

Administrative Data Research Facility -A Secure Environment to Integrate Data and Build Data Science Skills

Julia Lane
julia.lane@nyu.edu


Context


H.R. 1831: Evidence-Based Policy Act of 2016

Introduced: Apr 16, 2015
114th Congress, 2015–2017

Status: **Enacted — Signed by the President**
This bill was enacted after being signed into law.



Law: Pub.L. 114-140



Sponsor:  **Paul Ryan**
Representative for Wisconsin
Republican

Text:  **Read Text »**
Last Updated: Mar 18, 2016
Length: 5 pages




Color Key


- Thermal Overlay ☒
- Point Overlay ☒
- Selection Outline ☒

 Identify  Zoom to Selection

 Clear Overlays  Animate Overlays

Report/Map Outputs

-  Detailed Report
-  Export Geography
-  Print Chart/Map

 **Change Settings**

- 2020 Census design, however, on the new design
- CEDCaP (\$78M)
- American Corps continuing our
- Geographic S smarter geogra
- Administrativ and federally s linkage technic
- Economic & C are relevant to



Foundations for Evidence-Based Policymaking Act of 2018

The bipartisan Foundations for Evidence-Based Policymaking Act of 2018 builds off the work of the U.S. Commission on Evidence-Based Policymaking to strengthen data privacy protections, improve secure access to data, and enhance the federal government's capacity for producing and using evidence.

Strengthens Privacy Protections

Maintains Strong Confidentiality Protections for Sensitive Data. Reauthorizes the Confidential Information Protection and Statistical Efficiency Act (CIPSEA), an existing law that gives the American public strong privacy safeguards and legal protections for appropriate uses of confidential data.

Institutes Processes to Assess Data Risks. Strengthens efforts to protect confidentiality while making data accessible for evidence building and transparent to the public by requiring comprehensive risk assessments for certain publicly released data.

Enhances Public Trust in Data. Improves public trust in statistical activities by codifying language directing certain agencies to establish procedures to protect trust in data activities by appropriately maintaining objectivity, independence, and confidentiality.

Establishes Consistent Leadership on Key Data Issues. Ensures a senior leader in each agency is responsible for protecting privacy and ensuring confidentiality protections are appropriately applied by creating chief data officers.

Improves Secure Data Access

Encourages Agencies to Make Data Public and Open When Possible. Takes steps to improve the public information about what data government currently holds and make data publicly available when possible and in the public interest.

Requires Development of Data Inventories. Enables researchers and evaluators to better identify what government-collected data are available by directing agencies to create and maintain data inventories and publicly provide details about those datasets.

Makes Administrative Records Available for Evidence Building.

Under a strong set of confidentiality protections, encourages that government data can and should be used to generate evidence about policies and programs, unless otherwise restricted by law.

Creates a Common Portal for Researcher Applications to Access Restricted Data. Reduces burden on researchers for applying to access government data by establishing a common application system for qualified individuals to access restricted, confidential data for approved projects.

Facilitates Continuous Feedback about Data Coordination. Promotes the use of data for evidence building by establishing a government advisory committee to review existing coordination and availability of data.

Enhances Government's Evidence Capacity

Directs Agencies to Develop Evidence Plans. Enables agencies to better prioritize evidence building by requiring that agencies document their key research questions, data needs, and planned activities.

Prioritizes Evaluation Activities in Agencies. Improves agency capacity to engage in and use program evaluation by establishing evaluation officers in government agencies and requiring agencies to develop written evaluation policies.

Develops Baseline Information about the Resources Available for Evidence Building. Directs government agencies to periodically assess and report on their capabilities to engage in statistical, evaluation, and policy analysis activities and use the corresponding evidence for day-to-day government operations.

Documentation Reload Text-Only

Click a Characteristic link in the Summary Report to see more detail.

