Sci² Tool: Download and Run
also distributed on Memory Stick

- These slides
- Sci2 Tool Manual v0.5.1 Alpha, updated to match v1.0 Alpha tool release
  http://sci2.wiki.cns.iu.edu
- Sci2 Tool v1.0 Alpha (June 13, 2012)
  http://sci2.cns.iu.edu

- Additional Datasets
  http://sci2.wiki.cns.iu.edu/2.5+Sample+Datasets
- Additional Plugins
  http://sci2.wiki.cns.iu.edu/3.2+Additional+Plugins

Or copy them from the DVD or memory stick.

Postscript Viewer: Please try opening ‘chessboard.ps.’ You should see
Content and Instructional Approach

VIVO is an excellent system for creating and managing faculty and researcher profiles, but it is also capable of much more. VIVO's ability to obtain and interlink data from a variety of high quality sources, including institutional systems of record and online databases, coupled with VIVO's use of open semantic web technologies, makes VIVO an exciting and powerful data platform. This platform makes VIVO itself capable of many new features, and also makes it possible for outside companies and researchers to develop their own VIVO applications and perform their own analyses using VIVO data.

This hands-on workshop aims to explore the possibilities of VIVO technology and data by way of the Cyberinfrastructure for Network Science Center's VIVO visualization work. It starts with an overview of VIVO's architecture, and explores how the architecture makes it possible to expand and build on VIVO. We then describe the Cyberinfrastructure for Network Science Center's visualization work, including an overview of the visualizations created for VIVO thus far. Taking a closer look at the VIVO visualizations, we explore their individual motivations and features, and how we use VIVO to make these visualizations possible. This is followed by a hands-on component where participants are guided through the creation of a simple visualization using live data from VIVO instance as well as workflows that use the Science of Science (Sci2) Tool (http://sci2.cns.iu.edu) and Gephi (http://gephi.org) to perform detailed analyses of VIVO data. The workshop concludes with a general question-answer session.

Target Audience and Learning Objectives

- Anyone interested in the visualization capabilities of VIVO, and the visualization work of the Cyberinfrastructure for Network Science Center.
- Software developers interested in an overview of the VIVO environment from a technical perspective, and a more in-depth exploration of VIVO through the lens of visualization development.
- Librarians and Science Administrators interested in gaining a deeper understanding of how VIVO works, exploring some of what it is already capable of, and understanding what it could be used for in the future.
- Companies that plan to offer value-added services for VIVO.
- Researchers that would like to utilize VIVO data in their scientific work.
• **Individual level.** Statistics and ego-centric scholarly networks on VIVO Profile pages.

• **Institutional level.** Analyses and visualizations of funding intake and publication output for departments and centers accessible via the VIVO Index page. Download of relevant data in tabular and network formats for further analysis using the Network Workbench tool.

• **National level.** Visualization of VIVO installations and their profile holdings together with web page access and general VIVO information requests. Plus, services that use VIVO URIs to access data across different VIVO instances.

