



2014 VIVO Conference

August 6-8, 2014 Hyatt Regency Austin, Austin, TX



VIVO Conference Program

Panels/Long Papers

Collaboration institutionally and beyond: status, partners, and next steps

William Barnett, David Eichmann, Griffin Weber, Eric Meeks, Amy Brand and Holly Falk-Krzesinski

Abstract: A key component of success to VIVO and other research networking platforms is their successful use by and utility for science communities. Such success necessarily involves stakeholders representing investigators, institutions, service and software providers, and industry partners. This panel will explore progress and plans to advance the use of research networking by researchers, explore project dynamics around a public private partnership between academic institutions and industry, and discuss how these activities can advance the mission of translational science (or science in general) to the benefit of all. This panel will focus on the following:

- The adoption of research networking by translational sciences communities to date,
- The intellectual merits of a FutureResearchNet project in terms of improving our understanding of how researchers would discover collaborators
- How this effort will help better understand and serve the successful formation of science teams, and
- How it represents a collaboration between academia and industry.

Panelists represent national efforts to advance collaboration, institutions that have assessed the success of these efforts to date, and plans for the future.

Promoting ORCID Adoption in the Academic Institutional Ecosystem: Progress Reports from Adoption & Integration Partners

Katherine Chiang, Kristi Holmes, Violeta Ilik and Christopher Shanahan

Abstract: With funding from the Alfred P. Sloan Foundation, the ORCID Adoption and Integration Program has provided external funding for universities and science and social science professional associations to integrate ORCID identifiers and support the collaborative elicitation and documentation of use cases in the scholarly domain. As awardees, Boston University, Cornell University, and Texas A&M University have engaged in complementary integration projects addressing different key elements of the academic research networking ecosystem -- faculty members, college-level administrators, and graduate students through the thesis and dissertation process.

ORCID provides a path forward at a global scale for addressing disambiguation issues that have significantly reduced opportunities for networking researchers across projects, departments, and institutions. ORCID is an important part of

VIVO

connect • share • discover



any researcher's professional identity, and it is especially important for early career researchers to claim and use their ORCID iD to help ensure they get credit for all of their work. ORCID iDs can also benefit graduate schools, and postdoctoral affairs offices seeking better information about career outcomes.

Panelists will present their institutional experience(s) with ORCID integration and adoption ranging from proactive creation of ORCID iDs for all graduate students and/or faculty members to low-key, distributed engagement at the individual college and department level to raise awareness across a heterogeneous research university. Each panelist will also describe the role of a research networking system (Profiles or VIVO) with respect to ORCID adoption at their institution.

Long Paper Still to be Named

Reserved one hour slot for additional long paper

Short Papers

The Griffith Scholars Hub - VIVO Extensions

Arve Solland

Abstract: In 2012, when Griffith University went live with its Research Hub, profiling research data at the University, the site quickly rose to be one of the top 3 visited sites at the University. Its relevant linked data and search functionality made it a big success, winning 2 major awards in the space of a few months.

Now almost successful 2 years after going live, the Research Hub is changing into the Scholars Hub, aiming to profile not just research data, but also academics at the university. This phase will see a massive increase in the data ingested into the hub, and this has led to further development, optimization and extension of our VIVO instance to handle the new load, content and features required for the Scholars Hub.

The following topics will be discussed:

- Current Status - Some stats and facts on how the Scholars Hub implements VIVO and how it is performing.
- Dynamic Micro Portals / Lenses - Using rdf and SPARQL to create micro portals within VIVO.
- Data Ingest methods - Loading data using an ETL server, creation of difference models and post ingest enhancements applied to the data.
- External Data consumption - Identifying external data that is related to objects in the Hub, harvesting this data, and then notifying and allowing users to confirm these relationships in a way ala the Facebook timeline.
- Visualizations - Our creation of useful, interactive javascript visualizations based on data in the hub.
- Presentation – Optimizing the front-end of VIVO, making it fully responsive.
- Edit Interface – The creation of a new and snappy ajax driven edit interface, removing the long load times of past.
- Integration - How we currently are integrating, and are planning to integrate the Scholars Hub with other university and external systems.
- Planned additions and improvements – Our roadmap to improving the Scholars Hub VIVO instance.

VIVO: Bringing Together Cancer Researchers in India and South Asia

Anil Srivastava, Hemant Darbari, R.A. Badwe and Purvish Parikh.

Abstract: Open Health Systems Laboratory (ohsl.us) has been working over the years to deploy semantic web tools (VIVO and Eagle-i) to create team science collaborations for biomedical research coupled with its effort to create a global cyberinfrastructure— ICTBioMed: International Consortium for Technology in Biomedicine (ictbiomed.net).

The pilots initiated by OHSL have lead to two major projects in India with VIVO and Eagle-I in India supported by CDAC's team. CDAC and OHSL has shared with participants in VIVO conferences their effort to build capacity and support use of VIVO for creating teams for biomedical research using the semantic web technologies represented by VIVO and Eagle-I and the principles of team science. The present paper will report and describe the use of these semantic web tools in these projects.

Tata Medical Hospital (TMH) and its affiliate research center, ACTREC, are implementing a portal will be using VIVO to develop and maintain the digital curriculum vitae of its clinicians and researchers. Together TMH and ACTREC constitute Tata Medical Centre (TMC) and their portal will .

This is being expanded to include the oncologists and biomedical research resources across India's National Cancer Grid (NCG) which consists of 41 cancer centers across India being connected by the National Knowledge Network (NKN) and led by TMC. CDAC is working with TMC and OHSL to integrate technology tools including semantic web tools.

Further developments being integrated are:

- (a) Toolkit of open source harvesting tools for extraction of data of non-semantic and unstructured web sites;
- (b) Integrating human mediation and machine learning in the harvesting process;
- (c) Big Data to Knowledge to enrich digital curriculum vitae; and
- (d) South Asia being a multilingual environment, MAchiNe assisted TRAnslation Tool (MANTRA) for translation.

OHSL is discussing with the SAARC Federation of Oncologists and South Asian Journal of Cancer the deployment of VIVO and Eagle-I to cover oncologist and biomedical research resources across South Asia as serve that will reinforce the cooperation between oncologists in the region and create a map of cancer research and treatment resources in South Asia.

This effort to use VIVO and Eagle-i to map cancer researchers and resources, combined with the ICTBioMed: International Consortium for Technology in Biomedicine (ictbiomed.net) and the Indo-US Cancer Research Grid (IUCRG) is beginning to serve as a platform for collaboration in cancer research and treatment between South Asia and the United States.

The presentation will describe the status and lay out the plans for further and fuller implementation of semantic web technologies to create international team science for cancer research.

Enriching researcher profiles with Altmetric data

Euan Adie and Catherine Chimes

Abstract: The growth in the online dissemination and discussion of academic research, and the ability of altmetrics to monitor them, present a valuable opportunity for research managers and institutions as a whole.

Altmetrics tools can capture not only the downloads and citation counts for a specific article, but also enable us to understand the broader and societal impact of research. Through data such as mainstream media mentions and cites in patents and policy documents we can gather a much richer picture of the eventual outcomes of the work.

By tying this in with the structured staff and bibliographies in VIVO it is possible to explore the data by person, department or across a whole institution - allowing you to collate, monitor and report out on the attention that research published by your faculty is receiving.

In this session we will show examples of just such an integration, and explore the ways in which administrators and the authors themselves can use the data to maximum benefit.

New Tools for VIVO Developers

Jim Blake

Abstract: This session looks at two tools for VIVO developers: the VIVO API and the VIVO Developer panel. The API helps developers who write tools that work with VIVO. The Developer Panel helps developers who are working with VIVO itself.

