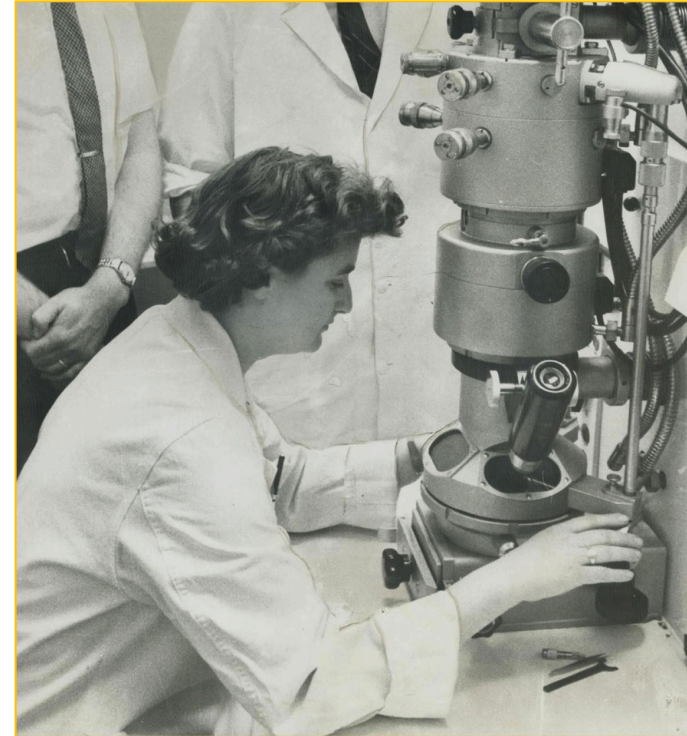


Gifts for WDQI Thinkers & Doers

- **The Great Mental Models Project** ([Farnam Street](#)) explores everything you need to know about mental models. By the time you're done, you'll think better, work better, live better.
- [Alternate Script](#): **Coronavirus reminds us what functioning communities look like.** #MutualAid (Quartz)
- June Almeida (1930-2007). A Scot without a college degree, she discovered coronaviruses in 1964. **Today, we're using the techniques she developed to identify COVID-19 a half century ago.** ([NatGeo](#))



“Without doubt [June Almeida] is one of the outstanding Scottish scientists of her generation.”



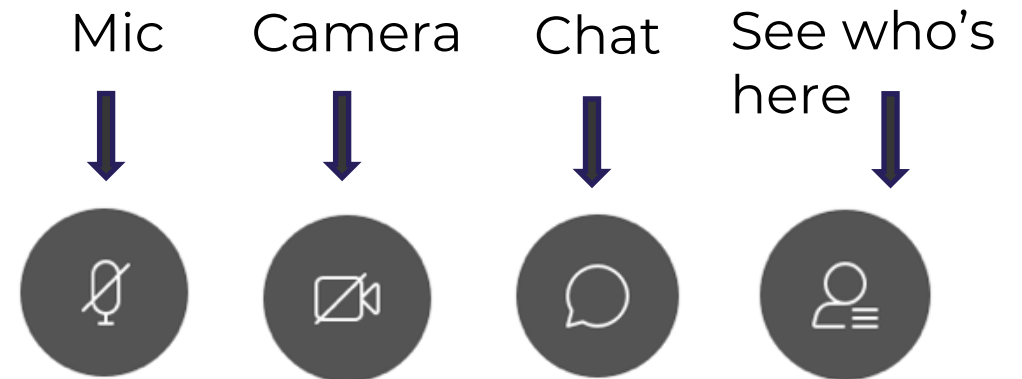
Let's Talk About Collaboration

(in a COVID-19 World)

WDQI Community Convening
April 22, 2020

Before we start

1. Please turn off mic & camera if not speaking.
2. Type questions into chat.
3. We may call on you.
4. We are recording.



We're glad you could join today.



Agenda

- A deep breath
- Collaboration and data sharing (thanks IL team!)
- Q&A



Take a breath. Chat.

What's one thing you've learned
(personally or professionally) since March
1, 2020?



