



Open Research Online

Citation

George, Macgregor; Knoth, Petr; Walk, Paul; Dowson, Nicola; Eadie, Mick; Jones, Beverley and Martínez-García, Agustina (2024). Exploring the concept of 'custodianship' in harvesting repository resources and graphing their relations: Rioxx version 3.0. In: The 19th International Conference on Open Repositories, 03-06 Jun 2024, Göteborg, Sweden.

URL

<https://oro.open.ac.uk/102475/>

DOI

<https://doi.org/10.5281/zenodo.12579240>

License

(CC-BY 4.0) Creative Commons: Attribution 4.0

<https://creativecommons.org/licenses/by/4.0/>

Policy

This document has been downloaded from Open Research Online, The Open University's repository of research publications. This version is being made available in accordance with Open Research Online policies available from [Open Research Online \(ORO\) Policies](#)

Versions

If this document is identified as the Author Accepted Manuscript it is the version after peer review but before type setting, copy editing or publisher branding

Exploring the concept of 'custodianship' in harvesting repository resources and graphing their relations: Rioxx version 3.0

George Macgregor

(Petr Knoth, Paul Walk, Nicola Dowson, Michael Eadie, Beverley Jones + Agustina Martínez-García)

University of Glasgow

2024-06-04

<https://purl.org/g3om4c>

<https://orcid.org/0000-0002-8482-3973>



Docs available: <https://eprints.gla.ac.uk/320745/>

Overview

1. Brief history + introduction to Rioxx
 - Sequel to [OR2023](#) [1]
2. Prior attempts to model expressions, relations, etc. of scholarly works + burgeoning scholarly graph
3. Rioxx: Research Output Metadata Schema (v3.0):
 - PID-centric + relational data reality
 - 'direct' and 'external custodianship'
 - discovery + interoperability
 - implementations (DSpace) + aggregation
4. Concluding remarks

Repository metadata context

- Repositories generally good at making content discoverable
 - [OAI-PMH](#) still de facto machine interface despite alternatives
 - Structured data within HTML, sitemaps + other animals...
- Folly of 'simple' Dublin Core [[2](#)]
 - Metadata profiles critical
 - Harvesting, aggregation, discovery and... compliance

Rioxx...

Rioxx: Research Output Metadata Schema

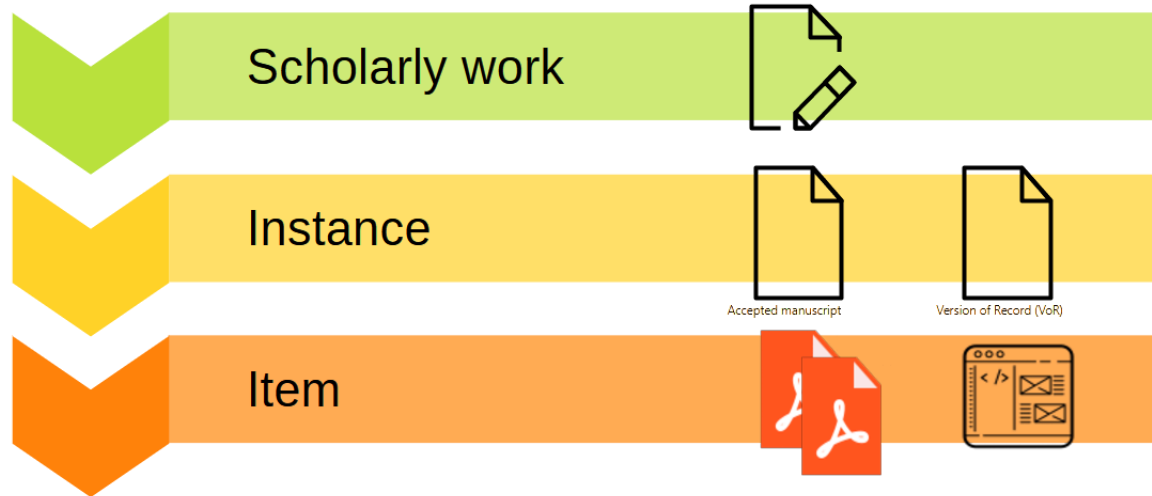
- [Version 2.0](#) widely adopted since 2016; Dublin Core with extensions
 - See Walk et al. (OR2023) for history [1]
 - UK centric + funder compliance
- Delivered discovery improvements [3]
 - Repositories default support of Dublin Core spectacularly ineffective for OAI-PMH harvesting of digital content
 - file location links critical [4]

