Collexis and VIVO

Paving the Collexis Pathway to the National Research Network

Brian Athey, Ph.D.
  Teri Grieb
  Mary Hill
  Michael Warden

August 13, 2010
Who We Are

Brian Athey, Ph.D. (bleu@umich.edu)
Professor, Biomedical Informatics, Psychiatry and Internal Medicine
Chair Designate, Computational Medicine and Bioinformatics
Associate Director, Michigan Institute for Clinical and Health Research (MICHRI)
Director, Academic Informatics
University of Michigan Medical School

Teri Grieb, Ph.D. (tgriev@umich.edu)
Sr. Director of Research; Managing Director, MICHRI
University of Michigan

Mary Hill (maryhill@umich.edu)
Manager, Research Information Services
University of Michigan Medical School

Michael Warden (warden@collexis.com)
Collexis Solutions Sales Manager
Elsevier, Inc.
Agenda

• Research Networking at Michigan
  • What is Collexis?
  • Collexis Implementation at Michigan
  • How it is used? Relation to CTSA.
  • Data Extension

• Linking Collexis to VIVO
  • Implementation
  • How it will be used
  • Data Extension
Elephants in the Room—to discuss

- Interoperability of Research and Social Networking Systems—especially commercial and open source systems
  - A theme of this talk
- IT Security and Privacy when using Research Networking
- Linking Resource Discovery with individuals and teams
- Data Sharing enhancement possibilities
- "Maureen Dowd Effect"—enabling chance encounters
Preface on Research Networking

The CTSA consortium facilitates researcher networking across national institutions and across topic domains, creating a virtual community where collaborations across institutions can arise easily, where expertise is mapped and can be located easily, and where matchmaking between collaborators or others, can expand beyond an institution’s walls as desired. (from 08/09 CTSA Key Function Committee [KFC])

• Professional connections
• Based on academic output and expertise mapping (bibliome)
• Connections on detailed conceptual level
• Built from passive connections
• Leverage ‘what we already know’
• **Who** knows **what**?
• Links to Resources is future goal (from CTSA Strategic Goal #3; e.g. eagle-i)
Agenda

• **Research Networking at Michigan**
  - What is Collexis?
  - Collexis Implementation at Michigan
  - How it is used?
  - Data Extension

• **Linking Collexis to VIVO**
  - Implementation
  - How it will be used?
  - Data Extension
Research Networking at Michigan

Needs Assessment and Realization (Researchers):
- Locate Collaborators and resources (locally & beyond)

Needs (Administrative):
- Better understanding of faculty expertise
- Non-financial view of faculty
- Research interests
- Central IT but no data population from faculty
- Portfolio Analysis (Future)

Considerations:
- Build a CV system? A research interests database?
  - Make vs. Buy
- Need an immediate solution, at scale, able to be implemented quickly at a justifiable price
Agenda

• Research Networking at Michigan
  • **What is Collexis?**
  • Collexis Implementation at Michigan
  • How it is used
  • Data Extension

• Linking Collexis to VIVO
  • Implementation
  • How it will be used
  • Data Extension
Some Collexis Customers
What is the Technical Crux?

Two key areas of expertise

1. “Fingerprinting” of unstructured text
   - Software mines text and performs detailed analyses and classification of key terms and concepts in entire document collections
   - Given any textual data and an ontology will assign a “fingerprint” (or classifier) – individual or group
   - Applied to PubMed document repository using MeSH to create BiomedExperts.com

2. Disambiguation of Authors
   - Process to match the correct publications with one unique author
   - Combination of automatic algorithm with recursive manual clean-up (i.e. semi-automated, adaptive)
Why Collexis?

• Needed to make not buy

• Desire to better understand our faculty expertise and interdisciplinary potentials (both administrative and researcher need).

• Administration wants a non-financial view of faculty

• Administrative advantages for UMichigan:
  1. Little or no faculty manual input required
  2. UM provides faculty listing
  3. Disambiguation tools provided to automatically associate new publications
  4. Web services allow downloads of data to warehouse
Profiling (using PubMed and NIH RePORTER)
Profiling (Scopus data)

Find the Expert

6 Experts found

<table>
<thead>
<tr>
<th>Name</th>
<th>Publications</th>
</tr>
</thead>
<tbody>
<tr>
<td>Peretti, Steven W.</td>
<td>34</td>
</tr>
<tr>
<td>Bullard, Lisa G</td>
<td>19</td>
</tr>
<tr>
<td>Fornara, Robert Joseph</td>
<td>10</td>
</tr>
<tr>
<td>Joines, Jeffrey Allen</td>
<td>35</td>
</tr>
<tr>
<td>Rouskas, Georgios</td>
<td>37</td>
</tr>
<tr>
<td>Wiebe, Eric N</td>
<td>42</td>
</tr>
</tbody>
</table>

