Open Archiving of the World’s Biomedical Literature  
A Geographic Analysis of PubMed Central Journals and Applications

Kathryn Funk, Program Manager, PMC // Amelia Llorens and Cecelia Vetter, NLM Associate Fellows // US National Library of Medicine, National Institutes of Health

1. Introduction

The Collection and Preservation Policy of the US National Library of Medicine (NLM), approved by the NLM Board of Regents on 17 September 2019, states that the NLM Collection, “represents the broadest and most comprehensive collection of biomedical literature.”

In keeping with NLM’s mandate to collect and preserve the biomedical literature, PMC focuses on selecting scientifically rigorous journals for the archive. NLM applies more stringent criteria to selecting content that is archived and indexed in the portal and does not include for journals selected for the general NLM collection.

Results

In applying visualizations and subsequent analysis to these datasets, the initial question NLM wanted to explore in the context of how PMC represents the diversity of the world’s biomedical literature was:

1. Are the countries or regions that are underrepresented in PMC due to low application volume?
2. Can we identify potential languages?
3. Are journals from certain countries or regions less successful than others in the PMC journal selection process?

In analysing visualizations and subsequent analysis to these datasets, the initial question NLM wanted to explore in the context of how PMC represents the diversity of the world’s biomedical literature was:

1. Are the countries or regions that are underrepresented in PMC due to low application volume?
2. Can we identify potential languages?
3. Are journals from certain countries or regions less successful than others in the PMC journal selection process?

In applying visualizations and subsequent analysis to these datasets, the initial question NLM wanted to explore in the context of how PMC represents the diversity of the world’s biomedical literature was:

1. Are the countries or regions that are underrepresented in PMC due to low application volume?
2. Can we identify potential languages?
3. Are journals from certain countries or regions less successful than others in the PMC journal selection process?

In applying visualizations and subsequent analysis to these datasets, the initial question NLM wanted to explore in the context of how PMC represents the diversity of the world’s biomedical literature was:

1. Are the countries or regions that are underrepresented in PMC due to low application volume?
2. Can we identify potential languages?
3. Are journals from certain countries or regions less successful than others in the PMC journal selection process?

In applying visualizations and subsequent analysis to these datasets, the initial question NLM wanted to explore in the context of how PMC represents the diversity of the world’s biomedical literature was:

1. Are the countries or regions that are underrepresented in PMC due to low application volume?
2. Can we identify potential languages?
3. Are journals from certain countries or regions less successful than others in the PMC journal selection process?

In applying visualizations and subsequent analysis to these datasets, the initial question NLM wanted to explore in the context of how PMC represents the diversity of the world’s biomedical literature was:

1. Are the countries or regions that are underrepresented in PMC due to low application volume?
2. Can we identify potential languages?
3. Are journals from certain countries or regions less successful than others in the PMC journal selection process?

In applying visualizations and subsequent analysis to these datasets, the initial question NLM wanted to explore in the context of how PMC represents the diversity of the world’s biomedical literature was:

1. Are the countries or regions that are underrepresented in PMC due to low application volume?
2. Can we identify potential languages?
3. Are journals from certain countries or regions less successful than others in the PMC journal selection process?

In applying visualizations and subsequent analysis to these datasets, the initial question NLM wanted to explore in the context of how PMC represents the diversity of the world’s biomedical literature was:

1. Are the countries or regions that are underrepresented in PMC due to low application volume?
2. Can we identify potential languages?
3. Are journals from certain countries or regions less successful than others in the PMC journal selection process?

In applying visualizations and subsequent analysis to these datasets, the initial question NLM wanted to explore in the context of how PMC represents the diversity of the world’s biomedical literature was:

1. Are the countries or regions that are underrepresented in PMC due to low application volume?
2. Can we identify potential languages?
3. Are journals from certain countries or regions less successful than others in the PMC journal selection process?

In applying visualizations and subsequent analysis to these datasets, the initial question NLM wanted to explore in the context of how PMC represents the diversity of the world’s biomedical literature was:

1. Are the countries or regions that are underrepresented in PMC due to low application volume?
2. Can we identify potential languages?
3. Are journals from certain countries or regions less successful than others in the PMC journal selection process?
InnOAccess
Innovative Open Access in Small Sciences (InnOAccess)

Academia – especially in the field of internet research – thrives on a dynamic and diverse publication landscape in which researchers can both use and publish results and findings free of major hurdles. The "InnOAccess" project (2016-2019 – 860-2026) takes this premise as a starting point and aims to substantially improve the situation of small science journals. Using Internet Policy Review as an example, the project will further develop editorial infrastructures, design open access financing models for small science contexts, and build new networks and cooperative models. The results will be published as white papers and discussed in workshops.

