**Open Citation: Definition --** Silvio Peroni and David Shotton

MUST​: the definition is an absolute requirement of the specification; MAY​: an item is truly optional.

**Preliminaries :** bibliographic citation: **conceptual directional link from a citing entity to a cited entity**, for the purpose of acknowledging or ascribing credit for the contribution made by the author(s) of the cited entity. A citation can be:

* **directly derived** ​by looking at the reference within the citing entity;
* **derived ​by querying a database** **containing bibliographic information** about the citing and/or cited entity, if the metadata for such an entity includes details of its outgoing references or of its incoming citation links;
* **retrieved from a database of citation information** (a citation index), which will typically contain a definition of the citation itself, information about the citing and the cited entity, and associated metadata.

**Definition of an open citation :** A bibliographic citation is an open citation **when the data needed to define the citation are freely available**, downloadable and reusable. Specifically, such data **MUST be compliant with the ‘SSO Principles’** introduced by the Initiative for Open Citations

* Structured​ – expressed in machine-readable formats;
* Separate – available without access the source bibliographic entity in which the citation is defined
* Open – freely accessible and reusable without restrictions

two further principles -- citing and cited entities must be:

* Identifiable –by a specific persistent identifier scheme (e.g. a DOI) or a URL
* Available – by resolving their identifiers basic metadata of both the entities must be reachable.

All such entity metadata MUST follow the same SSO principles. Additional metadata MAY be available.

**Open Citation Identifier: Definition --** Silvio Peroni and David Shotton

**Preliminaries**: A bibliographic citation is a **conceptual directional link from a citing entity to a cited entity**, for the purpose of acknowledging or ascribing credit for the contribution made by the author(s) of the cited entity. The citation itself, once the citing work is published, becomes an enduring component of the academic ecosystem. Analyses of citations can both reveal how scholarly knowledge develops over time and also be used to assess scholars’ influence and make wise decisions about research investment. Metadata describing a citation can be stored in a database containing bibliographic information. OpenCitations recently proposed an alternative view of the citation as a data entity in its own right, so that:

1. all the information regarding each citation can be stored in one place;
2. citations become easier to describe, distinguish, count and process;

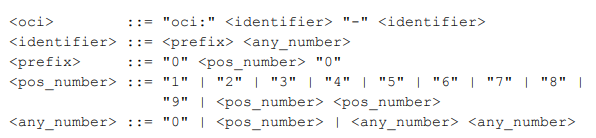
3. citations become easier to analyse using bibliometric methods (visualize citation networks / determine how citation time spans vary by discipline)

4. citations can be identified by means of globally unique and persistent citation identifiers, so to reference the citations within a specific bibliographic database. **This document defines of one particular identifier for citations, the Open Citation Identifier (OCI)**, used to provide identifiers for open citations .

**Definition of an Open Citation Identifier**:(OCI) is a globally unique persistent identifier (PID) for the identification of open bibliographic citations stored in a specific database or in other kinds of storages.

The Open Citation Identifier system, and a resolution service for OCIs that returns metadata about the identified citations, have been developed by OpenCitations.

OCI **structure**: the **lower-case letters “oci” + colon + two sequences of numerals separated by a dash**. Precisely, an OCI is defined as follows (in Backus-Naur form):



For an OCI to be valid, **both identifiers must refer to entities stored in the same database provided by a single supplier**. One cannot create an OCI by combining identifiers of entities recorded in different databases.

**Meaning:**

- **first identifier**: identifier for the citing bibliographic resource

- **second identifier**: identifier for the cited bibliographic resource

Each OCI MUST carry the following pieces of information:

1. **prefix** specifying the supplier’s database in which the bibliographic citation is recorded. encoded by the part of each half of the OCI, which MUST be assigned to the same supplier for both halves of the OCI, and which it is assigned by OpenCitations. This prefix is the same in both halves of the OCI.
2. **numerical identifiers** used in that database, or the encoded numerical equivalent of (part of) the identifiers used in that database. MUST clearly identify both the citing and cited entities of each citation by assigning them a particular internal identifier compliant with one and only one identifier scheme. the OpenCitations Corpus uses URLs following the pattern “https://w3id.org/oc/corpus/br/” to identify bibliographic entities

The sequence of number in each identifier part of an OCI is a numerical encoding of the original supplier’s identifier for the citing/cited entity. For instance: “010” is assigned to Wikidata, “020” is assigned to Crossref, “030” is assigned to the OpenCitations Corpus.

For every prefix identifying an individual database supplier of bibliographic and citation information, a particular conversion rule is applied to move from the unique identifiers used within the supplier’s database to the numerical identifiers of the citing and cited entities used in the OCI to identify that citation.

**Prefixes :**The prefix used for building an OCI defines the supplier of the original citation data (a database with bibliographic information). The list of currently supported : http://opencitations.net/oci.

**Software:** OpenCitations provides an Open Citation Identifier Resolution Service that takes an OCI and returns information about the identified citation: https://w3id.org/oc/oci.