



# Linked Data Glossary

## W3C Working Group Note 27 June 2013

**This version:**

<http://www.w3.org/TR/2013/NOTE-ld-glossary-20130627/>

**Latest published version:**

<http://www.w3.org/TR/ld-glossary/>

**Previous version:**

**Editors:**

[Bernadette Hyland](#), [3 Round Stones](#)

[Ghislain Atemezing](#), [EURECOM](#)

Michael Pendleton, [US Environmental Protection Agency](#)

Biplav Srivastava, [IBM](#)

Copyright © 2013 W3C® ([MIT](#), [ERCIM](#), [Keio](#), [Beihang](#)), All Rights Reserved. W3C [liability](#), [trademark](#) and [document use](#) rules apply.

---

## Abstract

This document is a glossary of terms defined and used to describe Linked Data, and its associated vocabularies and [Best Practices](#). This document published by the [W3C Government Linked Data Working Group](#) as a Working Group Note, is intended to help information management professionals, Web developers, scientists and the general public better understand publishing structured data using [Linked Data Principles](#).

## Status of This Document

*This section describes the status of this document at the time of its publication. Other documents may supersede this document. A list of current W3C publications and the latest revision of this technical report can be found in the [W3C technical reports index](#) at <http://www.w3.org/TR/>.*

This document was published by the [Government Linked Data Working Group](#) as a Working Group Note. If you wish to make comments regarding this document, please send them to [public-gld-comments@w3.org](mailto:public-gld-comments@w3.org) ([subscribe](#), [archives](#)). All comments are welcome.

Publication as a Working Group Note does not imply endorsement by the W3C Membership. This is a draft document and may be updated, replaced or obsoleted by

other documents at any time. It is inappropriate to cite this document as other than work in progress.

This document was produced by a group operating under the [5 February 2004 W3C Patent Policy](#). W3C maintains a [public list of any patent disclosures](#) made in connection with the deliverables of the group; that page also includes instructions for disclosing a patent. An individual who has actual knowledge of a patent which the individual believes contains [Essential Claim\(s\)](#) must disclose the information in accordance with [section 6 of the W3C Patent Policy](#).

## Table of Contents

1. [5 Star Linked Open Data](#)
2. [Apache License](#)
3. [API](#)
4. [CC-BY-SA License](#)
5. [Closed World](#)
6. [Connection](#)
7. [Conneg](#)
8. [Content Negotiation](#)
9. [Controlled Vocabulary](#)
10. [Comma Separated Values \(CSV\)](#)
11. [Creative Commons Licenses](#)
12. [CURIEs](#)
13. [cURL](#)
14. [Data Cloud](#)
15. [Data Hub, The](#)
16. [Data Market](#)
17. [Data Modeling](#)
18. [Dataset, RDF](#)
19. [Data Warehouse](#)
20. [DBpedia](#)
21. [Dereferenceable URIs](#)
22. [Description Logic](#)
23. [DCAT](#)
24. [DCMI](#)
25. [Directed Graph](#)
26. [Document Type Definition](#)
27. [Domain Name System \(DNS\)](#)
28. [Dublin Core Metadata Element Set](#)
29. [Dublin Core Metadata Initiative](#)
30. [Dublin Core Metadata Terms](#)
31. [Entity](#)
32. [ETL](#)
33. [FOAF](#)
34. [Fragment Identifier](#)
35. [Free/Libre/Open Source Software](#)
36. [Friend of a Friend](#)
37. [Government Open Data](#)
38. [Graph](#)
39. [HyperText Markup Language](#)

40. [HyperText Transfer Protocol](#)
41. [HTTP URIs](#)
42. [Inference](#)
43. [International Organization for Standards \(ISO\)](#)
44. [Internet Engineering Task Force \(IETF\)](#)
45. [Internationalized Resource Identifier](#)
46. [JSON](#)
47. [JSON-LD](#)
48. [Linked Data](#)
49. [Linked Data API](#)
50. [Linked Data client](#)
51. [Linked Data Platform](#)
52. [Linked Data Principles](#)
53. [Linked Open Data](#)
54. [Linked Open Data Cloud](#)
55. [Linked Open Data Cloud diagram](#)
56. [Linking Open Data Project](#)
57. [Linkset](#)
58. [Machine Readable Data](#)
59. [Message](#)
60. [Metadata](#)
61. [Metadata Object Description Schema](#)
62. [Modeling Process](#)
63. [N3](#)
64. [Namespace IRI](#)
65. [Natural Keys](#)
66. [Neutral URI](#)
67. [N-Triples](#)
68. [Object](#)
69. [Ontology](#)
70. [Open Government Data](#)
71. [ORG Ontology](#)
72. [Persistent Identifier Scheme](#)
73. [Persistent Uniform Resource Locator](#)
74. [Predicate](#)
75. [Protocol](#)
76. [Provenance](#)
77. [PURL](#)
78. [Quad Store](#)
79. [Query](#)
80. [R2RML](#)
81. [Raw Data](#)
82. [RDF](#)
83. [RDFa](#)
84. [RDF database](#)
85. [RDF-JSON](#)
86. [RDF Schema](#)
87. [RDF/XML](#)
88. [Representational State Transfer \(REST\)](#)
89. [Request](#)
90. [Resource](#)

