

# We Should Not Light an Open Access Lamp and then Hide it Under a Bushel!

Santiago CHUMBE<sup>a,1</sup>, Roddy MACLEOD<sup>b</sup> and Brian KELLY<sup>c</sup>

<sup>a</sup>ICBL (Institute for Computer Based Learning), School of Mathematical and Computer Sciences, Heriot Watt University, Edinburgh, UK

<sup>b</sup>Former subject librarian at Heriot Watt University, Edinburgh, UK

<sup>c</sup>Independent consultant at UK Web Focus

**Abstract.** The rapid growth of hybrid journals in the last few years has seen an unfortunate side effect: the majority of Open Access (OA) articles published in those journals cannot be recognized as OA beyond the publishers' websites, or by the discovery services used by researchers to access full-text articles. This reality has been demonstrated in the literature and solutions have been proposed. This paper explains the causes behind the problem, examines each of the proposed solutions, discusses the few implementations made with those solutions, and estimates whether the potential benefits merit the efforts required to implement the available solutions. Each of the solutions is analyzed from standardization and pragmatic perspectives. In particular, we critically analyze the solution proposed by NISO (RP-22-2015), and compare it with the solution offered by the JEMO project, which is based on using metadata elements from namespaces and XML schemas already being used by publishers. The contribution presents a number of case studies which show that research published as OA ends up erroneously being labelled as non-OA on the electronic services used by the end-user, when one of the components of the supply and delivery chain for e-journals fails to include OA information in its metadata. Furthermore, the case studies demonstrate that publishers of hybrid journals should not be the only ones being answerable for the problem. In fact, during the study, some publishers were actually not allowed to enable OA identification, at the article level, by key components of the supply chain. In those case studies, we worked with a sample of publishers that implemented the JEMO solution. From those experiences we draw answers to the main question of this presentation: which solution should be used to enable OA discovery from hybrid journals? What becomes apparent is that publishers are prepared and willing to implement any of the available solutions in their publishing workflow. The paper proposes that the simplest option is the best solution to provide standardized means to identify OA at the article level.

**Keywords.** Hybrid journals, Open Access articles, e-Publishing platforms, interoperability and integration, web feeds, metadata standards, e-Journal supply chain, discovery services, RSS.

## 1. Introduction

When researchers see that an article is published in a subscription journal for which they do not have full-text access, there is a high chance that they will give up instead of

---

<sup>1</sup> Corresponding Author. E-mail: S.Chumbe@hw.ac.uk.

trying to obtain the full-text (the probability is 70% according to S. A. Knowlton et al, [1].) What if the article in question is an OA article which happens to be published in a subscription-based journal? Are OA articles in hybrid journals read and cited less frequently because end-users are not properly made aware of the OA status of such articles? Unfortunately, this is happening and is a real problem. Figuratively, those OA articles are, in fact, being kept under a bushel. Authors wanting to publish OA in a hybrid journal are being penalized with higher APCs (Article Processing Charges) at the same time that the wider community is not being made aware of the availability of those OA articles. Those articles are not labeled as OA either beyond the publishers' websites, or by discovery services used by researchers to access full-text articles. This is a problem that has been noticed and identified by other studies [2, 3, 4, 5, 6, 7]. Accordingly the community has reacted by proposing solutions [8, 9, 10, 11]. However the problem is still unresolved. Our purpose is to describe the problem and to explain why it is still unresolved. Despite flaws in the hybrid model [12] the importance of hybrid journals cannot be disputed as most publishers are producing them. They attract authors who are interested in publishing OA articles in high ranking and well established journals [13]. Some may never become Gold journals [14]. If open access is intended to improve access to and dissemination of knowledge, it is crucial that any type of OA research output is identified as OA to everyone, everywhere, at all times.

The paper is organized as follows. In section 2, we use case studies to introduce the problem and describe its causes. Section 3 presents an analysis of available solutions and gives reasons why we endorse the solution proposed by JEMO. In section 4, through experiments carried out with a sample of hybrid journals, we demonstrate the benefits produced by a simple programmatic OA identification. The final section provides conclusions and recommendations.

