



PROV-Overview

An Overview of the PROV Family of Documents

W3C Working Draft 11 December 2012

This version:

<http://www.w3.org/TR/2012/WD-prov-overview-20121211/>

Latest published version:

<http://www.w3.org/TR/prov-overview/>

Latest editor's draft:

<http://dvcs.w3.org/hg/prov/raw-file/default/overview/prov-overview.html>

Editors:

[Paul Groth](#), VU University Amsterdam

[Luc Moreau](#), University of Southampton

[Copyright](#) © 2012-2012 [W3C](#)® ([MIT](#), [ERCIM](#), [Keio](#)), All Rights Reserved. W3C [liability](#), [trademark](#) and [document use](#) rules apply.

Abstract

Provenance is information about entities, activities, and people involved in producing a piece of data or thing, which can be used to form assessments about its quality, reliability or trustworthiness. The PROV Family of Documents defines a model, corresponding serializations and other supporting definitions to enable the inter-operable interchange of provenance information in heterogeneous environments such as the Web. This document provides an overview this family of documents.

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/>.

PROV Family of Documents

This document is part of the PROV family of documents, a set of documents defining

various aspects that are necessary to achieve the vision of inter-operable interchange of provenance information in heterogeneous environments such as the Web. These documents are:

- [PROV-OVERVIEW](#) (To be published as Note), an overview of the PROV family of documents (this document);
- [PROV-PRIMER](#) (To be published as Note), a primer for the PROV data model [[PROV-PRIMER](#)];
- [PROV-O](#) (Candidate Recommendation), the PROV ontology, an OWL2 ontology allowing the mapping of PROV to RDF [[PROV-O](#)];
- [PROV-DM](#) (Candidate Recommendation), the PROV data model for provenance [[PROV-DM](#)];
- [PROV-N](#) (Candidate Recommendation), a notation for provenance aimed at human consumption [[PROV-N](#)];
- [PROV-CONSTRAINTS](#) (Candidate Recommendation), a set of constraints applying to the PROV data model [[PROV-CONSTRAINTS](#)];
- [PROV-AQ](#) (To be published as Note), the mechanisms for accessing and querying provenance [[PROV-AQ](#)];
- [PROV-XML](#) (To be published as Note), an XML schema for the PROV data model [[PROV-XML](#)];
- [PROV-DC](#) (To be published as Note), describes a mapping between Dublin Core and PROV [[PROV-DC](#)].
- [PROV-LINKS](#) (To be published as Note), introduces a mechanism to link across bundles [[PROV-LINKS](#)].

This document was published by the [Provenance Working Group](#) as a First Public Working Draft. If you wish to make comments regarding this document, please send them to public-prov-comments@w3.org ([subscribe](#), [archives](#)). All feedback is welcome.

Publication as a Working Draft 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](#). The group does not expect this document to become a W3C Recommendation. 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

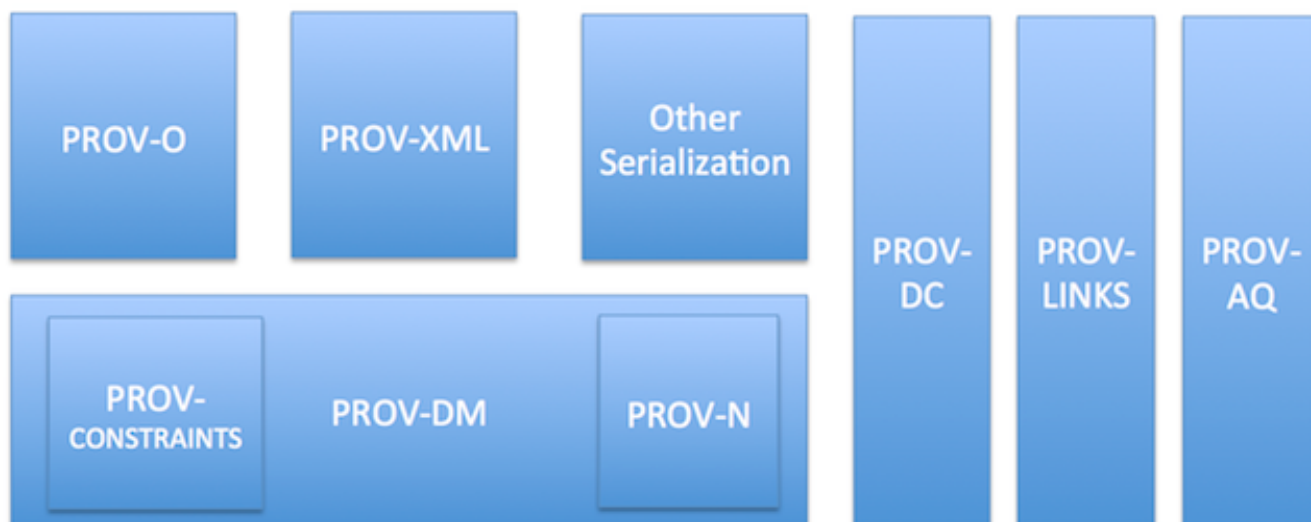
1. [Introduction](#)
2. [Document Roadmap](#)
3. [Additional Information](#)
 - A. [Acknowledgements](#)
 - B. [References](#)
 - B.1 [Normative references](#)