91. [Resource Description Framework \(RDF\)](#)
92. [Resource Description Framework in attributes \(RDFa\)](#)
93. [Response](#)
94. [REST](#)
95. [REST API](#)
96. [Schema](#)
97. [Semantic Technologies](#)
98. [Semantic Web](#)
99. [Semantic Web Search Engine](#)
100. [Semantic Web Standards](#)
101. [Sesame](#)
102. [Simple Knowledge Organisation System](#)
103. [Sindice](#)
104. [Skolemization](#)
105. [SPARQL](#)
106. [SPARQL endpoint](#)
107. [Subject](#)
108. [Taxonomy](#)
109. [Term](#)
110. [Triple](#)
111. [Triple store](#)
112. [Tuple](#)
113. [Turtle](#)
114. [Uniform Resource Identifier](#)
115. [Uniform Resource Locator](#)
116. [URI](#)
117. [URL](#)
118. [Validation Service](#)
119. [Vocabulary](#)
120. [Vocabulary Alignment](#)
121. [VoID](#)
122. [Web 2.0](#)
123. [Web 3.0](#)
124. [Web of Data](#)
125. [Web of Documents](#)
126. [Web Ontology Language \(OWL\)](#)
127. [Web Resource](#)
128. [World Wide Web Consortium \(W3C\)](#)
129. [eXtensible Hypertext Markup Language \(XHTML\)](#)
130. [eXtensible Markup Language \(XML\)](#)
131. [XML Schema](#)
132. [eXtensible Stylesheet Language Transformations \(XSLT\)](#)
- A. [Acknowledgments](#)
- B. [References](#)
  - B.1 [Normative references](#)
  - B.2 [Informative references](#)

## Scope

This glossary lists terms related to publishing and consuming either Linked Data in the enterprise or Linked Open Data on the public Web.

## 1. 5 Star Linked Open Data

5 Star Linked Open Data refers to an incremental framework for deploying data. Tim Berners-Lee, the inventor of the Web and initiator of the [Linked Data project](#), suggested a 5 star deployment scheme for Linked Open Data. The 5 Star Linked Data system is cumulative. Each additional star presumes the data meets the criteria of the previous step(s). 5 Star [Linked Open Data](#) includes an **Open License** (expression of rights) and assumes publications on the public Web.

Organizations may elect to publish 5 Star Linked Data, without the word "open", implying that the data **does not** include an Open License (expression of rights) and does not imply publication on the public Web.

☆ **Publish data on the Web in any format (e.g., PDF, JPEG) accompanied by an explicit Open License (expression of rights).**

☆☆ **Publish structured data on the Web in a machine-readable format (e.g., XML).**

☆☆☆ **Publish structured data on the Web in a documented, non-proprietary data format (e.g., CSV, KML).**

☆☆☆☆ **Publish structured data on the Web as RDF (eg Turtle, RDFa, JSON-LD, SPARQL)**

☆☆☆☆☆ **In your RDF, have the identifiers be links (URLs) to useful data sources.**

## 2. Apache License

[Apache License, version 2.0](#) is used for many Linked Data tools and projects. It is a popular Open Source license published by the Apache Software Foundation.

## 3. API

An Application Programming Interface (API) is an abstraction implemented in software that defines how others should make use of a software package such as a library or other reusable program. APIs are used to provide developers access to data and functionality from a given system.

## 4. CC-BY-SA License

A form of Creative Commons license for resources released online. Work available under a [CC-BY-SA license](#) means you can include it in any other work under the condition that you give proper attribution. If you create derivative works (such as modified or extended versions), then you must also license them as CC-BY-SA.

## 5. Closed World

A concept from Artificial Intelligence and refers to a model of uncertainty that an agent assumes about the external world. In a closed world, the agent presumes that what is not known to be true must be false. This is a common assumption underlying relational databases, most forms of logical programming.

## 6. Connection

A concept from computer networking. It refers to a transport layer virtual circuit established between two programs for the purpose of communication.

## 7. Conneg

Abbreviated term for content negotiation. See also [Content Negotiation](#).

## 8. Content Negotiation

Also called "conneg", refers to a phase in establishing a network connection. In the HTTP Protocol, the use of a message header to indicate which response formats a client will accept. Content negotiation allows HTTP servers to provide different versions of a resource representation in response to any given URI request. See also [\[HTTP Protocol 1.1\]](#). See also [Connection](#).

## 9. Controlled Vocabulary

Carefully selected sets of terms that are used to describe units of information; used to create [taxonomies](#), thesauri and [ontologies](#). In traditional settings the terms in the controlled vocabularies are words or phrases, in a linked data setting then they are normally assigned unique identifiers (URIs) which in turn link to descriptive phrases.

## 10. Comma Separated Values (CSV)

A tabular data format in which columns of information are separated by comma characters. CSV files are a non-proprietary format and are considered 3-star data on the 5-star scale.

## 11. Creative Commons Licenses

Licenses that include legal statements by the owner of copyright in intellectual property specifically allowing people to use or redistribute the copyrighted work in accordance with conditions specified therein. See also [About Creative Commons Licenses](#).

## 12. CURIEs

Compact URI expressions (CURIEs) are an RDFa approach for shortening URIs.

## 13. cURL

A command line Open Source/Free Software client that can transfer data, including machine readable RDF, from or to a server using one of its many supported protocols.

## 14. Data Cloud

Data cloud, also called the [Linked Data Cloud](#), is a visual representation of datasets published as Linked Data. Many academic institutions republish data from their respective governments as [Linked Data](#), often enhancing the representation in the process.

## 15. Data Hub, The

[The Data Hub](#) is a specific site offering a community-run catalogue of data sets of data on the Internet, powered by the open-source data portal platform CKAN. [The Data Hub](#) is an openly editable open data catalogue in the style of Wikipedia.