Search Facility: The Search Facility is a pilot access to analytical discovery services, and is for users, including government analysts. The Census Bureau developed the laborative Training analytics sponsored by New York University, Maryland.¹ It is currently with users accessing training program. The data-based computing security approvals, selected confidential element of Housing and the Census Bureau, as well as state agencies, and an

2015	Count	Share
	954	100.0%

2015	Count	Share
	782	21.9%
	296	57.4%
	876	20.6%

[Contact Us](#) | [Home](#)
[Feedback@census.gov](#)

Two year pilot activities

Technical

- Create secure environment where data providers can share their data across agency and jurisdictional lines

Operational

- Link disparate data
- Analyze data

Legal & Practical

- Document value associated with the data linkage
 - Consistent with the agency mission
 - Useful enough to engage decision-makers

ADRF Safe Data Strategy

- Safe People
 - Approved and trained researchers
- Safe Projects
 - Approved projects, consistent with agency mission
- Safe Settings
 - Secure environment, GovCloud, FedRamp Moderate
- Safe Outputs
 - Disclosure reviews and export controls


= SAFE DATA

Build technical environment

Users: Federal, state and local data owners
 Analysts and researchers
 Federal, state and local program
 managers

Technical Needs:

Security
Access, Discovery and Collaboration
Analysis and Dissemination



Secure
Reusable
Scalable
Extensible
Interoperable

Security

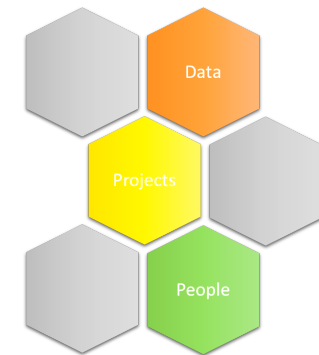
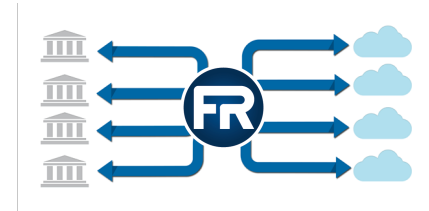
Federal Risk and Authorization Management Program

Provides a standardized cloud-based approach:

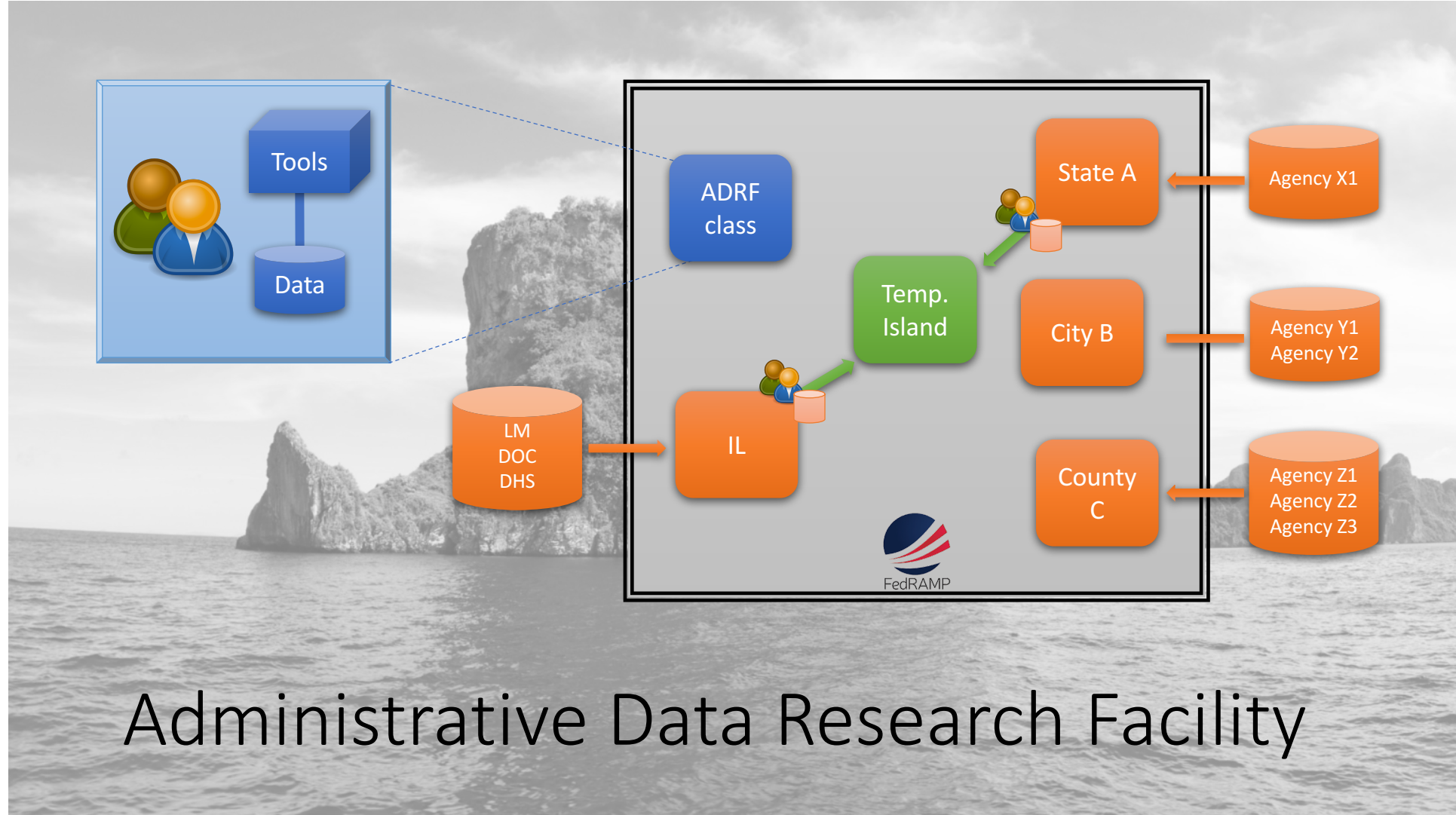
- security assessment
- authorization
- continuous monitoring

