

Data-Sharing and Administrative Data

JULIA LANE

THE COLERIDGE INITIATIVE

NEW YORK UNIVERSITY

Context



BIPARTISAN POLICY CENTER

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 enhance the federal government's capacity for producing evidence-based policy.

Agency Actions			
	Foundations for Evidence-Based Policymaking Act of 2018 and Associated OMB Guidance	Executive Order on Maintaining American Leadership in Artificial Intelligence	Improving Implementation of the Information Quality Act (M-19-15)
1. Identify Data Needs to Answer Priority Agency Questions	✓		✓
2. Constitute a Diverse Data Governance Body	✓		✓
3. Assess Data and Related Infrastructure Maturity			✓
4. Identify Opportunities to Increase Staff Data Skills	✓		
5. Identify Priority Data Assets for Agency Open Data Plans	✓	✓	✓
6. Publish and Update Data Inventories	✓		

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

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

Status: Enacted — Signed by the President
This bill was enacted after being passed by both the House and Senate.

Law: Pub.L. 114-140

Sponsor:  **Paul Ryan**
Representative for Indiana's 6th District, Republican

Text:  [Read Text](#)
Last Updated: Mar 2016
Length: 5 pages

- **2020 Census** design, how to use the new data.
- **CEDECaP** (Center for Evidence-Based Decision Making).
- **American Data Science** continuing to improve data science.
- **Geographic Information Systems** smarter geospatial data.
- **Administrative Data** and federal data linkage technology protection.

• **Economic & Government Censuses (\$144M)**: Data products drive economic activity and are relevant to the needs businesses, policymakers, and the public. \$10.1 million increase

well as state, city, and county agencies, and an array of public use data.

Open Records Available for Evidence
A strong set of confidentiality protections, government data can and should be used to learn about policies and programs, unless prohibited by law.

Open Portal for Researcher Applications to Government Data. Reduces burden on researchers for government data by establishing a common portal for qualified individuals to access restricted, approved projects.

Continuous Feedback about Data Coordination. Improves coordination of data for evidence building by establishing a biennial committee to review existing coordination data.

Government's Evidence Capacity

Require Agencies to Develop Evidence Plans. Enables agencies to prioritize evidence building by requiring that agencies submit their key research questions, data needs, and evidence plans.

Standardize Evaluation Activities in Agencies. Improves agency evidence in and use program evaluation by establishing standards in government agencies and requiring agencies to develop written evaluation policies.

Improve Information about the Resources for Evidence Building. Directs government agencies to assess and report on their capabilities to engage in evidence building, and policy analysis activities and use evidence for day-to-day government.

[evidence](#)

Context

Approach

Achievements

Future

Lessons Learned

Approach: Technical and Human

Technical

- Create secure environment where data providers can share their data across agency and jurisdictional lines
- Census and USDA Authorization to Operate; HHS in process

