



# Designing the W3C open annotation data model

**DOI:**

[10.1145/2464464.2464474](https://doi.org/10.1145/2464464.2464474)

[Link to publication record in Manchester Research Explorer](#)

**Citation for published version (APA):**

Soiland-Reyes, S., Sanderson, R., Ciccarese, P., & Van De Sompel, H. (2013). *Designing the W3C open annotation data model*. (W3C Community Draft). Association for Computing Machinery. <https://doi.org/10.1145/2464464.2464474>

**Citing this paper**

Please note that where the full-text provided on Manchester Research Explorer is the Author Accepted Manuscript or Proof version this may differ from the final Published version. If citing, it is advised that you check and use the publisher's definitive version.

**General rights**

Copyright and moral rights for the publications made accessible in the Research Explorer are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

**Takedown policy**

If you believe that this document breaches copyright please refer to the University of Manchester's Takedown Procedures [<http://man.ac.uk/04Y6Bo>] or contact [uml.scholarlycommunications@manchester.ac.uk](mailto:uml.scholarlycommunications@manchester.ac.uk) providing relevant details, so we can investigate your claim.





# Open Annotation Data Model

Community Draft, 08 February 2013

---

**This version:**

<http://www.openannotation.org/spec/core/20130208/index.html>

**Latest version:**

<http://www.openannotation.org/spec/core/>

**Previous version:**

<http://www.openannotation.org/spec/core/20130205/index.html>

**Editors:**

Robert Sanderson, Los Alamos National Laboratory  
Paolo Ciccarese, Massachusetts General Hospital and Harvard Medical School  
Herbert Van de Sompel, Los Alamos National Laboratory

**Contributors (in alphabetical order):**

Shannon Bradshaw, Dan Brickley, Leyla Jael García Castro, Timothy Clark, Timothy Cole, Phil Desenne, Anna Gerber, Antoine Isaac, Jacob Jett, Thomas Habing, Bernhard Haslhofer, Sebastian Hellmann, Jane Hunter, Randall Leeds, Andrew Magliozzi, Bob Morris, Paul Morris, Jacco van Ossenbruggen, Stian Soiland-Reyes, James Smith, Dan Whaley.

## Abstract

The Open Annotation Core Data Model specifies an interoperable framework for creating associations between related resources, annotations, using a methodology that conforms to the Architecture of the World Wide Web. Open Annotations can easily be shared between platforms, with sufficient richness of expression to satisfy complex requirements while remaining simple enough to also allow for the most common use cases, such as attaching a piece of text to a single web resource.

An Annotation is considered to be a set of connected resources, typically including a body and target, where the body is somehow about the target. The full model supports additional functionality, enabling semantic annotations, embedding content, selecting segments of resources, choosing the appropriate representation of a resource and providing styling hints for consuming clients.

## Status of this Document

*This section describes the status of this document at the time of its publication. Other documents may supersede this document.*

Copyright © 2012-2013 the Contributors to the Open Annotation Core Data Model Specification, published by the [Open Annotation Community Group](#) under the [W3C Community Contributor License Agreement \(CLA\)](#). A human-readable [summary](#) is available.

This specification was published by the [Open Annotation Community Group](#). It is not a W3C Standard nor is it on the W3C Standards Track. Please note that under the [W3C Community Contributor License Agreement \(CLA\)](#) there is a limited opt-out and other conditions apply. Learn more about [W3C Community and Business Groups](#).

This document has been made available to the Open Annotation Community Group for review, but is not endorsed by them. This is a working draft, and it is not endorsed by the W3C or its members. It is inappropriate to refer to this document other than as "work in progress".

Please send general comments about this document to the public mailing list: [public-openannotation@w3.org](mailto:public-openannotation@w3.org) ([public archives](#)).

---

## Table of Contents

1. **[Introduction](#)**
  - 1.1. [Aims of the Model](#)
  - 1.2. [Namespaces](#)
  - 1.3. [Terminology](#)
  - 1.4. [Examples](#)
2. **[Open Annotation Core](#)**
  - 2.1. [Body and Target Resources](#)
    - 2.1.1. [Typing of Body and Target](#)
    - 2.1.2. [Embedded Textual Bodies](#)
    - 2.1.3. [Tags and Semantic Tags](#)
    - 2.1.4. [Fragment URIs Identifying Body or Target](#)
    - 2.1.5. [Annotations without a Body](#)
    - 2.1.6. [Multiple Bodies or Targets](#)
  - 2.2. [Annotation Provenance](#)
    - 2.2.1. [Agents](#)
  - 2.3. [Motivations](#)
3. **[Module: Specifiers and Specific Resources](#)**
  - 3.1. [Specifiers and Specific Resources](#)
  - 3.2. [Selectors](#)
    - 3.2.1. [Fragment Selector](#)
    - 3.2.2. [Range Selectors](#)
      - 3.2.2.1. [Text Position Selector](#)
      - 3.2.2.2. [Text Quote Selector](#)
      - 3.2.2.3. [Data Position Selector](#)
    - 3.2.3. [Area Selectors](#)
      - 3.2.3.1. [SVG Selector](#)
  - 3.3. [States](#)
    - 3.3.1. [Time State](#)

