

Against the Grain

Volume 20 | Issue 5

Article 8

November 2008

VIVO: Simplifying Research Discovery in the Life Sciences

Jaron Porciello

Cornell University, jat264@cornell.edu

Medha Devare

Cornell University, mhd6@cornell.edu

Jon Corson-Rikert

Cornell University, jc55@cornell.edu

Follow this and additional works at: <https://docs.lib.purdue.edu/atg>

 Part of the [Library and Information Science Commons](#)

Recommended Citation

Porciello, Jaron; Devare, Medha; and Corson-Rikert, Jon (2008) "VIVO: Simplifying Research Discovery in the Life Sciences," *Against the Grain*: Vol. 20: Iss. 5, Article 8.

DOI: <https://doi.org/10.7771/2380-176X.5186>

This document has been made available through Purdue e-Pubs, a service of the Purdue University Libraries. Please contact epubs@purdue.edu for additional information.

students why publishers offer unlimited access to journal articles but only limited access to books and book chapters is futile. Though some vendors are now offering site licenses to eBooks, they come at a hefty price considering that the site license will only eliminate a few hours of heavy, peak traffic each day.

The opportunity for the library to partner with the UASOM has been beneficial but several inconsistencies with the way publishers digitize and sell eBook content to institutional customers remain in the way. For instance, **Digital Rights Management (DRM)** technologies prevent students from downloading and printing an entire chapter of an eBook. In their 2007 *Global Faculty E-book Survey*, **Ebrary** found a deterrent to using eBooks was the inability to download or print. Many prefer to print rather than read the eBook online. Studies have shown that resistance to reading books online is not generational but is in fact due to the physical limitations of the eye.⁴ So, an impasse remains. Until publishers allow

users greater ability to print and manipulate content, eBooks will only be useful for gathering quick facts on a specific topic. This is not entirely different from what free Web-based encyclopedia projects like **Wikipedia** are currently providing. Maybe what is needed is the condensed version of electronic content on a portable device.

Conclusion

While eBooks are certainly a good investment, they definitely cannot yet be seen as equivalent alternatives to print books. As one faculty member phrased it, "physical books help with [the students'] visual memories as well, it's easier to remember where, just visually how you think, of where I read that piece of information; it was near the front, so it was an introductory idea."⁵ Many users want to be able to highlight passages, write notes in the margin, and loan the book to a friend. Print books also open up the possibility of serendipitous discovery when browsing through the pages of the book. eBooks simply do not allow for such engagements. However, the eBook landscape is changing rapidly every day, and it will be interesting to watch what develops

next. As more and more schools offer distance education courses and professors want the ability to integrate textbook content into their online course management systems, eBooks will undoubtedly play a part. 🌸

Endnotes

1. **Levine-Clark, Michael**. "Electronic Book Usage: A Survey at the University of Denver," *Portal* vol.6, no.3 (July 2006): 285-299.
2. **Naomi Dillon**. "The e-Volving Textbook," *American School Board Journal* 196, no.7 (July 2008): 23.
3. **Cleto, Cynthia**. "ATG Special Report — 10 Steps to Implementing an eBook Collection: A Guide for Librarians," *Against the Grain* v.20, no.1 (2008): 47.
4. **Sottong, Stephen**. "The Elusive E-book," *American Libraries* 39, no.5 (2008): 47.
5. **Carlock, Danielle M. and Anali Maughan Perry**. "Exploring Faculty Experiences with E-books: a Focus Group," *Library Hi Tech* 26, no.2 (2008): 251.

VIVO: Simplifying Research Discovery in the Life Sciences

by **Jaron Porciello** (Assistant Librarian, Cornell University, Albert R. Mann Library)
<jat264@cornell.edu>

and **Medha Devare** (Senior Assistant Librarian, Cornell University, Albert R. Mann Library)
<mhd6@cornell.edu>

and **Jon Corson-Rikert** (Head, Information Technology Services, Cornell University, Albert R. Mann Library) <jc55@cornell.edu>

Introduction

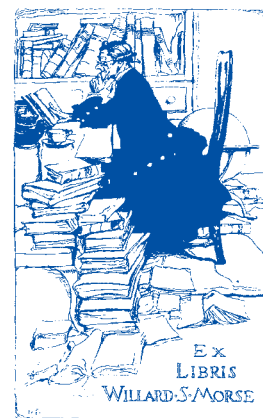
In a 21st century environment where users expect up-to-date information delivered by means of a single point of access, academic institutions, particularly those involved with life sciences, medical, or health science research, are increasingly engaged with interdisciplinary research communication technology and dissemination in an effort to remain competitive. As a 2007 **National Science Foundation (NSF) Impact of Proposal and Award Management Mechanisms (IPAMM)** report states, the **Foundation** has "actively fostered a shift [towards an interdisciplinary culture] both through the use of solicitations requiring interdisciplinary teams of researchers and by simplifying the mechanism for submitting collaborative projects to any NSF program through **FastLane**."¹ Certainly, NSF is not alone in expressing preference for interdisciplinary collaboration; other funding agencies, as well as **University** reports and task forces, have similarly articulated commitment to strengthen collaborative work.² At **Cornell University**, the creation of the 2001 **New Life Sciences Initiative (NSLI)**, the most far-reaching initiative in **Cornell's** history, is a major