## 2. Why Does the Problem Remain Unresolved?

The JEMO Project received funding from the Engineering and Physical Sciences Research Council (EPSRC) to find out why the problem introduced above remains unresolved. While some business aspects of the problem were beyond our control, the metadata used across the production, discovery and delivery chain of e-journals emerged as an important factor. We concluded that metadata is at the root of the problem of OA articles from hybrid journals being wrongly identified as non-OA articles. Metadata is important; it can enhance the results produced by retrieval and discovery systems and increase the usefulness and value of delivery systems such as link resolvers [15]. It can also enable the development of new services. But what matters is metadata quality [16]. Metadata has to be FAIR (Findable, Accessible, Interoperable and Re-usable)<sup>2</sup> otherwise it can even be harmful or misleading [17]. We will show that the problem is manifested in the diversity and poor quality of the metadata used in the e-journal supply chain, and in the amount of redundant and sometimes conflicting metadata specifications.

Information from 14 publishers is used in the case studies. Five of those publishers were official project partners and the other nine were invited to implement the project recommendations. The following table consolidates the number of Gold OA, Hybrid and Subscription-only journals currently being published by those publishers. Further

---

<sup>2</sup> <https://www.force11.org/node/6062/> (visited on 30 Apr 2015)

detailed analysis removes their individual identifications as some publishers preferred to remain anonymous. Specific identification of the commercial hosting platforms has also been removed from the discussion<sup>3</sup>.

**Table 1.** JEMO Participating Publishers. Showing hosting type (whether they have outsourced the hosting of their content to external e-publishing platforms or not) and type of journals (Hybrid: includes both OA and non-OA articles; Gold: includes OA articles only; Subscript: includes subscription-based articles only.)

| Publisher                 | Hosting Type | Journals | Hybrid | Gold | Subscript. |
|---------------------------|--------------|----------|--------|------|------------|
| IUCr                      | internal     | 9        | 7      | 2    | 0          |
| BioMed Central            | internal     | 278      | 0      | 273  | 0          |
| Cambridge U Press         | internal     | 447      | 151    | 5    | 291        |
| Edinburgh U Press         | outsourced   | 39       | 39     | 0    | 0          |
| IGI-Global                | internal     | 146      | 13     | 0    | 133        |
| Inderscience Publishers   | internal     | 397      | 397    | 0    | 0          |
| Libertas Academica        | internal     | 86       | 0      | 86   | 0          |
| Maney Publishing          | outsourced   | 201      | 200    | 1    | 0          |
| MDPI                      | internal     | 136      | 0      | 136  | 0          |
| Oxford U Press            | outsourced   | 342      | 301    | 28   | 13         |
| Walter de Gruyter         | outsourced   | 678      | 328    | 350  | 0          |
| Taylor & Francis          | outsourced   | 1920     | 1810   | 38   | 72         |
| The Geological Society    | outsourced   | 10       | 10     | 0    | 0          |
| Thieme Medical Publishers | outsourced   | 158      | 140    | 18   | 0          |
| Total                     |              | 4847     | 3396   | 932  | 509        |

Publishers of hybrid journals know that metadata is important. Yet, some fail to appreciate that what is critically important is that it has to be *fit-for-purpose* metadata. Being fit-for-purpose means providing reusable (interoperable), consistent, accurate and complete information about the article associated with the metadata [18]. The 14 publishers understood the benefits and importance of producing quality metadata. However, in the implementation stages differences started to emerge. Five of the publishers that were using in-house hosting were able to incorporate OA elements in their metadata. The other two publishers chose to wait for the NISO RP-22-2015 recommendations to be released. The situation with publishers using external hosting platforms was contrasting. They faced an additional situation whereby their outsourced content on e-publishing hosting platforms is outwith the control of the original publishers. Despite their willingness to do so, publishers who have outsourced the hosting of their content to external platforms could not implement the required changes as quickly as done or scheduled by the other publishers. Being the bridge between publishers and the rest of the components of the e-journal supply chain, e-publishing hosting platforms play an important role in the transport of OA metadata. A further analysis of those platforms showed that to be cost-effective they cannot implement on-demand software changes on an individual publisher basis. The changes have to be made globally; usually as part of one or two annual software updates. Also, as in most cases the publisher's production system is not integrated with the external platform, the metadata used by the publisher to feed their platforms with new content, needs to be updated to incorporate new OA elements. The hosting platforms were not aware of the importance of those OA elements for hybrid journals. With the support of three publishers, it required discussions with one of the largest platforms to agree to change their metadata to accept OA elements. The metadata is based in the JATS tag suite<sup>4</sup>, which is a schema increasingly being adopted by e-Publishing platforms to ingest

<sup>3</sup> Those platforms are Atypon, PublishingTechnology and PubFactory

<sup>4</sup> NISO Z39.96-2012 Journal Article Tag Suite (<http://jats.nlm.nih.gov/>, visited on 30 Apr 2015)

content from publishers. JATS already has elements to identify OA at the article level<sup>5</sup>. Figure 1 shows an example of using those elements assuming the copyright is retained by the publisher. Figure 2 represents the values for non-OA cases.

