

The Challenge of Disseminating Metadata on Library Published Open Access Books

Vincent W.J. van Gerven Oei @ontakragoueke



copim

a collaborative project
building infrastructure
to support Open Access books



Research
England



ARCADIA
A CHARITABLE FUND OF
LISBET RAUSING & PETER BALDWIN

No open access
without open infrastructure.



UC SANTA BARBARA
LIBRARY

doab
directory of
open access
books

ScholarLed



MATTERING PRESS

BRITISH
LIBRARY



Loughborough
University



Digital Preservation Coalition



Birkbeck
UNIVERSITY OF LONDON

Coventry
University



punctumbooks

OpenBook
Publishers



meson press

Lancaster
University



TRINITY
COLLEGE
CAMBRIDGE



OPEN HUMANITIES PRESS

Jisc

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



master 5 branches 52 tags Go to file Add file Code

ja573 Merge branch 'release/v0.8.4' 59332ea 6 days ago 1,880 commits

| | | |
|---------------------|--|---------------|
| .github | Add question template | 3 months ago |
| src | Fix new Clippy errors (needless borrow) under 1.54.0 | 10 months ago |
| thoth-api-server | Bump v0.8.4 | 6 days ago |
| thoth-api | Bump v0.8.4 | 6 days ago |
| thoth-app-server | Bump v0.8.4 | 6 days ago |
| thoth-app | Bump v0.8.4 | 6 days ago |
| thoth-client | Bump v0.8.4 | 6 days ago |
| thoth-errors | Bump v0.8.4 | 6 days ago |
| thoth-export-server | Bump v0.8.4 | 6 days ago |
| .dockerignore | Set docker environment | 12 months ago |
| .env.example | Set docker environment | 12 months ago |
| .gitignore | initial commit | 2 years ago |
| CHANGELOG.md | Update changelog | 6 days ago |
| CODE_OF_CONDUCT.md | Use COPIM code of conduct | 3 months ago |
| Cargo.lock | Bump v0.8.4 | 6 days ago |
| Cargo.toml | Bump v0.8.4 | 6 days ago |
| Dockerfile | Add ca certificates to docker image | last month |
| LICENSE | Add license | 2 years ago |

About

Metadata management and dissemination system for Open Access books

thoth.pub

- graphql
- rust
- metadata
- books
- webassembly
- openaccess
- bibliographic-data

- Readme
- Apache-2.0 license
- Code of conduct
- 22 stars
- 9 watching
- 6 forks

Releases 51

v0.8.4 Latest 6 days ago

+ 50 releases

Contributors 5



Home

Edit New Page

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
- YMI

Pages 206

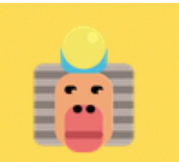
Find a Page...

- Home
 - Data and Metadata
 - Data Formats
 - Metadata Formats
 - Feed Formats
 - Citations and Bibliographies
 - Persistent Identifiers
 - Object Identifiers
 - DOI Registration Agencies
 - ISBN Registration Agencies
 - Institution Identifiers
 - Project Identifiers
 - Maker Identifiers
 - Subject Identifiers
 - Metadata Standards
 - Content Platforms and Distributors
 - Content Platforms
 - Ebook Aggregators
 - OA Platforms and Repositories
 - Shadow Libraries

