

Emerging Technical Tools

Moderator:

Gregory H Suprock
Nature Publishing Group
New York, New York

Panelists:

Mark Doyle
American Physical Society
Ridge, New York

Brian McMahon
International Union of
Crystallography
Chester, England

Reporter:

Laura Altobelli
American College of Physicians
Philadelphia, Pennsylvania

With reading no longer a passive activity, technology increasingly defines the momentum of scientific publishing. Authors, reviewers, readers, and libraries all must find appeal in new strategies for communicating research findings. This session's panelists introduced the audience to some of the emerging tools gaining popularity.

The hunger for raw data has led to the publishing of more research directly at its source, and some institutes make dataset repositories available through open access. Brian McMahon, research and development officer for the International Union of Crystallography (IUCr), explained, "Fields like crystallography need tight integration of data with discourse for the purpose of validation and quality control." Consequently, IUCr's Crystallography Journals Online (journals.iucr.org) now requires that any author who is reporting a new crystal structure submit all data accumulated during research in a crystallographic information file (CIF) annotated with comments.

The CIF format enables the author to

provide data deposited by instruments and allows reviewers to use an automated program for identifying major outlying data. As a result, inconsistencies can now be resolved more simply before peer review, and authors can easily annotate the data to highlight interesting new discoveries. McMahon explained that CIF files also aid readers, who are able to download data to a viewing program of their choice and build three-dimensional representations of crystallographic and molecular structures. The advance carries far-reaching implications; chemistry and physics publications are likely benefactors.

Mark Doyle, of the American Physical Society, suggested that publishers investigate several other emerging file types that are simplifying archiving and providing users improved access to figures and data. The new JPEG2000 format, capable of storing multiple image resolutions in a single file, gives users more options for creating, viewing, and storing graphics. MPEG21 uses extensible markup language (XML) to create one complex object containing text, figures, and even full datasets.

Large volumes of data are of greatest use when distributed universally in a timely fashion. Digital archiving under the Open Archives Initiative Protocol for Metadata Harvesting (OAI-PMH) reflects an effort to broaden distribution by bridging the gap between isolated digital library collections. The OAI-PMH initiative unites libraries in much the same way that the Internet united computer networks. It is the same protocol used by the open-access crystallographic data repositories and several publishing houses and has the potential to redefine the nature of online publishing.

According to Doyle, publishers are also making greater use of RSS, or Really Simple Syndication, to keep the scientific community informed of new research. This simple XML-based format allows short, frequently updated newsfeeds to be

disseminated from scientific Web sites. Syndication gives users access to highlights of new articles, often before the print publication date. Doyle also spoke of the emergence of wiki server software, which has given more power to users to create and edit Web content; this enables quick "publication" of new data.

Unlike other Web-based forms of open content creation (like message boards and forums), wikis offer the distinct advantage of enabling users to reorganize content continually by using keyword tags, and this enhances searching. Another user-driven approach, the folksonomy (also known as "social bookmarking"), is characterized by groups of people who cooperate spontaneously to organize information into categories. "Because the organizers of the information are usually its primary users", Doyle explained, "folksonomies more accurately reflect the population's conceptual model of the information." Both wikis and folksonomies avoid the drawbacks of hierarchic presentation of information on sites that don't evolve fast enough to accommodate the addition of new material. "Tags keep found things found", Doyle said.

Doyle and McMahon recommended several Web sites at which users can see some of the emerging technologies put to use:

- www.flickr.com—a simple photo-sharing site that makes use of keyword tags.
- del.icio.us—manages social bookmarks to enable users to create a collection of links organized by keywords.
- www.citeulike.org—allows academicians to organize and share papers.
- www.connotea.org—enables scientists to create and share "libraries" of links to frequently used references. 