Operational

Link disparate data

Analyze data

Legal & Practical

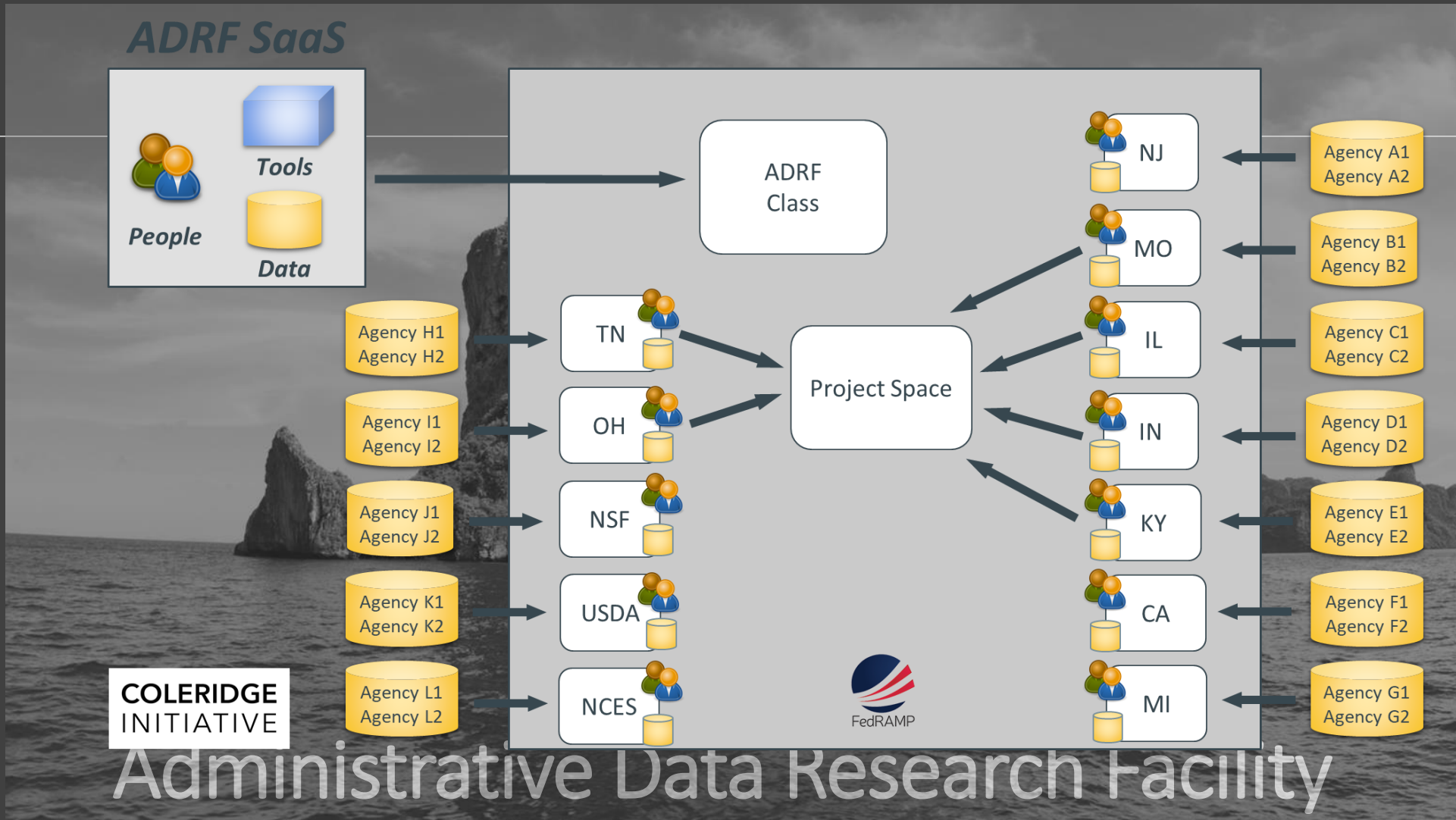
- Document value associated with the data linkage
