“VIZ”ualizing the Scholarly Record

Sandy Payette and Muhammad Javed
Cornell University Library
Presented at CNI Fall Meeting 2016
Data

- People, publications, organizations
- VIVO ontology as data model
- Exposed as linked open data

Graphs for knowledge

- What are hot research areas?
- What are the patterns of collaboration?
- Who are expert faculty in what areas?
- Who co-authors in which areas?

Visualizations

- Navigate the scholarly record
- Dynamic interactive views
The Expanding Landscape of Research Information Management
Many systems have a piece of the action

- VIVO
- Symplectic Elements
- Pure
- ResearchGate
- Academia.edu
- Many More…
Profile proliferation

- VIVO profiles
- ORCID records - “profiles?”
- PURE profiles (Elsevier)
- Scopus author profiles (Elsevier)
- Mendeley profiles (Elsevier)
- Symplectic Elements profiles
- OpenScholar faculty websites
- SciENcv (NCBI/NIH)
- ResearchGate
- Academia.edu.
How to make sense of this?

• The messiness of emerging knowledge infrastructure
• Private institutional vs. public perspective
• Proprietary vs. open data
• Commercial vs. open source vs. hybrid technology
• Isolated systems vs. interconnected networks
Scholars Positioning in Ecosystem

External Data Sources
- ORCID
- Web of Science
- Crossref
- PubMed

Internal Data Sources
- Elements
- Cornell University Division of Human Resources
- Cornell University Office of Sponsored Programs

Web and Linked Open Data
- data/viz mediator
- VIVO

Grants, Pub Data, People Data
How we motivate our work

- The messiness of emerging knowledge infrastructure
- Private institutional vs. public perspective
- Proprietary vs. open data
- Commercial vs. open source vs. hybrid technology
- Isolated systems vs. interconnected networks
**Backend**

- Authority Records
- Graph
- Extendable/Reusable Model
- Inference
- OpenAccess to Data
- Links to external knowledge
- Inference
- ListView 2 VizView
- Data Download
- Fresh look & feel
- User-friendly
- Navigate through LD
- Frontend
"I would found an institution where any person can find instruction in any study."

Ezra Cornell, 1868
Scholars@Cornell
Highlights from Demo

★ VIZ-VIVO
  ★ semantic graph meets dynamic web
  ★ turning data into a knowledge

★ Uberization
  ★ best of the best citation data
Chen, Y-R (Cornell) co-authored an article with Blader, S.L. and Shirako A. (New York).

The article is published in *Personality and Social Psychology Bulletin*. 

Uberization Channel of Citation Entries

(VIVO) Harvester API

Linked Open Data

Clean and Complete Data
Inconsistent/Incomplete Data

Data Validation

Curation Bins

Ranked List of Citation Data Sources

Articles Bin

Journals Bin

SUMPLECTIC Elements

Uber Record

Option 1

Option 2

WEB OF SCIENCE
PREFERRED

PubMed

Crossref

arXiv

SSRN

Europe PMC
Pilot Phase - Key Questions

• Library nudge a university-wide bottom-up coordinated process?
  • Cornell highly decentralized in research info mgmt
  • Role of library? Role of academic units?
  • No central mandate or common RIM system

• Data Quality
  • Automated curation - How much can we do?
  • Human curation - how much? who?

• Many user stories - sweet spot for Scholars@Cornell?
  • Deans, department chairs, university administration
  • University communications and outreach
  • Faculty and researchers
  • Prospective students and faculty

• Ongoing investment to sustain?
QUESTIONS?

Sandy Payette, Jon Corson-Rikert
Muhammad Javed, Tim Worrall, Jim Blake, Joseph McEnerney
Jill Wilson, Mary Beth Martini-Lyons, Kathy Chiang
George Kozak, Holly Mistlebauer, Jason Kovari, Adam Smith

Learn more....
http://about.scholars.cornell.edu

username: scholars
password: libvivo1