B.2 Informative references

1. Introduction

This document provides a non-normative overview of the PROV Family of Documents and provides a roadmap to using these these documents. Provenance is information about entities, activities, and people involved in producing a piece of data or thing, which can be used to form assessments about its quality, reliability or trustworthiness. The goal of PROV is to enable the wide publication and interchange of provenance on the Web and other information systems. PROV enables one to represent and interchange provenance information using widely available formats such as RDF and XML. In addition, it provides definitions for accessing provenance information, validating it, and mapping to Dublin Core.

Below is the organization of PROV. At its core is a conceptual data model, which defines a common vocabulary used to describe provenance. This is instantiated by various serializations. These serializations are what are used by implementations to interchange provenance. To help developers and users create valid provenance, a set of constraints are defined, which can be used to create provenance validators. Finally, to further support the interchange of provenance, additional definitions are provided for protocols to locate and access provenance, connect sets of provenance descriptions and define how to interoperate with the widely used Dublin Core vocabulary.



2. Document Roadmap

PROV consists of 10 documents (including this one). In order to use PROV, one need not be familiar with all of these documents. Indeed, PROV was specifically designed so that users and developers may get started quickly with basic usage and then incrementally progress to more advanced usage scenarios To help navigate PROV, each document is broadly classified as being intended for a specific audience.

1. Users - this audience wants to understand PROV and use applications that support PROV.
2. Developers - this audience wants to develop or build applications that create and

consume provenance using PROV.

3. Advanced - this audience aims to create validators, new PROV serializations, or other advanced provenance systems.

In the table below, we also denote the track that each document on whether it is intended to be a W3C Recommendation or a Working Group Note.

Part	Audience	Type	Document
1	Users	Note	PROV-PRIMER is the entry point to PROV offering an introduction to the provenance model. This is where you should start and for many may be the only document needed.
2	Developers	Rec	PROV-O defines a light-weight OWL2 ontology for the provenance model. This is intended for the Linked Data and Semantic Web community.
3	Developers	Note	PROV-XML defines an XML schema for the provenance model. This is intended for developers who need a native XML serialization of PROV
4	Advanced	Rec	PROV-DM defines a conceptual data model for provenance including UML diagrams. PROV-O and PROV-XML are serializations of this conceptual model.
5	Advanced	Rec	PROV-N defines a human-readable notation for provenance. This is used to define the conceptual model as well as PROV-CONSTRAINTS.
6	Advanced	Rec	PROV-CONSTRAINTS defines a set constraints that define a notion of valid provenance. It is specifically aimed at the implementors of validators.
7	Developers	Note	PROV-AQ defines how to use Web-based mechanisms to locate and retrieve provenance information.
8	Developers	Note	PROV-DC defines a mapping between Dublin Core and PROV.
9	Advanced	Note	PROV-LINKS Defines extensions to PROV to enable linking provenance information across containers for provenance

3. Additional Information

In addition, to these specifications, the [PROV FAQ](#) page addresses common questions as well as sets PROV in a broader context. This page is continually updated. Working group members have also given several [tutorials](#) about PROV including hands-on exercises, these may be a useful place to start. In addition, one can find a variety of blog posts and web pages on PROV - a short list can be found [here](#). For a broader review of provenance that led to the creation of PROV, there are several reports produced by the [W3C Provenance Incubator group](#) including:

1. [An Overview of Provenance on the Web \(slideshow - pdf\)](#)
2. [Requirements for Provenance on the Web](#)
3. [State of the Art Report](#)

Finally, the simplest way to use PROV is through one of the many [applications](#) that support it.

A. Acknowledgements

This document has been produced by the PROV Working Group, and its contents reflect extensive discussion within the Working Group as a whole.

