

**Report to SfN Council
August 8, 2007**

**PubMed Plus: New Directions in Publishing and Data Mining
Leadership Conference**

June 18-19, 2007
Eric P. Newman Education Center
Washington University Medical Center
St. Louis, Missouri

Supported by the Society for Neuroscience, the NIH Blueprint for Neuroscience Research, the International Neuroinformatics Coordinating Facility, and Science Commons

Organizers: David Van Essen, President, Society for Neuroscience, and
Robert Williams, Chair, Neuroinformatics Committee

Moderators: Floyd Bloom and David Van Essen

OVERVIEW

The pace at which experimental results are generated in biomedical research has increased dramatically, thanks to recent advances on scientific, technical, and computational fronts. In parallel, the growth of online journals, databases, and powerful search engines have led to enormous changes in the ways that scientific findings are communicated, archived, and used in new research. The scientific community stands to benefit greatly by actively planning how best to exploit informatics technologies to further accelerate the pace of discovery through more effective communication and collaboration.

In October 2006, SfN Council approved a proposal to organize a leadership conference in June 2007 to discuss how neuroscience journals and databases might operate more collaboratively, to permit more powerful mining of text and data published in the neuroscience literature. Participants in the resulting PubMed Plus conference included working scientists, journal editors and publishers, experts in computer science and semantic web technology, librarians, and representatives from scientific societies and funding agencies; many, but not all, have some association with neuroscience.

The proposal to Council defined four major topics for discussion at Pubmed Plus, and invitees were assigned to corresponding working groups several months before the meeting. By the time of the June conference, most invitees already had participated in informal discussions of at least one of the following topics [[Exhibit 1. PMP Working Groups](#)]:

- 1) Capturing metadata to enhance research retrieval and data mining;
- 2) Improving the linkages between journal articles and data repositories;
- 3) Sustainability and standardization of journal supplementary materials and associated databases;
- 4) Feasibility of sharing manuscript reviews among neuroscience journals with different publishers.

Each group was tasked with generating a set of desired outcomes for their specific topic, including both "low-hanging fruit" (specific goals that can be implemented within 2 years) and long-term objectives (that can be implemented within the next 2-10 years).

The specific recommendations that emerged from the four working groups are summarized below.

Working Group 1: Capturing metadata (key descriptors of data and design) to enhance research retrieval and data mining

- Encourage neuroscience journals to develop mechanisms for gathering common types of metadata on experimental methods using a process that is not onerous to authors;
- Define core experimental metadata to be collected (minimally, species, gender, and experimental technology used);
- Encourage development of text-mining tools that automatically extract major data categories prior to manuscript publication;
- Encourage journals and software developers to work together to make stored metadata searchable using emerging semantic-web tools;
- Encourage journal editors and publishers to make full text available for data mining.

Working Group 2: Improving the linkages between journal articles and data repositories

- Request that the 'LinkOut' feature in PubMed be modified so that a prominent icon appears next to publications associated with data in an external database;
- Encourage journal to incorporate concise 'anchor links' that link online articles to specific datasets within an external database;
- Encourage journals to promote use of '3D PDF' files and other resources that enhance the information content of journal articles;
- Encourage journal editors and publishers to improve data descriptions in figures and tables to facilitate data mining by humans and machines;
- Encourage authors to submit their data to a database whenever appropriate: The SFN Neuroinformatics Committee or some other entity could serve as an honest broker, promoting awareness of suitable databases directly to authors and indirectly through editors and reviewers;
- Encourage federal funding agencies to include data sharing as an issue to be addressed during the review process.

Working Group 3: Sustainability and standardization of journal supplementary materials and associated databases

- Encourage journals to adopt common standards for 'supplementary material' for peer reviewed materials;

- Encourage adoption of a new category called ‘related material’ that includes non-peer reviewed materials;
- Encourage publishers to insure that links from supplementary material and related material back to the online article are consistently available;
- Recommend that publishers affirm intellectual property principles, i.e., publishers do not own data published as supplementary material; depositing data in a database does not preclude later publication; readers have the right to mine data in supplementary material;
- Recommend that publishers insure the sustainability of supplementary material and make public statements of their preservation policies.

