Mapping the Boundaries of Bioinformatics

Lily Farris and Janet Atkinson-Grosjean

Introduction

How can we characterize bioinformatics? Is it a discipline, a sub-discipline, a profession, a methodology, a field of research, a field of practice, or simply a set of research tools and techniques? Should bioinformaticians be classified as scientists or as technicians? Do the cultures of biology and computer science become hybridized in bioinformatics? Or, do bioinformaticians remain culturally identified with their original training? In order to develop an understanding of these questions, we conducted participant-observation to explore the role of bioinformatics within a large-scale research network. This brief paper describes our approach and preliminary findings. We argue that the processes through which bioinformatics informs biological research signals a change, first, in approaches to biological research (e.g. from "wet" to "dry") and second, in the skill-sets that future researchers must develop to frame and answer biological questions. We argue that examining the impact of emerging fields, such as bioinformatics, may challenge taken-for-granted assumptions about the development and outcomes of science.[1]

Bioinformatics in Practice

Bioinformatics is classically understood as the use of computational and mathematical methods to solve issues in molecular biology. A broader definition recognizes bioinformatics as a property of the convergence of large-scale biosciences and ICTs[2] in the post-genomic era. The accelerated production of raw biological data has increased the need for computational techniques and systems that can manage, collect, store, analyze, manipulate, and interpret vast datasets of genomic information and molecular pathways.[3] The tools of biological science, information science, and computer science — harnessed together in bioinformatics — has made it possible to generate new scientific insights from raw data.[4]

In order to arrive at such insights, however, and as a necessary component of their scientific work, those who practice bioinformatics must constantly negotiate the challenges of disciplinary heterogeneity and jurisdictional expansion. By heterogeneity we mean the disciplinary and cultural diversity characteristic of large-scale research teams; such teams need to forge common goals which transcend these boundaries.[5] By jurisdictional expansion we refer to the boundary-work[6] through which bioinformatics competes with biology and computer science for control of knowledge stocks.

In our empirical work we examined the development of a curated biological database as a case of bioinformatics as civic translation.[7] We observed distinctive modes of interaction, communication and purpose among team members and concluded that the heterogeneity of the project team and development of common understandings was central to the successful completion of the work.

The case of InnateDB

Our interest in the practice of bioinformatics has been shaped by our work to understand the conduct of scientific inquiry and the logic of disciplinary boundaries[8] in the development of a particular type of biological tool, the manually curated public database. The bioinformatics team we studied was working to create InnateDB, a database of genes, interactions, and pathways involved in the innate immune response, and the bioinformatics tools needed to integrate, analyze and visualize the data.[9] InnateDB, is a component of Genome Canada's Pathogenomics of Innate Immunity [PI.2] project, intended to make a significant civic contribution to the research community studying innate immunity. The InnateDB team is embedded within the wider PI.2 research team in the sense that the database builders interact with and receive direct requests and feedback from the wider project community. The high degree of embeddedness distinguishes this particular bioinformatics approach from other models.[10] Knowledge stocks of team members range from exclusive training in computer science to advanced degrees in biology with the recent addition of qualifications in computational biology.

In order to track the database development process we installed a participant observer within the InnateDB team thereby achieving first-hand awareness of the diverse perspectival intersections that characterize the practice of bioinformatics. Goal coherence, milestones, timelines, and role definitions varied according to the stage of the database development process and the personal development of team members as they continued to be involved. Perspectival differences go a long way to explaining the challenges faced by research teams from diverse disciplinary backgrounds. Each team member will interpret the direction of their work slightly differently; this will impact their ability, willingness and understanding of the tasks to be completed.

Discussion

The heterogeneous team working to develop InnateDB had to function cooperatively in order to produce a resource that will aid and expand the work of "wet lab" scientists. To understand the process of team cooperation, it is necessary first to understand the perspectives that team members bring to the table.

InnateDB's developers are attempting to combine methods and knowledge stocks[11] from distinct scientific jurisdictions: computer science, mathematics and biology. While common objectives were established for the project as a whole, each team member approached the completion of their particular contribution with a unique set of problem-solving tools imported from their home disciplines. Their characteristics, values, goals, methods, capabilities, expertise, and understandings of other knowledge cultures also varied according to their disciplinary legacies. These distinguishing factors may challenge agents from different disciplinary domains to claim jurisdiction over institutional turf and knowledge stocks, since the power an individual wields within such a heterogeneous team may be specific to the intellectual field in which they function.[12]

When a community's common goal is knowledge production, determination of who is most capable of contributing to the objective lends significant weight to the chosen discipline.[13] The case of bioinformatics drives home the reality that intellectual niches or ecosystems are not permanent but evolve alongside the scientific projects in which agents are embedded.