ADRF Status

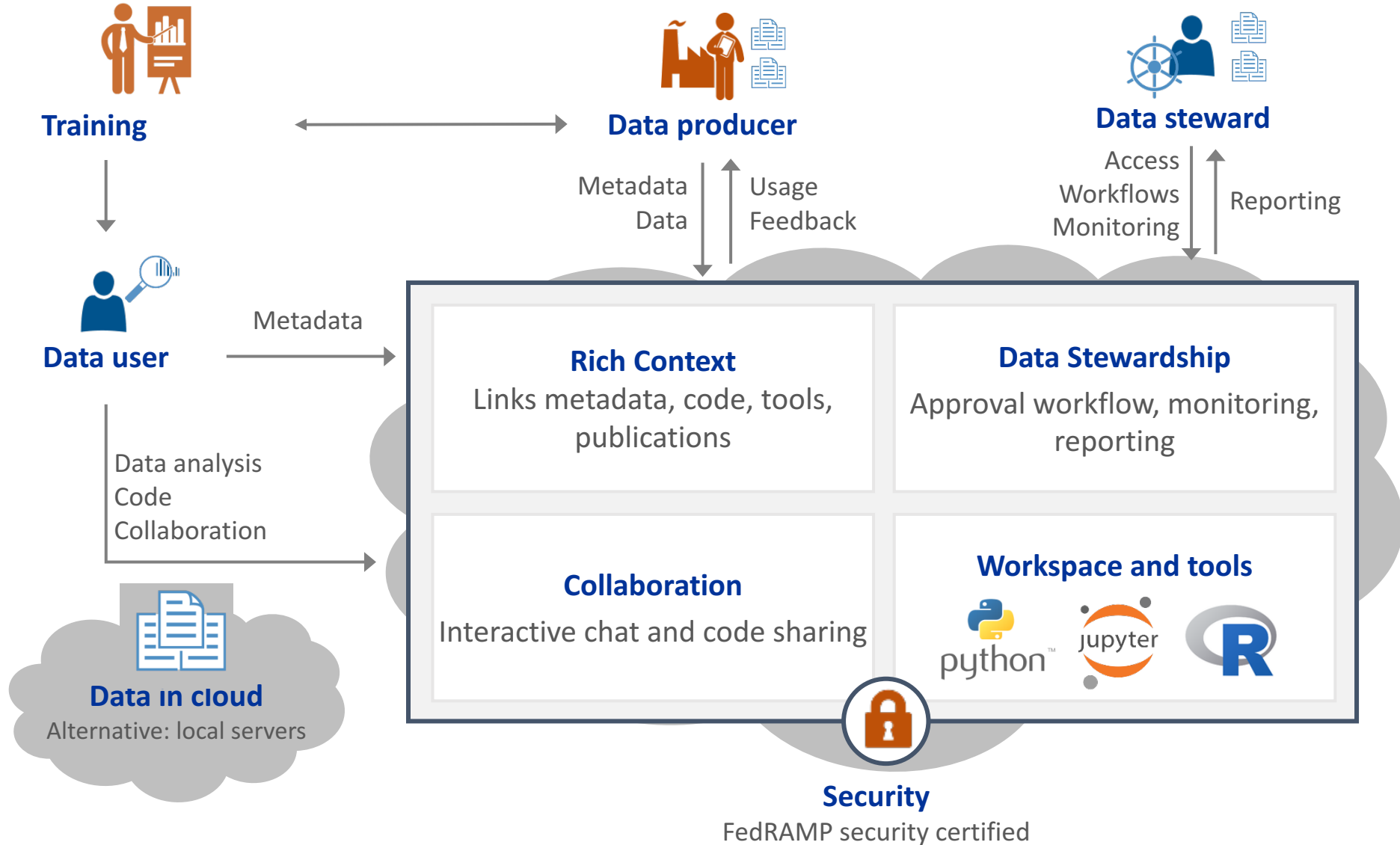
May:	Successful readiness assessment
June:	Census Authority to Test
July:	Title 13 Census data ingested
September:	Full Assessment
February 2018:	Census Authorization to Operate
October 2018:	FedRAMP marketplace