Your search terms
- Multidisciplinary teams

Refine search by adding Concepts

Activities & Behaviors
- Communication

Engineering
- Assessment processes
- Communication processes
- Computing tasks
- Delphi methods
- Generic computers
- Strategic decision makings
- Writing across the curriculums
- Cognitive skills
- Computing educations
- Integral parts
- Preliminary assessments
Agenda

- Research Networking at Michigan
  - What is Collexis?
  - Collexis Implementation at Michigan
  - How it is used?
  - Data Extension

- Linking Collexis to VIVO
  - Implementation
  - How it will be used?
  - Data Extension
Implementation at Michigan

• Scope:
  – Current: Medical School+ (2,260 faculty)
  – Future: All interested schools & colleges (We have 19, starting with 5)

• Timeline:
  – Contract June of 2008
  – Manual verification from fall 2008 to spring 2009
  – Full implementation, Phase I, June 2009
Manual Verification

• Collection of 2,144 CVs in November 2008

• Validation of data using CVs completed March 2009
  - Non-QA: Initial accuracy 60-70%
  - After QA: Accuracy >95%

• Cleanup process defined

• Production release in June 2009

• Well-received by faculty and administration
Where We Are At This Moment

• Version 3.5 released March 2010
  - Research network and trends
  - H-Index by faculty
  - Sort by citation counts
  - External collaborators link
  - Improved Google analytics
  - Improved Google search placement

• Upgrade to Scopus data currently underway, expansion from Medical School to non-life sciences

• Faculty refresh quarterly

• Growing number of data requests
Agenda

• Research Networking at Michigan
  • What is Collexis?
  • Collexis Implementation at Michigan
  • How it is used?
  • Data Extension

• From Collexis to VIVO
  • Implementation
  • How it will be used?
  • Data Extension
Administrative Mission

• Publication counts and data

• Who is publishing where?

• School/department/faculty areas of expertise
  (Identification of recruitment and retention issues)

• H-Index and other metrics by faculty

• Identification of faculty working on areas of strategic interest
Enhancing the Research Mission

- Determine global collaborations – where can we start new collaborations?
- Target faculty for funding requests – ARRA
- Identify differences in publications versus grant expertise
- Verify “stories”
- Where should we invest?
- Who are our top collaborators?
Department Use

Profiles
• Areas of expertise
• Link to profiles from their web sites
• Who are my experts?
• Top journals
• Who should we hire?

Departmental Use – Data
• Publication listings for marketing materials
• Who is publishing? / Who is not?
• Impact factors
• Participation in Team Science
Faculty

Research Use
• Look at my areas of interest
• Find and investigate potential collaborators
• Let sponsors find me…
• Discover new relationships and new areas of study

Administrative Use
• Track my citations
• Track my H-Index
• Feed publications to my CV
• Identify contact information
Trainees—Students, Fellows, Residents

- Find a mentor
- Research the history and connections of a mentor discretely
- Find areas of interest
- Explore research areas and key individuals
- Find friends and colleagues (locally and nationally)
Agenda

- Research Networking at Michigan
  - What is Collexis?
  - Collexis Implementation at Michigan
  - How it is used?
  - Data Extension

- Linking Collexis to VIVO
  - Implementation
  - How it will be used?
  - Data Extension
Data from Collextis

• Map Collextis ID to University of Michigan Employee ID

• Use Web Services to access data driving Collextis

• Output:
  • Concepts by faculty w/ranking
  • IDs of publications by faculty (PMID)
  • First/Co/Last author of publication indicator
  • Coauthor list and counts
Publication Counts

PubMed Publications from Current University of Michigan Medical School Faculty

**Publication Year**

- 2004
- 2005
- 2006
- 2007
- 2008*
- 2009 YTD*

**Count of Unique PubMed Publications**

- 0
- 500
- 1000
- 1500
- 2000
- 2500
- 3000
- 3500
- 4000

* Data Captured August 2009

[Long Term Trend]
Global Impact

What countries are we collaborating with?

• Mapped “location” to country

• Limitation: Location is for first author only so identifies where UM is co-author.