Photo by [Food Photographer Jennifer Palian](#) on [Unsplash](#)

So many questions

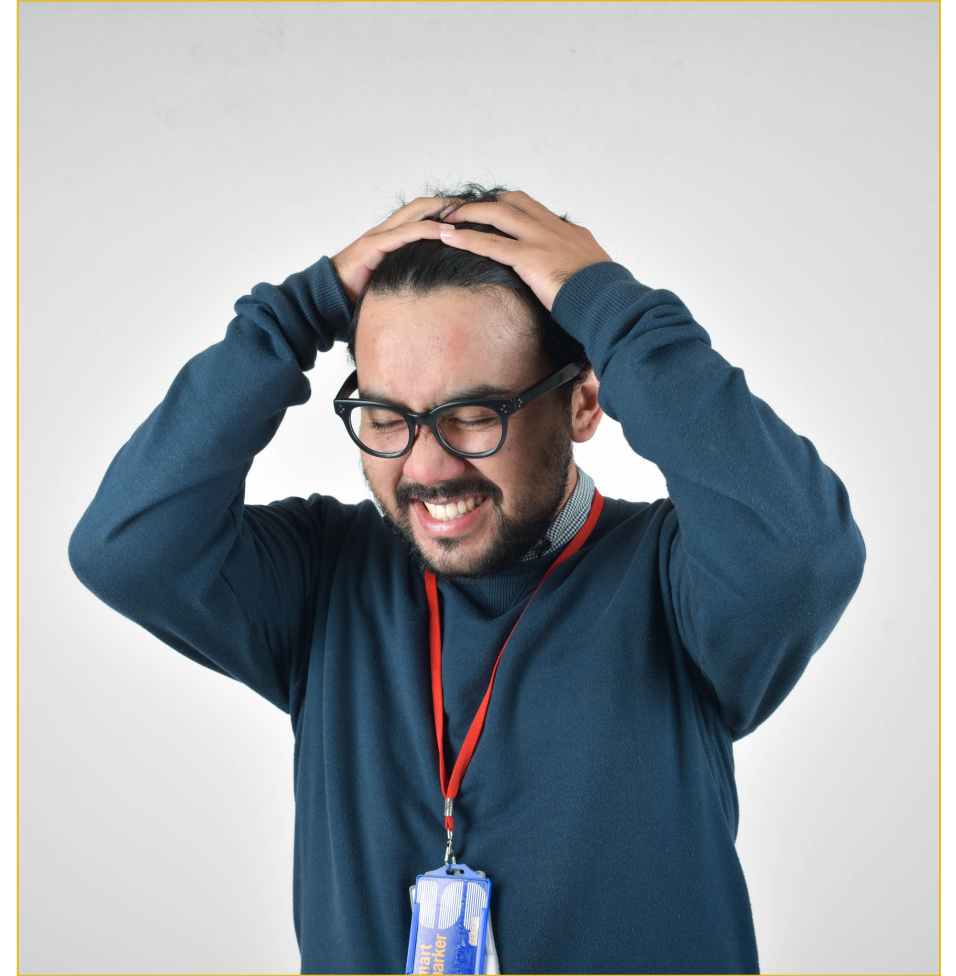
Over 40 of you RSVP'd, which is a lot 🙌.
We're officially stopping at the hour, but if you don't have to rush off, Maureen, Kristin, the Illinois team and Julia can stay on the line for a few extra minutes to answer any additional questions.

Or email them to
Maureen_Sarver@spra.com.



Why? Why now?

- UI at home?
- How many claims would be filed?
- Who has what data and how does it get shared?



Collaboration in a COVID World

The Emerging Midwest
Data Collaborative



Patti Schnoor & George Putnam,
IL WDQI Team



Julia Lane,
Coleridge Initiative,
NYU

Illinois – WDQI Round VII



Common Program Information

Strategic objectives address significantly improving the quality and breadth of workforce training program data via the ProPath Illinois initiative and the Common Program Information (CPI) Classifications

ProPath Illinois Universal Integrated Provider Platform

Utilizing Southern Illinois University, Illinois workNet technology in tandem with the National Credential Engine Registry innovation schematics and profiles to create a cloud based statewide platform which links program data with pathways, credentials, competencies, and outcomes data.

