Managing Data for Scholarly Communications

PART 1: Supplemental Materials

October 12, 2011

Speakers: Linda Beebe and Mike Buschman

Collision of 2 Worlds

Supplemental Journal Article Materials and the NISO-NFAIS Working Group Recommendations

Linda Beebe
October 12, 2011

Explosion of—
- Research
- Data
- Accrued Knowledge

Increased Requirements
- Funding Bodies
- Reporting Standards
Supplemental Materials—it sounded like such a good idea!

- Give author opportunity to expand on their research.
- Improve science by giving easy access to data needed to verify or replicate study.
- Enhance reporting of science with multimedia.
- And do it all at little or no additional cost!

Outcome for the user?

- Lack of descriptive metadata
- Discoverability issues
- Lack of context
- Concern about persistence
- No clarity on citations
- Some mystery in the main article about what is supplemental—a maze, maybe not value-added taken as a whole.

Outcome for the publisher?

- Direct costs
- Diverted energies—already crisis in peer review
- Tough decisions—
  - What is value-add?
  - Peer review dilemma, quality vs workload?
  - Plan for migration?

Two Indicators of the Problems

- **Journal of Clinical Investigation**
  - 2% of journal articles in 2000 had supplemental information.
  - 87% in 2009, then 93% in 2010
  - Projecting 95% in 2011.

- **Journal of Neuroscience**
  - Stopped accepting supplemental in August 2010.
  - Cited severe drain on peer review system.
  - Average size of content grown from under .6 MB in 2001 to 1.2 MB in 2009.
### NISO-NFAIS Working Group

- American Psychological Association
- Thomson Reuters
- AGU
- ACS Chemistry for Life
- NFAIS
- MIT
- ASCI
- ITHAKA
- IEEE
- NLM
- JAMA
- CDL
- AIP
- HighWire
- Crossref
- ACM
- NCBI

### Actually 2 Working Groups

- Business Working Group, co-chaired by Marie McVeigh and Linda Beebe
- 10 other members including Bonnie Lawler, ED of NFAIS
- Task—the “what”—policies and practices

- Technical Working Group, co-chaired by Dave Martinsen and Sasha Schwarzman
- 15 other members
- Task—the “how”—technical aspects of implementation

### The Vision

- “Reader’s Choice” article
- Modular with what is now supplemental material available in some modules
- Clearly identified and easy to select
- Emilie Marcus of Cell: “a more modern concept of a hierarchical or layered presentation in which the reader can define which level of detail best fits their interests and needs.”

### The Current Reality

- That maze I showed you earlier—
  - No standards
  - Inconsistent identifiers
  - Poor metadata
  - Lack of discovery tools
  - Abuse of readers and reviewers
Recommended Practices

- Intended to help scholarly community develop more standardized approaches.
- Lessen the burden on all parties.
- Assure that Supplemental Materials add substance to the scholarship.
- Make them more discoverable.
- Aid in preserving Supplemental Materials.

Challenges & Principles

- Wide variance in disciplines, types of content, citation systems
- Evolutionary time—technology changing so rapidly that this must be a living document.
- Readers vary in need for information.
- Need to look to the future, not limitations of the past.
- Recommended practices, not rules.

First, Definitions—

- Article
- Data
- Metadata
- Multimedia
- Supplemental Materials
  - Integral Content
  - Additional Content
- In addition, Related Content.

1. Integral Content

- Sounds like an oxymoron—but publishers currently treat some content essential to the understanding of the article as if it is supplemental for technical or logistical reasons.
- Examples: detailed results or methods, multimedia, extensive tables or figures.
- “Pseudo-Supplemental” or perhaps a transitory phenomenon.
- Generally publisher is responsible for hosting and curating in same way they do the article.
2. Additional Content

- Provides additional, relevant, and useful expansion of article, but not necessary to understand the core article.
- Examples: expanded methods sections, bibliographies, additional supporting data or results, copies of instruments, multimedia, interactive representations.
- Author has created, and generally publisher hosts or places on open web.

Related Content

- Often there is other work author wants to call attention to, but is not the typical cited reference.
- Publishers treat differently—some include in cited references, some link to it outside the citation list.
- Generally content in a repository or data center. Publisher has no authority.
- No recommended practices, but some comments on preservation and linking.

