

# Scalable Electronic Publishing in a University Library

Scholarly Publishing Office • University Library University of Michigan • <http://www.lib.umich.edu/spo/>

## SPO in brief

### History

Formed as a division of the University Library in 2001.

### Mission

Extends the Library's commitment to the distribution of scholarship by experimenting with innovative methods for publishing to serve the needs of scholars, both at the University of Michigan and around the world.

### Staff

7.5 FTE; approximately 1.5 FTE in part-time student employees.

### Scale

In 2007, published 2,000 articles in journals, reviews, and conference proceedings, plus a handful of monographs, image collections, and other digital projects.

### Infrastructure

DLXS middleware for search and display interface with XPAT search engine for text content.

### Content ingestion

See handout for workflow.

### Business models

SPO encourages open-access publishing but publishes subscription-based resources as well. SPO works with SPARC, the Scholarly Publishing Academic Resources Coalition, to develop business plans for journal editorial boards, especially those making a transition from print subscriptions.

### Author rights

SPO believes authors should retain strong rights to their content. Instead of copyright transfer, SPO requires only a non-exclusive perpetual license to publish the content online: this is obtained either from the author or the editorial board.

Feature	Technical details
Content preservation	Content receives preservation pledge like the Library collection. NAS in redundant data centers used to store: •XML encoded text (schema derived from TEI Lite P3) •TIFF page images (when applicable) •TIFF or JPEG 2000 high-resolution images (when applicable) •Other formats (PDF, GIF, JPEG, multimedia) as bitstreams
Content crawled by major search engines	Publicly available full-text content crawlable by Googlebot, Yahoo! Slurp, and msnbot (Live Search).
Bibliographic exposure	Resources cataloged by University Library, uploaded to OCLC. Metadata shared through OAI data provider (OAIProvider) and OAI service provider (OAIster).
Low barriers to access	Server-side processing and rendering of content for user: little performance degradation on older computers, no plugins required.
Robust searching of content	DLXS uses XPAT for word, phrase, and stemming search on metadata, full text, or particular XML elements.
Robust browsing of content	Browse by title, author, or issue of a journal (when applicable). High-resolution images can be opened in a separate window with zooming and panning capabilities.
Customizable interface	Color scheme, fonts, images, entry pages all highly customizable for individual publications. System for browsing and searching has some flexibility.
Usage statistics	Locally gathered usage statistics and reports in COUNTER format. Code for Google Analytics can be embedded in pages.
User tools	Content can be flagged using bookbag (like a shopping cart) or tagged using MTagger (similar to del.icio.us).
Web feeds	Atom 1.0 feed for new content in a particular journal of collection.
Authentication and authorization systems	Publishing partners have access to development server for previewing content. Subscription-based resources can be restricted by IP range or username and password.
Printing services	Books can be sold through print-on-demand technology. SPO can sell print copies from short-run digital printing.
Editing, indexing, and graphic design	While SPO doesn't routinely provide editing, indexing, or graphic design services, we can refer publishing partners to freelancers who do such work.
Digitization of print	SPO can arrange for back issues to be scanned and OCRd.
Persistent URLs	Handle System used for persistent identifiers.

## Future plans

### Single metadata management interface

Develop web interface for metadata creation and editing and for content ingestion. This will replace outdated workflows for SPO staff and allow publishing partners to create their own metadata and submit content to SPO in a more regularized way.

### Improve user experience

Better navigation, cleaner display of metadata and encoded text, spellchecking of search queries.

### New revenue models

Pay-per-download option for subscription-based content.

### Web 2.0 tools

Better integrate social bookmarking and community-building tools into the delivery system.

### Better, more explicit user rights

Renegotiate licenses with our publishing partners to allow for Creative Commons licensing.

### XML for users

Allow users to download XML.

### Bibliographic integration

Add OpenURL COinS and AB Meta to content.

### Improved printing options

Dynamic generation of PDFs from XML (for printing by users).

### Improved ingestion workflow

Implement better conversion methods for generating XML from unstructured electronic source files (Microsoft Word, Adobe InDesign, etc.).