Creating value at the National Institutes of Health

James King [a,*], Christopher W. Belter [b], Bridget Burns [c] and MaShana Davis [c]

[a] *Branch Chief & Information Architect, National Institutes of Health Library, Division of Library Services, Office of Research Services, National Institutes of Health, 10 Center Drive, Bethesda, MD 20892, USA*

[b] *Informationist, National Institutes of Health Library, Division of Library Services, Office of Research Services, National Institutes of Health, 10 Center Drive, Bethesda, MD 20892, USA*

[c] *Information Architect, National Institutes of Health Library, Division of Library Services, Office of Research Services, National Institutes of Health, 10 Center Drive, Bethesda, MD 20892, USA*

Abstract. The National Institutes of Health (NIH) Library meets the needs of the diverse NIH research community through a range of innovative services, resources, and knowledge. Based upon an understanding of the information industry and the mission and goals of NIH, the NIH Library offers a number of services that exploit data of various types to support assessment and create value. Understanding our users' engagement with content (e.g., citations, behavioral data, research funding) allows us to provide personal and customized services including bibliometrics, collection assessment, and custom information solutions.

Keywords: Information services, information solutions, Drupal, portfolio analysis, custom information services, application program interfaces, libraries, virtual research environments, digitization, bibliometrics, collection assessment, innovation

1. Introduction

The National Institutes of Health (NIH), a part of the Department of Health and Human Services (HHS), is the world's largest biomedical research agency. Comprised of twenty-seven institutes and centers, the NIH is a principal funder of biomedical research, with 80% of the NIH's $32B annual budget devoted to extramural research across the country and 10% devoted to intramural research performed by NIH staff. Nearly 6,000 researchers work in NIH's own research laboratories primarily on the main campus of NIH in Bethesda, MD. These researchers work to prevent disease and improve health for millions of people.[1]

The NIH Library's vision is "to be the premier provider of information solutions to enable discovery and improve health," supporting both the intramural research community and the NIH scientific administrators who manage the research grants awarded each year. Distinct from the National Library of Medicine (NLM) on the same campus which serves the entire nation's health needs and delivers vital services, including PubMed and PubMed Central, the NIH Library focuses on the information needs of the NIH community and also provides primary library services for most components of HHS. The NIH

---

*Corresponding author. E-mail: james.king@nih.gov.

[1] What is NIH? [Internet], National Institutes of Health (NIH), 2015 [updated 2015 September 14; cited 2 May 2016]. Available from: http://www.nih.gov/about-nih/explore-nih/what-nih.

Library actively collaborates with the libraries at our two major sister agencies in HHS, the Center for Disease Control and Prevention (CDC) and the Food and Drug Administration (FDA). A key component to the success of the NIH Library is the personalized service delivered through the Informationist program. A foundation of the Informationist program is evidence-based decision making, the process of systematically finding, appraising, and using the best-available research for decisions related to clinical care, biomedical research, public health, and health policy.

The NIH Library's Informationists program started in 2001 as one of the first in the country to merge scientific excellence with professional librarianship in order to more precisely anticipate and respond to customers' needs. Based on the changing needs and opportunities at NIH, the NIH Informationist program has expanded from its base of subject matter-based informationists to also include a growing number of skills-focused professionals in the areas of bibliometrics, bioinformatics, data services, editing, and information architecture.[2]

2. Assessment at the NIH Library

Evidence-based decision making is a core aspect of the NIH and the NIH Library and has driven change and innovation at NIH for many years. For this paper we focus on three areas: bibliometric analysis, collection analysis, and custom information solutions.

2.1. Bibliometric analysis

Bibliometrics is a scientific discipline focused on the quantitative analysis of academic publications.[3] Using these publications as a proxy for research, bibliometrics attempts to understand the production, producers, and evolution of academic research. One of the key aspects of bibliometrics is assessing the impact, or usefulness, of certain publications – or publications by certain authors, institutions, or countries – on subsequent research. Many academic institutions are increasingly using bibliometrics to evaluate the scientific research of authors, laboratories, institutions, and grant programs to make decisions about future funding.

