

Roundtable on Best Practices for Supplemental Journal Article Materials

Co-Sponsored by NFAIS and NISO

January 22, 2010

Washington, DC—American Psychological Association Building

The results of a recent survey on how publishers handle supporting materials in scientific journals by Alexander (Sasha) Schwarzman of the American Geophysical Union (AGU) has generated considerable interest within the information community. In recognition of the importance of this topic, on January 22, 2010 the National Information Standards Organization (NISO) and the National Federation of Advanced Information Services (NFAIS) cosponsored a roundtable discussion on the need for standardized bibliographic and publishing policies for supplemental material. Additional meeting planning and support was provided by the American Psychological Association (APA), who also hosted the meeting at the APA headquarters. Invited participants represented various journal publishers, scholarly organizations, and libraries. A full list of participants is provided at the end of this report.

This group met to discuss the issue of supplemental materials in journals in order to discover if there might be an opportunity to reach some consensus on how to work with supplemental materials in a more standardized fashion for improved management, access, and discoverability. This report summarizes the meeting and provides recommendations for further action.

Updates on Ongoing Work

The day began with updates from three attendees on extant work and projects in this arena to set the stage.

Supplemental Survey and White Paper

The impetus for the joint NISO/NFAIS meeting came from a white paper written by Sasha Schwarzman, Information Systems Analyst-Designer, AGU, that focused on the question of “how to handle supporting material in a scientific journal.”¹ Schwarzman began by providing a review of his report and some of the main issues he found related to supplemental materials. First, he stressed the need to clarify terminology. Although for this meeting the term “supplemental” is used to describe these materials, many other similar terms are used in this context, such as “supporting,” “supplementary,” “auxiliary,” “ancillary,” and so on.

Schwarzman identified two main groups of issues that need to be examined before we can formulate criteria for what constitutes supplemental material.

1. Distinguishing characteristics of a scientific research article as a literary genre.

Schwarzman noted that both electronic and print publishers want to retain an uninterrupted narrative flow and cogent presentation, so we often see print publishers “dumping” supplementary material on the Web or electronic publishers “hiding” the material in the lower hierarchical levels of the layered article presentation. What precipitates these actions is the need to ensure that the scientific

¹ Schwarzman, Alexander (Sasha). *Supporting Material*. American Geophysical Union: November 3, 2009. Available at <http://www.agu.org/dtd/Presentations/supporting-material.htm>

articles as a literary genre will preserve its integrity. Thus, the issue of how supplemental materials fit in this genre comes to bear. “Where do supplemental materials belong *structurally*?” Do they fit at the entire article level (for instance, as a supplementary section, such as Extended Methodology) or at the individual component level (where, for instance, Table 1 is part of the “core” article, whereas Table S1 is supplemental), or both? Further, are there content types that are inherently supplemental, such as datasets? Schwarzman noted, by answering these questions in a consistent, agreed-upon way we also will be able to help prevent some current abuses that we see:

- Abuse of reviewers, who have to review the many pieces that are included in the catchall area of “supplemental materials,” which at times are used by authors as a vehicle to smuggle an entire new article into publication.
- Abuse of readers, who must wade through an unreadable jumble of material, unsure how carefully it was reviewed and whether this is worth their time and effort.

2. *Costs and benefits of adding value to content.*

Schwarzman said it is essential to consider the business case of supplemental materials—that is, how to balance the real value that they add to content with the costs of adding that value. Do we apply the same criteria that we use for the core “article” to the supplemental materials? Costs may arise during the various work processes associated with making the materials available, including:

- Peer review
- Readability: copyediting, navigability, providing user friendly interfaces
- Usability: multi-format and multi-channel publishing
- Discoverability: applying metadata, creating linking references, building citation indices, providing links to related content, assigning DOIs, and so on
- Dissemination: depositing metadata to various channels and paying the associated fees to do so
- Preservation and Longevity: markup, migration, and so on

Schwarzman pointed out some questions for the group to consider, noting that they will need to be answered in any steps forward:

- 1) *Who?* If best practices or policies are developed, who should be creating them?
- 2) *What?* What is supplemental versus core or related? How does custodianship of the article/materials play into this?
- 3) *How?* How to implement the policies and ensure uniformity of their application across the industry, a publishing house, or a journal?
- 4) What *business models* can we use to help recuperate costs?