## 16. Data Market

A data market, also called a Data Marketplace, is an online (broker) service to enable discovery and access to a large collection of datasets offered by a range of data providers. Examples include Infochimps, Azure Marketplace and Factual. Data Markets may include open as well as paid-for data, and may offer value added services such as [APIs](#) and visualizations and programmatic data access.

## 17. Data Modeling

Data modeling is a process of organizing data and information describing it into a faithful representation of a specific domain of knowledge. Linked data modeling applies modeling techniques based on [Linked Data Principles](#).

## 18. Dataset, RDF

A collection of RDF data, comprising one or more RDF graphs that is published, maintained, or aggregated by a single provider. In SPARQL, an RDF Dataset represents a collection of RDF graphs over which a query may be performed.

## 19. Data Warehouse

A data warehouse is one approach to data integration in which data from various operational data systems is extracted, cleaned, transformed and copied to a centralized repository. The centralized repository can then be used for data mining or answering analytical queries. By contrast, Linked Data assumes a *distributed approach* of data management using HTTP URIs to describe and access information resources. A Linked Data approach is seen as a valid alternative to the centralized data warehouse approach especially when integrating datasets available on the public Web.

## 20. DBpedia

DBpedia is a community effort to extract structured information from Wikipedia and make it available on the Web. DBpedia is often depicted as a hub for the [Data Cloud](#). An RDF representation of the metadata held in Wikipedia and made available for SPARQL query on the World Wide Web.

## 21. Dereferenceable URIs

When an HTTP client can look up a [URI](#) using the HTTP protocol and retrieve a description of the resource, it is called a dereferenceable URI. Dereferenceable URIs applies to URIs that are used to identify classic HTML documents and URIs that are used in the Linked Data context [[COOL-SWURIS](#)] to identify real-world objects and abstract concepts.

## 22. Description Logic

Description Logic (DL) is a family of knowledge representation languages with varying and adjustable expressivity. DL is used in artificial intelligence for formal reasoning on the concepts of an application domain. The Web Ontology Language (OWL) provides a standards-based way to exchange ontologies and includes a Description Logic semantics as well as an RDF based semantics. Biomedical informatics applications often use DL for codification of healthcare and life sciences knowledge.

## 23. DCAT

Data Catalog Vocabulary (DCAT) is an RDF vocabulary. It is designed to facilitate interoperability between data catalogs published on the Web. See also [Data Catalog Vocabulary \(DCAT\)](#).

## 24. DCMI

See [Dublin Core Metadata Initiative](#)

## 25. Directed Graph

A directed graph is a graph in which the links between nodes are directional, i.e., they only go from one node to another. RDF represents things (nouns) and the relationships between them (verbs) in a directed graph. In [RDF](#), links are labelled by being assigned unique [URIs](#).

## 26. Document Type Definition

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

## 27. Domain Name System (DNS)

Domain Name System (DNS) refers to the Internet's mechanism for mapping between a human-readable host name (e.g. [www.example.com](http://www.example.com)) and an Internet Protocol (IP) Address (e.g. 203.20.51.10).

## 28. Dublin Core Metadata Element Set

Dublin Core Metadata Element Set refers to a [vocabulary](#) of fifteen properties for use in resource descriptions, such as may be found in a library card catalog (creator, publisher, etc). The Dublin Core Metadata Element Set, also known as "DC Elements", is the most commonly used vocabulary for Linked Data applications. See also [Dublin Core Element Set, Version 1.1 Specification](#). [DCMI]

## 29. Dublin Core Metadata Initiative

A public, not-for-profit organization with a mission to promote interoperable metadata design and innovative practice. [The Dublin Core Metadata Initiative \(DCMI\)](#) manages the long-term curation and development of metadata standards such as the Dublin Core Element Set and DCMI Metadata Terms.

## 30. Dublin Core Metadata Terms

A vocabulary of bibliographic terms used to describe both physical publications and those on the Web. An extended set of terms beyond those basic terms found in the Dublin Core Metadata Element Set. See also [Dublin Core Metadata Terms](#)

## 31. Entity

In the sense of an entity-attribute-value model, an entity is synonymous with the [Subject](#) of an RDF Triple.

## 32. ETL

ETL is an abbreviation for extract, transform, load. Linked Data developers routinely extract data from a relational database, transform data to RDF Triples, and load it into an RDF database.

## 33. FOAF

See Friend of a Friend.

## 34. Fragment Identifier

The part of an HTTP URI that follows a hash symbol ('#'). Fragment identifiers are not passed to Web servers by Web clients such as Web browsers.

## 35. Free/Libre/Open Source Software

Free, also known as Libre or Open Source, is a generic and internationalized term for

software released under an Open Source license.

## 36. Friend of a Friend

A Semantic Web vocabulary describing people and their relationships for use in resource descriptions. Commonly called [["FOAF"](#)].

## 37. Government Open Data

Many government authorities have mandated publication of data to the public Web. The broad intention is to facilitate the maintenance of open societies and support governmental accountability and transparency initiatives. To realize the goals of improved efficiency, transparency and accountability, re-use of structured content available on the Web is enhanced by following [Linked Data Principles](#).

## 38. Graph

A collection of objects (represented by "nodes") any of which may be connected by links between them. See also [Directed Graph](#).

## 39. HyperText Markup Language

The predominant markup language for hypertext pages on the Web. HyperText Markup Language (HTML) defines the structure of Web pages and is a family of W3C standards.

