

Creating an open, community-driven and resilient database of life science literature metadata: a brief overview

Miriam Albers and Konrad U. Förstner

German Central Library of Medicine (ZB MED) – Information Centre for Life Sciences, Cologne, Germany

Abstract

This article presents a brief overview of the proposed OLSPub (Open Life Science Publication database) project, which seeks to develop a continuously expanding database of life-science metadata. Designed as a resource built by the community and for the community, OLSPub aims to provide an open, interoperable, and sustainable infrastructure that enhances access, supports discovery, and ensures long-term resilience in the management of biomedical literature.

Key words: databases, bibliographic; information storage and retrieval.

Introduction

Biomedical researchers worldwide rely heavily on PubMed as the primary database for life-science literature. The National Library of Medicine (NLM), which operates PubMed, reports more than 35 million individual bibliographic entries indexed through the PubMed system and 3.66 billion searches in PubMed in 2023 (1). In addition, PubMed's data is not only used via the search platforms, but also provides the basis for numerous other databases such as Open Alex (<https://openalex.org/>), Europe PMC (<https://europepmc.org/>), LIVIVO (<https://www.livivo.de/>) and many further services. This centrality creates a single-point dependency for discovery services used by clinicians, researchers, librarians and the general public. Recent policy developments in the United States, such as proposals to cut federal research budgets (2), including in the life sciences, and the abrupt dismissal of the Director of the Library of Congress (3), underscore the vulnerability of depending on a single national system for worldwide scientific access. Several of these decisions even with larger magnitude were made rapidly, leaving little time for the international research community to prepare or respond (one example is the sud-

den end of funding for the FlyBase database, a central resource for research on *Drosophila*, FlyBase, <https://flybase.org/>). The prospect that services like PubMed could be scaled back, restructured, or even withdrawn poses a serious risk for the global science community. For many researchers, clinicians, and librarians, PubMed is an indispensable daily tool. The fear of losing or losing access to such a resource already diverts time, attention, and energy away from scholarship and service (4).

OLSPub: a project proposal

OLSPub (<https://www.zbmed.de/en/research/current-projects/olspub>) is a proposal led by the German Central Library of Medicine (ZB MED) – Information Centre of Life Sciences, to establish an open-source, resilient, Europe-based life-science literature database intended as an interoperable and sustainable complement or alternative to PubMed with the MEDLINE database forming the core of PubMed. The project proposal, submitted to the German Research Foundation (DFG) on 28 May 2025 and publicly archived on Zenodo (5), aims to combine large-scale metadata harvesting, MeSH-aware subject indexing (including Ger-

Address for correspondence: Konrad U. Förstner, German Central Library of Medicine (ZB MED), Information Centre for Life Sciences, Gleueler Str. 60, 50931 Köln, Germany, Cologne, Germany

man translations), community governance, and hosting on national bioinformatics infrastructure (de.NBI). The proposal is currently under review.

To implement the proposed solution, OLSPub has articulated three primary objectives that collectively address the technical, user-facing, and organisational dimensions of building a sustainable open literature database for the life sciences. These objectives can be summarised as follows:

- (1) data integration: develop the legal and technical foundations required to ingest and harmonize current and future publications from publishers, with an initial focus on journals already indexed in MEDLINE;
- (2) Access and discovery: create a user-facing platform that mirrors the core functionalities of PubMed while extending them through MeSH-based indexing and multilingual support to enable more powerful search and retrieval;
- (3) community and governance: Build a durable collaborative network of stakeholders spanning individual researchers, libraries, and institutions to guide the long-term evolution of an open life-science literature database and to secure sustainable European oversight and management.

Rather than seeking to replace existing platforms, OLSPub aims to complement them by offering an independent and openly governed alternative that reflects European language priorities and infrastructure needs. Europe PMC (operated by EMBL-EBI) already delivers access to millions of biomedical abstracts and full texts with extensive support for text mining, while OpenAlex provides a comprehensive, openly accessible bibliographic index across research domains. LIVIVO is a discovery service that builds on MEDLINE/PubMed by expanding it to include additional sources of information on biomedicine, health and agriculture. However, Europe PMC, OpenAlex and LIVIVO all ultimately depend on PubMed as a core source of biomedical metadata and could not function in their current form without it.

The OLSPub project, whose proposal was submitted to the DFG for funding in May 2025, presents a concrete and technically feasible plan to develop an open, MeSH-enabled, Europe-hosted life-science publication database. The proposal was drafted following several community consolidation sessions and was made pub-

licly available immediately after submission (5). Its main strengths lie in a strong focus on interoperability, community integration, solid institutional expertise, and a commitment to multilingual metadata and transparent governance. Major challenges remain, namely legal agreements for data redistribution, operational funding, and community governance, but the project addresses these at the planning stage.

If successfully implemented, OLSPub has the potential to enhance diversity and strengthen sovereignty in life-science literature discovery. It would reduce dependency on a single source or jurisdiction for essential literature services, expand accessibility for non-English speakers through regionally curated multilingual subject metadata, and foster the development of open programmatic interfaces and community-driven tools to support text mining and reproducible bibliometrics. Rather than duplicating the functions of existing platforms such as Europe PMC and OpenAlex, these contributions would complement them, provided that strong interoperability agreements are established.

A decision on the project approval will be made this autumn. Regardless of this, ZB MED is already committed to building a community, for example through regular online information meetings. If the project proposal is rejected, less formalised financing options will be examined and utilised.

Submitted on invitation.

Accepted on 17 September 2025.

REFERENCES

1. National Institutes of Health. Medline PubMed production statistics. US National Library of Medicine, 2024. Available from: https://www.nlm.nih.gov/bsd/medline_pubmed_production_stats.html [cited 2025 Sept 14].
2. Cutler DM, Glaeser E. Cutting the NIH: The \$8 trillion health care catastrophe. JAMA Health Forum. 2025;6(5):e252791. Available from: <https://doi.org/10.1001/jamahealthforum.2025.2791> [cited 2025 Sept 14].

3. Trump abruptly fires librarian of Congress in latest purge of government (2025, May 9th). The Guardian. Available from: <https://www.theguardian.com/us-news/2025/may/09/trump-congress-librarian-fired>
4. Mallapaty S. "Omg, did PubMed go dark?" Blackout stokes fears about database's future. Nature. 2025 Mar;639(8054):288. Available from: <https://doi.org/10.1038/d41586-025-00674-3>. [cited 2025 Sept 14].
5. Förstner KU, Albers M, Rebholz-Schuhmann D, Seidlmayer E, Castro LJ, Lippert K. Open Life Science Publication database (OLSPub) - Strengthening the biomedical research community by building a resilient and sustainable solution. Zenodo; 2025. Available from: <https://doi.org/10.5281/zenodo.15533303>. [cited 2025 Sept 14].