Note: Schwarzman would like to acknowledge his debt of gratitude to Emilie Marcus, Cell’s Editor-in-Chief, on whose editorial “Taming Supplemental Material” (published in Cell 139, October 2, 2009, p. 11) he based his considerations about “abuse” of reviewers and readers, as well as his points about where supplemental materials may belong structurally.

EU PARSE Project

Eefke Smit, Director of Standards and Technology, International Association of Scientific, Medical and Technical Publishers (STM Association), provided attendees with an outline of the PARSE.Insight Project², a two-year initiative funded in part by the European Union. PARSE (Permanent Access to the Records of Science in Europe) “aims to highlight the longevity and vulnerability of digital research data and concentrates on the parts of the e-Science infrastructure needed to support persistence and understandability of the digital assets of EU research.”³

Smit explained that PARSE started with looking at preservation, but now has expanded to also consider reuse and access, from gray literature and datasets to official publications. The scope of the project is international, and the deliverable is recommendations for a roadmap to the necessary infrastructure for preservation of the scientific record. Engagement in the project includes research institutes, national libraries, and publishers.

For the past two years, PARSE has been making inventory with surveys and case studies. Highlights of findings:

- Preservation is best organized around official publications. About 97% of scientific journals have a preservation policy in place, including disaster recovery. Among smaller publishers, though, less than 60% have preservation measures in place.
- Datasets/supplemental materials are at the least organized end of the spectrum. Only about 20% of the 1,500 researchers surveyed are sharing their datasets outside of their own research group. When asked where they wanted their datasets housed/archived, the initial response was that they would like to see them in a reliable, central repository, for example, associated with a government or institution. Publishers shouldering this role were a secondary choice (given by approximately 50% of the authors).
- When publishers were asked where they thought that datasets should be housed, about 70% responded that they should be included in a repository. However, less than 1% of publishers (more than 100 of which were in the survey, representing 8,000 journals) were willing to assume that responsibility.
- About 20-25% of the publishers surveyed treat the datasets/supplemental materials in the same way as the article themselves (in terms of preservation approach).

PARSE has a number of follow-up projects being worked on at the moment; most are about six months out. At this time, however, they have identified some characteristics for an ideal system:

- Good linking system
- Reliable metadata in place
- Certification of repositories
- Registration of datasets

² PARSE.Insight. <http://www.parse-insight.eu/>

³ PARSE.Insight. “About the Project.” <http://www.parse-insight.eu/project.php>

Interactive Science Publishing (ISP) Initiative

Scott Dineen, Deputy Senior Director of Publications, Optical Society of America (OSA), then provided an update of the Interactive Science Publishing (ISP) Initiative⁴, an experiment that OSA undertook with publishing large databases with its journals, done in partnership with the U.S. National Library of Medicine (NLM). Although OSA had been publishing movie files with its journals since 1997, the ISP Initiative began with an interest in sharing 3D lung cancer datasets. With the new imaging techniques available for measuring tumors, providing the datasets with the article gives context and validity—however, the sets are very large, usually several terabytes in size.

To achieve their objective, OSA worked with colleagues at NLM to set up a DSpace architecture (called “MIDAS”) that would let them accept author datasets, solving the first part of the problem. The second part, however—readers’ seamless access to the datasets and display of those datasets in a way that the reader could interact—still needed to be solved. This was done with a viewing software that allows readers to rotate, crop, zoom in/out, run cross-cut analysis, and so on

Some struggles in the process were:

- *Difficulty getting authors to give their datasets to OSA.*
Although some authors were hesitant to provide their datasets because of potential misuse, most of the difficulty was because authors didn’t know *how* to provide the datasets. A great deal of staff resources was spent in collecting, converting, organizing, and producing the datasets for display. Over the course of 18 months, the staff time amounted to approximately 1 FTE senior manager plus 1.5 FTEs in production. The datasets are open access, as are the articles.
- *Integrating supplemental materials in the peer review process*
- *Lack of metadata.*
Not only is there a scarcity of good models to work with, but authors weren’t providing terms for them to use, despite some built-in opportunities for harvesting.