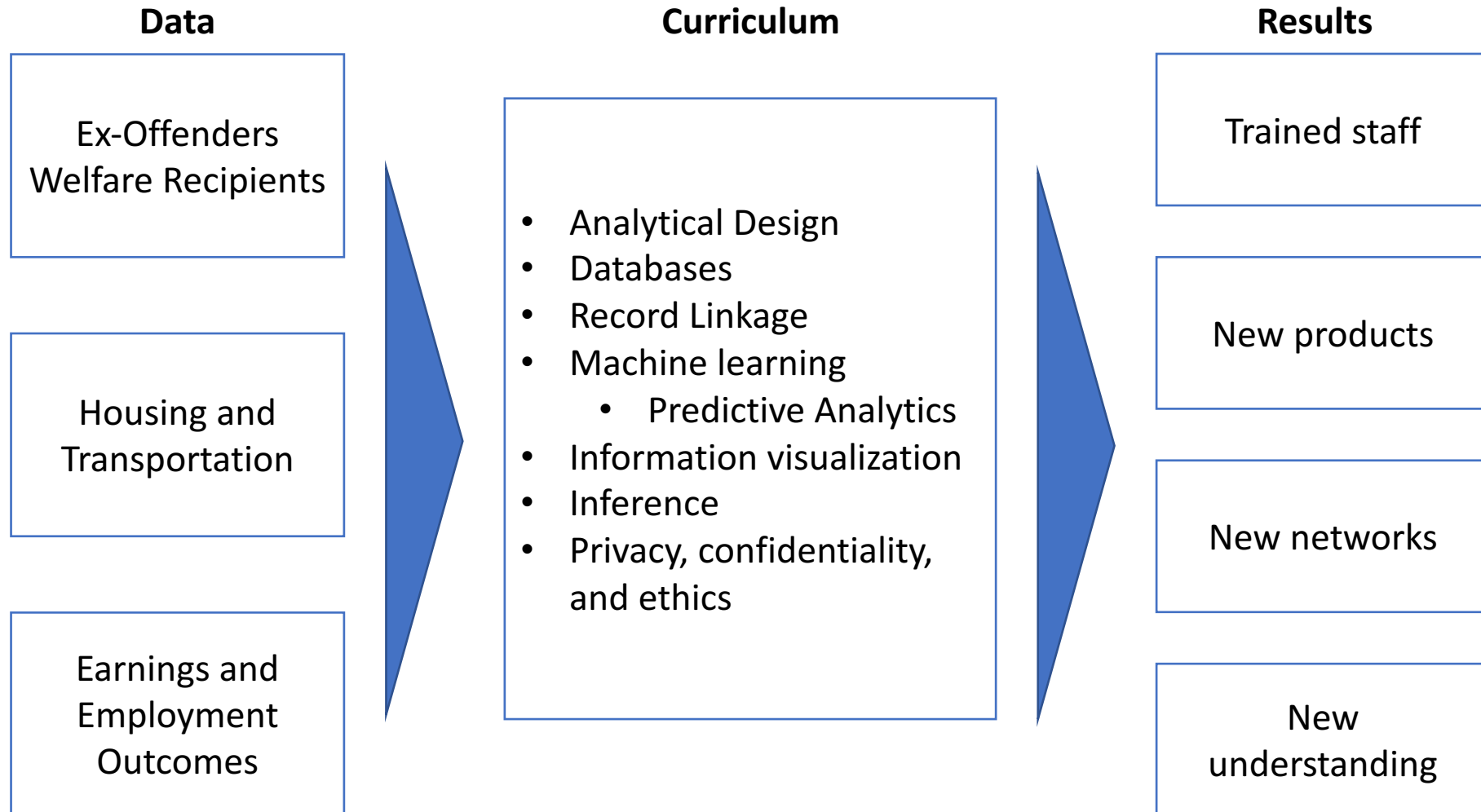
Access



Operational



Legal and Practical



Pilot Results



almost 300 participants from over 90 agencies have completed almost 60 projects

Pilot Results: Interstate data

1. Comparing Employment Dynamics Across Borders

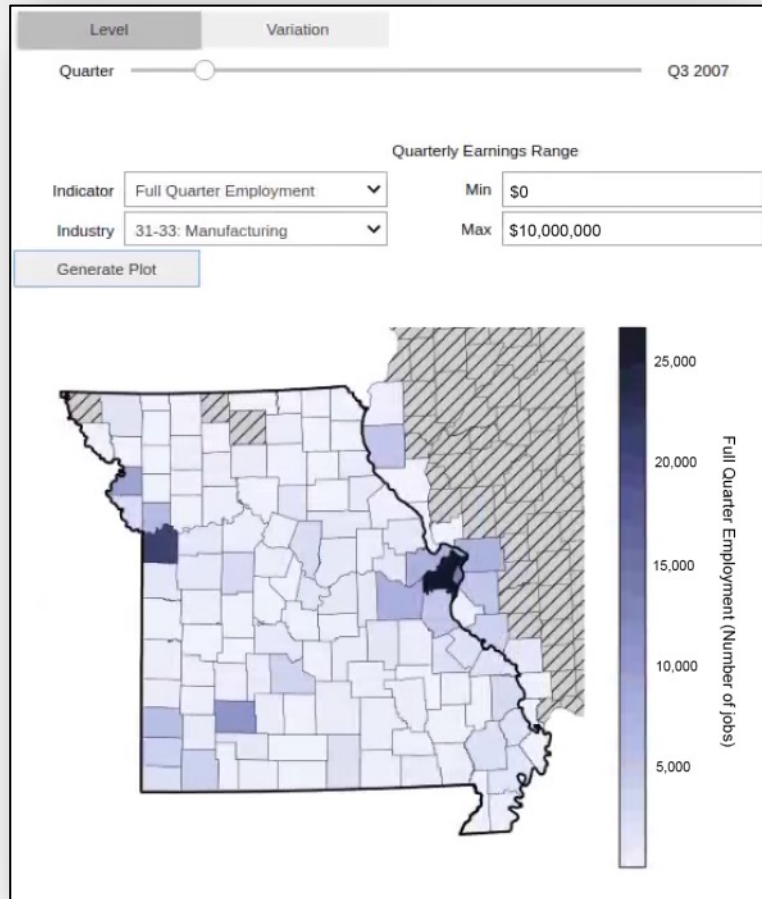


Fig. 3: Comparing total earnings with Illinois border counties

The dashboard can include border counties from the states that provide data to the ADRF.

UPCOMING TRAINING PROGRAMS

Columbus, OH - Winter 2019

APPLY

The Ohio State University is hosting the Winter, 2019 Coleridge Initiative Applied Data Analytics training program, sponsored by the Overdeck Family Foundation. Participants will work in teams to define and complete a project related to Transitions in Education and Workforce. The program will provide up-to-date perspective on the use of administrative data for policy analysis, and hands-on instruction of using micro data in SQL and Python for the following tasks: data management, record linkage, data visualization, and machine learning.

Application: **Rolling**
1st Admission Decision: **December 15**
SQL & Python tutorials: Complete by **February 1**
In-person training in Columbus, OH
• Module 1: **February 6-8**
• Module 2: **March 13-15**
Remote Presentations: **April 23**

Outcomes from Transitions in Education and Work: Developing a Scalable Regional Approach

Summary

There is enormous interest in building a better understanding of how people transition across different educational and work experiences to sustainable jobs. Transitions of people as they age can be nonlinear and include transitions within and across secondary and post secondary institutions, as well as the use of government services such as disability services, criminal justice interventions, or workforce development. Individuals can also move across political jurisdictions such as state or county lines, making it incredibly difficult to understand the regular patterns in service utilization. The sheer variation on types of transitions individuals can make and the geographical movement across borders opens up opportunities for governments to intervene in more productive ways. The data sharing across states and agencies also provides increasing possibilities for more efficient and effective utilization of government services among vulnerable populations.