VIVO is working to enhance its API, with improvements in releases 1.6 and 1.7. The Linked Open Data API has been streamlined and enhanced. SPARQL APIs have been added to permit updates as well as queries. The ListRDF API has been formalized. All of these APIs have improved content negotiation and compliance to standards, including JSON-LD output.

The VIVO Developer Panel, introduced in release 1.6, helps developers see what is happening inside their VIVO, in real time. Follow authorization flow, obtain timing information, add diagnostics to the generated HTML pages - all without stopping VIVO.

In this session, we will see how these tools are used. We also welcome discussion about tool for future releases.

VIVO and current research information systems in Germany: projects, trends and challenges

Lambert Heller, Gabriel Birke, Ina Blümel, Martin Mehlberg, Christian Hauschke and Robert Jäschke

Abstract: Current research information systems have existed in Germany for a few years now: some of the major universities and a number of facilities belonging to the four large German scientific associations (Max Planck, Fraunhofer, Helmholtz and Leibniz) are in the process of introducing institutional current research information systems. Exchange and standardisation in this field are being fostered by DINI's (German Initiative for Network Information) Current Research Information Systems working group (established in 2013) and the "Research core dataset" project initiated in 2014 by the German Council of Science and Humanities (Wissenschaftsrat).

Unlike in a number of Scandinavian and Eastern European countries, a centralised national infrastructure for research information is hardly conceivable in the highly federalised German research landscape. As a leading research nation with such a diverse landscape of autonomous scientific institutions, Germany is a test case for distributed research information that has to be combined nationally, interlinked and evaluated.

Some scientific libraries in Germany are well prepared for this challenge, thanks to intense activity in developing linked open data applications and online university bibliographies. For this reason, VIVO aroused considerable national interest last year, particularly due to activities undertaken by TIB's Open Science Lab, not least in the context of the Leibniz Research Association "Science 2.0", co-founded by TIB.

After giving an overview of the situation in Germany and highlighting trends, three research and development projects currently being realised at Hannover University of Applied Sciences and Arts, L3S Research Centre and TIB will be presented.

- VIVO as a current research information system at a large German university of applied sciences: Christian Hauschke, a librarian at Hannover University of Applied Sciences and Arts, began to map the entire HEI's research landscape in VIVO within a very short space of time; the search for an efficient data management and linked data enhancement workflow is currently under way.
- VIVO as a harvester for researcher profiles in "Science 2.0": Ina Blümel and Gabriel Birke are developing a prototype for a VIVO harvester on the topic of "Science 2.0" as a project involving students from Hannover University of Applied Sciences and Arts and Hamburg University of Applied Sciences.
- "German Academic Web" (GAW) application to the German Research Foundation (DFG): Led by Robert Jäschke (L3S), L3S Research Centre and TIB are planning to crawl all 425 state universities and 269 institutes of scientific associations in Germany on a monthly basis. The data obtained will be processed in an automated workflow using information extraction methods, and the results made available as open data.

Building VIVO-based research management infrastructure for Higher Education: the IREMA case

Anastasios Tsolakidis, Cleo Sgouropoulou, Evangelia Triperina and Panos Kakoulidis

Abstract: Higher education institutions need to capture the entire activity that happens within their academic units for several reasons, including quality assurance, strategic planning and dissemination of their results. Apart from the facilitation of the various academic processes, this information is also used for networking purposes as well as for deducing meaningful results.

In this paper, we present a system that aggregates research and education records and services. The aim of our system is to offer a solution for the management and the manipulation of academic information for evaluation and quality assurance of our institution and its departments. Our approach builds upon the VIVO ontology and introduces an elaborate research management information system that is called IREMA (Institutional REsearch MANagement).

Within IREMA we have implemented and profiled VIVO for our academic institution, Technological Educational Institution (TEI) of Athens, Greece. We have adopted the VIVO-ISF ontology and extended it in order to meet the needs for quality assurance in education and research. This has led to the creation of the AcademIS ontology. All the information of our system is stored in the VIVO instance and the web services are built upon VIVO to access the data. The academic networking requirements of TEI of Athens are also covered by VIVO, while IREMA accesses the institutional data stored in VIVO by executing SPARQL queries in our instance and the outputs are acquired in .json format. IREMA then reuses the data of our VIVO instance, applies metrics on the data and visualizes the results. Moreover, based on the data retrieved by the VIVO installation, we implement a decision support mechanism aided by the aforementioned visualization of the data. IREMA provides efficiency measure techniques, data mining methods (cluster analysis, association rules and bayesian networks), as well as social network analysis (including community detection, identification of research hubs and discovery of the most important researchers based on specific characteristics).

Our system builds upon the renowned ontology VIVO-ISF, uses the sharable and reusable linked data produced by the VIVO semantic web application, and implements web services that are supplied by VIVO to meet the existing needs in the academic setting. As a direct consequence, our solution is both sustainable and extendable.



Situating VIVO within a Research Analytics Architecture

Simon Porter

Abstract: The story of a VIVO implementation is often the story of the integration of previously disconnected sets of information: students, research, HR, finance, facilities and research infrastructure. Public exposure of this information also places demands on the quality and completeness of this information.

A byproduct of this integration is that many implementations find themselves able to answer strategic research questions that no other university system is capable of answering, such as the structure of interdepartmental collaboration patterns.

It is tempting then to re-imagine the research profiling system as the beginning of a research-focused data warehouse. Its value is realised through not only the public syndication of a university's research activity, but also in terms of analytic insights that this data can produce for strategic dashboards.

This approach, however, can lead to tensions internally with other data warehouse and reporting activities that might be ongoing, but also externally in terms of defining what can be communicated publicly rather than only for internal reporting purposes.

This presentation draws on the research analytics experience at the University of Melbourne. It intends to present a three-tiered approach to Research Analytics that situates VIVO within the context of the broader university analytics functions:

- Public Research Profiling (Business Intelligence for the public, industry and open research community)
- Internal Research Analysis (Business Intelligence for the broadest possible internal research community focusing on rapidly constructing research narratives)
- Research Performance Reporting (strategic dashboards for performance measurement)

As part of this presentation, specific examples of research analysis will be drawn out at each level, including capability mapping over profile data, the integration of Altmetrics and VIVO Google Analytics into university research reporting, and the creation of research dashboards for performance assessment.

OpenSocial in Practice – Easy Customization Benefits Everyone

Brian Turner and Eric Meeks

Abstract: Background: With the OpenSocial framework, UCSF has transformed Profiles and VIVO into extendable web platforms. OpenSocial plugins are relatively simple and quick to build and deploy. Locally, UCSF has deployed many plugins that are in heavy use and have made a real impact to our research networking system and researcher community.

Discussion: OpenSocial has proven itself valuable to research networking systems (RNS) in numerous ways:

- With OpenSocial we have added valuable and visually appealing “grey literature” from commercial media sites into our RNS.
- OpenSocial plugins are a great way for institutions to represent institutional specific data in their RNS because they do not require altering the RNS code or ontologies.
- OpenSocial plugins can facilitate active collaboration by connecting to external tools (e.g. Doodle polls, Chatter, or Dropbox).

Results from the 2014 survey of VIVO implementation sites

Paul Albert, Alex Viggio, Kristi Holmes and Jonathan Corson-Rikert.

Abstract: VIVO implementation sites have made significant resource and time commitments to the project. As with many open source projects, contributors at local sites have occasion to define their needs through a variety of channels: listservs, discussion in regular working group calls, and at the annual workshop and conference. This dialogue has helped the project amass an impressive wish list of features and functionalities. At the 2014 VIVO sponsors’ meeting, Dean Krafft suggested seven competing but not mutually exclusive visions of the VIVO project.

