VIVO
connect • share • discover

Wednesday, August 17th

MORNING PRE-CONFERENCE WORKSHOPS (Separate registration required.)

8:30 AM – 12:00 PM
Introduction to VIVO: Planning, Policy, and Implementation
Presenters: Paul Albert, Brian Lowe, Andi Ogier, Michaeleen Trimarchi, Julia Trimmer, and Alex Viggio
Room: Colorado Ballroom G

New to VIVO as a new team member or part of a new implementation? What is VIVO all about? How did VIVO evolve and what benefits does it offer to researchers, to institutions, and to the global community?

This workshop provides an institutional perspective from others who have worked with VIVO. You'll meet six VIVO community members with years of experience with VIVO at different institutions. The presenters will talk about how VIVO is used in their organizations, where the data come from, how VIVO is managed, and how to feed downstream systems. You'll learn how to find the right resources as a new VIVO implementation — data sources, team members, governance models, and support structures. This workshop brings best practices and "lessons learned" from mature VIVO projects to new implementations. We'll help you craft your messages to different stakeholders, so you'll leave this workshop knowing how to talk about VIVO to everyone from your provost to faculty members to web developers.

8:30 AM – 12:00 PM
Getting Data from Your VIVO: An Introduction to SPARQL
Presenter: Michael Conlon
Room: Colorado Ballroom H

One of VIVO's greatest strengths is its ability to provide all its data for reuse. This workshop will introduce the attendees to SPARQL (SPARQL Protocol and RDF Query Language), the W3C standard for querying RDF data. VIVO comes with SPARQL ready for use. SPARQL will be introduced as well as the basic concepts needed to get data from VIVO, such as Resource Description Framework (RDF), URI (Uniform Resource Identifier) and the VIVO ontologies. Each will be explained in simple terms and reinforced by example.

Attendees will work in groups, and the groups will work through simple examples, to moderately complex examples for SPARQL queries for real-world examples of using VIVO data. Attendees will learn how to export data returned by SPARQL queries to spreadsheets for

subsequent data analysis, tabulation or visualization. This is an introductory workshop, no prior knowledge of SPARQL, RDF, or the VIVO ontologies is needed to participate. Following the workshop, attendees will be able to read VIVO ontology diagrams, and use these diagrams to write and run SPARQL queries on their VIVOs.

8:30 AM – 12:00 PM
Getting Started with VIVO 1.9: Introduction to the Maven Build Process
Presenter: Graham Triggs
Room: Colorado Ballroom I

Following the 1.8.1 release of VIVO, the build environment was migrated from the existing Ant scripts, to using the Maven project descriptors. This brings a number of benefits: it's a more immediately-familiar environment for many Java developers. Modern IDEs can read the project description and immediately set up the environment automatically, and we're able to better declare and manage the dependencies. However, in order to make the project look and feel familiar to a new user approaching the project from a Maven perspective, with an expectation of a standard project layout, the structure of the Vitro/VIVO projects needs to change slightly.

This half-day workshop will help people understand how the project structure has changed in VIVO 1.9, and to show them how to adapt their existing codebases when upgrading. It will also provide an introduction to the Maven project layout for new users, and show them how they can make effective use of Maven when creating a new VIVO implementation.

LUNCH (On Own)
12:00 PM – 1:30 PM

AFTERNOON PRE-CONFERENCE WORKSHOPS (Separate registration required.)

1:30 PM – 5:00 PM
Data Integration with Karma
Presenters: Pedro Szekely and Violeta Ilik
Room: Colorado Ballroom G

The VIVO platform has been designed to lower barriers for data interchange and re-use by standard data formats, ontologies, and identifiers consistent with Semantic Web best practices. The workshop will introduce the basic functionalities of Karma data integration tool and provide attendees with hands on training. Attendees will learn how to provide ontologies to Karma, how to load data, how to define URIs, how to transform data using Python scripts, how to map the data to the ontology, how to save, reuse and share mapping files, and how to produce RDF and JSON. No prior knowledge of semantic technologies will be assumed.

Implementing a researcher profile system involves aggregating data from a variety of sources. Data needs to be mapped, cleaned and maintained. Participants will utilize the presented tools to:

- Model data in a variety of formats with the help of established ontologies (FOAF, FABIO, CITO, BIBO, SKOS, VIVO-ISF)
- Understand the use of Web Ontology Language (OWL)
- Create RDF data for use in ontology driven applications

We will begin with lectures and continue with hands-on demonstration and experimentation. The lectures are designed to help participants gain experience and knowledge of researcher profiling systems, importance of ontologies, a language for expressing customized mappings from relational databases to RDF datasets (R2RML), the advantage Karma data integration tool offers in transforming data into semantic web compliant data.

The workshop will help participants in planning for an organization's efforts to move existing data into ontology driven applications, like VIVO, to uniquely represent the scholarly outputs of researchers in their institutions and beyond.

1:30 PM – 5:00 PM
Linked Data Fragments: Hands on Publishing and Querying
Presenter: Ruben Verborgh
Room: Colorado Ballroom H

Linked Data on the Web—how can we use it, and how can we publish it?
This workshop explores different interfaces to Linked Data using the Linked Data Fragments conceptual framework. There are two main aims of this session: learning to consume existing Linked Data from the Web, and publishing your own dataset using a low-cost interface. Additionally, we will build small applications in the browser that make use of Linked Data.

This session is aimed at participants with a technical background, as we will get into the details of Linked Data publication and querying. However, people with a broader interest are also welcome, as participants can work together in groups. If you want to learn what roles Linked Data can play in your organization on a very practical level, this workshop is definitely for you.

If there is interest, this workshop can be extended with a hackathon later, in which people can build prototype applications on top of live Linked Data on the Web.

1:30 PM – 5:00 PM
How to Make Your Research Networking System (RNS) Invaluable to your Institution
Presenters: Brian Turner, Anirvan Chattejee, and Eric Meeks

VIVO 2016 Conference Schedule (Tentative)

Room: Colorado Ballroom I

This workshop is designed to help institutions build, leverage, and deploy the information within their RNS across the institution. The goal is to increase awareness of, engagement with and dependence on your RNS to solidify the RNS' roles in supporting researchers. Note that the takeaways from this workshop can be applied to your RNS regardless of the underlying product, and will work for a VIVO, Profiles, "home grown," or commercial RNS installation.

This workshop will provide a mix of lecture, discussion, exercises, and templates to enable participants to replicate the successful engagement at UCSF — and avoid our biggest mistakes. Use of this approach has garnered 1.3 million visits a year, 2,800 customized profiles, and 38 outbound data reuse integrations for UCSF Profiles.

This workshop will discuss ways to make your RNS indispensable in each phase of the implementation:

Setting the Stage
- Auto-added data: add as much data as possible to your RNS - publications, grants, photos, news stories, etc.
- Google Analytics - the crucial substrate
- Pre-packaging of publications, people and connections
- APIs to make the data available
- Targeting and messaging decision makers appropriately

Dress Rehearsal
- User support - answer the emails!
- Senior leadership support
- Deploy APIs to power other websites
- Search engine optimization

Opening Night & The Season
- Engagement email campaigns
- Bootcamps and department meetings
- Finding and catering to the "power users"
- Iterative process of growing user base/traffic and providing more data/features

After Party
- Wrapping it all up to show the value of your RNS in an executive level report for the higher administration

WELCOME RECEPTION
5:30 PM – 6:30 PM

OPEN EVENING – ENJOY DENVER!

Thursday, August 18th

REGISTRATION

7:30 AM – 7:00 PM
Room: Colorado Ballroom Foyer

NETWORKING BREAKFAST

7:30 AM – 8:30 AM
Room: Colorado Ballroom Foyer

OPENING WELCOME & KEYNOTE PRESENTATION

8:30 AM – 10:00 AM
Room: Colorado Ballroom E-F

Towards Digitizing Scholarly Communication
Presenter: Sören Auer

Room: Colorado Ballroom E – F
Abstract: Despite the availability of ubiquitous connectivity and information technology, scholarly communication has not changed much in the last hundred years: research findings are still encoded in and decoded from linear, static articles and the possibilities of digitization are rarely used. In this talk, we will discuss strategies for digitizing scholarly communication. This comprises in particular: the use of machine-readable, dynamic content; the description and interlinking of research artifacts using Linked Data; the crowd-sourcing of multilingual educational and learning content. We discuss the relation of these developments to research information systems and how they could become part of an open ecosystem for scholarly communication.

BREAK

10:00 AM – 10:30 AM

PANELS AND PRESENTATIONS

10:30 AM – 12:00 PM

Room: Colorado Ballroom G

An Early Look at Share Link: VIVO, Share, Hydra, and Fedora Together for Research Tracking and Discovery

Presenter: Rick Johnson
Track: Linked Data

As an extension of the Linked Data for Libraries project (LD4L) and parallel work to integrate VIVO and SHARE, the Hesburgh Libraries at the University of Notre Dame is working to link Fedora 4, Hydra, VIVO, and SHARE together to further realize the paradigm proposed within LD4L with a focus on research lifecycle events. This project dubbed SHARE Link is focused towards two primary aims desired by many institutions: institution level research activity tracking, curation, and sharing; enhanced search, browse and discovery of research events and related materials.

Subsequently, bringing SHARE, VIVO, and Fedora 4 together makes it possible to create a dashboard of research events including links to repository materials. In turn, the metadata in each system are enriched in ways not possible with any one solution. The underlying unified research and repository information graph (between Fedora and VIVO) will also allow browsing the virtual shelves of repository materials across venue, subject, and other person affinities like specialization, institution, or co-authors. It will allow creating a recommendation list when viewing a related work. While wider adoption within the Fedora community will be sought, this effort is initially focused on developing a reference implementation of VIVO paired with an Institutional Repository utilizing the Hydra Framework on top of Fedora 4.

This presentation will describe progress to date in the early life of this project. It will explore the relevant use cases and reasons why the pairing of these solutions are stronger together than apart. It will also cover possible collaborations, the project timeline, and current and future goals such as:
- Enhance serendipitous discovery of research materials.
- Collect and display research activity by researchers at a given institution or within an academic department.

- Create a widely accessible module for an institution to harness the SHARE data feed of broadly aggregated research events.
- Create a compelling integration between SHARE, VIVO, and Fedora 4 that harnesses the strengths of each platform.

Collaboration Through Linking: A Bilateral Project Between University of Florida and Elsevier for Access to Articles by UF Authors

Presenters: Sara Gonzalez, Valarie Minson and Judith Russell
Track: Product Information

VIVO implementations have long been closely tied to the library and yet local VIVO systems have prioritized the integration of institutional repository data far below other internal or external data sources, thereby bypassing the local infrastructure designed to preserve institutional research for a university. A new partnership between the University of Florida (UF) George A. Smathers Libraries and Elsevier makes the UF Institutional Repository (IR@UF) an excellent internal data source for UF author publications.

As background, in 2014 and 2015 the Smathers Libraries and Elsevier embarked on a bilateral pilot project that identified and collected metadata for over 30,000 articles written by UF staff, faculty and students and published by Elsevier from 1949 forward. The Smathers Libraries are placing the metadata and full-text into the IR@UF for indexing with the full text document available through the Elsevier ScienceDirect platform for supported aggregated metrics and article tracking. The partnership between Elsevier and the Smathers Libraries is a natural choice as Elsevier has a large volume of content and a high citation impact and UF authors publish 1,100-1,300 articles per year in Elsevier journals.

This bilateral partnership is now being expanded to include other publishers through CHORUS -- a non-profit cooperative effort involving publishers and funding agencies to expand access to scholarly publications reporting on funded research— designed to provide high-quality citation information as a tool to support reporting compliance through an alerting and reporting dashboard integrated with the IR@UF. CHORUS is developing a compliance dashboard for integration with the IR@UF and other academic institutional repository systems.

