

# Open Infrastructures: Open Metadata Management and Dissemination with Thoth

Vincent W.J. van Gerven Oei @ontakragoueke



No open access  
without open infrastructure.



# 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 6 branches 80 tags

Go to file Add file Code

 ja573 Merge branch 'release/v0.11.0'	38807b3 last week	2,653 commits
.github	Update build dependencies	last month
src	Allow setting actix keep alive via cli	2 months ago
thoth-api-server	Bump v0.11.0	last week
thoth-api	Bump v0.11.0	last week
thoth-app-server	Bump v0.11.0	last week
thoth-app	Bump v0.11.0	last week
thoth-client	Bump v0.11.0	last week
thoth-errors	Bump v0.11.0	last week
thoth-export-server	Bump v0.11.0	last week
.dockerignore	Ignore dev files	9 months ago
.env.example	Set docker environment	2 years ago
.gitignore	Ignore dev files	9 months ago
CHANGELOG.md	Update changelog	last week
CODE_OF_CONDUCT.md	Use COPIM code of conduct	2 years ago
Cargo.lock	Bump v0.11.0	last week
Cargo.toml	Bump v0.11.0	last week
Dockerfile	Update build dependencies	last month
Dockerfile.dev	Update build dependencies	last month

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
- 37 stars
- 9 watching
- 7 forks

Report repository

Releases 79

v0.11.0 Latest last week

+ 78 releases

Contributors 5



# Home

[Edit](#) [New page](#)

Rupert Gatti edited this page on Mar 12 · 301 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](mailto: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](#)
- [XML](#)

#### Pages 215

#### 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](#)[Other platforms and repositories](#)

# 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

2099 Works

[See all](#)

828 Publications

[See all](#)

1 Series

[See all](#)

508 Books

[See all](#)

4749 Contributors

[See all](#)

23 Imprints

[See all](#)

1591 Chapters

[See all](#)

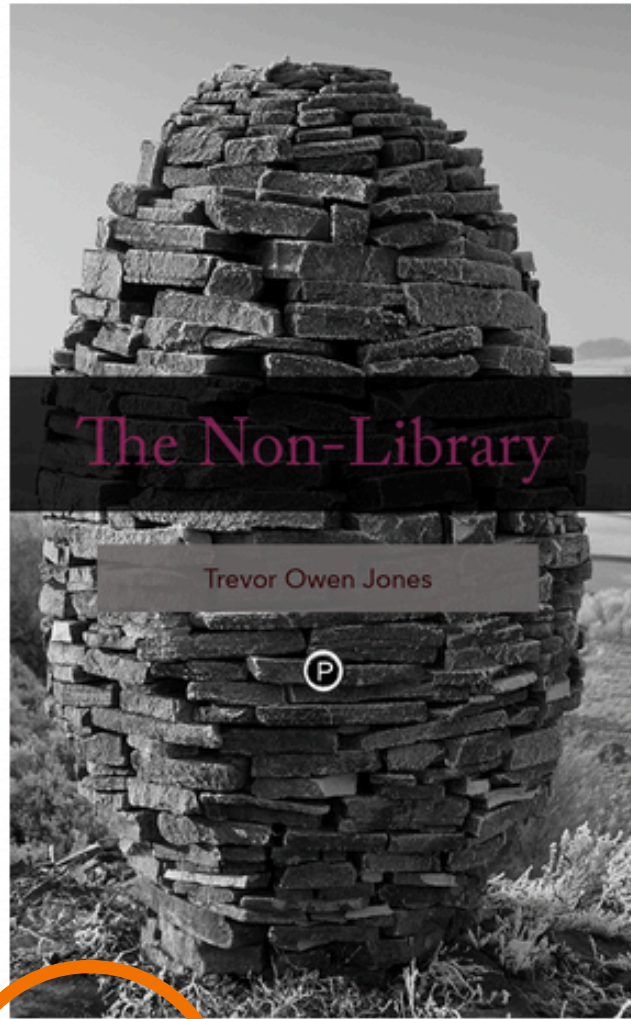
1 Publishers

[See all](#)

103395 Institutions

[See all](#)





