



PORTICO

Enriching Online Journal Articles with
Supplementary Materials:
Opportunities and Challenges

SSP 2008 Pre-Meeting Seminar

Opening and Closing Remarks

Evan Owens

Chief Technology Officer, Portico

www.portico.org

This Presentation

PART I: Setting the Scene

- Some Key Questions
- Historical Examples
 - Print Supplements
 - Early E-Supplements
 - Evolving Articles

PART II: Making it Happen

- Practical Considerations
 - Changing Modes of Communication
 - Content Management Issues
 - Preservation Issues
- Future Directions
 - Scientific Data as a Special Case

Part I: Setting the Scene

Key Questions

- Is this revolution or evolution?
 - “Moving Camera”? Old wine in new bottles?
- Supplemental or essential?
 - Non-textual? Media? Not printed? Online only?
- Survey of journal editors 1999-2000
 - Ecology, Astronomy, Infectious Diseases, Anthropology, etc.
- Is more better?
 - Is it a service to reader?
 - Removing artificial limitations of print?
 - Layered approaches?
 - From articles to monographs?
- What is the role of the journal?
 - Should it be a mini-data center?
- Answers will vary by discipline
 - The most interesting aspect of this topic!

Content Examples, Page 1

Due to copyright restrictions, the e-journal examples in this section are not available for redistribution.

- **1954: Astrophysical Journal Supplement Series**

- A print publication of supplemental content: Wilson, O. C.; Abt, Helmut A. Chromospheric Structure of the K-Type Component of Zeta Aurigae. *Astrophysical Journal Supplement*, vol. 1, p.1 (1954)

- **1976 AIP Physics Auxiliary Publications Service**

- A paper repository of supplemental content; Print-only, no URL

- **1996 AIP E-PAPS Supplemental File FTP site**

- An electronic repository built for supplemental files:
<http://www.aip.org/pubservs/epaps.html>

- **2000 Current Anthropology**

- Early use of multimedia supplemental files: (1) Editorial, *Current Anthropology*, Volume 41, Number 1 (February 2000) and (2) Richard W. Stoffle et al, Ghost Dancing the Grand Canyon: Southern Paiute Rock Art, Ceremony, and Cultural Landscapes. *Current Anthropology* Volume 41, Number 1, February 2000

Content Examples, Page 2

- **2001: Astronomical Journal**
 - (1) Supplemental data tables in machine readable form and (2) large supplemental data sets rendered as graphics in PDF files: Carpenter, John M.; Hillenbrand, Lynne A.; Skrutskie, M. F. Near-Infrared Photometric Variability of Stars toward the Orion A Molecular Cloud. *The Astronomical Journal*, Volume 121, Issue 6, pp. 3160-3190 (June 2001)
- **2005: Astrophysical Journal Supplement Series**
 - Article with 1616 online figures, called a "Figure Set": K. V. Getman et al., Chandra Orion Ultradeep Project: Observations and Source Lists. *The Astrophysical Journal Supplement Series*, 160:319–352 (October 2005)
- **2005: BePress Journal Examples**
 - (1) Video Interview with questions in print PDF: Bruce Mizrach (2005) "A Video Interview of Buz Brock", *Studies in Nonlinear Dynamics & Econometrics*: Vol. 9: No. 1, Article 1. <http://www.bepress.com/snede/vol9/iss1/art1>
 - (2) traditional supplemental files with software and data files: Min-Hsien Chiang (2007) "A Smooth Transition Autoregressive Conditional Duration Model", *Studies in Nonlinear Dynamics & Econometrics*: Vol. 11: No. 1, Article 5. <http://www.bepress.com/snede/vol11/iss1/art5>

Content Examples, Page 3

- **2008: Astrophysical Journal Supplement Series**
 - Article pointing at data held entirely external to the journal: Owen W. Westbrook et al. X-Atlas: An Online Archive of Chandra's Stellar High-Energy Transmission Grating Observations. *The Astrophysical Journal Supplement Series*, 176:218–261 (May 2008)

Part II: Making it Happen

Changing Modes of Communication

- New conventions take time to emerge
 - Best practices arise from experience
- Adrian Johns, *The Nature of the Book*
 - Closing speaker at 2008 SSP Annual Meeting
- ENTS (Essential Non-Text Stuff) 2004
 - Study group organized by AIP
 - Key Recommendations:
 - Each version (print/online) should be understandable on its own
 - File formats should be chosen for longevity and availability
 - Print versions should have representations of non-print components
 - Additional guidance regarding color, audio, animations, videos, interactive illustrations...
 - Concern regarding size of files and bandwidth requirements
 - Project renamed EIO (Enhanced Information Objects)
- Have times changed? How much?
 - Do we worry that reader of printer is missing online stuff?
 - Is anyone reading the print?
 - Bandwidth concern is a different order of magnitude?

Content Management Issues

- Is it in the custody of the publisher?
 - Early experiments with supplemental content on author's home pages
 - Entire articles not on publisher's web site
 - Robust links to data centers as an option
- Is it processed in the same way as the article proper?
 - Peer reviewed?
 - Edited and proofed? By the author and the publisher?
 - Same or different workflow?
- Is it truly part of the article or something separate?
 - Links between HTML, PDF, XML and supplemental files can be problematic
 - Post-publication additions by authors
 - A completely new mode of publication
- Does your content management system really support this?
 - From author to reader (peer review, production, delivery, archiving)
 - Risk of error or loss
 - Horror stories abound

Preservation Issues

- Good content management is key
 - If you can't find it, you haven't preserved it!
- What formats and technologies?
 - Supplemental text in MS-Word? PDF? XML?
 - Data in proprietary and/or discipline-specific formats?
- Best formats are
 - Widely used
 - Well documented
 - Have multiple independent tools that support the format
- How much discipline-specific knowledge is needed?
 - Syntax versus semantics
 - Does this point us toward data in data centers?
- Are we assuming that supplemental information has a shorter lifespan than the article proper?
 - Caveat lector, as it were

Scientific Data as a Special Case

- Where does data really belong?
 - With the journal?
 - Discipline-based data centers?
 - Institutional repositories?
- International Digital Curation Conferences
 - Started in 2004
 - Sponsored by UK Digital Curation Centre and JISC
- Data trends:
 - Contrasts between “Big Science” and “Small Science”
 - Amount of data is increasing in all disciplines
 - Small science will soon be producing data in much larger quantities
- What will the impact be on publications?
 - Is the journal article the best vehicle for distribution and access to larger and larger amounts of data?
 - Answer will vary by discipline
- Are supplemental files already passé?

Is the Camera Moving?

- From (broadly) generic publishing to highly discipline-specific
 - From syntax of articles (section, paragraph, heading)
 - To semantics of information (disease, observation, measurement)
- “Supplemental” files are a barometer of change
- Interesting times ahead!