INTERNET POLICY REVIEW
was founded in 2012 as a scholar-led open access journal. It does not charge publication fees, and features an accelerated and open peer review procedure. Internet Policy Review tracks legal changes as well as developments in internet regulation and norms that have a long-lasting impact on European societies.

ISSN: 2197-6773
ur: policyreview.net

INNOVATIVE PUBLICATION INFRASTRUCTURES
Based on preparatory work开展 now via OpenSesame funding and ongoing cooperation with, e.g., OpenAIRE and OAI, new technical features will be developed for the Internet, documented, and made available to the wider publishing community for subsequent use. These include various Open Access plug-ins, the implementation of the identity service DOISID, subscriptions to OA-ready small science journals (via DOISID module). Further elaborations on newly introduced alternative metrics as well as accessibility improvements.

SUSTAINABLE FINANCING MODELS
Open Access journals need stable, transparent, and sustainable financing to-wise quality. Based on a systematic evaluation of existing (OA) financing models and the requirements of new-IR, scholar-led small science journals, the project will design and test modular financing models. This budget tool will foster continuous journals and promote a new generation of OA journals. In addition, thematic publication cooperations will create further financing incentives.

SUSTAINABLE NETWORKS
In order to position academic journals thematically and establish them institutionally, robust peer community networks need to be initiated and maintained. In the project, existing cooperations between the journal and research societies (e.g., ASIS, DOISID) will be intensified and new cooperations will be established. For example, boot paper awards and special issues. Experiments with community building concepts (e.g., "open abstracts") will also be carried out. Along the way, Internet Policy Review will be communicating and multiplying successful collaborations, so as to become a hub for inspiration.

OUTPUTS & TRANSFER
- WHITE PAPERS on 1) the use of technical solutions for the publication workflow, and sustainable, financing models for small science journals.
- WORKSHOPS on 1) technical solutions, 2) OA financing models, 3) small science journals, and 4) community.
- PUBLIC EVENTS to discuss the results with other journals as well as interested members of the public.

CONTACT
For updates on outputs and events, scan the QR code or drop your email at innoadc.nig.de anzeigen.

MARCEL WERESKINSKI
Open Access Officer
wereskinski.de
https://doi.org/10.6084/M9.FOR1.798.5
Dryad is a community-led, standards-based, open-source research data curation and publishing platform. For the last ten years, Dryad has supported the research community with publishing close to 30,000 datasets associated with articles.

To best align with the researcher, publisher, and funder communities, Dryad has re-launched onto a new platform. The scope of Dryad is no longer limited to datasets supporting articles, and Dryad is focused on making curated data publishing as seamless as possible and a common practice for researchers.

Key Features & Benefits

- **Standards Based**
  Persistent identifiers and standardized vocabulary for: people, organizations, funders, journals

- **Best Practices for Data Stewardship**
  Dryad relies on a CoreTrustSeal certified preservation repository
  Dryad supports best practices in research data usage metrics and data citation

- **Data Curation**
  Dryad has a team of professional curators who check every submission to ensure the validity of files and metadata. Curators enrich research data to help make it FAIR

- **Seamless Publishing**
  We are building direct integrations with publishing platforms and researcher tools to enable uploads to Dryad within various researcher workflows

Dryad Member Community

Support your researchers and join the Dryad community: https://datadryad.org/join
punctum books & UCSB Library are working together to ensure that research funded by the public is openly available to a global readership, that the tools for Open Access publishing are open source and community built, and that the future of open books in the Humanities and Social Sciences is economically sustainable, without financial barriers to authors.

Our Supporting Library Membership Program helps to subsidize:
- the publication of high-quality open-access books in the Humanities, Social Sciences, Fine Arts, and Architecture & Design, with no publication fees imposed upon authors; and
- the collaborative development and sustainable maintenance of open-source and community-owned infrastructure for open-access books in the Humanities and Social Sciences.

What membership enables:
- MARC/ONIX records
- COUNTER-compliant statistics
- Long-term preservation of all e-books and associated metadata
- Cross-referenced metadata (DOIs for all books)
- Annotation through Hypothes.is
- Voting position on the Library Advisory Board
- And much more... ask us!

Investing in the Supporting Library Membership Program supports the operations of a press that shares values with the research library community and wants to work in direct relationships with librarians and other knowledge workers in order to ensure its catalog is fully legible and integrated into universal research systems valued by librarians and other research-intensive organizations, including repositories and digital learning environments, while also ensuring its books is available to a global readership without financial barriers to authors or readers worldwide.

