Crafting the third century of the National Library of Medicine

RECEIVED 18 July 2016
ACCEPTED 20 July 2016

Patricia Flatley Brennan

AMIA INFORMATICS PROFESSIONALS. LEADING THE WAY.
OXFORD UNIVERSITY PRESS

EDITORIAL

In 20 years the National Library of Medicine (NLM) will enter its third century of service to the nation. From a small bookshelf in the office of an army surgeon to a dynamic suite of resources for discovery and care, the NLM assures that the information needed for discovery and care is available where needed, when needed, and, increasingly, in the format needed. This week I begin my tenure as the 19th appointed director of the NLM. I am using this opportunity to reach out to colleagues in the biomedical and health informatics community to solicit your guidance about how to invest in and grow the NLM.

What kind of library and library services are needed to determine whether a foodborne pathogen is a new species or a known species? How many infrastructure standards, predictive algorithms, natural language processing–guided tumor interpretations, and vocabulary cross-mappings must be in place to support an automated query soliciting the next step in cancer screening for a specific patient in a given community? Is it possible to describe and instantiate the reference genome datasets necessary to characterize zoonotic diseases? Where does a participant in the Precision Medicine Initiative cohort turn to get help with understanding the results of a recently returned genetic assay?

Each of these poses significant and important challenges that draw on the current human and technical resources of the NLM. Complex data structures, analytical tools that allow exploration of huge databases, meaningful information presentation, and the infrastructure to support inquiry comprise but a few of the innovative efforts found among the intramural and extramural programs of the NLM. I would like the NLM to provide the answers to all of these questions in an authoritative, timely, human- and/or machine-interpretable manner. For this to happen, all stakeholders, everywhere, must envision a bold future of data-driven care enabled by a trusted, cost-effective, federated resource.

Come along with me and help craft this third-century vision. What we lay down now will guide investments over the next 20 years, leading up to the launch of the third century. Input is needed now for near-term investments in methodologies and demonstration projects as well as the analytical and retrieval tools necessary to make use of data.

To start your thinking, let me share with you some of the advice I've already received as we envision a third-century NLM.

The push is to "store everything." Not only is this not possible, it has within it its own limitations. What should be stored? How? Where? Is the cloud the solution? In what ways can patient privacy be preserved as we increasingly can know who you are, not simply because of the presence of personal health information, but because of the configurations of personally unique organization of one's genome and epigenome?

Preprints are necessary to accelerate knowledge discovery, even at the expense of preserving intellectual property rights of discovery. Visions of data sharing, stewardship, and reuse of data must align with presently held ideas and values of reproducibility and rigor. The federal government is the trusted party overseeing research dissemination.

How do we conceptualize data as an information resource? It should be possible to apply the same knowledge formalizations used in curating the published literature to effectively curate published datasets. Decisions about data definitions, metadata, and access/authorization rights should remain within the purview of the individual investigator.

Think about these questions. Are they presaging critical directions for health and biomedical informatics? What's missing? What should be emphasized?

Here are some other questions I would like to take on: What is the nature of journals in a digital future? How do we preserve the context of individual study reports in a manner similar to volumes of a journal in a world where direct access to a singular article. Importantly, what kind of training is needed not only to build the data science workforce, but also to prepare data-intensive clinical practitioners and data-informed citizens? In addition to fair use and intellectual property protection, what public policy issues should claim the attention of the NLM?

I bring to the directorship 30 years of experience in a vibrant biomedical informatics community, where my work spanned the application of game theory to the development of health information exchanges, the alignment of clinical information systems with extant and emerging professional practice models, and the design, deployment, and evolution of specialized computer tools to support self-management and self-care in the home.

In my first 90 days, I will work with the NLM Board of Regents to launch a strategic planning process. In addition, I will work with the NIH leadership to foster a sustainable strategy for data science. My other priorities include fostering our education and training initiatives, using the Institute of Medicine report on social and behavioral domains to guide new information models, and meeting key NLM-funded researchers and trainees. In addition, I plan to continue my more than 20-year affiliation with *JAMIA* as an associate editor for special projects.

Crafting a vision for the third-century NLM requires participation from every stakeholder community. The biomedical informatics community brings special knowledge to this discussion, and its members must help guide the evolution of this essential institution. Watch for opportunities to participate, or send me your thoughts directly at patti.brennan@nih.gov.

AUTHOR AFFILIATIONS

Wisconsin Institute for Discovery, Email: pbrennan@discovery.wisc.edu, Effective August 15, 2016: Director, National Library of Medicine, Building 38 NIH, 8600 Rockville Pike, Bethesda, MD 20894 USA

 For numbered affiliations see end of article.