 - Consistent with the agency mission
 - Useful enough to engage decision-makers

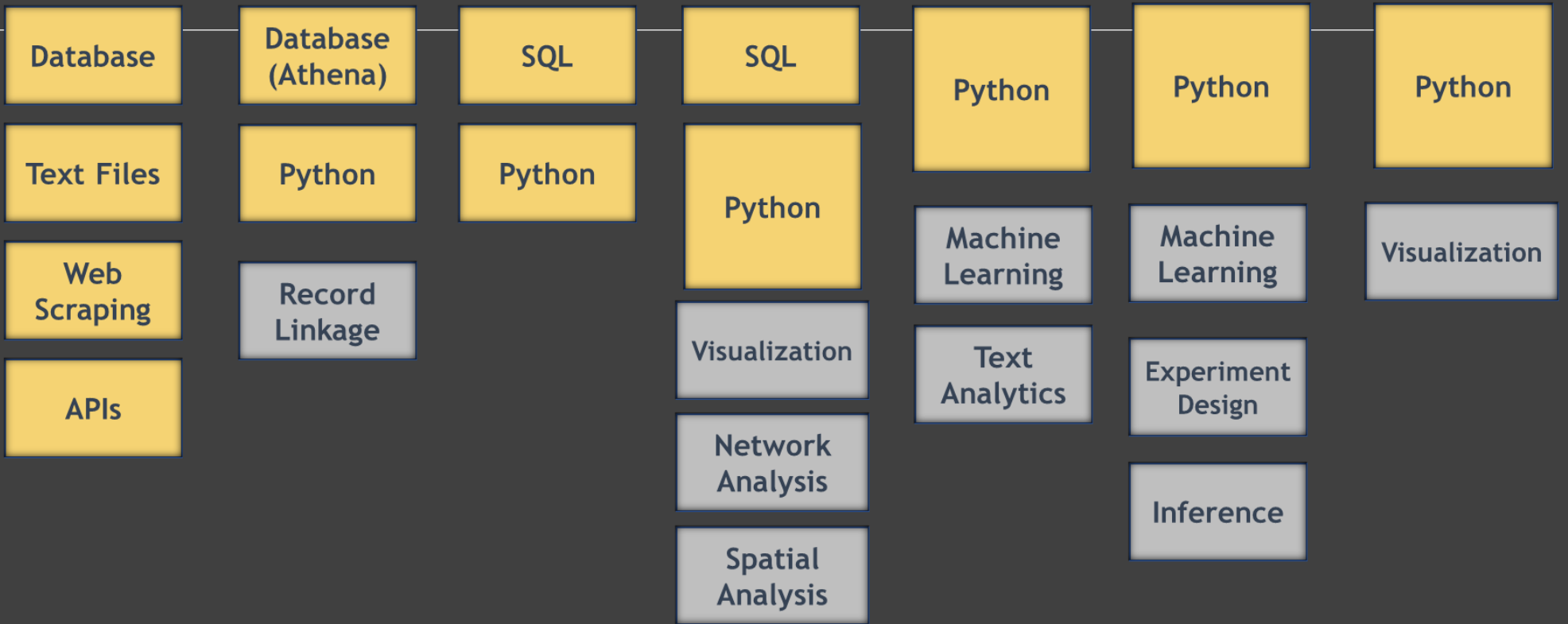
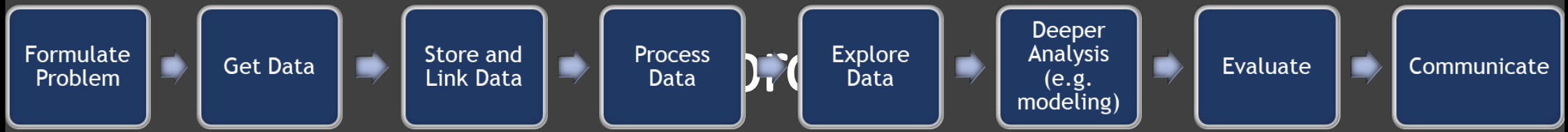
Technical Approach: Five Safes