## 40. HyperText Transfer Protocol

HyperText Transfer Protocol (HTTP) is the standard transmission protocol [[RFC2616](#)] used on the World Wide Web to transfer hypertext requests and information between Web servers and Web clients (such as browsers). It is an IETF standard.

## 41. HTTP URIs

See [Uniform Resource Identifier](#).

## 42. Inference

Inference is the process of deriving logical conclusions from a set of starting assumptions. Using Linked Data, existing relationships are modeled as a set of (named) relationships between resources. Linked Data helps humans and machines to find new relationships through automatic procedures that generate new relationships based on the data and based on some additional information in the form of a vocabulary.

## 43. International Organization for Standards (ISO)

ISO refers to a network of the national standards institutes of over 160 countries that

cooperate to define international standards. It defines many standards including, in the linked data context, formats for dates and currency. See also [ISO website](#).

## 44. Internet Engineering Task Force (IETF)

IETF is an open international community concerned with the evolution of Internet architecture and the operation of the Internet. It has defined standards such as HTTP and DNS.

## 45. Internationalized Resource Identifier

A global identifier standardized by joint action of the World Wide Web Consortium and Internet Engineering Task Force. An IRI may or may not be resolvable on the Web. A generalization of URIs that allow characters from the Universal Character Set (Unicode). Slowly replacing URIs. See also [Uniform Resource Identifier \(URI\)](#), [Uniform Resource Locator \(URL\)](#). Also called IRI.

## 46. JSON

JavaScript Object Notation (JSON) is syntax for storing and exchanging text based information. JSON has proven to be a highly useful and popular object serialization and messaging format for the Web. See also: the application/json Media Type for JavaScript Object Notation (JSON) [[RFC4627](#)].

## 47. JSON-LD

JavaScript Object Notation for Linking Data (JSON-LD) [[JSON-LD](#)] is a language-independent data format for representing Linked Data, based on JSON. JSON-LD is capable of serializing any RDF graph or dataset and most, but not all, JSON-LD documents can be directly transformed to RDF. JSON-LD Syntax is easy for humans to read and write as well as, easy for machines to parse and generate. JSON-LD is an appropriate Linked Data interchange language for JavaScript environments, Web service and NoSQL databases. See also: [[JSON-LD](#)] JSON-LD Syntax 1.0

## 48. 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 Data is intended for access by both humans and machines. Linked Data uses the RDF family of standards for data interchange (e.g., RDF/XML, RDFa, Turtle) and query (SPARQL). If Linked Data is published on the public Web, it is generally called [Linked Open Data](#). See also [[Linked Data Principles](#)].

## 49. Linked Data API

A REST API that allows data publishers to provide URLs to lists of things and clients to retrieve machine-readable data from those URLs.

## 50. Linked Data client

A Web client that supports HTTP content negotiation for the retrieval of Linked Data from URLs and/or SPARQL endpoints. A Linked Data client understands standard REST API, for example the Linked Data REST API. Examples of Linked Data clients include: Tim Berners-Lee's early [Tabulator browser](#), [gFacet](#), and the [Callimachus Shell \(CaSH\)](#).

## 51. Linked Data Platform

A specification that defines a REST API to read and write Linked Data for the purposes of enterprise application integration. The Linked Data Platform describes the use of a REST API for accessing, updating, creating and deleting resources from servers. See also [[LDP-ONE](#)]

## 52. Linked Data Principles

Provide a common [API](#) for data on the Web which is more convenient than many separately and differently designed APIs published by individual data suppliers. Tim Berners-Lee, the inventor of the Web and initiator of the [Linked Data](#) project, proposed the following principles upon which Linked Data is based:

1. Use URIs to name things;
2. Use HTTP URIs so that things can be referred to and looked up ("dereferenced") by people and user agents;
3. When someone looks up a URI, provide useful information, using the open Web standards such as RDF, SPARQL;
4. Include links to other related things using their URIs when publishing on the Web.

## 53. Linked Open Data

Linked Data published on the public Web and licensed under one of several open licenses permitting reuse. Publishing Linked Open Data enables distributed [SPARQL](#) queries of the data sets and a "browsing" or "discovery" approach to finding information, as compared to a search strategy. See also: "Linked Data: Structured Data on the Web" [[LD-FOR-DEVELOPERS](#)] and "Linked Data: Evolving the Web into a Global Data Space" [[HOWTO-LODP](#)]

## 54. Linked Open Data Cloud

A colloquial phrase for the total collection of Linked Data published on the Web. See also: [Data Cloud](#), [Linked Open Data Cloud diagram](#)

## 55. Linked Open Data Cloud diagram

A diagram representing datasets published by the Linking Open Data project from 2007-2011. The diagram stopped being updated when individual datasets could no longer be meaningfully represented in a single diagram due to the number of total

datasets. See the pictorial depiction of the [Linked Data Cloud](#).

## 56. Linking Open Data Project

A community activity started in 2007 by the W3C's [Semantic Web Education and Outreach \(SWEO\) Interest Group](#). The project's stated goal is to "make data freely available to everyone".

## 57. Linkset

A collection of RDF links between two datasets.

## 58. Machine Readable Data

Data formats that may be readily parsed by computer programs without access to proprietary libraries. For example, CSV, TSV and RDF formats are machine readable, but PDF and Microsoft Excel are not. Creating and publishing data following Linked Data principles helps search engines and humans to find, access and re-use data. Once information is found, computer programs can re-use data without the need for custom scripts to manipulate the content.

