 27, 29, 31, 38, 39, 51, 52, 58, 67, 69, 70, 113, 129, 183, 192, 194, 211, 217
 Open Government License, 33
 Open Knowledge Foundation, 22, 29, 108
 OpenStreetMap, 115
 OWLIM, 37
- Pedantic Web Group, 36
 Provenance, 37, 58, 66, 158
 Pubby, 37, 43, 47
 Public Domain Dedication and License, 33
- R2RML, 35, 115, 128
 RDF Recommendation, 14
 RDFa, 4, 18, 22, 218
 Rensselaer Polytechnic Institute, viii, ix, 13, 51
 REST, 19, 52, 69, 205, 219
- SchemaCache, 33
 SchemaWeb, 33
 Sesame, 37
 SILK, 36, 43, 46
 Sindice, 22, 38, 44, 47
 SIOC, 16, 66, 117
 Sitemap, 38
 SKOS, 15, 143, 144, 150, 160, 161, 175, 212
 Socrata, 52, 69, 108
 Spain, ix, 4, 39, 40
 SPARQL, 19–22, 37, 38, 54, 56, 58, 61, 66–68, 78, 82, 84, 95, 98, 100, 101, 104, 106, 107, 114–116, 147–149, 184, 194, 202, 211, 213, 215, 216, 218
- SQL, 8, 21, 115, 116, 131
 Sunlight Foundation, 52, 69
 Swoogle, 22, 33, 96
- Tabulator, 39
 TBox, 32, 41, 45
 Tim Berners-Lee, 4, 7, 10, 18, 29, 138–140, 158
 TopBraid, 34, 35
- U.S. Government Work, 23
 UK Cabinet Office, 12, 32
 UN Fundamental Principles of Official Statistics, 136
 United Kingdom, viii, ix, 4, 28, 51, 75, 96, 156, 162, 214, 215
 United States, ix, 4, 23, 28, 51, 61, 96, 166, 175, 206, 211, 212, 215
 USAID, 215, 216
- vCard, 15, 16
 Virtuoso, 35, 37, 43, 47, 54, 114, 128
 Vivek Kundra, 5
 VoID, 15–17, 22, 23, 37, 44, 47, 54, 58, 59, 62, 63
- W3C eGovernment Activity, 6, 28
 W3C Government Linked Data Working Group, 6, 19, 22, 23
 W3C Interest Group on eGovernment, 15
 W3C RDB2RDF Working Group, 115
 Web of Data, 39, 44, 54, 73, 74, 83, 84, 89
 Wikipedia, 10, 68, 140, 212, 213
 World Wide Web Consortium, vii, 3, 42
- Yahoo, viii
 Yandex, viii