Building a common framework across states requires (i) identifying key issues of interest (ii) bringing together data from many sources across both state and agency lines and (iii) bringing together staff from those agencies. These three activities are necessary to work through the important data and measurement issues that are necessary to build operational metrics for decision-making. This workshop will build on a successful collaborative approach developed by the State of Illinois, New York University's Coleridge Initiative, Chapin Hall and Ohio State University that uses an inter-agency and inter-state collaborative, classroom environment to address the issues. It will bring together key stakeholders from the Midwest to develop an operational agenda for the spring of 2019. We expect to discuss how to develop such metrics as:

1. What are the main educational and work flows of interest for policy?
2. What data are available?
3. What kind of metrics would be of greatest use to stakeholders?

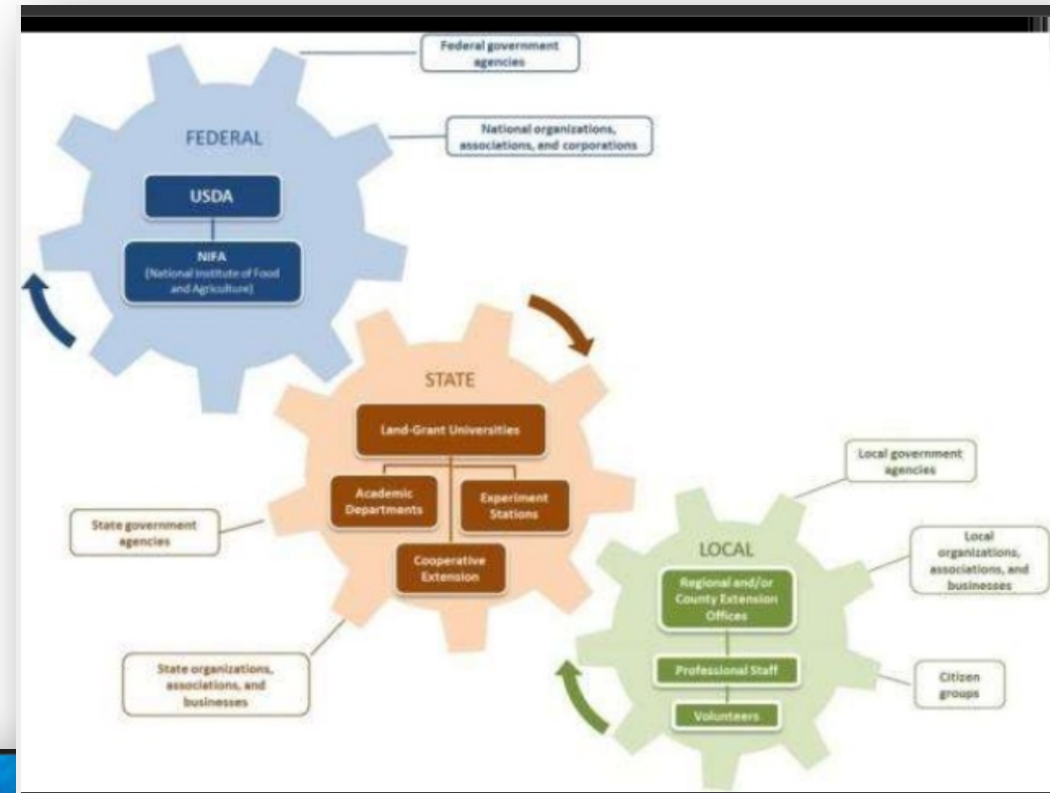
The National Academies of
SCIENCES • ENGINEERING • MEDICINE

CONSENSUS STUDY REPORT

DATA SCIENCE FOR UNDERGRADUATES OPPORTUNITIES AND OPTIONS



Vision



Contact and more information

- [Website: https://coleridgeinitiative.org/](https://coleridgeinitiative.org/)
- [E-mail](#)
 - dataanalytics@coleridgeinitiative.org
 - julia.lane@nyu.edu
 - clayton.hunter@nyu.edu
- [GitHub organization: https://github.com/Coleridge-Initiative](https://github.com/Coleridge-Initiative)