Open Metadata for OA Publishing

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





GENERAL

Dashboard

METADATA

Works

Books

Chapters

Publications

Publishers

Imprints

Contributors

Series

Institutions

449 Works
[See all](#)

759 Publications
[See all](#)

2 Series
[See all](#)

449 Books
[See all](#)

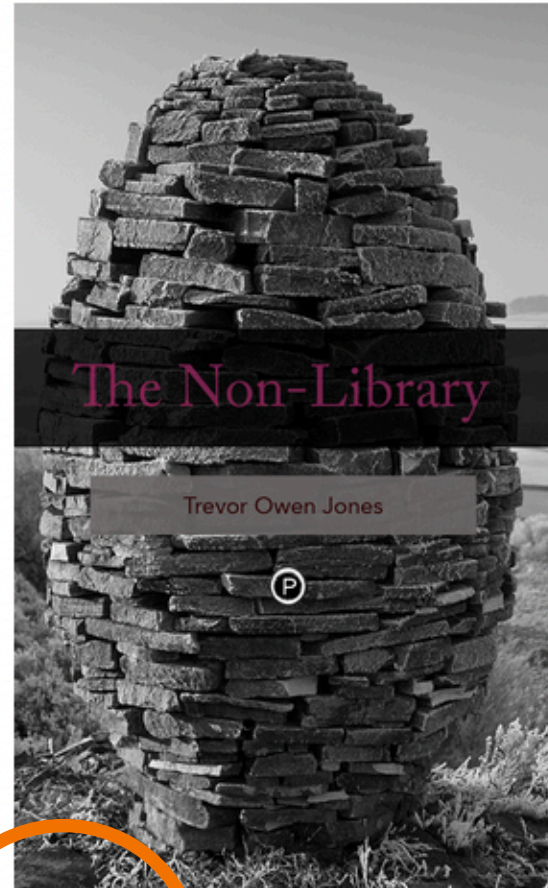
1274 Contributors
[See all](#)

21 Imprints
[See all](#)

0 Chapters
[See all](#)

1 Publishers
[See all](#)

414 Institutions
[See all](#)



Cover URL

https://punctumbooks.com/punctum/wp-content/uploads/2013/08/Jones_Non-Library_Cover_WEB.png

Cover Caption

Cover Caption

DOI

<https://doi.org/>

LCCN

10.21983/P3.0065.1.00

LCCN

OCLC Number

OCLC Number

Internal Reference

0065

Page Count

104

Page Breakdown

Page Breakdown

First Page

First Page

Last Page

Last Page

Image Count

0

Table Count

0

Audio Count

0

Video Count

0

License

<https://creativecommons.org/licenses/by-nc-sa/3.0/>

Copyright Holder

Jones, Trevor Owen

Landing Page

<https://punctumbooks.com/titles/the-non-library/>

Edit publication

Edit

Delete

Publication Type




PDF

ISBN

Relations below are saved automatically upon change.

Locations

Add Location

| | Landing Page | Full Text URL | Platform | Canonical | |
|---|---|---|-----------------|------------------|------------------------|
|  | https://muse.jhu.edu/book/76457 | | Project MUSE | No | Remove |
|  | https://library.oapen.org/handle/20.500.12657/25553 | https://library.oapen.org/bitstream/handle/20.500.12657/25553/1004542.pdf?sequence=1&isAllowed=y | OAPEN | No | Remove |
|  | https://punctumbooks.com/titles/the-non-library/ | https://cloud.punctumbooks.com/s/78FgRcB7qWAPMKH | Other | Yes | Remove |

Edit institution

Delete

Member(s) contributed to: [Lapidari 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 Ethos Theory in Russia and Beyond](#), from: Open Book Publishers
Member(s) contributed to: [Lapidari 3](#), from: punctum books
Member(s) contributed to: [Pedagogies of Disaster](#), from: punctum books
Member(s) contributed to: [Lapidari 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

<https://ror.org/> 016476m91

Country

United Kingdom

Save

Edit contributor

Delete

Contributed to: [Love Don't Need a Reason](#), from: punctum books

Given Name

Matthew J.