**Cover URL**

[https://punctumbooks.com/punctum/wp-content/uploads/2013/08/Jones\\_Non-Library\\_Cover\\_WEB.png](https://punctumbooks.com/punctum/wp-content/uploads/2013/08/Jones_Non-Library_Cover_WEB.png)

**Cover Caption**

Cover Caption

**DOI**

<https://doi.org/>

10.21983/P3.0065.1.00

**LCCN**

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	
 <a href="https://muse.jhu.edu/book/76457">https://muse.jhu.edu/book/76457</a>		Project MUSE	No	<a href="#">Remove</a>
 <a href="https://library.oapen.org/handle/20.500.12657/25553">https://library.oapen.org/handle/20.500.12657/25553</a>	<a href="https://library.oapen.org/bitstream/handle/20.500.12657/25553/1004542.pdf?sequence=1&amp;isAllowed=y">https://library.oapen.org/bitstream/handle/20.500.12657/25553/1004542.pdf?sequence=1&amp;isAllowed=y</a>	OAPEN	No	<a href="#">Remove</a>
 <a href="https://punctumbooks.com/titles/the-non-library/">https://punctumbooks.com/titles/the-non-library/</a>	<a href="https://cloud.punctumbooks.com/s/78FgRcB7qWAPMKH">https://cloud.punctumbooks.com/s/78FgRcB7qWAPMKH</a>	Other	Yes	<a href="#">Remove</a>



## 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 Etnos 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

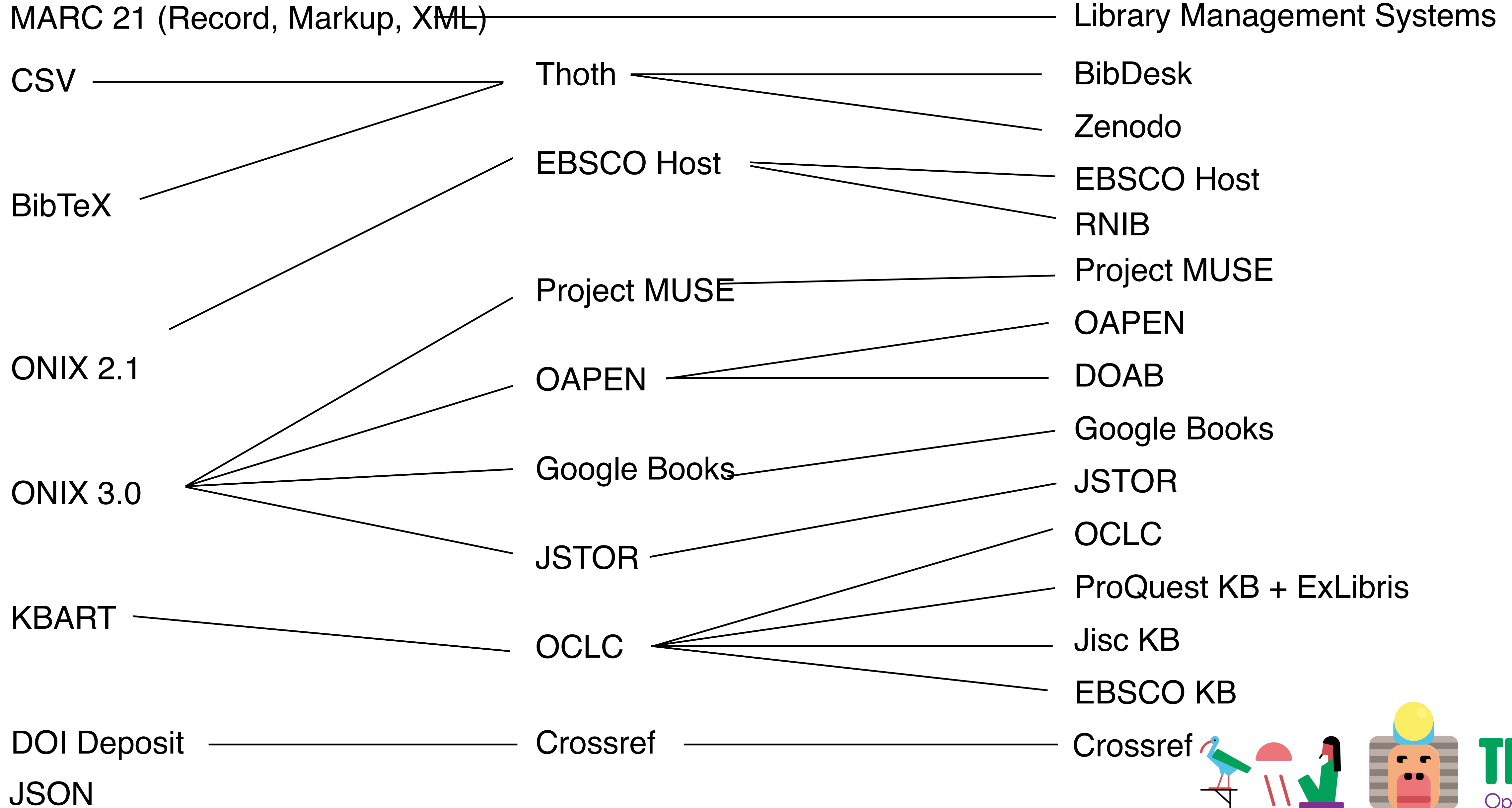
- Thoth: Open Metadata Management & Dissemination
  - Open source
  - Tailor-made for Open Access book metadata
  - Multiple output formats and specifications for 10+ platforms