The overall goals of this project are for the Libraries to provide an infrastructure to assist with researcher compliance, provide CHORUS and the participating publishers a better understanding of the information needed by institutional repositories for tracking and reporting on compliance, and further increase discoverability of scholarship through potential linkages with VIVO@UF. This presentation will give an overview of how embargo periods will be handed

and accessibility addressed when a user does not have access to subscribed content, and will also outline the benefits of the partnership which includes reduced faculty burden, facilitated compliance reporting, increased visibility and accessibility, and additional institutional linkages supporting contextualization of research.

How to Read 100 Million Publications: VIVO and Comprehensive Open Publication Databases

Presenters: John Mark Ockerbloom
Track: Strategy

A number of commercially developed publication databases such as Web of Science and Google Scholar aim to provide a comprehensive view of the scholarly literature. Such databases are quite large in scale, needing to handle metadata on the order of about 100 million publications, and to grow by more than 1 million new publication records every year. There is ongoing interest in creating more open comprehensive databases in the community as well, for various purposes, ranging from open access support to preservation to various kinds of researcher analysis. Is it worth creating and supporting such open databases in the community, and if so, for what purposes? How could they scale up as they would need to, in technical, political, and participatory terms? How could they use cooperation to thrive, rather than being starved as unwanted competition? What can we learn from experiences with existing and proposed comprehensive and specialized publication knowledge bases, whether run by commercial firms, academic institutions, interested amateurs, or bootleggers? How could VIVO sites help build up comprehensive publication databases, and how would the availability of such databases affect the sort of data and services that local VIVO sites would focus on? The goal of this session is not to propose Yet Another Big Database for the VIVO community to develop, but rather to provoke discussions on how efforts in the VIVO community can best support and take advantage of a growing ecosystem of open publication data, at the global scale.

Room: Colorado Ballroom H

Using VIVO and Other Tools to Achieve a Strategic Vision

Presenters: Bruce Herbert, Laurel Haak, Andrea Michalek and Marianne Parkhill
Track: Strategy

A globally interconnected research world is the norm. This offers great opportunities for communication and collaboration across fields of research, but it also comes with great

challenges, in particular for systems interoperability. Many technology products, systems and tools have been created to help universities and other research organizations meet these challenges while achieving their strategic goals. This is creating a new challenge: a confusing stack of technology for users to assemble into a functional research ecosystem.

This panel will explore how one university, Texas A&M, selected and implemented a technology stack and is using it to achieve their strategic plan, "Vision 2020: Creating a Culture of Excellence." Their stack includes:

- VIVO
- ORCID
- Repositories, Vireo & DSpace
- PlumX

The panel will discuss the components and the interactions between them, to provide the audience concrete examples of a working, interconnected research ecosystem.

Texas A&M will describe its process for defining its technology needs, who was engaged, and the questions that drove decision-making. They will describe their current technology stack, and the connections between the components. One cross-platform connector used by Texas A&M is the ORCID ID, which is allowing them to integrate information about researcher works and affiliations across various systems and present this information in their VIVO instance.

ORCID will describe how ORCID identifiers are being used at universities to assert researcher affiliation, and how universities may benefit from integrations by publishers and funders. Consortia approaches to implementation and adoption will be discussed, as well as the use of ORCID in federated identity management systems such as eduGAIN, and how ORCID interacts with identifiers for works and organizations.

Green open-access institutional repositories are a key component of a research ecosystem. It is the container for the research output conducted at the university. Recent studies have shown that opening up access to research creates more usage and citations for that research. Texas A&M implemented Vireo for theses and dissertations and DSpace for faculty research.

How is all of this enabling Texas A&M to measure progress toward its strategic goals? As a component of its tech stack, Texas A&M has integrated PlumX to track research metrics in five categories: usage, captures, mentions, social media and citations. Sometimes this is referred to as altmetrics. This part of the panel will include a discussion of what Texas A&M has discovered by tracking metrics across disciplines including output for the humanities. It will also showcase

some other use cases of metrics in institutional repositories, analytics reports and researcher dashboards.

Streamlining Assessment Exercises and Faculty Activity Reviews

Presenters: Jonathan Breeze, Jeff Dougherty and Michael Metcalf

Track: RMS Workflows

Ask a researcher to populate an institutional webpage and you'll see them roll their eyes and move on to something more enjoyable.

Ask a researcher to complete their faculty activity report and they'll grudgingly do it, but complain about the duplicated effort.

However: ask a researcher to keep a system up-to-date throughout the year, so that they can avoid the annual form-filling common with other systems, and they might not only do it, but tell their colleagues about how easy it was!

One of the core concepts behind Symplectic's mission is the reduction of the administrative burden from research institutions and their staff. The next logical step in this calling is the tackling of faculty activity and performance-related assessments: parts of a much larger data-gathering exercise that institutions partake in every year and one that remains laborious to all required to contribute.

The latest major release of Elements, Symplectic's leading Research Information Management System, contained a highly-anticipated solution to this issue - the Assessment Module. Assessment-related data can now be reused, repurposed, and resurfaced on a VIVO profile, with researchers having the freedom to pick what's pertinent for public use.

This presentation will explore how this module offers a significant new repurpose of institutional data. Complementing our long-standing relationship with VIVO, Symplectic is continually committed to semantically linking previously siloed data; knowing that this data is only useful when it can be reused throughout an institution freeing faculty from burdensome reentry.

Examining Federal Expert Networking and the Economics of Scale: Moving the "HHS Profiles" Pilot Towards "Experts.gov"

Presenters: James King, Jessica Hernandez Berrellez, Nichole Rosamilia, Ben Hope, Mashana Davis, and Bridget Burns

Track: VIVO Implementations

The federal government tackles America's hardest problems; yielding decisions with important consequences for our nation's future. Thomson Reuters recently named the U.S. Department of Health and Human Services (HHS) as the most innovative publically-funded research organization in the U.S., and fourth in the world. The challenges we face are worthy of mobilizing our nation's best minds across multiple sectors. And yet we lack a systematic way to locate and match expertise for participatory problem-solving.

The question of how best to leverage collective intelligence for better governance remains unanswered. In 2015, expert networking implementers from the U.S. Food and Drug Administration (FDA) and National Institutes of Health (NIH) forged a collaboration to launch the HHS Profiles pilot. Funded by 18F, General Services Administration (GSA), the goal was to test the viability of a public-facing, multiorganizational expertise matching platform called "HHS Profiles."

The team presented the project plan at the VIVO 2015 Conference, and returns to report preliminary results and lessons learned from this effort to design a platform-agnostic, integrated framework to harness intra-government expertise. We will then discuss the scope and requirements of the next project phase, and the vision for "HHS Profiles" scales to "Experts.gov." The presentation will close with a reflection on future pathways for cross-sector collaboration, and address the unique challenges to expert discovery in the federal space.

Room: Colorado Ballroom I

Update on Reciter: An Open Source Author Name Disambiguation System that Leverages Institutionally – Maintained Metadata

Presenters: Michael Bales, Jie Lin, Steve Johnson, and Paul Albert

Track: VIVO Tools

Administration at Weill Cornell Medicine has called upon the Library to track authorship for thousands of individuals including faculty, postdocs, and alumni. Until identifiers like ORCID are

adopted more widely and used more consistently, author disambiguation remains an important challenge towards producing valid and real-time reports of publication activity.

At the 2015 VIVO Conference, we presented ReCiter, a Java-based tool which uses institutionally-maintained data to perform author name disambiguation in PubMed. Since last year's conference we have continued to improve the performance of ReCiter. In our set of 63 randomly selected existing faculty members, ReCiter can now assert author identity at over 97% accuracy. In a randomly selected sample of 20 alumni, ReCiter performed at 93% accuracy, and, for active students this figure was 80%.

ReCiter employs 15 separate strategies for disambiguation. The range of data used spans six systems of records: Office of Faculty Affairs, Human Resources, alumni database, student database, physician profile system, and the grants management system. For our test set of data, we have established which data are most powerful in contributing to overall accuracy. These include: known co-investigators, known department, and year of doctoral degree.

Our team has designed ReCiter to be generalizable. Our goal is to determine if the high accuracy we note can be achieved at other sites. The code will be shared with the community.

The VIVO Pump: A New Tool for Data Ingest

Presenters: Mike Conlon, Christopher Barnes and Kevin Hanson

Track: VIVO Tools

The VIVO Pump is a new tool for managing data in VIVO. The Pump allows data to be managed in spreadsheets -- simple rows and columns corresponding to attributes of entities in VIVO. Using the Pump, a data manager can "get" data from VIVO into a spreadsheet, modify the values, and/or add rows, and then "update" VIVO using the improved spreadsheet. Enterprise data can be loaded into VIVO using and update from a spreadsheet.

The Pump uses definition files in JSON format to define the mapping between spreadsheet rows and columns, and VIVO's graph data models. Definition files are included with the Pump for managing people, positions, educational background, organizations, publications, grants, teaching, service, mentoring, dates, journals and concepts. No knowledge of the VIVO ontology is required to use delivered definition files. Definition files are used to "round trip" data -- the same definition is used to perform the "get" as to perform the "update."

The Pump can be used to load data to new VIVOs as well as for managing data of established VIVOs. Additional features of the Pump include "enumerations" -- simple translation tables that allow the Pump to convert between terminologies and identifiers, and "filters" which can be assigned to columns to perform routine standardization of values. The Pump is completely domain agnostic. Definition files can be created for any set of ontologies to manage any graph data as spreadsheets. In this talk, we will describe the design goals of the Pump and provide examples of its use to manage data in VIVO. The Pump is fully documented in on-line resources, as well as PDF and eBook formats. The Pump is open source software, made freely available under the Apache License.

Publication Data Enrichment at the University of Florida with the Web of Science™

Presenters: Christopher P. Barnes, Ted Lawless, Nicholas Rejack, and Kevin S. Hanson
Track: Product Information

This talk will focus on a publications data enrichment project at the University of Florida. We will describe how the team has utilized web services from the Web of Science™ to automate processes and to add unique identifiers to existing publications. We will also explain modifications to VIVO's core queries and templates that allow for adding important contextual links to VIVO, which enhance browsing and discovery. Data processing code, methods and lessons learned will be shared with attendees.

LUNCH (ON OWN)/ BIRDS OF A FEATHER DISCUSSIONS
12:00 PM – 1:30 PM

INVITED SPEAKERS

1:30 PM – 2:15PM
Room: Colorado Ballroom E – F
The Future is Federated
Presenter: Dr. Ruben Verborgh

For years, scholars, scientists, and practitioners in various fields have been told about the importance of open data. Increasingly, many of us have been asked to publish Linked Data as well. While the publication of Linked Data becomes easier—not the least because of VIVO—consumption often remains difficult. As such, we might start to wonder: for whom are we

creating Linked Data? In this talk, I will argue the importance of decentralization and federation on the Web. I'll show with practical examples how we can consume Linked Data and make it work for us, today.

2:15PM – 3:00 PM
Room: Colorado Ballroom E – F
Building Semantic Web Applications Without SPARQL
Presenter: Dr. Pedro Szekely

The Semantic Web offers an approach for integrating heterogeneous data and publishing it using widely-used terminologies to facilitate reuse.Even though SPARQL has been a standard for publishing integrated datasets to application developers since 2010, adoption has been low. Two critical barriers to adoption are the fragility of SPARQL endpoints to heavy loads, and the complexity of the RDF/SPARQL tool-chain. In this talk we present an alternative approach for publishing RDF data using JSON-LD and ElasticSearch, a NoSQL data store. Our experience shows that this approach scales to datasets of billions of triples using commodity hardware, supports sub-second query response times for queries commonly used in Web portals, even under heavy loads, and uses a simpler JSON-based tool chain familiar to most software developers.

