NISO RP-22-2021

Access & License Indicators

A Recommended Practice of the
National Information Standards Organization

About NISO Recommended Practices

A NISO Recommended Practice is a recommended "best practice" or "guideline" for methods, materials, or practices in order to give guidance to the user. Such documents usually represent a leading edge, exceptional model, or proven industry practice. All elements of Recommended Practices are discretionary and may be used as stated or modified by the user to meet specific needs.

This recommended practice may be revised or withdrawn at any time. For current information on the status of this publication contact the NISO office or visit the NISO website (www.niso.org).

This specification is provided "as is," and neither the developer nor the user of this specification makes any representation or warranty whatsoever in connection with its use, including with respect to the currency, authenticity, accuracy, machine readability, or completeness of any rights or restrictions that may be found at the destination of the noted HTTP URI provided in the rights metadata field, the assessment and use of which are the sole responsibility of the person or entity making use of this specification.

Published by
National Information Standards Organization (NISO)
3600 Clipper Mill Road, Suite 302
Baltimore, MD 21211
www.niso.org

ISBN: 978-1-950980-17-8

DOI: https://doi.org/10.3789/niso-rp-22-2021

Contents

Foreword iv

Section 1 : Introduction 1

1.1 Purpose and Scope 1

1.2 Terms and Definitions 1

Section 2 : Summary of Recommendations 3

Section 3 : Metadata Elements 5

3.1 free_to_read 5

3.2 license_reference 5
3.2.1 Description 5
3.2.2 `start_date` Attribute 6
3.2.3 `applies_to` Attribute 7
3.2.4 Use of `license_reference` 7
3.2.5 License URI Expression 7

Section 4 : Recommended Mechanisms for Distributing Metadata 9

Section 5 : Use Case Review 10

5.1 Use Case: End User Seeks to Discover, Identify, and Access Free-to-Read Items 10

5.2 Use Case: End User Seeks to Know the Readability Status of an Item 10

5.3 Use Case: End User Seeks to Know Reuse Permissions of an Item 11

5.4 Use Case: End User Seeks to Know Reuse Permissions of a Sub-Component of an Item 11

5.5 Use Case: Repositories Seek to Expose Free-to-Read Items 11

5.6 Use Case: End User Seeks to Text Mine Content 12

5.7 Use Case: Ensure Author/Publisher Rights Assertions Align with License Statements 12

5.8 Use Case: Funding Agency Seeks to Track Compliance of Research Outputs to Open Access Mandates 12

5.9 Use Case: Institution Seeks to Report on Open Access Compliance of Research Outputs 13

5.10 Use Case: Platform Seeks to Know the Sharing Terms of an Article 13

Appendix A: Means of Expressing Metadata Elements 14

Bibliography 22

Foreword

About this Recommended Practice

In 2013, a new work item proposal was approved by NISO members to develop metadata and indicators that would provide information on whether a specific article is openly accessible (i.e., can be read by any user who can get to the journal website over the internet) and what reuse rights might be available to the reader.

Many offerings are available from publishers under the banner of open access, increased access, public access, or other names. The terms used 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 authors of an article can pay a fee to make their paper freely available to readers while the rest of the content in that journal remains under subscription control.

Currently, publishers provide articles that are "free to read" under a wide range of reuse terms and licenses. Some publishers and organizations favor the Creative Commons licenses, specifically the Creative Commons Attribution license (CC BY) to provide reuse terms, while others use proprietary or modified licenses for this purpose.

Funders find that the lack of information and cooperation between stakeholders creates difficulty in determining whether a specific article is compliant with their policies. Publishers of hybrid journals have no simple mechanism for signaling the "free to read" status of specific articles or the reuse rights of downstream users. Authors have difficulty confirming whether they are compliant with a given funder policy. Readers face the burden of figuring out what they can and cannot do with specific articles. Aggregators and service providers have no single mechanism for identifying articles that can be legitimately harvested.

The objectives of the project were to develop:

1. A specified format for bibliographic metadata and, possibly, a set of visual signals describing the readership rights associated with a single scholarly work.
2. Recommended mechanisms for publishing and distributing this metadata.
3. A report on the feasibility of including clear information on downstream reuse rights within the current project and, if judged feasible, inclusion of these elements in outputs 1 and 2.
4. A report stating how the adoption of these outputs would answer (or not) specific use cases to be developed by the working group.

About the 2021 Revision

In September 2020, a new NISO work item proposal was approved by NISO members to develop metadata and indicators that enable metadata users—including content platforms—to further qualify or scope the license being conveyed. For example, allowing publishers and platforms to utilize ALI tagging to establish sharing policies for a given research article using a specific permissioning framework. To meet this need, an additional attribute definition (`applies_to`) is added to this Recommended Practice.

NISO Information Discovery & Interchange Topic Committee Members

The following individuals served on the NISO Information Discovery & Interchange Topic Committee that had oversight for this work and approved the 2021 Revision of this Recommended Practice:

Robert Boissy
Springer Nature

Tamir Borensztajn
EBSCO Information Services

Mark Dehmlow
University of Notre Dame Libraries

Peter Murray, Co-chair
Index Data

Michael Roberts
Emerald Publishing Group

Doralyn Rossmann
Montana State University

Bob Schulz
OCLC Online Computer Library Center

Christine Stohn, Co-chair
Ex Libris, Inc.

Jan Waterhouse
University at Albany Libraries

NISO Access & License Indicators Working Group Members

The following individuals served on the NISO Access & License Indicators Working Group (originally named Open Access Metadata and Indicators), which developed and approved this Recommended Practice:

Tim Devenport
EDItEUR

Gregg Gordon
Social Science Research Network (SSRN)

Julie Hardesty
Indiana University Bloomington Libraries

Cecy Marden
The Wellcome Library

Cameron Neylon, Co-chair
Public Library of Science (PLoS)

John Ochs
American Chemical Society (ACS)

Edward Pentz, Co-chair
CrossRef

Heather Reid
Copyright Clearance Center (CCC)

Jill Russell
University of Birmingham

Chris Shillum
Reed Elsevier

Ben Showers
JISC Collections (through November 2014)

Eefke Smit
International Association of STM Publishers

Christine Stohn
Ex Libris, Inc.