## Format

## Specification

## Platform





```

# 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
# typeaheads 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

<b>*specification_id</b> string	<input type="text" value="onix_3.0::oopen"/>
<b>*publisher_id</b> uuid	<input type="text" value="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

non-library

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

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



# 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
  - Integration with third-party platforms





## Collective Catalogue

The collective catalogue displays books that are in the Thoth system, which is an open source metadata management service used by many of the publishers supported by the Open Book Collective.



SEARCH

[Search the  
catalogue](#)

INITIATIVE

Select an Initiative

SORT

Date Newest

## Latest Titles



**AFRICAN MINDS**

### Transformative innovation in times...

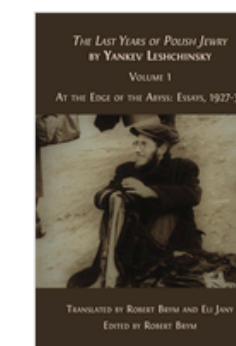
Erika Kraemer-Mbula,  
Rebecca Hanlin, Rob Byrne,  
Chux Daniels, and Ann Kingiri



**OPEN BOOK PUBLISHERS**

### Introduction to Systems Biology: W...

Thomas Sauter, and Marco  
Albrecht



**OPEN BOOK PUBLISHERS**

### The Last Years of Polish Jewry: Vo...

Yankev Leshchinsky, Robert  
Brym, Robert Brym, and Eli  
Jany



# Catalogue

 0 Filters

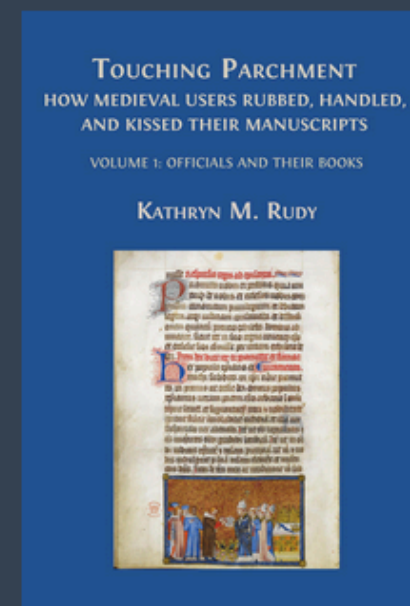
 Search...

Showing 1 to 20 of 344 results

20 per page ▾ Sort By: Publication date desc. ▾

## Category

- African Studies 5
- American and Latin American Studies 9
- Anthropology, Archaeology and Religion 38
- Asian Studies 7