Panel session: *How to make repository content indexed and discoverable* Hope you were there! 😊

Entities + their relations

SWAP: Scholarly Works Application Profile [5]

- Prescient!
- Recognized importance of relations between entities, esp. funding
- Used [FRBR](#) as conceptual tool



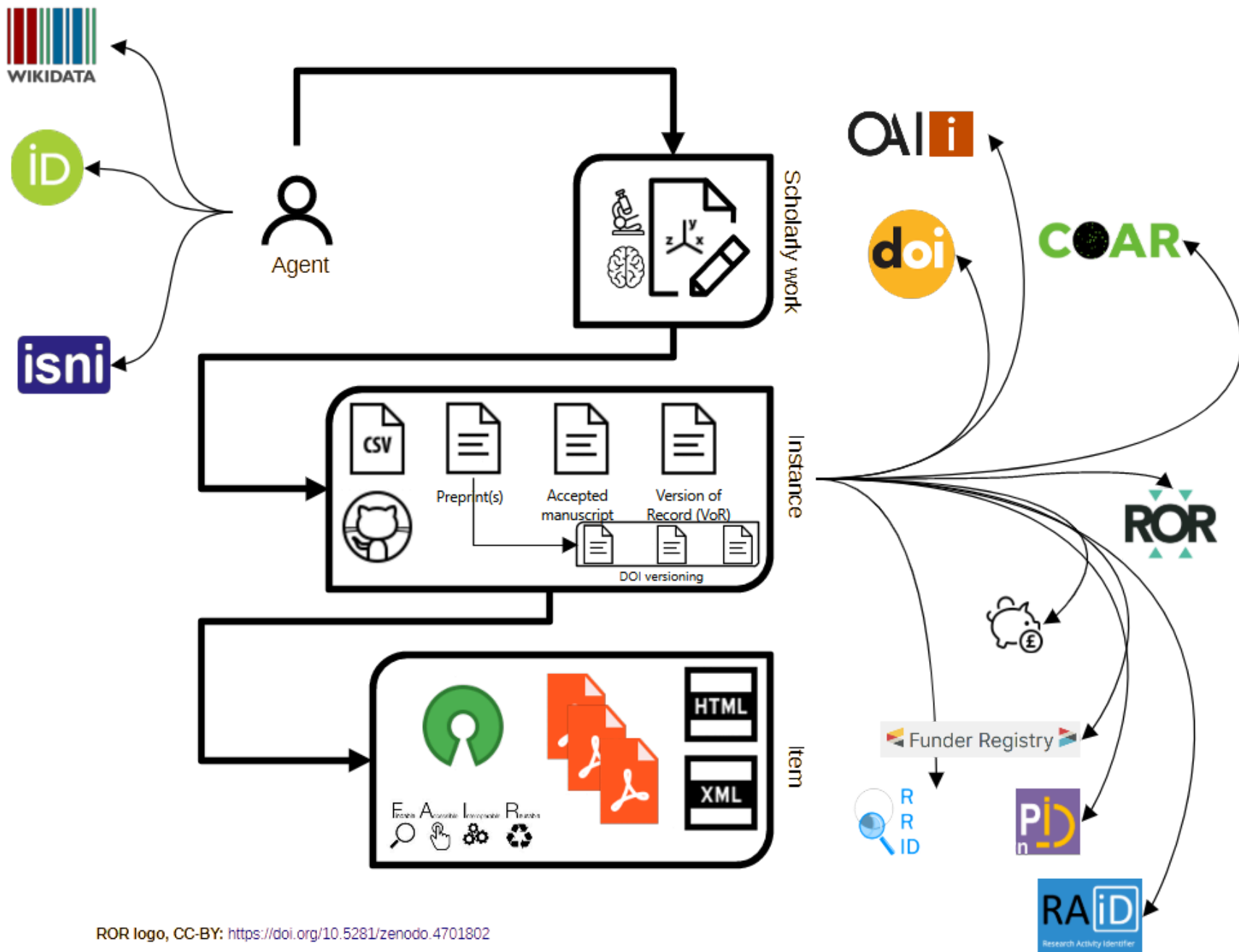
...

