

## HHuLOA project RIOXX review

### Mandatory

#### ali:licence\_ref

##### Notes

- The proposal that the latest start\_date attribute takes precedence is not built into the metadata, and will be left to systems to interpret. Is additional context required to make this clearer in the metadata itself (e.g., as with use of first\_named\_authot)?
- What happens if the start\_date is the same for different licences?
- Publisher licences are not normally captured, so this is a new piece of metadata to hold. We need to know we are capturing the right information.
- Not all publishers have licence pages that can be used. Many of these will be HTML pages, which could change during the period of the embargo (or disappear, leading to broken links). This might lead to incorrect information being conveyed.
- How should this element be used if there are multiple documents with different licence terms? Should a separate RIOXX record be generated for each document?

**Proposed practice – It is noted that NISO request that this information be made available by publishers. As this is new and there is some uncertainty about capturing and holding this information, it is proposed to default to the RIOXX ‘all rights reserved’ URLs in lieu of publisher information, and map licence information held already where feasible (e.g., CC licences). Addition of publisher URLs will be reviewed as the NISO request is implemented.**

#### dc:identifier

##### Notes

- What if you are referencing a file hosted elsewhere where the available URL is a landing page? It is noted that linking to the resource itself is recommended, so it is assumed that if only a landing page is available this will suffice.

**Proposed practice – This is a straightforward mapping of the link to the resource content.**

#### dc:language

##### Notes

- This will map from language metadata (if held).
- Is it feasible to do a look-up from a standard ISO list?

**Proposed practice – Map from information held in the repository. If not held as standard, default to ‘eng’ and amend as required for records in other languages.**

## dc:source

### Notes

- How to select which identifier to use if more than one (ISSNs/ISBNs are rarely unique now)? Does this matter so long as there is a reference?
- EISSN/EISBN preferred (though note print journal alternative need if item is only published in this way).
- It is unclear that Sherpa acts as an authority for ISSNs, as it is more a convenient reference point.

**Proposed practice – Map from content held in the record where this is available. Add where it is missing.**

## dc:title

### Notes

- The recommended format may not be how the original title/subtitle is quoted. It is assumed that we hold the title as is, making use of the recommendation where there is uncertainty.

**Proposed practice – Map from title as held.**

## dcterms:dateAccepted

### Notes

- Although not yet held in many cases, it is recognised that this is an element we will all have to capture from now on. Systems enhancements underway should provide this.
- Retrospective addition may be tricky (and manual).
- Entering date elements to create ISO standard format may require separate fields.

**Proposed practice – Hold and provide as needed using appropriate system (repository and/or CRIS) following any enhancements required.**

## rioxterms:author

- ORCID option – this was not felt to be not widely captured yet, but capability is coming along.
- HR departments are looking at adding this, and would then feed the ORCID into the repository for use in RIOXX records. There will be local solutions for this.
- It is noted that institutional creation now no longer seen as default – leave it to individuals.
- The use of an ISNI is unclear. How are these generated and at what organisational level?
- Why is ‘author’ the preferred term (as opposed to ‘creator’ from the original DC element)? Many research outputs do not have ‘authors’ as such.
- First-named-author needs to be flagged programmatically somehow. This will depend on how systems are set up to make use of it.
- The format of names is noted, and may require separate capture of parts to create this.

**Proposed practice – Map from author information held in the repository. Add ORCiDs as local solution enables.**

#### **rioxxterms:project**

##### Notes

- An EPrints array to capture this information will need coding into the capture form. Ditto for Hydra.
- FundRef can be used for authority records for funder name and id
- This may be better mapped from a RIS rather than capturing it directly in the repository.
- What about unfunded research? Should we default to institution as funder behind the scenes?

**Proposed practice – Systems will need to be enhanced to hold the information required in separate elements to create the RIOXX information. Information will be taken from other systems holding it already rather than create it within the repository.**

#### **rioxxterms:type**

##### Notes

- Many such lists are in existence – what is the origin of this list, and why wasn't another one re-used?
- Need to map types between local and RIOXX lists and put in place process changes required.
- Will there be a mapping created between this list and REF categories?
- It will be useful to have a controlled list to pick from. Can the list be provided as an ontology to reference?