BREAK
3:00 PM – 3:30 PM

PANELS AND PRESENTATIONS

3:30 PM – 5:00 PM

Room: Colorado Ballroom G
Optimizing VIVO: Under the Hood of 1.8.1
Presenter: Graham Triggs
Track: Technical VIVO
Whilst the 1.8 release of VIVO had brought a number of usability improvements, there had been a slight catch: general page rendering was slower. It was important to address this, and so optimization was the key priority of a minor point release.

The release of 1.8.1 managed to not just reverse the performance loss, but to gain around 30% improvements on the 1.7 release - without changing any of the data model, or triple store implementation.

In this talk, I will show how using a profiler identified areas of concern, the strategies that were applied to improve the performance - including some significant changes to the models cached in the application layer - and some tips for thinking about performance when customizing or developing for VIVO.

Maintenance Queries, Automating the Identification and Resolution of Problematic Data in VIVO

Presenters: Steven McCauley and Ted Lawless
Track: Technical VIVO

After deploying VIVO, adopters will find that, overtime, data can become inconsistent, incomplete, or missing. This talk will focus on methods for identifying and resolving problematic data in an automated fashion. We will describe a reusable toolkit we have developed that utilizes SPARQL based rules for identifying data and VIVO's SPARQL API for updating, or correcting, these problems. Further, we will explain how web services are used to augment incomplete data. The toolkit and methods used will be extendable and reusable by other sites.

Deploying VIVO as A Static Site

Presenters: Ted Lawless, Alexandre Rademaker and Fabricio Chalub
Track: Technical VIVO

VIVO's rich data model can be ideal for publishing information about research activity and making connections among these activities. However the technical architecture can be a challenge to deploy as a public facing website for potential adopters. This talk will describe a process for building a research discovery website from VIVO's data model but deployed as a static site, using just HTML5, CSS, and JavaScript. This talk will identify the data mapping process, publishing steps, lessons learned, and identify future directions of pursuit.

Room: Colorado Ballroom H

The SEO State of the Union 2016: Five Data-Driven Steps to Make Your Research Networking Site Discoverable by Real Users, Based on Google Results

Presenter: Anirvan Chatterjee

VIVO 2016 Conference Schedule (Tentative)

Track: Outreach

The second annual SEO State of the Union dives back into the world of search engine optimization for research networking sites. Search engines like Google are, in many cases, the most critical pathway for discoverability for research networking platforms. UCSF Profiles, for example, receives over 80% of its visits (about 100k of 125k visits per month) via Google. But not all research networking platforms perform equally well on search engines.

The first annual SEO State of the Union 2015 broke down real-world SEO performance for over fifty public research networking sites. We will present updated findings for 2016, including a detailed overview of how different implementations and platforms rate in terms of search engine discoverability, stratifying by platform and type of deployment.

We will present five key findings into the real-world factors that impact search engine traffic, both at the level of the entire site, and also between different profile pages on the same site. These findings will be helpful at every stage of the research networking process, from product evaluation/selection to the promotion of mature platforms.

Towards Deeper RNS User Engagement: How to Get Users to Sign in for Personalized Content
Presenter: Brian Turner
Track: Outreach

UCSF proposes to present an update to last year's presentation on Leveraging Personalized Google Analytics Information for Greater RNS Engagement. We have launched a personalized dashboard and promoted it in various ways. We would like to share our learnings with the RNS community including progress on getting profile page owners to sign in for personalized content and what we can learn from their signed-in behavior. We'll also discuss ideas for next steps.

Pimp my VIVO: Creating an "Expertise Hub" with Focused, Responsive Design
Presenters: Simon Porter, Sabih Ali, and Michael Metcalf
Track: Outreach

Over the past few years, VIVO has gained global appeal amongst institutions. However, many institutions struggle to get past the early adopter phase and 'go live' with their own VIVO. As Registered Service Providers of VIVO, we are often asked for 'showcase' VIVO systems to

display. We will typically send them examples where institutions have undertaken large customisation projects on top their VIVO, like Griffith University.

Institutions who adopt VIVO still struggle to find a focus for the 'stock' version. For it to be an effective showcase of an institution's data, there needs to be consideration of usability and design to realise its full potential, to make it a true centre of the expertise of an institution. Many institutions do not always have the resource or expertise to achieve this.

Room: Colorado Ballroom I

Linked Data for Libraries (LD4L): Expanding the Linked Data Ecosystem
Presenter: Dean Krafft
Track: Linked Data

The Linked Data for Libraries (LD4L) team, consisting of librarians, ontologists, metadata experts, and developers from Cornell, Harvard, and Stanford libraries with support from the Andrew W. Mellon Foundation, has recently completed its first two years of work on adapting and developing LOD standards for describing and sharing information about scholarly information resources. In this presentation, we will describe how to access and use the LOD created by the project, representing some 29 million scholarly information resources cataloged by the three partner institutions. We will also describe the demonstration Blacklight search operating over the combined dataset.

We will then describe the follow-on work currently underway in two closely related efforts. LD4L Labs is a partnership of Cornell, Harvard, Iowa, and Stanford focused on creating tools to support original cataloging of scholarly information resources using linked data, as well as tools to support using linked data to organize, annotate, visualize, browse, and discover these resources. The LD4P (Linked Data for metadata Production) project is a partnership of Stanford, Columbia, Cornell, Princeton, Harvard, and the Library of Congress to do original and copy cataloging of a wide range of collections and materials, including unique collections of Hip-Hop LPs, performed music, cartographic materials, audiovisual and sound recordings, two and three-dimensional art objects, and the personal library of a famous author and scholar.

Finally, we will draw on the use cases and examples embodied in this work to discuss some of the opportunities to engage directly with VIVO profiles, the VIVO community, and the broader researcher profiling ecosystem. Several of the efforts within LD4L Labs and LD4P will be looking at using VIVO profiles as local authorities during the process of cataloging scholarly resources.

There are also potential opportunities for VIVO instances to take advantage of some of the ontology refinements that LD4L uses in describing scholarly works. We will explore some of the implications that these developments might have for how VIVO is used and evolves at academic institutions.

Service Driven Profile Linkage – The CTSAsearch/CrossLinks Experience

Presenters: David Eichmann and Eric Meeks
Track: VIVO Tools

While multi-site search of research profiling systems has substantially evolved in recent years, the deployed instances of profiling systems substantially remain disconnected islands of Linked Open Data (LOD). CTSAsearch harvests VIVO-compliant LOD and provides a number of data integration services in addition to search and visualization services. Seventy-eight institutions are currently included, spanning nine distinct platforms. In aggregate, CTSAsearch has data on 168-415 thousand unique researchers and their 7 million publications. The public interface is available at http://research.icts.uiowa.edu/polyglot. This paper presents our experiences connecting CTSAsearch data integration services into the UCSF CrossLinks service, providing a means of directly interlinking researcher profiles across sites and platforms.

Data Integration Services
CTSAsearch integration services have been designed to support human exploration of the identity of profiled persons and automated linkage of LOD across systems. Services hence fall roughly along a human-system interface spectrum:

Person by name query: This service accepts a last name and a first name prefix and returns two discrete lists, one of "real" persons (i.e., profiles deemed as not stubs) and one of "stubs" (i.e., profiles appearing to only be placeholders in some profiling system). The returned URIs can then be used as input for sameas queries below.

Person by publication query: This service accepts a DOI or PMID and returns a list of profiled authors and their rank in the author list of the target publication. Output formats include HTML, JSON and XML. This service is specifically designed to support inter-site linkage of coauthors.

Person sameas query: This service accepts a person URI and returns a list of URIs referring to the same person. Currently, two URIs are asserted to refer to the same individual if they share one or more publications with the same PMID or DOI, have the same family name and either the same first name or one first name is a single initial that matches the first name of the other. (All name comparisons are case insensitive.)

Cross-linking Profiles
Almost all research profiling sites currently provide only internal links. In the case of extra-institutional co-authors, either no information is provided or stub profiles are generated containing only an author name generated from the citation. UCSF Profiles now harvests CTSASearch data to provide links to the non-institutional co-authors, and is working with Harvard to make this feature a part of the common Profiles RNS software. Benefits arising from this feature include a richer web user experience, the ability to "crowd source" disambiguated data (an author at UCSF noticed a link to an incorrect author with a similar name at another institution, and has notified the administrators at the other institution), as well as improved SEO due to links from many top level domains.

Conclusion
CTSAsearch and CrossLinks demonstrate that substantial value can be added to the current research networking landscape through integration of these data. Our future work in this area will include enhanced ability to interconnect these systems and to visualize the resulting aggregated information space.

Connect UNAVCO, a VIVO for a Scientific Community
Presenters: M. Benjamin Gross, Linda R. Rowan, Matthew Mayernik, Michael D. Daniels and Dean B. Krafft
Track: Technical VIVO

Science, data, and research products are the focus of two new VIVO implementations, Connect UNAVCO (connect.unavco.org) and EOL Arctic Data Connects (vivo.eol.ucar.edu). The sites are being developed as part of EarthCollab, a project funded by The National Science Foundation. EarthCollab is a collaboration between UNAVCO, the University Corporation for Atmospheric Research (UCAR)/National Center for Atmospheric Research (NCAR), and Cornell University. UNAVCO is a non-profit university-governed consortium that facilitates geoscience research and education using geodesy. NCAR's Earth Observing Laboratory (EOL) manages observational data related to a broad range of research fields in the geosciences. Both organizations support a large community of researchers and collaborators, many of them at U.S. and international universities. VIVO, with the science-focused application and ontology extensions we are developing, connects the research output of our diverse communities with datasets, grants, and instruments managed by our organizations. Unique identifiers for people (i.e., ORCID), publications (i.e. publication DOIs), data (i.e. data DOIs) are included whenever possible to avoid confusion and to yield the most stable and effective connections.

To date, Connect UNAVCO includes active pages for 650+ community members. The members are connected to 4000+ academic articles, 3,800+ datasets, and 3,600+ GPS stations managed by UNAVCO. The VIVO-ISF ontology captures many of the connections in the data model. Local ontology extensions capture organization membership roles and their representatives, as well as roles related to stations and instruments. We utilize parts of the WGS84 and Global Change Information System (GCIS) ontologies to capture geospatial and scientific instrument concepts, respectively. UNAVCO's implementation also includes a handful of application extensions built to highlight data, instruments, and the people who manage them. For example, individual pages for stations have been extended to plot the station on a map, include a link to the data archive, and the station's principal investigators. Another extension allows staff and community members to select their expertise and research areas from a curated list of terms. In addition to appearing on a person's profile page, a community-wide word cloud is generated and displayed on the home page.

POSTER RECEPTION

5:00 PM – 7:00 PM

Room: Colorado Ballroom A – D

Join us for our annual networking reception while you view and discuss the many Poster Presentations on display.

Graph-Based Modeling of the Bibliographic Data and Linked Open Data for the Citation Records

Authors: Muhammad Javed, Joseph McEnerney, Katherine Chiang, Jon Corson-Rikert and Sandy Payette

VIVO is an open source web application used by researchers across the globe to model, record and publish their scholarly work in a linked data environment. In this poster, we present our ongoing work on VIVO-ISF ontology-driven modeling of scholarly data originally recorded in Sympletic Elements. Sympletic Elements (SE) is a web-based framework that is used by research organizations to harvest, analyze and report their research outputs over time. Research outputs may include bibliographic items such as journal articles, books, patents, presentations, chapters, reports, thesis, and other publication types. In Elements, bibliographic items can either be manually added by researchers or harvested from internal or external data sources.