---

**Workshop Attendees**

Registered by Aug. 13, 2012

<table>
<thead>
<tr>
<th>Participant_First_Name</th>
<th>Participant_Last_Name</th>
<th>Participant_Title</th>
<th>Participant_Organization</th>
<th>Participant_Email</th>
</tr>
</thead>
<tbody>
<tr>
<td>Radha</td>
<td>Allam</td>
<td>Managing Director</td>
<td>NETE</td>
<td><a href="mailto:nasim@nete.com">nasim@nete.com</a></td>
</tr>
<tr>
<td>Melissa</td>
<td>Benton</td>
<td>Program Analyst</td>
<td>EPA</td>
<td><a href="mailto:benton.melissa@epa.gov">benton.melissa@epa.gov</a></td>
</tr>
<tr>
<td>Terry</td>
<td>Bishop</td>
<td>Hematology Program Director</td>
<td>National Institute of Health</td>
<td><a href="mailto:bishop@extra.niddk.nih.gov">bishop@extra.niddk.nih.gov</a></td>
</tr>
<tr>
<td>Richard</td>
<td>Bookman</td>
<td>Sr. Adv for Igm Dev &amp; Sci Policy</td>
<td>UM</td>
<td><a href="mailto:rbookman@miami.edu">rbookman@miami.edu</a></td>
</tr>
<tr>
<td>Will</td>
<td>Coliante</td>
<td>Librarian</td>
<td>University of Florida</td>
<td><a href="mailto:wjcol3@ufl.edu">wjcol3@ufl.edu</a></td>
</tr>
<tr>
<td>Hillary</td>
<td>Corbett</td>
<td>Scholarly Communication Librarian</td>
<td>Northeastern University</td>
<td><a href="mailto:h.corbett@neu.edu">h.corbett@neu.edu</a></td>
</tr>
<tr>
<td>Bryan</td>
<td>Dennis</td>
<td>Consultant</td>
<td>American Psychological Association</td>
<td>bj <a href="mailto:dennis@gmail.com">dennis@gmail.com</a></td>
</tr>
<tr>
<td>Holly</td>
<td>Falk-Krajeski</td>
<td>Director</td>
<td>Northwestern University</td>
<td><a href="mailto:h-falk@northwestern.edu">h-falk@northwestern.edu</a></td>
</tr>
<tr>
<td>Alexander</td>
<td>Garcia-Castro</td>
<td>Assistant in Research</td>
<td>Florida State University</td>
<td><a href="mailto:alexgarcia@gmail.com">alexgarcia@gmail.com</a></td>
</tr>
<tr>
<td>Laura</td>
<td>Guazzelli</td>
<td>Web Developer</td>
<td>University of Florida</td>
<td><a href="mailto:laura2@ufl.edu">laura2@ufl.edu</a></td>
</tr>
<tr>
<td>Jeffrey</td>
<td>Horon</td>
<td>Consultant</td>
<td>Elsevier, Inc.</td>
<td><a href="mailto:j.horon@elsevier.com">j.horon@elsevier.com</a></td>
</tr>
<tr>
<td>Nicholas</td>
<td>Rejack</td>
<td>Ontologist &amp; Data Steward</td>
<td>University of Florida</td>
<td><a href="mailto:nrejack@ufl.edu">nrejack@ufl.edu</a></td>
</tr>
<tr>
<td>Arve</td>
<td>Sollard</td>
<td>Senior Programmer/Web Design</td>
<td>Griffith University</td>
<td><a href="mailto:a.sollard@griffith.edu.au">a.sollard@griffith.edu.au</a></td>
</tr>
<tr>
<td>Bradley</td>
<td>Taylor</td>
<td>Manager, Bioinformatics</td>
<td>Medical College of Wisconsin</td>
<td><a href="mailto:btaylor@mcm.edu">btaylor@mcm.edu</a></td>
</tr>
<tr>
<td>Michaelaen</td>
<td>Trimarchi</td>
<td>Senior Reference Librarian</td>
<td>The Scripps Research Institute</td>
<td><a href="mailto:mtrimarc@scripps.edu">mtrimarc@scripps.edu</a></td>
</tr>
</tbody>
</table>
Workshop Schedule

Hands-On (Katy Borner)
- Analysis & Visualization of VIVO Data using the Science of Science Tool (http://sci2.cns.iu.edu)

15 min break

Guided Tour (Chin Hua Kong)
- Customizing existing VIVO visualization

Guided Tour (Katy Borner)
- Analysis & Visualization of VIVO Data by others

Q&A

Please also consult Chapter 7 in the VIVO Book.

Analyzing and Visualizing VIVO Data

Chintan Tank, Indiana University and General Sentiment
Micah Linnemeier, Indiana University and University of Michigan
Chin Hua Kong and Katy Börner, Indiana University
Workshop Schedule

Hands-On (Katy Borner)
- Analysis & Visualization of VIVO Data using the Science of Science Tool ([http://sci2.cns.iu.edu](http://sci2.cns.iu.edu))