**Proposed practice – Map local lists to RIOXX list and generate appropriate mappings to create RIOXX records**

#### **rioxxterms:version**

##### Notes

- EPrints has created a mapping from default EPrints version labels. Hydra will use as set out as no default entries are currently added.
- How do you determine which version is which? It is likely this will vary by publisher depending on policy, but policy is often unclear and not always present.
- NA will be used if the item is a metadata only record.
- It is assumed VoR is if publisher version used.

**Proposed practice – Default to NA unless there is a clear indication of which version is being used. Additional guidance on how to apply these options is urgently needed.**

## Recommended

### dc:description

#### Notes

- Always recorded, but primarily for discovery rather than reporting.
- Use in RIOXX a matter of preference. It could support discoverability through Gateway to Research.

**Proposed practice – Provide where held otherwise, but don't add specifically for RIOXX purposes.**

### dc:format

#### Notes

- Should be fine, as mapped from the file referenced by dc:identifier.

**Proposed practice – Map from information held within repository. Should be fairly straightforward to provide.**

### dc:publisher

#### Notes

- Should be captured already, but it should be from a controlled list as recommended.
- It is unclear whether Sherpa can act as an authority for these names.
- Entries are most likely to be free text unless a controlled list can be used.

**Proposed practice – Provide information as held. Implement a controlled list as this becomes available.**

### dc:subject

#### Notes

- Likely to get what has been entered. This will vary from record to record.
- Not often controlled – this needs more work.

**Proposed practice – Provide information where held. Work towards finding ways of making such entries standard from a controlled vocabulary.**

### rioxxterms:version\_of\_record

#### Notes

- DOI or whatever else is available.
- Make sure the full HTTP version is used.

**Proposed practice – Map whichever link to the version of record is held in the repository, using the DOI as default but a URI where the DOI is not available.**

## Optional

### ali:free\_to\_read

#### Notes

- Start date – this will be the same as that used in ali:licence\_ref.
- Use the deposit or embargo end date, whichever is later.
- Only applicable if the item is free to read.
- Doesn't feel like a priority given other fields to focus on – but could be useful.
- Will work best if using information captured elsewhere already.
- Could provide an answer to questions on how much is openly available.
- Information may need to be added later after accepted date deposit.

**Proposed practice – This needs more investigation prior to implementation. Not a priority given the need to make other elements available.**

### dc:coverage

#### Notes

- Not sure this is adding much value.
- Can use what is recorded where this is the case.

**Proposed practice – Map information if this is held.**

### dc:relation

#### Notes

- If links are recorded then they can be output as part of this
- Funder links will be made better via rioxxterms:project
- Should this record the landing page or content in URI?

**Proposed practice – Map information if this is held.**

### rioxxterms:apc

#### Notes

- Should be used, would love to, but don't have the information easily accessible – something to work towards.
- Data is likely to be held and managed elsewhere rather than as a flag in the repository.

**Proposed practice – Omit for now, and work towards adding this as information becomes more easily accessible.**

### rioxxterms:contributor

#### Notes

- Very unlikely to be used for texts (although more likely for other materials).
- Hard to control the meaning of what is entered.

**Proposed practice – Map this information where there is obvious relevance, but unlikely to be common.**

## **rioxterms:publication\_date**

### Notes

- Fairly straightforward, so long as you have clarity about the publication information you are collecting.
- For use with bibliographic purposes, but of less value in funder reporting.
- Can be useful for internal reporting.
- Has to be added after the initial record has been created due to acceptance deposit process.
- Note that this should map to the start\_date in the ali elements, but is held here as free text.

**Proposed practice – Add this information where known.**

Chris Awre, University of Hull  
Graham Stone, University of Huddersfield  
Paul Stainthorp, University of Lincoln

This work is licensed under a [Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License](https://creativecommons.org/licenses/by-nc-sa/4.0/).