But the energies of the contributors to the project are finite and not all improvements to the code or community would have the same positive impact. In order to better define what should be the priorities of the project going forward, we circulated a survey containing over one hundred questions to all VIVO implementation sites in April and May. We asked institutional representatives for VIVO sites questions covering areas from server configuration, apps and tools, core development, outreach, and performance.

We will present on sites’ experiences and most pressing needs in the hopes that we can clarify the priorities for the project going forward.

Using OAI-PMH Protocol for Data Ingest into VIVO Instances

Alexandre Rademaker and Violeta Ilik

Abstract: In this article and presentation we aim to demonstrate the tools developed by FGV and Texas A&M for data retrieval from DSpace instances using OAI-PMH protocol, transforming the data into VIVO compatible RDF using XSLT and ingesting the data into VIVO using the VIVO SPARQL Update API. We plan to discuss the benefits of such tools, the current workflows at our two institutions, and the limitations.

De-Facto Standards Interoperating in the Real World - VIVO/ORCID/CASRAI

David Baker, Thorsten Höllrigl and Rebecca Bryant.

Abstract: A compelling and dominant use case for the research management domain is the discoverability of researchers based on their expertise. An approach to meet the requirements of a larger research information landscape needs to apply business context to constrain the whole-world into smaller semantic packages. These packages (CVs, finance reports, HR reports, outputs reports, impacts reports, etc.) need to be exchangeable with relevant systems in multiple domains such as HR, Finance and Research Information systems. To guarantee a scalable approach, these packages need to align with community-driven efforts at standardizing IDs and models for these package exchanges. VIVO and ISF already align with ORCID for the standard ID part of the problem and we plan to focus on the standard model part of the

problem.

In particular, there is an increased need to communicate research information between universities and funders - in our experience HR and finance systems are typically a poor source of information about researchers and their research. Instead, the CRIS (Current Research Information System) provides the framework for consolidating, enriching, and communicating: to the public and other researchers (research portal) and funders (through funding applications). In order to communicate researcher biosketch information to a multitude of different funding systems, we need standards. Several funders now require structured CVs for grantees, such as the Canadian Common CV and the U.S. SciENcv, offering exciting opportunities to standardize and exchange data, improving reporting capabilities, and reducing redundant data entry. Being able to communicate standardized researcher information in a variety of formats and contexts would benefit from a standardized approach to allow users to correctly interpret and map individual data elements in the target system.

In this talk we discuss the general contributions of ORCID and CASRAI, identify points of collaboration, and provide examples of how organizations are using these standards within research administration solutions such as CONVERIS.

The Semantic CV - CASRAI in VIVO/ISF

David Baker, Thorsten Höllrigl and Albert Bokma

Abstract: There is an increased need to communicate research information between universities and funders - several funders now require structured CVs for fundholders and that entails data exchange between institutions and their respective systems. In our experience HR systems tend to be poor when it comes to rich information about individuals and particularly about their research. Typically one can get basic person details and funding details out of HR and finance systems, but detailed research activity data is usually not supported. The CRIS (Current Research Information System) is the place where this basic information gets enriched and from which one can begin to generate fuller records such as detailed CVs that can then be communicated to funders for the purpose of funding applications. There are thousands of funders, with different systems and data models, and thus to communicate CVs effectively we need standards. Being able to communicate conceptual 'packages' in a variety of formats and contexts would benefit from a semantic approach to allow to reliably interpret and map the individual data elements correctly in the target system. Thus not only providing the data but alongside communicate the meaning through the use of ontologies.

We will discuss how CASRAI profiles can help provide standard 'message' formats and models and therefore bridge the gap on inter-organisational integration use cases. Using the example of the exchange of CV data we will present an ontological model of an underlying branch of the CASRAI research administration dictionary (canonical domain model - CDM) and an RDF model of a business exchange (canonical messaging model - CMM) derived from the CDM.

In this talk we discuss the general challenges and opportunities of expressing the CASRAI dictionary components in an ontological model and, specifically ones that would integrate with VIVO/ISF. We will identify overlapping parts, make clear how those efforts do complement each other and demonstrate how to implement and use them in combination within a typical research administration solution such as CONVERIS.

We are proposing to enhance CASRAI to provide semantically rich information for the purpose of messaging. We present

a model for a semantically enhanced CASRAI standard that is capable of integrating at a semantic level. We present a model from which we can communicate CV details to a VIVO installation out of a system such as Converis using CASRAI as the standard.

VIVO Dashboard: a Drupal-based tool for harvesting and executing sophisticated queries against data from a VIVO instance

Paul Albert, Miles Worthington and Don Carpenter

Abstract: Administrators at Weill Cornell Medical College (WCMC) have repeatedly expressed interest in making decisions based on evidence. One key piece of evidence is publication output. Any requests for reports are often more sophisticated than simply a count of publications. For example, the Dean's Office has asked to regularly see a quarterly list of journal articles in high-impact publications in which an institutional author was first or last rank.

The VIVO implementation team at WCMC uses its instance of VIVO to authoritatively track the College's record of publication output, but these records are tedious to query using SPARQL and are not outputted in a variety of user-friendly formats.

In 2012, Weill Cornell began the VIVO Dashboard project in order to harvest publication metadata from VIVO and display it within a Drupal-based faceted interface. The goal of VIVO Dashboard is to empower the untrained user to easily produce reports in whatever format they desire including spreadsheet, bar graph, HTML list, and Word document. End users can also facet publication records by publication type, journal name, author name, author type, and organizational affiliation.

Because publication metadata is always changing, we also wanted to provide a means by which administrators could recurrently update VIVO Dashboard with new data. Administrators of VIVO Dashboard have several ways to import records. The most straightforward way is to point it at your VIVO. Over the course of several days, VIVO Dashboard will routinely retrieve batches of RDF data from the VIVO and turns it into native Drupal content using Drupal's Feeds module.

The VIVO application goes to great lengths to expose open linked data, and VIVO Dashboard serves as an example by which applications can easily consume structured data via VIVO's intended mechanisms.

An example instance of VIVO Dashboard can be viewed here:

<http://dev-vdb.gotpantheon.com/>

The open source code follows:

<https://github.com/paulalbert1/vivodashboard>

Never Throw Anything Away! Using External Triplestores for Data Ingest, Archiving and Improved Data Quality

Chris Westling, Brian Lowe, Jon Corson-Rikert and Joseph McEnerney

Abstract: VIVO is a useful tool for aggregating data from disparate ingest sources, and providing linked data to the world. What happens to the data that you'd like to keep, but don't want in your production VIVO? At Cornell University, we are experimenting with the use of "off-line" triplestores that can provide a staging area for ingest RDF, and provide a method for archiving triples as they are added to or removed from the production database. We will discuss the use of Sesame triplestores that allow us to better manage multiple ingest processes, and we will demonstrate methods for cross-referencing data before creating the "finished product" for our production instance. Topics discussed will include the implementation of D2R Server, Joseki, Fuseki, and Sesame at Cornell, as well as the creation of custom ingest processes that utilize these external stores.