• Where are we collaborating on specific diseases such as – Lung Neoplasms

Concepts->Faculty->Publication->MeSH Terms
World Heat Map
Facility Expansion

Expansion into vacated Pfizer Global Research HQ Space (177 Acres; 30+ Buildings, 2.3M sq ft) -- called North Campus Research Complex)

- Blur departmental boundaries (IDR)

- Determine which faculty could create programs around their expertise and collaborations

- Looked at who is publishing in top journals
Significant journals

Faculty Publishing in *Nature* (1970 - Present)

**Year / Author / Title / PubMed ID**

2009 Chinnaiyan, Arul M Transcriptome sequencing to detect gene fusions in cancer. 19136943

2009 Chinnaiyan, Arul M Metabolomic profiles delineate potential role for sarcosine in prostate cancer progression. 19212411

2009 Laxman, Bharathi Metabolomic profiles delineate potential role for sarcosine in prostate cancer progression. 19212411

.....
Centers and Departments

• How academically effective are our centers? *Keep historical downloads for comparison*

• Looking at concept relationship--are there new centers that might be useful?

• Are collaborations occurring in and outside of our department, school, state, region, country?

• P30 Application support
Visualizations - Top 100 Diseases

Semi-Quantitative Observation: Centers increase collaboration opportunities.
For areas that bridge clinical and basic science concepts, how are we as an organization facilitating those exchanges? How does this influence building design?
Welcome! You have successfully installed VIVO!

Your next step is to log in and change your password. You can find the login link in the header on the right above.

Please see the VIVO User Guide for help. The user guide includes information on Site Administration. You may want to read about the Site Configuration topics next:

- Edit site information
- Manage tabs
- Manage user accounts
VIVO for Michigan

63,631 Articles
12,209 Grants
VIVO for Michigan

Center

Medical School Basic Sciences Center | on ResearchProfiles.Collexis.com
Medical School Clinical Sciences Center | on ResearchProfiles.Collexis.com
Med School Admin Core Roll-up Center | on ResearchProfiles.Collexis.com
MSA Oversight Units Roll-up Center | on ResearchProfiles.Collexis.com
Other Med School Departments Center | on ResearchProfiles.Collexis.com
Alan Tait

on ResearchProfiles.Collexis.com

positions

Anesthesiology Department

principal investigator on

ANESTHESIA AND UPPER RESPIRATORY TRACT INFECTIONS
INFORMED CONSENT FOR STUDIES INVOLVING SURGICAL PATIENTS
Strategies to Optimize Communication of Informed Consent for Pediatric Research

selected publications

Can we improve the assessment of discharge readiness?: A comparative study of observation
Plasma levels and cardiovascular effect of nitroglycerin in pregnant sheep.
The role of neutrophils, oxidants, and proteases in the pathogenesis of acid pulmonary injury.
Risk factors for perioperative adverse respiratory events in children with upper respiratory tract
Children who refuse anesthesia or sedation: a survey of anesthesiologists.

108 more
Search Results for 'plasma'

Show only results of this type: people, publications

Plasma-calcitonin in man.
... title Plasma-calcitonin in man. year and month 196903 volume 1 start page 443-6 abstract authors ...

Distribution of inorganic sulfate between plasma, liver, and bile.
... title Distribution of inorganic sulfate between plasma, liver, and bile. year and month 197408 volume 15 start page 741-9 abstract authors ...

Inverse relationship between plasma vasopressin and intracranial pressure.
... title Inverse relationship between plasma vasopressin and intracranial pressure. year and month 199203 volume 24 start page 141 abstract authors ...

The respiratory tract in amyloidosis and the plasma cell dyscrasias.
... title The respiratory tract in amyloidosis and the plasma cell dyscrasias. year and month 198604 volume 21 start page 113-27 abstract authors ...

The Plasma Membrane-Granule interface in Exocytosis
... start date 1998-12-15 title The Plasma Membrane-Granule Interface in Exocytosis end date 2011-02-28 principal investigator Ron ...

Exploring the human plasma proteome.
... title Exploring the human plasma proteome. year and month 200508 volume 5 start page 3223, 3225 abstract authors ...

Plasma 3,4-dihydroxyphenylalanine (dopa) and catecholamines in neuroblastoma or pheochromocytoma.
... title Plasma 3,4-dihydroxyphenylalanine (dopa) and catecholamines in neuroblastoma or pheochromocytoma. year and month 199103 volume 16 start page 83-97 abstract authors ...

The Plasma Membrane-Granule Interface in Exocytosis
... start date 1998-12-15 title The Plasma Membrane-Granule Interface in Exocytosis end date 2011-02-28 ...