Publishing machine readable data using Linked Data principles provides a human and machine readable version. For example, Wikipedia includes a Web page about the color [Red](#). DBpedia, the database containing structured content contained in Wikipedia, allows a Linked Data client to look up "Red" [<http://wikipedia.org/wiki/Red>] by changing "wiki" to "data" and appending the appropriate file extension.

```
$ curl -L http://dbpedia.org/data/Red.ttl
```

## 59. Message

The basic unit of HTTP communication. It consists of a structured sequence of octets matching the syntax defined as an [HTTP Message](#) and transmitted via the connection.

## 60. 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 may be further subcategorized into several types (including general, access and structural metadata). Linked Data incorporates human and machine readable metadata along with it, making it self describing.

## 61. Metadata Object Description Schema

It is a bibliographic description system intended to be a compromise between [MARC](#) and [DC metadata](#). It is implemented in [XML Schema](#). See DC, MARC, XSD.

## 62. Modeling Process

Modeling process in the context of RDF refers to the act by subject matter experts to work with developers to capture the context of data and define the relationships of the data. By doing so, high quality of Linked Data is obtained since capturing organizational knowledge about the meaning of the data within the RDF data model means the data is more likely to be reused correctly. Well defined context ensures better understanding, proper reuse, and is critical when establishing linkages to other data sets.

## 63. N3

N3 is an abbreviation for Notation3. It has a readable [RDF](#) syntax used for expressing assertion and logic. N3 [N3] is a superset of RDF, extending the RDF model by adding formulae (literals which are graphs themselves), variables, logical implication, and functional predicates. See also [\[Turtle\]](#).

## 64. Namespace IRI

A namespace IRI is a base IRI shared by all terms in a given vocabulary or ontology.

## 65. Natural Keys

Human-readable categories and sub-identifiers within a URI that reflect what the identifier describes. They are recommended when creating URIs so that people reading RDF in its source format (mostly developers) will be able to more quickly understand it.

## 66. Neutral URI

A URI that avoids the exposure of implementation details within the URI itself.

## 67. N-Triples

A subset of Turtle that defines a line-based format to encode a single RDF graph. Used primarily as an exchange format for RDF data. See also [N-triples](#)

## 68. Object

In the context of [RDF](#), the object is the final part of an RDF statement. See also [\[Subject\]](#) [\[Predicate\]](#)

## 69. Ontology

A formal model that allows knowledge to be represented for a specific domain. An ontology describes the types of things that exist (classes), the relationships between them (properties) and the logical ways those classes and properties can be used together (axioms).

## 70. Open Government Data

Refers to content that is published on the public Web by government authorities in a variety of non-proprietary formats.

## 71. ORG Ontology

ORG is an RDF vocabulary to enable publication of information about organizations and organizational structures, even at governmental level. See also [<http://www.w3.org/TR/vocab-org/>]

## 72. Persistent Identifier Scheme

A persistent identifier scheme is a mechanism for resolution of virtual resources. [Persistent Uniform Resource Locator \(PURLs\)](#) implement one form of persistent identifier for virtual resources. PURLs are valid URLs and their components must map to the URL specification. The scheme part tells a computer program, such as a Web browser, which protocol to use when resolving the address. The scheme used for PURLs is generally HTTP. Other persistent identifier schemes include [Digital Object Identifiers \(DOIs\)](#), [Life Sciences Identifiers \(LSIDs\)](#) and [INFO URIs](#). All persistent identification schemes provide unique identifiers for (possibly changing) virtual resources, but not all schemes provide curation opportunities.

## 73. Persistent Uniform Resource Locator

URLs that act as permanent identifiers in the face of a dynamic and changing Web infrastructure. Persistent Uniform Resource Locators (PURLs) redirect to the current location of or proxy specific Web content. A user of a PURL always uses the same Web address, even though the resource in question may have moved or changed ownership.

## 74. Predicate

The middle term (the linkage, or "verb") in an RDF statement. For example, in the statement "*Alice knows Bob*" then "*knows*" is the predicate which connects "Alice" (the subject of the statement) to "Bob" (the object of the statement).

## 75. 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. One of the most common Internet protocols is the Hypertext Transfer Protocol (HTTP).

## 76. Provenance

Data related to where, when and how information was acquired.

## 77. PURL

See [Persistent Uniform Resource Locator](#)

## 78. 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.

## 79. Query

Programmatic retrieval of resources and their relationships. Using the SPARQL language, developers issue queries based on (triple) patterns.

## 80. R2RML

[R2RML](#) (RDB to RDF Mapping Language) is a language for expressing customized mappings from relational databases to RDF datasets. Such mappings provide the ability to view existing relational data in the RDF data model, expressed in a structure and target vocabulary of the mapping author's choice.

## 81. Raw Data

Raw data refers to machine-readable files from the wilderness released without any specific effort to make them applicable to a particular application. Raw data typically requires additional scripts or programs to process the data.

## 82. RDF

See [Resource Description Framework](#)

## 83. RDFa

See [Resource Description Framework in Attributes](#)

## 84. RDF database

A type of database designed specifically to store and retrieve RDF information. May be implemented as a triple store, quad store or other type.

## 85. RDF-JSON

A name for one of the early proposals for serializing RDF in JavaScript Object Notation (JSON) [[RFC4627](#)]. RDF-JSON is still widely used. Originally proposed as the [Talis Platform API Output Type](#). See also a concrete syntax in JSON [[RFC4627](#)] for RDF as defined in the RDF Concepts and Abstract Syntax [[RDF-CONCEPTS](#)] W3C Recommendation and [JSON-LD](#) which are more recent W3C documents.