An overview of existing data ingest strategies will be covered, and several different types of data quality problems will be discussed. Following the overview, we will introduce the results of our research about how using multiple external triplestores help us populate our production VIVO instance with more reliable data, and allow us to retain a historical record of triples that previously existed in our production instance. We can use these flexible external triplestores as a way to process our ingest data from several sources, identify problems with the source data, fix what's broken *before* it gets to our production instance, and establish methods for correcting bad data at the source. We'll also discuss the possible problems and complications of "data hoarding", and some strategies that can help alleviate problems before they appear in VIVO

Integrating Symplectic Elements with VIVO 1.6

Daniel Grant, Mary Walters, Michael Mitchell and Tim Morris

Abstract: Background: The scope of the Symplectic Elements application has increased dramatically. Originally a system that served as a research management information system that focused primarily on faculty's publication data, it has now grown to be a faculty information system that can support virtually any and all data related to a faculty member and their related accomplishments. From grants to teaching to other professional activities, Symplectic Elements is become a source in which faculty information can be managed. However, the VIVO platform is the perfect choice to publicly showcase the faculty infrastructure we have developed in Elements.

Problem: Emory needs to develop a means for the Elements data to feed into VIVO. In addition, this is a need shared by many other schools who are adopting Symplectic Elements.

Solution: While there are existing connectors for bringing information into VIVO, our goal is to address the need from both a functional and technical perspective to facilitate the use by other universities.

From a functional perspective, we are going to identify step-by-step guidelines for implementing the new harvester as well as develop a set of supplemental documents to facilitate structural analysis, including the most essential VIVO data elements to map to a Symplectic system and examples of the XSLTs and RDFs.

From a technical perspective, we are going to build on the existing VIVO Harvester (developed by University of Florida, <https://github.com/vivo-project/VIVO-Harvester>) and the Symplectic Harvester Extension (developed by Symplectic, <https://github.com/Symplectic/vivo>). The mapping will need significant rework to account for the new VIVO 1.6 ontology and the additional information modules Elements has since built. Also, a limitation of the Harvester in its

current form is that the index needs to be manually rebuilt once the information has been added. With the new SPARQL endpoint API introduced in VIVO 1.6, we are hoping to modify the harvester such that the index will not require rebuilding, and thus make the Harvester solution more attractive to other schools, especially those interested in importing data from Elements to VIVO.

Representing Arts Scholarship at Duke: Design, Development and Migration of Artistic Works Data

Julia Trimmer, Damaris Murry and Patrick McElwee

Abstract: Duke University recently rolled out their VIVO implementation, Scholars@Duke, to faculty in the arts and humanities. At the first Steering Committee meeting, Duke's provost stated his strong commitment to serving the needs of the arts departments. He made it clear that the system must represent arts scholarship in a way that was analogous to grants and publications. Knowing that the tools available for harvesting information for faculty profiles supported STEM fields and other disciplines that publish journal articles, the project team accepted this challenge without a clear idea of the way forward.

This presentation will outline the process of analysis, design, and development of the ontology and functionality for developing Artistic Works in VIVO/Scholars@Duke. We will discuss our experiences for engaging departmental users in the development process and incorporating their feedback. We will also outline the methodologies for ingesting the artistic works data from departmental web pages. In this presentation, we'll talk about what seemed to work well and what we learned that might be applicable to other rollouts.

Building a Knowledge Base for the APA Digital Library, Semantic Enrichment of Literature in Psychology

Alexander Garcia, Leyla Jael Garcia, Bryan Dennis, Eva Winer, Gary R. Vandenbos and Hal Warren

Abstract: We are building a knowledge base for scientific literature in psychology and behavioral sciences. We have generated a semantic model representing structural elements in the document; the content is annotated by using robust techniques from natural language processing (NLP) including named entity recognition (NER), word sense disambiguation (WSD) and deep semantic analysis. The RDF output is embedded in the Linked Data cloud by using vocabularies from the Semantic Publishing and Referencing Ontologies (SPAR), Dublin Core (DC) and VIVO. Domain knowledge is annotated using the APA ontology. The APA ontology is the product of decades of working with psychological documents; it has been built upon the APA Thesaurus of Psychological Index Terms. Expert curators manually extracted the terminology and extensively discussed the nuances of the classification schemata. Although the APA ontology covers the domain of psychology, we are also using CHEBI for chemicals, NDFRT for drugs and other ontologies such as FMA, and ICD10. We are experimenting with other biomedical ontologies (e.g., SNOMED and the NCI Thesaurus) to better define our domain model.

The application of semantic-based annotation over large digital libraries delivers a more expressive query-model than that based on simple keyword-based methods. Not only can semantic queries be formed, but axioms describing relations among concepts make it possible for a user to derive information that has been specified only implicitly. Also, as the resulting data set of self-describing documents is fully interoperable, content enrichment is a simple task; some of the datasets we are exploring for content enrichment include DBpedia, VIVO, and LinkedCT. In addition, we are enriching VIVO profiles with topics extracted from the annotated content. In this way, VIVO profiles are able to accurately represent topic expertise for a given researcher; these are linked to the papers. In coordination with VIVO data, our semantic annotated data set also makes it possible to produce maps illustrating clusters of research that can

be granularly filtered by topics (e.g., where is the research about Post-traumatic stress disorder (PTSD) concentrated?) We have initially focused our efforts on the Archives of Scientific Psychology; as our exploration of feasible uses of Linked Data evolves, we expect to have more content and user-oriented functionalities making use of this format.

The stewardship challenges of long term open scholarship: An open access institutional repository perspective

Simone Sacchi and Amy Nurnberger

Abstract: While in the past institutional repositories (IRs) have been primarily concerned with the availability of traditional research outputs in the immediate time frame, in the ‘now’ —with a limited set of stewardship concerns, mostly related to textual resources— they are currently faced with challenges in curating a potentially high heterogeneity of complex digital resources, and their relationships, in support of ‘long term open scholarship’.

One commonly understood role of an IR is to archive the research products and outputs of its institution. When an IR is also dedicated to making these resources openly accessible it acquires a role in supporting a broader open scholarship agenda.

The open scholarship paradigm has expanded the notion of research output to include research products such as datasets, software, scripts, and other types of complex and multimedia resources. These additions to the traditional research outputs of preprint and postprint publications pose unprecedented stewardship challenges, in particular when IRs are motivated to support the meaningful use and reuse in research and science in the long term. IR technical infrastructure, and policy and organization capabilities, must adapt to new scenarios.

Lessons learnt from the digital stewardship communities can inform this paradigm shift with regard to resource characterization (e.g., descriptive, provenance, technical and rights metadata), and the representation information necessary to enable an effective use and reuse of submitted resources outside their original creation context. We believe, however, that a broader discussion about these new requirements calls for redefining the intersection between scholarly communication activities, stewardship activities, and the role of IRs as long term resource management platforms. Here we propose a discussion around the following points:

- Are there specific technical and policy requirements for open-access-oriented long term digital stewardship?
- How should IR infrastructures evolve in order to support an effective representation of complex research outputs, their relationships in support of their preservation and meaningful use and reuse in the long term?
- Repository initiatives oriented to specific domain or research or specific types of research outputs (e.g. datasets, code, etc.) are already addressing some of these issues: how can IRs leverage these initiatives in an institutional context?

Improving accessibility and cost-effectiveness of research publications: The role of public data

John Mark Ockerbloom and Manuel De La Cruz Gutierrez

Abstract: Sponsors of research are increasingly adopting or enforcing mandates to make the articles and data produced from that research openly accessible. They also have an interest in controlling the costs of making these research outputs available. Compiling and publicizing information on openly accessible resources and their costs can promote both of these goals. The Wellcome Trust's recent publication of their open access payments to publishers, for instance,

has allowed the research

community to monitor publisher compliance with open access licenses and mandates, as well as assess the cost-effectiveness of different publication venues. Likewise, the creation of live links to repository copies of papers in some VIVO instances allows readers and funders to easily access and verify open access copies of research results.