But never adopted by repositories at scale

- Too prescient (circa 2008)...?
- Too esoteric for those working with scholarly digital content [2]
- Useful conceptual exercise but did not address machine discovery satisfactorily

Developments in scholarly works + open research...

1. Need to respond to complexity while ensuring discovery advantages
 - i. *Provenance + contextualization*
2. Open research requirements of funders (Plan S, UKRI, G7)
3. *Transparency = reproducibility, verification, replication*
 - i. Maintenance of the scholarly record + the "reproducibility crisis"
 - ii. Growth of [rights retention](#), [FAIR](#), data management planning (DMPs)
4. Supporting the burgeoning, diverse 'PID graph'
5. Scholarly works as (unofficial) multi-part or multi-object outputs
 - i. Being 'inclusive' about nature of works



ROR logo, CC-BY: <https://doi.org/10.5281/zenodo.4701802>

Rioxx v3.0

Version 3.0 - published! 😊

- Improves modelling of scholarly entities + relations
 - Borrows some conceptual thinking from [BIBFRAME](#)
 - Jurisdiction agnostic
- Capitalizes on discovery potential
- Better supports productive contribution to PID graph
 - *PID-ification* -- greater URI referencing + semantics
- Alignment with [Signposting](#) and [ResourceSync](#)

Since OR2023...

- Introduced: `<rioxxterms:grant>` + `<rioxxterms:project>`
- Improved identification: `<rioxxterms:name>` + `<rioxxterms:id>`, sub-properties of `<rioxxterms:creator>` + `<rioxxterms:contributor>`
- Introduced: notion of *direct* + *external custodianship* via `dc:relation` + `rioxxterms:ext_relation`
- Introduced: better modelling of dates in lifecycle
 - `<rioxxterms:record_public_release_date>`, attributes within `<dc:relation>` + `<rioxxterms:ext_relation>`
- Major change to the use of `<dc:identifier>`
 - HTTP(S) URI identifying *primary resource* resolving to repository landing page

v3.0: Vocabularies, semantics, + PID types

(Beyond structure) language independent semantics conveyed by SKOS:

- [COAR Resource Type Vocabulary](#)
- [COAR Access Rights Vocabulary](#)
- [COAR Version Type Vocabulary](#)

Resource Type Label: '**observational data (English)**', '**gözlemsel veri (Türkçe)**', etc.

http://pur1.org/coar/resource_type/FF4C-28RK

Broader concept: '**dataset**'

http://pur1.org/coar/resource_type/c_ddb1

...

Referral to entities by URIs widely supported:

- **Creators/Contributors:** *ORCID*, ISNI, VIAF, WikiData
- **Organizations:** ISNI, VIAF, WikiData - *ROR*, *FundRef*
- **Research activity:** *RAiD*

Optimum use of URIs (PIDs) for reference + relational associations between related works + instances to enrich scholarly graph, support discovery, contextualization

- Issues with 'authority of assertion' [6]

Example snippets...

```
<rioxxterms:creator>  
  <rioxxterms:name>Fry, Hannah</rioxxterms:name>  
  <rioxxterms:id>https://isni.org/isni/0000000446254946</rioxxterms:id>  
  <rioxxterms:id>http://viaf.org/viaf/314908506</rioxxterms:id>  
  <rioxxterms:id>https://orcid.org/0000-0003-0601-9100</rioxxterms:id>  
</rioxxterms:creator>
```

```
<dc:relation
  rel="item"
  type="application/pdf"
  coar_type="https://purl.org/coar/resource_type/c_6501"
  coar_version="https://purl.org/coar/version/c_970fb48d4fbd8a85"
  deposit_date="2023-10-18"
  resource_exposed_date="2023-10-18"
  access_rights="https://purl.org/coar/access_right/c_abf2"
  license_ref="https://creativecommons.org/licenses/by/4.0/">
  https://eprints.whiterose.ac.uk/204315/1/amt-16-4375-2023.pdf
</dc:relation>
```