Bioinformatics has become such an essential component of scientific work that this hybrid of biological and computational knowledge stocks has emerged as its own skill set.[14] Thus, a struggle ensues to demarcate the territory of bioinformatics and capture the rights to regulate this knowledge base.[15] Examination of the ways in which new bioinformaticians are trained and certified may help define whether the skill-set is deemed academic, professional, or technical. This social process of occupation creation occurs through the establishment of qualifications, enactment of governing bodies, and a system of qualifying amateurs giving prestige and position to individuals with this knowledge stock[16]

By institutionalizing bioinformatics, organizational structures may constrain the practices and functions of individuals. Power is accumulated in the definition of academic capital. The location of bioinformatics within the scientific landscape impacts the opportunity to cultivate scientific prestige with the support of peers. The potential of bioinformatics to continue as a hybrid skill-set requires the professionalization of this skill set and the recruitment of new students.

Finally, the case of bioinformatics proves interesting as a point from which to view the redrawing of scientific boundaries. It represents a case of technology driving science; the need for scientific understanding to keep pace with the raw ability of technology to produce new data. The choice to move beyond disciplinary boundaries is justified by the need to bridge scientific gaps in order to further collective understanding. Bioinformatics exists in many ways as a technological tool (obligatory passage point[17]) through which to arrive at the opportunity for scientific discovery. Bioinformaticians may serve as the instruments of analysis, taking biological science beyond the processing of raw data and into the realm of rethinking analytical results. This continuing process of computational biological research inspired by technological development may render current disciplinary boundaries obsolete.

As bioinformatics continues to be an essential aspect of biological research, we begin to recognize the unique migration of agents who function at the crossroads of multiple institutional and intellectual intersections. While the majority of career development trajectories propel these trainees toward either an academic or a professional career, bioinformatics produces researchers who move fluidly from one side of the academy-industry boundary to the other.

Conclusion

In reflecting briefly on the development of bioinformatics and the agents who engage in its practice, we recognize the opening of new spaces of intellectual action. Through witnessing the ways in which scientific boundaries are being redrawn, we attempt to further our understanding of the factors that impact scientific development.

Lily Farris is a Research Assistant and Janet Atkinson-Grosjean is a Senior Research Associate e at the W. Maurice Young Centre for Applied Ethics, University of British Columbia, Vancouver, BC. Corresponding author: janetat@interchange.ubc.ca

1. Pierre Bourdieu, *Homo academicus* (Paris: Les Éditions de Minuit, 1984).
2. Information and computer technologies.
3. Michael Huerta *et al.*, *NIH Working definition of bioinformatics and computational biology* (*July 17, 2000)*, online: Biomedical Information Science and Technology Initiative <http://www.bisti.nih.gov/CompuBioDef.pdf>.
The NIH Biomedical Information Science and Technology Initiative Consortium agreed on the following definitions of bioinformatics and computational biology recognizing that no definition could completely eliminate overlap with other activities or preclude variations in interpretation by different individuals and organizations.
Bioinformatics: Research, development, or application of computational tools and approaches for expanding the use of biological, medical, behavioral or health data, including those to acquire, store, organize, archive, analyze, or visualize such data.
Computational Biology: The development and application of data-analytical and theoretical methods, mathematical modeling and computational simulation techniques to the study of biological, behavioral, and social systems
4. Robert Stevens, Carole A. Goble & Sean Bechhofer, "Ontology-based knowledge representation for bioinformatics" (2000) 1:4 Briefings in Bioinformatics 398. N.M. Luscombe, D. Greenbaum & M. Gerstein, "What is bioinformatics? A proposed definition and overview of the field" (2001) 40 Methods of Information in Medicine 346.
5. Susan Leigh Star, & James R. Griesemer, "Institutional ecology, `Translations' and boundary objects: Amateurs and professionals in Berkeley's museum of vertebrate zoology, 1907-39" (1989) 19:3 Social Studies of Science 387.
6. Thomas F. Gieryn, *Cultural boundaries of science: credibility on the line* (Chicago: University of Chicago, 1999).
7. In our larger project, of which this study forms a part, we are examining three translational pathways along which genomics travels between the laboratory and the world: commercial translation, clinical translation, and civic translation. Included in the latter is the creation of open-source/ open-access public databases. See Janet Atkinson-Grosjean & Nichole Dusyk (in review) "Translating translational science: towards a conceptual model," *Science, Technology and Human Values.*
8. *Supra* note 1.
9.For more description, see <http://www.genomebc.ca /research_tech/research_projects/health/pi2.htm> and <http://www.pathogenomics.ca/~dlynn/>.
10. One model will bring in a bioinformatician in a consulting role towards the end of a biological project. Another employs a stand-alone team of bioinformaticians who receive little in the way of direct feedback from the scientists using their research tools (e.g. Wormbase)
11. Thomas F. Gieryn, "Boundary-work and the demarcation of science from non-science: Strains and interests in professional ideologies of scientists" (1983) 48:6 American Sociological Review 781.
12. *Supra* note 1.
13. Andrew D. Abbott, *Chaos of disciplines* (Chicago: University of Chicago Press, 2001).
14. *Supra* note 6.
15. Andrew D. Abbott, *The system of professions: An essay on the division of expert labor* (Chicago: University of Chicago Press,1988).
16. *Supra* note 1.
17. Bruno Latour, *Science in action: how to follow scientists and engineers through society* (Cambridge, Mass.: Harvard University Press, 1987).