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

= SAFE USE

Safe Setting





Collaboration: "shared" folders in your project

Privacy, Confidentiality, Security

Context

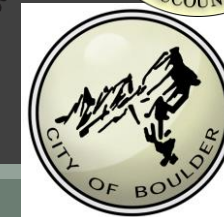
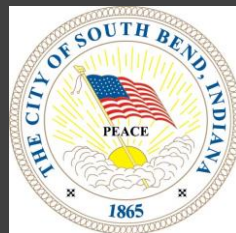
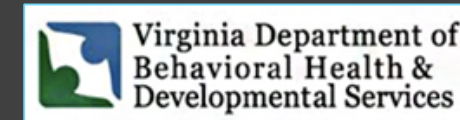
Approach

Achievements

Future

Lessons Learned

Work in teams



Context

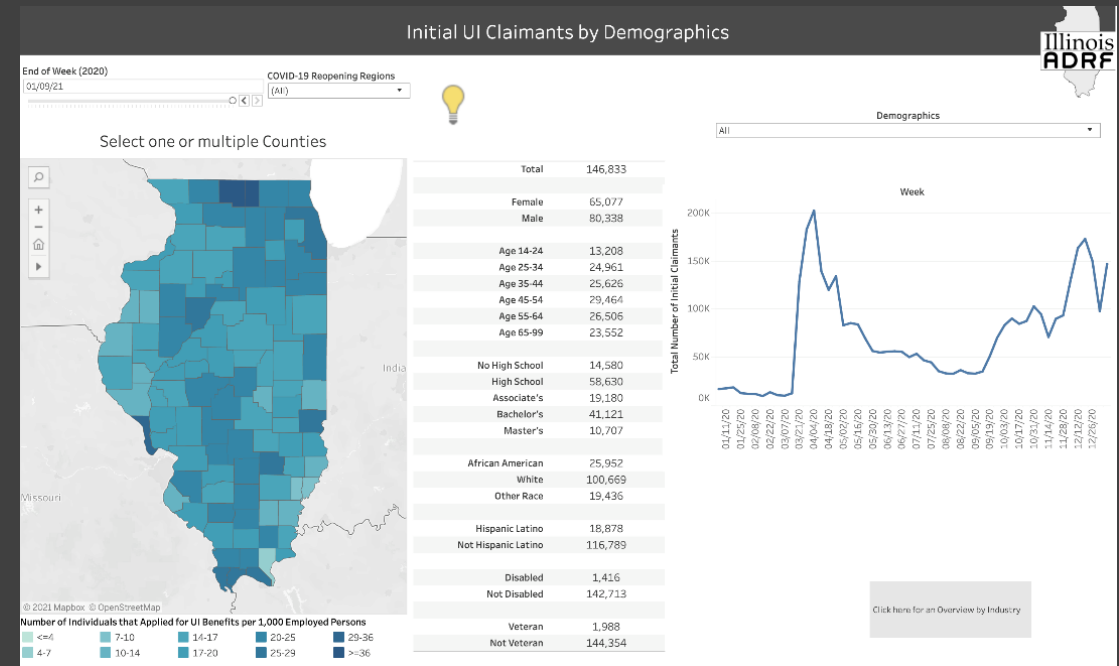
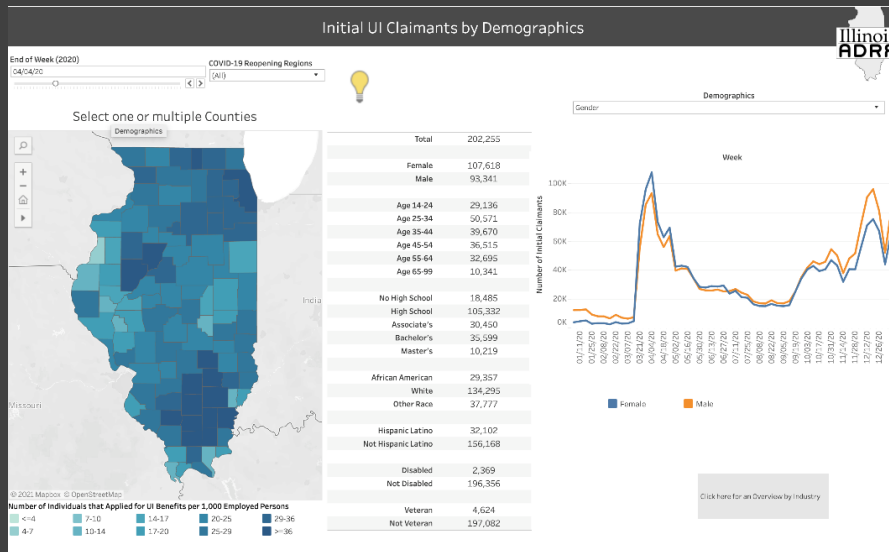
Approach