In addition, by supporting punctum, libraries also invest in a more diverse, scholar-led, community-owned, and not-for-profit publishing ecosystem that we believe is crucial for the cultivation of more creative and diverse modes and forms of scholarship and their open dissemination and preservation as public knowledge. Libraries can also be assured that they are investing in a press that has established its role as a transformative actor in the scholarly communications landscape, as well as its economic and operational sustainability and integrity.

Finally, the Supporting Library Membership Program adopts a consortial funding model whereby small annual contributions from libraries form a robust revenue pool that enables an extremely cost-effective method for funding Open Access in which no single institution bears a disproportionate cost, and each institution determines for itself what is an appropriate level of support.

punctum books is a scholar-led open-access publisher proudly partnering with:
An independent report commissioned by

UK Research and Innovation

The Association of Learned and Professional Society Publishers

Alicia Wise and Lorraine Estelle, Information Power Ltd

Final Project Report and Transformative Agreement Toolkit

Transformative Agreement Toolkit

Principles for a model OA transformative agreement

1. Authors retain copyright and control over the use of their work.
2. Authors may retain the right to make their own content available in OA form.
3. Open access publication may be subject to a fee.
4. OA articles must be citable and indexable.
5. OA content must be interoperable and re-usable.

Tool Kit

Data template

Overview

A brief summary of what the tool kit does and how it can be used.

The tool kit can also be used as an educational resource.

Additionally, the tool kit includes a guide on how to submit a proposal.
In 2018, MIT Press identified the need for a "comprehensive and critical analysis of OS publishing systems in use" that "could prove to be durable alternatives to complex and costly proprietary services." The Andrew W. Mellon Foundation funded the research and development of this report.

Part I: High-level analysis of the landscape and the projects within it.
Part II: Catalogue of the projects themselves.

Key themes: Sustainability, scale, collaboration, coordination, integration.

Examines software developed in such a way that its source code is open and available online, and explicitly licensed as such.

Part of an ethical imperative to keep academic work open and in free circulation.

With the support of: MELLON FOUNDATION

Get the full report at: mindthegap.pubpub.org
The language we use to describe Open Access publishing can activate bias against Open Access.

The Status Quo Bias and the Uptake of Open Access Publishing

INTRO
Information provided online about open access is for the most part prohibitively complex and introduces contradictory interpretations that increase the cognitive load of readers; high cognitive load activates a phenomenon known as the status quo bias. The only reliable method of counteracting this bias in order to bolster the uptake of open access is to re-frame the language commonly employed in association with open access and to minimize the tiers of decisions expected of authors, which create a barrier rather than a gateway to open access engagement.

METHODS
1. Scraped text from publisher, library, and advocacy org websites.
2. Text analysis for complexity using Textstat.
3. Definitions coding for
   1. Parameters of Open Access mentioned.
   2. Whether “choice” language is used.
   3. Number of options for cost recovery.

RESULTS
- Over half (52%) of sources scored a graduate or above reading level for their OA webpages; all but one source scored at college level or higher. Higher reading level increases cognitive load for readers.
- Many library and advocacy group definitions include “squishy” language such as without most permission barriers.
- More parameters also increase cognitive load. The majority (55%) of sources included 8+ parameters for definitions of OA.
- Nearly three-quarters (72%) of sources described OA as a “choice” or “option”; more choices lead to high cognitive load and bias towards the status quo.

This poster is based on our publication at DOI: 10.5210/fm.v24i7.10089
A Sustainable Scholar-led Model for Open Access without Publication Fees

Open Library of Humanities

> 200 Libraries Financially Supporting the OLH

OLH: A Scholar-led, gold open-access publisher with no author-facing charges. With funding from the Andrew W. Mellon Foundation, the platform covers its costs by payments from an international library consortium, rather than any kind of author fee.

27 Journals on or Supported by the Platform

small sums

Library

x100s

$
THE OUTPUTS

Guidance
Outputs that represent our thinking about metadata and our suggestions on how it should be characterized, managed, thought about and valued.

Metadata principles
How we create ‘richer metadata,’ taking into account how context affects this equation.

For Metadata to support the community, it should be:

<table>
<thead>
<tr>
<th>COMPATIBLE</th>
<th>COMPLETE</th>
</tr>
</thead>
<tbody>
<tr>
<td>It should provide a guide to content, for both machines and people.</td>
<td>It should reflect the context, components and relationships as published.</td>
</tr>
<tr>
<td>So, metadata must be as open, interoperable, portable, machine-readable, and human understandable as possible within its context.</td>
<td>So, metadata must be as complete and comprehensive as possible within its context.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>CREDIBLE</th>
<th>CURATED</th>
</tr>
</thead>
<tbody>
<tr>
<td>It should maintain content discoverability and longevity.</td>
<td>It should reflect updates and new elements.</td>
</tr>
<tr>
<td>So, metadata must be of higher provenance, trustworthiness and assurance within its context.</td>
<td>So, metadata must be maintained over time within its context.</td>
</tr>
</tbody>
</table>

Metadata persona
Considering the key players that will follow the principles to ensure the benefits.

