Editorial **Comments**

JAMIA

Human Brain Program Research Progress in Bioinformatics/ Neuroinformatics

The Human Brain Program (HBP) organized by the Office of Neuroinformatics of the National Institute of Mental Health, National Institutes of Health, is a broadly based federal research initiative sponsored, in a coordinated fashion, by 15 federal organizations from four federal agencies (http://www.nimh.nih.gov/neuroinformatics/index.cfm). The scientific goals of this initiative are to enable the progress of neuroscience research through the creation of a Web-based set of distributed and federated databases, analytical and modeling tools, and simulators.

Currently, neuroscientists collect complex data in ever-increasing amounts, fostering increased specialization, with resultant challenges to integrate data between and across levels of interaction, control, and function. The sheer quantity and complexity of the data are such that the field of neuroscience would benefit considerably from an information management system for its experimental data. The field should enhance its wealth of ever-increasing empirical data, accumulated from its many disciplines and experimental approaches, by developing appropriate databases and a greater capability for both theory development and simulation models.

The focus section in this issue of *JAMIA*, on HBP research progress in neuroinformatics, collects three representative works of HBP grantees on the tools and methodology to support the bioinformatics aspects of neuroscience research.

The purpose of analyzing brain ultrastructure is to understand the normal synaptic communication pathway of neurons and supporting cellular elements and the alternations of such pathways and cellular elements caused by diseases. The prevalent analytic approach is based on one or paired sections and the use of electron microscopy. Working with volumes rather than single sections, however, provides a richer and more accurate representation of brain ultrastructure. Recent advances in informatics are making volume reconstruction and three-dimensional analysis of brain ultrastructure increasingly practical and cost effective.

The research reported by John Fiala and Kristen Harris[1] describes such a new reconstruction system and new algorithms capable of operating on personal computers for analyzing in three dimensions the location and ultrastructure of neuronal components, such as synapses. Specifically, a volume of brain tissue from stratum radiatum of the hippocampal area CA1 is reconstructed and analyzed for synaptic density to demonstrate and compare the techniques. On the basis of the findings, these authors also propose general rules for performing synaptic density analysis on reconstructed volumes of brain ultrastructure.

The amount of neuroscience information increases significantly as the HBP and related clinical research move forward. The paper by Daniel Gardner et al.[2] addresses one critical issue in interoperability, i.e., facilitate neuroscience research and information exchange through a common data model (CDM), which is generalized from two prototype neurophysiology databases. These authors apply the emerging extensible markup language (XML) standard to implement data exchange between any CDM-derived data models. The authors state nine design goals that meet the essential criteria of a good data

model, and adopt the data-driven design philosophy, which, they explain, is well suited to brain information. The CDM has five root classes, or "super classes"—data, site, method, model, and reference—that are a reasonable abstraction of the problem domain. The authors believe that when fully implemented using biophysical description markup language (BDML), as well as the hierarchic attribute value implementation of controlled vocabulary, the proposed CDM will have the capacity to mediate among disparate neuroscience database projects and similar resources with compatible data and data models.

The experimental work of the Yale research group on neuroscience data analysis provides yet another perspective on data modeling and knowledge representation. The traditional paradigm of reductionism works reasonably well in physical sciences, but the biological domain requires a different research paradigm. A major goal for neuroscientists is to have a sound theoretic foundation that can cut across multiple biological levels, from the genetic level up through the synaptic, neuronal, network, and brain pathway levels, ultimately reaching the behavioral level. The ability to capture, relate, and analyze diverse types of data at multiple levels of abstraction is one of the central challenges of neuroinformatics.

The work reported by Perry Miller et al.[3] on the integrated data analysis of a particular model system of neurobiology, i.e., the olfactory system, explores new paradigms of biological research to study new relations among findings obtained at individual levels of representation. Their research paper provides an overview of SenseLab; in particular, the introduction of a flexible data model called EAV/CR (entity-attribute-value with classes and relationships) to integrate olfactory and associated data of four different databases for neuroscience research and clinical studies later on. The modeling overhead of EAV/CR, however, may not be suitable to model large amounts of very homogeneous data that have been mass-produced by high-throughput instruments, such as microarray analyzers. Although the current focus of SenseLab's activities is on basic neuroscience and neuroinformatics research, the work has a range of potential clinical and other real-world applications.