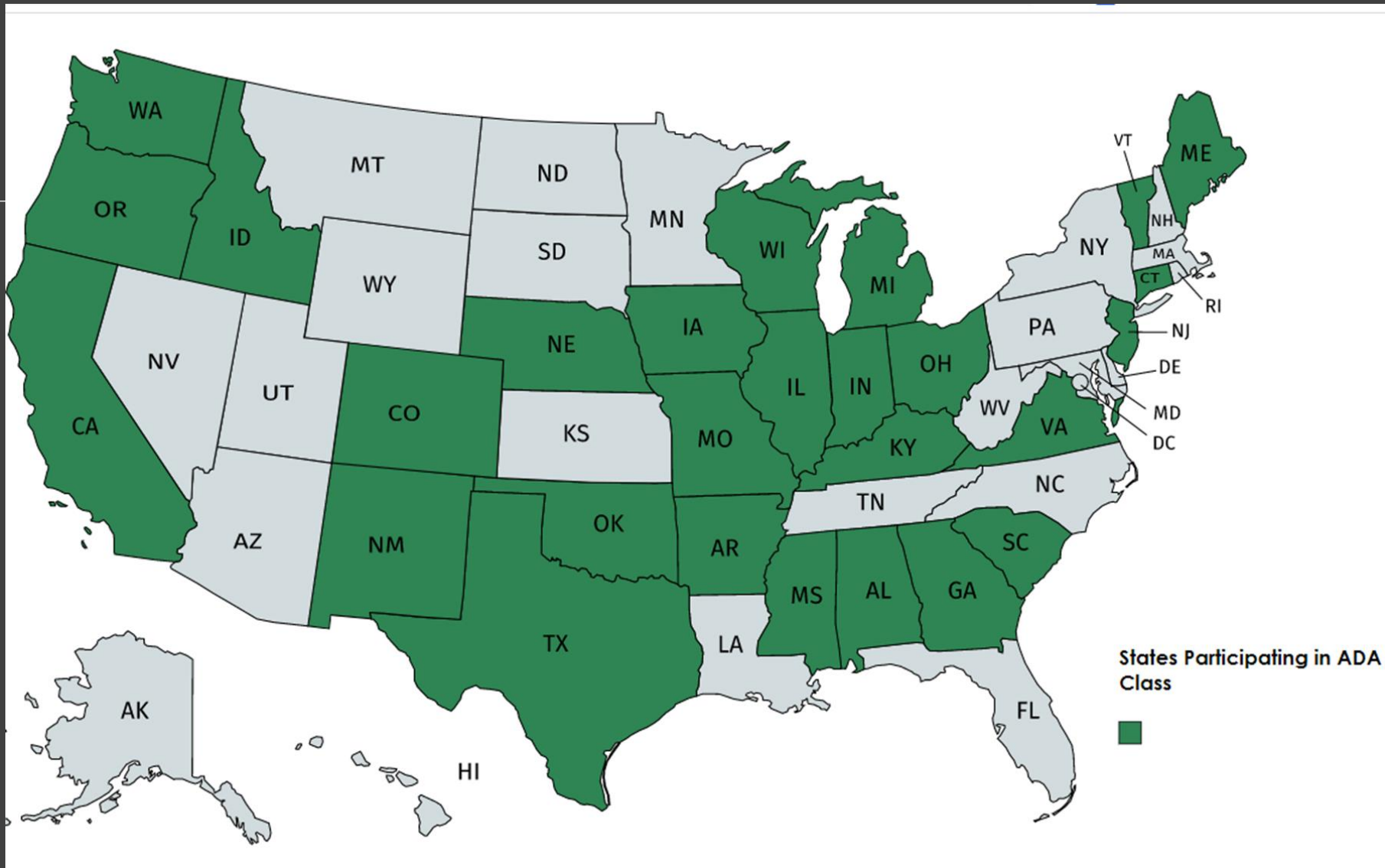
Achievements

Future

Lessons Learned

Unemployment to Reemployment Portal: Demographics





Context

Approach

Achievements

Future

Lessons Learned

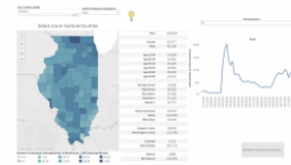
Register

Featured Data Products

Unemployment to Reemployment Portal

States in the Midwest Collaborative are working across state borders to create real-time labor market information that is critical for local pandemic response. This interactive tool draws in the latest de-identified UI claims data, processed through the secure ADRF environment. Through visualizations, state staff can:

- Easily identify which local labor markets and categories of workers are being hardest hit,
- Estimate the economic impacts of job loss, and
- Describe the demographic, education, industry, and occupational characteristics of UI claimants.

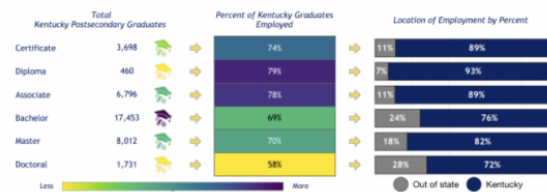


Enlarge Image
Initial Unemployment Insurance Claims Dashboard

Learn more by viewing a [presentation that demonstrates the uses of the Unemployment to Reemployment Portal](#), with the Labor Market Information Director for the State of Illinois, Dr. George Putnam, and the Assistant Director of Research and Analysis for the State of Indiana, Diana Barrett.

Multi-State Post-Secondary Dashboard

The dashboard will empower states to better understand the employment outcomes of individuals who participated in their State's education programs. The information presented in the dashboard will inform state decision-making in education and workforce policy by providing employment outcome measures from multiple states' workforce data.