They conducted a reader survey, and thus far they have received feedback that the dataset availability has been an added value, though about 15% of people who were surveyed indicated that they had technical problems viewing the data.

At this point, OSA is looking at grant opportunities and licensing models, but is also reaching out to other players, such as Microsoft to see about alternative ways to fund infrastructure as well as to give OSA opportunities to market their products. Dineen noted that the business model at this point is still not clear.

Defining What Is Supplemental

Following the three updates that provided some background and context for the day’s discussion, Linda Beebe, the meeting facilitator, led the group in discussing how to define what supplemental materials are. Members noted that a particular format doesn’t automatically relegate the item to a supplemental role; for instance, a figure may be a part of the core article, or may be considered supplemental. At this time, there is no uniform treatment of supplemental materials. To gain a better

⁴ *Interactive Science Publishing*. The Optical Society of America. <http://www.opticsinfobase.org/isp.cfm>

understanding of the kinds of materials that are being considered, the participants created the following short list of content types:

- Datasets
- Figures (including high-resolution figures)
- Tables
- Movies
- Software/scripts (or network files in format)
- Videos
- Appendices
- Audio files
- Images
- Text

AGU’s position is that supplemental materials should not be anything that the reader needs to make basic conclusions. That is, the article should stand alone, and the supplementary material should provide the next step for more information, if desired.

Science uses the criterion that supplemental materials should be “essential to the scientific integrity of the article,” and exist as a tool for the authors to make their full case. Articles published in *Science* are frequently very short; the associated “supplemental” materials can be quite extensive.

All participants agreed that supplemental materials should be peer-reviewed at the same level as the article with which they are affiliated.

The scope of this problem was highlighted by Ken Beauchamp, Online Production Editor of the Journal of Clinical Investigation. He presented a slide (see Figure 1) showing the percentage of articles JCI has published that included some form of supplemental material. The percentage has increased from 2% a decade ago to 87% in 2009. Beauchamp reported that so far in 2010 only one article did not include some form of supplemental materials.

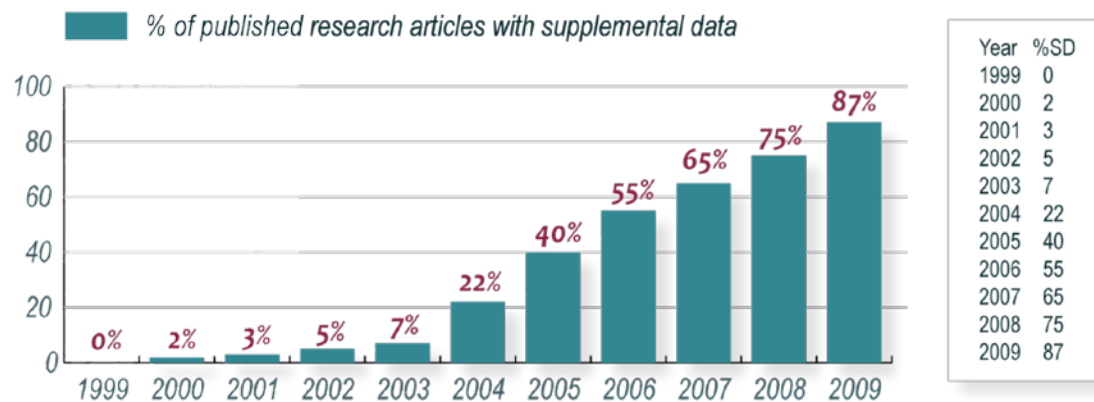


Figure 1: Journal of Clinical Investigation % of published research articles with supplemental data

Determining Roles

Determining what is supplemental has thus far been largely subjective. The following roles for decision-making were discussed:

Authors

Although it was acknowledged that at times authors use supplemental materials as a way to get around length rules and other article requirements, many authors put a lot of thought into what they include. There is a need to educate authors about what should be included or excluded.

