



Biological Informatics: A Comparison of Biodiversity Informatics and Neuroinformatics

by Bryan Heidorn, Guest Editor

This special section of the *Bulletin of the American Society for Information Science and Technology* is devoted to two branches of biological informatics that warrant special attention from the information science community: biodiversity informatics and neuroinformatics. Biologists are turning to information technology to produce critically needed efficiencies in their work, a hope conveyed by E. O. Wilson (see box). Unfortunately, most of the information science community is unfamiliar with either of these fields of study so definitions and explanations are in order.

The relationship between *bioinformatics* and the term that we have chosen to use here, *biological informatics*, is not as subtle as it might seem. Over the past 10 years *bioinformatics* has come to mean “information on molecular biology” and in particular gene and protein sequences. This use of the term in the popular press, associated with the great progress and successes in that field, has served to cement this definition into the psyches of the general population and scientists alike. As a result, a new term (*biological informatics*) to cover the science of information about all levels of biologi-

P. Bryan Heidorn is assistant professor, GSLIS, University of Illinois, Urbana-Champaign. He can be reached by mail at 501 E. Daniel St., Champaign, IL 61820; by phone: 217-644-7792; or by e-mail: heidorn@alexia.lis.uiuc.edu

“New electronic technology, increasing exponentially in power, is trimming the cost and time required for taxonomic description and data analysis. It promises to speed traditional systematics by two orders of magnitude. What is lacking and needed now is a concerted effort, comparable to the Human Genome Project (HGP), to complete a global biodiversity survey — pole to pole, whales to bacteria, and in a reasonably short period of time.” (Wilson, E.O. (2000). A global biodiversity map. *Science* 289, p. 2279.)

cal analysis was in order. Health informatics (medical informatics), neuroinformatics, biodiversity informatics and biomolecular informatics (bioinformatics) all fall under this broader concept. Because of the length of this special section dealing with biological informatics, it will be split between this issue and the next one.

Both biodiversity informatics (BDI) and neuroinformatics (NI) are brought together into one category because, while they differ broadly in other aspects of their social, political and historical context, it has become clear that progress in both fields is dependent on advances in information gathering and access. Both have reached the point of becoming “integrative sciences.” In this issue of the *Bulletin*, the paper “Toward Integrative Science,” by Melissa Cragin, addresses how expansion and fragmentation of knowledge has led to the opportunity and need for integration. This integration and sharing of data is a new concept for many of the scientists at the center of this process. Geoffrey Levin and Melissa Cragin address the role of data gathering and integration across laboratories and institutions in “The Role of Information Science in Gathering Biodiversity and Neuroscience Data.”

Progress has been made in biological informatics, including the formation of the Global Biodiversity Information Facility (www.gbif.org). Meredith Lane discusses this facility and related global projects. Neuroscience has a similar push for integration of information to answer new questions that could not be addressed using the traditional paradigm of disposable data.

Both BDI and NI are marked with a critical need for advancement, an abundance of new research and new sources of information, but a lack of a cultural and technical infrastructure to share and use this abundance of information. As is true in many fields, electronic media and the Internet are changing the nature of publishing in BDI and NI. Publishing projects now can include hundreds of authors spread across

the globe. In endeavors like *The Flora of North America*, *The Flora of Australia* and *The Flora of China* the goals are frequently to document all species in a geographic region or all species in a particular group such as the grasses. In either case the species can number in the tens of thousands. Information science needs to develop information management methods to help scientists coordinate their efforts to document all species on earth. These same forces of scale and integration are breaking down the culture of “one lab, one data collection.” Bryan Heidorn will address the promise and perils of this change in “Publishing Digital Flora and Fauna” in the December/January 2004 issue while Gwen Williams, in the paper “Intellectual Property and Biological Knowledge,” will discuss how the shift to electronic delivery has made issues of intellectual property in biological science more complicated.

Assuming that we can solve all of the above issues, we still have a major problem with the physical communication of biological information. Projects springing up around the world are being designed to use existing information retrieval standards to record, encode and deliver biological information. Chulin Meng will cover some of the emerging standards with familiar sounding names but new twists, such as “Darwin Core” (tsadev.speciesanalyst.net), Z39.50 (ZBIG) and Open Archives Initiative style treatment of natural history style holdings in “Biological Information Standards” (December/January 2004).

We are still just beginning a biological science revolution. Global environmental change, extinction, alien species invasion, over fishing and many other natural and man-made phenomena are making it essential that we understand our world’s biodiversity so that we can preserve the pieces that we need for survival. Advances in neuroscience are leading the way to our understanding normal brain function and treatment of human diseases of the nervous system. Information science and technology will play a central role in this revolution.