The NIH, as both a funder and a producer of scientific research, is one of these institutions. Over the past few years, the NIH Library saw an increased demand for bibliometrics information from NIH staff, so in June 2014 it launched the Bibliometric Services program to provide expert advice, training, and custom services to its customers. In the same way that Library customers could request literature searches or data management services from the Library, Library customers can now request bibliometric services.

One of our major initiatives over the first year of the program was to create bibliometric profiles of the intramural research at each of the NIH's twenty seven institutes and centers (ICs). These profiles not only served as an example of the kinds of analyses we could perform, but also as a marketing tool that we sent to the scientific directors of each institute and center to advertise the new service. In addition to these profiles, we have received a growing number of requests for custom analyses. Many of these custom analyses attempt to evaluate the research produced by specific laboratories or grant funding

---

[2]S.F. Grefsheim, S.C. Whitmore, B.A. Rapp, J.A. Rankin, R.R. Robison and C. Canto, The informationist: Building evidence for an emerging health profession, *Journal of the Medical Library Association* **98**(2) (2010), 147–156.

[3]C.W. Belter, Bibliometric indicators: Opportunities and limits, *Journal of the Medical Library Association* **103**(4) (2015), 219–221.

programs at NIH. Others attempt to trace the evolution of research on a specific topic that was funded by a particular institute, such as the evolution of obesity research funded by the National Heart, Lung, and Blood Institute.[4] Other analyses attempt to study the intellectual structure and development of entire disciplines, like drug discovery for Alzheimer's disease or cancer immunotherapy, to determine what the major research directions in these disciplines are and where they might be headed.

In the nearly two years since the program launched, we have provided, or are currently working on, over two hundred consultations, training sessions, and bibliometric analyses. As word of our services has spread through NIH, we have shifted our focus from the self-initiated profile reports of our first year to demand-driven analyses in our second year. Both the number and complexity of these demand-driven analyses continue to increase, suggesting that not only was there a need for these services at NIH, but also that NIH staff feel comfortable asking the Library for them. The success of our bibliometrics program and the formation of similar programs by librarians at Washington University in St. Louis, Northwestern University, Cornell University, the University of Massachusetts, the University of Waterloo, the University of Alberta, the Johns Hopkins Applied Physics Laboratory, and others also suggests that bibliometrics represent a significant opportunity for academic and special libraries in the U.S. and Canada.

2.2. Collection assessment

The NIH Library provides desktop or tablet access to over ten thousand individually-subscribed scholarly journals (no aggregator databases included), most with archives all the way back to the first issue. Our online journal collection is also popular (with our usage often compared against an entire country's usage – we use an average of twenty-three thousand articles per day).

Given the importance and cost of our annual subscription budget, we have continued to refine our collection management and analysis strategy. Over the past couple years, we have shifted into an active collection management approach by evaluating our collection regularly to identify low use, low impact titles and removing them rather than waiting for a budget exercise to force us to find potential cancellations.

Like most other libraries, we start our collection evaluation with price, usage, and impact factor. We wanted additional and more accurate ways of measuring the value of our collection so we also analyze our document delivery requests and turn-away reports from publishers, and have switched our book collection from subject-matter selection to Demand-Driven Acquisitions (DDA). Starting in 2016, we began applying bibliometric analysis to our collection.[5] Just looking at 2014 data, we now have a dataset of nearly a half-million citations from nearly 10,000 NIH publications to work through. We are looking to see what journals our NIH authors are most often publishing in and which ones they are most often citing – light years ahead of keeping track of which print volumes were being re-shelved in the old days.

Another important aspect of our collection management strategy is maintaining a database to manage all of our purchase and usage information. We are moving from a home-grown database and spreadsheets to a commercial Electronic Resources Management (ERM) system from ProQuest/Ex Libris by the end of 2016.

---

[4]H.L. Nicastro, C.W. Belter, M.S. Lauer, S.A. Coady, L.J. Fine and C.M. Loria, The productivity of NHLBI-funded obesity research, 1983–2013, *Obesity* **24**(6) (2016), 1356–1365.

[5]C.W. Belter and N.K. Kaske, Using bibliometrics to demonstrate the value of library journal collections, *College & Research Libraries* **77**(4) (2016), 410–422.