#### For OA articles:

```
<permissions>
<copyright-statement>Copyright © Publication_Year Publisher_Name
</copyright-statement>
<copyright-year>Publication_Year</copyright-year>
<copyright-holder>Publisher_Name</copyright-holder>
<license license-type="open-access"
xlink:href="http://creativecommons.org/licenses/by-nc-nd/4.0" />
<license-p>This is an open-access article distributed under the terms of the
Creative Commons Attribution License, which permits NonCommercial use,
distribution, and reproduction in any medium, provided the original work is
properly cited and you do not distribute the modified material.</license-p>
</license>
</permissions>
```

**Figure 1.** OA elements included in a JATS file to enable OA identification at the article level.

#### For non-OA articles:

```
<permissions>
<copyright-statement>Copyright © Publication_Year Publisher_Name
</copyright-statement>
<copyright-year>Publication_Year</copyright-year>
<copyright-holder>Publisher_Name</copyright-holder>
</permissions>
```

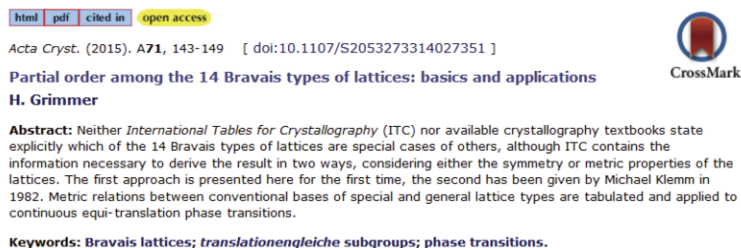
**Figure 2.** Optional copyright elements included in a JATS file for non-OA articles.

Six months after the five publishers were producing metadata with OA elements; we tested the discovery services that researchers are likely to use when trying to access full-text. Were those services taking advantage of the changes made by the publishers to provide OA identification, at the article level? The answer was no.

Discovery services are at the end of the supply chain and are supposed to be the main full-text access points for end-users. The problem with those services is that they can only identify OA at the journal level. Those services have implemented their own solutions. SerialSolutions and ExLibris for example, are addressing the problem using OA packages from OA aggregators that in theory would allow users to discover OA content published in any journal. However, the following example shows that this approach is not working.

*Acta Crystallographica Section A* is an hybrid journal published by the International Union of Crystallography (IUCr), which was one of the five publishers that immediately implemented OA identification at the article level in their metadata (in March 2014.) The journal published one OA article in its Volume 71, Issue 2 (2015). Figure 3 shows how the article is identified as OA on the journal's website.

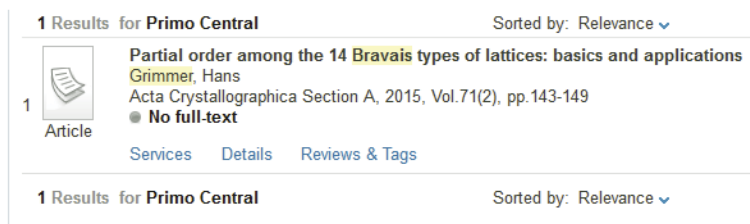
<sup>5</sup> JATS will add support for the new NISO RP-22-2015 elements too (<http://jats.nlm.nih.gov/1.1d3/>)



**Figure 3.** OA article identified on the journal's website. Source: <http://scripts.iucr.org/cgi-bin/paper?eo5044> (visited on 1 May 2015)

The Wiley Online Library database also hosts articles of the journal and has no problem identifying the OA article<sup>6</sup> or any OA article published in this hybrid journal. The same could happen with any aggregator or discovery service that supports OA identification at the article level, for example on JournalTOCs<sup>7</sup>.

We tried to access this same OA article from Ex-Libris Primo Central, without success. Figure 4 shows that Primo is erroneously labelling this OA article as non-OA (“No full-text.”) If you select the “Services” link to gain full-text, you will be suggested to use the Inter Loan Library (ILL) service to read this OA article<sup>8</sup>. The first screenshot in Figure 5 is from EbscoHost, where the user is advised to request ILL to be able to read the OA article. The last screenshot comes from Summon, which includes the “Full Text Online” link; giving the appearance that through this link you could get full-text access. However the link will send you to the SerialSolutions OpenURL landing page<sup>9</sup>, which will point you to the OA Digital Library<sup>10</sup> aggregator from where you will need to start your search again, only to find out at the end that this aggregator doesn't include articles from *Acta Crystallographica Section A*.