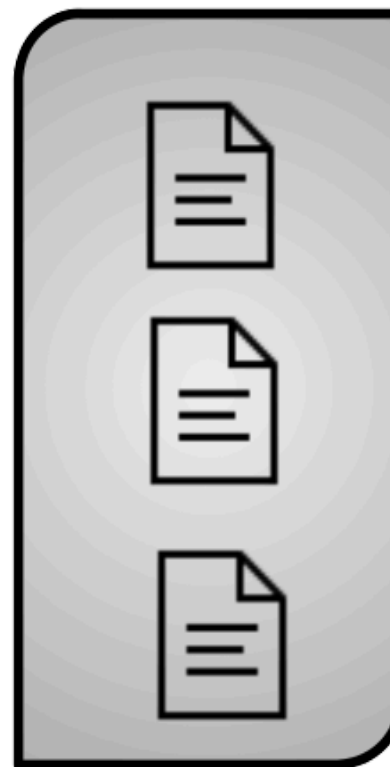
```
<dc:relation
  rel="item"
  type="application/pdf"
  coar_type="http://purl.org/coar/resource_type/YZ1N-ZFT9"
  deposit_date="2023-10-18"
  resource_exposed_date="2023-10-18"
  access_rights="https://purl.org/coar/access_right/c_abf2"
  license_ref="https://creativecommons.org/licenses/by/4.0/">
  https://eprints.whiterose.ac.uk/204315/9/amt-16-4375-2023-supplement.pdf
</dc:relation>
```

```
<dc:relation rel="cite-as">https://oai.core.ac.uk/oai:eprints.whiterose.ac.uk:204315</dc:relation>
```