- Biography 19
- Cambridge Semitic Languages and Cultures 19
- Classics 15
- Classics: Greek Textbooks 1
- Classics: Latin Textbooks 10
- Digital Humanities 11
- Economics, Politics and Sociology 57
- Education 20
- Environmental Studies 15
- European Studies 82
- European Studies: Eastern European

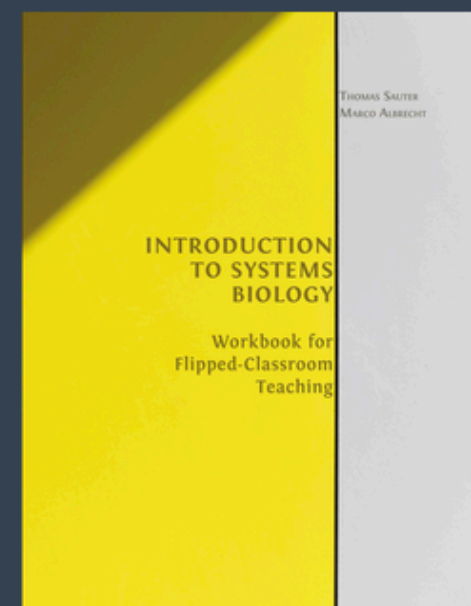


HISTORY OF THE BOOK

## TOUCHING PARCHMENT: HOW MEDIEVAL USERS RUBBED, HANDLED, AND KISSED THEIR MANUSCRIPTS: VOLUME 1: OFFICIALS AND THEIR BOOKS

Kathryn M. Rudy

The Medieval book, both religious and secular, was regarded as a most precious item. The traces of its use through touching and handling during different rituals such as oath-taking, is the subject of Kathryn Rudy's research in Touching Parchment.



TEXTBOOKS AND LEARNING GUIDES • SCIENCE

## INTRODUCTION TO SYSTEMS BIOLOGY: WORKBOOK FOR FLIPPED-CLASSROOM TEACHING

Thomas Sauter · Marco Albrecht

This book is an introduction to the language of systems biology, which is spoken among many disciplines, from biology to engineering. Authors Thomas Sauter and Marco Albrecht draw on a multidisciplinary background and evidence-based learning to facilitate the understanding of biochemical networks, metabolic modeling and system dynamics.



Search blog posts

Search ...

Search

Categories

About SO

Aggregation

Altmetrics

Announcements

Author Interviews

Board Interviews

Books and Book Chapters

Collections

Early Career

Events

Guest Blog

Impact Factor

Licenses

Open Science Interviews

ORCID

Peer Review

Profiles

Publisher

Research

Services and Features

In: [Announcements](#), [Services and Features](#) ▶