Peer reviewers

Although peer reviewers should look at proposed supplemental material when reviewing a manuscript, they often do not. Along with authors, peer reviewers need education about what should be included or excluded.

Editors

Editors set policy and make the final decisions. However, they need to balance many issues and do not necessarily think about downstream processes and needs when making decisions about the use of supplemental material.

Publishers

Publishers have a key role to play in this process, particularly in educating editors about those things that happen after an article is accepted—downstream in workflow—so that the editors may be able to make better decisions.

Libraries/Data Centers

Actual data, particularly data sets that support multiple lines of research and publication, will most likely be gathered together in repositories that are managed and curated by libraries or independent data centers.

Defining Recommended Practices

There was considerable discussion about findability. A&I services are not all able to capture the availability of supplemental materials because of inconsistencies in notification. Some publishers assign DOIs to supplemental material; others do not. It appears that most publishers are marking up a minimal amount of metadata describing their supplemental materials. However, they are generally not adding mark up to the actual content.

Some significant questions were raised:

- Are supplemental materials considered a part of the main article, and embedded as such? Or are the two related works? What are the consequences for the preservation record, citations, and copyright agreements?
- What about supplemental material, such as a datasets, that can be shared across different articles, and may be submitted by different authors who share or reuse the data for new content? What impact do shared data have on impact storage, citation, and so on?

- Who is the “author” of shared supplemental material content? How is authorship differentiated from “contributors”?
- How does the treatment of supplemental materials factor into the identification of horizontal versions of an article, such as print version, e-version with supplemental materials, special audience version?
- How are concerns with sharing data managed? For example, will embargoes be put in place? What permissions and/or use restrictions might be needed? How are sensitive data, such as patient data, dealt with? What about growing requirements to share data, for example, in the case of federally funded research?
- How do publishers weigh competing user needs in this arena? How do their decisions have an impact on access to content? Some readers will want to have access to as much information as possible—for example, the article, supplemental materials, related links, and so on—in one file, whereas others will want to look only at the key information available in a summary or in the article itself, without needing to work through the other content.

The group identified some general aspects related to potential Recommended Practices:

- It would be best to avoid strict definitions of what should be considered “supplemental,” because they may vary from one discipline to another and across publishers.
- It would be useful to formulate a recommendation that publishers define what formats are acceptable.
- Peer review, preservation, and interaction with repositories should be included.
- Universal agreement on how to cite supplemental materials is necessary. It could be that “citability” serves as a test for determining inclusion as a supplemental material. It may be that citing supplemental materials and citing datasets will be considered/treated differently.
- Are DOIs needed for components?
- Costs should be considered, as well as the need to educate editors about those costs. A statement along the lines of “Publishers expect to cover the costs of publishing/making available supplemental materials” might be discussed.
- A common vocabulary is needed for terms such as:
 - Dataset⁵
 - Supplemental materials
 - Supporting materials
 - Related materials
 - Article
 - Essential materials
 - Citation

Other areas to be considered are the following:

- Clear, consistent indicators that supplemental materials exist, and where to find them. These are needed for use both by readers and by systems for discoverability.

⁵ See *Ensuring the Integrity, Accessibility, and Stewardship of Research Data in the Digital Age* (by Committee on Ensuring the Utility and Integrity of Research Data in a Digital Age; National Academy of Sciences) for a definition of data. Available at http://www.nap.edu/openbook.php?record_id=12615&page=1; for purchase at http://www.nap.edu/catalog.php?record_id=12615

- Metadata needs, such as:
 - File descriptions and requirements (video, PDF, JPG, any plug-in required for viewing/use).⁶
 - Descriptive content (what it is, why it is included, and so on)
 - Bibliographic information (to support DOI registration and discovery)
- Archiving:
 - The roles of A&I services, libraries, data centers
 - Issues of continued access (for example, updating viewers) and long-term preservation.
- Clearly defined specific responsibilities of the author/publisher/editor/peer reviewers in the process of delivering supplemental materials. (An example given was “if there is an EXE file, then it should be opened/tested by the reviewer.”)