'Externally' curated objects



'Internally' curated objects



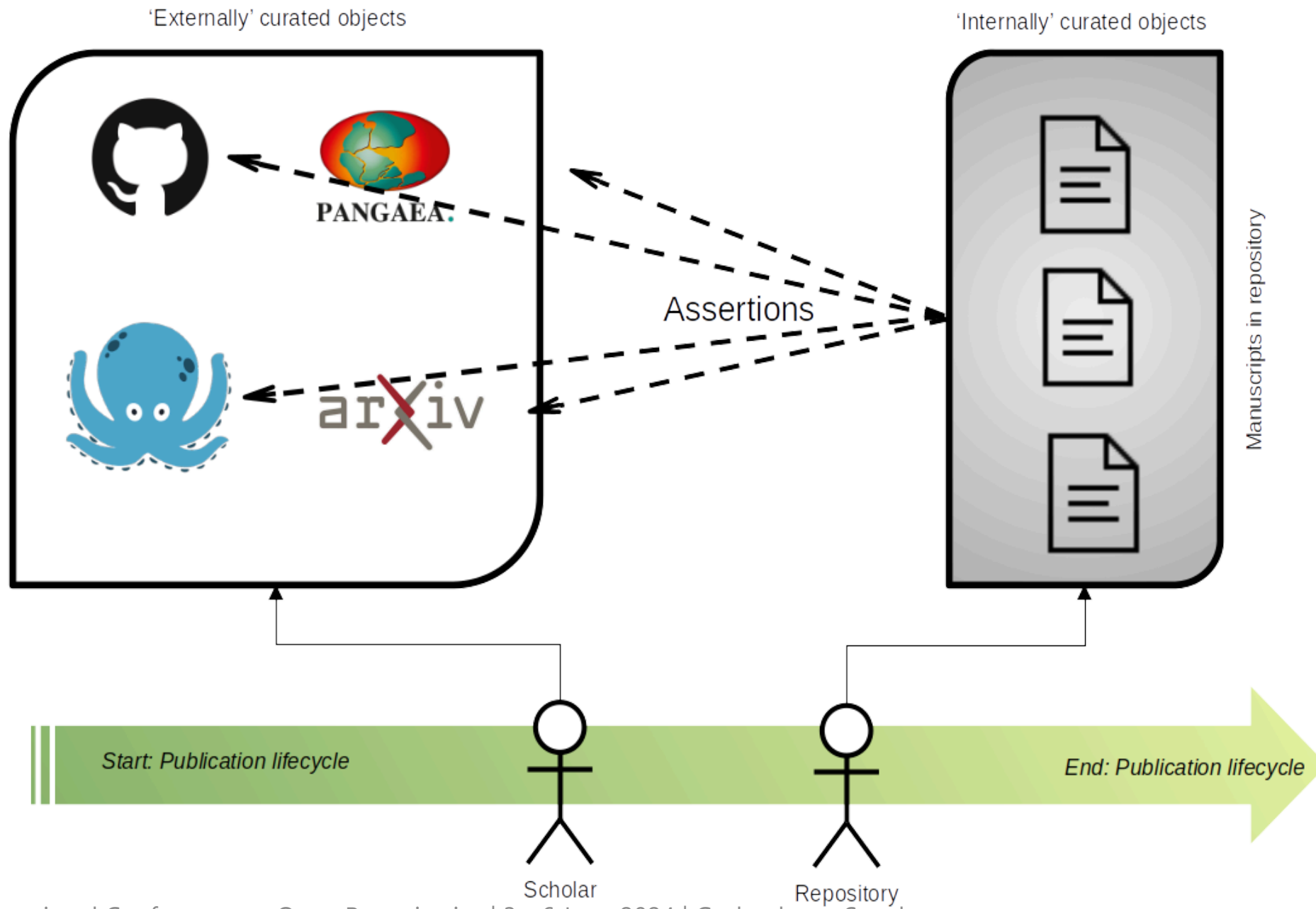
Manuscripts in repository



Scholar



Repository actor



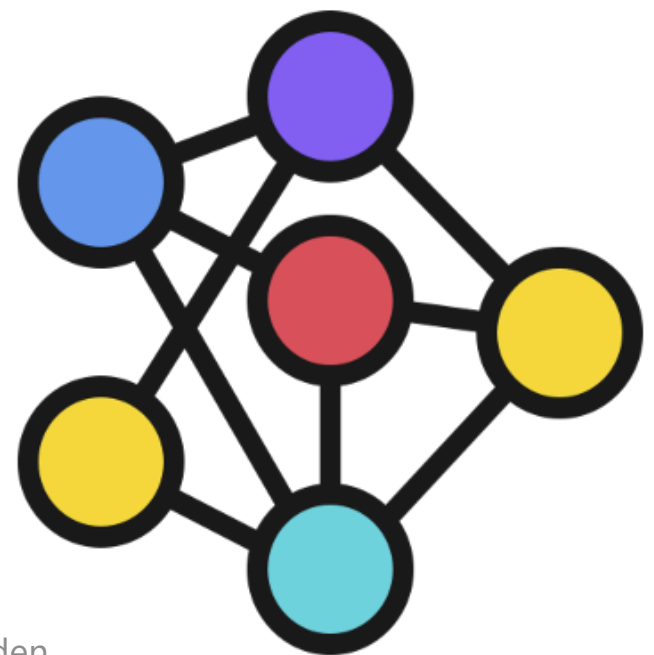
Unification of metadata – curation host irrelevant to end user