15 min break

Guided Tour (Chin Hua Kong)
- Customizing existing VIVO visualization

Guided Tour (Katy Borner)
- Analysis & Visualization of VIVO Data by others

Q&A

---

Type of Analysis vs. Level of Analysis

<table>
<thead>
<tr>
<th>Type of Analysis</th>
<th>Micro/Individual (1-100 records)</th>
<th>Meso/Local (101–10,000 records)</th>
<th>Macro/Global (10,000 &lt; records)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Statistical Analysis/Profiling</td>
<td>Individual person and their expertise profiles</td>
<td>Larger labs, centers, universities, research domains, or states</td>
<td>All of NSF, all of USA, all of science.</td>
</tr>
<tr>
<td>Temporal Analysis (When)</td>
<td>Funding portfolio of one individual</td>
<td>Mapping topic bursts in 20-years of PNAS</td>
<td>113 Years of physics Research</td>
</tr>
<tr>
<td>Geospatial Analysis (Where)</td>
<td>Career trajectory of one individual</td>
<td>Mapping a state’s intellectual landscape</td>
<td>PNAS publications</td>
</tr>
<tr>
<td>Topical Analysis (What)</td>
<td>Base knowledge from which one grant draws.</td>
<td>Knowledge flows in Chemistry research</td>
<td>VxOrd/Topic maps of NIH funding</td>
</tr>
<tr>
<td>Network Analysis (With Whom?)</td>
<td>NSF Co-PI network of one individual</td>
<td>Co-author network</td>
<td>NSF’s core competency</td>
</tr>
</tbody>
</table>
Temporal Analysis (When) Temporal visualizations of the number of papers/funding awarded at the institution, school, department, and people level
Geospatial Analysis (Where?) Where are what NRN instances and what data holdings do they have?
An interdisciplinary network

Enabling collaboration and discovery among scientists across all disciplines.

The network of scientists will facilitate scholarly discovery. Institutions will participate in the network by installing VIVO, or by providing semantic web-compliant data to the network.
Topical Analysis (What) Science map overlays will show where a person, department, or university publishes most in the world of science. (in work)
Network Analysis (With Whom?)  Who is co-authoring, co-investigating, co-inventing with whom? What teams are most productive in what projects?

Sci² Tool – “Open Code for S&T Assessment” to run replicable workflows


Sci² Tool

Sci² Tool: Documentation
Sci2 Tool: Download and Run
also distributed on Memory Stick

- These slides
- Sci2 Tool Manual v0.5.1 Alpha, updated to match v1.0 Alpha tool release
  http://sci2.wiki.cns.iu.edu
- Sci2 Tool v1.0 Alpha (June 13, 2012)
  http://sci2.cns.iu.edu

- Additional Datasets
  http://sci2.wiki.cns.iu.edu/2.5+Sample+Datasets
- Additional Plugins
  http://sci2.wiki.cns.iu.edu/3.2+Additional+Plugins

Or copy them from the DVD or memory stick.

Postscript Viewer/Converter: Please try opening ‘chessboard.ps.’
You should see the chessboard.
Adobe Distiller or http://ps2pdf.com work well.

Using Data from VIVO

General Statistics
- 36 publication(s) from 2001 to 2010 (.CSV File)
- 80 co-author(s) from 2001 to 2010 (.CSV File)

Co-Author Network
(GraphML File)

Save as Image (.PNG file)

Tables
- Publications per year (.CSV File)
- Co-authors (.CSV File)

http://vivo.iu.edu/vis/author-network/person25557
36 publication(s) from 2001 to 2010 (CSV File)

80 co-author(s) from 2001 to 2010 (CSV File)

Co-author network (GraphML File)

Save as Image (PNG file)

Publications per year (CSV File), see top file.

Co-authors (CSV File)

http://vivo.iu.edu/vis/author-network/person751182
Run Sci2 Tool. Use 'File > Load' to select 'conlon-michael_co-author-network.graphml'

Network Analysis Toolkit
Nodes: 177
Edges: 771

Visualize the file using 'Visualization>Networks>Radial Tree/Graph (prefuse alpha)'

Click on node to focus on it.
Hover over a node to highlight its co-authors.

Code and tutorials are linked from http://sci2.wiki.cns.edu
Network Extraction: Examples

Author co-occurrence network

Paper-author 2-mode network
Using the Sci2 Tool to analyze VIVO Data: Faculty-Courses Bipartite Networks