Enlarge Image

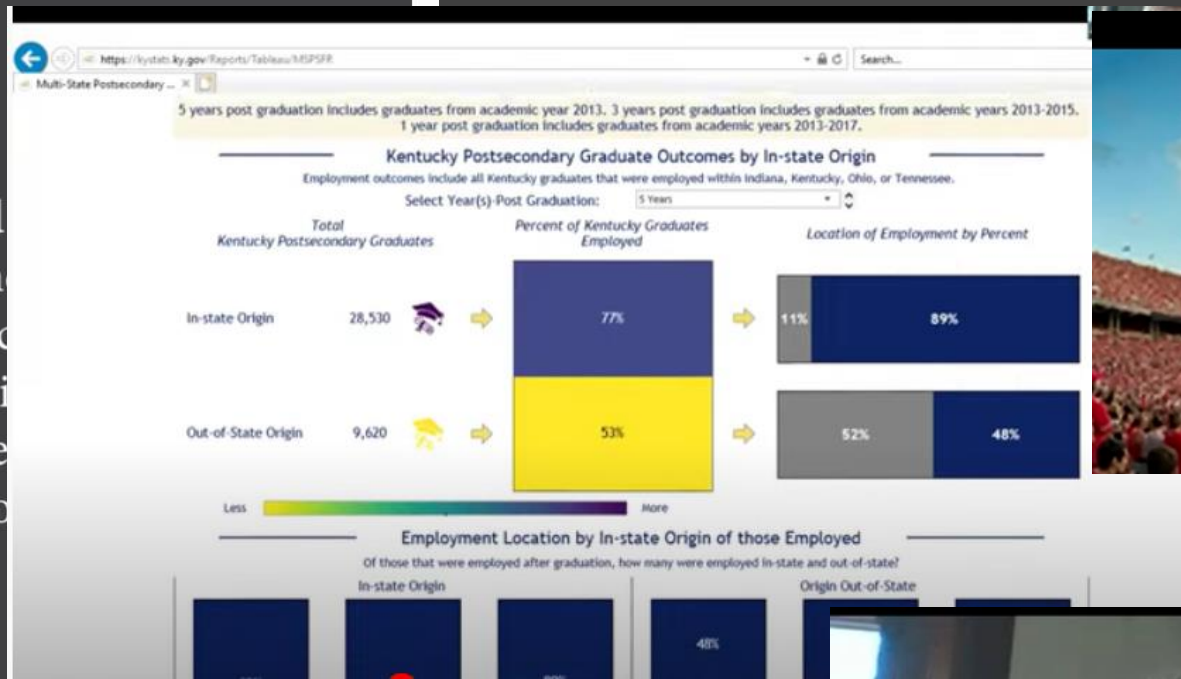
Meeting passwords will be provided following registration.

Questions? [Contact us](#)

Sponsored by:



COVID-19 and the resulting budget crisis have heightened the importance of training and improve outcomes for America at the forefront of making training decisions — and they need be evidence to make wise decisions