Working Group 4: Feasibility of sharing manuscript reviews among neuroscience journals with different publishers

- Establish a pilot Neuroscience Publishing Consortium, initially sponsored by the Society for Neuroscience, with journals represented at PubMed Plus initially invited to join;
- Develop a system of cascading submissions, in which reviews from one journal in the consortium could be passed to the next journal if the paper were not accepted by the first journal, and test this system over the next year;
- Identify a task force charged with implementing this plan (WG 4 members are willing to serve in this role);
- If the Consortium proves successful, formalize its existence, invite other journals to join, and move sponsorship to a permanent “home” (the International Neuroinformatics Coordinating Facility in Stockholm has been suggested);
- If the Consortium proves successful, develop a front end system to automate the review sharing process.

***Staff Note:** Working Group 4’s recommendations subsequently were enthusiastically endorsed by the Publications Committee. A separate, detailed formal proposal to establish a Neuroscience Publishing Consortium has been submitted jointly to Council by the Chairs of the SFN Neuroinformatics, IT, and Scientific Publications Committees, and the current and incoming Editors-in-Chief of The Journal of Neuroscience. Jan Bjaalie, NIC member and Executive Director of the INCF, headquartered in Stockholm, has offered to provide computational and administrative support for this effort.*

SUMMARY

The PubMed Plus Leadership Conference addressed a key aspect of the Neuroinformatics Committee charter by positioning SfN as an influential “honest broker” in a dialogue on journals and data sharing that successfully engaged the broader information community. By bringing together a diverse group of experts who normally don’t attend the same meetings, SfN was able to promote frank and constructive conversations about how to increase the usefulness and sharing of data published in online journals and databases, and to generate considerable enthusiasm and momentum around these issues [[Nature Neuroscience, 10\(8\): 931, August 2007](#)].

The PubMed Plus working groups defined many key issues and generated many specific recommendations. Some are straightforward ‘bite-sized’ steps, while others are more ambitious and might take years to implement. An underlying philosophy was to encourage but not to

coerce changes that will benefit the neuroscience community as well as the broader biomedical research community.

For many of the recommendations, implementation will fall primarily in the domain of stakeholders other than SfN and can be carried out without SfN's approval or action. For example, a subgroup of leading journals is already considering next steps in developing standards for supplementary material and data deposition. This resonates with SfN's objective that the PubMed Plus meeting should catalyze diverse efforts among a broad range of stakeholders. On the other hand, continued progress on many recommendations will require, or at least benefit from, SfN's continued involvement.

In a post-meeting survey, participants reported that the meeting was highly productive and stimulating ([Exhibit 2](#)). Many participants commented that SfN had distinguished itself in by assuming a leadership role and by taking a long and broad view on the issues. They expressed the desire to build on the PubMed Plus meeting, and recommended that SfN continue to play a lead role in addressing these issues ([Exhibit 3](#)).

Remarkably, Working Group 4 was able to complete its deliberations and generate a final proposal in a very short time frame. The ideas and recommendations of the other working groups will in general need additional development before they are ready for implementation as concrete action items.

The organizers hope that Council will consider the ideas raised at PubMed Plus as worthy of continued attention. If so, a next step might be for Council to ask the Neuroinformatics Committee to form task forces that include working group volunteers, which would prioritize and develop recommendations via "virtual" meetings (online workspace, conference calls). Concrete proposals that emerge could be funneled through NIC (and other committees, if appropriate) to Council. For example, a task force involving Working Group 1 might recommend specific types of experimental metadata that could be collected by *The Journal* using the new manuscript submission/peer review system under development. If endorsed by the Scientific Publications Committee and approved by Council, the Central Office staff could proceed with implementing this process and communicating it to authors, editors, and other journal publishers.