We hope that the ideas and results reported in this focus section will suggest new and better ways to develop tools to support neuroinformatics research and will lead us to the next generation of bioinformatics tools and management systems. Another HBP focus section, on the progress of enabling technology for brain imaging aspects of neuroscience research, will appear in a future issue of the Journal.—STEPHEN T.C. WONG, STEPHEN H. KOSLOW

References ■

1. Fiala JC, Harris KM. Extending unbiased stereology of brain ultrastructure to three-dimensional volumes. J Am Med Inform Assoc. 2001;8:1–16.
2. Gardner D, Knuth KH, Abato M, et al. Common Data Model for neuroscience data and data model exchange. J Am Med Inform Assoc. 2001;8:17–33.
3. Miller PL, Nadkarni P, Singer M, Marenco L, Hines M, Shepherd G. Integration of multidisciplinary sensory data: a pilot model of the Human Brain Project approach. J Am Med Inform Assoc 2001;8:34–48.

Affiliation of the author: University of California, San Francisco, California (STCW) and National Institute of Mental Health, Bethesda, Maryland (SHK)..

Correspondence and reprints: Stephen T.C. Wong, PhD, PE, UCSF, Department of Radiology, Box 0628, UCSF 505, Parnassus Avenue, San Francisco, CA 94131; e-mail: <swong@radiology.ucsf.edu>.

Received for publication: 9/27/00; accepted for publication: 9/28/00.

The Challenge of Bridging Between Disciplines

The September 2000 issue of *JAMIA* presented three papers on the intersection of bioinformatics and biomedical informatics.[1–3] That set of papers was written by individuals who are grounded in biomedical informatics and working in areas of bioinformatics. Their papers explain aspects of bioinformatics research in the language of biomedical informatics.

This issue of *JAMIA* contains the first of two sets of papers[4] coming out of the Human Brain Project . With one exception, these papers are written by people who are grounded in neuroinformatics (the intersection of neuroscience and bioinformatics). They are working on problems that are analogous to the ones being tackled by researchers in biomedical informatics. Nonetheless, they come at these problems from a different perspective. Although they use many of the same words to describe what they do, the careful reader with a background in biomedical informatics will find that the words mean something different to them. For example, Gardner et al.[5] use metadata as "the neurobiological descriptors characterizing neurophysiology datasets." In other words, their metadata describe the experimental context in which the data were acquired. This difference in use is explicitly stated in the text, but it may be missed by someone who has a different expectation about the meaning.

The review and revision process was lively. As Editor-in-Chief, I suggested that the work be reframed to communicate clearly to the journal's audience. The authors disagreed, pointing out that these were the primary archival publications about this work and that they had to communicate clearly to people working directly in their area. On reflection, I came around to their view.

We are fortunate to have these papers in the mainstream of biomedical informatics. Otherwise, most of our readers would not get an in-depth view of this important area of research. But, read the papers carefully. If you see a statement that you think is naive, you have probably come across a place where the author uses words differently than you do. After you finish the papers, think about how you might apply the techniques to problems in your area of work. I think you will find the extra effort worthwhile. If you take the time to bridge the communication gap, you always learn more from people tackling problems from a new perspective.—WILLIAM W. STEAD, MD

References ■

1. Miller PL. Opportunities at the intersection of bioinformatics and health informatics: a case S=study. J Am Med Inform Assoc. 2000;7:431–8.
2. Altman RB. The interactions between clinical informatics and bioinformatics. J Am Med Inform Assoc. 2000;7: 439–43.
3. Kohane IS. Bioinformatics and clinical informatics: the imperative to collaborate [editorial]. J Am Med Inform Assoc. 2000;7:512–6.
4. Wong S (Editorial). Human Brain Project Research Progress in Bioinformatics/Neuroinformatics. J Am Med Inform Assoc. 2001;8:105–6.
5. Gardner D, Knuth KH, Abato M, et al. Common Data Model for neuroscience data and data model exchange. J Am Med Inform Assoc. 2001;8:17–33.

Affiliation of the author: Vanderbilt University, Nashville, Tennessee.

Correspondence and reprint requests: William W. Stead, MD, Associate Vice Chancellor for Health Affairs, Eskind Biomedical Library, Vanderbilt University Medical Center, 2209 Garland Avenue, Nashville, TN 37232-8340; e-mail: <bill.stead@mcmail.vanderbilt.edu>.

Received for publication: 9/27/00; accepted for publication: 9/28/00.