

Characteristics of a Preprint Server

Last Update Jan 31, 2016

*Anyone with this link can comment on the document and we encourage you to do so. The document will form part of the prereading to the **Accelerating Science and Publication in Biology (ASAPbio) Workshop to be held on February 16-17, 2016.***

What are the characteristics of preprint servers if they are to be adopted by the biomedical community?

- Governance
 - Oversight by all stakeholders - authors, reviewers, editors, readers, funders - with proportional representation in making decisions as to the development and operation of the preprint server. In short, a sense of shared governance.
 - Final decisions relating to the preprint server made by an individual/organization/committee trusted by the community.
 - [CC-BY](#) license or optionally [CC0](#)/PD for content - copyright retained by the authors but with unlimited access by for-profit and non-profit communities with attribution to the original authors.
- Community
 - Trusts the process of managing content and the content itself. Community has the sense they collectively own the content and the process by which it is managed. Those responsible for the oversight of the preprint server are simply the custodians of the scientific content. Custodians can change while the content remains available and immutable, though updatable.
 - Accepts that preprints - whether posted on this service or elsewhere - “count” in terms of establishing scholarly priority.
 - Is comfortable with the idea of multiple preprint servers that effectively compete for their papers, with the understanding that a preprint is only submitted to one server.
 - Embraces the idea of peer review of those preprints.
 - Accepts that posting a preprint does not imply pursuing a printed publication.
 - In the case where the preprint is not formally published, weighs in on how the preprint is treated by aggregation services such as PubMed, Google Scholar etc.
 - Defines the types of materials to be posted as preprints, which may go beyond what becomes traditional journal publications (e.g., workflows, large data sets, animations).
- Features (technical)
 - Platform is open source under an [OSI-approved license](#).
 - Easy registration service using an existing persistent author identifier.
 - Easy to use deposition interface that supports common document formats.
 - Assignment of a unique identifier for the deposition, which is persistently resolvable to the original deposition.
 - [FundRef](#) for affiliations.

- Support for versioning and recovery of the history of a deposition.
- Easy to use, fast and accurate search and retrieval of appropriate content.
- Simple views and download options in keeping with the leading current article dissemination platforms, including support for the [JATS format](#).
- Avoidance of feature creep - making the platform more complex to use with gains for only a small segment of the stakeholder population.
- Robust open application programming interface (API), so that features can be built around it by 3rd parties.
- Evolves to meet the needs of the community as scholarly communication itself evolves.
- W3C-compliant annotations using [Hypothes.is](#) or a similar tool.
- Various types of outputs for use by humans (e.g., PDFs) and machines (e.g., XML).
- Features (scholarly)
 - Ability to automatically and manually pre-filter submissions for material that is not-suitable as defined by the governing body, for example, material that is clearly not scientific or has been plagiarized.
 - Some form of editorial oversight defined by the governing body to handle retractions, ethical issues and other matters relating to the content of the server.
 - Built in post-submission review process and display of open reviews that are transparent and persistent and have their own unique identifier.
 - Ability to comment at any time.
 - Bibliometrics that fairly and accurately provide access statistics for each article - so called Article-Level Metrics (ALMs).
 - Ability to link to the version appearing in a peer-reviewed journal.
- Persistence
 - Archiving of both the code for and the content of the platform in a trusted long-term archival system.
 - Archival as both digital and hardcopy.
 - Established workable and transparent business model.
 - Understanding by all stakeholders as to the anticipated longevity of the preprint server, but with a clear and transparent plan for shutting down the service, should it not be needed any more.
 - Clear delineation of the responsibilities for the preprint server at all times during its lifecycle.
 - Clear and transparent disaster recovery plan, which is periodically reviewed and tested.