In this talk, we will discuss how open access and data provision requirements have been influencing the development and deployment of VIVO and Symplectic Elements at Penn. In particular, we will discuss how we, and other institutions, are collecting and presenting metadata relevant to these requirements, and consider reports and interface improvements in VIVO that could improve the visibility and effectiveness of open research and data.

Using an Enterprise Research Management System to maintain accurate and up-to-date scholarly data for campus re-use

Katy Borner, Ann Beynon and Timothy Otto

Abstract: This case study will explore recent developments to Indiana University's VIVO installation, focused on automated integration of publication and citation data from Web of Science Core Collection using the Converis research information system. Indiana University was a partner institution on the original VIVO project. The university had created a VIVO installation for its NetSci Institute, but maintenance of the data was problematic. In 2014, Indiana University conducted a pilot project with Thomson Reuters to help solve these challenges. Through this pilot, Thomson Reuters migrated data from the existing VIVO installation to Converis, and augmented the faculty's publication lists with further journal articles from Web of Science Core Collection. Faculty members were able to validate and augment their publication lists through Converis. Then, the data from Converis were imported into a sandbox VIVO environment. Through this pilot, we demonstrated how data can be exported from VIVO and used in a research information system, and how additional data can be added in the system and exported back into VIVO. Also, we are exploring how the combination of campus-wide faculty activity reporting with publication management and VIVO can help solve a variety of research data challenges. Those who are just starting to use VIVO can also benefit by utilizing profile management tools like Converis, which offer profile seeding services. Lessons learned from this pilot will be discussed, as well as the next steps.

VIVO Team Recommender Based on Linked Open Data Conforming to the VIVO Ontology

Noshir Contractor, Anup Sawant and Harshad Gado

Abstract: In earlier work, we developed a suite of heuristics for recommending collaborations between researchers. These heuristics are informed by empirical studies that test theories from the social sciences regarding the formation of effective collaborations and teams. In the current work, the initial prototypes of these heuristics were ported to operate over survey data collected in a class or conference settings. The 'VIVO Team Recommender' is a web application on top of VIVO data at Northwestern University.

Our efforts demonstrate that the architectures and programming techniques of the semantic web are well suited to the problem of building practical software tools that can be leveraged to apply to diverse sources of data. We found that the interoperability between researcher networking systems (RNSs) from diverse institutions and vendors offered by the VIVO ontology is a sound basis on which to build researcher team recommender tools.

The 'VIVO Team Recommender' embraces the World Wide Web Consortium (W3C) standard SPARQL query language for

real-time retrieval of semantic web data. We found the SPARQL implementation available in open source software to be robust. Further, because programmers can target only the particular data needed, performance is enhanced by reducing unnecessary network traffic. We found the learning curve and technical skills needed for SPARQL programming to be similar to that of more traditional, relational, SQL-based programming competencies which are more generally available.

“Give me one good reason . . .” A case study of rolling out VIVO across disciplines

Julia Trimmer and Lamont Cannon

Abstract: What’s it like to implement VIVO across all disciplines, including the arts, humanities and social sciences? In this presentation, we’ll present our own experiences rolling out Duke University’s VIVO implementation, Scholars@Duke, to Arts & Sciences departments in spring of 2014. This presentation covers lessons learned from this rollout, including our attempts to confront attitudes and fears about academic technology in traditional humanities and social science departments.

In this presentation, we’ll set the stage by describing the particular climate for these departments at Duke. We’ll state our plan and good intentions, and contrast them with the events as they actually unfolded. Finally, we’ll talk about how we’re addressing these adoption challenges and plans for moving forward, since we continue to look for ways to increase engagement with certain groups of users.

Is Presentation and Institutional Branding Really More Important than Linked Research Open Data?

Alex Viggio and Jonathan Breeze

Abstract: Despite the best efforts of VIVO’s growing community of developers, researchers and data scientists, today’s leading research institutions continue to spend significant sums of money and time developing and implementing static academic profile pages which lack the research discovery capabilities found in the VIVO linked open data platform.

Symplectic has been an active member of the VIVO community for several years, and in 2013 shared an open source extension to the VIVO Harvester project to help streamline the VIVO implementation process for research institutions by generating VIVO RDF from its Symplectic Elements platform.

Symplectic continues to support the adoption of VIVO by its institutional clients but in 2013 saw one of its clients elect to replace their public VIVO instance with a home-grown Content Management System due to concerns over the presentation of data to visitors of the their VIVO site.

In 2014, Symplectic launched an internal pilot project to explore how they might extend other aspects of VIVO 1.6. A second pilot project was also undertaken to understand how Symplectic’s earlier VIVO integration could be re-purposed to provide support for passing Elements data to Profiles RNS implementations.

In this session, the presenters will explore lessons learned from the past few years of engagement with the VIVO project, attempt to explore the reasons behind the continuing trend for selecting CMS styled profile solutions and present the findings of their recent VIVO 1.6 and Profiles RNS 2.x pilot projects.

Comparing author- and institution-based bibliographic search results across databases and their viability for populating researcher profiles in VIVO

Jeremy McLaughlin

Abstract: The VIVO project has exposed the inherent difficulties found in working with researcher and research-output related data. Institutions currently using or considering an instance of VIVO must decipher the process of not only creating but also managing and continually updating person and bibliographic data from multiple sources. It is clear that one size does not fit all when it comes to bibliographic data harvesting needs, and one source of data is not sufficient for most VIVO instances.

This paper builds off of the work done at the University of New Mexico testing the accuracy of bibliographic results in the VIVO Harvester compared to researcher publications in PubMed. Using a set of researchers from several institutions, we compare the necessary search strategies and the search results of popular bibliographic databases to results from the VIVO Harvester, public VIVO instances, and researcher's "gold" list of publications. The discussion will focus on the process of searching and finding accurate publication lists in various databases, a comparison of results across databases and institutions, and the applicability of publications harvesting from specific databases across disciplines. The results are of particular importance to organizations currently working to populate a VIVO instance as well anyone interested in author identification, bibliographic data management, and the viability of using data from popular databases in the VIVO community.

The Case for Stable VIVO URIs

Melissa Haendel, Brian Lowe and Violeta Ilik

Abstract: VIVO instances typically are populated with authoritative institutional data about their researchers, including departmental affiliation, research specialization, research products, and other publicly available information. The scholar's web presence is fragmented based on discipline, and also based on the fact that sometimes scholars deliberately maintain distinct identities (e.g., publishing in different subject areas, writing under pseudonyms, change institutions, etc.). Therefore, there exists a pressing need for standardizing and distinguishing authoritative researcher profiles, above and beyond researcher identification (such as ORCID).

In the Spring of 2013, the Program for Cooperative Cataloging (PCC) received the "Report for PCC Task Group on the Creation and Function of Name Authorities in a Non-MARC Environment." The most significant contribution of this report was the acknowledgement of the need for use of non-MARC authority records in bibliographic data. Not every researcher is represented in the Library of Congress Name Authority File and that is an important reason to open the door for use of non-MARC authority records in bibliographic data. VIVO is positioned as one of the most promising linked data platforms for managing researcher profiles and as such holds promise to establish itself as an authoritative source. We propose that the use of VIVO URIs can provide quality authoritative data as one of the main data sources for the Virtual International Authority File (VIAF) and also in other bibliographic contexts. Librarians as well as authority vendors look for stable representations or identifiers for controlled headings (names, conferences, corporate bodies), and the VIVO community can take advantage of this opportunity offered by the expansion of the PCC procedures to allow for incorporation of authenticated sources beyond the Library of Congress Name Authority File.