We performed an empirical case study to analyze and investigate the bibliographic items harvested by Elements. For different bibliographic item types, Elements captures the title of the

article, a list of the involved persons and their contribution roles (such as authors, editors, translators etc.), the name of the journal, volume number, issue, ISSN number, DOI, publication status, pagination, external identifiers named as "Citation Items." The data for these citation items may or may not be available in every internal/external data source. This leads to recording multiple different versions of the same journal article, named as "Version Entries." For the case study, we concentrated on the "journal article" category of the bibliographic items only.

We formalize the version entries of a journal article using a graph-based approach. We used attributed graphs. Graphs with node and edge attribution are typed over an attribute type graph (ATG). Attributed graphs (AG) ensure that all edges and nodes of a graph are typed over the ATG and each node is either a source or target, connected by an edge.

In this poster, we present our graph-based model, outcomes of the analysis of the version entries graph and the process of the merge of the different version entries graph (of a bibliographic record) into a single record, known an Uber Record. The core objective of our research work is to represent all the data given in SE bibliographic records and make them available (to be searched, used or extended etc.) in the form of Linked Open Data (LOD) using the VIVO software. To do so, we built a VIVO-ISF ontology-driven model for the journal article category and extended the VIVO authorship (relationship) model. Our plan in ongoing work is to continue the same process for other types of the bibliographic items.

New Methods for Capturing and Surfacing Research Impact Narratives

Authors: Jonathan Breeze, Julia Hawks and Michael Metcalf

For many disciplines, the citation frequency of scholarly articles, alongside measures of the outlet in which they are published, remains a common, albeit controversial, means of evaluating the academic impact of research. However, evidence of the wider effects of impact, e.g. societal, economic, or environmental, can't always be captured via traditional metrics. In recent years, a growing number of funding agencies (such as the NIH and NSF) have started to ask institutions and researchers to provide qualitative evidence of the impact generated as a result of externally funded research, and this trend is expected to continue.

Previously, there hasn't been a standardized method for recording this qualitative data. At the end of 2015, Symplectic released an update to their leading Research Information Management System, Elements, which aims to tackle this. The Impact Module offers a system to help capture a narrative of emerging evidence of research impact.

This presentation will explore how an impact narrative can be both captured and optionally displayed in a public-facing profile tool. Capturing and reusing impact data means researchers no longer have to rewrite the proof of the impact of their work during grant applications.

VIVO and Profiles RNS, Sharing Applications as Well as Data Through Open Research Networking Gadgets (ORNG)

Authors: Keith Brophy, Eric Meeks and Brian Turner

The University of Wollongong is using ORNG/OpenSocial applications to add additional content to their VIVO installation, which is planned to go public in May. This represents the first production VIVO instance to adopt ORNG, and demonstrates the value of the research networking community to not only share data across platforms due to the adoption of a common VIVO Linked Open Data standard, but to additionally share applications across platforms that conform to the ORNG/OpenSocial standard that UCSF introduced into the research networking community.

Within their VIVO implementation, the University of Wollongong is enhancing the standard profile by enabling the addition of YouTube and Twitter content feeds along with Slideshare profile links by adopting ORNG gadgets built by UCSF on Harvard Catalyst Profiles.

Additionally, the University of Wollongong has developed their own ORNG gadget to display LinkedIn content, which is now being adopted by UCSF for local use and is available for other ORNG-compatible research networking sites such as Harvard, Baylor and Boston University. ORNG applications are well suited for bringing in content and "grey literature" that is hosted at external web sites. The grey literature is typically faster to show a researcher's latest work than a publication or citations, and also often more consumer-friendly than formal research outputs such as peer reviewed publications. As such, grey literature can provide valuable content for research networking systems.

The researchers at our institutions recognize the value in external collaborations. This is reflected in the publications and grants that we represent in our VIVO-based systems, and is part of why it is important for us to adopt a common data representation with our research networking tools. The research networking community also recognizes the need to collaborate at a larger scale, and the ability to share applications across our systems is a clear way to strengthening the collective value of our research networking systems, and to easily expand the content our systems provide.

From Postdoc to Provost: Seven Simple Ways UCSF Profiles has Been Used to Win Funding, Find Collaborators, and Get the Job Done

Authors: Leslie Yuan, Eric Meeks, Brian Turner and Anirvan Chatterjee

After many years of promotion, using a myriad of engagement strategies, UCSF will share real-life stories of how UCSF Profiles has made life easier for many people across our enterprise. We will describe how the system has been used by junior researchers to our senior-most faculty members, including the Executive Vice Chancellor & Provost and the Dean of the UCSF School

of Medicine. Back in the day, our goal was to create a system that was "relied upon by the campus" and we will share examples that now justify this status.

An Initial Strategy of Texas A&M University's VIVO Project

Authors: Dong Joon Lee, Bruce Herbert and Michael Bolton

Texas A&M University (TAMU) has recently identified six interdisciplinary and societally-relevant research themes as strengths of the institution (today.tamu.edu/category/topics/grand-challenges) and is dedicated to help increase research collaboration within the themes and the institution. As one part, TAMU Libraries are interested in adapting and extending the VIVO ontology in order to create a dashboard that presents research related to each of these themes. This would allow the libraries to highlight the researchers and research products (i.e., papers and grants) associated with each of the themes. This would also help support and demonstrate TAMU's impact on society. In addition, implementation of the dashboard would motivate researchers to engage with the VIVO and repository systems and contribute to the systems development in digital library settings.

VIVO community encourages adapting and extending the VIVO ontology for sustaining institutions' local needs. The extension that maintains interoperability with core VIVO ontology structure facilitates customized data visualization, analysis, and search for the local institutions. Different extensions and applications from many institutions can greatly enhance the community knowledge/resource for effective and efficient VIVO ontology modeling. Initial efforts on this project start with literature analysis of best practices in VIVO ontology development. Before ontology modeling phase, the project team analyzes the six different themes and their subjects, existing controlled vocabularies, and other ontologies related to this project. In the modeling phase, the relationships to build RDF triples (i.e. subjects, objects, and predicates) are analyzed by the team in order to develop linked ontology structure.

The outcomes of this project will include but not be limited to sharing the practices of the VIVO ontology extension. The extension will especially discuss a method to display a local institution-tailored VIVO dashboard including a different classification view. This project can provide a solution for many research universities that have similar lists and interests to communicate their efforts to serve society.

Accelerating Semantic Web Content Using PULSAR and Amazon Web Services

Author: Nick Benik

Semantic Web technologies are being increasingly adopted in various fields and roles. VIVO and Profiles Researcher Networking Software (PRNS) are two Researcher Profile systems that power specific examples of large, public-facing websites. The task of efficiently running Semantic Web-powered websites that service significant Internet traffic presents an interesting technical

problem. The future of the Semantic Web cannot be fully realized unless technologies exist that are capable of handling web-scale traffic patterns in a reliable manner and at a reasonable cost.

The author previously developed PULSAR (Platform for Ultra-Large Semantic Archiving and Retrieval) in researching the economics of generating, hosting and querying large amounts of Semantic Web data. PULSAR uses the well-known principle of time/space tradeoff to enable high-performance hosting of massive static (or slowly changing) Semantic Web data sets in a cost-effective manner.

For this presentation, the data layer of PULSAR was rewritten to use the Amazon Simple Storage Service (Amazon S3), a cloud-based storage solution, to archive Semantic Web content from a Researcher Profile system. Further developments to PULSAR were completed to use the Semantic Web's URI dereferencing action to perform traffic offloading based on the requested content type. Future changes are envisioned to allow PULSAR to act as a full reverse proxy server, offloading nearly all web traffic from the primary web server and its backing Semantic Web triple store. Such significant offloading of non-dynamic content should greatly improve the consumer-viewed performance and the data provider's hosting costs in providing the public with free access to said data.

An Open Source Tool to Verify and Import Publication Data into VIVO

Authors: Rodney Jacobson and Amar K. Das

One of the challenges when implementing a VIVO instance is to populate VIVO with accurate faculty profile data which may not be available from another local source. One of the most valuable categories of faculty profile data is publication information. Harvard's Profiles RNS Disambiguation Engine provides a web service to help identify a researcher's publications from MEDLINE. In our poster, we show how we integrated a system with VIVO to allow faculty members to easily add publication data to their profiles by confirming or correcting the accuracy of suggested publications identified by the Disambiguation Engine. Our poster will break down the components of our system as follows:

User Interface
Our application was written in Python using the Django web framework. A VIVO user is taken to our application when they click on the standard icon within VIVO for adding publications to their profile. The application is styled to match our VIVO theme in order to provide a seamless experience for the user.

Interface with our VIVO Instance
We connect to VIVO's SPARQL Query and SPARQL Update API's to query and update VIVO data. We have developed tools that allow us to import, export, and update our VIVO's data using simple CSV data formatting. These tools can be configured to work with different VIVO ontologies by manually entering sample data into VIVO and feeding RDF exports from before

and after the data entry into our system. Using this data, the system learns what triples are needed to add a row of CSV data.

Interface with Profiles RNS Disambiguation Engine
We send the following information to the Disambiguation Engine: Researcher's name, email address, PubMed article numbers of the researcher's known articles, and PubMed article numbers of the researcher's rejected articles. Using this information, the Disambiguation Engine returns a list of suggested PubMed article numbers. We run a new query of the Disambiguation Engine whenever we have either exhausted the list of suggested articles or have had a suggestion rejected by the user. This allows the Disambiguation Engine to provide more accurate suggestions as it learns from confirmed or rejected articles.

Interface with PubMed / MEDLINE
Using PubMed's web service, we retrieve article information about suggested articles to present to the user, and import that information into VIVO for confirmed articles.

Our poster shows how our user interface, Profiles Disambiguation Engine, and MEDLINE data are combined to allow our VIVO users to quickly identify and import their publication data into VIVO. This is an open source project which other institutions can use to add this functionality to their VIVO instances. We hope that this tool will help other institutions obtain more complete publication data for their researchers' profiles.

SHARE: Maximizing Research Impact by Making Research Widely Accessible and Discoverable

Authors: Rick Johnson, Linda Plunket and Judy Ruttenberg

SHARE is an academic, community-driven initiative which is addressing the need to maximize research impact by developing an open, structured dataset which links to a variety of outputs across the research lifecycle. This poster addresses the goals of SHARE, the development of its first service, SHARE Notify, upcoming work within the next phase of SHARE, and related efforts to the VIVO community.

During Phase One (2015-2016) of its funding, SHARE released a beta version of SHARE Notify, an open, searchable dataset about the publication of articles, the posting of data, the deposit of presentations, etc. SHARE Notify harvests metadata from content providers including data, institutional, and disciplinary repositories and databases such as CrossRef, Clinical Trials, and PubMed Central. SHARE Notify also includes a feed that can be consumed by any institution to filter and track research release events as they are harvested by SHARE. During Phase Two (2015-2017) of its funding, SHARE is enhancing the dataset by harvesting from more sources; adding identifiers to augment searching and facilitate disambiguation; further curating the metadata; working with similar international initiatives on interoperability; and promoting SHARE Notify.

SHARE is funded by Sloan and IMLS; led by ARL and COS; and co-sponsored by the AAU and APLU. SHARE is collaboratively building a dataset of linked metadata about research release events for all. This standards-based initiative is being developed by participants representing libraries, repositories, university administrations, publishers, and non-profit organizations.

Populating and Maintaining VIVO Data by Integrating Converis with VIVO

Authors: Ann Beynon and Ted Lawless

Converis provides a complete research management solution for aggregating, curating and validating information about research and scholarly activities. VIVO is well-positioned to provide a public-facing research portal to information stored in Converis. This poster will focus on identifying, mapping, and transferring selected data from Converis into VIVO for purposes of discovery and reuse. We will describe the technical toolkit that makes this integration possible, examples of implementations, lessons learned, and future directions.