The Multi-State Post-Secondary Dashboard

Unlisted



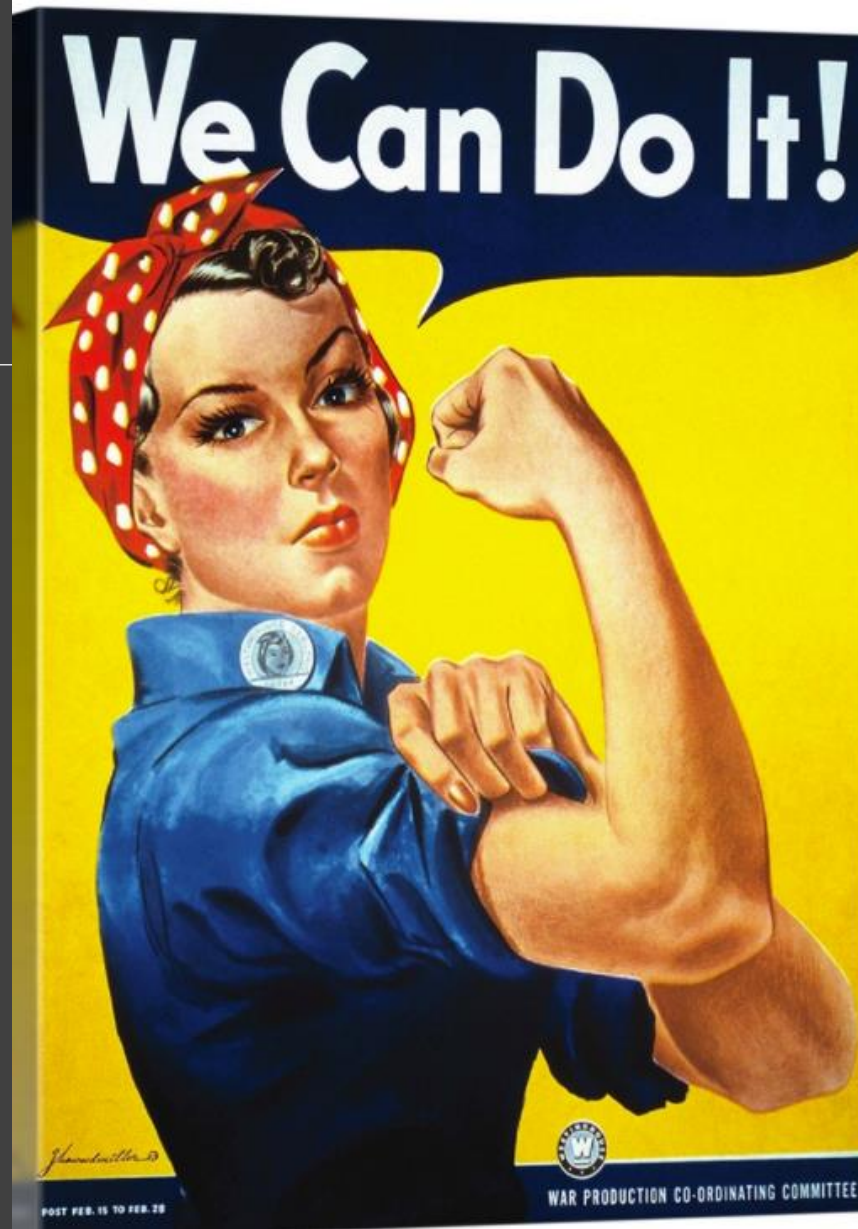
Context

Approach

Achievements

Future

Lessons Learned



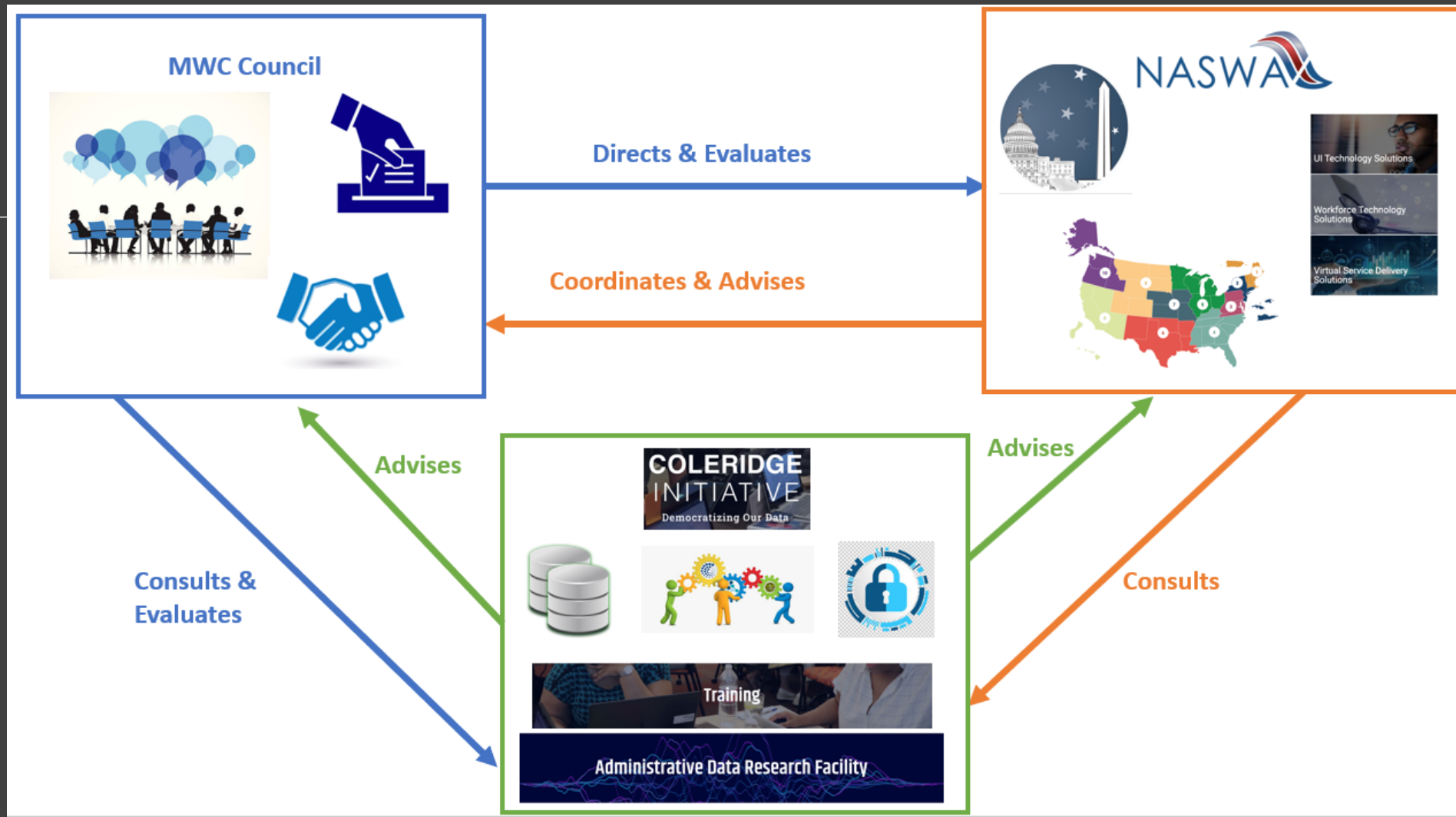
Context

Approach

Achievements

Future

Lessons Learned



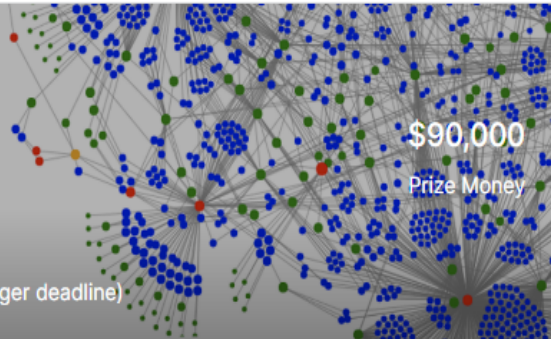
Coleridge Initiative - Show US th x +

← → ↻ 🏠 🔒 https://www.kaggle.com/c/coleridgeinitiative-show-us-the-data

🔍 Search Sign In Register

- ☰ kaggle
- 🏠 Home
- 🏆 Competitions
- 📁 Datasets
- 🔗 Code
- 💬 Discussions
- 📖 Courses
- ⌵ More

📄 View Active Events

Featured Code Competition


Coleridge Initiative - Show US the Data

Discover how data is used for the public good

Coleridge Initiative · 1,134 teams · a month to go (a month to go until merger deadline)

\$90,000

Prize Money

Overview Data Code Discussion Leaderboard Rules

Join Competition

Overview

Description	This competition challenges data scientists to show how publicly funded data are used to serve science and society. Evidence through data is critical if government is to address the many threats facing society, including; pandemics, climate change, Alzheimer's disease, child hunger, increasing food production, maintaining biodiversity, and addressing many other challenges. Yet much of the information about data necessary to inform evidence and science is locked inside publications.
Evaluation	
Timeline	
Prizes	Can natural language processing find the hidden-in-plain-sight data citations? Can machine learning find the link between the words used in research articles and the data referenced in the article?
Code Requirements	
Coleridge Initiative	Now is the time for data scientists to help restore trust in data and evidence. In the United States, federal agencies are now mandated to show how their data are being used. The new Foundations of Evidence

Quick wins and visibility key (training programs have been magic)

Security necessary but not sufficient

Strong, trusted leadership

Value proposition is key

- Create value for agencies -
- Operational and implementable (timeliness often key)