VIVO

connect • share • discover



Ensuring long-term stability of VIVO URIs may suggest the need for some centralized infrastructure, or at the very least, common data retention policies and URI redirection implementations. While the VIVO community's current decentralized approach has had the advantage of not requiring specific ongoing financial resources, the goal of truly linked interoperable data that uses common identifiers across institutional boundaries has remained somewhat elusive. It will become increasingly important for the community to consider the possible advantages of a centralized URI registry (perhaps with owl:sameAs links to local linked data URIs) and of commitment to maintaining the availability of historical data that can aid future efforts in name disambiguation.

Poster

Mirror, Mirror, Who's Looking at My Profile? (or How to Harness the Power of Google Analytics and Vanity)

Nooshin Latour and Anirvan Chatterjee

Abstract: Professional social networks like LinkedIn send users smart data-driven emails to increase engagement and stay on the forefront of their minds. LinkedIn's personalized email content include the total number of times a user's profile has been viewed (paid members can see exactly who) to suggestions on others to connect with based on similar skills, career and educational history. Delivering personalized product updates and useful information right to the user's inbox help make the social network "sticky," and incentivizes users to keep their LinkedIn profiles up to date. UCSF informatics and communications professionals have leveraged UCSF Profiles web analytics data to deliver a customized "2013 UCSF Profiles Annual Report" to individual researchers at UCSF and other campaigns to increase user adoption.

CONVERIS 5 - Next Generation Research Management System - Unifying Standards and Flexibility

Thorsten Hoellrigl and Jan Maier

Abstract: CONVERIS is a state-of-the-art Research Management System that provides all levels of the organisation with quality assured information on its research activities and results. Its high degree of connectivity with other systems opens the door for managing all data along the Research Life Cycle in one integrated system.

Research institutions are facing an increasing competition for funding, research excellence and talent. It has thus become essential to keep track on publications, projects and other research information in order to facilitate monitoring and reporting, to strengthen grant applications, to enhance the public presentation and to improve research performance.

CONVERIS supports the complete research life cycle from an initial idea, over grant applications, through to the day-to-day management of ongoing projects, and their results, including publications, patents and many other outputs. On top of this, CONVERIS offers an extensive support for exploiting the information through several channels, including exports, reports, CVs, public presentation over the web, and much more.

For an efficient data handling CONVERIS collects data from external and internal sources. Web of Science, PubMed, Scopus, MS Academic Search, EuroPubMed and contract information from funding bodies are a few examples of external systems that CONVERIS integrates with. Internally, connections are made to login servers, HR, Finance, Pricing & Costing, Student records and Institutional repositories to continuously keep key information up-to-date in the right place.

As an active member of euroCRIS, ORCID, CASRAI and of course VIVO – organisations dedicated to improve the use of Research Information Systems and their interoperability – AVEDAS has the proliferation of the CERIF data model (Common European Research Information Format) as well as of CASRAI (Consortia Advancing Standards in Research Administration Information) as one of its primary goals.

Through our active participation in standardisation events and in relevant Task Group, we contribute to the standards with a focus on keeping them as practical, usable standards. At the same time this allows us to ensure compatibility of

CONVERIS with those important standards. In fields where application areas and functionality extend beyond the standards, CONVERIS architecture of modular and freely definable entities, relations and attributes allows radical extensions of the data model. It can thus be easily adjusted to an institution's organisational and structural characteristics providing the flexibility a modern and state-of-the-art research information system required for the dynamic and heterogeneous research administration domain.

Visualization for Research Management

Jeff Horon

Abstract: Universities and funding bodies are placing increasing emphasis on return on investment (ROI) related to research. Research managers at all levels need objective metrics and data, further developed into visualizations, that provide insights to support decisions about investments and that also promote understanding of the outcomes of those decisions.

It is critical to have visibility to both inputs and outputs related to research, and VIVO-compatible data can be used for these purposes. Examples demonstrated will include:

- An organizational dashboard used by top-level administrators at a large research organization (>\$0.5 Billion annual research expenditure)
- Benchmarking
- Faculty activity reporting
- Recruitment and retention analysis

(examples will be primarily focused on end-user presentation, but some data translation will be discussed, along with supplementation of open information resources with confidential 'internal' data resources)

Seeing is Believing: Promoting Cross-Institutional Collaborations By Interlinking Research Networking Systems

Eric Meeks, Praveen Angyan, Ed Ward, Aditya Arun Vaidya, John Burgos, Brian Turner, Leslie Yuan and Katja Reuter

Abstract: Large research organizations have become increasingly interested in Research Networking Systems (RNSs) as a means of addressing multidisciplinary research challenges, fostering collaboration, and increasing the visibility of individual investigators. Many RNSs are locally installed, as this gains the benefit of ownership and provenance, but our collaborations are not limited to our local institutions. Additionally, efforts to tie together multiple RNSs that manage and discover knowledge about researchers across institutional and disciplinary boundaries have historically been focused on web-based search as evidenced by projects such as Direct2Experts and VIVO Search. An additional approach is to connect our systems through simple but relevant hyperlinks on the profile pages themselves. In promoting RNSs that produce Linked Open Data we have solved half of this problem. We now need to consume the data for features that are both useful and used in order to complete the solution.

It's All About the Links: An Automated Approach to More Efficiently Promote Research Networking Systems

Anirudha Kumar, Praveen Angyan and Katja Reuter

Abstract: *Introduction:* Large research organizations have begun deploying Research Networking Systems (RNSs) that allow for easy search of their research experts and related networks as a means of fostering research and collaborations. However, increasing the adoption of RNSs has posed challenges. Previous data show that cross-linking (i.e., links to the RNSs on other websites) serves as an effective tool to significantly increase the traffic to and adoption of RNSs. A strategy to embed those cross-links on other websites requires the establishment of partnerships and is a time intensive effort that may take months or years before yielding significant results. We wanted to provide a more effective solution and developed an automated approach (i.e., NLA, Name Linking Application) that identifies researchers' names on web pages (e.g., University news, directory, department pages) and links them to their respective researcher page (if it exists). The tool can be extended for any RNS (including VIVO) with minimal code modifications.

Methods: Our NLA utilizes natural language processing based on Named Entity Recognition (NER). We use an NER library based on CRF (Conditional Random Fields) originally developed at Stanford University to identify names in text. It has been shown that Stanford's NER Library gives an accuracy of around 91%. Our application then searches the RNS database for the found names. Every match in the RNS database gets a score based on the quality of the match. We apply a threshold score to filter out distant (poor) matches, which further filters out false positives. To make NLA easy to use, we have given it a RESTful interface to which any entity on the Internet can send content scanning requests. Thus, the NLA can be used with any website i) dynamically, by injecting a client side script into the website's pages, or ii) statically, by integrating NLA with a Content Management System (CMS) used by the website.

Results: The effectiveness of the tool is being assessed focusing on two key questions: How efficient and accurate are the linkages NLA creates? How effective is the tool to increase referral traffic back to an RNS?

Conclusion: Previous data show that the RNS search traffic is heavily dominated by commercial search engines such as Google and Bing, which is why we believe that this application will serve research organizations well to efficiently promote and increase the adoption of their RNSs without requiring time-intensive manual cross-linking.

International Standard Name Identifier (ISNI) and Linked Data

Marie Linvill and Laura Dawson.