The plasma cofactor and anticaldilipin antibodies.
... title The plasma cofactor and anticaldilipin antibodies. year and month 199011 volume 8 start page 613-5 abstract authors ...

The Plasma Membrane-Granule Interface in Exocytosis
... start date 1998-12-15 title The Plasma Membrane-Granule Interface in Exocytosis end date 2011-02-28 ...

Reversal of cyclosporine-associated hemolytic-uremic syndrome by plasma exchange with fresh-frozen plasma replacement in renal ...
... title Reversal of cyclosporine-associated hemolytic-uremic syndrome by plasma exchange with fresh-frozen plasma replacement in renal ...

Plasma measures of B-endorphin-like immunoreactivity in depressives and other psychiatric subjects.
... title Plasma measures of B-endorphin-like immunoreactivity in depressives and other psychiatric subjects. year and month 198211 volume 12 start page 851-2 abstract authors ...

Mechanism of plasma catecholamine increases during surgical stress in man.
... title Mechanism of plasma catecholamine increases during surgical stress in man. year and month 197711 volume 45 start page 115-21 abstract authors ...

Blood and plasma substitutes--plasma expansion and oxygen transport properties.
... title Blood and plasma substitutes--plasma expansion and oxygen transport properties. year and month 199807 volume 19 start page 314-9 abstract authors ...
Two distinct groups in this department – Spin out Health Service Research Unit?
Limitations

- Standard limitations of any data-driven tool
- Limited by source:
  - Only as clean as data that has been validated
  - Currently includes only PubMed information (adding Scopus)
- Limited by representation of data:
  - Limited by effectiveness of algorithm to associate terms with MeSH taxonomy (addressed through new Scopus implementation)
  - Weighting and visualizations limited by validation of data as well as data limitations
- Limited by novelty/newness:
  - Ongoing exploration of information - a completely new way to manage information and research
  - Time issues
  - Concepts vs faculty vs publications vs grants—not really quite the space of inquiry and analysis.
Agenda

• Research Networking at Michigan
  • What is Collexis?
  • Collexis Implementation at Michigan
  • How it is used
  • Data Extension

• Linking Collexis to VIVO
  • Why VIVO?
  • Implementation
  • How it will be used?
  • Data Extension
Why VIVO?

Collexis is used for...

- Research Networking inside Michigan and with other Collexis user sites/communities
- Data source for internal analysis of research and feeding other systems with publication and expertise data

... but there is further interest

- Connect with other institutions not using Collexis
- Connect other data – events, coursework, trainees, etc
- Participate with other major research networking initiatives
- Connect national (and international) initiatives with research networking
Considerations

In participating with VIVO: Michigan Centric View

- Must connect to, and help to, leverage existing research networking solutions at Michigan
- Must be complementary to other initiatives and systems
- Must allow Michigan to be part of the ‘conversation’ nationally
- Must take advantage of efficiencies established with Collexis
- Must continue to meet original focus of Collexis project:
  - Little or no work impact on faculty for unnecessary data entry
  - Disambiguation of data and automatic updates
  - Non-financial view of faculty
  - Connection to data to extend to other applications
Extension of Collexis features

Collexis Community

With the latest release of Collexis, we are tying together the community of organizations that are using Collexis Research Profiles at their institution(s). In order to make each implementation more powerful, we have connected them all together through the Collexis Community. When searching throughout the application you can switch from just the internal experts and data or from the entire community at any time. Please let us know what you think at community@collexis.com.

Search the Community

Browse the Community

- Albert Einstein College of Medicine
- Johns Hopkins University
- University of Maryland
### Find the Expert

Use the features on this search results page to adjust the source of the information and/or the search terms. Selecting Community vs. Internal allows you to choose results from just this institution or from the full Collexis community, and adjusting "Experts based on..." allows you to choose from publications or other data sources. The additional concepts allow you to narrow and expand your search to explore these results.

