

ICPSR



Role of Federal Statistical Research Data Centers in Promoting Transparency and Reproducibility in Federal Statistics

Margaret Levenstein

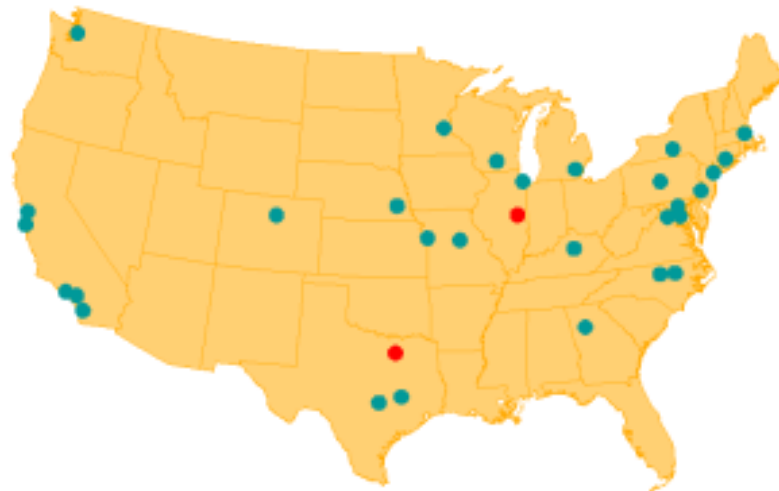
Director

Michigan FSRDC and ICPSR: The Data Consortium

What are the FSRDCs?

Joint projects between universities and federal statistical agencies to provide support and secure locations for researchers who need access to restricted data

