

# Open archiving of open access scholarly content

Dr Gareth Cole

- Open Research Development and Discovery Lead. Loughborough University, UK
- Co-Work Package Lead (on Open Book Futures) for Archiving and Preservation



# Today's session

- Background of wider project
- Thoth Open Archiving Network
- Our plans/thoughts
- Discussion/critiques/push-back



# Open Book Futures

- Research England and Arcadia funded project: <https://copim.ac.uk/> and <https://copim.pubpub.org/open-book-futures-project>
- Three years until April 2026
- “Open Book Futures aims to initiate a step-change in the ambition, scope and impact of community-led Open Access book publishing.”
- Outputs include:
  - [Thoth open metadata management and dissemination system](#)
  - [Opening the Future: A Funding Model for OA Books](#)
  - [Open Book Collective](#)



# Thoth

- Thoth
  - [Open] Metadata management for books and chapters.
  - Automated distribution to various platforms.
  - Archiving and preservation of scholarly works (TOAN).
  - DOI registration and management.
  - Open APIs for seamless integration with other platforms.



# Thoth Open Archiving Network

- Thoth Open Archiving Network (TOAN)
  - An open archiving network partnering with open repositories online, such as the Internet Archive and Zenodo.
  - Open monographs are archived in an openly accessible way, including book and metadata files.
  - Automated for participating presses, with monitoring via checksums in active development.
- Rationale:
  - Supporting small and scholar-led presses with little or no archiving & preservation in place.
  - Working towards a completely open, and auditable, archiving network solution.
  - Part of a larger set of open workflows and infrastructure within Thoth.



# Open Archiving? Must have?

- We identified the following features as essential to an open archiving solution:
  - Openly accessible content
  - Openly accessible metadata
- Openly verifiable processes:
  - Publishing [checksums](#) to allow verification of content integrity
  - Transparent version control (for both content and metadata)
  - Clear mechanisms for checking and maintaining the content
- Reliability and long-term sustainability
- Multiple geographically-redundant copies

# Open Archiving? Highly desirable?

And these additional features as highly desirable:

- Support for retrieving and archiving associated content:
  - “Additional materials” provided to supplement the main content
  - Web pages represented by URLs within the main content
- Clearly-stated policies around removal of content
- Collation of usage statistics
- Independence from private or government-controlled entities

Please tell us what you think?

**Be Honest!**

# Further reading

- Blog Post (Putting the 'Open' in 'Thoth Open Archiving Network'):  
<https://doi.org/10.21428/785a6451.76e96572> or  
<https://copim.pubpub.org/pub/thoth-open-archiving-network/release/1?readingCollection=78821e91>
- Project documentation site: <https://copim.pubpub.org/>