Members of the PROV Working Group at the time of publication of this document were: Ilkay Altintas (Invited expert), Reza B'Far (Oracle Corporation), Khalid Belhajjame (University of Manchester), James Cheney (University of Edinburgh, School of Informatics), Sam Coppens (IBBT), David Corsar (University of Aberdeen, Computing Science), Stephen Cresswell (The National Archives), Tom De Nies (IBBT), Helena Deus (DERI Galway at the National University of Ireland, Galway, Ireland), Simon Dobson (Invited expert), Martin Doerr (Foundation for Research and Technology - Hellas(FORTH)), Kai Eckert (Invited expert), Jean-Pierre EVAÏN (European Broadcasting Union, EBU-UER), James Frew (Invited expert), Irini Fundulaki (Foundation for Research and Technology - Hellas(FORTH)), Daniel Garijo (Universidad Politécnica de Madrid), Yolanda Gil (Invited expert), Ryan Golden (Oracle Corporation), Paul Groth (Vrije Universiteit), Olaf Hartig (Invited expert), David Hau (National Cancer Institute, NCI), Sandro Hawke (W3C/MIT), Jörn Hees (German Research Center for Artificial Intelligence (DFKI) GmbH), Ivan Herman, (W3C/ERCIM), Ralph Hodgson (TopQuadrant), Hook Hua (Invited expert), Trung Dong Huynh (University of Southampton), Graham Klyne (University of Oxford), Michael Lang (Revelytix, Inc.), Timothy Lebo (Rensselaer Polytechnic Institute), James McCusker (Rensselaer Polytechnic Institute), Deborah McGuinness (Rensselaer Polytechnic Institute), Simon Miles (Invited expert), Paolo Missier (School of Computing Science, Newcastle university), Luc Moreau (University of Southampton), James Myers (Rensselaer Polytechnic Institute), Vinh Nguyen (Wright State University), Edoardo Pignotti (University of Aberdeen, Computing Science), Paulo da Silva Pinheiro (Rensselaer Polytechnic Institute), Carl Reed (Open Geospatial Consortium), Adam Retter (Invited Expert), Christine Runnegar (Invited expert), Satya Sahoo (Invited expert), David Schaengold (Revelytix, Inc.), Daniel Schutzer (FSTC, Financial Services Technology Consortium), Yogesh Simmhan (Invited expert), Stian Soiland-Reyes (University of Manchester), Eric Stephan (Pacific Northwest National Laboratory), Linda Stewart (The National Archives), Ed Summers (Library of Congress), Maria Theodoridou (Foundation for Research and Technology - Hellas(FORTH)), Ted Thibodeau (OpenLink Software Inc.), Curt Tilmes (National Aeronautics and Space Administration), Craig Trim (IBM Corporation), Stephan Zednik (Rensselaer Polytechnic Institute), Jun Zhao (University of Oxford), Yuting Zhao (University of Aberdeen, Computing Science).

B. References

B.1 Normative references

No normative references.

B.2 Informative references

[PROV-AQ]

Graham Klyne; Paul Groth; eds. [*Provenance Access and Query*](#). 19 June 2012, Working Draft. URL: <http://www.w3.org/TR/2012/WD-prov-aq-20120619/>

[PROV-CONSTRAINTS]

James Cheney; Paolo Missier; Luc Moreau; eds. [*Constraints of the PROV Data Model*](#). 11 December 2012, W3C Candidate Recommendation. URL: <http://www.w3.org/TR/2012/CR-prov-constraints-20121211/>

[PROV-DC]

Daniel Garijo; Kai Eckert; eds. [*Dublin Core to PROV Mapping*](#). 11 December 2012, Working Draft. URL: <http://www.w3.org/TR/2012/WD-prov-dc-20121211/>

[PROV-DM]

Luc Moreau; Paolo Missier; eds. [*PROV-DM: The PROV Data Model*](#). 11 December 2012, W3C Candidate Recommendation. URL: <http://www.w3.org/TR/2012/CR-prov-dm-20121211/>

[PROV-LINKS]

Luc Moreau; Timothy Lebo; eds. [*Linking Across Provenance Bundles*](#). 11 December 2012, Working Draft. URL: <http://www.w3.org/TR/2012/WD-prov-links-20121211/>

[PROV-N]

Luc Moreau; Paolo Missier; eds. [*PROV-N: The Provenance Notation*](#). 11 December 2012, W3C Candidate Recommendation. URL: <http://www.w3.org/TR/2012/CR-prov-n-20121211/>

[PROV-O]

Timothy Lebo; Satya Sahoo; Deborah McGuinness; eds. [*PROV-O: The PROV Ontology*](#). 11 December 2012, W3C Candidate Recommendation. URL: <http://www.w3.org/TR/2012/CR-prov-o-20121211/>

[PROV-PRIMER]

Yolanda Gil; Simon Miles; eds. [*PROV Model Primer*](#). 11 December 2012, Working Draft. URL: <http://www.w3.org/TR/2012/WD-prov-primer-20121211/>

[PROV-XML]

Hook Hua; Curt Tilmes; Stephan Zednik; eds. [*PROV-XML: The PROV XML Schema*](#). 11 December 2012, Working Draft. URL: <http://www.w3.org/TR/2012/WD-prov-xml-20121211/>