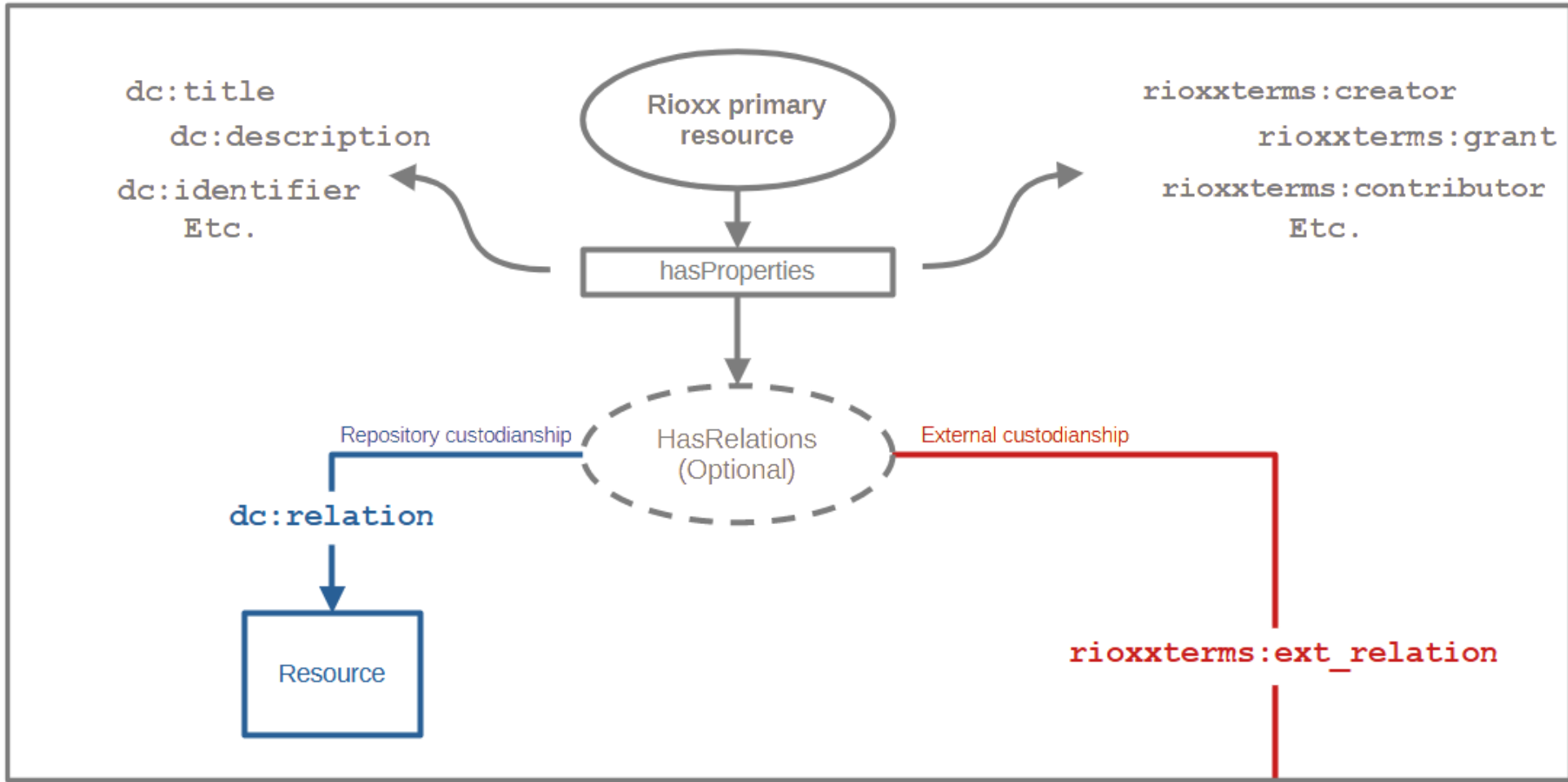
Manuscript in repository
/OR/
E-resource holdings



User,
Human or
machine



Meaningful, reliable, verifiable, truth
contributions to and from the PID
graph



- = direct custodianship (repository)
- = external custodianship

'Custodianship'

custodianship - notion that location of resources significant to understanding nature of thing being described

- relevant to machine processing of Rioxx data
- *primary resource* + other *resources*, inc. URIs
 - *direct custodianship* - those under direct management of a local repository or publication platform
 - local system controls + maintains file content and/or other resources
 - *external custodianship* - resources outside *direct custodianship*, such as those hosted by 3rd party services

Introduction of `rioxterms:ext_relation` :

```
<!-- Other expressions (or 'instances') - publisher version -->
<rioxterms:ext_relation
  rel="cite-as"
  coar_type="https://purl.org/coar/resource_type/c_6501"
  coar_version="https://purl.org/coar/version/c_970fb48d4fbd8a85">
  https://doi.org/10.1007/s11229-020-02724-x
</rioxterms:ext_relation>

<!-- related dataset -->
<rioxterms:ext_relation
  rel="cite-as"
  coar_type="https://purl.org/coar/resource_type/c_ddb1">
  https://doi.org/10.15129/5d28213e-8f9f-402a-b550-fc588518cb8b
</rioxterms:ext_relation >

<!-- related software -->
<rioxterms:ext_relation
  rel="cite-as"
  coar_type="https://purl.org/coar/resource_type/QH80-2R4E">
  https://doi.org/10.5281/zenodo.3478185
</rioxterms:ext_relation>
```

Repositories on front line of graph data contributions

distribution + fragmentation + fluidity

Repository important component of making scholarly works:

Interpretable....