effort to boost multi-disciplinary collaboration. One system initially designed to respond to **NSLI's** collaborative mission is the **Virtual Life Science Library**, or **VIVO**. **VIVO** is a library-developed system dedicated to providing on-line users, both internal and external to the university, with an overview of life science research and scholarship at **Cornell**. An ontology and semantic Web application, **VIVO** has introduced a new approach for individuals interested in medical and life sciences research to discover grant information, facilities and services, publications, research areas, and more, in one virtual space. This article will provide an overview of **VIVO's** development as a library initiative and discuss its structure, sources of content and next steps that ensure the information in the system remains accurate, current, and accessible.

Background

VIVO's impetus stemmed from **NSLI's** mission "to strengthen existing premier departments and programs as well as promote new and collaborative initiatives."³ Specific **NSLI** disciplines include a range of fields such as, but not limited to, ecology and evolutionary biol-

ogy, nutritional sciences, plant sciences, biomedical engineering and computational biology. During preliminary **NLSI** meetings, university faculty and academic staff articulated that they had difficulty identifying intra-institutional research collaborators, facilities, and resources. Since faculty engaged in multi-disciplinary collaboration in life sciences and biomedicine receive major sources of funding from foundations and agencies keen to award research endeavors spanning multiple disciplines, this lack of information was a major impediment.⁴ **Mann Library** representatives who attended these meetings realized such concerns constituted a void in access to information, and that the library, as a bastion of information stewardship and dissemination, was perhaps the best candidate to examine this need in greater detail. Librarians, as well as many developers in library information technology departments, are trained to evaluate both resources and information management practices by considering vast informational landscapes, including how to best classify, represent, and deliver the myriad



of outputs such as books, articles, Websites, and audio-visual material, that comprise a field of study. In addition, many academic libraries have embraced technology that can expedite discovery of substantial quantities of information. To further explore the institutional complexities as expressed by the faculty, the **Cornell University Library** formed the **Life Sciences Working Group**. This group, whose expertise range from agriculture to medical and veterinary science to chemistry, identified a number of key areas for impact and examination, including the need to present a unified view of the **Life Sciences at Cornell**, via a single point of entry, where research and scholarship are emphasized.⁵ As the group considered approaches to connect life sciences research activities, events, and scholarly output from a diverse academic community, it became clear that an index providing links to Websites and resources would not reveal the kind of associations faculty sought. Thus, instead of preserving a linear hierarchy where research is affixed to person via their college, department or campus, members of the **Life Sciences Working Group** were determined to find a new way to cross-reference research and activities spanning multiple departments, centers, institutes and campuses to connect scholarship to its most basic element: the researcher.

Other academic institutions also active in scientific research have similarly created new research-focused discovery models. Similar Web-based efforts from institutions committed to fostering inter-disciplinary collaboration are taking place across the United States and abroad. **HealthLinks**⁶ at the **University of Washington**, and **Bio-X** at **Stanford University**, are two such systems that highlight research and communication in an effort to link individuals to their areas of research in an online environment that encourages scholastic connection. In addition, **University of Florida's Marston Library** has implemented the **VIVO** technology to develop their own comprehensive research tool, **Gator Scholar**. One non-U.S. based system, **Find an Expert**, from Australia's **University of Melbourne**, is a University-wide system that "turn[s] administrative data inside out" to provide the public with a user-friendly interface to discover experts on a particular topic.⁷ Creation of these, as well as many other research discovery models not mentioned here, seems to clearly demonstrate the need for inter-disciplinary scientific research retrieval.

System

VIVO was created by **Mann Library** developers using an ontology blueprint from the **Advanced Knowledge Technologies (AKT)** project, an early Semantic Web project designed to discover and promote connections among computer science research activities across multiple universities in the United Kingdom. Ontology is, historically-speaking, a philosophical concept, but has more recently been adopted by information science as a process to express relationships within a domain; in this instance, the domain is **Cornell**, and the relationships

against the grain people profile

Assistant Librarian, Special Projects
Albert R. Mann Library, Cornell University
206 Mann Library

Phone: (607) 255-9958 • <jat264@cornell.edu>

Jaron Porciello

PROFESSIONAL CAREER AND ACTIVITIES: Jaron focuses on institutional and international collaborative projects, and is also active on **Cornell University Library's Instruction Committee**.