- 3.3.2. [Request Header State](#)
  - 3.4. [Styles](#)
    - 3.4.1. [CSS Style](#)
  - 3.5. [Scope of a Resource](#)
  - 4. **[Module: Multiplicity Constructs](#)**
    - 4.1. [Choice](#)
    - 4.2. [Composite](#)
    - 4.3. [List](#)
  - 5. **[Module: Publishing](#)**
    - 5.1. [Serialization](#)
    - 5.2. [Embedding Resources](#)
    - 5.3. [Embedding RDF Graphs](#)
    - 5.4. [Equivalence of Resources](#)
  - 6. **[Appendices](#)**
    - A. [W3C Provenance Model Mapping](#)
    - B. [Extending Motivations](#)
    - C. [References](#)
    - D. [Acknowledgements](#)
    - E. [Change Log](#)
- 

## 1. Introduction

Annotating, the act of creating associations between distinct pieces of information, is a pervasive activity online in many guises but currently lacks a structured approach. Web citizens make comments about online resources using either tools built in to the hosting web site, external web services, or the functionality of an annotation client. Comments about photos on Flickr, videos on YouTube, people's posts on Facebook, or mentions of resources on Twitter could all be considered as annotations associated with the resource being discussed. In addition, there are a plethora of closed and proprietary web-based "sticky note" systems and stand-alone multimedia annotation systems. The primary complaint about these types of systems is that the user-created annotations cannot be shared or reused due to a deliberate "lock-in" strategy within the environments where they were created. The minimum requirement for any solution is a common approach to expressing these annotations.

The Open Annotation data model provides an extensible, interoperable framework for expressing annotations such that they can easily be shared between platforms, with sufficient richness of expression to satisfy complex requirements while remaining simple enough to also allow for the most common use cases, such as attaching a piece of text to a single web resource.

An annotation is considered to be a set of connected resources, typically including a body and target, and conveys that the body is related to the target. The exact nature of this relationship changes according to the intention of the annotation, but most frequently conveys that the body is somehow "about" the target. Other possible

relationships include that the body is an identifier for the target, provides a representation of the target, or classifies the target in some way. This perspective results in a basic model with three parts, depicted below. The full model supports additional functionality, enabling content to be embedded within the annotation, selecting arbitrary segments of resources, choosing the appropriate representation of a resource and providing styling hints for consuming clients. Annotations created by or intended for machines are also considered to be in scope, ensuring that the Data Web is not ignored in favor of only considering the human-oriented Document Web.

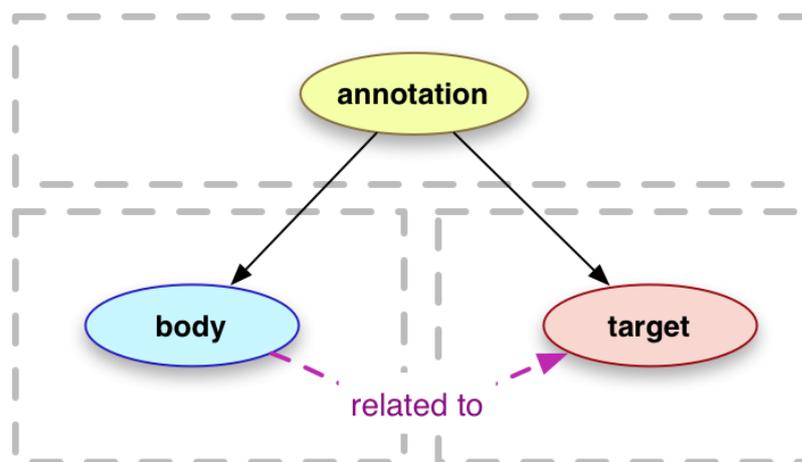


Figure 0.1. Annotation, Body and Target

Unlike previous attempts at annotation interoperability, the Open Annotation system does not prescribe a transport protocol for creating, managing and retrieving annotations. Instead it describes a web-centric method, promoting discovery and sharing of annotations without clients or servers having to agree on a particular set of network transactions to communicate those annotations.

The specification is divided into the essential core plus distinct modules that add functionality. The modules cover cases where the exact nature of the body or target cannot be sufficiently captured in a URI, explicit semantics for multiplicity and recommendations for publishing best practices.

## 1.1. Aims of the Model

The primary aim of the Open Annotation Data Model is to provide a standard description mechanism for sharing Annotations between systems. This interoperability may be either for sharing with others, or the migration of private Annotations between devices. The shared Annotations must be able to be integrated into existing collections and reused without loss of significant information. The model should cover as many annotation use cases as possible, while keeping the simple annotations easy and expanding from that baseline to make complex uses possible.