A Social Machine for Distributed Sense-Making Networks

Authors: David Lebow and Gaetano Lotrecchiano

The evolution of the Web has led to the emergence of Social Machines (SMs) that enable new and productive forms of interaction between humans and machines. An SM is a Web-based environment where machines help humans extend their cognitive capabilities and creativity (Berners-Lee, 1999). Examples of existing SMs include Wikipedia (a crowdsourcing approach for publishing a democratic encyclopedia), GalaxyZoo (a crowdsourcing approach for distributing the massive task of classifying galaxies and stars), and OpenRov (a distributed sensor network composed of underwater drones and their human handlers for exploring the ocean).

HyLighter is a new Social Machine that enables distributed sense-making networks for two primary use cases including: (a) research teams and workgroups performing literature reviews and related activities and (b) students seeking to develop key learning skills required for the 21st century information economy. In both of these contexts, sense-making involves the deliberate effort to understand information from multiple sources (i.e., intertextual thinking) and recognize non-obvious associations, frequently between contexts that seem distant from each other (i.e., far transfer). Further, sense-making represents a set of skills that remain difficult for machines to perform. These include the integration of observations and evidence with human values, intuitions, and story-telling capabilities in generating new knowledge and making decisions when faced with challenging problems.

HyLighter integrates various technologies including multilingual text analytics and visual analytics with the HyLighter collaborative sense-making engine. This combination of technologies and affordances serves as a sense-making assistant for individuals and groups. From a related perspective, HyLighter is an external memory technology. It helps users to find, organize, and synthesize high-value information from multiple sources, create new knowledge

of value, and maintain an historical record of conversations tied to important sections of documents for auditing past decisions and informing future actions.

With machines acting as specially-abled teammates, users highlight fragments of text and images, add relevant comments and tags, and create links between related pieces of information. HyLighter captures these activities as a layer of color-coded HyLights (i.e., highlighted fragments and related comments, tags, and links), provides users with a mechanism for searching, sorting and filtering the accumulating layer from many contributors and large collections of documents, displays the results in a number of formats and visualizations, and provides feedback to smart machines for machine learning.

HyLighter enables individuals and teams to produce curated conversations, a new form of publication made possible by combining various machine and human capabilities. Curated conversations address the problems of information overload and content graveyards. Curators identify sets of high-value HyLights (including links from HyLights to exact locations in their sources) on various topics and arrange the HyLights in sequences that reveal important relationships and tell meaningful stories. These sets of HyLights and their sources are available to others for navigating collections of documents and accelerating the sense-making process.

VIVO Counts: Results of the Almost-Annual VIVO Survey

Authors: Mike Conlon and Paul Albert

Each year (except for 2015), the VIVO Project has surveyed the VIVO community regarding their reasons for adopting VIVO, their experience in implementing VIVO, their effort in managing VIVO, and their view of VIVO's future. In 2016, the survey was resumed, with a greatly simplified set of questions, a new technology (Google Forms) and an effort to improve response rate. This poster will present the approach to the survey, as well as survey results, and compare results to the results of previous surveys. Through the survey, the VIVO Project comes to understand the nature of VIVO sites, and the needs of the community. The survey informs strategy, the roadmap and development.

Start Here: An On-Ramp for the VIVO Project

Authors: Mike Conlon and Graham Triggs

In this poster, Mike Conlon, VIVO Project Director, and Graham Triggs, VIVO Technical Lead, will present an on-ramp for the VIVO project -- how the project works and how anyone can become involved. VIVO can be bewildering to newcomers with its task forces, interest groups, governance groups, Duraspace stewardship, and membership. VIVO also has various communication channels, including Twitter, Facebook, and Google Groups. Add to that, VIVO is a busy place -- there are on-going projects, meetings, and on-line resources. This poster will present a unified view of the project -- its current status, how to keep abreast of VIVO, and how

to participate. The VIVO community is growing. Stop by this poster and learn more about how you can be involved.

Homegrown Altmetrics: How Could VIVO use Raw Data that Tracks Scholarly Outputs on the Social Web?

Authors: Jennifer Lin, April Ondis and Rosa Clark

Part of the promise of the Internet has always been to make vast amounts of data about organizations, people, and their activities available to all. Yet, even as we adapt to drinking from the firehose of big data, we find that much of the data we seek is still unstructured or hidden behind paywalls. The VIVO platform offers researchers the opportunity to incorporate the full breadth of their contributions into their profile and share these details within an institution.

To facilitate the availability of structured, open data about research outputs, a new service will locate and identify activity on research outputs in the Crossref corpus from both scholarly and non-scholarly sources on the web. These references will include web bookmarks, discussions, comments, social shares, and links to associated research entities from a growing list of web sources such as Wordpress blogs, Wikipedia, Twitter, Facebook, Reddit, and Mendeley. Rather than gate-keeping access to such data, Crossref is making all data publicly available for unlimited use and reuse. Crossref Event Data will serve as the clearinghouse and ensure data consistency across all the researchers, publishers, funders, institutions and service providers that use it.

This poster will focus on how the VIVO community could use this auditable and replicable raw data with an open standards approach to build its own tools to connect institutional research outputs, share faculty accomplishments, and create data visualizations.

EOL Arctic Data Connects: a Prototype Based on VIVO for Improved Discovery and Access to the Bering Sea Project Archive

Authors: Don Stott, John Allison and C. Brooks Snyder

The Bering Sea Project was a research program from 2007 through 2012 that sought to understand the impacts of climate change and dynamic sea ice cover on the eastern Bering Sea ecosystem. Over the six-year period of the program, hundreds of multidisciplinary datasets coming from a variety of instrumentation and measurement platforms within thirty-one categories of research were processed and curated by the National Center for Atmospheric Research (NCAR) Earth Observing Laboratory (EOL). For the investigator proposing a field project, the researcher performing synthesis, or the modeler seeking data for verification, the easy discovery and access to relevant data is of prime importance. The heterogeneous products of multidisciplinary field programs such as the Bering Sea Project challenge the ability of researchers to identify which datasets, people, or tools might be relevant to their research, and

to understand how certain data, instruments, or methods were used to produce particular results.

EOL, as a partner in the NSF funded EarthCollab project, is using linked open data to permit the direct interlinking of information and data across platforms and projects. We are leveraging VIVO, the open source semantic web application, to address connectivity gaps across distributed networks of researchers and resources and identify relevant content, independent of location. During the Bering Sea Project, more than 100 scientists engaged in field data collection, original research, and ecosystem modeling to connect climate, physical oceanography, plankton, fishes, seabirds, marine mammals, humans, traditional knowledge and economic outcomes. In the first round of development of EOL Arctic Data Connects (vivo.eol.ucar.edu), we transformed ISO standard metadata in the EOL production database to RDF triples. The EOL Arctic Data Connects VIVO presents another view into the Bering Sea Project scientific data archive, grouping and connecting relevant data, people and research topics for the user. In the second round of development, we are working on connecting to sources beyond our relational database in order to link hundreds of publications and authors to grants and datasets.

The poster will present our approach in connecting ontologies and integrating them within the VIVO system, using the Bering Sea Project datasets as a case study, and will provide insight into how the geosciences can leverage linked data to produce more coherent methods of information and data discovery across large multi-disciplinary projects.

The Mycroft Cognitive Assistant®: Helping Researchers find Funding and Collaborators

Authors: Paul Courtney and Anil Srivastava

The US National Institutes of Health (NIH) awards 30 billion dollars each year through competitive peer-review process to mostly academic medical institutes in the United States. There are over 2,500 universities and other academic research institutions around the world, many of them in the United States, that regularly apply for and receive grants from NIH. Increased federal oversight is driving universities and colleges to transform the way they conduct research. The array of regulations and challenges requires academic leaders and individual researchers to act with strategic and innovative vision in order to seize the grants opportunities. At the same time, the grant application process is both rigorous and extremely time-consuming. American research universities make substantial investments in people and resources to prepare and submit grant proposals. The expenditure on pre-award application process in the United States alone is estimated to be over 1 billion dollars. Less than 20% of the accepted applicants actually end up receiving awards.

At Open Health Systems Laboratory (OHSL), we have spent the last five years studying the grant application process both from inside the NIH and from the perspective of those institutions applying for funding, in the US and abroad. We have experience developing Research Network Systems and of preparing and submitting grant applications in partnership with leading

resources such as labs or computational capacity; (3) draft of applications for grant opportunities created by Mycroft. Our poster presents our progress on developing this cognitive assistant.

VIZ-VIVO: Towards Visualizations-Driven Linked Data Navigation

Author: Muhammad Javed

Abstract: Driven by different needs, users look at VIVO from different viewpoints. Colleges require its structured data for reporting needs. Deans of the colleges are interested in identifying where collaborations are happening. The library is interested in preservation, delivery and open access of its scholarship data. Researchers are interested in reporting their contributions and expose their scholarly work to the world. Such viewpoints do not fully overlap with each other. The current VIVO user interface represents the scholarship data in a list view format. Such representation of the scholarship data is not easy to use or consumable by the users. Here, we opt for a bottom up approach where we start from a list of questions that different users would like the VIVO application to answer. Such questions include "list of contributions/research output of a unit/person," "list of potential collaborators," "list of faculty members with similar research interests," "impact of a scholarly work," etc. In this poster, we present our ongoing work of using D3 visualizations in VIVO profile pages. These visualizations provide interactability and allow a user to navigate through the linked data in an intuitive manner. Visualizations are constructed on the fly, based on the underlying RDF data. This visualization-driven approach provides efficient overview of the network of interconnected resources.

A Person-to-Research Area (P2RA) map view demonstrates the overlap of research interests among different faculty members of a department. The P2RA map is helpful for the identification of i) the list of research interests of a person and ii) the list of potential collaborators. A Word Cloud (WC) view goes one step deeper and presents the domain expertise view of a faculty member. At the person level, the WC consists of the keywords that we accumulate from the scholarly works being published by the researchers. Clicking on a keyword allows a user to view the list of the work titles where such keyword is actually used. The user can further click on one of the work titles from the list in order to navigate to the specific VIVO page of such work. We use sunburst D3 visualization to present the inter/cross unit collaborations' view. The collaboration view consists of three layers: the college-level layer (innermost), departmental layer (middle), and the person-level layer (outermost). A user can start navigation from the college level layer and, while traversing through the department (of his/her interest), can reach to a specific person who has co-authored some of the collaborative work. Clicking on a person's name area displays the list of the collaborative work titles. The user can further click on one of the work titles from the list in order to redirect to the specific VIVO page of such work. We use a bubble chart for the demonstration of the active grants of a department. In addition to the traditional metrics (i.e., citation count, journal ranking, etc.), an Altmetrics badge is used to demonstrate the impact of a scholarly work. We discuss the architecture and the integration of these visualizations into the VIVO framework.

Profiles Research Networking Software: Overview, Recent Enhancements, and Future Plans

Authors: Nick Brown and Griffin Weber

Profiles Research Networking Software (RNS) is a free open source semantic web application which uses the VIVO ontology to generate searchable online profiles of an organization's investigators (http://profiles.catalyst.harvard.edu). It is used in dozens of institutions worldwide, including academic health centers, universities, pharmaceutical companies, and federal agencies. In addition, it has an active developers' community, which frequently contributes to the software. This poster describes (1) several unique features of Profiles RNS, (2) enhancements to Profiles RNS made in the last year, and (3) plans for future features.

PROFILES RNS: A unique aspect of Profiles RNS is its Name Disambiguation Engine, which automatically finds and links information about investigators from local administrative databases and external sources such as Medline. It then uses these data to discover connections between investigators, such as prior co-authorship of publications, organizational relationships and geographic proximity. The resulting researcher networks are displayed on the website as interactive visualizations. Other features of Profiles RNS include granular privacy control, integration with ORCID, and support for OpenSocial (ORNG) plugin "gadgets."