## 86. RDF Schema

The simplest RDF vocabulary description language. It provides much less descriptive capability than the Simple Knowledge Organization System (SKOS) or the Web

Ontology Language (OWL). A standard of the W3C [[RDFS](#)]

## 87. RDF/XML

An RDF syntax encoded in XML. A standard of the W3C. [[RDF](#)]

## 88. Representational State Transfer (REST)

An architectural style for information systems used on the Web. It explains some of the Web's key features, such as extreme scalability and robustness to change. REST is the foundation of the World Wide Web and the dominant Web service design model. The term "Representational State Transfer" was introduced and defined in 2000 by Roy Thomas Fielding in his doctoral dissertation. See also "[Architectural Styles and the Design of Network-based Software Architectures](#)" by Roy Thomas Fielding.

## 89. Request

Request refers to a stage in the HTTP protocol. A request message from a client to a server includes, within the first line of that message, the method to be applied to the resource, the identifier of the resource, and the protocol version in use.

## 90. Resource

In an RDF context, a resource can be anything that an RDF graph describes. A resource can be addressed by a [Unified Resource Identifier \(URI\)](#). See also Resource Description Framework (RDF) 1.1 Concepts and Abstract Syntax [[RDF11-CONCEPTS](#)]

## 91. Resource Description Framework (RDF)

A family of international standards for data interchange on the Web produced by W3C. Resource Description Framework (RDF) is based on the idea of identifying things using Web identifiers or [HTTP URIs](#), and describing resources in terms of simple properties and property values. See also [[RDF 1.1 Concepts and Abstract Syntax](#)]

## 92. Resource Description Framework in attributes (RDFa)

An RDF syntax encoded in HTML documents. RDFa provides a set of markup attributes to augment the visual information on the Web with machine-readable hints. It is a standard of the World Wide Web Consortium. See also RDFa Primer - Rich Structured Data Markup for Web Documents [[RDFa-PRIMER](#)]

## 93. Response

An action by a server taken as the result of a request by a client. In HTTP, a response provides a resource representation to the calling client. See also [[RFC2616](#)] bis for

an [HTTP Response](#) message.

## 94. REST

See [Representational State Transfer](#)

## 95. REST API

An application programming interface (API) implemented using HTTP and the principles of REST to allow actions on Web resources. The most common actions are to create, retrieve, update and delete resources. See also [Representational State Transfer](#).

## 96. Schema

Schema refers to 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.

## 97. Semantic Technologies

The broad set of technologies that related to the extraction, representation, storage, retrieval and analysis of machine-readable information.

## 98. 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)

## 99. Semantic Web Search Engine

A search engine capable of making use of semantic technologies to model its knowledge base and to deliver content.

## 100. Semantic Web Standards

Standards of the World Wide Web Consortium relating to the Semantic Web, including RDF [[RDF](#)], RDFa [[RDFa-PRIMER](#)], SKOS [[SKOS-REFERENCE](#)], OWL [[OWL2](#)] and SPARQL 1.1 Overview [[SPARQL-11](#)].

## 101. Sesame

Sesame is an Open Source Software implementation of a Semantic Web development framework. It supports the storage, retrieval and analysis of RDF information. See also [[Open RDF](#)].

## 102. Simple Knowledge Organisation System

Simple Knowledge Organisation System (SKOS) [[SKOS-REFERENCE](#)] is a vocabulary description language for RDF designed for representing traditional knowledge organization systems such as enterprise taxonomies in RDF.

## 103. Sindice

Sindice is a search engine for Linked Data. It offers search and querying capabilities across the data it knows about, as well as specialized APIs and tools for presenting Linked Data summaries. See also <http://sindice.com>.

## 104. Skolemization

Skolemization is a process whereby some RDF databases and other systems implementing the SPARQL query language automatically assign URIs to blank nodes so that they are more easily operated upon.

## 105. SPARQL

SPARQL Protocol and RDF Query Language (SPARQL) defines a query language for RDF data, analogous to the Structured Query Language (SQL) for relational databases. A family of standards of the World Wide Web Consortium. See also SPARQL 1.1 Overview [[SPARQL-11](#)].

## 106. SPARQL endpoint

A service that accepts SPARQL queries and returns answers to them as SPARQL result sets. It is a best practice for datasets providers to give the URL of their SPARQL endpoint to allow access to their data programmatically or through a Web interface. A list of some endpoints status is available at <http://labs.mondeca.com/sparqlEndpointsStatus/>

## 107. Subject

The subject is the first part of an [RDF](#) statement. A subject in the context of a [triple](#) <?s ?p ?o> refers to who or what the RDF statement is about.

## 108. Taxonomy

Taxonomy is a formal representation of relationships between items in a hierarchical structure. Also see [Ontology](#).

## 109. Term

An entry in a controlled vocabulary, schema, [Taxonomy](#) or [Ontology](#).

## 110. 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 such statements).

## 111. Triple store

A colloquial phrase for an [RDF](#) database that stores RDF triples.

## 112. Tuple

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

## 113. Turtle

An RDF serialization format designed to be easier to read than others such as RDF/XML. The term "Turtle" was derived from Terse RDF Triple Language. Turtle allows an RDF graph to be written in a compact and natural text form, with abbreviations for common usage patterns and datatypes. Turtle provides levels of compatibility with the existing N-Triples format, as well as the triple pattern syntax of the SPARQL W3C Recommendation. See Terse RDF Triple Language [[TURTLE-TR](#)].