Use ‘File > Load’ to load cleaned SLIS-Faculty-Courses.csv

<table>
<thead>
<tr>
<th>A</th>
<th>B</th>
</tr>
</thead>
<tbody>
<tr>
<td>personName</td>
<td>courseTitle</td>
</tr>
<tr>
<td>Park, Taemin</td>
<td>SLIS L597 - Topics On Libr &amp; Info Science (Advanced Bibliographic Control)</td>
</tr>
<tr>
<td>Park, Taemin</td>
<td>SLIS S504 - Cataloging</td>
</tr>
<tr>
<td>Park, Taemin</td>
<td>SLIS L520 - Bibliographic Access &amp; Control</td>
</tr>
<tr>
<td>Park, Taemin</td>
<td>SLIS S631 - Advanced Cataloging</td>
</tr>
<tr>
<td>Milojevic, Stasa</td>
<td>SLIS S503 - Representation &amp; Organization</td>
</tr>
<tr>
<td>Milojevic, Stasa</td>
<td>SLIS L505 - Org &amp; Rep Of Knowledge &amp; Info</td>
</tr>
</tbody>
</table>

Total of 52 people teaching 319 courses.
Pivot table of #unique courses per faculty is on right.
No course is taught by more than 3 faculty:

<table>
<thead>
<tr>
<th>A</th>
<th>B</th>
</tr>
</thead>
<tbody>
<tr>
<td>SLIS G901 - Advanced Research</td>
<td>3</td>
</tr>
<tr>
<td>SLIS L524 - Information Sources &amp; Services</td>
<td>3</td>
</tr>
<tr>
<td>SLIS L528 - Collection Development &amp; Mgmt</td>
<td>3</td>
</tr>
<tr>
<td>SLIS L546 - User-Centered Database Design</td>
<td>3</td>
</tr>
<tr>
<td>SLIS L554 - Education Of Information Users</td>
<td>3</td>
</tr>
<tr>
<td>SLIS L571 - Info Architecture For The Web</td>
<td>3</td>
</tr>
<tr>
<td>SLIS L574 - Comm In Elec Info Environments</td>
<td>3</td>
</tr>
<tr>
<td>SLIS L579 - Information Visualization</td>
<td>3</td>
</tr>
<tr>
<td>SLIS L597 - Topics On Libr &amp; Info Science (Structural Datamining &amp; Modeling)</td>
<td>3</td>
</tr>
</tbody>
</table>

Run ‘Data Preparation > Extract Bipartite Network’
With parameter values:

Visualize resulting ‘Bipartite network from personName and courseTitle’ using ‘Visualization > Network > GUESS’ and ‘Layout > GEM’, ‘Layout > Bin Pack’
Using the Sci2 Tool to analyze VIVO Data: Faculty-Courses Bipartite Networks

Pan:
“grab” the background by holding left-click and moving your mouse.

Zoom:
Using scroll wheel, press the “+” and “-” buttons in the upper-left hand corner, or right-click and move the mouse left or right.
Center graph by selecting ‘View -> Center’.

Select to select/move single nodes. Hold down ‘Shift’ to select multiple.
Right click to modify Color, etc.

Using the Sci2 Tool to analyze VIVO Data: Faculty-Courses Bipartite Networks

Graph Modifier:
Select
Object: nodes based on
Property: bipartitetype
Operator : ==
Value: personName
then click ‘Show Label’ button and Colour to pick a color.

Select “Format Node Labels”, replace default text \{originallabel\} with your own label in the pop-up box ‘Enter a formatting string for node labels.’
Using the Sci2 Tool to analyze VIVO Data: Faculty-Courses Bimodal Networks

Graph Modifier:

IF node Degree was calculated in Sci2, select ‘Resize Linear > Nodes > totaldegree’ in drop-down menu, then type “5” and “20” into the From” and To” Value box separately. Then select ‘Do Resize Linear’.

Select ‘Colorize> Nodes>totaldegree’, then select white and enter (204,0,51) in the pop-up color boxes on in the “From” and “To” buttons.

Using the Sci2 Tool to analyze VIVO Data: Faculty-Courses Bipartite Networks

Graph Modifier:

IF node Degree was calculated in Sci2, select ‘Resize Linear > Nodes > totaldegree’ in drop-down menu, then type “5” and “20” into the From” and To” Value box separately. Then select ‘Do Resize Linear’.