A single, consistent model that can be used by all interested parties is the goal of the standardization process. The number of RDF triples

required or bytes needed for serializations, while a consideration, is less important than the coherency of the model. All efforts are made to keep the implementation costs for both producers and consumers to a minimum. A single method of fulfilling a use case is strongly preferred over multiple methods, unless there are existing standards that need to be accommodated or there is a significant cost associated with a method that is otherwise necessary.

The methods of storage and maintenance for Annotations are not specified by the model: databases do not need to be restructured, RDF triplestore technologies do not need to be used, and existing websites and interfaces do not need to be re-scripted or re-engineered. The only requirement is that at least one serialization of the model describing the Annotations be made available for other systems to retrieve, potentially along side serializations into other models. Every effort has been made to ensure this mapping from internal structures to the Open Annotation Data Model is as clear and straightforward as possible.

## 1.2. Namespaces

The Open Annotation model defines a namespace for its classes and properties, and uses several others as listed below. The namespace URI will always remain the same, even if the ontology changes. All versions of the ontology will remain available from version-specific URLs, and the namespace URI will provide access to the most recent version.

The following namespaces are used in this specification:

Prefix	Namespace	Description
oa	<a href="http://www.w3.org/ns/oa#">http://www.w3.org/ns/oa#</a>	The Open Annotation ontology
cnt	<a href="http://www.w3.org/2011/content#">http://www.w3.org/2011/content#</a>	<a href="#">Representing Content in RDF</a>
dc	<a href="http://purl.org/dc/elements/1.1/">http://purl.org/dc/elements/1.1/</a>	<a href="#">Dublin Core Elements</a>
dcterms	<a href="http://purl.org/dc/terms/">http://purl.org/dc/terms/</a>	<a href="#">Dublin Core Terms</a>
dctypes	<a href="http://purl.org/dc/dcmitype/">http://purl.org/dc/dcmitype/</a>	<a href="#">Dublin Core Type Vocabulary</a>
foaf	<a href="http://xmlns.com/foaf/0.1/">http://xmlns.com/foaf/0.1/</a>	<a href="#">Friend-of-a-Friend Vocabulary</a>
prov	<a href="http://www.w3.org/ns/prov#">http://www.w3.org/ns/prov#</a>	<a href="#">Provenance Ontology</a>
rdf	<a href="http://www.w3.org/1999/02/22-rdf-syntax-ns#">http://www.w3.org/1999/02/22-rdf-syntax-ns#</a>	<a href="#">RDF</a>

rdfs	<a href="http://www.w3.org/2000/01/rdf-schema#">http://www.w3.org/2000/01/rdf-schema#</a>	<a href="#">RDF Schema</a>
skos	<a href="http://www.w3.org/2004/02/skos/core#">http://www.w3.org/2004/02/skos/core#</a>	<a href="#">Simple Knowledge Organization System</a>
trig	<a href="http://www.w3.org/2004/03/trix/rdfg-1/">http://www.w3.org/2004/03/trix/rdfg-1/</a>	<a href="#">TriG Named Graphs</a>

Editor's Note: The "Content in RDF" specification is considered stable according to its editors, but at time of publication is still a Working Draft. The status of the "TriG Named Graph" ontology is unknown. The use of these ontologies may be revisited in the future to take into account activities that impact them.

### 1.3. Terminology

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in [RFC 2119](#).

### 1.4. Examples

The examples throughout these documents will be conveyed as both a diagram and in the [Turtle](#) RDF format, and do not represent specific use cases with real resources. The Turtle examples do not provide namespace declarations, and should be considered to follow the [Namespaces table](#) above. Usage examples in [SPARQL](#) are given for each section, based on a query expressed in natural language. Additional examples of how to model and implement specific situations are available in the [Tutorial](#) and the [Annotation Cookbook](#).

The diagrams in the specification use the following style:

- Instances are depicted as colored ellipses
  - Instances with a resolvable URI have a single line border
  - Instances with a non-resolvable URN or are a blank node have a double line border
- Classes are depicted as white rectangles
- Literals are depicted as white lozenges
- Relationships are depicted as straight, black lines. Relationships are RDF predicates where the range is a Resource, and equivalent to OWL object properties.
- Properties are depicted as curved, black lines. Properties are RDF predicates where the range is a Literal, and equivalent to OWL datatype properties.
- Class instantiation (`rdf:type`) is depicted as a straight black line with white arrow head, following UML.
- Example instance identifiers are lowercase and end in a

number.

For example, `anno1` is a specific instance of an Annotation, whereas `oa:Annotation` is a class

- Example literals follow the requirements for the model and, thus, must not be interpreted as the only possible value
- Conceptual resource boundaries not explicit in the model, but considered important for understanding, are depicted as grey dashed boxes around the components. They are used to convey spatial parts of the diagrams and may be safely ignored.

---

[contents next](#)

---