Abstract: ISNI (the International Standard Name Identifier) is the ISO-certified global standard number that seeks to identify the names of researchers, scientists, publishers, writers, artists, musicians, politicians, and other public figures and contributors to creative or scientific works. The objective of ISNI is to provide links and interoperability among as many data sets as possible, disambiguate the names of researchers and authors, and provide accurate URIs on the open web. This poster provides detailed information on where ISNI is being used, member organizations, and benefits to the research community. It also explains how the role of a 'bridge identifier', a standard suited to cut across all portions of human activity, not just scholarly research, could benefit a software system like VIVO.

The REACH NC Resource Finder: a multi-institutional, statewide implementation of eagle-i

Sharlini Sankaran, Hong Yi and Michael Cherry

Abstract: The Research, Engagement, And Capabilities Hub of North Carolina (REACH NC) is a publicly-accessible, searchable, web-based portal that enables quicker and easier location of research expertise and assets in North Carolina. REACH NC currently contains profiles from over 9,000 university researchers from 19 North Carolina higher education and research institutions. REACH NC has recently launched a statewide resource finder tool, furthering its

vision of becoming a one-stop hub to connect research expertise and assets across North Carolina with economic development and other collaborative opportunities. The resource finder tool will contain searchable listings of core facilities, instrumentation, and other university assets across the 15-campus University of North Carolina system. The resource finder is implemented using eagle-i, an open-source, open access tool originally developed by the eagle-i Consortium, and currently supported by Harvard Catalyst, the Harvard Clinical and Translational Science Center and now adopted by 28 institutions across the country. Although eagle-i uses its own resource ontology, the recent release of the Integrated Semantic Framework (ISF), which aligns and merges eagle-i and VIVO ontologies into a single ontology, allows our eagle-i based statewide resource finder to link to other ISF-based researcher networking systems including VIVO. The REACH NC resource finder is the first implementation of an eagle-i network consisting of multiple institutional nodes and one central search node outside of the original eagle-i network. As such, the implementation of the resource finder has proved challenging and required extensive customization and set-up. The goal of this poster is to share lessons learned with others considering similar multi-institutional instances of eagle-i. We will discuss the process of implementation and adoption of the statewide REACH NC resource finder as well as potential opportunities for better integrating the Resource Finder with existing researcher networking tools such as Scival experts and VIVO.

Deconstructing a Dashboard: Inside the UCSF Profiles Team's Monthly Key Metrics

Anirvan Chatterjee and Brian Turner

Abstract: As the research networking community starts to establish usage tracking standards, it becomes more important to explore how different RNS management teams actually implement the standards to understand their platforms and drive product decisions. By August 2014, the CTSA Research Networking Affinity Group's Recommendations for RNS Usage Tracking will have been publicly available for over a year, but there has been little public discussion of what it means to adopt and implement the recommendations.

The UCSF Profiles team at UC San Francisco's Clinical & Translational Science Institute has published a monthly dashboard of site usage for over three years. The dashboard is compiled monthly based on Google Analytics data, stored in an Excel file, and emailed to about a dozen people every month. In addition to statistics, the monthly dashboard emails contain relevant commentary about site usage and traffic patterns.

Our poster will:

- share the UCSF Profiles team's internal monthly usage dashboard
- explain how each field is calculated and why it exists, e.g.:
- the importance of distinguishing between internal and external usage
- why break out heavy users
- when to include site performance characteristics, and when that may not be necessary
- provide context about the evolution of the dashboard over time

We anticipate that sharing the structure, content, and story of our dashboard will help spur conversations and encourage greater transparency and consistency around usage tracking throughout the RNS community.

Enhancing VIVO profiles by using a content-topic-based altmetrics

Alexander Garcia, Leyla Jael Garcia, Bryan Dennis, Eva Winer, Gary R. Vandenbos and Hal Warren

Abstract: We are generating annotations for American Psychological Association (APA) content; initially, we are focusing our efforts on the Archives of Scientific Psychology. Users annotate documents and comment on annotations. Annotations and comments are semantically processed using the APA ontology for Psychology; we are identifying domain knowledge topics. We use the Open Annotation Framework for representing annotations and a custom made ontology for structuring conversations. The metadata and bibliographic references for the annotated content is captured and represented by means of the Semantic Publishing and Referencing Ontologies (SPAR). Annotations are locally stored and also propagated over social networks such as Twitter, Mendeley, Facebook, and Research Gate. We record all the activity that every annotation generates over these social networks. We are facilitating interaction based on topics in annotations; as annotations are also exposed over existing social networks we are facilitating community engagement that we record and store as activity generated from the initial annotation. By using natural language processing (NLP) and domain ontologies, all the information is semantically processed. VIVO profiles are enriched with annotations, topics and generated activity. In this way, we are characterizing topics of expertise, interests and relations over social networks that enrich our VIVO profiles. By relying on VIVO and ORCID, and anchoring on publications, the semantic data we are generating is more authoritative; the enrichment of the profiles makes it easy to describe expertise in a granular way. VIVO profiles become aggregators of social-annotation-based activity creating a richer and more expressive discovery mechanism.

We believe open annotations offer a unique new capability that has the potential to radically transform the way we engage and discover scientific content and colleagues with similar interests across the web. Scientific annotations, we argue, are central in the empowerment of communities of practice by making content-based conversations possible. Unlike tweets, annotations are not limited in length or mechanisms of conversation. Similar to wall-based conversations in social networks, annotations are rooted in contextualized interests referencing and enriching existing content. One step ahead of tweets and wall-posts, annotations have a rich and well-structured provenance; by combining the benefits of annotations with VIVO profiles and ORCID we are enhancing the characterization of researchers and facilitating content-topic-based discovery; thus, supporting collaboration across social networks. In addition, annotations facilitate content-based exploration, generating a novel alternative metric. As the annotation is specific to a part of the text, it allows for granular analysis of the paper. Such a metric tells us not just the number of tweets or LIKES for a given document, it also allows us to identify the topics that are raising interest and how these are being discussed.

Import of Legacy Faculty Expertise Data into Symplectic Elements Application Software

Paul Grossman

Abstract: The Perelman School of Medicine at the University of Pennsylvania has selected Symplectic Elements to become its new system of record for providing pertinent information to the public about its faculty, along with curriculum vitae maintenance for the faculty members themselves.

Our legacy data has been housed in a proprietary application that has been developed and elaborated over time. The need to migrate this legacy data into Elements was an essential requirement, and our focus was to accomplish this taking into account that there were significant differences between our legacy data model and Symplectic's.

Although our approach was employed for transforming legacy faculty expertise data into metadata for Symplectic Elements, (a much larger scale effort as compared to the one we had undertaken for migrating subsets of that data into

VIVO), this approach can also be modified, and adapted for similar efforts where it is necessary to migrate data between two disparate data models, (such as a legacy data model with VIVO, or any other two applications in which the data models are not the same).

We will discuss the data migration process we used, including gap analysis, preparation of data mappings, data verification, and quality control.

Strategies for VIVO Adoption: Collaboration, Partnerships, and Support

Mary Walters, Timothy Morris, Daniel Grant and Michael Mitchell

Abstract: A successful adoption of VIVO at a university requires collaboration across many divisions. From different units such as Faculty Affairs to Research to Scholarly Communications, there are many varied interests to integrate into a successful adoption of the system and the area of primary advocacy may vary from one institution to another. This paper will explore the successes and challenges in bringing VIVO to the Emory Community and, in particular, identify the primary motivators from the respective units to promote engagement, lend support, and develop governance.