Greg Tananbaum, Co-chair
Scholarly Publishing & Academic Resources Coalition (SPARC)

Timothy Vollmer
Creative Commons

Maarten Zeinstra
Kennisland

NISO Access & License Indicators 2021 Revision Working Group Members

Jeffrey Beck
National Center for Biotechnology Information (NCBI), U.S. National Library of Medicine

Patricia Feeney
Crossref

Esther Heuver, Co-chair
Mendeley/Elsevier

Vincent Lizzi
Taylor & Francis

Dan O'Brien, Co-chair
American Chemical Society

Acknowledgements

The Access & License Indicators Working Group wishes to acknowledge the following persons outside the formal working group membership who contributed to this effort.

- Jeff Beck, National Center for Biotechnology Information, National Library of Medicine
- Geoffrey Bilder, CrossRef
- Paul Mostert, Elsevier
- Evan Owens, Cenveo Publisher Services
- Members of the NISO *oa-indicators-info* interest list, who provided significant input to the development of the use cases listed in Section 5.

The Access & License Indicators 2021 Revision Working Group wishes to acknowledge the following person who contributed to this effort:

- Miriam Kessler, formerly of Springer Nature

Trademarks, Service Marks

Section 1: Introduction

1.1 Purpose and Scope

This recommended practice defines metadata indicators to be used to indicate free-to-read content and a link to license terms for the use/reuse of that content. Humans and machines will be able to assess the status of the content based on these indicators, and in many cases the combination of the `free_to_read` and `license_reference` metadata will indicate open access content. The indicators include a date component so that content with access and reuse rights that change over time can be adequately understood by both humans and machines using the metadata.

This recommended practice also defines metadata indicators that may be used to be used to specify a specific framework or context in which permissions are being conveyed.

1.2 Terms and Definitions

The following terms, as used in this recommended practice, have the meanings indicated.

Term	Definition
DOI	Digital Object Identifier: a framework for uniquely identifying and digitally linking content, solving for persistence (avoiding broken links) and locations of authority versions. Defined as ISO 26324.
embargo	A period during which a work is available only to users who have paid for or licensed access. Outside of the embargo period, the work may become free to read.
free to read	A work that is accessible to read online without charge or authentication (including registration) to any person with access to the internet. Note: The work may not necessarily have this status throughout its lifetime.
hybrid journal	A journal title that contains some open access or free-to-read articles, while other articles require payment and/or a subscription to access and read.
license	Agreement related to the use of a work. Generally includes payment requirements, if any, and terms of use and reuse—potentially including access, reproduction, adaptation, and distribution, among others.
namespace	A prefix added to XML element names to prevent conflicts when XML documents from different XML applications are combined.

Term	Definition
permissioning framework	A method to enable a rightsholder to assert, and a content platform to determine, the policies that convey the effective permissions and context(s) for making content available.
paywall	An access control system that limits access to electronic content to those users with a paid subscription or license.
public license	A license (or waiver) offered to any potential user of a resource. Note: The `license_reference` attribute defined in this Recommended Practice covers only public licenses.
text mining	The process of analyzing text to derive patterns, trends, or other meanings.
URI	A string of characters used to identify the name of a resource. Such identification enables interaction with representations of the resource over the web. The most common form of a URI is the uniform resource locator (URL), frequently referred to informally as a web address.
waiver	Language that releases a content owner's interests in particular works and places them as completely as possible in the public domain. The CC0 tool, created by Creative Commons, is an example of a waiver.

Section 2: Summary of Recommendations

The NISO Access & License Indicators ("ALI") Working group was chartered to develop protocols and mechanisms for transmitting the access status of scholarly works. This is a contentious area where political views on modes of access lead to differing interpretations of what constitutes "open access." The working group included a wide range of stakeholders with very different views of this area.

What was agreed among all stakeholders was that there is value in transmitting two key pieces of information relating to a scholarly work as a full-text online document. The first of these relates to whether a specified work is free to read—meaning that any person with access to the internet can access the work via the web. The second issue is that of what reuse rights are granted to a reader of a work.

A wide range of mechanisms for transmitting this information, in addition to the scope of information that would be transmitted, was considered. Taking into account the political, legal, and technical issues involved, a simple approach for transmitting a minimal set of information was preferred. The Working Group therefore recommends the following:

1. Adoption of two core pieces of metadata that can be transmitted through existing channels:
    a) `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 status can optionally carry a start date to define the time point when the work will become free to read. It may also specify an end date when it will cease to be free to read, where appropriate. (See Section 3.1.)
    b) `license_reference` – A reference to a URI that carries the license terms specifying how a work may be used. There will be 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 date. (See Section 3.2.)
2. Eschewal of a newly created logo or icon to be displayed along with the `free_to_read` element. The group discussed this issue at length and decided that creating a logo would be out of scope. The goal of the group is to create a simple means for metadata to be exchanged and each site or system will determine how best to convey the status of content to their users.

It is the view of the Working Group that these two metadata elements can cover most current use cases of delayed access (embargoes) and of license terms (see Section 5) that activate at a particular time post publication. While not completely covering all use cases, this simple approach provides a framework for conveying essential information that addresses the most common use cases.

The Working Group recommends that the `free_to_read` and `license_reference` metadata will be encoded in XML, that a namespace be declared for these elements, and that they are included in existing metadata distribution channels and with the content itself where appropriate. Rather than expecting all existing metadata formats and schemas to adopt equivalent elements, this Recommended Practice declares an XML namespace for the `<free_to_read>` and `<license_ref>` XML elements that determines how they should be added to existing schemas and workflows. The `free_to_read` and `license_reference` metadata can be used by systems to display appropriate status icons to users.

For more on specific use cases, see Section 5.