Family Name

Jones

Full Name

Matthew J. Jones

ORCID

<https://orcid.org/> 0000-0002-9707-6655

Website

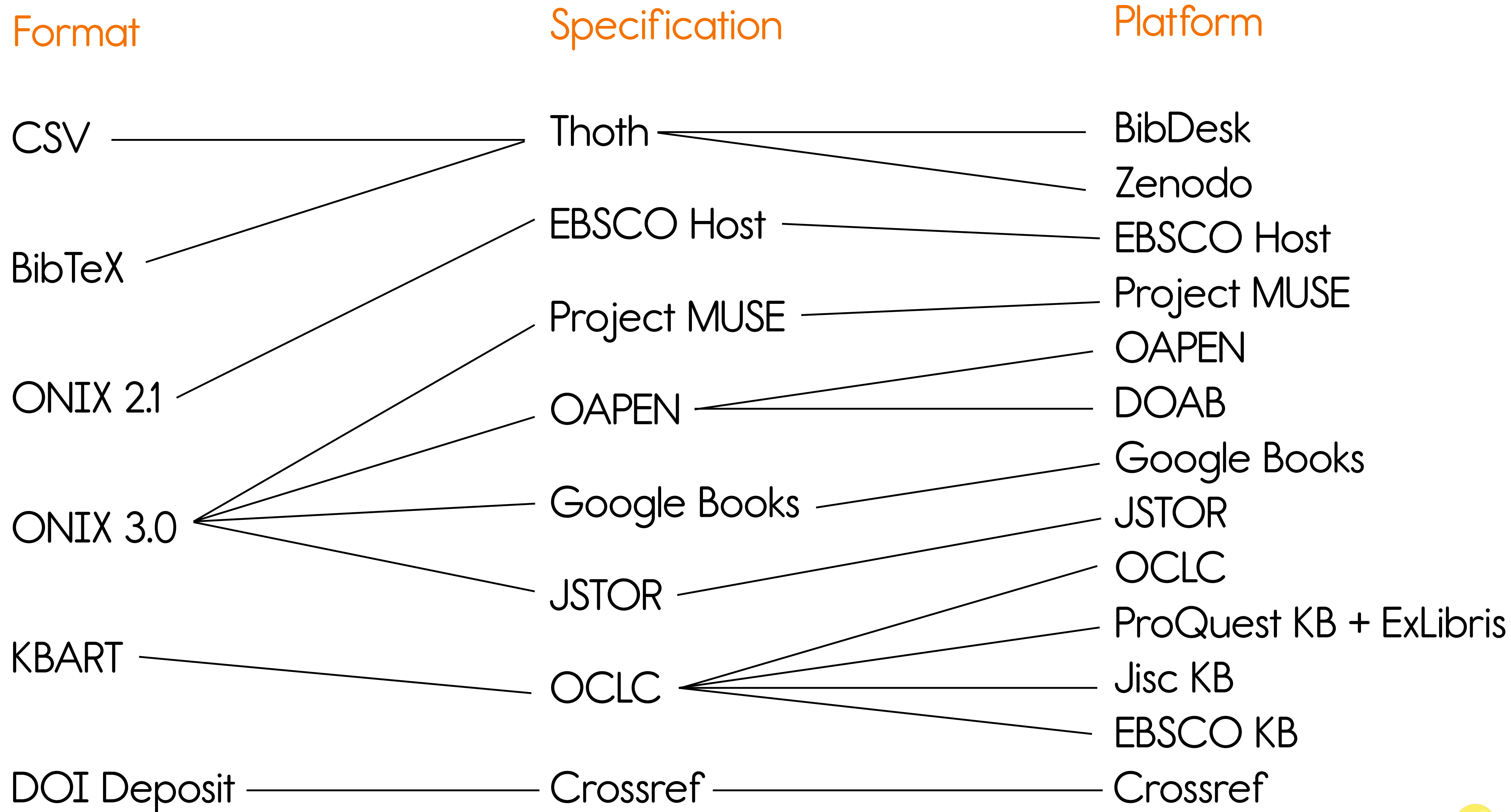
Website

Save

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).
#
# GraphiQL 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 aware of the current GraphQL type schema and live 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 starts
# with a # are ignored.
#
# An example Thoth GraphQL query might look like:
#
#   {
#     works(workType: MONOGRAPH, workStatus: ACTIVE) {
#       fullTitle
#       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):
#
{
  works(order: {field: PUBLICATION_DATE, direction: ASC}) {
    fullTitle
    doi
    publications {
      publicationType
      isbn
    }
    contributions {
      contributionType
      fullName
    }
  }
}

```