**Figure 4.** Primo hiding the OA article behind subscription walls. Seen at <http://goo.gl/OBvYNN>, May 2015

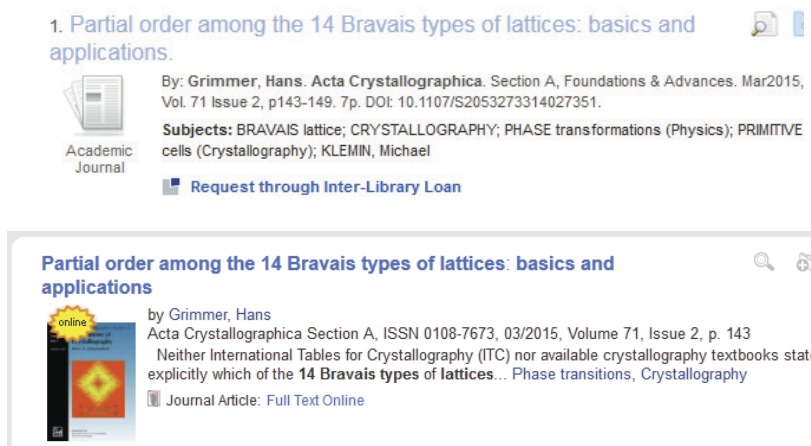
<sup>6</sup> As seen at <http://onlinelibrary.wiley.com/doi/10.1111/aya2.2015.71.issue-2/issuetoc> on 01 May 2015

<sup>7</sup> As seen at <http://www.journaltoCs.ac.uk/?issn=2053-2733> on 15 Feb 2015

<sup>8</sup> As seen at <http://goo.gl/ajitF8> on 1 May 2015

<sup>9</sup> As seen at <http://goo.gl/5xtSZ6> on 1 May 2015

<sup>10</sup> As seen at <http://grweb.coalition.org/oal/oal.html> on 1 May 2015



**Figure 5.** EBSCOhost and Summon discovery services showing the OA article erroneously hidden behind subscription walls. Sources: <http://goo.gl/bjhscZ> and <http://goo.gl/CUWsu0> (visited on 1 May 2015)

The scenario illustrated by the previous case study is not acceptable. The fact that OA identification is still done at the journal level across the supply chain needs to change. OA articles will continue to be erroneously labelled as non-OA on the electronic services used by the end-user if at least one of the components of the supply chain fails to embed the publication's OA status in the metadata shared across this chain. OA identification at the article level requires cooperation between all parties involved and the use of common and standard metadata elements. This lack of cooperation and interest is one of the underlying causes of why this problem is still unresolved. Furthermore, if the solution passes for embedding OA information in the metadata, it cannot be a responsibility of the publishers or publishing platforms only. We believe that, as long as discovery services don't use metadata with OA elements at the article level, any effort made by the publishers will fail.

As mentioned before, two of the publishers decided to wait until a standard solution was agreed across the publishing industry; specifically the one that NISO was preparing at that time. They didn't want to implement something that may not be interoperable with the other components of the e-journal supply chain. This "insecurity" hints at the second cause of the problem. While existing standards could solve the problem, they were not considered and formulated as a consensual solution to identify OA articles. Consequently, publishers are waiting, implementing their own solutions, or just ignoring the problem. The large number of standards, schemas and namespaces to produce metadata for research resources adds more uncertainty<sup>11</sup>. Well-intentioned machine-readable solutions instigated by publishers, such as the *Open Access Collection* of the Geological Society<sup>12</sup> and the *Get New Open Access Article Feed* of Elsevier<sup>13</sup> as well as using HTML meta-tags [10], shows the publishers' willingness to enable programmatic identification of OA, but they are still far from being efficient solutions for aggregators, databases and discovery services as these services would need to shoulder a greater demand to perform the normalizations and transformations required when dealing with diverse types of feeds and metadata elements.

<sup>11</sup> E.g. <http://goo.gl/kD3kSS> or <http://goo.gl/629OqO> (visited on 2 May 2015)

<sup>12</sup> <http://cct.highwire.org/svc/getfile?fileId=283&publisherId=gs1> (visited on 22 April 2015)

<sup>13</sup> For example <http://www.sciencedirect.com/science/journal/15708705> (visited on 22 April 2015)

### 3. Analysis of Available Metadata Standard Solutions

Regarding standard solutions proposed by the community to resolve this imbalanced situation, M. Van Ballegooie [19] identified two options that are currently available, the NISO RP-22-2015 recommendation [20] and the elements proposed by JEMO [21].