2.3. Custom Information Solutions

The Custom Information Solutions service (http://nihlibrary.nih.gov/Services/Pages/Custom InformationSolutions.aspx) functions as the "geek squad" or the technical arm for the Library's Informationist program. By leveraging information architecture, which blends technology expertise with librarianship and information science, this group offers services to determine the best approach for solving the information and technology problems currently facing research groups at the NIH. The Custom Information Solutions service falls into four primary areas: consulting and project management of data-intensive websites, API and commercial/government dataset access, digitization of government documents, and technology community building. Since launching in 2009, several projects have been produced in partnership with the various NIH institutes and centers. Outlined here are some specific examples.[6]

2.3.1. Consulting and project management of data-intensive websites

Supporting scientific collaboration through virtual research environments. The Custom Information Solutions service was first launched through a project to digitize and index a collection of over five thousand historical publications covering the influenza pandemic of 1918. The collection, consisting primarily of books and journals, spanned back over five hundred years and contained seventeen different languages. The project, later named the Pandemic Influenza Digital Archive or PIDA (http://pida. nihlibrary.com), sparked a partnership between the Library and the National Institute of Allergy and Infectious Diseases (NIAID). NIAID approached the Library to turn this paper-based collection into a virtual collaboration space that would support researchers around the globe studying pandemics. Interested researchers would be able to browse and search the collection and visualize the publications on maps and timelines. Authenticated users would be able to save and share custom search results and update records. In 2014, the Library hired a subject matter expert to index the online collection to include information about associated disease outbreaks, complications and coexisting diseases, historical organisms and diseases, age groups, surrounding events and locations, weather conditions and seasons, epidemiology, and etiology of the influenza pandemic of 1918. The end result is a rich dataset of the influenza pandemic of 1918 that can potentially inform discovery, prevention, and treatment of future pandemics. Future development will unleash the collaborative components of PIDA, enabling users to comment on, rate, and update records. The Library is now the primary manager of this digital archive and will continue to fine tune the infrastructure and dataset in preparation for the centennial of the 1918 pandemic.[7]

Analyzing Alzheimer's disease and pain research grant funding. To ensure the most effective use of government funding, research groups at NIH are taking a deeper dive into analyzing grant funding data. Two examples of this include the International Alzheimer's Disease Research Portfolio or IADRP (http://iadrp.nia.nih.gov) and the Interagency Pain Research Portfolio or IPRP (http://paindatabase.nih. gov). Faced with a congressional mandate to make Alzheimer's disease research funding more visible and global in perspective,[8] the National Institute on Aging (NIA) Office of Planning, Analysis, and

---

[6]J. King, Geek squad for libraries: Custom information solutions through the NIH Library informationist program, *Public Services Quarterly* **10**(4) (2014), 350–354.

[7]J. King, Building the Pandemic Influenza Digital Archive (PIDA) at the National Institutes of Health Library, *Sci-Tech News* [Internet] **64**(3) (2010). Available from: http://jdc.jefferson.edu/scitechnews/vol64/iss3/6.

[8]National plan to address Alzheimer's disease [Internet], Office of the Assistant Secretary for Planning and Evaluation (ASPE) at the U.S. Dept. Of Health and Human Services, 2015 [updated 2012 May 14; cited 2 May 2016]. Available from: http://aspe.hhs.gov/daltcp/napa/natlplan.shtml.

Evaluation (OPAE) approached the Library to convert six thousand rows of printed Alzheimer's disease grant funding data into a searchable website. The Library took this opportunity to expand the concept into the area of portfolio analysis. Facilitating a relationship with a technology partner, we were able to craft a proof-of-concept with searching, exporting, and data analysis capabilities. Today, IADRP provides access to grant funding data from over thirty public and private organizations in the U.S. and abroad, allowing users to search and export data across sixteen different fields including by taxonomy term, funding organization or year, and principal investigator. Through IADRP, NIA now has a powerful tool to inform the public about Alzheimer's disease research funding. In 2016 the online portfolio will be expanded to include flexible charting and analysis tools to support review of gap analysis and funding overlaps, for both NIH and its global partners.