Moving Forward

It was agreed that the group would like to move forward with a defined proposal to create Recommended Practices as a joint NISO/NFAIS effort under the NISO Recommended Practice publication series and NFAIS’s Best Practice Series. NISO would be able to provide the required infrastructure support for such an effort.

Working on Recommended Practices would indicate this community’s willingness to be collaborative and demonstrate their interest in doing well by this issue in both the short- and long-term. It is hoped that the final downstream output would be recommended best practice statements to be agreed to by the major societies and publishers in this arena. These could then be separately turned into policies and incorporated into editorial handbooks.

To achieve the goal of creating such Recommended Practices, the group suggested that it would be best to establish three groups:

- *Stakeholders Interest Group*—comprising stakeholders to be kept apprised of the development of a recommended practice. Members would serve as a source of feedback on document drafts, and they would provide community vetting of a final document.
- *Business Working Group*—a small group (no more than 12–15 people) who would draft recommendations related to the semantic aspects of the Best Practices document. What are they? What are the definitions? What are the recommended roles? What are the business practices? And so on.
- *Technical Working Group*—another small group that would look at the syntactic, structural issues related to supplemental materials. Members would address issues such as syntax, linking, interoperability, markup, metadata, and so on.

The latter two groups would need to come together to formulate drafts to be presented to the Stakeholder’s Group.

⁶ See *NFAIS Best Practices for Article Publishing* for more information on file naming conventions, labeling, etc. in the section on Supplemental Materials. Available at www.nfaais.org/files/file/Best_Practices_Final_Public.pdf

Conclusion

The organizers—Todd Carpenter and Karen Wetzel of NISO, Bonnie Lawlor and Jill O’Neill of NFAIS, and Linda Beebe of APA—will draft a report to be circulated to all attendees. They will also create a structure for the groups and initiate their implementation. NISO will develop a space on the NISO website for the development of the Recommended Practices. Expected kick-off for the working groups is late March.

**NISO-NFAIS Roundtable on Best Practices for Supplemental Journal Article Materials
January 22, 2010
Washington, DC—American Psychological Association Building**

Meeting Participants

The following people participated in this meeting. Attendance was both virtual via teleconference line and in-person.

Name	Title	Organization	E-mail Address
Linda Beebe (Host/Meeting Chair)	Senior Director, PsycINFO	American Psychological Association	lbeebe@apa.org
Bonnie Lawlor (Meeting Co-Sponsor)	Executive Director	NFAIS	blawlor@nfais.org
Todd Carpenter (Meeting Co-Sponsor)	Managing Director	NISO	tcarpenter@niso.org
Stewart Wills	Online Editor	AAAS/Science	swills@aaas.org
Jon Sears	Secondary Services, AGU Publications	American Geophysical Union	jsears@agu.org
Sasha Schwarzman	Information Systems Analyst/Designer	American Geophysical Union	sschwarzman@agu.org
Carter Glass	Assistant Director, Electronic Publishing Development	American Geophysical Union	cglass@agu.org
Tshawna Byerly	Manager, Editorial Services	American Geophysical Union	tbyerly@agu.org
David Martinsen	Senior Scientist, Strategic Planning and Analysis	American Chemical Society, Publications Division	mailto:d_martinsen@acs.org
Craig Rodkin	Digital Library Production Manager, Office of Publications	ACM	rodkin@hq.acm.org
Alison Loudon	Senior Product Manager for Journals	American Institute of Physics	aloudon@aip.org