Illinois Administrative Data Research Facility (I-ADRF)

Extending research capacity and longitudinal performance evaluation for program effectiveness by joining the New York University - Administrative Data Research Facility (ADRF) to use the most innovative data sharing strategies available to foster evidence-based policy making and work toward regional vision of workforce data sharing.



System Fundamentals for Architecture

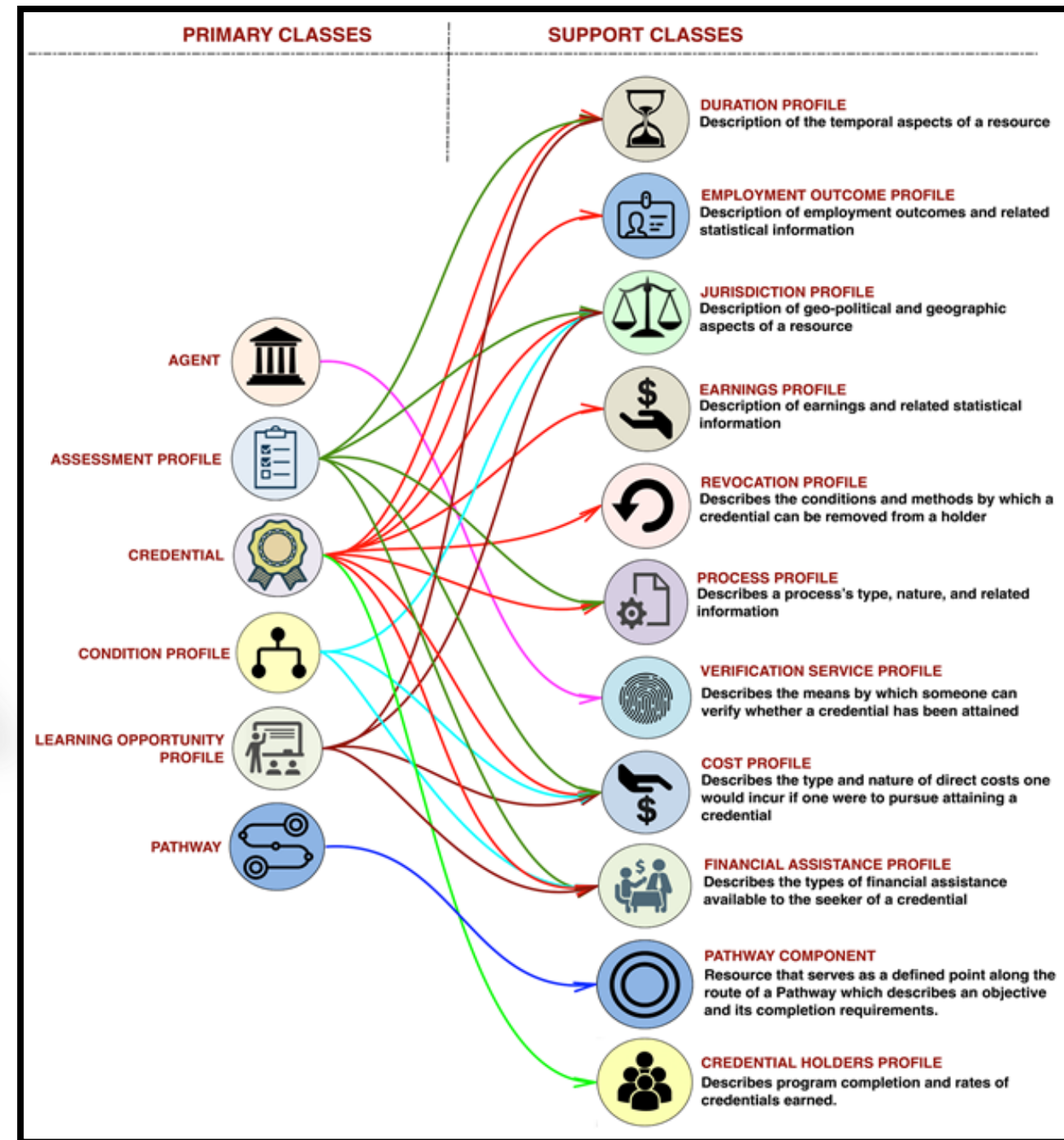
Fundamentals	Description
Open Licenses	Use open specifications and licenses wherever possible (e.g., Creative Commons for content, Apache for software, CEDS, Credential Engine, Schema.org, and World Wide Web Consortium for specifications).
Schema/Data Dictionary	Define the terms (i.e., classes, properties and concept schemes/vocabularies) to define expected values. Where reasonable, use terms from established schemas (e.g., schema.org, CTDL, GoodRelations).
Data Specifications and Models	Determine best solution for modeling data (CPI - Common Program Information elements) and how data needs to be expressed externally. Resource Description Framework (RDF) is the World Wide Web Consortium's specification for Linked Data. This supports applications that need to consume data and for making data usable on the Web.
Process Workflows	End-to-end progression for how users will need to interact with publishing, managing, and consuming data.
Protocols and Permissions.	How data is accessed and exchanged.
Data Quality Policy	APIs, written policy requirements for minimum data and currency of data.
Serializations	How the data is programmatically encoded and exposed for: exchange formats between systems, data on the Web, search engine crawling.
APIs	Web services to support integrations and 3 rd party applications.