In particular, it is growing more important and critical for higher education IT to:

1. Understand stakeholders in a new way and look for synergies on projects across the organization.
2. Understand and reevaluate workflow closely from a user's perspective to determine how it can best be incorporated across projects
3. Understand how to position ourselves better for future needs, which in this case includes enhancements and elevation and discovery of research through VIVO
4. Understand appropriate support models to facilitate genuine adoption of a product or service

To adapt to the ever-changing environment of higher education, strong partnerships and collaboration are more critical than ever. Emory University has gone through an organizational restructuring forming Library and Information Technologies (LITS), where the library units and IT now exist within the same division. We will also explore the relevance of this restructuring and the value brought by the organizational realignment in supporting collaboration and building synergy across multiple projects related to VIVO. This restructuring has also created an enhanced platform for building partnerships on campus.

A successful partnership is based upon trust and an understanding of mutual needs and a shared goal. A system like VIVO provides a great case study for how essential it is to navigate organizational dynamics to create an enterprise-level solution.

Network Imputation in Predicting Researcher Collaboration

Yun Huang, Chuang Zhang, Maryam Fazel-Zarandi, Hugh Devlin, Alina Lungeanu, Stanley Wasserman and Noshir Contractor

Abstract: Linked Open Data supplies rich information on researchers such as their interests and expertise, previous collaboration, and interactions in different relation networks. This provides a big opportunity and demand for

discovering and facilitating potential collaboration in virtual research networks. Expert recommendation requires attributes of pairs of individuals (called dyadic attributes, Borgatti and Everett 1997) and different methods compared to product recommendation which is mostly based on attributes of individuals. This study proposes a statistic framework based p^* models to utilize the dyadic attributes in Linked Open Data and recommend potential collaborators and teams for academic researchers.

p^* models, the exponential random graph models used to study social network formation, are applied to network imputation in predicting collaboration links among researchers. We use p^* models to characterize the relational patterns among 147 researchers that competed in a 2009 grant competition funded by the Clinical and Translational Science Awards (CTSA) program of the National Institute of Health. Researchers' attributes such as gender, seniority and experience, dyadic bibliography information such as co-authorship and citation relations, and network structures such as degree distribution and network transitivity are used to predict team collaboration relationships, i.e. researchers working together to submit grant proposals.

We perform an experiment to examine the prediction power of the p^* method and compare its outcome against three prediction techniques traditionally used in the literature: node-wise similarity, the Katz method, and Relational Bayesian Networks. Our results show that network structures are more informative than node-wise similarity and covariate relations such as previous co-authorship. As an ensemble approach, p^* models outperform the benchmark methods by integrating all dyadic attributes including similarities, covariate relations, and network topology. Additionally, p^* models provide a unifying, multi-theoretical interpretation of the prediction models and the recommended results. As a result, the p^* method allows the customization of prediction heuristics towards desirable recommendations.

Our findings help to build a better foundation to develop effective recommendation tools.

[Extend Your VVIO Platform with Research Impact Metrics](#)

Andrea Michalek and Marianne Parkhill

Abstract: Add value to your VIVO implementation by integrating alternative research impact metric information. These metrics includes article downloads, presentation views, book holdings among many others. PlumX is an impact dashboard that provides information on how research output is being utilized, interacted with, and talked about around the world. PlumX integrates bi-directionally with VIVO. That is, PlumX starts with VIVO profiles and gathers alternative metrics about the researchers and their articles and other research artifacts such as posters and presentations. Then, PlumX creates robust information about the research with visualizations, analysis and metric data. Finally, via a PlumX widget, you can integrate the metric data and information right back into your own VIVO profiles. This bi-directional integration gives your VIVO implementation more power, makes it more useful to the researchers who utilize it, and allows you to discover even more knowledge about researchers across all boundaries. This presentation will describe how the VIVO and PlumX integration works, what kinds of impact information and visualizations you can expect, and a demonstration of both. There will be plenty of time for questions.

[Linked Open Data and interoperability of DOI metadata to improve scholarly communications](#)

Carol Anne Meyer

Abstract: 100 million Digital Object Identifiers (DOIs) have been assigned by 10 registration agencies to a variety types of content from films to genomic data. This presentation will focus on the challenges of providing resolution,

discoverability, metadata, and enhanced applications for these documents and data without driving scholars mad by requiring them to know the arcane rules of the DOI universe. Examples of interoperability will include collaborations between CrossRef, mEdRA, EZID, DataCite, ORCID, SHARE, and CHORUS to make the information that has already been collected by DOI Registration Agencies usable for institutions, libraries, scholars, funders, and publishers. Projects such as the JISC Open Citation Corpus, Citation formatting (transforming metadata into styled references for multiple style manuals), reporting published output for funding bodies, and supporting altmetrics are among the useful applications already built on linked data from DOIs.

Research Networking to Enhance Multi-institutional Expertise Discovery and Collaboration

Holly Falk-Krzesinski, Daniel Calto and Jeff Horon

Abstract: Research networking is facilitated through web-based expertise profiling systems—research networking tools (RN tools)—that aggregate research and scholarly information about faculty members and investigators to enable the rapid discovery and recommendation of experts to address new or existing research challenges and to facilitate new interdisciplinary collaborations. While most RN tool implementations focus on harvesting and displaying expertise from a single institution and across traditional academic organizational structures, this session will examine the flexibility of Elsevier’s Experts Portal to deliver innovative multi-institutional, cross-sector, and international semantic expertise portals in partnership with universities and other research institutions. These multi-institutional expertise portals are a key to stimulating networking and collaboration across typical research silos.

Northwestern Scholars at Northwestern University was the first to implement a university-wide research networking tool, profiling faculty researchers across all of its schools, not just those in the STEM fields. Moreover, they organized their researchers by research centers/institutes and graduate programs to foster expertise discovery across institutional boundaries and promote broad interdisciplinary and collaborative research.

Using the same Experts platform, the Chicago Collaboration for Women in STEM portal represents the first instance of an RN tool focused on making discoverable members of an underrepresented minority group. A joint initiative of Northwestern University and The University of Chicago, it also profiles women researchers at two regional Department of Energy national laboratories, Argonne and Fermi.

The Solar Fuels Institute (SOFI) is a global research consortium of universities, government labs, and industry united around the goal of developing and commercializing a liquid solar fuel. The SOFI portal represents one of the first uses of a research networking portal to connect researchers internationally across a single initiative and demonstrate a research institute’s collective expertise. Similarly, the Experts portal at MD Anderson Cancer Center includes profiles of cancer researchers from all of its sister institutions around the globe.

The Michigan Corporate Relations Network (MCRN) Expertise Portal is a statewide university research network specifically designed to connect corporations and SME’s to relevant university-based researcher expertise and core facilities to promote innovative cross-sector research and grow Michigan’s economy. The MCRN Experts Portal profiles researchers from five public research universities in the state and aims to build on findings that university-industry publications are generally of higher impact and university-industry patents are more likely to be successfully brought to market and drive revenues.

VIVO

connect • share • discover



Like all public Experts portals, these sites are connected via multiple federated searches (SciVal Community, Direct2Experts, and CTSAssearch), offering the ability to expand the search for experts and collaborations even more broadly. And the Northwestern Scholars and Women in STEM implementations broadcast data in their profiles openly in RDF triplestore, VIVO-compatible format, which can be accessed freely through any SPARQL endpoint. This provides analysts and developers access to the rich profile information, enabling sophisticated collaboration and networking studies.

Collectively, these multi-institutional research networking portals serve as models for other institutions to extend expertise discovery and networks beyond their institutional boundaries to promote collaborative and global research.

Integrating VIVO with ORCID - first steps

Jim Blake

Abstract: VIVO 1.7 includes the option of a basic integration with ORCID. VIVO site administrators may register their VIVO with ORCID, and offer users the opportunity to link their ORCID identity to their VIVO identity.

We encourage discussion of ways to expand the integration in future releases.