NISO/NFAIS Supplemental Journal Materials Roundtable

Name	Title	Organization	E-mail Address
Monica Mungle	Editorial Systems Manager	American Medical Association	monica.mungle@jama-archives.org
Annette Flanagin	Managing Deputy Editor	American Medical Association	annette.flanagin@jama-archives.org
Karl Elvin	Manager, Composition	American Medical Association	karl.elvin@ama-assn.org
Ken Beauchamp	Online Production Editor	The Journal of Clinical Investigation	ken.beauchamp@the-jci.org
Beverly Jamison	Director, Electronic Publishing Systems	American Psychological Association	bjamison@apa.org
Kathleen Sheedy	Serials Manager, Full-Text Production	American Psychological Association	ksheedy@apa.org
Skip Maier	Director, Journal Services	American Psychological Association	smaier@apa.org
Greg Suprock	Head of Solutions Architecture	Apex CoVantage	gsuprock@apexcovantage.com
John Kunze	Preservation Technologist	California Digital Library	jak@ucop.edu
Dave Yaron	Associate Professor of Chemistry	Carnegie Mellon University	yaron@cmu.edu
Maureen C. Kelly	Independent Consultant in e-Publishing	Consultant/CODATA	mckelly@ix.netcom.com
Chuck Koscher	Director of Technology	CrossRef	ckoscher@crossref.org
Todd Vision	Associate Director of Informatics	datadryad.com	tjv@bio.unc.edu
Jozef Paulik	Senior Product Manager, Platform and Content	Elsevier	j.paulik@elsevier.com
Adam Philippidis	Manager, Abstracting & Indexing Products	IEEE	a.philippidis@ieee.org
Bratati Biswas	Manager, Content Engineering	IEEE	b.biswas@ieee.org

NISO/NFAIS Supplemental Journal Materials Roundtable

Name	Title	Organization	E-mail Address
Linda Custor	Database Production Supervisor, Publications	IEEE	l.custor@ieee.org
Melita Dixon	Senior Administrator, Information QA	IEEE	m.dixon@ieee.org
Nancy Hollabaough	Database Production Specialist	IEEE	n.hollabaough@ieee.org
Ray Kern	Database Production Specialist	IEEE	r.kern@ieee.org
Elisabeth Moscara	Senior Information Analyst	IEEE	e.moscara@ieee.org
Christina Moxley	Associate Information Analyst	IEEE	c.moxley@ieee.org
Christina Noguerras	Department Assistant	IEEE	c.noguerras@ieee.org
Ira Polans	Project Leader	IEEE	
John Sack	Associate Publisher and Director	HighWire Press	sack@stanford.edu
Bruce Rosenblum	CEO	Inera Incorporated	bruce@inera.com
Eefke Smit	Director of Standards and Technology	International STM	eefke.smit@xs4all.nl
Lois Polese	Journal Content Management Director	John Wiley & Sons	lpolese@wiley.com
Laura Bartolo	Professor and Director, Center for Materials Informatics	Kent State University	lbartolo@kent.edu
Debbie Lapeyre	Senior XML/XSLT Consultant	Mulberry Technologies, Inc.	dalapeyre@mulberrytech.com
Carlo Nuss	Indexing Supervisor	National Agricultural Library	carlo.nuss@ars.usda.gov
Michael Esman	Head, Bibliographic Services Branch	National Agricultural Library	michael.esman@ars.usda.gov
David Gillikin	Chief, Bibliographic Service Division	NIH, Department of Health and Human Services	gillikd@mail.nih.gov
Kathy Kwan	Staff Scientist	NCBI, NIH	kwan@ncbi.nlm.nih.gov

NISO/NFAIS Supplemental Journal Materials Roundtable

Name	Title	Organization	E-mail Address
Tara Packer	Head of Author & Referee Services	Nature Publishing Group	t.packer@us.nature.com
Jill O'Neill	Director of Planning and Communication	NFAIS	jilloneill@nfais.org
Karen Wetzel	Standards Program Manager	NISO	kwetzel@niso.org
Scott Dineen	Deputy Senior Director of Publications	Optical Society of America	sdinee@osa.org
Marie McVeigh	ISI Product Development Manager	Thomson Reuters	Marie.mcveigh@thomsonreuters.com