

NO-BOUNDARY THINKING IN BIOINFORMATICS

JASON H. MOORE

*Institute for Biomedical Informatics
University of Pennsylvania
Philadelphia, PA 19104, USA
Email: jhmoore@upenn.edu*

STEVEN F. JENNINGS

*Sector3 Informatics
Marana, AZ 85658, USA
Email: drsfjennings@gmail.com*

CASEY S. GREENE

*Department of Systems Pharmacology and Translational Therapeutics
University of Pennsylvania
Philadelphia, PA 19104, USA
Email: csgreene@upenn.edu*

LAWRENCE E. HUNTER

*Computational Bioscience Program,
University of Colorado School of Medicine
Aurora, CO 80045, USA
Email: larry.hunter@ucdenver.edu*

ANDY D. PERKINS

*Department of Computer Science and Engineering,
Mississippi State University
Jackson, MS 39762, USA
Email: perkins@cse.msstate.edu*

CLARLYNDA WILLIAMS-DEVANE

*Department of Biological and Biomedical Sciences,
North Carolina Central University
Durham, NC 27707, USA
Email: clarlynda.williams@ncsu.edu*

DONALD C. WUNSCH

*Department of Electrical and Computer Engineering
Missouri Science and Technology University
Rolla, MO 65409, USA
Email: dwunsch@mst.edu*

ZHONGMING ZHAO

*Center for Precision Health
University of Texas Health Science Center
Houston, TX 77030, USA
Email: zhongming.zhao@uth.tmc.edu*

XIUZHEN HUANG

*Department of Computer Science
Arkansas State University
Jonesboro, AR 72467, USA
Email: xhuang@astate.edu*

1. Bioinformatics is a Mature Discipline

Bioinformatics had its origins in the 1970s with the convergence of DNA sequencing, personal computers, and the internet. The field rapidly evolved as biotechnology improved making it critical to store, process, retrieve, and analyze bigger and bigger data to address important questions in the biological and biomedical sciences. Bioinformaticians throughout the 1980s and 1990s were often seen as consultants that provided a data service that represented one step in the process of asking a question, formulating a hypothesis, carrying out an experiment, analyzing the results, and making an inference. Much of bioinformatics at that time was about developing the capacity for providing this service. As the discipline matured in the 2000s it quickly became apparent that bioinformaticians were needed as collaborators and not just consultants. This facilitated the integration of bioinformatics into every aspect of a research project. We are at yet another turning point in the evolution of bioinformatics that will see in the coming years bioinformaticians transition from collaborators to leaders that bring interdisciplinary teams together to solve a complex problem. In other words, bioinformaticians will be able to ask the questions, define the hypotheses, and orchestrate the scientific study. This is the natural result of interdisciplinary training, the public availability of data, open-source software, the widespread availability of core facilities for conducting experiments and, importantly, the ability to integrate and synthesize knowledge sources to ask more impactful questions.

2. The Golden Era of Bioinformatics Has Begun

The turning point in the maturity of bioinformatics as a discipline has led some to speculate that we are entering a golden where the focus on computational approaches to biomedical research will be front and center¹. There are several reasons for this speculation. First, big data is now the norm rather than the exception and computational methods are critical for successful storage, management, retrieval, analysis, and interpretation for answering scientific questions. Bioinformatics has never been so important for moving research forward. Bioinformatics areas such as databases, machine learning, and visualization are in high demand. Second, high-performance computing (HPC) is inexpensive and widely available in different technologies such as cloud computing and parallel computing using graphic processing units (GPUs) that bring thousands of compute core to a single desktop computer. Third, artificial intelligence and machine learning have matured and are now routinely being used to solve complex problems in the biomedical sciences. This is the result of decades of research on intelligent algorithms and software and the HPC resources necessary to apply them to big data. Fourth, the power of combining computational intelligence with statistical methods has emerged in the form of data science that allows the integration of different philosophical and quantitative schools of thought to solve biomedical problems. Fifth, visual analytics that brings visualization technology together with data science and human-computer interaction is maturing quickly with areas such as virtual reality, augmented reality, and 3D printing. Visual analytics will be essential for allowing human to interact with and understand data and research results that are too big and too complex to understand. Sixth, data and knowledge integration is maturing quickly as we have seen with electronic health records, data warehouses, and knowledge resources such as PubMed. Seventh, there is an increasing recognition of the importance of bioinformatics by federal funding agencies, biotechnology and pharmaceutical companies, and academic institutions. Investment in bioinformatics personnel and technology has never been greater and is expanding quickly. Now is the time for bioinformatics to have a substantial impact on biological and biomedical research.

3. No-Boundary Thinking in Bioinformatics

The purpose of this workshop is to introduce and discuss the future of bioinformatics as a mature discipline. We have previously defined this evolution and its impact as No-Boundary Thinking (NBT) in Bioinformatics^{2,3}. The NBT philosophy provides bioinformaticians with the unique opportunity to move past being service providers to asking and answering research questions. This is because they

are in the best position to integrate and synthesize knowledge across many disciplines to articulate a question that might have broader impact than one formulated from the knowledge of a single discipline. This allows them to be an equal contributor to the motivation and design phases of research studies. NBT puts the emphasis on knowledge-based question definition with big data serving a secondary and supporting role. This is counter to the current philosophy of letting big data drive the questions that are asked³. The workshop will introduce and define the NBT approach and will provide several scientific examples. An important component the workshop is providing examples of how NBT can be moved into the classroom to prepare bioinformatics students for a future where they are leading scientific studies. Panel discussions around NBT in science and education will allow for a robust discussion about these new ideas.

References

1. J.H. Moore, J.H. Holmes, *BioData Mining* **9**, 15 (2016).
2. X. Huang, B. Bruce, A. Buchan, C.B. Congdon, C.L. Cramer, S.F. Jennings, H. Jiang, Z.Li, G. McClure, R. McMullen, J.H. Moore, N. Nanduri, J. Peckham, A. Perkins, S.W. Polson, B. Rekepalli, S. Salem, J. Specker, D. Wunsch, D. Xiong, S. Zhang, Z. Zhao, *BioData Mining* **6**, 19 (2013).
3. X. Huang, S.F. Jennings, B. Bruce, A. Buchan, L. Cai, P. Chen, C.L. Cramer, W. Guan, U.K. Guan, U.K. Hilgert, H. Jiang, Z. Li, G. McClure, D.F. McMullen, B. Nanduri, A. Perkins, B. Rekepalli, S. Salem, J. Specker, K. Walker, D. Wunsch, D. Xiong, S. Zhang, Z. Zhao, J.H. Moore, *BioData Mining* **8**, 7 (2015).