Role-based personas are a useful construct

<table>
<thead>
<tr>
<th>CREATORS</th>
<th>CURATORS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Individuals that provide descriptive information (metadata) about research and scholarly outputs.</td>
<td>Individuals that classify, normalize and standardize this descriptive information to increase its value as a resource.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>CUSTODIANS</th>
<th>CONSUMERS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Individuals that store and maintain this descriptive information and make it available for consumers.</td>
<td>Individuals that knowingly or unknowingly use the descriptive information to find, choose, connect and cite research and scholarly outputs.</td>
</tr>
</tbody>
</table>

Understanding
Outputs that provide insights into how we have classified and made sense of what currently exists and the comparison of these works to each other.

Existing best practices
What are the key metadata practices in use today?

Use survey
What do researchers and others think about metadata and how we can add value?

Literature review
How do academics think about metadata and characterize its effectiveness?

Glossary tool
How can we provide support for commonly used terms?

Schemas index
What dictionaries, metadata models, context standards, and value lists are commonly used?

Evaluation
What tools would enable us to assess effectiveness?

About the Project
Established in 2017, Metadata 2020 is a collaboration that advocates richer, connected, and reusable, open metadata for all research outputs, which will advance scholarly outputs for the benefit of society.

Our volunteer project team and sub-groups have built several new resources, reports and publications.

www.metadata2020.org

‘Most people wouldn’t think:’

‘Well, if we can fix this metadata we can find a cure for cancer.’

Nobody is asking:

‘What is the cost to society?’

If we can find a way to connect those dots, that would be huge.’
PLOS supports transparency in the peer review process to surface deeper context from more perspectives, and increase trust in science. We have combined manual and automated processes, including the development of new software, to enable us to publish peer review histories at “PLOS” scale.
Creating an Open, Inclusive, and Sustainable Scholarly Publishing Landscape

About the LPC
- An independent, community-led membership association
- Two purposes: to support and advance a range of scholarly publishing practices and to further the interests of libraries involved in publishing activities
- Founded in 2017
- Current membership: 130 academic libraries and library consortia in North America
- Strategic affiliates include OAIP, Association of University Presses, Open Textbook Network, DOAJ, CCB, SNAP, and more

Mission
- The Library Publishing Coalition (LPC) extends the impact and sustainability of library publishing and open access by providing a professional forum for developing best practices and shared expertise.

Values
- Professionalism, openness, diversity, collaboration, innovation

Strategic Plan 2018-2023
- Promote best practices in library publishing
- Strengthen the community of library publishers
- Act as a focal point and a resource multiplier for library publishing

Opportunities for Involvement
- Committees and task forces (member institution staff only)
  - Directory Committee
  - Professional Development Committee
  - Program Committee
  - Research Committee
  - Task Force on Diversity and Inclusion (2019-2020)
  - Publication Award Task Force (2019-2020)
  - Strategic Affiliates Program
  - Publishers & Service Providers Program
  - LPC Fellowship Program (for non-member libraries)

Resources
- Professional Development Offerings
  - 18 webinars offered since 2015 covering diverse topics such as ...
  - Publishing open educational resources
  - Best practices in OA publishing
  - Publishing with open licenses
  - DOAJ indexing
  - Publication agreements
  - Professional Development Guide: https://lpcguides.org/lphub

Freely Available Resources
- The Library Publishing Directory — an annual snapshot of the publishing activities of academic and research libraries
- An Ethical Framework for Library Publishing — supports best practices by providing resources and guidance in a number of ethical areas of importance to library publishers:
  - Publishing practice
  - Accessibility
  - Diversity, equity, and inclusion
  - Privacy and analytics
  - Academic and intellectual freedom
- The Library Publishing Curriculum
- M2S-Funded projects
- Four modules: Content, Impact, Sustainability, Policy
- How-To Guide for Library Publishers: DOAJ Application
- The Library Publishing Bibliography
- Recorded Webinars
- Job board

Member-Only Resources
- Shared Documentation Library
- Active and engaged email discussion list
- Webinars, roundtable discussions, other opportunities to learn and network
- DOAJ access

Library Publishing Forum
An annual conference for representatives from libraries engaged in or considering publishing initiatives. Goals:
- To define and address major questions and challenges
- To identify and document collaborative opportunities
- To strengthen and promote the community of practice