## ScienceOpen and Thoth join forces on book metadata

April 13, 2023 \* Author: Nina Tscheke

To increase **the discoverability and visibility for open books with richer and more interoperable metadata**, ScienceOpen is happy to announce **a new collaboration with Thoth**, an open metadata management and dissemination platform. With this collaboration, **open access book publishers and the entire book publishing ecosystem** will benefit from a more **interconnected digital publishing environment**.

When we first started working with books and aggregating publicly available Crossref records via book DOIs to the ScienceOpen citation index, we quickly learned that the potential of Crossref records as actual metadata source was hardly recognized making for rather basic publicly available records. Rich metadata was available mostly in ONIX, but while being the default format for distribution channels for book sellers and libraries, there tend to be data gaps in the ONIX information required for best practice usage in digital environments—speaking here esp. about persistent identifiers which are essential to keep records unified, connected and discoverable throughout the publication’s lifecycle.

**Thanks to this new collaboration with Thoth, publishers will now be able to create a record on BookMetaHub, send it to Thoth for further enhancement, and then gain access to additional output formats for future export.** We are also working on sending data created on Thoth to the BookMetaHub, so stay tuned!

In response to the **current data gap** in the book publishing landscape **between rich publication records used by book vendors within their sales channels and commonly available data**, we are **putting together two innovative and complementing digital infrastructures** that will support publishers to share their works even further throughout the entire ecosystem—by offering simple DIY-tools for the creation and maintenance of metadata that



Tags

- [Academic Publishing](#)
- [Aggregation altmetrics Book Publishing Books Climate Change Collection Collections](#)
- [conference Context Digital Publishing Discovery education Hosting Impact Factor Indexing Journal Journal Hosting Journals metadata Monthly Digest Open Access Open data Open Peer Review Open Science ORCID Peer Review Post-Publication Preprints Psychology Publishers Publishing Publishing Platform Publishing Solutions Research Research Discovery Review Scholarly Publishing Science Communication ScienceOpen SDGs Sustainable Development Goals Technology UCL Press University Presses](#)

Archives

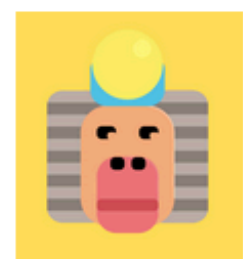
- [April 2023 \(7\)](#)
- [March 2023 \(10\)](#)
- [February 2023 \(7\)](#)
- [January 2023 \(7\)](#)
- [December 2022 \(19\)](#)
- [November 2022 \(11\)](#)
- [October 2022 \(6\)](#)
- [September 2022 \(4\)](#)
- [August 2022 \(10\)](#)

# 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
  - Integration with third-party platforms
  - Thoth Archiving Network







# Thoth Archiving Network

- Share
- Favorite
- RSS
- Play All

