NEW APPROACHES TO MEASURING THE LINKS BETWEEN RESEARCH AND INNOVATION

Cathy Buffington (U.S. Census Bureau)
Nathan Goldschlag (U.S. Census Bureau)
Ron Jarmin (U.S. Census Bureau)
Barb McFadden Allen (Committee on Institutional Cooperation)
Julia Lane (New York University)
Jason Owen Smith (University of Michigan)
Rebecca Rosen (American Institutes for Research)
Bruce A. Weinberg (Ohio State University)
Nik Zolas (U.S. Census Bureau)
Basics: How do Ideas Spread?

- We focus on people / networks as the transmission vectors for ideas to enter the broader economy
- This approach contrasts on bibliometric approaches focusing on disembodied ideas
  - Like looking for your keys under the light even if it’s dim
    - Geographic concentration of industries
    - People go to bricks and mortar schools
    - Travel to be face-to-face
Big Data: Follow the Money

- Identify graduate students employed on research projects
  - Use STAR / U METRICS data on CIC Universities to identify
  - Research teams (Undergrad, Grad, Postdoc, Staff, Faculty)
  - Purchases
  - Topics
  - Funders
- Use Census data to match researchers to subsequent employers
  - Location, Industry, Size, Age, Growth, Wages
- Match people to publications, grants, patents,…
- Trace flows of ideas from earnings, firm performance, even the areas that firms research
Establishment of new Institute

Institute for Research on Innovation and Science (IRIS) founded 01/01/2015

- Goal – leverage existing data to both serve university data and generate new research
- Core facility at University of Michigan
- 3 years seed funding for infrastructure from Sloan & Kauffman

Engagement by CIC, AAU and APLU

- Ongoing engagement with VPRs
- Links to Federal Statistical Agencies (e.g. Census)
Operational Structure

Universities contribute data, support infrastructure and receive campus-specific and aggregate products.

Approved nodes materially improve data, develop research/stakeholder communities.

Approved researchers securely access de-identified, aggregate datasets.

Create New research and reports.

“NPR membership” Model

63 researchers have accessed data.

Diagram showing the flow of data and interactions between different entities.
Figure 2. Career Trajectory

Pre / Post Doctoral Training Environment
U METRICS data on:
• Team / Network (size, composition, equipment)
• Funding mechanism
• Gender, Racial, Ethnic match with PI

Doctoral Degree
Linked data on:
• Thesis topics
• Advisors

Placement Outcomes
Census data on:
• Startups
• Sector, Location
• Firm Characteristics
• Earnings

Research Outcomes
Algorithmically linked data on:
• Publications, Citations, Text
• Patents
Data Architecture

Scientific Outputs and Biographic Data
- Publications
- Author Names, Keywords, Abstracts, Text, Acknowledgments, Affiliations
- Citations Co-Author Names, Affiliations, Altmetrics
- Dissertations
- Author Names
- Publications
- Biographic Information

Inputs
- Funding (Fed agencies, Foundations, Industry)
- Abstracts Resources
- STAR METRICS Level I
- Researcher time
- Human interactions
- Equipment

Large-Scale, Disambiguated, Longitudinal Researcher Data
- Web Searches
  - Age (Bio & Career)
  - Affiliations
  - Fertility

Results
- Utilization
  - Clinical Trials, Drug Approvals
- Patents
  - Inventor Names
  - Technology Classes
  - Citations

LEHD
- Census data
- Student/postdoc Employment History (UI Wage Records)
- Employer variables

ILDB
- Census data
- New Businesses

IPO Data
- Company ID
- Value
Researcher Characteristics

- Data tagged with job titles
- We can link our data to researcher characteristics
  - Gender, Race, Ethnicity, Age, National origin
- Look at how they are related to training environments
  - For instance, my grad student Wei Yang Tham finds
    - Substantial gender concentration on projects
    - Female grad students work on projects with female PIs
    - Female grad students employed on more faculty-intensive projects
- Can look at differences in Placement, Earnings, Firm performance, publications, patenting
Occupational distribution

Workforce Distribution for Grants Funded in Different Scientific Areas

 NIH Institute
- Faculty
- Graduate
- Postgraduate
- Staff
- Undergraduate

NSF Division
- Chemistry
- Materials Research
- Astronomical Sciences
- Computer and Network Systems
- Information & Intelligent Systems
- Advanced CyberInfrastructure
- Organisational Systems
- Mathematical Sciences
- Aging
- Mental Health
- Neurological Disorders & Stroke
- General Medical Sciences
- Kidney Diseases
- Cancer
- Diabetes & Digestive & Human Dev.
- Child Health & Human Dev.
- Allergy & Infectious Diseases
- Heart, Lung & Blood
Analysis of Teams and Networks

Figure 7. Sample Walktrip Community

Grad 17638759

Grad 29908482

Faculty member 43737344

Grad 17638759

Grad 29908482

Faculty member 43737344
Gender and Networks