Based on the success of the IADRP, the Library also formed a partnership with the National Institute of Neurological Disorders and Stroke (NINDS) to develop IPRP, an online portfolio of pain research grant funding and training activities supported by six government agencies, including NIH. IPRP proved its value when NINDS was called on to respond to a multi-agency data request with an urgent congressional deadline. Recently, the website was enhanced to include a distributed coding workflow to support contributing agencies' coding efforts. Funding records can now be modified or coded with the correct pain conditions and themes through the website, ultimately speeding up the coding process that was previously done by exchanging multiple spreadsheets via email. Given the success of these projects, the Library is embarking on a third online portfolio related to epilepsy research. The examples highlighted here have allowed NIH Institutes and Centers (ICs) to discern areas of overlap, gaps in research, and collaborative opportunities. These resources that the NIH Library has developed are now enabling NIH ICs to respond to congressional inquiries and provide the public with insight into how federal research funding is being spent.

A more recent project in the Custom Information Solutions service portfolio is the Alzheimer's Disease Preclinical Efficacy Database or AlzPED (http://alzped.nia.nih.gov). This project was done in partnership with NIA's Drug Discovery and Development Program to address the challenge of successful development of drugs and treatment and/or prevention of Alzheimer's disease related to poor translation of preclinical efficacy from Alzheimer's disease models to the clinic. At the outset, the Library hired a subject matter expert to help shape project requirements, identify appropriate fields, promote the end product, and curate submitted data, and they also partnered with a technology firm to implement a proof-of-concept. AlzPED is a publically-available web-based portal that aims to promote efficiency, transparency, reproducibility, and accuracy of research targeted at preclinical therapy development for Alzheimer's disease. It offers centralized storage of, and access and visibility to, integrated preclinical efficacy data from published and unpublished preclinical studies in a variety of fields including study design, therapeutics targets and agents, outcome measures, and conclusions. Users can search for studies, browse by topic, and view information in detailed records. AlzPED also integrates with PubMed, PubChem, Clinical Trials, and patent databases to connect users to related and relevant data.[9]

2.3.2. API and commercial/government dataset access

Given the evolving information needs and use cases of our customers, the Library has gathered information and expertise on application programming interfaces (APIs) and procedures for bulk data purchases. Building on contractual relationships that the Library has with major bibliographic data providers such as Elsevier and Thomson Reuters, we are able to provide access to a variety of datasets through

---

[9]About AlzPED/AlzPED Alzheimer's Preclinical Efficacy Database [Internet], National Institutes of Health, 2016 [cited 2 May 2016]. Available from: http://alzped.nia.nih.gov.

these "side-doors," offering NIH staff programmatic access to journals, books, citation data, and abstracts. The Custom Information Solutions team has grown its consulting efforts with developers and Information Technology staff across NIH to help them take advantage of API information and other specialized tools, in addition to helping them purchase data dumps or evaluate analysis tools available from specialty firms. Going beyond major vendors, the Library plans to expand to cover other internal resources available at NIH including NIH Reporter, a research portfolio online reporting tool website that provides a central point of access to reports, data, and analyses of NIH research.[10]

2.3.3. Digitization of government publications

Digitization, although not new to libraries, is becoming a popular service to migrate government-created publications onto the open web. The NIH Library has consulted with different groups across NIH to identify their digitization requirements and make recommendations on the best path forward, assisted with the contracting and evaluation process, and managed the digitization workflow. Most of our digitization efforts have used the joint partnership between the Library of Congress and Internet Archive, called FedScan. FedScan provides non-destructive digitization services to Federal agencies at ten cents per page image, with a multitude of output formats.[11] All scans are publicly-available on the Internet Archive site (archive.org). For organizations looking to digitize public information or other content not under copyright, FedScan is an affordable option.

Over the past several years, the NIH Library has been able to digitize over two thousand publications (https://archive.org/details/nihlibrary) from its print collection, primarily NIH and institute annual reports and out-of-copyright medical texts. We have also been able to partner with specific institutes and the NIH History Office to digitize various newsletters and other NIH-produced publications. In addition, we had the opportunity to partner with the Centers for Medicare & Medicaid Services (CMS) Library, a part of HHS, to digitize over six thousand government publications from its print collection (https://archive.org/details/cmslibrary).