Example Publishing Management

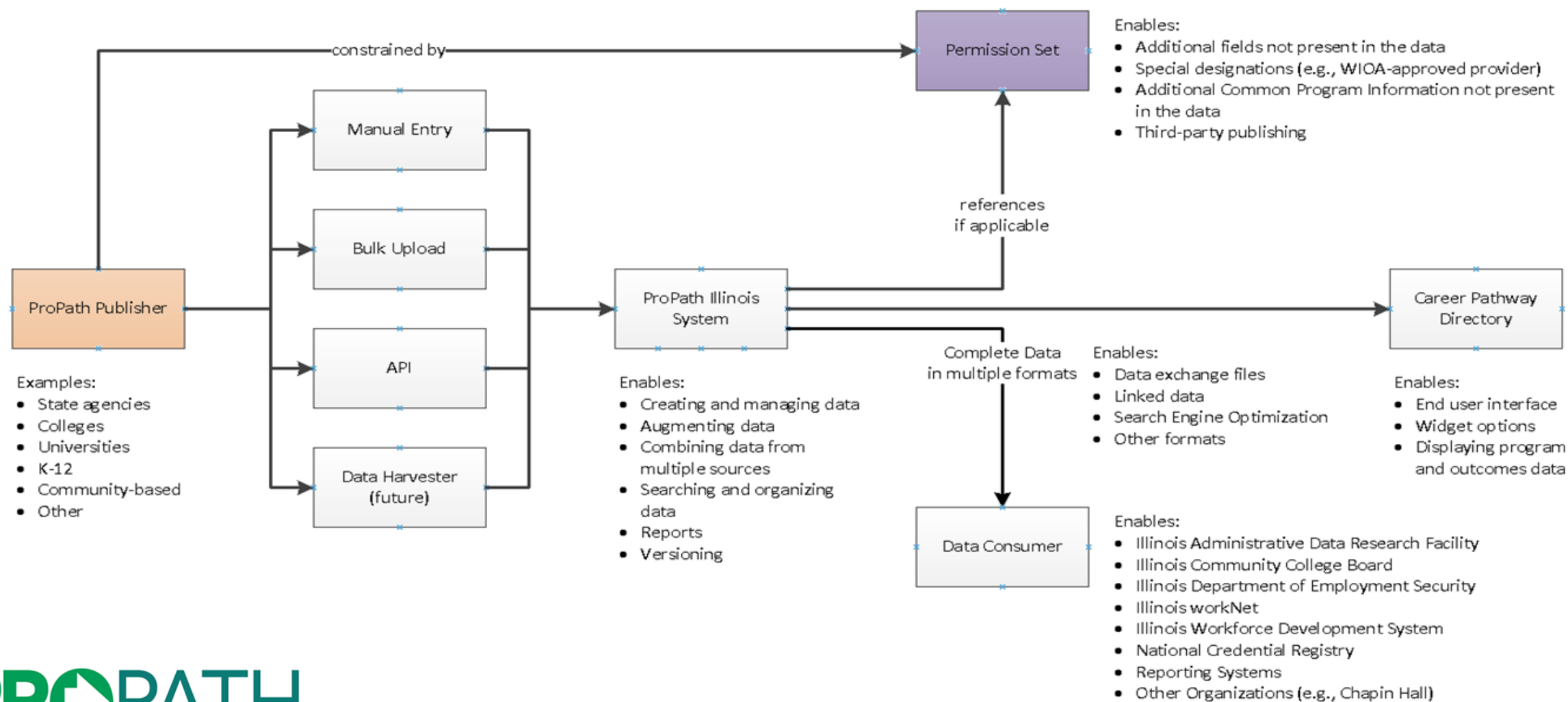
The ProPath System will enable multiple organizations to provide and manage education and training information via a defined set of methods.