Rationale

The Working Group decided against proposing metadata items that were labeled or named "open access," as there are different definitions of this term. The approach agreed to by the Working Group was to provide factual metadata to be disseminated to enable people and machines to make decisions according to their own criteria.

Another important decision of the original ALI Working Group, for a number of reasons, was that reuse rights should not be expressed in metadata. Reuse rights will vary depending on who the user is (e.g., someone affiliated with an institution with a site license versus an unaffiliated/unrecognized user). If rights are described both in metadata and in a license, there are almost certain to be incompatibilities of expression and therefore inconsistencies. No one will be able to trust the metadata, because they know they are bound by the legal document and the metadata could be useless or actively misleading. For corporate or institutional users, no legal team is going to agree to allow any specific use based on metadata unless they have agreed that the license allows it. Therefore, the approach agreed upon is to have a reference in the metadata to a license.

In the context of user reuse rights it is also important to note that the question of which parties have legal rights over various versions of scholarly manuscripts remains a point of contention both in terms of the legal situation and community views. For the purpose of this Recommended Practice we assume as a foundational principle that the party acting as publisher and/or host of the digital object is the source of or authority for transmitted metadata on access status and that it has legal standing and authority to assert the rights specified. However, it should be clear that neither NISO, nor the Working Group, nor the Recommended Practice can require that this be the case. As a matter of underlying basic best practice, all stakeholders should only assert or offer rights for which they reasonably believe they have authority and standing.

The Working Group also elected not to recommend adoption of particular visual indicators to signify the `<free_to_read>` and `<license_ref>` elements. Input from the community indicated that there is a plethora of badges within the ecosystem already, and the stakeholders we caucused were not overly enthusiastic about adding additional visual indicators for the proposed elements. Additionally, and with respect to license conditions, there was consensus that existing visual indicators such as the Creative Commons icon sufficiently overlap with this project that the addition of a complementary badge might cause confusion.

The Working Group for the 2021 revision of the ALI recommendation continues to uphold the basic premise of not directly expressing reuse rights within metadata, for the same reasons already expressed. However, the Working Group also recognizes a need has emerged to allow for providing additional context or qualification for the conveyed license indicators. For example, identification of the larger permissioning framework from which the license identifiers were assigned.

Section 3: Metadata Elements

Two core pieces of metadata should be adopted that can be transmitted through existing channels, as defined in this section.

3.1 free_to_read

A number of the key use cases can be addressed by a simple metadata field, which will often be expressed as an XML element, to indicate that content can be read or viewed at its current location by any user without payment or authentication—the `<free_to_read>` element. The absence of the `free_to_read` metadata element doesn't indicate anything, and the user needs to look at other information to ascertain if s/he is able to read a copy.

The presence of the `free_to_read` element would provide a very simple indication of the status of the content without making statements about any additional reuse rights or restrictions. In some situations, the presence of the `free_to_read` element would be enough to provide a visual indication to the user that the content was available to read without payment or authentication. Whatever system is using the metadata needs to decide what visual indicator to show a user based on the presence of the `free_to_read` element.

The `free_to_read` element 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, special offers, etc., 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 indicates a permanent state of free to read access. Any date elements must follow ISO 8601. To maximize interoperability and ensure compatibility with IETF RFC 3339, the format YYYY-MM-DD is strongly recommended. Times and time zones should not be included.

- `start_date` – Use of this attribute indicates that the content is not publicly available immediately and/or indefinitely. For example, it may not be available from its date of publication but becomes available at some future time; i.e., there is a temporary embargo on its free availability. Or alternatively, it might be available in the first few months of publication, but then go behind a paywall. In the latter case, the `start_date` attribute would need to be used in combination with the `end_date` attribute.

  For example, if an article published on February 3, 2014, does not become freely available to all readers until one year after publication (i.e., on February 3, 2015), the metadata would be encoded as:

  ```
  <free_to_read start_date="2015-02-03"/>
  ```

- `end_date` – This attribute will be used where there is a temporary `<free_to_read>` period. The end date attribute indicates when the content will no longer be freely available.

  For example, an article that is freely available from February 3, 2013 through October 3, 2013 would be encoded as:

  ```
  <free_to_read start_date="2013-02-03" end_date="2013-10-03"/>
  ```

3.2 license_reference

3.2.1 Description

The `license_reference` element should be included in any metadata for specific journal articles or other scholarly content (book chapters, articles from conference proceedings, etc.) to point to a public

license or waiver—human and/or machine readable—that explains the terms for use or reuse of the content. "Public" means that the license or waiver is universally applicable to users (i.e., is offered generally and not privately). The data within this element must include a stable identifier expressed as a URI that resolves to a resource that is publicly accessible on the web (see examples in Table 1). The supplied URI in this element would be the responsibility of the platform making the content available. While no mechanism is available to require that these URIs be persistent, it is strongly recommended that URIs should be constructed and managed as license identifiers so as to maximize their long-term availability and interpretability.

Table 1: Examples of `license_reference` URIs

Publisher	Possible existing URIs that might be used
American Chemical Society	http://pubs.acs.org/page/policy/authorchoice_termsofuse.html
University of California Press	http://ucpressjournals.com/assets/ucp_sample_auth_agr.pdf
BioMed Central (CC BY v 2.0)	http://creativecommons.org/licenses/by/2.0/
Rockefeller University Press	http://www.rupress.org/site/subscriptions/terms.xhtml and http://creativecommons.org/licenses/by-nc-sa/3.0/
British Institute of Radiology	http://www.bjrpublications.org/page/copyright

3.2.2 `start_date` Attribute

For some publishers, it is important to be able to express how usage rights change over time. To enable the simplest cases, a `start_date` attribute (see 3.1) should be used with the `license_reference` element, as shown in the following examples:

```
<license_ref start_date="2014-02-03">http://www.psychoceramics.org/
license_v1.html</license_ref>

<license_ref start_date="2015-02-03">http://www.psychoceramics.org/
open_license.html</license_ref>
```