## 114. Uniform Resource Identifier

A global identifier standardized by joint action of the World Wide Web Consortium and Internet Engineering Task Force. A Uniform Resource Identifier (URI) may or may not be resolvable on the Web. URIs play a key role in enabling Linked Data. URIs can be used to uniquely identify virtually anything including a physical building or more abstract concepts such as colors. See also [Internationalized Resource Identifier \(IRI\)](#) and [Uniform Resource Locator \(URL\)](#). See also Uniform Resource Identifier (URI): Generic Syntax [[RFC3986](#)] and <http://www.w3.org/DesignIssues/Architecture.html>.

URIs have been known by many names: Web addresses, Universal Document Identifiers, Universal Resource Identifiers. If you are interested in the history of the many names, read Tim Berners-Lee's design document [Web Architecture from 50,000 feet](#). See also Uniform Resource Identifier (URI): Generic Syntax [[RFC3986](#)].

## 115. Uniform Resource Locator

A global identifier for Web resources standardized by joint action of the World Wide Web Consortium and Internet Engineering Task Force. A URL is resolvable on the Web and is commonly called a "Web address". All HTTP URLs are URIs however, not all URIs are URLs. See also [Internationalized Resource Identifier](#) and [Uniform Resource Identifier](#).

## 116. URI

See [Uniform Resource Identifier](#)

## 117. URL

See [Uniform Resource Locator](#)

## 118. Validation Service

The W3C offers an RDF validation service to check and validate RDF files. It is considered a best practice to validate RDF files prior to publishing them on the Web. See <http://www.w3.org/RDF/Validator/>. See also [http://www.w3.org/People/Barstow/#online\\_parsers](http://www.w3.org/People/Barstow/#online_parsers).

## 119. Vocabulary

A collection of "[terms](#)" for a particular purpose. Vocabularies can range from simple such as the widely used RDF Schema, FOAF and Dublin Core Metadata Element Set to complex vocabularies with thousands of terms, such as those used in healthcare to describe symptoms, diseases and treatments. Vocabularies play a very important role in Linked Data, specifically to help with data integration. The use of this term overlaps with Ontology.

## 120. Vocabulary Alignment

The process of analyzing multiple vocabularies to determine terms that are common across them and to record those relationships.

## 121. VoID

Vocabulary of Interlinked Datasets, an [RDF Schema](#) vocabulary for expressing metadata about RDF datasets and a standard of the World Wide Web Consortium. VoID is intended as a bridge between the publishers and users of RDF data, with applications ranging from data discovery to cataloging and archiving of datasets. VoID can be used to express general metadata based on Dublin Core, access metadata, structural metadata, and links between datasets. See also Describing Linked Datasets with the VoID Vocabulary [[VOID-GUIDE](#)]

## 122. Web 2.0

Web 2.0 is 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.

## 123. Web 3.0

Web 3.0 is 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. It is considered synonymous with the phrases "Semantic Web" and "The Web of Data".

## 124. Web of Data

A subset of the World Wide Web which contains machine readable data represented as [Linked Data](#).

## 125. Web of Documents

The original, or traditional, World Wide Web in which published resources were nearly always documents as opposed to machine readable data.

## 126. Web Ontology Language (OWL)

OWL is a family of knowledge representation and vocabulary description languages for authoring ontologies, based on RDF and standardized by the W3C. See also OWL 2 Web Ontology Language Document Overview (Second Edition) [[OWL2](#)].

## 127. Web Resource

A web page addressed by a URL. Examples include: an HTML web page, an image offered by a web server, or a dataset accessible by a URL. A Web Resource may have different representations. For example, an RDF database might be accessed at a single URL using multiple syntaxes, such as RDFa, JSON-LD, and Turtle. See also Hypertext Transfer Protocol HTTP/1.1 [[RFC2616](#)].

## 128. World Wide Web Consortium (W3C)

An international community that develops and promotes protocols and guidelines for the long-term growth for the Web. W3C's standards define key parts of the World Wide Web, including [Web Design](#), [Web Architecture](#) and the [Semantic Web](#). See also [W3C Mission](#).

## 129. eXtensible Hypertext Markup Language (XHTML)

eXtensible Hypertext Markup Language (XHTML) is a family of versions of HTML based on XML and standardized by the W3C. See also XHTML 1.0 The Extensible HyperText Markup Language (Second Edition) [[XHTML1](#)].

## 130. eXtensible Markup Language (XML)

XML [[XML](#)] is a specification for creating structured textual computer documents, subset of SGML enabling such documents to be served, received and process on the Web in the same way as HTML documents . There are many thousands of XML formats, including XHTML. It is part of a family of standards from the W3C. See also [[XHTML1](#)].

## 131. XML Schema

XML Schemas provide a means for defining the structure, content and semantics of

XML documents as defined in [XMLS-SCHEMA0].

## 132. eXtensible Stylesheet Language Transformations (XSLT)

eXtensible Stylesheet Language Transformations (XSLT) is a declarative program to transform one XML document into another XML document.

### A. Acknowledgments

The editors are grateful to [David Wood](#) for contributing the initial glossary terms from [Linking Government Data](#), (Springer 2011). The editors wish to also thank members of the [Government Linked Data Working Group](#) with special thanks to the reviewers and contributors: Thomas Baker, Hadley Beeman, Richard Cyganiak, Michael Hausenblas, Sandro Hawke, Benedikt Kaempgen, James McKinney, Marios Meimaris, Jindrich Mynarz and Dave Reynolds who diligently iterated the W3C Linked Data Glossary in order to create a foundation of terms upon which to discuss and better describe the Web of Data. Thank you!