In December 2012, NISO formed a group to recommend a specification for the accessibility of journal articles. In January 2015, the group released the NISO RP-22-2015 recommendation. Initial reactions endorsed an eagerly expected specification, but concerns were also expressed by the community.<sup>14 15 16 17</sup> It was noted that the recommendation does not aim to specifically solve the problem of OA articles published in hybrid journals<sup>18</sup>. Aiming to cover all scenarios, the recommendation prefers to use the “Free to Read” term instead of Open Access. From a pragmatic perspective, NISO created the *free-to-read* and *license\_ref* elements and a new namespace<sup>19</sup> specifically designed to support these new elements.

On the other hand, JEMO draws on the wealth of experience provided by the simple yet effective CC (Creative Commons) and DC (Dublin Core) metadata schemas<sup>20</sup> that have been adopted by publishers and are widely used in the e-journal supply chain. The implementation, presented at the NASIG 2014 Conference, aims to resolve the machine-readable or programmatic identification of OA at the article level. It proposes using the *dc:rights* and *cc:license* elements to embed OA information in the metadata already being used by publishers.

The elements proposed by NISO and JEMO are described and assessed from the implementation perspective in Tables 2, 3 and 4.

**Table 2.** Metadata elements proposed to identify OA at the article level. *free-to-read* and *license\_ref* are new elements created by NISO RP-22-2015. *dc:rights* and *cc:license* are elements of the DC and CC metadata standards, respectively. Applicable or needed attributes are noticed.

| Element             | Purpose                                                                                                                                        | Attributes                           | Namespace                                                               |
|---------------------|------------------------------------------------------------------------------------------------------------------------------------------------|--------------------------------------|-------------------------------------------------------------------------|
| <i>dc:rights</i>    | To inform about the ownership of, or rights held in and over, an article                                                                       | None                                 | DC, implemented since 2000 <sup>22</sup>                                |
| <i>cc:license</i>   | To provide a reference to a URI that defines the associated license, indicating the restrictions and how the article may be used and accessed. | <i>rdf:resource</i>                  | CC, de facto license for OA publications <sup>22</sup>                  |
| <i>free-to-read</i> | To define whether the article is accessible, without charge or other restriction to read online.                                               | <i>start_date</i><br><i>end_date</i> | New Access and License Indicators (ALI) <sup>21</sup> to be implemented |
| <i>license_ref</i>  | To provide a reference to a URI that carries the license terms specifying how the article may be used.                                         | <i>start_date</i>                    | ALI                                                                     |

<sup>14</sup> [http://rioxn.net/guidelines/RIOXX\\_Metadata\\_Guidelines\\_v\\_3.0.pdf](http://rioxn.net/guidelines/RIOXX_Metadata_Guidelines_v_3.0.pdf), pp 6-8 (visited on 2 May 2015)

<sup>15</sup> <http://goo.gl/ZQkjbL> (visited on 2 May 2015)

<sup>16</sup> In some way NISO RP-22-2015 blurs the term Open Access by stating that publishers use the terms Open Access, Increased Access, Public Access and other names to identify their offerings; which is not the case as no publisher or entity of the e-journal supply chain uses “Increased Access” or “Public Access” to name Open Access articles or to identify any type of journal.

<sup>17</sup> <http://goo.gl/FkhkV3> (visited on 4 May 2015)

<sup>18</sup> The NISO document tangentially mentions “Open Access” a few times only. In some way, it explains its stand by stating that “this is a contentious area where political views on modes of access lead to differing interpretations of what constitutes ‘open access.’” [20]

<sup>19</sup> <http://www.niso.org/schemas/ali/1.0/> (visited on 2 May 2015)

<sup>20</sup> <http://creativecommons.org/ns> and <http://purl.org/dc/elements/1.1/> (visited on 2 May 2015)

**Table 3.** Remarks for the *free-to-read*, *license\_ref*, *dc:rights* and *cc:license* metadata elements that should be considered by implementers.

| Remark                                                                                 | dc:rights | cc:license | free-to-read | license_ref |
|----------------------------------------------------------------------------------------|-----------|------------|--------------|-------------|
| It is part of a mature and widely adopted standard                                     | YES       | YES        | NO           | NO          |
| Can provide information on whether a specific article is Open Access (OA)              | NO        | YES        | Partially    | Partially   |
| Can provide information on the restrictions and re-use rights of a specific OA article | NO        | YES        | NO           | YES         |
| It is already being used in the e-journal supply chain.                                | YES       | Partially  | NO           | NO          |
| Can indicate the period of time when access to an article is delayed                   | NO        | NO         | YES          | NO          |
| Can indicate how the license's terms change over time                                  | NO        | NO         | NO           | YES         |
| It has already been used to provide copyright metadata                                 | YES       | NO         | NO           | NO          |