Select ‘Colorize> Nodes>totaldegree’, then select white and enter (204,0,51) in the pop-up color boxes on in the “From” and “To” buttons.
Using the Sci2 Tool to analyze VIVO Data:
Faculty-Courses Bipartite Networks

Interpreter:
Uses Jython a combination of Java and Python.
Try
colorize(totaldegree, white, red)

Using the Sci2 Tool to analyze VIVO Data:
Co-Author networks

Use ‘File > Load’ to load cleaned YingDing-CoAuthors.csv or KatyBorner-Co-Authors.csv

Run ‘Data Preparation > Extract Co-Occurrence Network’
With parameter values:

Visualize resulting ‘Bipartite network from personName and courseTitle’ using ‘Visualization > Network > GUESS’ and ‘Layout > GEM’, ‘Layout > Bin Pack’
Using the Sci2 Tool to analyze VIVO Data:
Co-Author networks
Using the Sci2 Tool and Gephi to analyze VIVO Data: Co-Author networks

**Using the Sci2 Tool to analyze VIVO Data:**

**Paper to Co-Author Bipartite Networks**

Use ‘File > Load’ to load cleaned *KatyBorner-Co-Authors.csv*

<table>
<thead>
<tr>
<th>A</th>
<th>B</th>
</tr>
</thead>
<tbody>
<tr>
<td>PublicationTitle</td>
<td>Authors</td>
</tr>
<tr>
<td>113 Years of Physical Review: U Borner, Katy:Penunmarthy, Shashikant:Herr Jr, Bruce W.;Hardy, Elisha F.;Duhon, Russell J.</td>
<td>KatyHolloway, Todd:Bozickevic, Miran</td>
</tr>
<tr>
<td>Analyzing and Visualizing the S Borner, Katy: :Holloway, Todd:Bozickevic, Miran</td>
<td></td>
</tr>
</tbody>
</table>

Run ‘Data Preparation > Extract Bipartite Network’

With parameter values:

Extract Bipartite Network

- **First column:** PublicationTitle
- **Second column:** Authors
- **Text Delimiter:**

Visualize resulting bipartite network using ‘Visualization > Network > GUESS’

and ‘Layout > GEM’, ‘Layout > Bin Pack’
Using the Sci2 Tool to analyze VIVO Data:
Geoffrey Fox Funding Data

Use ‘File > Load’ to load cleaned GeoffreyFox-Funding.csv

<table>
<thead>
<tr>
<th>Title</th>
<th>Start</th>
<th>End</th>
<th>Amount</th>
</tr>
</thead>
<tbody>
<tr>
<td>XROADS: Planning the future eXtreme Digital Cyb</td>
<td>5/1/2009</td>
<td>7/31/2010</td>
<td>154949</td>
</tr>
<tr>
<td>ERDC PET Support (Computer Science Corp.)</td>
<td>7/1/2001</td>
<td>9/30/2001</td>
<td>41574</td>
</tr>
<tr>
<td>Earthquake Data Enhanced Cyber-Infrastructure</td>
<td>10/12/2009</td>
<td>9/30/2010</td>
<td>45000</td>
</tr>
</tbody>
</table>

Run ‘Visualization > Temporal > Temporal Bar Graph’ with parameter values:
Using the Sci2 Tool to analyze VIVO Data:
Geoffrey Fox Funding Data

Result will be rendered into a Postscript file (red icon)

Save as ps file, convert into pdf and view:
Temporal Visualizations

Temporal Visualization
Generated from NSF csv file: C:|UniversityDesktop|TOOLs|Sec-2012-05-04-N|NAW|sampled|as|scientometrics|@|indiana.net
June 05, 2012 | 4:50 PM EDT

Legend
Area in Heat of Number
Minimum = 200, 500
Maximum = 25,000,000
Text label: Title
Color: NSF Organization
See end of PDF for color legend.

How To Read This Map
This temporal bar graph visualization represents each record as a horizontal bar with a specific start and end date and a text label on its left side. The area of each bar encodes a numerical attribute value, e.g., total amount of funding. Bars may be colored to present categorical attribute values of records.