### B. References

#### B.1 Normative references

##### [OWL2]

[OWL 2 Web Ontology Language Document Overview](#), W3C OWL Working Group, 27 October 2009. W3C Recommendation. URL: <http://www.w3.org/TR/owl2-overview/>

##### [RDF]

[RDF/XML Syntax Specification \(Revised\)](#), Dave Beckett (eds), 10 February 2004, W3C Recommendation. URL: <http://www.w3.org/TR/REC-rdf-syntax/>

##### [RDF-CONCEPTS]

[Resource Description Framework \(RDF\): Concepts and Abstract Syntax](#), Graham Klyne, Jeremy J. Carroll, 10 February 2004. W3C Recommendation. URL: <http://www.w3.org/TR/2004/REC-rdf-concepts-20040210/>

##### [RDF11-CONCEPTS]

[RDF 1.1: Concepts and Abstract Syntax](#), Richard Cyganiak, David Wood, 15 January 2013. W3C Recommendation. URL: <http://www.w3.org/TR/rdf11-concepts>

##### [RDFS]

[RDF Vocabulary Description Language 1.0: RDF Schema](#), ed. Dan Brickley, R.V. Guha, 10 February 2004. W3C Recommendation. URL: <http://www.w3.org/TR/rdf-schema/>

##### [RDFa-PRIMER]

[RDFa Primer](#), Ben Adida, Ivan Herman, Manu Sporny, 07 June 2012. W3C Note. URL: <http://www.w3.org/TR/2012/NOTE-rdfa-primer-20120607/>

##### [RFC2616]

[Hypertext Transfer Protocol -- HTTP/1.1](#), R. Fielding; et al. June 1999. Internet RFC 2616. URL: <http://www.w3.org/Protocols/rfc2616/rfc2616.html>.

##### [RFC3986]

[Uniform Resource Identifier \(URI\): Generic Syntax](#), Berners-Lee, et al. January 2005. Internet RFC 3986. URL: <http://tools.ietf.org/html/rfc3986>.

**[RFC4627]**

[The application/json Media Type for JavaScript Object Notation \(JSON\)](#), D. Crockford, July 2006. Network Working Group. URL: <http://www.ietf.org/rfc/rfc4627.txt>

**[SKOS-REFERENCE]**

[SKOS: Simple Knowledge Organization System Reference](#), Sean Bechhofer, Alistair Miles (eds), 18 August 2009, W3C Recommendation. URL: <http://www.w3.org/TR/2009/REC-skos-reference-20090818/>

**[SPARQL-11]**

[SPARQL 1.1 Overview](#), The W3C SPARQL Working Group, 21 March 2011. W3C Recommendation. URL: <http://www.w3.org/TR/sparql11-overview/>

**[TURTLE-TR]**

[Turtle: Terse RDF Triple Language](#), Eric Prud'hommeaux, Gavin Carothers, 19 February 2013. W3C Candidate Recommendation. URL: <http://www.w3.org/TR/turtle/>

**[XHTML1]**

[XHTML 1.0 The Extensible HyperText Markup Language \(Second Edition\)](#), Steven Pemberton, Daniel Auster, et al., 26 January 2000. W3C Recommendation. URL: <http://www.w3.org/TR/xhtml1/>

**[XML]**

[Extensible Markup Language \(XML\) 1.0 \(Fifth Edition\)](#), Tim Bray, Jean Paoli, C.M. Sperberg-McQueen, Eve Maler, François Yergeau, 26 November 2008. W3C Recommendation. URL: <http://www.w3.org/TR/REC-xml/>

**[XMLS-SCHEMA0]**

[XML Schema Part 0: Primer Second Edition](#), David C. Fallside, Priscilla Walmsley (eds), 28 October 2004, W3C Recommendation. URL: <http://www.w3.org/TR/xmlschema-0/>

## B.2 Informative references

**[COOL-SWURIS]**

[Cool URIs for the Semantic Web](#), L. Sauermann and R. Cyganiak, W3C Interest Group Note 03 December 2008. URL: <http://www.w3.org/TR/cooluris/>

**[HOWTO-LODP]**

[Linked Data: Evolving the Web into a Global Data Space](#). 2011, Chris Bizer, Tom Heath URL: <http://linkeddatabook.com/editions/1.0/>

**[JSON-LD]**

[JSON-LD Syntax 1.0](#), Many Sporny, Gregg Kellogg, Markus Lanthaler (eds), 12 July 2012, W3C Working Draft. URL: <http://www.w3.org/TR/json-ld-syntax/>

**[LD-FOR-DEVELOPERS]**

[Linked Data: Structured Data on the Web](#). David Wood, Marsh Zaidman, Luke Ruth, with Michael Hausenblas; 2013 URL: <http://www.manning.com/dwood/>

**[LDP-ONE]**

[Linked Data Platform 1.0](#). Steve Speicher, John Arwe. 07 March 2013. W3C Working Draft, Linked Data Platform Working Group. URL: <http://www.w3.org/TR/ldp/>

**[N3]**

[\*Notation3 \(N3\): A readable RDF syntax\*](#), Tim Berners-Lee, Dan Connolly, 28 March 2011. W3C Team Submission. URL: <http://www.w3.org/TeamSubmission/n3/>

**[VOID-GUIDE]**

[\*Describing Linked Datasets with the VOID Vocabulary\*](#), K. Alexander, R. Cyganiak, M. Hausenblas, and J. Zhao, W3C Interest Group Note 03 March 2011. URL: <http://www.w3.org/TR/void/>