

IMPROVING SCHOLARLY PUBLISHING METADATA

CHORUS FORUM

29 September 2022

Overview of weak spots of scholarly publishing metadata: Read-and-Publish agreements

Alexander ('Sasha') Schwarzman

aschwarzman@optica.org

We Are Now

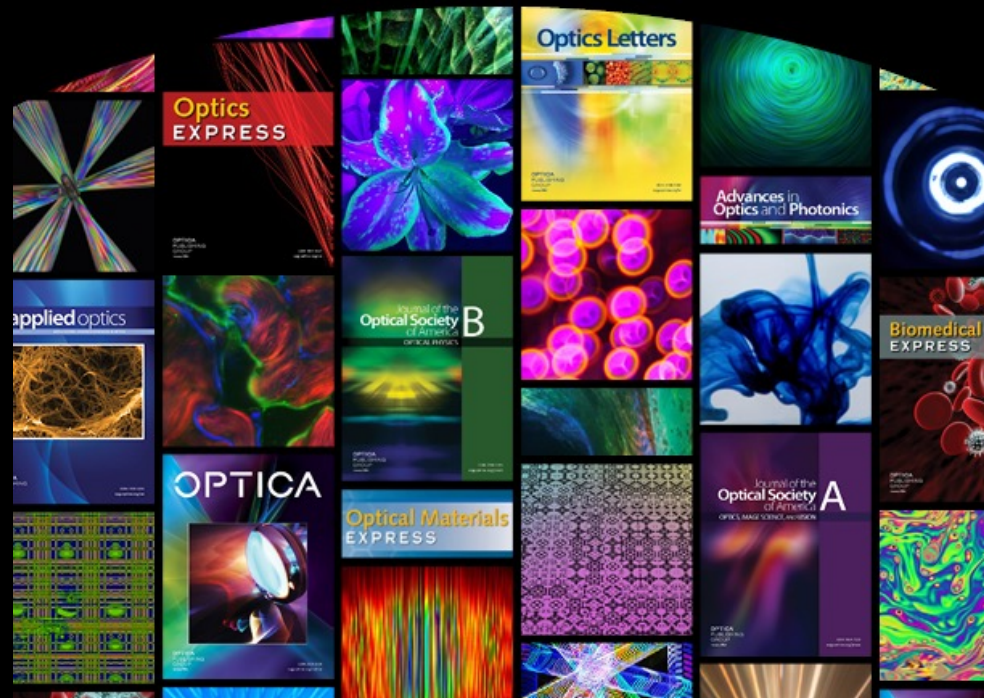
OPTICA PUBLISHING GROUP

Formerly
OSA

OSA has a new name: Optica, the Society advancing optics and photonics worldwide.

Optica Publishing Group, a division of Optica, publishes the most-cited and largest peer-reviewed collection of optics and photonics content.

opg.optica.org



Contents

- Read-and-Publish agreements
- Metadata
 - Authors
 - Affiliations
 - Article types
 - Licensing
 - Funding
- Metadata retention and preservation
- Conclusions

Read-and-Publish agreements

Read-and-Publish (R&P) agreement: the publisher receives payment for reading and payment for publishing bundled into a single contract.

(Lisa Janicke Hinchliffe)

- **Read:** author's institution has access to the publisher's subscription content.
- **Publish:** the Article Processing Charges (APC's) for Open Access are defined in the contract between the publisher and a library consortium that includes author's institution.
- Institution's expenditures on subscriptions (**Read**) are redirected to paying APC's (**Publish**).
- APC's are paid by the consortium, rather than by an individual author.

Authors: eligibility and identifiers

- **Eligible authors: only corresponding or any?**
 - ❖ **Metadata/markup: two ways of indicating “corresponding” in JATS**
- **How to indicate author’s R&P benefit eligibility?**
 - ❖ **Metadata/markup: no best practice, no recommendation**
- **Identifiers**
 - **ORCID**
 - **name + email**

Authors: eligibility and contribution type

- Contributor's role in producing the scholarly output (investigation, writing, project administration, data curation, visualization, etc.)
 - ❖ Metadata/markup: CRediT taxonomy, JATS4R recommendation
- Contribution level
 - contributed equally
 - shared contribution
 - joint contribution
 - joint first authors
 - co-first authors
 - co-last authors
 - ❖ Metadata/markup: no taxonomy, no recommendation

Affiliations: timing, identifiers, primary affiliation

- **Timing**
 - Affiliation at the time research was done
 - Current affiliation
- **Identifiers**
 - ROR, GRID, Ringgold, Crossref Funder ID/DOI, ISNI
 - Granularity (ID at the level of R&P-participating institution)
- **Primary affiliation**
 - How to indicate primary affiliation for a R&P-eligible author?
- ❖ **Metadata/markup: no recommendations**

Article types. Licensing. Funding

- **R&P-ineligible article types?**
 - Editorial, Correction, Comment, Reply, Retraction, ...
- **Licensing**
 - CC BY (what type? CC BY-SA, CC BU-NC, CC BY-NC-SA, CC BY-ND, CC BY-NC-ND)
- **Funding**
 - Library consortium (must or may)

Metadata retention and preservation

- Production metadata versus long-term preservation metadata
- Is R&P metadata worth preserving or should it be removed upon publication?
- Retain for auditing (librarians, funders, publisher's internal)
- Retain for business intelligence (analyses for renewing current R&P deals and negotiating future ones)
- Share with Crossref? (But – community best practices are still emerging)
- Retain in the publisher's version/Crossref databases but not transmit to Portico/LOCKSS/CLOCKSS for long-term preservation?

Conclusions

- **Diversity of R&P business models**
 - **Differences in what metadata is considered significant**
 - **Lack of recommended markup practices**
 - **Absence of metadata retention guidelines**
- ❖ **Challenges in:**
- **findability**
 - **accessibility**
 - **interoperability**
 - **reusability**
 - **building community-wide metadata validation tools**