Part A, The Business Practices

- Selecting Content
- Maintaining Links
- Editing Content
- Providing Context
- Managing & Hosting
- Preserving Material
- Assuring Discoverability
- Rights Management
- Referencing Materials
- Appendixes A & B

Selecting Supplemental Materials

Cautions
- Are they relevant?
- Are they useful additions to the scholarly record?
- Can they be preserved?
- Consider costs of vetting, delivery, and maintenance.

Best Practices
- Integral content: review at same level as article. Be sure reviewers get all content.
- Additional content: ideally the same, but may not be possible.
- Provide clear guidelines.
**Editing Supplemental Materials**

- Technical editing is determined by editor and publisher.
- Integral Content should be edited at same level as core article.
- Ideally, Additional Content also. If it is not, indicate it is as author provided it.

**Managing and Hosting Materials**

**Who does it?**
- Generally, publishers manage and host their content.
- Some outsource. Contracts should be crafted to assure the level of protection needed.
- Journal content hosted by an aggregator or other host must include supplemental material.

**External Repositories**
- Data hosted by an external repository may be deemed essential. The publisher must preserve the link between core article and data in repository.
- Repositories may be unstable.
- Author web sites are verboten as repository!

**Assuring Discoverability—Consistency**

- Use consistent naming conventions for Supplemental Materials in TOC and article.
- Reference the core article from the Supplemental Materials consistently.
- Locate pointers to the Supplemental Materials in the same place always.
- Streamline navigation.
- Practice consistency from article to article and issue to issue.

**Consistency, Part 2**

- **Tables of Contents**—Indicate clearly presence of Supplemental Materials in online TOCs; use easy-to-understand and consistent naming convention.
- **Display in Article**—locate links upfront on first page of article in PDF and HTML. Link to Integral Content from within article. Important to be readable by machines and humans.
- **Display in Supplemental Materials**—match navigational elements within Supplemental Materials to those in core article. Use consistent locations as much as possible.
Discoverability with A&I Services

**Why?**
- Scholars rely on A&I services
- Assures ongoing use

**How?**
- Good metadata
- Purpose & format
- Difficult to find
- Need new fields

**Obstacles**

Referencing Supplemental Materials Within the Article

- Describe (cite and link) at the same level as table or figure in core article.
- Do not add bibliographic citation for integral content, as it is part of the article.
- Provide in-text citation to additional content and link to it at appropriate point in text.

Reference in Other Publications

**Integral Content**—cite the article as whole; do not cite separately. Assign unique DOI to materials or use article DOI with sub-suffix.

**Additional Content**—may cite separately. Some style systems advise using words “Supplemental Material” in citation.

Maintaining Links

- Links to information owned, published, or contracted by publisher must be maintained.
- Broken links to content outside their control should be resolved when possible or deactivated.
- If there are alternative access points, publisher should indicate.

- Bidirectional linking for integral and additional content is essential.
- DOI registration assures persistence & findability.
- Use separate DOIs for integral and additional content to enable separate linking.

Best Practice
Providing Context

- Among the most egregious of current problems—what am I looking at and why?
- Need an indication of the nature of the content and its connection to the article.
- Need reminder of what they are seeing or hearing.

- Include following on landing page or in content:
  - Article citation and DOI
  - Title and/or succinct statement about content
  - File extension and indication of size for multimedia
  - If multiple files, a list is helpful.
  - Player information for multimedia
  - If separate DOI, list.

Preserving Supplemental Materials

**Integral Content**
- Preserve same standard as core article.
- Give same metadata markup.
- Tag text accompanying multimedia appropriately.

**Additional Content**
- Consider preservation before accepting.
- If unsure of ability to preserve, ask author to put in trusted repository.
- To extent possible, tag as for core article.

These constitute an important part of the scholarly record. Consider distributed repositories. Specify acceptable repositories or outline criteria for evaluating them. Ideas: Accessibility, perpetual archiving, and bi-directional linking.

Rights Management

**Intellectual Property**
- Treat rights for Integral Content same as for core article.
- Rights for Additional Content may differ—but should be transparent to users.