30 FSRDCs across the country



Participating agencies

Agency for Healthcare Research & Quality

Bureau of Economic Analysis

Bureau of Labor Statistics

Census Bureau

National Center for Health Statistics

National Center for Science and Engineering Statistics

Data from BJS, BTS, EIA, EPA, ERS, HUD, IRS, SSA

FSRDC network

Each FSRDC has

- Executive Director

 - University professor

- RDC Administrator

 - Census employee

- Physical Data Enclave

 - Secure link to U.S. Census Bureau computing

Daily management: Project Management Office

- Meet weekly with Directors and Admins

Governance: Executive Committee

- Co-chairs: Chief Statistician and RDC Director

- Members: Rotating reps of statistical agencies and RDC directors

FSRDC Network

Vibrant network that has created and supported specific human capital

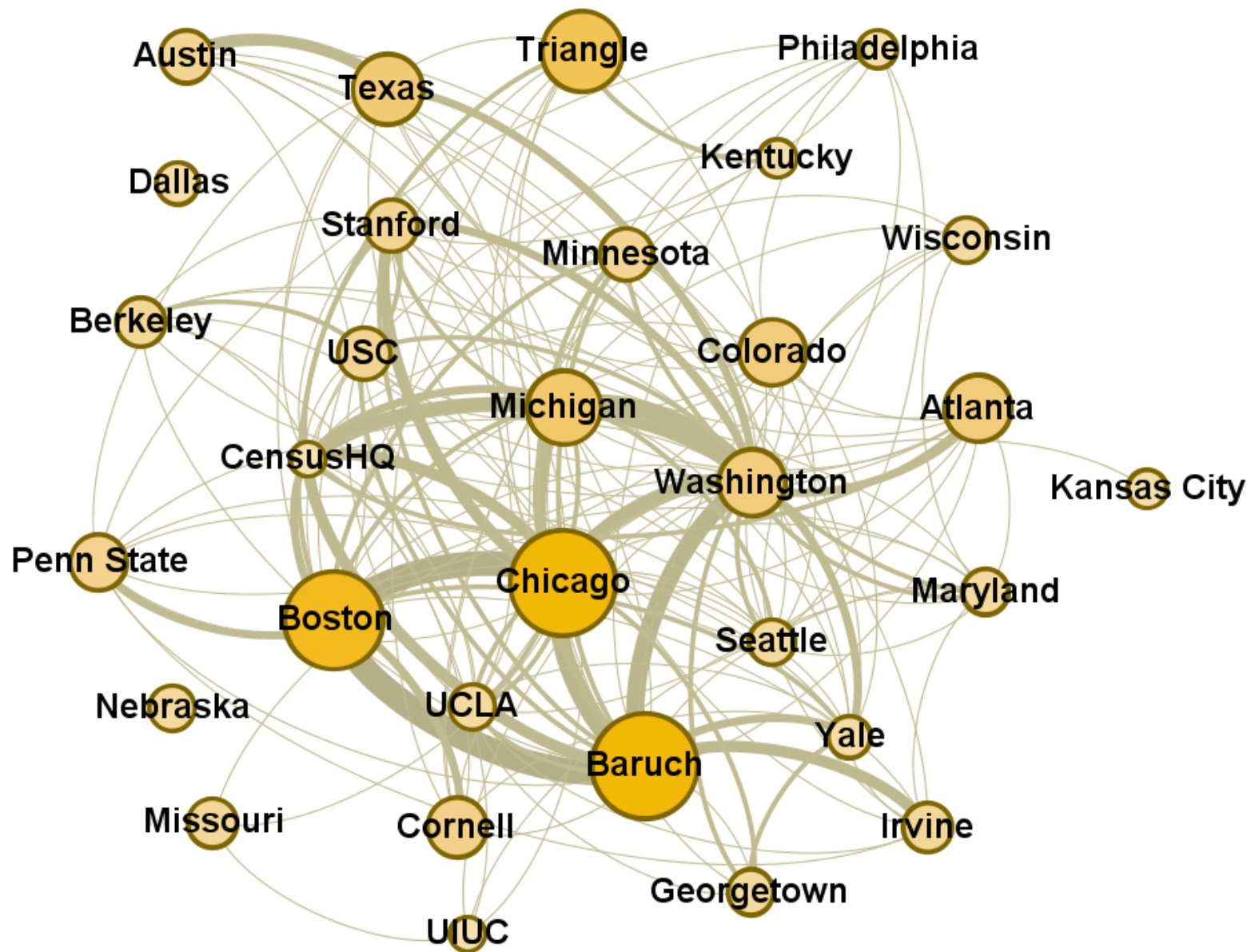
374 active projects

855 researchers

Lots of tacit and embedded knowledge

Network has expanded access to that knowledge, but it is still not perfectly transparent

Increasingly open network supports collaboration



Legal basis drives impact on transparency

Started in 1994 to facilitate access to Census's business microdata, for which there is no public use product

All projects accessing Census data must benefit the Census Bureau

Projects evaluate the quality of the data and the data generation process

Often business data use combination of survey and administrative data, so the "data generation" process is not simply the processed survey instrument

FSRDC network and transparency of federal statistics

Thousands of researchers have done significant analyses of data quality, as part of a quid pro quo for access to data for research

Improved understanding of public use data products

Increased interrogation of methods used to produce public use data products

- Imputation, disclosure avoidance

Challenge: use and sharing of this knowledge, even within a statistical agency

- Because everything is locked down

- Because agencies don't have absorptive capacity

Reproducibility and the FSRDCs

Reproducibility of federal statistics

- Enabled by availability of underlying raw data

- Limited by lack of code, detailed methods documentation

Reproducibility of analyses using FSRDC data

- Everything is archived

- Principled commitment to allowing replication projects

 - Plan to create replication project

- In practice, very challenging

Two initiatives to increase FSRDC transparency

Virtual Data Enclave pilot

Initiated in anticipation of increased need for RDC access with increased concern about protecting confidentiality of public use micro data

Single Portal

Impetus from Evidence Based Policy Act

Inventory, single point of entry to federal data

Virtual Data Enclave pilot

Leveraging Michigan governance and Census VDI

- Institutional contracts

- Census employee WFH technology

Based at Michigan FSRDC

Originally limited to ten projects

- Expanded rapidly because of pandemic

 - All RDCs closed in March, about 1/3 have re-opened

 - 38 projects

 - All but one are Census projects

 - 61 researchers

Evaluation underway

Virtual Pilot: What have we learned?

Researchers are elated that this is an option

Questions as we got off the ground

Who is in charge of what?

What's allowed and what isn't?

Can you talk on the phone? No

Do you have to work at home? Yes

Beta version of instant messaging inside FSRDCs to facilitate collaborative communication

Virtual Pilot: Process logistics

Drop in security checks

- Turn on camera and show inspector the set up

New online training

Revised agreement between researcher's institutions and University of Michigan, representing Michigan FSRDC

- Complement agreement between Census and researcher

Streamlined processes over time

- Who needs to be notified, approve to make each step happen

- New processes as more agencies participate

Single Portal: Research Data Gov

LOG IN/CREATE ACCOUNT

BROWSE BY AGENCY ▾ BROWSE ALL DATA ABOUT CONTACT FAQ



ResearchDataGov

Application Portal for Restricted Data for Federal Statistics



New data and agencies are being added! Don't see the data you want? Check back later or contact the agency directly.



Interagency Council on Statistical Policy

[About](#) | [Contact Us](#) | [Privacy Policy](#)

© 2019 The Regents of the University of Michigan

Research Data Gov increases transparency

Metadata for all FSRDC statistical agencies

And BJS

181 data sets, easily searchable, easy to add documentation, citations to publications and projects

On-line application form

Communication to agency and FSRDC

Review and approval by agency

Increases transparency

Available data

Process for access

Sheds light on some of that tacit knowledge

Single Portal: a work in progress

Missing metadata from some agencies

Harmonization of application processes

Easy, but then again not

Expansion to include non-FSRDC agencies

Funding, control

FSRDCs, Transparency and Reproducibility

Over almost 30 years, FSRDCs have increased transparency

Allowing researchers to see inside the sausage factory → Increased transparency of federal statistics themselves

Creating a human, and increasingly an online, resource describing data resources and application process → Increased transparency to data access

FSRDCs have been key institutions in realizing, and creating capacity to further, the goals of the EBP Act and the National Secure Data Service

**Data Jeff will take
your questions**