RECENT ENHANCEMENTS: This past year, Active Directory and Shibboleth authentication options were added; a new Activity Log keeps track of changes users make to their profile pages; Flash based visualizations were converted to HTML5; and Altmetric Badges, which indicate the number of times articles have been cited in social media, were added to publication lists.

FUTURE PLANS: The Harvard development team is building a new grants and funding module, which like publications will enable both manual entry of funding and automatic import of NIH grants. USCF is contributing two modules for displaying education/training and custom keywords.

This project was funded by NIH grants 8UL1TR000170 and 1UL1TR001102, and Harvard University and its affiliated academic health centers.

How Can We Help Researchers Increase the Reach and Impact of Their Work?

Author: Charlie Rapple

The Kudos project was set up in 2013. We have been providing researchers with a platform for explaining their work in plain language, and for centrally managing their sharing of that work – we give them a trackable link and can then see how they share (email, social media, academic networks, etc.) and map this against publication metrics (downloads, citations, and altmetrics). By building up a central dataset in this area we are able to provide insight (to researchers,

institutions, publishers, societies, funders, etc.) into which sharing media are most effective, and the extent to which metrics are improved. For example, the altmetrics team from Nanyang Technological University have analysed the data we've built up so far, and determined that when researchers actively explain and share their work via Kudos, it increases downloads by 23% on average. The study has also found that researchers are more likely to use Facebook than Twitter, for sharing their work, but that shares via LinkedIn are most likely to be clicked on. We are eager to share these results with those who support research dissemination as the evidence we are building can (a) help libraries (and other units) provide more tailored guidance to researchers, and (b) help early career researchers demonstrate to their supervisors that small outreach efforts are a good use of time, as they can substantially increase readership and from there, impact.

"VIVO-Lab" at the German National Library of Science and Technology: Review and Outlook

Authors: Tatiana Walther, Lambert Heller and Ina Bluemel

Looking back at 2015, we started with the VIVO implementation for the German National Library of Science and Technology (TIB), within the scope of a student project. A prototype was set up and complemented with an in-house developed feature, "VIVOPaths," which visualizes the semantic connections between researchers and their activities on an external web page through a SPARQL-based process – see https://github.com/hornmo/vivopaths.

In September 2015, we organized the first German VIVO workshop – "VIVO as a Research Information System in Practice" – at the TIB in Hannover. About 60 participants from various German and European research organizations, universities, and libraries used the opportunity to learn more about VIVO or to exchange their user experience. Visitors included also representatives of publishing companies (Elsevier and Thomson Reuters), software companies (Symplectic), and Brian Lowe (Ontocale SLR) with his opening keynote address.

Today we continue our work on the VIVO TIB. The application is considered to be a public research information system, which reproduces the scientific output and activities of the TIB – primarily of the library's "Research and Development" division. Currently, we are about to finalize the implementation of the VIVO TIB and upgrade it to VIVO 1.8.1.

Beyond this, we are investigating the ways of inserting Open Metrics data, in particular from the CrossRef and DataCite DOI "event data" services, into VIVO. The use of the GRID and Wikidata data sets and services for less redundant, but more consistent and enriched information is planned as well. Regarding Wikidata, it is particularly worth mentioning, that Ina Bluemel and Lambert Heller, together with the Hanover University of Applied Sciences and Arts, initiated a project for the automatic harvesting and indexing of technology pictures and graphics, using the infrastructure of Wikimedia, Wikisource and Wikidata – see http://doi.org/bhmz.

As the Kerndatensatz Forschung (KDSF), the new national standard for representing research infrastructure and activities of academic and research organizations in Germany, is gaining traction, research information systems are recommended to be aligned to the specification of KDSF. In this context we explore the scope of the internal usage of VIVO for aggregating sensitive administrative data in a separate non-public VIVO instance in alignment with the KDSF – see https://github.com/VIVO-DE/kdsf-alignment.

According to our recent observations the demand for open source alternatives to proprietary Current Research Information Systems (CRIS) among academic institutions in Germany has significantly increased. With regard to the continuously growing number of requests for VIVO from external organizations, we intend to set up a strategic concept for VIVO services at the TIB.

Several VIVO applications for external organizations are planned to be implemented by the TIB's "VIVO Lab" – e.g. a VIVO for the international working group at the Leibniz Association. In the meantime, the VIVO Team at the TIB has been growing. In February 2016 it was joined by a linked open data librarian, followed by a web developer. A project coordinator will soon be completing our "VIVO Lab."
Regarding our future prospects, we are looking forward to further expanding of our team, exploring new research areas, and facing the next challenges around VIVO.

Smashing Silos Through Community Engagement Maps: a VIVO Extension

Authors: Tessa McKenzie, Audrey Trussell, Valerie Holton, Stephen Guilford and Jon Deshazo

How can VIVO address silos while bridging academia, industry, and the community? By forging connections and collaborations across the disciplines through the exploration of community-engaged partnership activity.

Virginia Commonwealth University's (VCU) strategic principles focus on community engagement to provide high quality learning and research, and generate innovative solutions to societal challenges. VCU is actively involved in the Richmond region through community-engaged partnerships in research, teaching, service/outreach, and patient care. With this comes the need to develop specific mechanisms to identify, track and evaluate these activities and the partnerships involved. No existing system satisfied our requirements and fit within our current infrastructure, so we developed a community engagement ontology that extended the existing VIVO ontologies and built an interactive search interface and map display. The software consists of freemarker code, Google Map API, and ontology additions – and is drop-in ready for later VIVO installations.

The domain ontology is descriptive of community engagement activities, partnerships, and organizations. A Community Program is an activity that involves one or more Partner Organizations and VCU Unit. The Community Program is mapped to one or more geographic

locations (typically at the neighborhood level) and is associated with activities and issues of focus. The ontological connectivity within the VIVO database links scholarship, organization structure, and other data together enabling deep computer reasoning and powerful reports.

Navigation can be done either through the conventional VIVO interface, including text search and menu browsing or a new map interface. The Interactive Community Map has multiple drill down menu features and displays community partner locations as pindrops, and activities through heatmap shading.

VCU's Community-Engaged Partnership Map is a tool that will serve as a pan-institutional resource to connect, coordinate, and collaborate community-engaged activities and identify potential areas of opportunity. The Map uses an extension of VIVO, an open source web application that enables the discovery of research and scholarship and promotes networking across disciplines. Operationally, we define partnerships as a sustained level of interaction and/or ongoing engagement between institutions of higher education and communities for 2+ years.

VIVO and the Map showcase our community partnerships for internal and external audiences and give community partners a simple pagefront presence, encouraging more partnerships. VIVO and the Map provide flexible reporting and data aggregations at the university, college, school, department, and faculty levels. This effort aligns with institutional commitment to serve as a resource to "advance focused and strategic university-community engagement that addresses critical needs and opportunities in the region."

This poster presentation will review the development of the Community-Engaged Partnership Map, an extension of VIVO. Opportunities and challenges will be highlighted, along with recommendations for other VIVO users seeking to adopt this extension or develop similar mechanisms including: 1) Community-engaged partnership definitions and ontology development, 2) University-wide data collection methodology, 3) System development and data integration, 4) Collaboration and ownership, and 5) Maintenance resources and sustainability.

Enhancing SciENcv Through Semantic Research Profile Integration with the VIVO-ISF Ontology

Authors: Marijane White, Matthew Brush, Shahim Essaid, Robin Champieux, Adrienne Zell, Colin Grove, Syeda Momina Tabish, David Eichmann, and Melissa Haendel

SciENcv, the US Federal Science Experts Network Curriculum Vitae, is an online system for simplifying the creation of researcher biographical sketches or biosketches, which are required when applying for federal funding. SciENcv profiles are curated and controlled by researchers themselves – they own the data, they control what data are public, and they edit and maintain the information contained within the profiles, which includes expertise, employment history, educational background, and professional accomplishments. The system leverages data from myNCBI and eRA Commons, and includes links to ORCID researcher identifiers. The structure of

SciENcv biosketches is defined by an XML Schema Definition and profiles can be downloaded in XML format. The system aims to eliminate the need for researchers to repeatedly enter biosketch information and reduce the administrative burden associated with federal grant submission and reporting requirements, as well as creating a repository of researcher profile data where researchers can describe their scientific contributions in their own language.

The VIVO Integrated Semantic Framework (VIVO-ISF) is an ontology for representing people, works, and the relationships between them and was developed as a merger of the VIVO and eagle-i ontologies. VIVO-ISF is the designated standard for research profiling data in the Centers for Translational Science across the US, and is utilized by a variety of research profiling tools such as Harvard Profiles. However, many research profiling tools are on different versions and/or have not yet adopted VIVO-ISF as a standard representation. SciENcv is in this category.

The SciENcv integration project aims to create interoperability between the SciENcv XML Schema and the VIVO-ISF ontology. This will facilitate generation and consumption of standard, compliant data, expand the role of SciENcv in exchanging data within the research landscape, and broaden the definition of researcher impact by facilitating the emergence of new standards relating to the contribution and attribution of researchers to software, datasets, and other scholarly products. Achieving this integration will support improved research profiling analytics across a much wider set of data sources and allow researchers to more effectively and specifically describe their contributions.

The project enables mapping between the SciENcv XML Schema and the OWL constructs in VIVO-ISF. The mapping is being validated via CTSASearch, which ingests VIVO-ISF, ORCID, PubMedCentral, Medline, and more, allowing for a fully integrated and curated set of institutional data. CTSASearch is also adding community detection visualization features for search results.

Here, we showcase the interpretation of the SciENcv schema, its alignment to the VIVO-ISF ontology, highlight some of the competency questions being used to evaluate exchanged profiles, and some example queries that can be used to access SciENcv profile data via a VIVO-ISF version 1.6 SPARQL endpoint.

Who We Are and What We Contribute: the Open Researcher Information Framework as a Means to Understand Scholarship

Authors: Marijane White, Shahim Essaid, Matthew Brush, Karen Gutzman, Alexandre Rademaker, Tenille Johnson, Muhammad Javed, Jon Corson-Rikert, George Chacko, Mike Conlon, Kristi Holmes, Melissa Haendel, and Violeta Ilik

OpenRIF, the Open Researcher Information Framework, is an open source community devoted to representing the scholarly research ecosystem – all the things we do and all the things we contribute. The community works on developing and promoting interoperable and extensible semantic infrastructure, such as the VIVO Integrated Semantic Framework (VIVO-ISF), an

ontology for representing people, works, and the relationships between them; federated databases modeled on PARDI, the Portfolio Analysis and Reporting Data Infrastructure, for research impact and evaluation; and eagle-i, which aims to make discoverable research resources and their relationships to scholarly activities. OpenRIF seeks to align with other scholarship knowledge representation efforts in order to facilitate the scalability and interoperability of information about scholarly efforts, such as CASRAI, DataCite, EuroCris, LD4L, NIH BD2K, ORCID, SHARE, and the SPAR Ontologies.

One area of particular focus is the attribution of contributions to scholarship beyond manuscript authorship. Advances in technology have had significant impact on the structure of research teams, broadening the participation in scientific discoveries of researchers from a wide variety of backgrounds and in forms that may not be recognized by traditional measures of scholarly impact. Limiting metrics to grants and publications means these contributions can go unrecognized, which in turn limits researchers' career advancement, financial incentives, funding opportunities, and scholarly recognition. Representation of all scholarly contributions and activities and their relationships to persons and organizations is critical to the health of the scholarly ecosystem as a whole.