**Table 4.** Issues particularly relevant for OA articles published in hybrid journals.

| Question                                                                            | Answer |
|-------------------------------------------------------------------------------------|--------|
| Do OA articles published in hybrid journals have embargo dates?                     | NO     |
| Do OA articles published in hybrid journals have "moving wall" dates?               | NO     |
| Do CC licenses have end or expire dates?                                            | NO     |
| Could the use of end dates inadvertently create gaps between applicable licenses?   | YES    |
| Are any of the hybrid journals exclusively using licenses different to CC licenses? | NO     |
| Have publishers been in the past quick in implementing new metadata specifications? | NO     |
| Can an OA article already published as OA using a CC license, become non-OA?        | NO     |

It is noticeable that, from the OA perspective, the function of the *license\_ref* element can be provided through use of the *cc:license* element. *cc:license* can state the OA status of an article, plus its associated re-use rights. In the NISO case, *free-to-read* can only tell us whether an article can be freely read or not, but this can be an OA article, a free sample, a temporary promotion, etc. *free-to-read* alone is not enough to know the re-use rights of OA articles. A second new element (*license-ref*) is needed to complete the article's OA status. Consequently, *cc:license* resolves the specific OA problem caused by hybrid journals, while *license-ref* is a general-purpose solution, which needs to be combined/analyzed with *free-to-read* to indicate OA as a function: OA =  $f(\text{free-to-read, re-use rights, embargo-period})$ .

The new NISO elements provide an embargo period; a concept relevant to subscription-based journals but alien to OA. OA means full-text access without any delay, forever. The *start\_date* and *end\_date* attributes of those two new elements do not apply in OA; if used, they would need to semantically be analyzed by the services trying to identify OA articles and ignored for OA articles, a process that would introduce additional complexity to the handling of terms used in Open Access. In contrast, when the value of *cc:license* is a valid CC URI, the risk of identified a non-OA article as OA is null. CC licenses are not revocable<sup>21</sup>. Therefore, an OA article licensed with any CC license is perpetually OA. Furthermore, an OA article published under any CC license is immediately OA upon publication. The unanimous praxis among OA stakeholders is that OA means immediate open access<sup>22,23</sup>. Certainly, if

<sup>21</sup> [https://wiki.creativecommons.org/Frequently\\_Asked\\_Questions](https://wiki.creativecommons.org/Frequently_Asked_Questions) (visited on 2 May 2015)

<sup>22</sup> "Delayed Access is neither Green OA nor Gold OA" (<http://goo.gl/gGIxWv> visited on 2 May 2015)

publishers use different custom licenses instead of CC licenses, the perpetuity and immediateness concepts associated with OA wouldn't apply and the *cc:license* element wouldn't be enough to identify OA. However, CC is universally accepted by hybrid journals. Data analyzed by the project shows that every hybrid journal accepts CC licenses<sup>24</sup>. Therefore, using *cc:license*, together with *dc:rights*<sup>25</sup>, becomes a suitable, less onerous and low-barrier solution to identification of OA articles published in those journals; with the ease of implementation illustrated by the experiments run with JournalTOCs, an aggregator of scholarly journal RSS feeds.

#### 4. Results of Prototyping Programmatic OA Identification

Five participating publishers added the *cc:license* and *dc:rights* elements to their RSS feeds in a matter of weeks. Three of them, whose feeds were already following the CrossRef recommendations for scholarly feeds<sup>26</sup>, needed only a week. When NISO RP-22-2015 became available, publishers were given the choice of implementing either NISO or JEMO elements or both. At the end of the project over 20 publishers were using the *cc:license* in their RSS metadata, including SpringerOne and Biomed Central. No publisher had implemented the new *free-to-read* and *license\_ref* elements yet. Once RSS feeds providing OA elements in their metadata became available, JournalTOCs was able to create an API exposing OA articles collected from different gold and hybrid journals<sup>27</sup> and demonstrate the benefits of those new elements. The experience has shown that the maturity of the metadata specifications in question, the level of support from experts and validation services are important factors for adoption. As long as the e-journal supply chain components are unable to parse new elements, metadata providers will use what is easier and convenient for them.