Geospatial Visualizations

Geospatial Visualization (Choropleth Map)
Generated from CSV file: Preprocessed-southflorida-0316730090-13743704.csv
Jun 05, 2012 | 05:45:50 PM EDT

Legend
Country Color (Linear)
Times Cited

How to Read this Map
This choropleth map shows 226 countries of the world using the equal-area Eckert IV projection. Each country may be color coded in proportion to a numerical value. Minimum and maximum data values are given in this legend.
Geospatial Visualizations

---

**Topical Visualization**

**Data:** WoS and Scopus paper level data for 2001–2010, about 25,000 separate journals, proceedings, and series.

**Similarity Metric:** Combination of bibliographic coupling and keyword vectors.

**Number of Disciplines:** 554 journal clusters further aggregated into 13 main scientific disciplines that are labeled and color coded in a metaphorical way, e.g., Medicine is blood red and Earth Sciences are brown as soil.
Workshop Schedule

Hands-On (Katy Borner)
• Analysis & Visualization of VIVO Data using the Science of Science Tool (http://sci2.cns.iu.edu)

15 min break

Guided Tour (Chin Hua Kong)
• Walk through Co-author Network visualization

Guided Tour (Katy Borner)
• Analysis & Visualization of VIVO Data by others

Q&A

Workshop Review

• 2011 VIVO visualization workshop
  ▪ http://wiki.cns.iu.edu/display/PRES/VIVO+Presentation
  ▪ Visualizing organization hierarchy data
  ▪ Creating a new VIVO visualization

• 2012 VIVO visualization workshop
  ▪ http://wiki.cns.iu.edu/display/PRES/VIVO+Presentation
  ▪ Learn from 2011 visualization workshop and new requests from vivo development mailing list
  ▪ Visualizing data downloaded from VIVO visualization page
  ▪ Visualizing publication and teaching data from VIVO
  ▪ Learn how to customize the ego-centric network visualization (used for co-author network and co-investigator network)
VIVO Visualization

- Accessing VIVO Data
- Architecture
- Pipeline Explanation

Accessing VIVO Data

- Read “The Semantic Web: An Introduction”
  http://infomesh.net/2001/swintro

author: Chris Bizer
taken from: http://linked databook.com/editions/1.0/images/lod-cloud_2010.png
Linked Open Data (via RDF or N3)

- Accessible to anyone on the Web.
- It can be a bit tedious to work with large amounts of data quickly/easily.
- N3 example:
  - http://vivo.iu.edu/individual/person25557/person25557.n3
- RDF example:
  - http://vivo.iu.edu/individual/person25557/person25557.rdf
- VIVO Ontology

SPARQL Endpoints

- Working with data is easier/faster (using SPARQL queries).
- But may not be accessible to everyone.
VIVO Visualization

• Accessing VIVO Data
• Architecture
• Pipeline Explanation

VIVO Visualization Architecture

1. User requests the visualization
2. Request is received by the VIVO application
3. Specific controller gets control
4. Controller delegates the control of flow to the handler of the requested visualization, if permission validated
5. The handler passes request information to the Query Handler
6. Query Handler queries the semantic web data store (cached or live)
7. Results of the query are converted into Java objects
8. Java objects are used to generate response in the requested format
9. Request handler renders the generated response
VIVO Visualization

- Accessing VIVO Data
- Architecture
- Pipeline Explanation

Map of Science Visualization Pipeline

- Breakdown of serving the *map of science* visualization request received at, http://vivo.iu.edu/vis/map-of-science/IndianaUniversity

- Short URL Request
- It has following parameters,
  - /vis – Short URL prefix
  - /map-of-science – Visualization type
  - /IndianaUniversity – URI of subject of the visualization

- Long form looks like
  http://vivo.iu.edu/visualization
    ?vis=map-of-science
    &uri=http://vivo.iu.edu/individual/IndianaUniversity
Map of Science Visualization Pipeline

- /vis – Short URL Visualization Controller assumes the control of flow
- Parse URL to get visualization type – Map Of Science
- Gets permission requirements for Map of Science visualization
- Validates it against the requesting user’s privileges

- After validation captured parameters are passed to the Map of Science Visualization
- Check cache for all models pertaining to “IndianaUniversity” entity
- If not present create models and store in cache
Map of Science Visualization Pipeline

- Models used are,
  - ORGANIZATION_MODEL_WITH_TYPES
    - no input uri
    - all sub-organizations recursively
  - ORGANIZATION_TO_PUBLICATIONS_FOR_SUBORGANIZATIONS
    - specific for an input uri
    - all publications for entity-associated people
  - ORGANIZATION_ASSOCIATED_PEOPLE_MODEL_WITH_TYPES
    - specific for an input uri
    - e.g. President of University
  - PEOPLE_TO_PUBLICATIONS
    - no input uri
    - all people associated with publication