OpenRIF reuses and integrates existing semantic infrastructure as much as possible, with content derived from OBO, SPAR, FOAF, and SKOS, etc., enabling shared semantics across domains and maximal extensibility in different scholarly domains. Recently, OpenRIF has been working on expanding this infrastructure to include a wide array of contributor roles in the context of the FORCE11 Attribution Working Group, an OpenRIF community member. Here, we have reviewed existing scholarly contribution taxonomies and have expanded the CRediT taxonomy beyond authorship to create a contributorship model that covers a diverse selection of fields of research and that is specific enough to describe contributions in a more meaningful fashion. This contribution role taxonomy is currently being evaluated within OpenVIVO and the OpenRIF community welcomes feedback.

This poster describes the contribution role taxonomy, provide some analysis of the usage of the contribution roles in OpenVIVO, and discuss plans to broaden the scope of OpenRIF to industry research and to improve coordination between OpenRIF and the VIVO project.

Integrating and Building on Elasticsearch and VIVO-ISF Data

Authors: Don Elsborg, Nate Prewitt and Alex Viggio

The VIVO development team at the University of Colorado Boulder has replaced the default People list view with a faceted search solution developed by the Deep Carbon Observatory team at Rensselaer Polytechnic Institute. This 100% open source integration uses Elasticsearch and Cottage Labs' facetview2. Elasticsearch is seeing widespread adoption in industry, and has become the most popular enterprise search solution. Our poster will explore what we learned from this integration, and how we are reusing Elasticsearch to power a modified version of University of Melbourne's Find an Expert Capability Map. We're also interested in discussing

several new ideas for presenting and sharing VIVO-ISF data using Elasticsearch, RESTful APIs, JSON-LD, and Javascript.

Comparing Measures of Research Output

Authors: Yunpeng Li, David Turner, Emily Maemura, Diego Serrano, Kelly Lyons and Eleni Stroulia

Multidisciplinarity and collaboration are increasingly recognized as necessary in order to exchange knowledge across disciplines, foster learning, and address problems that transcend a single distinct academic field. Existing bibliometric measures such as citation counts and the h-index tend to reinforce research output within a single field. Recently, new measures that attempt to capture the collaborative, multidisciplinarity and social aspect of research output have been developed. We seek to understand how new metrics compare with more traditional measures. In this poster, we present results of our analysis of several measures of research output from the GRAND Network of Centres of Excellence, a large-scale, Canadian, multidisciplinary research network conducting research on digital media with numerous academic and industrial partners.

Enhancing Collaboration using Research Study Information

Authors: Randall Alexander, Tami Crawford, Stephanie Gentilin, Signe Denmark, Leslie Lenert and Jihad Obeid

When looking for collaborators for Clinical Trials or Research Studies, Profiles Research Networking system is a great place to look. However, there is no way to find investigators who are currently engaged in clinical research who have not yet published their findings.

This submission describes how MUSC and Health Sciences South Carolina(HSSC) is addressing this issue. MUSC/HSSC has a statewide website, SCResearch.org, with a list of research studies, a summary in plain language describing the study and contact information for the study. This information along with the keywords for the each study is provided to the website in a nightly feed from the shared electronic IRB system. We are developing a similar feed directly into our statewide instance of Profiles. The feed is used to add keywords from the investigators' studies into searchable fields in Profiles. This will allow others to find active collaborators for research studies. As with publications, this data requires minimal input and/or adjudication by investigators.

A widget is being developed that publishes the list of studies for each investigator, for both actively recruiting and closed studies. Links will be provided to the active clinical studies in

SCResearch. Studies that are closed for recruitment will be listed with the title and the associated keywords. Future plans include adding reciprocal links from SCResearch to Profiles allowing the interested party to view the investigator's research profile.

Friday, August 19th

REGISTRATION

7:30 PM – 5:00 PM
Room: Colorado Ballroom Foyer

NETWORKING BREAKFAST

7:30 AM – 8:30 AM
Room: Colorado Ballroom Foyer

AWARDS & KEYNOTE PRESENTATION

8:30 AM – 10:00 AM
Room: Colorado Ballroom E – F
Verifiable, Linked Open Knowledge That Anyone can Edit
Presenter: Dario Taraborelli

15 years since the birth of the project, Wikipedia's vision of a free encyclopedia that anyone can edit has come of age. Today, as one of the top 10 sources of global traffic to DOIs, Wikipedia is not only a comprehensive, multilingual repository of the world's knowledge: It is also one of the primary entry points to the scholarly literature and a popular vehicle by which scientific knowledge is disseminated globally. In this talk, I'll showcase a number of ways in which this vision is augmented by Wikidata – the knowledge base that anyone can edit and Wikipedia's fastest growing sister project. By linking source metadata to facts and encyclopedic entries, Wikidata allows scholars, librarians, data curators, software developers and volunteer contributors to share in the sum of a *new kind* of knowledge. I'll review, in particular, the most recent efforts aiming to incorporate into Wikidata a bibliographic repository and an open citation graph, to ensure that linked open knowledge is also persistently and transparently verifiable.

VIVO 2016 Conference Schedule (Tentative)

BREAK
10:00 AM – 10:30 AM

PANELS AND PRESENTATIONS

Room: Colorado Ballroom G

Visualizing Virtual Communities in the VIVO Profile Ecosystem
Presenter: David Eichmann
Track: VIVO Tools

CTSAsearch (http://research.icts.uiowa.edu/polyglot) visualizes co-authorship connections between the matched profiles using a force graph implemented in D3. Connections are pre-computed at profile harvesting time using multiple alternative identifiers (DOI, PMID, and PMCID) present in the profile data. OCLC pmid2doi crosswalk data is used to span the identifier spaces. Useful force graph visualizations are possible for 'reasonable' result scales (n ~ 200). This presentation will address recognition of research collaboration communities whose identity beyond the default notion of institution.

Institution-level visualization
Labeling nodes (profiles) by institutional affiliation has proven useful for small-scale (n<200) graphs, particularly for topics relating to research communities relatively evenly distributed across institutions. However, for searches involving large numbers of profiles from a comparatively small number of institutions, the inherent substructure of the collaboration networks gets lost in the 'hairball' of interconnectivity.

Inter-institutional community visualization
Focusing on community detection in the network structure is proving to be a far more robust approach to untangling large networks. I use a user-selectable set of community detection algorithms [1,2] to identify community membership based upon characteristics of the local neighborhood. The resulting node coloring reveals natural substructure even in densely interconnected graph components.

[1] V.D. Blondel, J.-L. Guillaume, R. Lambiotte, E. Lefebvre, Fast unfolding of communities in large networks, J. Stat. Mech. Theor. Exp. 10, P10008 (2008)

[2] L. Waltman and N. J. van Eck, A smart local moving algorithm for large-scale modularity-based community detection, Eur. Phys. J. B (2013) 86: 471.

Using VIVO for Scientific Applications

Presenters: Matthew Mayernik, Anne Wilson and John Furfey

Track: VIVO Tools

Semantic Web technologies have been used in geoscience information and data applications for a couple of decades. Geoscience focused ontologies have been developed at many levels of detail, from controlled thesauri to formal ontology specifications with properties, logical constraints, and underlying axioms. This geosciences semantic community has largely existed separately from the VIVO community. In the last five years, however, a small number of VIVO instances have been deployed to manage and share information about geoscience projects, data, and other resources. Within the VIVO data model, almost everything can be represented as a first-order object – not just people, organizations, publications and grants, but instruments, projects and their components, work groups, datasets, methodologies developed, presentations, and any other items of interest declared using appropriate ontologies.

This panel will feature presentations that focus on using VIVO for scientific applications. The goal will be to bring about discussion of the benefits and drawbacks of representing scientific information in VIVO. Presentation and discussion topics may include: extending the VIVO application interfaces, coupling the VIVO-ISF ontology to domain-specific ontologies, and leveraging VIVO for supporting cross-organizational projects.

The panel will be organized to engage the audience in questions about VIVO's ability to support the use cases presented by scientific research organizations and projects. These use cases may include using VIVO to represent information about multi-organizational and multi-disciplinary scientific projects, and capturing the relationships between people, organizations, grants, datasets, publications, scientific instruments, and research sites. Speakers in this panel will present about ongoing VIVO initiatives to illustrate current efforts.

The Citation Impact Tool: Visualizing Research Impact

Presenters: Michael Bales, Paul Albert, Prakash Adekkanattu, and Terrie Wheeler

Track: VIVO Tools

Many metrics may be used to evaluate a researcher's scholarly impact. Because research articles are a prominent currency of academic scholarship, citation impact plays a central role in scholarly impact assessment. Times cited data are highly skewed, and it is well recognized in the bibliometrics literature that nonparametric approaches are appropriate for skewed distributions. By convention in bibliometrics the scholarly impact of an article is assessed against a reference set of other articles of the same publication type (i.e. academic article or review), in the same field, and published the same year. The articles in the reference set are used as a baseline to calculate the percentile rank of the number of times a given article has been cited. To our knowledge, no currently available systems are capable of displaying article-level impact based on percentile rank of times cited for the entire percentile distribution.

To support research impact assessment, we have developed the Citation Impact Tool, an interactive component of VIVO Dashboard. The system allows users to assess the percentile rank of the number of citations specific articles have received relative to peer publications. Publications are individually benchmarked against reference sets of 200 articles of the same type, topical category of journal, and year of publication. Times cited data comes from Elsevier's Scopus. The system uses an iconographic bar chart to portray article-level citation impact of multiple articles in one view, and allows users to filter by type of article, author order (first/last author rank), and topical category.

The system uses Thomson Reuters Web of Science journal categories to assign articles to individual fields. When an article is in more than one category, the percentile rank of times cited is calculated separately against a separate reference set of articles for each category. A mean percentile rank for the article across categories is then calculated and, in turn, used in the bar chart. Because recently published articles have had little time to accumulate citations, articles from the past two to three years are deliberately excluded from the charts.

The code will be shared with the community.

Room: Colorado Ballroom H

Open VIVO: Goals and Results

Presenters: Mike Conlon and Violeta Ilik
Track: Linked Data

Open VIVO is a demonstration of an open VIVO that anyone with an ORCID identifier can use. Users can join Open VIVO. The project engaged more than a dozen team members in creating infrastructure, branding, a new contribution ontology, new interactive elements, new datasets of RDF for VIVO, and new reuse elements. Open VIVO provides an opportunity for every scholar to participate in an immediate way, to add data and indicate contribution to scholarly works, and to provide data in a truly open and accessible manner on a daily basis. In this panel, participants in the Open VIVO project will share their views on the project, its goals and results. Elements of Open VIVO will be ported to future versions of VIVO. Implications of Open VIVO for VIVO will be discussed.

Semantically-Enabling Figshare as a Basis for Rapid Ontology

Presenter: Simon John Porter
Track: Linked Data

One of the challenges for the progression of research metadata standards is that technology often gets in the way. Bound by the pressure not to be weighed down by technical debt, and the desire to appeal to as broader an audience as possible, the adoption of metadata standards within technology platforms is conservative. Ideally, this is not how things should be. The adoption of standards and practices should be community led, with technology functioning as the enabler. This presentation will showcase one attempt to move closer towards this goal by semantically enabling Figshare as a basis for rapid ontology development and adoption.

As part of the work required to build the FORCE2016 OpenVIVO initial implementation, a mapping from the Figshare API to VIVO RDF was created. This work was then extended to expose all RDF documents that are part of a Figshare record as a single RDF graph. It is intended that this graph be accessible from the Figshare article URL when RDF is requested. With two developments:

- The use of the Figshare API to create custom metadata forms that can write additional triples to a Figshare article, and;
- The inclusion of RDF as a metadata output of Figshare's new OAI-PMH feed; Figshare records can now be annotated with new ontologies that are immediately discoverable.