Statewide Universal Training Provider Platform
with Common Program Information
and Credential Engine Schematics



ProPath Illinois Publishing Workflow



Illinois Administrative Data Research Facility (I-ADRF)

WIOA Performance Indicators *ENHANCED MEASURES*

Employment Rate 2nd Quarter After Exit

Employment Rate 4th Quarter After Exit

Median Earnings 2nd Quarter After Exit

Credential Attainment

Measurable Skill Gains

1. The I-ADRF is the Illinois component of a broader ADRF ecosystem, administered by New York University (NYU), containing microdata from several Midwestern states.
2. Leverage existing State data infrastructure to provision career outcome information to potential training participants and support the scaling of successful programs.
3. Through Pro Path Illinois, additional State program participant data on education, workforce training (Title I), and employment services (Title III) will be added to I-ADRF.
4. Regional projects will be supported, and new stakeholder products on workforce outcomes will be generated.
5. Chapin Hall and Urban Labs at the University of Chicago will partner with State agencies to develop evaluation frameworks and dashboards that enhance performance information (e.g., job stability, career job earnings, and industry job flows) for dissemination.
6. These evaluation frameworks and dashboards will address WIOA performance accountability requirements and provide learning pathways and career outcomes data for incorporation into evidence-based outreach methods and policy formulation.

Socializing the IL WDQI: The Midwest Collaborative



MidWest Collaborative (IA, IL, IN, KY, MI, MO, NJ, OH, TN)

Data-analytic environment for evidence-based operational and programmatic initiatives that draw from interagency/interstate research imperatives, Covid-19 impact

Socializing IL WDQI: standardization

Program evaluation framework (cohort structure)

Dashboard metrics (workforce outcomes, wage replacement)