- SPARQL query fired against the previous mentioned cache
- Using the query results create the java objects
  - Entity (for subject entity)
  - SubEntity (for child entities)
  - Activity (for publication info)
Map of Science Visualization Pipeline

- Java objects used to generate output
- FreeMarker object
  - HTML markup including CSS, JavaScript
- JSON
  - Used by JavaScript to render tables, map etc

Map of Science Visualization Pipeline

- Output from visualization (FreeMarker + JSON) sent to Response Handler
- Makes sure output formatted properly
  - E.g. For FreeMarker – compile HTML markup
Map of Science Visualization Pipeline

Workshop Schedule

Hands-On (Katy Borner)
- Analysis & Visualization of VIVO Data using the Science of Science Tool (http://sci2.cns.iu.edu)

Guided Tour (Chin Hua Kong)
- Walk through Co-author Network visualization

Guided Tour (Katy Borner)
- Analysis & Visualization of VIVO Data by others

Q&A
Walk through Co-author Network visualization


- Pre-implementation
  - Identify the data to be represented
  - Research Visualization to present the data
  - Visualization technologies

- Front-End module
  - Create Front-End module that renders the data created by the back-end

- Back-End module
  - Create SPARQL Queries to get the data
  - Edit Back-End module for new data

- Dependency Injection

Identify the data to be represented

- What data is available?

- Who is the target audience?

- What story do you want to tell?

- Why is the story important for your users?

- Integrity of the data.
Research visualization to present the data

- Existing visualizations
- Mockup + Sampling with Sci2
- Frontend technology
- Target users environment
- Is the visualization easy to understand?

Visualization Technologies

- Front-End
  - HTML
  - FLASH
  - JSON
  - CSV
  - GRAPHML
  - JavaScript

- Back-End
  - Java
  - FreeMarker
  - SPARQL
  - Ontology
Mock-up of the desired goal

Pre-implementation result
Front-End Module

- Co-author network related modules
  - `/coauthorship` – custom for Co-author Sparkline widget
  - `/publication` – generate Publication Sparkline widget
  - `/personlevel` – Main HTML page shared between co-author network and co-investigator network

- Create coauthorship folders in `images`, `css`, `js`, and `templates` folders
  - Copy images, js, and css libraries into newly created `coauthorship` folders

- Customize the HTML contents into freemarker `.ftl` files
  - Main HTML page content:
    `/productMods/templates/freemarker/visualization/personlevel/coAuthorPersonLevel.ftl`
  - Co-author Sparkline content:
    `/productMods/templates/freemarker/visualization/coauthorship/*`
  - Publication Sparkline content:
    `/productMods/templates/freemarker/visualization/publication/*`
Back-End Module


---

**Back-End module**

- Create SPARQL Queries to get co-author network information for a person

```sparql
PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
PREFIX core: <http://vivoweb.org/ontology/core#>

SELECT DISTINCT ?name ?coAuthorLabel
WHERE
{
    ?coAuthorship core:linkedAuthor ?coAuthor .
    ?coAuthor rdfs:label ?coAuthorLabel.
    FILTER (?http://vivo.iu.edu/individual/person751182> != ?coAuthor)
}
```

---

```
<table>
<thead>
<tr>
<th>name</th>
<th>coAuthorLabel</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ding, Ying</td>
<td><em>ben-05</em></td>
</tr>
<tr>
<td>Ding, Ying</td>
<td><em>ben-08</em></td>
</tr>
<tr>
<td>Ding, Ying</td>
<td><em>ben-05</em></td>
</tr>
<tr>
<td>Ding, Ying</td>
<td><em>ben-05</em></td>
</tr>
<tr>
<td>Ding, Ying</td>
<td><em>ben-08</em></td>
</tr>
<tr>
<td>Ding, Ying</td>
<td><em>ben-05</em></td>
</tr>
<tr>
<td>Ding, Ying</td>
<td><em>ben-05</em></td>
</tr>
<tr>
<td>Ding, Ying</td>
<td><em>ben-08</em></td>
</tr>
<tr>
<td>Ding, Ying</td>
<td><em>ben-05</em></td>
</tr>
<tr>
<td>Ting, Ying</td>
<td><em>ben-05</em></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>
```
Back-End Module