As a concrete example of this approach, a working supplementary metadata form will be demonstrated that allows authors to be annotated with organizational affiliations, and additional roles assigned using the draft CRediT ontology. In a similar fashion, this same approach can be used to extend the adoption of a greater part of the VIVO ontology beyond faculty profiles. Through the harvesting of Figshare records into Symplectic Elements, Figshare

records can be linked to grants and publications in Elements, and then this additional metadata pushed back onto the Figshare record as VIVO RDF through the reuse of Symplectic's VIVO harvester code.

In this presentation, we will demonstrate the first iteration of a user-focused, modern reimagining of the VIVO interface. Using Bootstrap and CSS, we will show how some layout and design improvements can work wonders on the stock VIVO interface. Coupled with some basic SEO knowledge, any institution can use these simple methods to drastically improve their own VIVO and make it the 'expertise' hub it's destined to be. We will also demonstrate a new set of integrations with products such as Altmetric, figshare and The Conversation to help further enrich and showcase institutional data and expertise.

Updates from a Registered Service Provider
Presenters: Jonathan Breeze and John Gieschen
Track: Product Information

As the first official DuraSpace Registered Service Provider for VIVO, Symplectic provides services that cover everything from installation, to support of VIVO on client servers, to providing our own hosted solution. Our continued involvement in user groups, the DuraSpace community and participation in VIVO working groups mean that we are well placed to support the needs of the VIVO community.

Symplectic is also a registered service provider for Profiles RNS, the open source Semantic Web research networking platform based on the VIVO ontology.

This presentation will reflect on Symplectic's second year as a Registered Service Provider, showcasing the projects we have supported for both VIVO and Profiles RNS, including institutions outside the US - covering installations, data analysis, data population, customisation, and hosting. As well as our current engagements, we will introduce a number of new initiatives we plan to introduce in support of both open source communities.

Room: Colorado Ballroom I

Voted Most Popular: Generating CVs from VIVO
Presenters: Julia Trimmer, Richard Outten, and Danielle Heckman
Track: VIVO Tools

A frequent request from Duke faculty members is, "Can I create a CV from Scholars@Duke?" Now we can say, "Yes!" Last fall, the School of Medicine's Appointments, Promotions and Tenure Committee asked for a CV generator to help streamline dossier submissions as well as to encourage maintenance of Duke's VIVO implementation, Scholars@Duke. With the help of the APT committee and the VIVO widgets, users can generate (client-side) a simple CV in a Word document. The Scholars@Duke team is working to enhance the CV by adding a number of user-maintained fields. Phase two of the Scholars CV will take many months and a lot of development effort -- will the APT committee like it? And more importantly, will faculty use it? Get the inside scoop at this presentation.

Faceted Searching in Scholars@Duke

Presenters: Ulysses Cannon and Julia Trimmer
Track: VIVO Tools

Users are searching Duke's VIVO, Scholars@Duke for many reasons: Duke faculty to identify collaborators for research efforts, prospective students and their families to learn about our faculty and scholarship, patients to learn more about our medical providers, and industry to locate expertise. But many of our users have told us that our current search functionality is insufficient to effectively and optimally facilitate those use cases.

In this presentation, we will discuss in more detail the feedback which prompted this proposed enhancement of our search functionality, as well as, the key elements of the analysis and considerations that informed our final design. Faculty and our broader user community have requested search functionality that enables them to filter results to the level of granularity required and by facets that are helpful. In addition, they would like to sort the search results by relevance or other factors to provide context for the results listed. Also, some users are seeking a very specific set of results and need an "advanced" search that's easy to use.

Our proposed enhancement of the Scholars@Duke search functionality seeks to address those limitations by enhancing our basic search to provide more effective filtering of search results, enabling the search results to be sorted by relevance, and adding new advanced search capabilities.

Usability vs. Purity: How UCSF and Duke Enabled Data Reuse by Going Beyond Linked Data

Presenters: Julia Trimmer, Anirvan Chatterjee, Eric Meeks, Richard Outten and Brian Turner

Track: Linked Data

Duke and UCSF have come up with ways to make their research networking platforms indispensable at their institutions, winning friends and allies, and strengthening the case for renewed funding. Others can do this too — but only if we retrofit VIVO, Profiles, and other platforms with simple data reuse mechanisms. In this session, we'll talk about techniques for developers, traffic growth on downstream sites, and researcher engagement. We'll look at examples of sites that consume VIVO and Profiles data and talk about what it's like to share data across nearly 50 local websites and apps.

LUNCH (ON OWN)/BIRDS OF A FEATHER DISCUSSIONS

12:00 PM – 1:30 PM

INVITED SPEAKERS

1:30 PM – 2:15PM
Room: Colorado Ballroom E - F
VIVO Futures: Scholarly Infrastructure, Sustainability, and Community
Presenter: Sandy Payette

While VIVO can be understood as a platform for creating and managing open linked data, paradoxically and in practice, VIVO can still exhibit elements of "silo" architecture, often a result of the institutional-centric orientation of many existing VIVO installations.   In an expanding ecosystem of scholarly systems, as well as the cross-institutional nature of scholarly communication, I ask the question, how does VIVO position itself for the future?   In its earlier history, VIVO identified itself an open source semantic web *application for scholarly networking and discovery* (Borner, et. al., 2012, p. 4).   What does this mean today? What are the present tensions and future opportunities for VIVO in an era when many new players have entered the field of scholarly and research infrastructure?

By both definition and by design, VIVO is compatible with Web infrastructure due to its embrace of semantic and open linked data fundamentals. Nevertheless, gaps and friction points exist that can interfere with end users' ability to experience the full potential offered by VIVO as a scholarly knowledge "graph."   For all the richness of the VIVO ontology and the flexibility of VIVO data encoded as RDF, it remains difficult for users, and even some software developers, to use VIVO to generate *answers* to key questions – sometimes basic user questions, but especially more subtle or complex ones.   One barrier may be the necessity of having specialized knowledge about ontologies and the SPARQL query language to be able to discover interesting patterns and unleash the meaning of relationships in the data.   Another

barrier may be path dependencies to the original VIVO user experience, reinforcing a paradigm of how VIVO data can be experienced.

These are challenges that can be surmounted. As part of emerging scholarly infrastructure, VIVO must be examined from multiple vantage points, including (1) the dynamics of the ecosystem of vendors and systems serving the scholarly domain, (2) the role of open systems and open source in promoting open knowledge, and (3) the health and growth of communities that invest in VIVO innovation and VIVO sustainability. In this talk, I will draw upon themes of emerging infrastructure, my experiences in open source and community leadership, and my recent work on "Scholars@Cornell," a new VIVO-based project.

2:15PM – 3:00 PM
INVITED SPEAKER PRESENTATION
Additional information will be available soon

BREAK
3:00 PM – 3:30 PM

FEATURED PRESENTATIONS
Room: Colorado Ballroom E-F

3:30 PM – 4:00 PM
Extending VIVO Infrastructure to Support Linking Information Between EarthCollab VIVO Instances
Presenters: Huda Khan, Matthew Mayernik, Keith Maull, M. Benjamin Gross, Mike Daniels, Steve Williams, Linda R. Rowan, Sandy Payette, Erica Johns, Dean Krafft and David Eichmann

Geosciences research, given its interdisciplinary and inter-organizational nature, is often conducted using distributed networks of researchers and resources including instruments and platforms. To better enable the discovery of the research output from the scientists and resources used within these organizations, UCAR[1], Cornell University, and UNAVCO[2] are collaborating on the EarthCollab[3] project which seeks to leverage semantic technologies to manage and link scientific data.

As part of this effort, we have been exploring how to leverage information distributed across multiple VIVO instances by working on mechanisms to enable a VIVO instance to lookup,

consume, and display information from another VIVO instance without having to ingest RDF. The challenges we encountered and the path we took to support inter-VIVO communication can provide useful input to the larger discussion of how to extend VIVO's infrastructure to more seamlessly integrate external information.

Our presentation will include the following:

- Demonstration of linking between VIVO instances: Using multiple VIVO instances setup as part of EarthCollab, we will show how we have extended VIVO's core infrastructure to enable discovery of information from an external VIVO instance and support display of this external information within a VIVO profile.

- Technical infrastructure: Central to the interconnectivity demonstrated are the abilities to: (1) designate multiple URIs from separate VIVO name spaces as equivalent to each other or to an independent unique identifier (such as an ORCID ID) using sameAs assertions, (2) retrieve the appropriate URIs that might designate the same person using a lookup service based on (1), and (3) display information for a URI from a different VIVO instance without having to copy or duplicate information. We will also discuss how these extensions can support other linked data lookups and sources of information. In addition, we have built mechanisms for displaying all the RDF underlying a VIVO profile. This RDF depiction can go far beyond the default linked data representation as it captures and presents the information about the related entities that is used to generate a particular profile.

As challenges and open questions, we will discuss how this mechanism of interconnectivity relies on reliable and open lookup options (e.g. Solr search indices) and how we have had to address how to open up our individual VIVO instances to enable communication. Additional questions include how to display the information obtained from an external VIVO instance, both in order to preserve the receiving VIVO instance's brand and to handle discrepancies between ontologies, content, and/or VIVO versions.

[1] http://www2.ucar.edu/
[2] https://www.unavco.org/
[3] http://earthcube.org/group/earthcollab

4:00 PM – 4:30 PM

VIVO for Historical Persons
Presenters: Thea Lindquist and Alex Viggio

This presentation will explore the motivations behind and report early results of the VIVO for

Historical Persons (VIVO4HP) project. This experiment seeks to reuse and extend VIVO and the VIVO-ISF ontology to represent and facilitate discovery of historical persons, a humanities use case that reflects VIVO's original purpose. We will address whether VIVO can be reasonably adapted to this purpose, which would be of interest to a wider digital humanities community including other projects that are using linked data.

VIVO4HP's initial use case focuses on historical persons belonging to a specific professional community – diplomats that served England in the seventeenth century. Our target data is derived from the standard British biographical source, the Oxford Dictionary of National Biography (ODNB). This authoritative data set will be supplemented by event data on diplomatic missions from A Handlist of British Diplomatic Representatives. VIVO presents Linked Open Data, which offers the potential to make our use case data available to other digital humanities projects and applications for subsequent analysis, evaluation, and visualization.

We are undertaking this experiment as a multi-stage project. The first step is to manually create profiles in VIVO for a limited number of diplomats to identify issues with the source data, default ontology, data mapping and transformation, and online display. The second is to make adjustments to the ontology and online display to address gaps. The VIVO-ISF ontology re-uses and extends a number of established ontologies, such as FOAF and BIBO, and we will evaluate the integration of other ontologies relevant to our use case. The third is to automatically ingest ODNB data into VIVO, using custom scripts where possible to address data mapping and transformation issues. The final step is to augment the profiles with other data sources, such as historical sources, other web sites, and linked data.

We anticipate a variety of challenges and will be prepared to discuss how we addressed these. General issues with historical data include dealing with data ambiguity (e.g., dates, spelling variations), incomplete biographical information, individuals' identities (e.g., using noble titles instead of their personal names), and historical geographies. The ODNB data was created with another purpose in mind and thus can be incomplete. It may be untagged or not tagged to the specificity necessary for certain types of data that might be desirable to represent for our use. Lastly VIVO is meant to represent a researcher's professional life, and does not incorporate some of the personal, political, or social aspects desirable for representing historical persons.

How our experiment is able to deal with these issues, and the level of intervention required, are key factors in our assessment of VIVO's utility for the discovery and representation of historical persons. This focused assessment of extending VIVO for a humanities purpose should be of interest to a segment of VIVO conference attendees and community members. We hope to

show that VIVO4HP provides an example that other humanists can build upon to represent, facilitate discovery of, and share linked data about historical persons.

VIVO CLOSING SESSION

4:30 PM – 5:00 PM
Room: Colorado Ballroom E – F