Content providers prefer to provide metadata with the minimal effort possible for them [22, 23]. For example, only 50% of journal TOC RSS feeds use the CrossRef recommendations for RSS feeds published in 2009 [24]. This fact should make us cautious when proposing new metadata elements. As the complexity and number of metadata specification increase, their adoption by metadata providers tends to proportionally decrease. Some publishers were reluctant to enrich their RSS feeds until the suggested metadata had reached a certain level of maturity and acceptance; confirming that new specifications create high barriers to adoption.

The new NISO elements suffer from the same problems affecting other standards that have low or incomplete adoption. They are rich in theory but demanding in practice. There is the over-optimistic assumption that aggregators will know how to fully implement the new specifications (e.g. the NISO recommendations don't provide any technical means of enforcement for its *start\_date* and *license* URI attributes of the new NISO elements, leaving the decision to aggregators.) The recommendation of NISO in some way contradicts its own advice that before creating new metadata elements, adapting existing schemas should be considered. "*We use standards to*

---

<sup>23</sup> Publishing open access makes your work immediately and permanently available online for everyone, worldwide, <http://www.springer.com/open+access> (visited on 2 May 2015)

<sup>24</sup> <https://openjemo.wordpress.com/2015/05/04> (visited on 5 May 2015)

<sup>25</sup> Stating the rights associated with the CC license is recommended because CC licenses are operative only when applied to material in which a copyright exists.

<sup>26</sup> [http://oxford.crossref.org/best\\_practice/rss/](http://oxford.crossref.org/best_practice/rss/) (visited on 9 May 2015)

<sup>27</sup> <http://www.journaltoCs.ac.uk/api/articles/oa/> (visited on 10 May 2015)

*improve interoperability and to reduce unnecessary variation. It is better and easier to adopt something that already exists, is well modelled, and comprehensively supported.”*<sup>28</sup>

What becomes apparent is that the participating publishers were actually prepared and willing to implement the simplest of the available initiatives in their publishing workflow. This response made sense because publishers will normally be more disposed to implement a new specification if it involves using elements with which they are already familiar. The fact that CC and its different licensing flavors are used by practically all the publishers of hybrid journals was an important factor in their quick understanding and adoption of the JEMO CC-based tagging scheme.

## 5. Conclusion

Open Access articles are being erroneously hidden behind subscription-access walls because the OA status of articles is not embedded in all of its metadata manifestations shared by the multiple databases and discovery services involved in the e-journals delivery chain. The confusing landscape of various standard metadata exchange specifications proposed to cover every free to read possibilities, without giving a particular solution for OA articles, escalates the problem. The JEMO project has shown that using Creative Commons and Dublin Core elements is an easy and effective option for metadata providers (e.g. publishers) and consumers (e.g. discovery services) to programmatically identify OA at the article level. OA identification will eventually fail if OA status is not embedded in all metadata manifestations in the e-journals delivery chain.

Instead of creating new general-purpose specifications, we argue that efforts should be directed to implement elements that are already part of schemas being used by publishers and to enable OA identification at the article level on any online service used to access full-text.

*cc:license* provides a framework for conveying essential information that addresses common OA use cases. All publishers of hybrid journals offer CC licenses.

The JEMO case studies demonstrate that publishers do not intend to hide their Open Access articles behind subscription walls; it is, rather, a question of whether the e-journals delivery chain is propagating the appropriate forms of access in the right places.

This study has tangentially uncovered some problems with discovery services. We have shown that because discovery services are not using OA elements in their metadata, users are being denied access to OA articles published in subscription journals. Articles that were tagged as OA on the publishers' websites are being kept undiscoverable as OA in discovery services.

Our study has demonstrated that enabling programmatic identification of OA at the article level would enhance current services; hence benefiting both the research community and the OA hybrid business model.

---

<sup>28</sup> ISO/TC 46/SC11N800R1 Recommendations (as seen at <http://goo.gl/Dv5zjp> on 5 May 2015)