Inflows, through flows, and outflows to unemployment and employment

Interstate and interagency value-add

Socializing the IL WDQI



The MidWest Collaborative

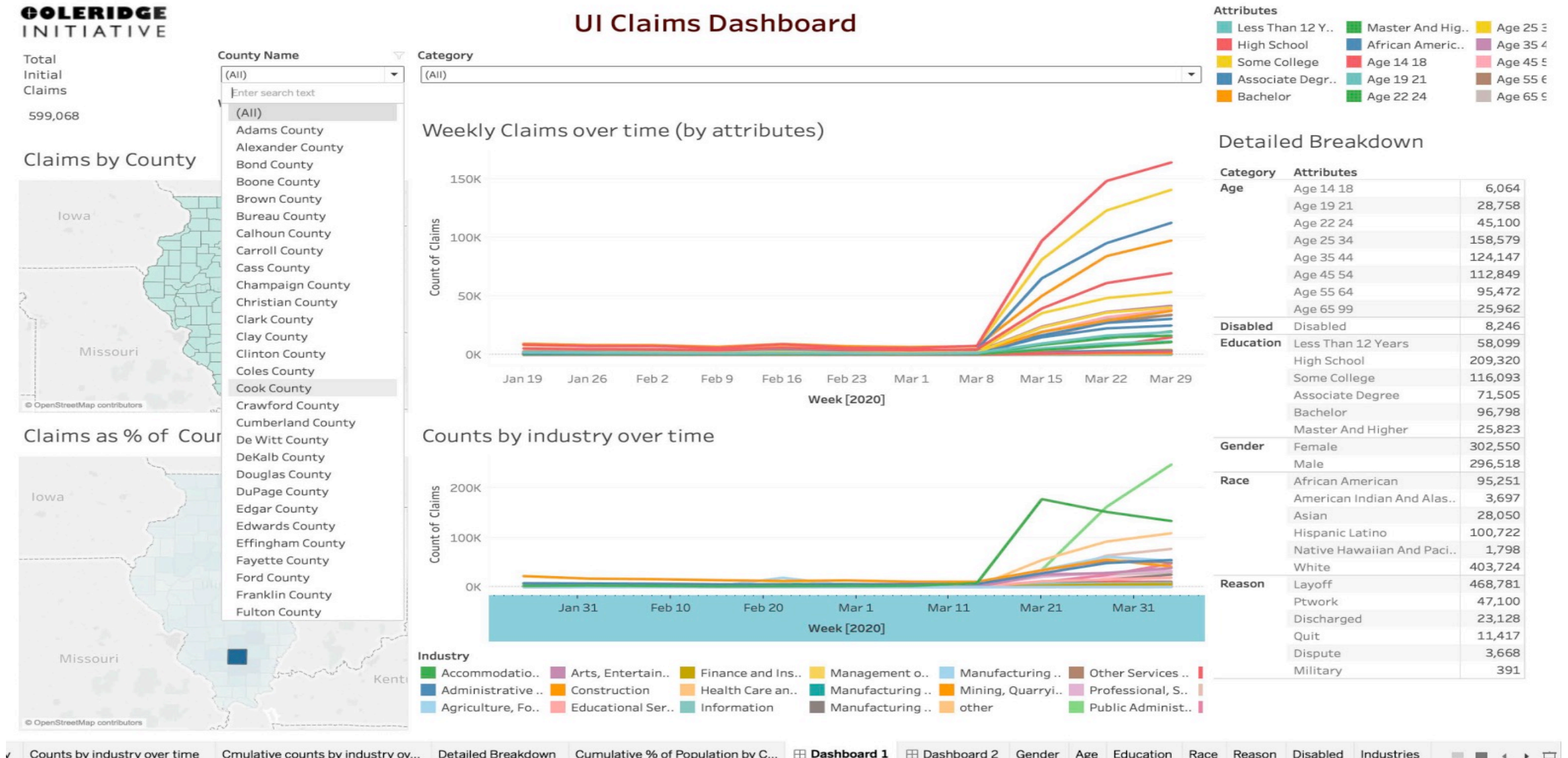
- Data-analytic environment for evidence-based operational and programmatic initiatives that draw from interagency/interstate research imperatives
- Covid-19 impact

IL WDQI Standardization

- Program evaluation framework (cohort structure)
- Dashboard metrics (workforce outcomes, wage replacement)
- Inflows, through flows, and outflows to unemployment and employment
- Interstate and interagency value-add

MidWest Collaborative: UI Portal (prototype)

Left Panel- County map views and selection; Center Panel- data time series display; Right Panel- detailed breakdown by claimant attributes



MidWest Collaborative: UI Portal (prototype)



Gender

Attributes

Female

Male

Count of Claims

0K 50K 100K 150K 200K

Disabled

Attributes

Disabled

Count of Claims

4K 6K 8K

Race

Attributes

African American

American Indian And Alas..

Asian

Hispanic Latino

Native Hawaiian And Paci..

White

Count of Claims

0K 100K 200K 300K 400K

Education

Attributes

Less Than 12 Years

High School

Some College

Associate Degree

Bachelor

Master And Higher

Count of Claims

0K 50K 100K 150K 200K

Age

Attributes

Age 14 18

Age 19 21

Age 22 24

Age 25 34

Age 35 44

Age 45 54

Age 55 64

Age 65 99

Count of Claims

0K 20K 40K 60K 80K 100K 120K 140K 160K

Reason

Attributes

Discharged

Dispute

Layoff

Military

Ptwork

Quit

Count of Claims

0K 100K 200K 300K 400K

Industry

Accommodatio..