2.3.4. Technology community building

The NIH Library strongly believes that libraries and librarians should serve as champions for technologies that will make a marked improvement in how government agencies operate and provide services internally and to the public. Technologies that we are currently focused on or are exploring include 3D printing/scanning/modeling, bioinformatics, data visualization, electronic lab notebooks and pens, gamification, mobile devices and applications, online mapping and spatial data analysis, web development and graphic design, and audio/video production. To highlight these technologies, the Library has held several conferences and training classes, and has provided access to equipment to facilitate and encourage the adoption of these technologies.

One specific example of how we are building community is the Library's involvement with Drupal. Given the dramatic growth and adoption of Drupal (http://www.drupal.org) at NIH and across federal government, the Library has hosted user group meetings and three national conferences, on the NIH campus, that focused on Drupal. The quarterly user group meetings, held in the library and online, usually are themed and have guest speakers from the NIH community, other government agencies, or vendors that provide information on their experience with Drupal. The Library's team has also offered

---

[10]Research Portfolio Online Reporting Tools (RePORT) [Internet], National Institutes of Health, 2016 [cited 2 May 2016]. Available from: http://alzped.nia.nih.gov.

[11]Federal Scanning Center at the Library of Congress (FedScan) frequently asked questions [Internet], Library of Congress, 2008. Available from: http://www.loc.gov/flicc/contracts/pdfs/FEDSCAN_IAFAQs.pdf.

Drupal training in the form of "Ask Me Anything" sessions, where a Drupal expert provides recommendations on a range of Drupal topics, from how to implement Drupal to migrating from another content management system, as well as more advanced topics like security and enterprise user authentication. In addition, the Custom Information Solutions team implemented a listserv and website to facilitate idea exchange and collaboration and to educate potential and existing adopters. Since 2013, the Library has served as host and a co-planner for the annual Drupal GovCon (http://drupalgovcon.org/), a free event that brings together individuals from the public and private sectors that use, develop, design, and support Drupal. In 2015, over eight hundred attendees enjoyed three dynamic keynote talks, over seventy breakout sessions, all day training classes, and coding sprints.

3. Conclusion

Assessment and evidence-based decision making are critically important to the effective management of the NIH Library. However, rather than primarily using those approaches to show the value of the library, we are focused on using these tools and approaches to show the value of NIH's scientific staff and to help them improve their research. By helping our customers to show their value to their stakeholders, that in turn shows our value.

We firmly believe that our future is wrapped-up in how well we can collaborate with our customers. The technology trends of the past twenty years can be summed up as a shift away from collections and towards innovative, collaborative services. Great service is built on a clear understanding of our customers' needs and delivering solutions, in some cases in advance of their asking for them. By building on our embedded relationships with our customers, the NIH Library has been able to "enable discovery and improve health" by adding value to the research and grant management processes at NIH.

Disclaimer

The information in this article does not necessarily reflect the opinions of the National Institutes of Health. Any mention of a product or company name is for clarification and does not constitute an endorsement by NIH or the NIH Library.

Acknowledgements

About the corresponding author

James King currently leads the Custom Information Services effort at the National Institutes of Health Library, NIH. Recent efforts have included digitization of a thousand volumes for public access, development of Drupal-based project web sites including one focused on historic pandemic influenza

research (http://bit.ly/12c8qVx), and consulting/development of disease space portfolio analyses. Previously, James spent eighteen years at the Naval Research Laboratory's Research Library and was fortunate to be a part of a forward-thinking digital library team during the birth of the Internet and web era. James assisted with and later managed one of the first and largest government digitization and Web-based efforts in federal libraries. James is very active in the Special Libraries Association (SLA), currently serving on the SLA Board of Directors. He has a Masters in Library & Information Science from Catholic University (2002) and a Bachelors of Computer Graphics from Salem International University. James has served on several industry library advisory boards and has received several recognitions including the CUA Library School Alumni Award (2015), the FEDLINK Federal Librarian of the Year (2015), SLA Fellow (2011), SLA Innovations in Technology Award (2006), and ALA's Federal and Armed Forces Libraries Round Table Achievement (2006) and Distinguished Service (2003) Awards.