```

{
  "data": {
    "works": [
      {
        "fullTitle": "That Greece Might Still Be Free: The Philhellenes in the War of Independence",
        "doi": "https://doi.org/10.11647/OBP.0001",
        "publications": [
          {
            "publicationType": "PAPERBACK",
            "isbn": "978-1-906924-00-3"
          },
          {
            "publicationType": "HARDBACK",
            "isbn": "978-1-906924-01-0"
          },
          {
            "publicationType": "PDF",
            "isbn": "978-1-906924-02-7"
          },
          {
            "publicationType": "HTML",
            "isbn": "978-1-80064-428-1"
          },
          {
            "publicationType": "EPUB",
            "isbn": "978-1-906924-02-7"
          },
          {
            "publicationType": "MOBI",
            "isbn": "978-1-906924-02-7"
          }
        ],
        "contributions": [
          {
            "contributionType": "AUTHOR",
            "fullName": "William St Clair"
          },
          {
            "contributionType": "INTRODUCTION_BY",
            "fullName": "Roderick Beaton"
          }
        ]
      },
      {
        "fullTitle": "The Altering Eye: Contemporary International Cinema",
        "doi": "https://doi.org/10.11647/OBP.0002",
        "publications": [
          {
            "publicationType": "PAPERBACK",
            "isbn": "978-1-906924-03-4"
          },
        ],
      }
    ]
  }
}

```

REQUEST

PATH PARAMETERS

| | |
|------------------------------------|--------------------------------------|
| *specification_id string | onix_3.0::oopen |
| *publisher_id uuid | 9c41b13c-cecc-4f6a-a151-be4682915ef5 |

API Server <https://export.thoth.pub>

Authentication Not Required

FILL EXAMPLE

CLEAR

TRY

[Response Status: \(CORS or Network Issue\)](#)

CLEAR RESPONSE

RESPONSE RESPONSE HEADERS CURL

