ISQ

TOPIC

OPEN ACCESS INFRASTRUCTURE

NISO

How the information world
CONNECTS

NR [NISO REPORTS]

Cameron
Neylon

Ed
Pentz

Greg
Tananbaum

CAMERON NEYLON, ED PENTZ, AND GREG TANANBAUM

Standardized Metadata Elements to Identify Access and License Information

Many journal articles are available from publishers under the banner of Open Access (OA), Public Access, or similar names. The meanings of these terms vary both between publishers and within publishers by journal—and in some cases, based on the funder. Adding to the potential confusion, a number of publishers also offer hybrid options, in which one or more articles in a journal are freely accessible, while the rest of the content in that journal remains under subscription control.

The guide *HowOpenIsIt?* from SPARC, PLOS, and OASPA depicts a continuum of openness that also varies by the rights accorded to readers, reuse rights, copyrights, author posting rights, automatic posting, and machine readability. Clearly, as the Guide points out, "not all Open Access is created equal." Currently, there is no standard metadata in use that succinctly defines these various levels of openness and licensing. As a result, readers are often unaware of the free-to-read status of specific articles and downstream users are unsure of the reuse rights, if any. Authors have difficulty determining what rights they will retain and whether they are compliant with a given funder policy. Aggregators and service providers have no machine-readable mechanism for identifying articles that can be legitimately harvested.

In January 2013, NISO Voting Members approved a new work item proposal to develop a Recommended Practice on Open Access Metadata and Indicators (later re-named Access and Licensing Indicators) to address this gap. The goal of the project was to identify a standardized set of metadata elements to describe both the accessibility of a specific article and the available reuse rights.

A draft for comments Recommended Practice was issued in January 2014 proposing the adoption of two core pieces of metadata that can be transmitted through existing channels:

» **Free-to-read** (<free_to_read>) –
A simple status that defines whether the work is accessible, without charge or other restriction (such as registration), to read online. This tag has two defined attributes that should be used, if applicable, to indicate start and end dates. Start and end dates would accommodate delayed access models (embargoes) and special offers where content was free-to-read for a period of time or after a particular date. The absence of both a start and end date would mean a permanent state of free-to-read access.

» **License reference** (<license_ref>) –
A reference to a URI that carries the license terms specifying how a work may be used. There are no limitations on the license specified or on the terms contained within the license. Multiple license reference elements can be provided. Each of these may have a different start dates to address embargoes or how usage rights change over time. There is no corresponding end date attribute for the <license_ref> element, because including end dates could introduce ambiguities. The data within this tag should be a stable identifier expressed as an HTTP URI, the maintenance of which would be the responsibility of the platform making the content available.

The Working Group specifically decided against proposing metadata items that were labeled or named "Open Access" due to the many different definitions of this term, as discussed above. Instead, the chosen approach was to provide factual metadata to be disseminated to enable people and machines to make decisions about how they can use the content. With widespread implementation of these recommended metadata tags, humans and machines will be able to assess the accessibility and reuse rights associated with a given article.

The Working Group considered and rejected the expression of reuse rights in the actual metadata. These rights could vary depending on who the user is and it could be difficult to fully and accurately express them in metadata, possibly creating a conflict or inconsistency with the actual license. Therefore,

> The Working Group specifically decided against proposing metadata items that were labeled or named "Open Access" due to the many different definitions of this term.

the agreed approach was to have a reference in the metadata to the license that would be posted separately and linked from the metadata reference.

It is the view of the working group that these two metadata elements can cover most current use cases of delayed access and of license terms that activate at a particular time post publication. Use cases fully addressed include:

» End user seeks to discover, identify, and access free-to-read items

» End user seeks to know the readability status of an item

» End user seeks to know reuse permissions of an item

» End user seeks to know reuse permissions of a sub-component of an item

» Repositories seek to expose free-to-read items

Use cases that are at least partially addressed by the new elements are:

» End user seeks to text mine content

» Ensure author/publisher rights assertions align with license statements

» Funding agency seeks to track compliance of research outputs to open access mandates

» Institution seeks to report on open access compliance of research outputs

While it was outside the scope of this Recommended Practice to determine how components of works (e.g., figures, images, datasets, etc.) should be identified, where such components are separately identified, the <free_to_read> and <license_ref> tags can be applied separately to those components.

Wherever possible, creation and population of these elements should become part of standard editorial/production workflows. The metadata should be made an integral part of the feeds to CrossRef and other DOI registration agencies, included alongside (or within) article/chapter content on hosting websites, and delivered in content feeds to third parties. The metadata should be embedded in the content itself along with other metadata; for example, in HTML META tags and in PDF files where bibliographic and other metadata are being included.

The Working Group is also recommending that the "free-to-read" and "license reference" metadata be encoded in XML and included in existing metadata distribution channels and with the content itself, where appropriate. Thus the <free_to_read> and <license_ref> tags would need to be added to existing schemas and workflows. Publisher or aggregator systems could be programmed to read the tags and display appropriate status icons to users.

It may also be worthwhile for content providers to consider including the metadata elements within other alerting channels, such as e-ToCs and RSS subscription feeds as well as information provided directly to abstracting and indexing services. Whatever channel is used, wider distribution of this (and other) article, chapter, or book metadata is likely to be helpful in driving discovery and usage for the materials concerned.

The Working Group is currently finalizing the Recommended Practice to address issues identified during the public comment period. The final document is expected to be published in the fall of 2014.

The Group recognizes that if the recommendations are adopted, there will need to be further work on implementation and an analysis done on the best way to incorporate the <free_to_read> and <license_ref> metadata into existing formats, such as ONIX, RDF, OAI-PMH, and Dublin Core (DC). NISO will be looking into the need for a Standing Committee to work on these follow-up items. I NR I doi: 10.3789/isqv26no2.2014.07

CAMERON NEYLON (cneylon@plos.org) is Advocacy Director for the Public Library of Science (PLOS).
ED PENTZ (epentz@crossref.org) is Executive Director, CrossRef.
GREG TANANBAUM (greg@scholarnext.com) is Consultant, SPARC Scholarly Publishing & Academic Resources Coalition

The three are the co-chairs of the NISO Access and License Indicators Working Group (formerly Open Access and Metadata Indicators).

CrossRef
http://www.crossref.org

HowOpenIsIt? SPARC, PLOS, and OSPA, 2013.
http://www.plos.org/open-access/howopenisit/

NISO Access & Licensing Indicators Working Group webpage
http://www.niso.org/workrooms/ali/

RELEVANT
LINKS