The publisher can identify the URI to a license from the date of publication and then indicate that a different license supersedes the proprietary license on an indicated start date. This addresses the issue of embargoes. By way of example, the following coding indicates that the content is under a proprietary license from its date of publication on February 3, 2014, and then is under a CC BY license a year later, beginning on February 3, 2015:

```
<license_ref start_date="2014-02-03">http://www.psychoceramics.org/
license_v1.html</license_ref>
<license_ref start_date="2015-02-03">http://creativecommons.org/licenses/by/
3.0/</license_ref>
```

Neither the `start_date` attribute, nor the license URI itself, provides any technical means of enforcement of these attributes. Any downstream user who is using the content will be responsible for identifying which specific terms apply at the time the content is being used. It will be the downstream

user's responsibility to check for the `start_date` attribute of each `<license_ref>` element to determine which terms currently apply. As noted above, the start date should adhere to ISO 8601 and be in the format YYYY-MM-DD without inclusion of a time or time zone.

Note that there is no corresponding `end_date` attribute for the `<license_ref>` element. This is because including end dates could introduce ambiguities. For example:

- Open licenses, such as Creative Commons (CC), do not have end dates.
- The use of end dates could inadvertently create gaps between applicable licenses.

3.2.3 `applies_to` Attribute

Some publishers and consumers of this license data have a need to further distinguish, qualify, or scope the license or permissioning framework being identified. For this purpose, the `applies_to` attribute may be used to provide additional context for the license or policy being conveyed.

For example, the following coding indicates that two licenses both apply concurrently (covering differing contexts), and are both related to STM's Article Sharing Framework:

```
<license_ref start_date="2014-02-03"
applies_to="https://doi.org/10.15223/asf-profile"
>https://doi.org/10.15223/policy-005</license_ref>

<license_ref start_date="2014-02-03"
applies_to="https://doi.org/10.15223/asf-profile"
>https://doi.org/10.15223/policy-025</license_ref>
```

While a URI value is recommended when applicable, a "name token" string containing only characters (a-z, A-Z), digits (0-9), periods (.), colons (:), and hyphens (-) is also permitted.

3.2.4 Use of `license_reference`

The `license_reference` approach will enable community norms to develop around recognized licenses. This could be done by a single organization or a group of organizations, establishing a "whitelist" of recognized licenses. This gives the flexibility required for tracking compliance with various funder and institutional mandates. For instance, the public access requirements of the Australian Research Council allow for a link to a free-to-read copy to be provided, whereas the Research Councils UK policy requires that, where a publishing fee is paid, the article be both free to read and available under a CC BY license.

The combination of `free_to_read` and `license_reference` metadata provides a mechanism for signaling or determining compliance with most funder and institutional policies that allow compliance through the article publisher's site.

Many publishers use custom licenses, and there is the potential for variant terms to proliferate further. The value generated by these recommendations is greatest when URIs are consistent across a large range of content and publishers. The use of the `license_reference` element should therefore encourage some convergence on license choices without requiring the application of licenses from a specified list. To gain the maximum value, it will also be necessary to maximize the consistency of URIs that refer to the same license.

3.2.5 License URI Expression

This recommended practice attempts to find a balance between allowing flexibility in license choices while providing a consistent and flexible means of transmitting the location of the license. While the

license URI is not required to act as an identifier in the strict sense, the value to downstream users is maximized when the URIs are easily recognizable, stable, and persistent.

To maximize the value of implementing this Recommended Practice, the following procedures are strongly recommended in creating and maintaining license URIs:

1. Where feasible, use a license that is associated with a stable and persistent URI. Where the license is provided by a third party, the form of URI recommended by the provider should be used.
2. When the URI is under control of the publisher, consider the longevity of the domain space and structure to maximize the long-term persistence of the URI. Maintain consistent use of this URI and create new URIs for new versions of the terms.
3. For the specific case of Creative Commons licenses, the URI should be in the form recommended by Creative Commons (or the license provider) for maximum usability. The recommended form (examples in Table 1 and below) includes both the license version, locale (where appropriate), and a trailing slash.
4. The license document itself should contain its own URI identifier. For example, the following text could be used:
   The following license is identified by the URI:
   http://bjr.birjournals.org/site/misc/terms.xhtml

In order to better serve end users and improve the quality and usefulness of the recommended indicators (`<free_to_read>` and `<license_ref>`), it is also recommended that licenses provided to end users be clearly and plainly understandable. To that end, the following best practices are recommended for providing license text to end users:

- If you are using a recognized license, e.g., Creative Commons, use the standard URI for the license. Example: http://creativecommons.org/licenses/by/2.0/
- If you are using a proprietary license, then:
    - Provide a human-readable version of the license, keeping the language simple.
    - Describe what end users can do with the content.
    - Describe if there are any conditions that apply (something the end user must do if s/he uses the content).
    - If possible, make this the only version of the license.
- Use `<license_ref>` to point to a human-readable license first, then link from the human-readable license to the legal license document text or a structured/tagged machine-readable license, if necessary.
- Provide section headers to break up the license text for easier reading/scanning of the license document (human-readable version and legal text).
- Provide definitions for named entities (for example, author, work, user, you).
- Provide the license (human-readable and actual) in well-formed HTML or as accessible PDF (tagged with headers) so that everyone can read it on any device that can access the content in question.

Section 4: Recommended Mechanisms for Distributing Metadata

To ensure the widest dissemination of metadata, publishers, aggregators, and other content providers are encouraged to include the `free_to_read` and `license_reference` elements in all of their standard metadata sets. 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.

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 use of a namespace for the new elements helps clarify how they should be incorporated into existing schemas and formats. The appropriate organizations will need to consider the best way to incorporate the `free_to_read` and `license_reference` metadata into existing formats, such as the Journal Article Tag Suite (JATS), ONIX, RDF, OAI-PMH, and Dublin Core (DC).

Section 5: Use Case Review

The Working Group identified and analyzed a number of potential use cases, both to aid in development of the recommendations and to test whether the recommendations would completely address the case.