```
curl -X GET "https://export.thoth.pub/specifications/onix_3.0%3A%3Aoopen/publisher/9c41b13c-cecc-4f6a-a151-be4682915ef5"
```

Copy

RESPONSE

200

OK

Catalogue

Displaying works 1-1 of 1

Previous

Search

Next page



The Non-Library

Trevor Owen Jones

Brooklyn, NY: punctum books, 2014

10.21983/P3.0065.1.00

Read

Metadata



- ONIX 3.0 (Project MUSE)
- ONIX 3.0 (OAPEN/DOAB)
- ONIX 3.0 (JSTOR)
- ONIX 3.0 (Google Books)
- ONIX 2.1 (EBSCO Host)
- CSV
- KBART
- BibTeX

CrossRef DOI deposit
x_3.0::oapen/work/33917b8f-775f-4ee2-a43a-6b5285579f84

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





THOTH

Open Metadata.

Open API
Try it with GraphiQL



Open Standards
Read the Export API docs



Open Metadata
Truly open data



Open Access
Built for OA books



Open Source
Code available on GitHub



Catalogue

Displaying works 1-10 of 882

Previous

Search

Next page



(((
Alessandro De Francesco

Earth, Milky Way: punctum books, 2021
10.53288/0370.1.00

Next Steps for Thoth

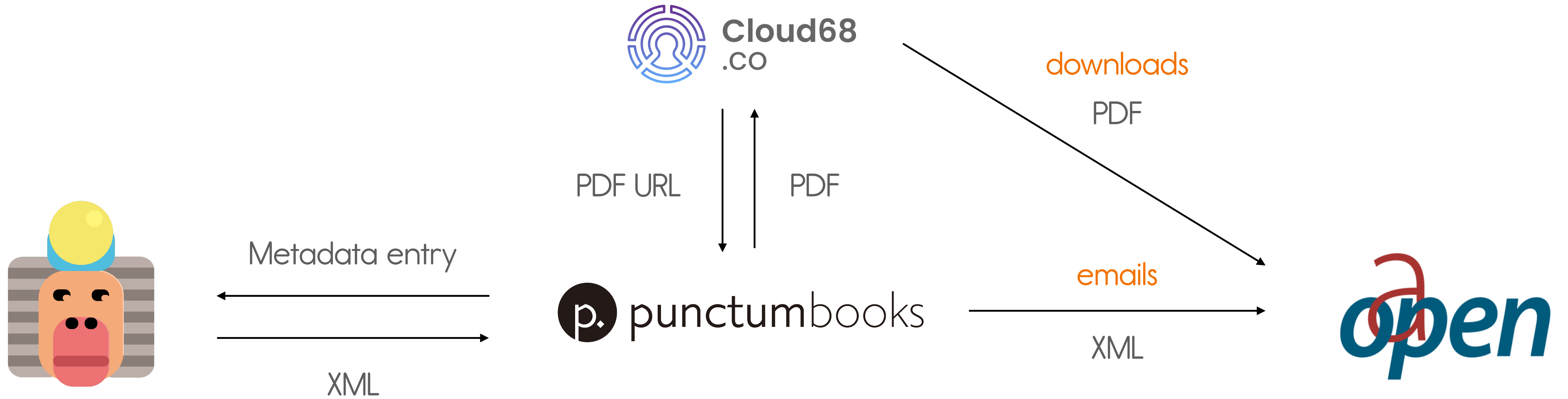
- Quality control via “traffic light” system

Next Steps for Thoth

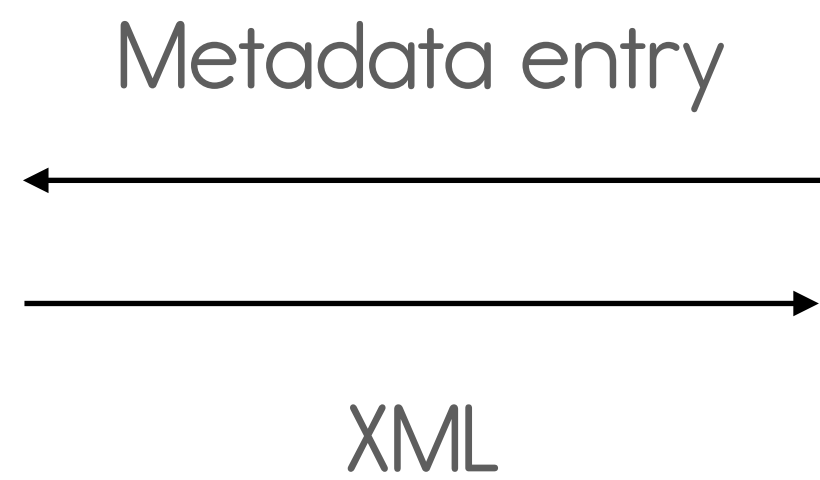
- Quality control via “traffic light” system
- Automated metadata dissemination to platforms



Simplifying Dissemination Pipelines



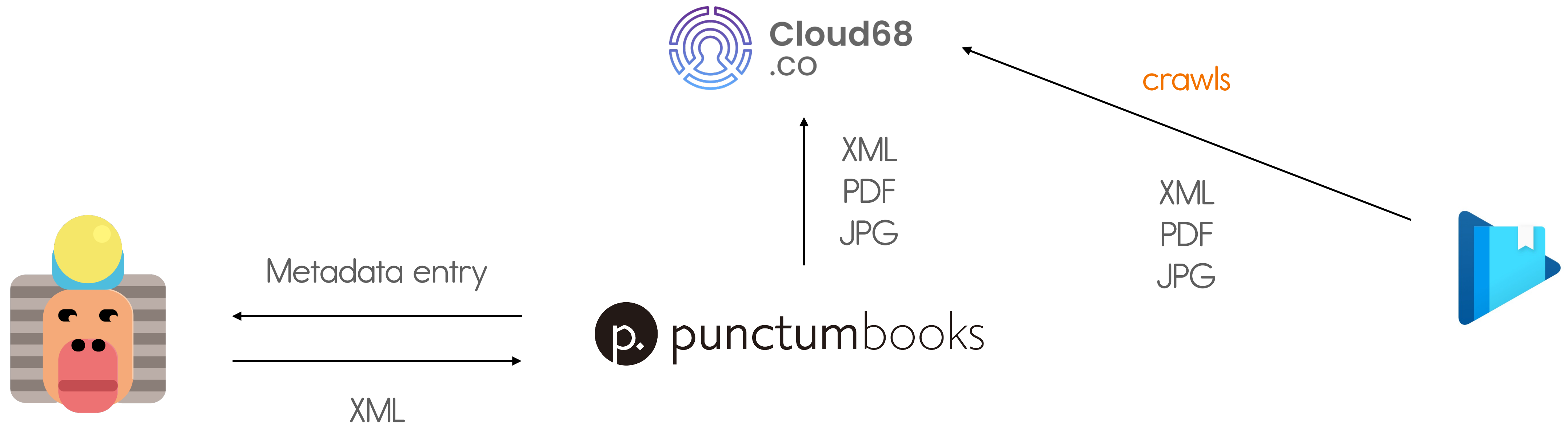
Simplifying Dissemination Pipelines



p. punctumbooks



Simplifying Dissemination Pipelines



Next Steps for Thoth

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



Ingest

Export

Free

- API
- UI
- ONIX

- GraphQL API
- Export API

Plus

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

- 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/>



Get in Touch

Vincent W.J. van Gerven Oei

punctum books

vincent@punctumbooks.com

Rupert Gatti

Open Book Publishers

rupert.gatti@openbookpublishers.com

