The Challenge of Disseminating Metadata on Library Published Open Access Books

Vincent W.J. van Gerven Oei
No open access without open infrastructure.
a collaborative project  
buiding infrastructure  
to support Open Access books
Open Metadata for OA Publishing

• Smaller, non-profit, and library-based OA publishers often struggle with metadata management and dissemination

• Commercial solutions are expensive/not tailored to OA books > metadata often managed in custom spreadsheets

• Dissemination channels require different output formats according to different specifications > technical knowledge required is often absent, leading to lower discoverability

• OA book metadata from commercial vendors often of low(er) quality

• Metadata often not CC0 licensed
Open Metadata for OA Publishing

- Thoth: Open Metadata Management & Dissemination
- Open source, open documentation on wiki
Home

Rupert Gatti edited this page on Mar 4 · 288 revisions

The following wiki lists all output formats, platforms, and other stakeholders that relate to the development of open source metadata management system Thoth. Those wikis marked with an asterisk (*) are identified in the WP5 Scoping Report and are indexed as an issue in the Thoth Project. Those wikis marked with a caret (†) have been successfully implemented in Thoth.

The transfer protocols and file specifications listed in this wiki are those which have current or future support by Thoth. Any updates are welcome at wp5@copim.ac.uk.

The Thoth User Manual, for publishers and other creators of metadata records in Thoth, can be found here.

Data and Metadata

Data Formats

Data formats are the formats in which the textual content of an object are commonly digitally encoded. The Thoth project is agnostic as to which data format is used.

- AZW3
- EPUB
- HTML
- KFX
- MOBI
- PDF
- XLM
Open Metadata for OA Publishing

• Thoth: Open Metadata Management & Dissemination
  • Open source
  • Tailor-made for Open Access book metadata
Edit institution

Member(s) contributed to: Lacidari 1, from: punctum books
Member(s) contributed to: The Anglo-Scottish Ballad and its Imaginary Contexts, from: Open Book Publishers
Member(s) contributed to: Manifesto for a Post-Critical Pedagogy, from: punctum books
Member(s) contributed to: Continent: Year 1, from: punctum books
Member(s) contributed to: Life Histories of Ethn Theory in Russia and Beyond, from: Open Book Publishers
Member(s) contributed to: Lacidari 3, from: punctum books
Member(s) contributed to: Pedagogies of Disaster, from: punctum books
Member(s) contributed to: Lacidari 2, from: punctum books
Member(s) contributed to: [Given, If, Then], from: punctum books

Institution Name
Aberdeen University

Institution DOI
https://doi.org/10.13039/501100007874

ROR ID
http://ror.org/016478e91

Country
United Kingdom
Contributed to: Love Don't Need a Reason, from: punctum books

Given Name
Matthew J.

Family Name
Jones

Full Name
Matthew J. Jones

ORCID
http://orcid.org/0000-0002-9707-6655
Open Metadata for OA Publishing

- Thoth: Open Metadata Management & Dissemination
  - Open source
  - Tailor-made for Open Access book metadata
  - Multiple output formats and specifications for 10+ platforms
# Welcome to Thoth's GraphQL API explorer (GraphiQL).
# GraphQL is an in-browser tool for writing, validating, and
testing GraphQL queries.
# Type queries into this side of the screen, and you will see intelligent
typeahead suggestions of the currently available GraphQL type schema and throw syntax and
validation errors highlighted within the text.
# Click on the QueryRoot in the Documentation Explorer (< docs >) on the
right hand side of the screen to navigate the API schema.
# GraphQL queries typically start with a "{" character. Lines that start
with a # are ignored.
# An example Thoth GraphQL query might look like:
#
```graphql
{
  works(workType: MONOGRAPH, workStatus: ACTIVE) {
    shortTitle
doi
  }
}
```
# Keyboard shortcuts:
# Run Query: Ctrl-Enter (or press the play button above)
# Auto Complete: Ctrl-Space (or just start typing)
# Run the following query (Ctrl-Enter):
#
```graphql
{
  works(order: [FIELD: PUBLICATION_DATE, direction: ASC]) {
    shortTitle
doi
    publications {
      publicationType
      isbn
    }
    contributions {
      contributionType
      fullName
    }
  }
}
```
REQUEST

PATH PARAMETERS

<table>
<thead>
<tr>
<th>*string specification_id</th>
<th>onix_3.0:oopen</th>
</tr>
</thead>
<tbody>
<tr>
<td>*uuid publisher_id</td>
<td>9c41b13c-cecc-4f6a-a151-be4682915ef5</td>
</tr>
</tbody>
</table>

API Server: https://export.thoth.pub
Authentication: Not Required

Response Status: (CORS or Network Issue)

RESPONSE

200
OK
Open Metadata for OA Publishing

• Thoth: Open Metadata Management & Dissemination
  • Open source
  • Tailor-made for Open Access book metadata
  • Multiple output formats and specifications for 10+ platforms
  • CC0 licensed
Next Steps for Thoth

• Quality control via “traffic light” system
Next Steps for Thoth

• Quality control via “traffic light” system

• Automated metadata dissemination to platforms
Simplifying Dissemination Pipelines

Metadata entry

XML

PDF URL

PDF

Cloud68.co

downloads

PDF

emails

XML

punctumbooks

open

THOTH
Open Metadata.
Simplifying Dissemination Pipelines

Metadata entry

XML

punctumbooks

uploads

XML
PDF
JPG

JSTOR
Simplifying Dissemination Pipelines

Metadata entry

XML

Cloud68.CO

XML
PDF
JPG
crawls
XML
PDF
JPG

punctumbooks

xml

THOTH Open Metadata.

THOTH
Next Steps for Thoth

- Quality control via “traffic light” system
- Automated metadata dissemination to platforms
- Business model launch to assure long-term sustainability
Ingest

- API
- UI
- ONIX

Export

- GraphQL API
- Export API

Free

- Metadata Services
  - Formatting
  - Backlist ingest
  - Validation/Enhancement

Plus

- Distribution Services
  - Curated (meta)data delivery
  - Archiving/Preservation
  - Support/Training
Expanding the Thoth Ecosystem

• Integration with the Open Book Collective
• Harmonization with WikiData & OpenAlex
• White-label website built on top of Thoth
• Potential for collaborative editing of metadata records
• Toward an open, community-managed universal book catalog
Resources

Thoth

https://thoth.pub/

WP5 Scoping Report: Building an Open Dissemination System

https://copim.pubpub.org/pub/wp5-scoping-report-building-open-dissemination-system/

Open Metadata in Thoth

https://copim.pubpub.org/pub/open-metadata-thoth/