5.1 Use Case: End User Seeks to Discover, Identify, and Access Free-to-Read Items

Actor(s): End users, indexing and abstracting services, institutions, publishers, metadata harvesters

Description: When an interested party is searching for an item, it would be helpful to have the item's accessibility status identified in the search results. For items that are openly accessible via a hybrid/"open choice" plan, library electronic resource management (ERM) systems may not, at present, convey this accessibility to the library's discovery system. This is because the availability status often comes from the knowledge base of a link resolver or the library ERM and is managed on a journal/volume/issue level, rather than an article level.

Use Case Addressed? Yes

Publishers should transmit a machine-readable `<free_to_read>` indication in their article metadata. Discovery systems, knowledge bases, and other third-party mechanisms can use those indicators to provide the correct availability status. Search engines can include this material in their harvesting programs.

5.2 Use Case: End User Seeks to Know the Readability Status of an Item

Actor(s): End users, indexing and abstracting services, institutions, publishers, metadata harvesters

Description: Readers, on arrival at a standard web page for a journal article, may not be aware that an article is free to read. One example of this is the scenario in which a reader who does not generally have access to subscription content runs a Google search and pay-for-use scholarly articles are among the returned results. For this user, reaching a paywall can be frustrating and might lead such a user to copies other than the definitive versions of record. In some cases, technical/procedural problems, in combination with confusion, can mean that readers are not aware they *should* be able to read an article when they are blocked for some reason (e.g., the end user is not using an authorized network connection to licensed content to which s/he has access). Clear identification of free-to-read content could help reduce time wastage as readers attempt to reach alternative versions. This identification will need to be understandable to an audience that is inexperienced with scholarly and subscription content.

Use Case Addressed? Yes

Publishers should transmit machine-readable `<free_to_read>` indication in the metadata. Discovery systems, knowledge bases, and other third-party mechanisms should use those indicators to provide the correct availability status. Search engines can include this material in their harvesting program. If a publisher maintains multiple versions of an item on its site (e.g., the Version of Record and the Accepted Manuscript), each item could transmit a distinct machine-readable `<free_to_read>` indication in the metadata.

5.3 Use Case: End User Seeks to Know Reuse Permissions of an Item

Actor(s): End users, publishers, repositories

Description: Researchers are unsure what they can and cannot legally do with content they find online; for example, whether they can distribute articles (or parts of them) to a class, include full text in their reference management tools or academic social networking profiles, or share material in open educational resources (OERs). The reuse of licensed material is limited because it is not easy to identify material that can be reused, especially if the reader wants to share it outside his/her institution. Clear identification of those materials with wide reuse rights may also discourage infringing uses of materials with limited or no reuse rights.

Use Case Addressed? Yes

Publishers should provide a machine-readable `<license_ref>` indicator in the article metadata. Note that interpretation of reuse rights would be the responsibility of the user. Based on the information in the elements, sites/systems should display the status of content in a human readable form.

5.4 Use Case: End User Seeks to Know Reuse Permissions of a Sub-Component of an Item

Actor(s): End users, authors, publishers

Description: A person wants to use a sub-component of an article (e.g., the abstract, an image, or a full poem quoted within a scholarly article, but not a full article), either in a single case or in some automated reuse pipeline. It is not clear if the sub-component has the same reuse rights and restrictions as the full article. The overall article (for instance, the original text contribution of the author) might be licensed CC BY, but the article may contain content (possibly, though not under all circumstances, from a third party) such as photographs, datasets, or figures under a different open license (or even a custom license) that is not clearly marked. This increases the likelihood that reusers will violate those terms, that licensors will refrain from incorporating valuable sub-components, and that third-party providers will object to having their content included in free-to-read articles.

Use Case Addressed? Yes

Where components of a content item (e.g., figures in a journal article or book chapter) are separately identified, the `<license_ref>` URI can be applied to the component. It is not currently common practice to uniquely identify sub-components of content, so for now there will be limited use of `<license_ref>` for sub-components. It is outside the scope of this Recommended Practice to determine how components should be identified, but the group hopes that the `<license_ref>` URI will become part of the standard practice for identification of sub-components as it develops.

5.5 Use Case: Repositories Seek to Expose Free-to-Read Items

Actor(s): Repositories, both institutional and subject-based; publishers

Description: Currently it is difficult and labor intensive for both authors and repository managers to be sure what they can and cannot do with potential deposits, including which articles (and which parts of articles) they can make directly available and which articles they can disseminate to other repositories.

Use Case Addressed? Yes

Publishers should provide a machine-readable `<license_ref>` URI in the article metadata that also references any part of the article licensed under a different license to that governing the article as a

whole. Interpretation of reuse rights would be the responsibility of the user. Repository managers replicating full text via their repositories are encouraged also to provide the same `<license_ref>` tag in the metadata relating to the copy of the item in their repository.

5.6 Use Case: End User Seeks to Text Mine Content

Actor(s): End users, text miners, repositories

Description: Currently it is difficult and labor-intensive for text miners to know what they can access for mining content (which articles and which parts of articles, for example), and what they can do with the information they mine.

Use Case Addressed? Partially

Publishers should provide a machine-readable `<license_ref>` URI in the article metadata that references any part of the article licensed under a different license to that governing the article as a whole. Licenses should indicate whether or not text and data mining is allowed and whether there are additional terms and conditions. Note that interpretation of reuse rights would be the responsibility of the user.

5.7 Use Case: Ensure Author/Publisher Rights Assertions Align with License Statements

Actor(s): Authors, publishers, end users

Description: A user wishes to use visual images from an article, either in a single case or in some automated reuse pipeline. Acting in good faith, the user seeks licensing information, e.g., at PubMed Central or a similar source, to ascertain his/her rights. However, in some cases the article licensing metadata is contradictory or incorrect. For example, an article might be properly licensed under CC BY, but the publisher (or whomever is adding metadata) is making conflicting licensing statements or identifies other restrictions not provided for in the license.[1]

Use Case Addressed? Partially

Publishers should provide a machine-readable `<license_ref>` license URI in the article metadata that references any part of the article licensed under a different license to that governing the article as a whole. There will always be mistakes and inconsistencies,[1] but users should be able to rely on information provided formally via the `<license_ref>` attribute. As this Recommended Practice only applies to public licenses, any other contractual extension (or limitation) of reuse rights by specific parties is not overridden by the `<license_ref>` attribute.

