

# PRODUCTION METRICS AND METADATA

---

Using Production Tools to Gather, Track, and  
Publish Content and Metadata

Shaun Halloran

Operations Manager

*Environmental Health Perspectives*

# About *EHP*



- Founded in 1972
- Published with support from the National Institute of Environmental Health Sciences, a division of the NIH
- Online only since 2013
- All content is Open Access and free to reuse
- Publish peer-reviewed research papers, original news articles, proceedings abstracts, and a Chinese Edition
- Recently moved to a continuous publication workflow

# Production Matters

- Editorial review differs from production review
  - Focuses on organization and completeness
- Opportunity to validate metadata ahead of publication
- Normalize content to a specific baseline, or converge data from various sources
- Provides a more consistent product for your readers (and Marketing staff)
- Good production processes improve timeliness and quality

# Production Tracking System

- Allows for smooth handoff of content and data collected during the peer-review process
- Keeps data electronic and validated, such as:
  - ORCIDs
  - DOIs
  - Preprint information
  - Funding data
  - Information about figures, tables, supplemental materials
  - Related article information
  - Permissions letters for reused content
  - Correspondence history
- Provides opportunity for capturing data relevant only to accepted papers

# XML: Built-in Metadata!

- XML can hold almost all relevant data/metadata from your tracking system
  - Aids in discoverability if your online systems surface this information
- Tracking system may output metadata in XML format, allowing for an easy and accurate transfer to composition
- Can be parsed for errors using DTD and Schematron
- Easy translation to other formats
- Can be customized for your particular use
  - Certain types of reuse and transformations may suffer, however

# One Size Does Not Fit All

- There are many types of XML DTDs available, depending on your needs:
  - JATS/NLM
  - DITA
  - BITS
  - EPUB
  - DocBook
  - MathML
  - BeerXML (seriously!)
  - And many more from countless sources – XML is a viable solution for almost all publication types



# Choosing Your Composition System

- If using a single-source XML workflow, consider systems that will preserve the integrity of your XML
- Some systems require conversions, which could introduce errors
- A good composition system will ingest XML, allow modifications or updates to any metadata, and output XML nondestructively
- Some systems also offer first pages in minutes, improving turnaround times

# Cloud-Based Author Proofing

- Pros:

- Keeps data electronic at all times
- Updated references can be validated in real time as the author introduces new content
- Additional opportunity to get ORCIDs, preprint IDs, funding data, etc.
- May be quicker than incorporating traditional PDF corrections

- Cons:

- Still requires review of author corrections
- Systems are still in development, and all features may not be available yet
- Latency and bandwidth issues may cause problems for international users



0	1	0	1	0
1	0	1	0	1
0	1	0	1	0
1	0	1	0	1
0	1	0	1	0
1	0	1	0	1
0	1	0	1	0
1	0	1	0	1
0	1	0	1	0

# Continuous Publication

- Not a requirement for maintaining good metadata, but will help your time-to-publication metrics
- Keeps the focus on the article instead of the issue
  - Easier to track article engagement via analytics
- Article metadata should still include issue publication information (if you continue to use issues)
- Citations count, even before the content is in an issue
  
- But, consider how this impacts embargoes and press releases if you regularly issue these
- Each paper follows its own production schedule, which can increase overhead

# Final Deliverables

- The XML is a deliverable
  - It should include any data/metadata needed for discoverability and SEO
  - Its content should match the other deliverables in both data and metadata, where possible
  - Tag as much as you can – it's cheap insurance
- Don't forget the Properties of your PDFs
- Talk to your platform provider

# Vendor Support and Accountability

- Typically not a good idea to build all systems on your own
- Numerous vendors can offer assistance
- Allows you to focus on content, while they focus on technological development
- You will still need to be a driver of change
- Build adequate oversight into vendor contracts
  - Live portals help you keep tabs on work



# Best Practices & Recommendations

- NISO Access and License Indicators (ALI)
  - <http://www.niso.org/workrooms/ali/>
- Dublin Core
  - <http://dublincore.org/>
  - For reference, at least
- NISO Journal Article Versions (JAV)
  - <http://www.niso.org/workrooms/jav>
- JATS for Reuse (JATS4R)
  - <http://jats4r.org/>



# Thank You!

“Metadata is the DNA of the data warehouse, defining its elements and how they work together.” - Ralph Kimball

Shaun Halloran

Operations Manager

*Environmental Health Perspectives*

<https://ehp.niehs.nih.gov>

[shaunh212@gmail.com](mailto:shaunh212@gmail.com)

