



# Scholix Metadata Schema for Exchange of Scholarly Communication Links

---

Version 2.0

21 February 2017

Adrian Burton, Wouter Haak, Martin Fenner, Paolo Manghi

Members of the Metadata Working Group

Amir Aryani

Adrian Burton

Geoffrey Bilder

Michael Diepenbroek

Wouter Haak

Martin Fenner

Paolo Manghi

Uwe Schindler

Gerry Ryder

Florian Graef

# Table of Content

<b>Introduction</b>	<b>3</b>
Scholix	3
The Scholix Community participation	3
The Objectives of this schema	3
The Scholix metadata schema	3
Schema updates	4
<b>Scholix Metadata properties</b>	<b>5</b>

# Introduction

## Scholix

The Scholix initiative is a high level interoperability framework for exchanging information about the links between scholarly literature and data. It aims to enable an open information ecosystem to understand systematically what data underpins literature and what literature references data.

Scholix is an evolving lightweight set of [Guidelines](#) to increase interoperability rather than a normative standard. It consists initially of:

- Conceptual model
- Information model
- Information standards and encoding guidelines
- Options for exchange protocols

## The Scholix Community participation

Scholix is the “wholesaler to wholesaler” exchange framework, to be implemented by existing hubs or global aggregators of data-literature link information such as DataCite, CrossRef, OpenAIRE, and EMBL-EBI. These hubs in turn work with their natural communities of data centres or literature publishers to collect the information through existing community-specific workflows and standards.

The current vehicle for the Scholix initiative is the Scholarly Link Exchange Working Group, a joint initiative of the Research Data Alliance and the World Data System.

## The Objectives of this schema

## The Scholix metadata schema

The Scholix metadata schema is a set of properties describing a *Link Information Package*, which carries information about a *link* between two scholarly communication objects. An high-level view of the metadata schema is illustrated in Figure 1.

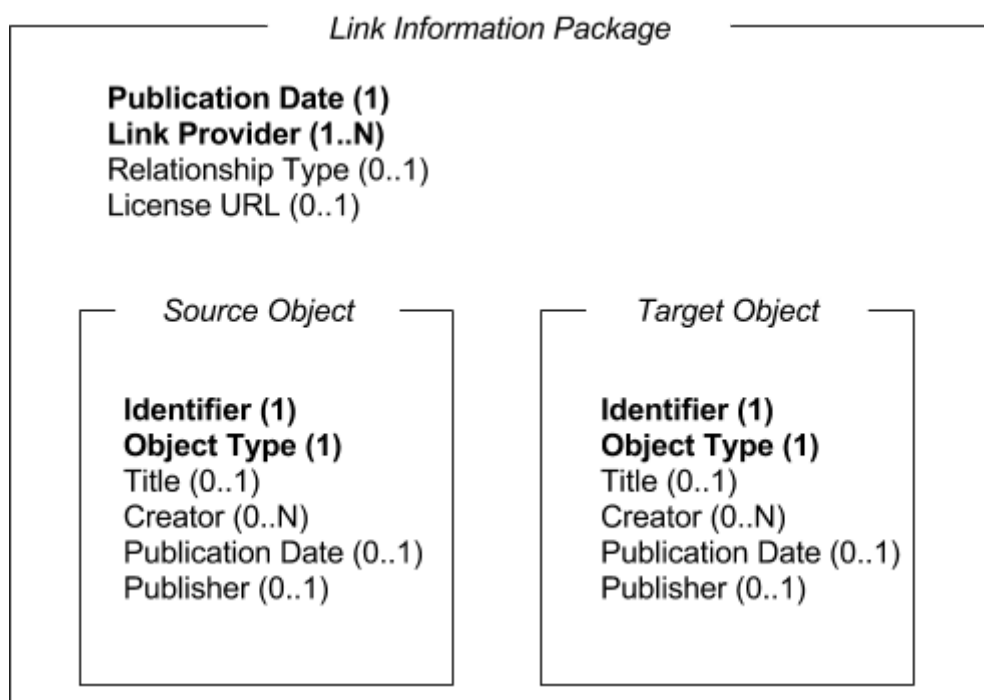


Figure 1 - Link Information Package: high-level view

## Schema updates

Author	Date	Update description	Version
Paolo Manghi	29/11/2016	<ul style="list-style-type: none"> <li>Changed property 1 in Table 1</li> <li>Changed property 2 of Table 2</li> </ul>	0.0
Paolo Manghi	20/12/2016	<ul style="list-style-type: none"> <li>One mandatory PID for source and target object</li> <li>Relationships semantics from mandatory vocabulary</li> </ul>	1.0
Paolo Manghi	21/01/2017	<ul style="list-style-type: none"> <li>Make Link Publisher optional and name changed to Link Provider</li> <li>Make Link Provider mandatory</li> <li>Add optional License URL field</li> <li>Drop glossary</li> </ul>	2.0

## Scholix Metadata properties

The properties of the Scholix Metadata Schema are presented in Table 1, for the Link properties, and Table 2, for the object properties. Each table provides, for each property, an identifier, a definition, domain, and multiplicity (Mul.), which can be:

- 0..N = optional and repeatable
- 0..1 = optional, but not repeatable
- 1..N = required and repeatable
- 1 = required, but not repeatable

Some of the properties are “structured”, which means they do not characterize a value but a nested record, with internal properties. Structured properties are indicated by a *[structured property]* tag.

**Table 1** - Link Information Package properties

ID	Property	Mul.	Definition	Domain (allowed values, potential application profiles, constraints, examples)
1	<b>Link publication date</b>	1	date of formal issuance (e.g., publication) of the resource; generally different from source object and target object publication dates	dc:date (qualifier “issued”)
2	<b>Link provider</b>	1..N	The source from which this record was collected. Can be one or multiple.	[structured property]
2.1	Provider Name	1		e.g. Crossref
2.2	Identifiers	0..N		[structured property]
2.2.1	ID	1		
2.2.1	ID schema	1		e.g. ISNI
3	<b>Relationship Type</b>	0..1	Semantics of the relationship from source to target	[structured property]
3.1	RelName	1		Term from DataCite vocabulary
3.2	InverseRelName	0..1		Term from DataCite vocabulary
4	License URL	0..1		

**Table 2** - Source and Target Object properties

ID	Property	Occ.	Definition	Allowed values, examples, application profiles other constraints
1	<b>Identifier</b>	<b>1</b>		[structured property]
1.1	ID	1		
1.2	ID schema	1		e.g. DOI, PDB, URL
2	<b>Object type</b>	<b>1</b>	Describes the nature of the object (its intended usage)	[structured property]
2.1	Type	1		e.g. publication, dataset
2.2	Sub-type	0..1		e.g. for literature type: article, report, pre-print, thesis
3	Title	0..1		dc:title
4	Creator	0..N		[structured property]
4.1	Name	1		Recommended template: Name, Surname, e.g. "John H., Smith"
4.2	Identifiers	0..N		[structured property]
4.2.1	ID	1		
4.2.1	ID schema	1		e.g. ORCID, VIAF, DAI
5	Publication Date	0..1		dc:date (qualifier "issued")
6	Publisher	0..1		[structured property]
6.1	Publisher Name	1		e.g. "PANGAEA"
6.2	Identifiers	0..N		[structured property]
6.2.1	ID	1		
6.2.1	ID schema	1		e.g. ISNI