5.8 Use Case: Funding Agency Seeks to Track Compliance of Research Outputs to Open Access Mandates

Actor(s): Funding agencies, institutions

Description: At present it is difficult and inefficient for funding agencies to track the published outputs of funded work and ensure they are published in or through compliant channels.

Use Case Addressed? Partially

[1] For a discussion of this issue, see: "Publishers Deliver Inconsistent XML to PubMed Central." *GLAM*, November 2012, volume II, issue XI. Available at: https://outreach.wikimedia.org/w/index.php?title=GLAM/Newsletter/November_2012/Contents/Open_Access_report&oldid=51485%20-%20Metadata_at_PubMed_Central#Metadata_at_PubMed_Central

Machine-readable `<free_to_read>` and `<license_ref>` metadata could be used by funding agencies to develop various automated reports to track and analyze funded research outputs, but it is beyond the scope of this working group to address the wider issue of tracking research outputs.

5.9 Use Case: Institution Seeks to Report on Open Access Compliance of Research Outputs

Actor(s): Institutions, funding agencies

Description: It is challenging for institutions to report to funders on policy compliance levels and for funding agencies to manage and monitor transition to wider access or report to their sponsors in government.

Use Case Addressed? Partially

Machine-readable `<free_to_read>` and `<license_ref>` metadata could be used by institutions to develop various automated reports to track and analyze funded research outputs, but it is beyond the scope of this working group to address the wider issue of tracking compliance.

5.10 Use Case: Platform Seeks to Know the Sharing Terms of an Article

Actor(s): Online Content-Sharing Service Providers (OCSSP), Scholarly Collaboration Networks (SCNs)

Description: A content-sharing platform seeks to determine whether a given article may be shared within a given context based on the audience, version of the article, portions of the article to be shared, and the type of platform.

Use Case Addressed? Partially

Publishers use ALI to assert machine-identifiable URI expressions for the sharing policies that apply to the article using a predefined permissioning framework, such as STM's Article Sharing Framework. By identifying the sharing policies using the `license_ref` URIs, and identifying the framework using `applies_to` attribute, platforms can determine whether a given article contains policy assertions related to a particular permissioning framework.

Appendix A:
Means of Expressing Metadata Elements

A.1 XML Implementation

To implement the Access & License Indicators (ALI) recommendation in XML, there are two strategies: incorporate the vocabulary directly into an existing schema or use a standalone schema and reference it via namespace.

Example: XML stand-alone schema with referenced namespace

```
<?xml version="1.0" encoding="UTF-8"?>
<xsd:schema xmlns:xsd="http://www.w3.org/2001/XMLSchema"
targetNamespace="http://www.niso.org/schemas/ali/1.0/">
    <xsd:element name="free_to_read">
    <xsd:complexType>
        <xsd:attribute name="end_date" use="optional" type="xsd:date"/>
        <xsd:attribute name="start_date" use="optional" type="xsd:date"/>
    </xsd:complexType>
    </xsd:element>
    <xsd:element name="license_ref">
    <xsd:complexType>
        <xsd:simpleContent>
            <xsd:extension base="xsd:anyURI">
                <xsd:attribute name="start_date" use="optional" type="xsd:date"/>
                <xsd:attribute name="applies_to" use="optional">
                        <xsd:simpleType>
                            <xsd:union memberTypes="xsd:anyURI xsd:NMTOKEN"/>
                        </xsd:simpleType>
                </xsd:attribute>
            </xsd:extension>
        </xsd:simpleContent>
    </xsd:complexType>
    </xsd:element>
</xsd:schema>
```

Example #1 of XML using above schema:

```
<document>
...
<free_to_read xmlns="http://www.niso.org/schemas/ali/1.0/" start_date="2014-01-01"
end_date="2014-12-31"/>
```

```
<license_ref xmlns="http://www.niso.org/schemas/ali/1.0/" start_date="2015-02-03">
    http://www.examplesite.org/open_license.html
</license_ref>
...
</document>
```

Example #2 of XML using above schema with `applies_to`:

```
<document>
...
<free_to_read
   xmlns="http://www.niso.org/schemas/ali/1.0/"
   start_date="2014-01-01" end_date="2014-12-31"/>
<license_ref
   xmlns=http://www.niso.org/schemas/ali/1.0/
   applies_to="http://www.examplesite.org/licenses"
>http://www.examplesite.org/open_license.html</license_ref>
...
</document>
```