**Access**
- Anyone with rights to core article must have rights to Integral Content.
- ILL is a particular concern for librarians—best practice is to provide both article and Supplemental Materials.

Some Additions

**Roles and Responsibilities**
- Maintaining and delivering the record of scholarship involves many parties:
  - Publisher
  - Editor & Peer Reviewer
  - Author
  - Libraries
  - A&I Services
  - Repository Administrators

**Linking to Related Content**
- Citing practices vary—some link from within text, some from reference list. Latter enables easy export to manager.
- Repositories now enabling specific links.
- Bidirectional links are strongly recommended.

Best Practice
Part B, Technical Considerations

Building Blocks
- Finish draft BWG Recommended Practices.
- TWG on packaging, metadata, etc.
- Meld the two.
- Open Review
- NISO & NFAIS approve.
- Circulate for other approval.

THE ARTICLE OF THE FUTURE MAY SOLVE SOME OF THE PROBLEMS.

Linda Beebe
Senior Director, PsycINFO
American Psychological Association
lbbeebe@apa.org

Summon™
Web-scale Discovery Service
MIKE BUSCHMAN, Director, Product Management

Serials Solutions
A ProQuest Company
A real digital front door for the library

Customers of the Summon™ service include
- 26% of ARL libraries
- 63% of Ivy League schools
- 50% of Group of Eight (Go8) libraries in Australia
- Several of the largest libraries in China
- Military/Government libraries
- The World Health Organization
- Many of the largest research libraries around the world

Content growth
New content sources added every week
Updates from existing sources added every day

Supplemental materials—what does that mean for the Summon™ service?
- Enhancement records
- Full text
- Tables, charts
- Data sets
- Photos
- References
- Publication and series records
- Tables of contents
- Book chapters
Most complete and richest representation of metadata

**Match and Merge Process**

- **DOI from CrossRef**
- **Abstracts, subject terms, etc. from Gale, ProQuest**
- **Peer-reviewed status from Ulrich’s**
- **Full Text from APA, Sage, Wiley**
- **Citation counts, subject terms, etc. from Web of Science**

**Summon™ Record**

---

**Data sets**

*Search Results: Your search for returned 4,566 results*

- **Higher Education Annual Enrollment Survey, 1974**
  - Enrollment statistics for the state of Connecticut are presented. Highlights indicate: (1) Total enrollment in all colleges and universities in the state area; State University, Enrollment Totals, Enrollment, Higher Education, State Colleges
  - **Data Set**: Data Set Online

- **Attitudes of South Dakota Elementary and Secondary Parents toward the Public Schools**
  - by Walker, Loraine and Wiedl, Robert W
  - This report presents the results of a modified version of the 1963 Gallup Poll survey of parents of school-aged children in South Dakota. Parents of elementary or State Surveys, Parent Attitudes, Discipline, Educational Attitudes
  - **Data Set**: Data Set Online

- **Time-Saving Tables for Computing the Harris-Jacobson Readability Formulas**
  - by Smith, Lawrence L
  - The ten tables presented in this paper were prepared in the belief that more educators would use the Harris-Jacobson Readability Formulas if the operations were... Reading Instruction, Readability Formulas, Scoring Formulas
  - **Data Set**: Data Set Online

- **Education in OECD Developing Countries: Trends and Perspectives**
  - by Organization for Economic Cooperation and Development, Paris (France)
  - 1974, 26
  - A major purpose of this publication is to describe and analyze educational development during the 1960s in the developing countries that are members of the... Educational Development, Educational Development, Educational Development, Educational Development
  - **Data Set**: Data Set Online

---

**Institutional repository and other local collections**

- **Summon™**

---

**Multimedia content**

- **Summon™**
Full text search of printed books

- HathiTrust to provide full text of over 9 million print items
- Search the pages of the books from the stacks
- Other publisher and open access full text sources

Winner of prestigious awards

Software & Information Industry Association (SIIA)

CODiE

2011 SIIA CODiE WINNER

Best Enterprise Search Solution

In great company... past CODiE winners include:

Thank You!

Mike Buschman

mike.buschman@serialssolutions.com