## References

- [1] S.A. Knowlton, I. Kristanciuik, M. Jabaily, Spilling Out of the Funnel: How Reliance Upon Interlibrary Loan Affects Access to Information, *Library Resources & Technical Services*. **59**(1) (2015), doi: 10.5860/lrts.59n1.4
- [2] C. Bullock and N. Hosburgh. OA in the library collection: The challenges of identifying and managing open access resources. *The Serials Librarian* (In Press)
- [3] Fair Prices for Article Processing Charges (APCs) in Hybrid Journals, *UK Open Access Implementation Group*, <http://goo.gl/cGBY2l> (as of April 2014)
- [4] Not all hybrid is equal, *Australian Open Access Support Group* (AOASG) Weblog, <http://aoasg.org.au/not-all-hybrid-is-equal/> (as of March 2014)
- [5] S. Pinfield, J. Salter and P.A. Bath, The 'total cost of publication' in a hybrid open-access environment: Institutional approaches to funding journal article-processing charges in combination with subscriptions, *J. of the Assoc. for Inf. Science and Tech.* (In Press) <http://goo.gl/WW59U4>
- [6] S.M. Shieber, Equity for Open-Access Journal Publishing, *PLoS Biology* **7**(8) (2009), doi: 10.1371/journal.pbio.1000165
- [7] Wellcome Trust calls for greater transparency from journals on open access publishing costs, *Europe PubMed Central Weblog*, <http://goo.gl/1BjFyJ> (as of December 2013)
- [8] C. Xiaotian, Journal Article Retrieval in an Age of Open Access: How Journal Indexes Indicate Open Access Articles, *Journal of Web Librarianship* **7**(3) (2013) doi:10.1080/19322909.2013.795426
- [9] S. Chumbe, B. Kelly and R. MacLeod, Hybrid journals: Ensuring systematic and standard discoverability of the latest Open Access articles, *The Serials Librarian* **68**(1-4) (2015), 143–155 doi: 10.1080/0361526X.2015.1016856.
- [10] C. Hutchens, Open access metadata: current practices and proposed solutions, *Learned Publishing* **26**(3) (2013), 159–165(7) DOI: 10.1087/20130302.
- [11] N. Lagace and G. Tananbaum, NISO Open Access Metadata and Indicators Working Group: Creating a Cross-Audience Solution, *The Serials Librarian* **65**(2) (2013), doi: 10.1080/0361526X.2013.813892.
- [12] B.-C. Björk, The hybrid model for open access publication of scholarly articles: A failed experiment? *Journal of the American Society for Information Science and Technology* **63**(8) (2012), 1496–1504, doi: 10.1002/asi.22709.
- [13] T. Koler-Povh, P. Južnič and G. Turk, Impact of open access on citation of scholarly publications in the field of civil engineering, *Scientometrics* **98**(2) (2013), 1033–1045.
- [14] S. Armato III, C. Baldock and C.G. Orton. "Hybrid gold" is the most appropriate open-access modality for journals like Medical Physics, *Med. Phys.* **42**(1) (2015) doi: 10.1118/1.4895979.
- [15] S. Glasser, NISO Webinar: It's Only as Good as the Metadata: Improving OpenURL and Knowledge Base Quality. *Serials Review* **37**(1) (2011), 58–60, doi: 10.1080/00987913.2011.10765348.
- [16] A. Tani, L. Candela and D. Castelli, Dealing with metadata quality: The legacy of digital library efforts, *Information Processing & Management* **49**(6) (2013), 1194–1205, doi: 10.1016/j.ipm.2013.05.003.
- [17] C.M. Yassera. An Analysis of Problems in Metadata Records. *Journal of Library Metadata*. **11**(2) (2011), 51–62, doi: 10.1080/19386389.2011.570654.
- [18] J-R Park. Metadata Quality in Digital Repositories: A Survey of the Current State of the Art, *Cataloging & Classification Quarterly*, **47**(3-4) (2009) 213-228, doi: 10.1080/01639370902737240
- [19] M. Van Ballegooiea. Knowledgebases: The Cornerstone of E-Resource Management and Access. *Serials Review*. **40**(4) (2014) pp. 259-266. doi: 10.1080/00987913.2014.977127
- [20] Access License and Indicators: A Recommended Practice of the National Information Standards Organization, NISO RP-22-2015, available at: <http://www.niso.org/publications/rp/rp-22-2015> (as of January 2015)
- [21] Step by Step Guide to enable OA identification from RSS feeds, *JEMO Project* weblog, <http://goo.gl/ZXWpPb> (as of April 2015)
- [22] D. Mietchen, C. Maloney and N.D. Moskopp, Inconsistent XML as a Barrier to Reuse of Open Access Content, *Journal Article Tag Suite Conference (JATS-Con) Proceedings* 2013 <http://www.ncbi.nlm.nih.gov/books/NBK159964/>
- [23] S.D. Shapiro, We are all aggregators (and publishers) now: how discovery tools empower libraries, *Library Hi Tech News* **30**(7) (2013), 7–9, doi: 10.1108/LHTN-07-2013-0041
- [24] Recommendations on RSS Feeds for Scholarly Publishers, [http://oxford.crossref.org/best\\_practice/rss/](http://oxford.crossref.org/best_practice/rss/) (as of May 2015)