A.2 PDF Metadata Implementation – XMP

The Extensible Metadata Platform (XMP) standard (ISO 16684) can be used to embed metadata within a content object, such as a PDF document. The example below shows how the `<free_to_read>` and `<license_ref>` can be incorporated.

```
<?xml version="1.0" encoding="UTF-8"?>
<rdf:RDF
    xmlns:ali="http://www.niso.org/schemas/ali/1.0/"
    xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
    xmlns:dc="http://purl.org/dc/elements/1.1/"
    <!-- more name spaces -->
    xmlns:xsd="http://www.w3.org/2001/XMLSchema#">
        <rdf:Description rdf:about="http://dx.doi.org/10.1234567890">
            <!-- Various metadata elements can go here... -->
            <dc:title>A Breakthrough Paper</dc:title>


                <ali:free_to_read>
                    <rdf:Bag>
                        <rdf:li rdf:parseType="Resource">
                            <!-- free to read period around X-Mas -->
```

```
                            <ali:start_date>2014-12-24</ali:start_date>
                            <ali:end_date>2014-12-31</ali:end_date>
                        </rdf:li>
                        <rdf:li rdf:parseType="Resource">
                            <!-- free to read period for Valentine’s Day -->
                            <ali:start_date>2015-02-14</ali:start_date>
                            <ali:end_date>2015-02-14</ali:end_date>
                        </rdf:li>
                    </rdf:Bag>
                </ali:free_to_read>


                <ali:license_ref>
                    <rdf:Bag>
                        <rdf:li rdf:parseType="Resource">
                            <!-- applying a license from 1 Nov 2014 onwards -->
                            <ali:start_date>2014-11-01</ali:start_date>
                            <ali:uri>http://www.psychoceramics.org/
license_v1.html</ali:uri>
                        </rdf:li>
                        <rdf:li rdf:parseType="Resource">
                            <!-- making the license more liberal one year later -->
                            <ali:start_date>2015-11-01</ali:start_date>
                            <!-- naming the source licensing or permissioning
framework -->
                            <ali:applies_to>http://www.psychoceramics.org/
licenses</ali:applies_to>
                            <ali:uri>http://www.psychoceramics.org/
open_license.html</ali:uri>
                        </rdf:li>
                    </rdf:Bag>
                </ali:license_ref>
        </rdf:Description>
</rdf:RDF>
```

If an article is fully `free_to_read` without any date restrictions, the following assertion must be inserted instead of the date-range version:

```
                <ali:free_to_read>true</ali:free_to_read>
```

Example of `<free_to_read>` without date restrictions:

```
                <ali:free_to_read>true</ali:free_to_read>
                <ali:license_ref>
                    <rdf:Bag>
                        <rdf:li rdf:parseType="Resource">
                            <ali:start_date>2014-01-01</ali:start_date>
                            <ali:uri>http://www.psychoceramics.org/
open_license.html</ali:uri>
                        </rdf:li>
                    </rdf:Bag>
                </ali:license_ref>
```

Note that fully `free_to_read`, and `free_to_read` with date ranges are mutually exclusive; you can have one or the other, but not both.

A.3 RDF

The following shows an example of the `<free_to_read>` and `<license_ref>` metadata in a Resource Description Framework (RDF) expression.

```
<?xml version="1.0" encoding="UTF-8"?>
<rdf:RDF
    xmlns:ali="http://www.niso.org/schemas/ali/1.0/"
    xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
    xmlns:dc="http://purl.org/dc/elements/1.1/"
    <!-- more name spaces -->
    xmlns:xsd="http://www.w3.org/2001/XMLSchema#">
        <rdf:Description rdf:about="http://dx.doi.org/10.1234567890">
            <!-- Various metadata elements can go here... -->
            <dc:title>A Breakthrough Paper</dc:title>


                <ali:free_to_read rdf:parseType="Resource">
                    <!-- free to read period around X-Mas -->
                    <ali:start_date rdf:datatype="xsd:date">2014-12-
24</ali:start_date>
                    <ali:end_date rdf:datatype="xsd:date">2014-12-31</ali:end_date>
                </ali:free_to_read>
                <ali:free_to_read rdf:parseType="Resource">
                    <!-- free to read period for Valentine's Day -->
                    <ali:start_date rdf:datatype="xsd:date">2015-02-
14</ali:start_date>
```

```
                    <ali:end_date rdf:datatype="xsd:date">2015-02-14</ali:end_date>
                </ali:free_to_read>



                <ali:license_ref rdf:parseType="Resource">
                    <!-- applying a license from 1 Nov 2014 onwards -->
                    <ali:start_date rdf:datatype="xsd:date">2014-11-
01</ali:start_date>
                    <ali:uri rdf:resource="http://www.psychoceramics.org/
license_v1.html"/>
                </ali:license_ref>
                <ali:license_ref rdf:parseType="Resource">
                    <!-- making the license more liberal one year later -->
                    <ali:start_date rdf:datatype="xsd:date">2015-11-
01</ali:start_date>
                    <!-- naming the source licensing or permissioning framework -->
                    <ali:applies_to rdf:datatype="xsd:anyURI"
                          >http://www.psychoceramics.org/licenses</ali:applies_to>
                    <ali:uri rdf:resource="http://www.psychoceramics.org/
open_license.html"/>
                </ali:license_ref>


        </rdf:Description>
</rdf:RDF>
```

If an article is fully `free_to_read` without any date restrictions, the following assertion must be inserted instead of the date-range version:

```
                <ali:free_to_read rdf:datatype="http://www.w3.org/2001/
XMLSchema#boolean">true</ali:free_to_read>
```

Example of `<free_to_read>` without date restrictions:

```
                <ali:free_to_read rdf:datatype="http://www.w3.org/2001/
XMLSchema#boolean">true</ali:free_to_read>
                <ali:license_ref rdf:parseType="Resource">
                    <ali:start_date rdf:datatype="xsd:date">2014-01-
01</ali:start_date>
                    <ali:uri rdf:resource="http://www.psychoceramics.org/
open_license.html"/>
                </ali:license_ref>
```

Note that fully `free_to_read` and `free_to_read` with date ranges are mutually exclusive; you can have one or the other, but not both.