- Analysis of outliers
- Analysis of small populations
- Analysis of regions and localities

Workforce capacity is critical

Long term investment necessary

<https://hdsr.mitpress.mit.edu/pub/0mb0zzlc/release/5>

<https://www.nature.com/articles/d41586-020-02733-3>

<https://www.scientificamerican.com/article/the-u-s-needs-a-national-data-service/>

https://gcn.com/Articles/2018/11/01/PSI_Administrative-Data-Research-Facility.aspx?m=1

The screenshot shows the MIT Press website page for the book "Democratizing Our Data: A Manifesto" by Julia Lane. The page features a navigation bar with links for Books, Journals, Open Access, Resources, Give, About, and Contact. Below the navigation bar, there is a breadcrumb trail: Home | Information Science | Data Science | Democratizing Our Data. The main content area displays the book's cover, which shows a person standing in a large, circular, data-like structure. The title "DEMOCRATIZING OUR DATA" is prominently displayed in large, bold, white letters, with "A MANIFESTO" and "JULIA LANE" below it. To the right of the cover, the book title "Democratizing Our Data" and subtitle "A Manifesto" are listed, along with the author's name "By Julia Lane". A green button with a shopping cart icon and the letter "B" is visible. Below the book information, there are tabs for "Overview", "Author(s)", and "Praise". The "Overview" tab is selected, showing a "Summary" section with the text: "A wake-up call for America to create a new framework for democratizing data." and "Public data are foundational to our democratic system. People need consistently high-quality information from trustworthy sources. In the new economy, wealth is..."

Contact and more information

Website: <https://coleridgeinitiative.org/>

Email

- julia.lane@coleridgeinitiative.org

GitHub organization: <https://github.com/Coleridge-Initiative>

Class information: <https://ada.coleridgeinitiative.org> (password adrf)

Textbook: <https://textbook.coleridgeinitiative.org/>

Kaggle competition: <https://www.kaggle.com/c/coleridgeinitiative-show-us-the-data>