- Create package for “coauthorship” along with other visualizations
- Query Runner
  - SPARQL Query
    edu.cornell.mannlib.vitro.webapp.visualization.coauthorship.CoAuthorshipQueryRunner
  - Java Objects:
    edu.cornell.mannlib.vitro.webapp.visualization.collaborationutils.CoAuthorshipData
- Visualization Request Handler
  edu.cornell.mannlib.vitro.webapp.controller.visualization.PersonLevelRequestHandler
  edu.cornell.mannlib.vitro.webapp.visualization.coauthorship.CoAuthorshipRequestHandler
- Get input from controller: PersonLevelRequestHandler
- Requests Query Runner to get data (Model to be cached?)
  - See jena model tutorial at http://jena.sourceforge.net/tutorial/
- Prepares data
  - FreeMarker config: PersonLevelRequestHandler
  - JSON output: CoAuthorshipRequestHandler
  - Graphml output (Specified for ego-centric): CoAuthorshipRequestHandler
Wire pieces to get a complete pipeline

RequestHandler injection

- File: productMods/WEB-INF/visualization/visualizations-bean-injection-fm.xml
- Register request handler with ID
  ```xml
  <bean id="person_level"
       class="edu.cornell.mannlib.vitro.webapp.visualization.personlevel.PersonLevelRequestHandler" />
  ```
- Define request handler for vis mode
  ```xml
  <entry key="author-network">
      <ref bean="person_level"></ref>
  </entry>
  ```
End Product

http://vivo.iu.edu/vis/author-network/person751182

Workshop Schedule

Hands-On (Katy Borner)
- Analysis & Visualization of VIVO Data using the Science of Science Tool (http://sci2.cns.iu.edu)

Guided Tour (Chin Hua Kong)
- Walk through Co-author Network visualization

Q&A
2008 collaboration patterns for medical institutions located close to Melbourne University
Source: Web of Science co-authorship information. Compiled by Simon Porter

Co-authorship network for the department of Information Systems
Source: Melbourne Research Windows. Contact Simon Porter simon.porter@unimelb.edu.au
Bimodal network of search terms and researchers extracted from research profile search results to show the **University’s capability in Disaster Management** to the Government

**Contact:** simon.porter@unimelb.edu.au

---

**Top MeSH Disease Concepts Appearing in PubMed Publications by the University of Michigan Medical School.** Links connect concepts where 100+ authors published about both concepts within the span of their careers.

**Contact:** Jeffrey Horon, J.Horon@elsevier.com
Top MeSH Disease Concepts Appearing in PubMed Publications by the University of Michigan Medical School. Links connect concepts where 100+ authors published about both concepts within the span of their careers. This visualization revealed that animal disease models were central to disease research at U-M which encouraged additional thought and attention to animal husbandry, animal expenses, and core/shared services overall.

Contact: Jeffrey Horon, J.Horon@elsevier.com

P30 Member Collaborations – Sponsored Project Co-Participation and Co-Authorship Network. Used in successful P30 funding application. Shows the PI’s relationships with various P30 members, conveying that the PI was not only the formal center of the group but also the informal center and the person who exhibited the highest betweenness centrality. Contact: Jeffrey Horon, J.Horon@elsevier.com
Inter-Institutional Collaboration Explorer
This visualization shows information about “collaborative publications” found at 2 or more Researcher Networking websites.

The idea that institutions don't work together and that biomedical research is conducted in silos is not true. Researchers, even when separated by great distances, are in fact willing to work together, and this visualization demonstrates that they often do.

Contact: Nick Benik (nbenik@gmail.com), Harvard Medical School, Boston, MA.
URL: http://xcite.hackerceo.org/VIVOviz

Inter-Institutional Collaboration Explorer
The outer solid colored arcs represent the 11 institutions. The size of the arc is proportional to the number of collaborative publications found on the site. The inner colored bands represent the number of collaborative publications found between the two institutions that each band connects. Clicking an institution's arc will hide any bands not connected to that institution and will display a timeline of when that institution's collaborative publications were written.
Workshop Schedule

Hands-On (Katy Borner)
• Analysis & Visualization of VIVO Data using the Science of Science Tool ([http://sci2.cns.iu.edu](http://sci2.cns.iu.edu))

Guided Tour (Chin Hua Kong)
• Customizing existing VIVO visualization

Guided Tour (Katy Borner)
• Analysis & Visualization of VIVO Data by others

Q&A