FAVORITE BOOKS: *Brideshead Revisited* by Evelyn Waugh; *Palace of Dreams* by Ismail Kadare.

HOW/WHERE DO I SEE THE INDUSTRY IN FIVE YEARS: I think librarians will increasingly seek collaborative tools and platforms to deal with content management challenges. 🐼



are people and their research activities. The ontology reflects people in academia, and the affiliations they are likely to have, such as his or her relationship to a department, a graduate field, a research grant, a publication, or an event. As relationships overlap and intersect through common associations, a network of connections builds that can be entered at any point (typically by discovery through a search engine) and navigated to provide users a much greater sense of context than typical top-down, administratively-organized Websites. An ontology structure is particularly useful not only because of the collaborative emphasis within life sciences and medical research, but also because it can accurately represent **Cornell's** field-oriented structure for graduate studies, which cuts across traditional department and college boundaries. **VIVO's** ontology is "home-grown" in part because of the specificity of information desired by faculty. In addition, very few commercial systems could, at the time, effectively showcase an array of material that was constantly changing, highly inter-linked, and that permitted direct entry of new content by individuals (faculty) with little incentive to adopt anything beyond a simple interface.

Content is initially entered into **VIVO** both through manual curation and automated or semi-automated processes, but developers are promoting direct self-update as an important additional means of updating narrative information in the terminology of the moment. Input, as well as discovery, is bi-directional; thus, the link between faculty and field of study could be made by either declaring a field of study relationship to the person, or by declaring a field of study and then adding the person. Conducting a **VIVO** search will yield a display that clusters entities into broad categories, such as people, events, and organizations.

Sources of Content and Library Outreach

VIVO's content is culled by two primary means: automated data ingests from units that

can provide standardized, machine readable data in a consistent format, and non-automated processes such as manual entry. **VIVO's** mission is to represent the breadth and depth of scholarship at **Cornell** and not to replace or replicate department or other university Websites which are designed for more thorough narratives, branding, multimedia, and document hosting. To this end, librarians identified and prioritized specific classes to populate first, such as people, research facilities, academic units, and graduate fields because cross-referenced content in these areas could quickly bolster the impact **VIVO** would have on the life sciences community. Populating each parent class with individual entities was a combination of manual and automated effort. Student editors, as well as librarians, surveyed publicly available Web and print resources from departments, laboratories, centers, and administrative units for information about institutional contributions to scholarship. This was complemented by certain content currently provided to **VIVO** though automated means by cooperation with several on-campus units including the **Office of Human Resources (OHR)**; the **Office of Sponsored Projects (OSP)**; the **System for Tracking Administrative Records for Students (STARS)**; and internal college reporting systems. Some journal citations have also been harvested from commercial and non-commercial vendors. Data from these sources offer integral updates to the faculty and academic staff's overall summary in the **VIVO** system, providing such information as faculty and academic staff members' official department appointment and title (**OHR**), and grants reported and administered (**OSP**). In addition, colleges that have employed their own annual reporting mechanisms can elect to provide non-sensitive faculty information, such as research areas, professional activities, and publications. Publications have required a three-part effort to gather and display in **VIVO**; journal citation information has been brought

continued on page 38

in from external commercial and open-source database vendors such as **Biosis** and **PubMed**, through manual curation of information found on faculty member's department Websites and online C.V.'s, as well as through internal reporting systems.

The initial content input for the life sciences and subsequent disciplines required a great deal of human entry. The efforts of both manual curator and automated processes have resulted in a system that reflects the electronic information currently available for faculty and staff, but manual entry alone is not a viable option for moving forward. **VIVO's** goal is to transition to a platform where nearly all of its content is harvested and delivered through automated or semi-automated processes. The sustainability of **VIVO's** content has remained at the forefront of the technological and administrative development of the system, particularly with regards to the profiles of individual researchers.

Librarians and developers on the **VIVO** project have focused much of their effort on outreach with faculty, staff, and administrative and academic units to unearth standardized information from any available sources, as well as to engage the university community in content development and management. The library recognizes its strong role as an impartial nucleus between multiple campus interests; it does not, however, have the expertise and resources to continually develop accurate class associations for the university's diverse community. Nor does it have the type of consistent foundation that a system like **Find an Expert (University of Melbourne)** has; Australia maintains several government-defined classification schemes, including **Research Fields, Courses and Disciplines (RFCD)** that offers a strong semantic framework on which to build. Engaging the university community, particularly administrative units and research centers and institutes, in content management has been met with some success; for example, a high-tech facility recently requested and received editor training in an effort to manage their own content. The **VIVO** system, in as much as it seeks to enrich the collaborative culture on campus, must also foster partnerships for itself to ensure that the system remains a powerful, accurate and highly utilized component of **Cornell's** overall online presence.