- to contextualize; to provide provenance + transparency

Support reproducibility....

- to facilitate reproducibility, verification, replication

Discoverable....

- to enable discovery through a variety of access points, some unconventional

Complete examples...

A number of examples available from Rioxx 3.0 documentation

		<p>source is referenced using a unique identifier from a recognised system e.g. the unique 8-digit International Standard Serial Numbers (ISSN) assigned to <i>electronic</i> periodicals, or the 13 digit International Standard Book Number (ISBN13) assigned to books. In the latter case, the ISBN13 for the electronic version of the book SHOULD be used if available.</p> <p>Use of this property is applicable where <i>the primary resource</i> is to be published as part of a larger resource. Typical examples might include a conference paper belonging to proceedings or a chapter of a book (but not a complete book).</p>
dc:subject	Zero or more	<p><code>dc:subject</code> is used to describe the subject of <i>the primary resource</i>. It is RECOMMENDED that a URI be used for the value, however a literal value that identifies the subject may be provided instead. Both should preferably refer to a subject in a controlled vocabulary.</p>

Example Rioxx Records (XML)

Example 1



Example 2



Example 3



Example 4



This website and blog are licensed under a [Creative Commons Attribution 4.0 International License](#)
Built & deployed with [Hugo](#) v. 0.124.1 on Thu, 02 May 2024 13:05:09 UTC.

Rioxx v 3.0 support: DSpace + CORE

- [DSpace integration work](#), developed + donated by [University of Cambridge](#)
 - Docs available: <https://wiki.lyrasis.org/x/IgBOEw>
 - Available in DSpace 8.0 (released in late May 2024)
 - Cambridge exploring release of DSpace 7.x patch
 - [Example article record](#), from Apollo OAI-PMH endpoint
 - [ListRecords](#), from Apollo
- [CORE: Rioxx validator via CORE Repository Dashboard](#)
- [EPrints development stirring on GitHub...](#)

Concluding remarks

- v3.0: exposing repository content in a relational way but...
 - supporting discovery + efficient harvesting + graphing
- Long road -- change always a certainty
 - 'Aggressively open'
- Advocate for adoption -- technical but also socio-technical
- JSON-LD serialization of Rioxx necessary....?

References

- [1] Walk, P., Macgregor, G., & Knoth, P. (2023). [Rioxx 3: A Modernised Metadata Profile](#). *The 18th International Conference on Open Repositories (OR2023)*, Stellenbosch, South Africa. Zenodo.
- [2] Knoth, P. (2013). [From open access metadata to open access content: two principles for increased visibility of open access content](#). *The 8th International Conference on Open Repositories (OR2013)*, Charlottetown, Prince Edward Island, Canada.
- [3] P. Knoth and B. Notay, '[UKRI OA policy requirements for repositories and how to meet them](#)', presented at the *Jisc Workshop, 2021*.

...

[4] P. Knoth, M. Cancellieri, M. Klein, '[Comparing the performance of OAI-PMH with ResourceSync](#)', *The 14th International Conference on Open Repositories (OR2019)* June 2019. Universität Hamburg, Hamburg.

[5] J. Allinson, '[Describing Scholarly Works with Dublin Core: A Functional Approach](#)', *Library Trends*, 57 (2), pp. 221–243, 2008. Accessed Jul. 18, 2023.

[6] Macgregor, G., Lancho-Barrantes, B. S., Rasmussen McAdie, D. (2023) [Authority of assertion in repository contributions to the PID graph](#). "Re-Discovery": *Metadata & Discovery Group (MDG) Conference 2023*, Sept. 2023, IET Birmingham: Austin Court, Birmingham.

Acknowledgements

Rioxx acknowledgements

- Rioxx Governance Group (chaired by Agustina Martínez-García, University of Cambridge)