A.4 JSON Implementation

Example of JSON implementation that specifies only that the resource is free to read:

```
{
    "@context": "http://www.niso.org/schemas/ali/1.0.1/jsonld.json",
    "free_to_read": true
}
```

Example of JSON implementation with start and end date for `free_to_read:`

Note that if the `free_to_read` key is present, it must take the value of "`true`" or a JSON object that specifies the date(s). It would be invalid for the value to be "`false`" or of any other data type.

```
{
    "@context": "http://www.niso.org/schemas/ali/1.0.1/jsonld.json",
    "free_to_read": {
        "start_date": "2013-02-03",
        "end_date": "2013-10-03"
  }
}
```

Example of JSON implementation that specifies `free_to_read`, as well as giving a license URI:

```
{
    "@context": "http://www.niso.org/schemas/ali/1.0.1/jsonld.json",
    "free_to_read": true,
    "license_ref": "http://creativecommons.org/licenses/by/3.0/"
}
```

Example of JSON implementation that specifies a start date for the license:

```
{
    "@context": "http://www.niso.org/schemas/ali/1.0.1/jsonld.json",
    "free_to_read": true,
    "license_ref": {
        "start_date": "2014-04-04",
        "uri": "http://creativecommons.org/licenses/by/3.0/"
  }
}
```

Example of JSON implementation of two licenses, each with different start dates:

```
{
    "@context": "http://www.niso.org/schemas/ali/1.0.1/jsonld.json",
    "free_to_read": true,
    "license_ref": [
        { "start_date": "2014-04-04",
        "uri": "http://creativecommons.org/licenses/by/3.0/" },
        { "start_date": "2015-04-04",
        "uri": "http://creativecommons.org/licenses/by/4.0/" }
  ]
}
```

Example of JSON implementation of two licenses, both from same permissioning framework identified by "`https://doi.org/10.15223/asf-profile`":

```
{
    "@context": "http://www.niso.org/schemas/ali/1.0.1/jsonld.json",
    "free_to_read": true,
    "license_ref": [
        { "start_date": "2015-04-04",
        "applies_to": "https://doi.org/10.15223/asf-profile",
        "uri": "https://doi.org/10.15223/policy-005" },
        { "start_date": "2015-04-04",
        "applies_to": "https://doi.org/10.15223/asf-profile",
        "uri": "https://doi.org/10.15223/policy-025" },
  ]
}
```

JSON Context

The context used above should point to a JSON-LD document that defines the names and the datatypes. (A @context in JSON-LD must point to a document that defines the names and datatypes, whereas a @vocab defines the namespace.) In the above examples, the following context is used. Note how it defines the same namespace prefix, `http://www.niso.org/schemas/ali/1.0/`, as is used by the XML examples. This ensures that the fully qualified name, which is comprised of the namespace and tag name, is the same whether it is encoded in XML or in JSON-LD. An example is `http://www.niso.org/schemas/ali/1.0/free_to_read`.

```
{
    "@context": {
    "@vocab": "http://www.niso.org/schemas/ali/1.0/",
    "xsd": "http://www.w3.org/2001/XMLSchema#",
    "free_to_read": { "@type": "xsd:boolean" },
```

```
    "license_ref": { "@type": "@id" },
    "uri": { "@type": "@id" },
        "applies_to": { "@type": "xsd:anyURI" },
        "start_date": { "@type": "xsd:date" },
        "end_date": { "@type": "xsd:date" }
  }
}
```

Bibliography

ARC Open Access Policy. Australian Research Council, last modified January 8, 2013. Available at: http://www.arc.gov.au/applicants/open_access.htm

Budapest Open Access Initiative [website]. http://www.budapestopenaccessinitiative.org/

Creative Commons [website]. http://creativecommons.org/

Creative Commons. About CC0 — "No Rights Reserved". Available at: http://creativecommons.org/about/cc0

Creative Commons. Attribution 4.0 International (CC BY 4.0). Available at: http://creativecommons.org/licenses/by/4.0/

Crossref [website]. http://www.crossref.org/

Data elements and interchange formats – Information interchange – Representation of dates and times. ISO 8601:2004. Geneva: International Organization for Standardization, 2004. Available at: http://www.iso.org/iso/home/store/catalogue_tc/catalogue_detail.htm?csnumber=40874

Date and Time on the Internet: Timestamps. RFC 3339. Internet Engineering Task Force, July 2002. Available at: http://www.rfc-editor.org/rfc/rfc3339.txt

DOI (Digital Object Identifier) Registration Agencies [webpage]. http://www.doi.org/registration_agencies.html

The Dublin Core Metadata Element Set. ANSI/NISO Z39.85-2012. Baltimore, MD: National Information Standards Organization, approved February 20, 2013. Available at: http://www.niso.org/standards/z39-85-2012

Extensible Markup Language (XML) [website]. World Wide Web Consortium. http://www.w3.org/XML/

Graphic technology – Extensible metadata platform (XMP) specification – Part 1: Data model, serialization and core properties. ISO 16684-1:2012. Geneva: International Organization for Standardization, February 9, 2012. Available at: http://www.iso.org/iso/home/store/catalogue_tc/catalogue_detail.htm?csnumber=57421

JATS: Journal Article Tag Suite. ANSI/NISO Z39.96-2012. Baltimore, MD: National Information Standards Organization, August 9, 2012. Available at: http://jats.niso.org/

The JavaScript Object Notation (JSON) Data Interchange Format. RFC 7159. Internet Engineering Task Force, March 2014. Available at: http://www.rfc-editor.org/rfc/rfc7159.txt

"Meta data." Section 7.4.4 in: *HTML 4.01 Specification*. W3C Recommendation, December 24, 1999. Available at: http://www.w3.org/TR/1999/REC-html401-19991224/struct/global.html#h-7.4.4

ONIX Family of Standards [webpages]. EDItEUR. http://www.editeur.org/8/ONIX/

Open Access Scholarly Publishers Association (OASPA) [website]. http://oaspa.org/

The Open Archives Initiative Protocol for Metadata Harvesting (OAI-PMH). Version 2.0. Open Archives Initiative, June 14, 2002; document last updated December 7, 2008. Available at: http://www.openarchives.org/OAI/2.0/openarchivesprotocol.htm

RCUK Policy on Open Access and Supporting Guidance. Research Councils UK, [2005]. http://www.rcuk.ac.uk/documents/documents/RCUKOpenAccessPolicy.pdf

Resource Description Framework (RDF) [website]. World Wide Web Consortium. http://www.w3.org/RDF/

STM Article Sharing Framework [website]. https://doi.org/10.15223/asf-profile

Uniform Resource Identifier (URI): Generic Syntax. RFC 3986. Internet Engineering Task Force, January 2005. Available at: http://www.ietf.org/rfc/rfc3986.txt