Conclusion and Next Steps

As interdisciplinary collaboration is imbedded into more university cultures, both in the life and medical sciences and beyond, adopting online systems that can provide new discovery models will likely increase. A case in point is **VIVO**; when **VIVO** was released to the **Virtual Life Sciences Library** in 2005, other **Cornell** colleges, departments, and administrative units realized the potential offered by **VIVO**. These entities, too, wanted a system that could provide a bird's eye view of their resources, research, events, and facilities. In 2006, with cooperation and funding from

against the grain people profile

Bioinformatics and Life Sciences Librarian
Albert R. Mann Library, Cornell University
160 Mann Library, Cornell University, Ithaca, New York 14850
Phone: 607 255 2199 • Fax: 607 255 0318
<mhd6@cornell.edu> <http://vivo.cornell.edu>

Medha Devare, PhD

BORN & LIVED: Born in Bombay, India.

EARLY LIFE: Bombay, India.

FAMILY: I have a husband, 7-year old son, one 8-year, and one 6-month old dog.

PROFESSIONAL CAREER AND ACTIVITIES: I am part scientist, part librarian and part mom.

IN MY SPARE TIME I LIKE: I read.

FAVORITE BOOKS: Too many to list.

PET PEEVES/WHAT MAKES ME MAD: People who're peevish with pets.

PHILOSOPHY: Aim high...so you hit middle ground at the very least.

MOST MEANINGFUL CAREER ACHIEVEMENT: Surviving PhD research in the paddy rice fields of Nepal without encountering interesting internal parasites.

GOAL I HOPE TO ACHIEVE FIVE YEARS FROM NOW: To have a well-adjusted and happy pre-teenager.

HOW/WHERE DO I SEE THE INDUSTRY IN FIVE YEARS: The successful academic library will be more intimately involved in facilitating the business of academia and research, with librarians closely involved in providing teaching support and in-depth subject expertise, while needs for reference services diminish due to improved access to collections and knowledge. The knowledge, skills, and creations of library information technology departments will be important resources not just for the library, but for the university as a whole. 🐾



Cornell's administration, the library agreed to expand what was initially a discovery tool to emphasize life sciences to a system that could showcase the research of over 1,200 faculty and academic staff who span **Cornell's** geographically distant campuses in New York State and abroad. In the latest phase of development, self-editing capabilities for the system have been developed and are currently under testing. This functionality will allow faculty and academic staff at **Cornell** to securely login using the university authorization system, and modify their current **VIVO** profile. This latest iteration, **VIVO Research and Scholarship**, will be released campus-wide in fall 2008. The library will continue to play its role as the technical arbitrator and outreach liaison between **VIVO** and the university community, as well as seeking new ways to strengthen access to content across not only disciplinary boundaries within **Cornell**, but across institutions as well.

Acknowledgments

The authors wish to acknowledge the dedication and innovation of the **VIVO** development team; **Brian Caruso** and **Brian Lowe**, Programming Development, and **Nick Cappadona** and **Miles Worthington**, Interface Development. 🐾

Endnotes

1. **National Science Foundation.** "Impact of Proposal and Award Management Mechanisms." (2007). <http://www.nsf.gov/pubs/2007/nsf0745/nsf0745.pdf>.
2. **Association of American Universities.** *Report of the Interdisciplinary Task Force.* (2005). http://www.aau.edu/reports/InterdisciplinarityReport_Final102705.pdf.
3. **Cornell University.** **New Life Sciences Initiative.** http://lifesciences.cornell.edu/about/initiative_article02.php.
4. **Sonnenwald, Diane.** 2007. "Scientific Collaboration." *Annual Review of Information Science and Technology* 41; 643-681.
5. **Devare, Medha, Jon Corson-Rikert, Brian Caruso, Brian Lowe, Kathy Chiang, and Janet McCue.** 2007. "VIVO: Connecting People, Creating a Virtual Life Sciences Community." *D-Lib Magazine* 13 no. 7-8. <http://www.dlib.org/dlib/july07/devare/07devare.html>.
6. **VIVO** development was inspired and adapted from many of the features found on **HealthLinks**.
7. **Porter, Simon.** 2006. "Turning Administrative Information Inside Out: A report on creating a University wide Public Research Portal based on Research, Finance, HR, and Library Systems." The Council of Australian University Directors of Information Technology (CAUDIT). http://www.caudit.edu.au/educauseaustralasia07/authors_papers/Porter-243.pdf.