Based on the recommendations set out in the [WP5 Scoping Report](#), the development of open metadata management and dissemination system [Thoth](#) has been going ahead, and we're now at the point that two scholar-led open-access publishers, [punctum books](#) and [Open Book Publishers](#), are using Thoth for their day-to-day metadata management. The next step in our process will be to create a variety of metadata output formats, such as different flavors of ONIX, MARCXML, and BibTeX, which will allow users of Th...

[More...](#)

ABOUT

COLLECTION

661 RESULTS

SORT BY VIEWS · TITLE · DATE PUBLISHED · CREATOR

Search this Collection

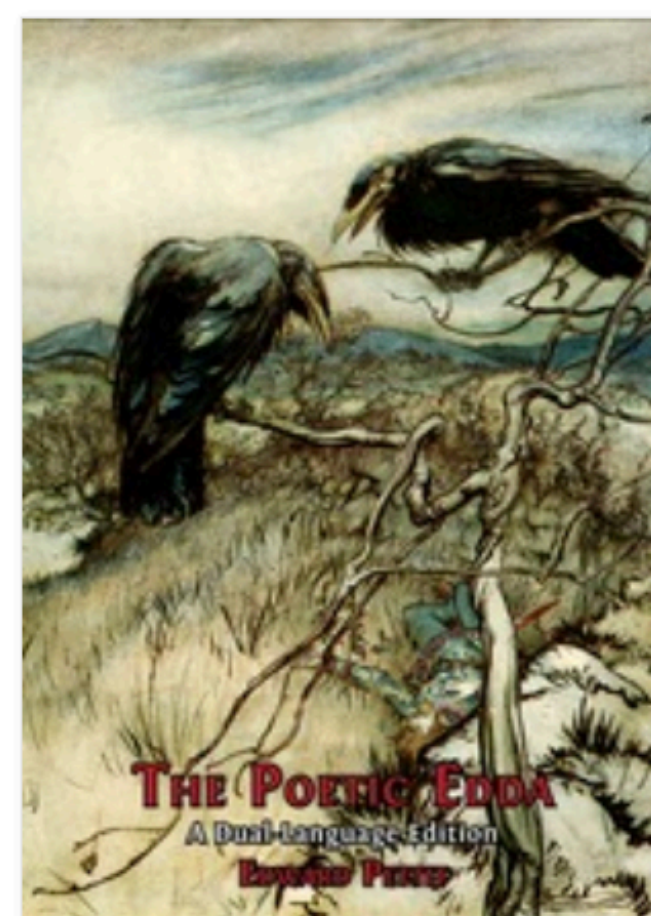
- Metadata
- Text contents

Media Type

- [Slider]
- [Slider]
- [Slider]
- [Slider]
- [Slider]

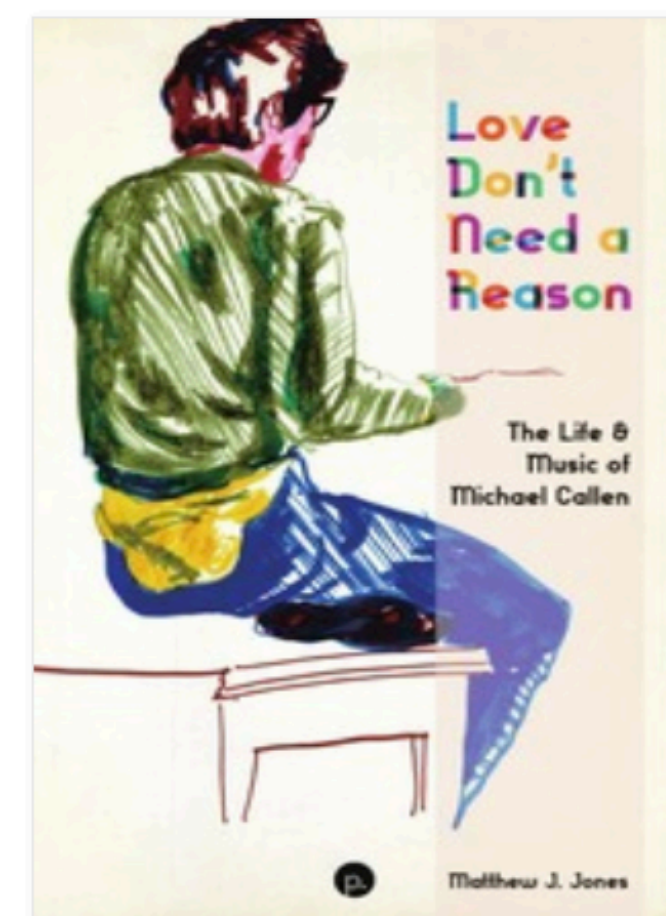
Year

- [Slider]



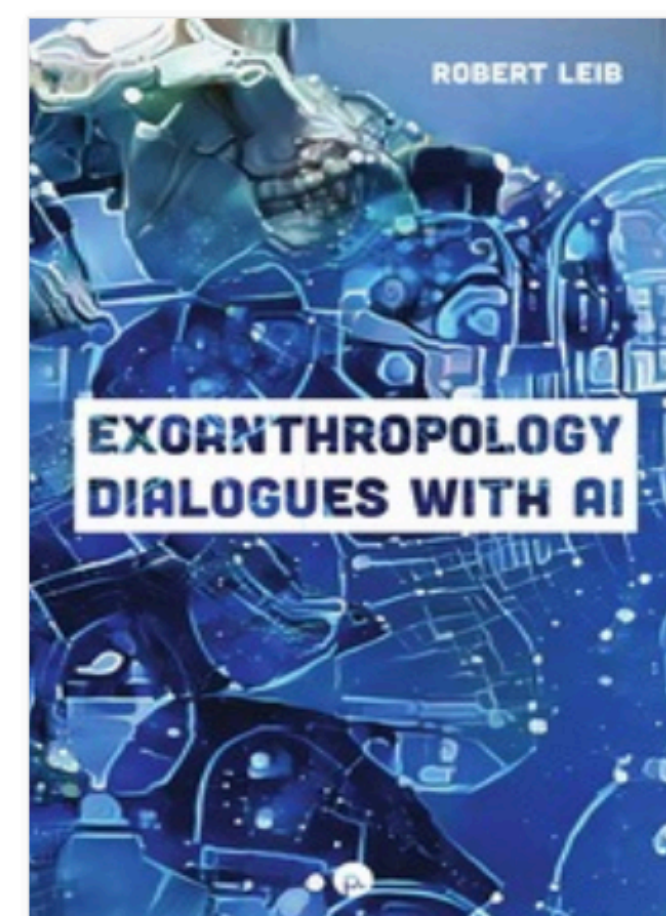
The Poetic Edda: A Dual-Language Edition  
by Edward Pettit

📖 21 👁️ 0 ⭐ 0 💬 0



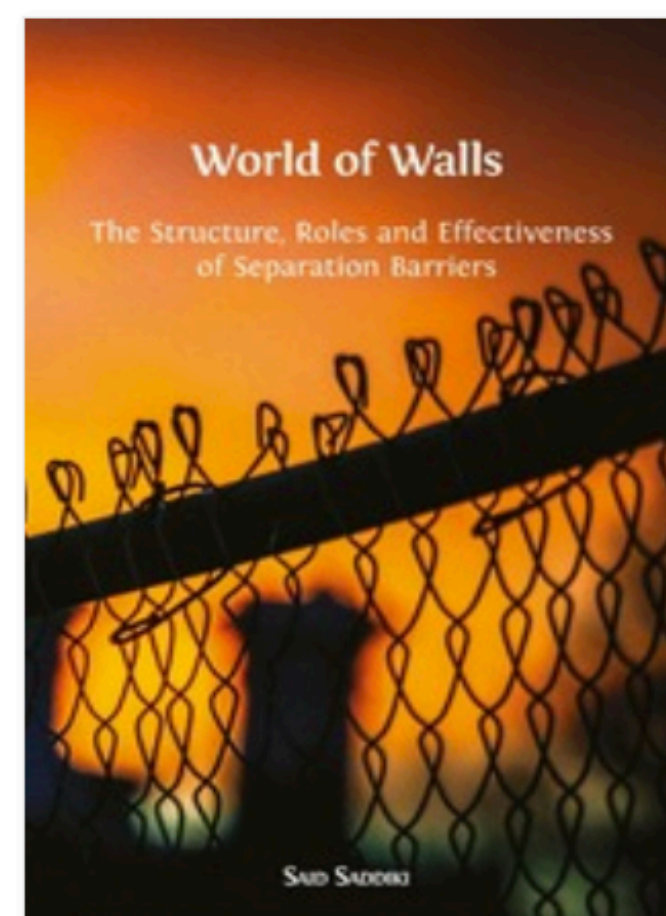
Love Don't Need a Reason: The Life & Music  
by Matthew J. Jones

📖 78 👁️ 1 ⭐ 0 💬 0



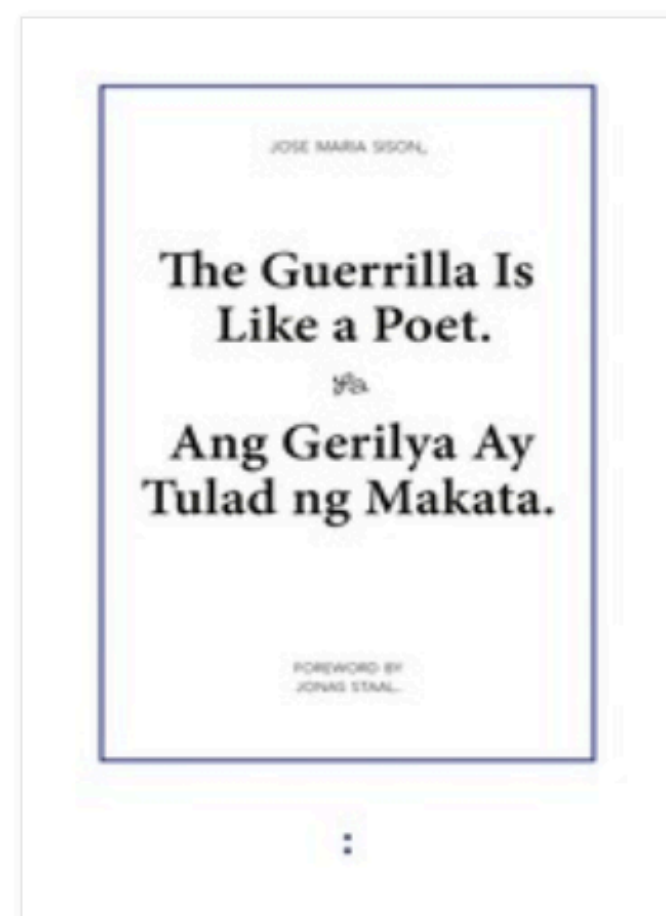
Exoanthropology: Dialogues with AI  
by Robert Leib

📖 47 👁️ 2 ⭐ 0 💬 0



World of Walls: The Structure, Roles and  
by Said Saddiki

📖 51 👁️ 1 ⭐ 0 💬 0



The Guerrilla Is Like a Poet – Ang Gerilya Ay  
by Jose Maria Sison

📖 76 👁️ 2 ⭐ 0 💬 0





# 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



# Next Steps for Thoth

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





# Next Steps for Thoth

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

# 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](mailto:vincent@punctumbooks.com)

Rupert Gatti

Open Book Publishers

[rupert.gatti@openbookpublishers.com](mailto:rupert.gatti@openbookpublishers.com)