<table>
<thead>
<tr>
<th>34 Experts found</th>
<th>Experts based on...</th>
</tr>
</thead>
</table>

**Community**

**34 Experts found**

<table>
<thead>
<tr>
<th>Name</th>
<th>Institution</th>
<th>Publications</th>
<th>Grants</th>
</tr>
</thead>
<tbody>
<tr>
<td>Walkley, Steven U.</td>
<td>Albert Einstein College of Medicine</td>
<td>79</td>
<td>7</td>
</tr>
<tr>
<td>Bellon, Douglas</td>
<td>Memorial Sloan-Kettering Cancer Center</td>
<td>59</td>
<td>4</td>
</tr>
<tr>
<td>Ropin, Isabelle</td>
<td>Albert Einstein College of Medicine</td>
<td>160</td>
<td>4</td>
</tr>
<tr>
<td>Chugani, Harry</td>
<td>Wayne State University</td>
<td>219</td>
<td>3</td>
</tr>
<tr>
<td>Cowell, John</td>
<td>Medical College of Georgia</td>
<td>221</td>
<td>8</td>
</tr>
<tr>
<td>Naidu, Sambhav</td>
<td>Johns Hopkins University</td>
<td>122</td>
<td>0</td>
</tr>
<tr>
<td>Green, William</td>
<td>Johns Hopkins University</td>
<td>673</td>
<td>0</td>
</tr>
<tr>
<td>Heckertively, John R.</td>
<td>University of Michigan Medical School</td>
<td>186</td>
<td>2</td>
</tr>
<tr>
<td>Nickel, Robert E.</td>
<td>Oregon Health &amp; Science University</td>
<td>12</td>
<td>0</td>
</tr>
<tr>
<td>Weleber, Richard</td>
<td>Oregon Health &amp; Science University</td>
<td>163</td>
<td>1</td>
</tr>
</tbody>
</table>

**Experts based on...**

- Publications

**Your search terms**

Neuronal Ceroid-Lipofuscinoses

**Refine search by adding Concepts**

**Disorders**

- Retinal Diseases
- Metabolic Brain Diseases
- Myoclonic Epilepsies
- Lafora Disease
- Retinal Degeneration
- Stiff Person Syndrome
- Atrophy
- Progressive Myoclonic Epilepsies

**Add**

**Chemicals & Drugs**

- Cereol
- Lipofuscin

**Add**
Agenda

• Research Networking at Michigan
  • What is Collexis?
  • Collexis Implementation at Michigan
  • How it is used
  • Data Extension

• From Collexis to VIVO
  • Why VIVO?
  • Implementation
  • How it will be used
  • Data Extension
Implementation for customers

Process for a Collexis customer to launch VIVO:

1. Provide a budget for services to launch application (cost: <=1 FTE)
2. Collexis implements and populates a VIVO instance using cleaned-up data from Expert Profiling
3. Begin using launched, hosted and populated VIVO instance
4. Collexis maintains current version of software and updates publications and grants automatically
Results for Michigan

• Pre-populated VIVO instance
• 1,800+ faculty
• 60,000+ publications
• 18,000+ grants
• Full Departmental structure
• Little effort by IT or Research staff
After turn-key implementation:

- Many options for ongoing use:
  - Michigan configures implementation, or
  - Collexesis configures based on Michigan directives
  - Michigan adds data, or
  - Collexesis adds data for Michigan
  - Partner on problem-solving, customization, etc.

- Time spent is on using and extending the tool, not launching it
Agenda

• Research Networking at Michigan
  • What is Collexis?
  • Collexis Implementation at Michigan
  • How it is used
  • Data Extension

• From Collexis to VIVO
  • Why VIVO?
  • Implementation
  • How it will be used?
  • Data Extension
Use by Office of Research

- Use to connect and share information on non-Collexis content:
  - Events and Seminars
  - Press releases
  - Connection to custom applications by people using VIVO as data-provider
Agenda

• Research Networking at Michigan
  • What is Collexis?
  • Collexis Implementation at Michigan
  • How it is used
  • Data Extension

• From Collexis to VIVO
  • Why VIVO?
  • Implementation
  • How it will be used
  • Data Extension
Use Case: M-Cores

• Current state:
  – M-Cores: Core Resource web database for Medical School and Health Center research core facilities
  – Used by a small number of cores, difficult to keep up to date, manual updates and tagging with expertise
  – National initiatives on resource repositories is ongoing but disconnected
Use Case: M-Cores → M-Resources → CTSA Resourcome/eagle-i

• Future state (possible):
  – Extend from Core facilities to any resource that benefits from a profile
  – Use VIVO ontology combined with Collexis expertise fingerprints to match researchers and resources and list them in VIVO
  – Use publication feed from Collexis to identify possible users of cores
  – Use VIVO to add more information about core resources in the same location as expertise data
  – Leverage future U24 initiatives on resource repositories together with Collexis expertise
Links

• University of Michigan - Research Profiles
  http://www.researchprofiles.collexis.com/umichigan/

• Public access to VIVO:
  – coming soon!

• Collexis
  http://www.collexis.com
Questions?