Administrative ..

Agriculture, For..

Arts, Entertain..

Construction

Educational Ser..

Finance and Ins..

Health Care and..

Information

Management of..

Manufacturing ..

Manufacturing ..

Mining, Quarryi..

other

Other Services ..

Professional, S..

Public Administ..

Real Estate and..

Retail Trade 44

Retail Trade 45

Transportation ..

Transportation ..

Unknown

Utilities

Wholesale Trade

Industry: Health Care and Social Assistance

Attributes: Female

Count of Claims: 48,667

Counts by industry over time

Cumulative counts by industry ov...

Detailed Breakdown

Cumulative % of Population by C...

Dashboard 1

Dashboard 2

Gender

Age

Education

Race

Reason

Disabled

Industries

The Context

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

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

Status: **Enacted — Signed by the President on Mar 30, 2016**
This bill was enacted after being signed by the President on March 30, 2016.

Law: Pub.L. 114-140

Sponsor:  **Paul Ryan**
Representative for Wisconsin's 1st congressional district
Republican

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

FY 2016 Significant Investments

- **2020 Census (\$663M):** We have the potential to save \$5 billion with the new 2020 Census design, however, we now have to build operations and systems for the 2020 Census, based on the new design.
- **CEDCaP (\$78M):** Smarter-IT Delivery Built on a Shared-Services Model.
- **American Community Survey (\$257M):** We must maintain the quality of the data while continuing our efforts to reduce respondent burden.
- **Geographic Support (\$81M):** We must make use of technology and partnerships to deliver smarter geographic solutions to our surveys and censuses.
- **Administrative Records Clearinghouse (\$10M):** Will expedite the acquisition of federal and federally sponsored administrative data sources, improve data documentation and linkage techniques, and leverage and extend existing systems for governance, privacy protection, and secure access to these data.
- **Economic & Government Censuses (\$144M):** Data products drive economic activity and are relevant to the needs businesses, policymakers, and the public. \$10.1 million increase

THE PROMISE OF EVIDENCE-BASED POLICYMAKING

Report of the Commission on Evidence-Based Policymaking

Transparency
Humility
Data
Privacy
Rigorous
Capacity

Administrative Data Research Facility: The Administrative Data Research Facility is a pilot project that enables secure access to analytical tools, data storage and discovery services, and general computing resources for users, including Federal, state, and local government analysts and academic researchers. The Census Bureau and academic partners developed the project as part of the collaborative Training Program in Applied Data Analytics sponsored by the University of Chicago, New York University, and the University of Maryland.¹ It is currently operating as a pilot with users accessing the Facility as part of the training program. The Facility operates as a cloud-based computing environment, with Federal security approvals, which currently hosts selected confidential data from the U.S. Department of Housing and Urban Development and the Census Bureau, as well as state, city, and county agencies, and an array of public use data.

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.

Learn more at bipartisanpolicy.org/evidence

The Foundation: Security



The ADRF Workspace

Using the Administrative Data Research Facility

SECURITY

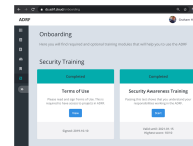
FIVES SAFES FRAMEWORK

Safe Projects

All data access is controlled on a project by project basis. Projects must be approved by the Data Owners and be consistent with the agency's mission.

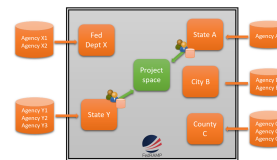
Safe People

Only approved researchers who have undergone security training are allowed to work on projects.



Safe Settings

The ADRF is hosted within AWS GovCloud and has received Authorization to Operate from the Census Bureau, has achieved a FedRAMP Moderate approval and is listed on the FedRAMP Marketplace.



Safe Data

The data in the ADRF is deidentified (hashed).

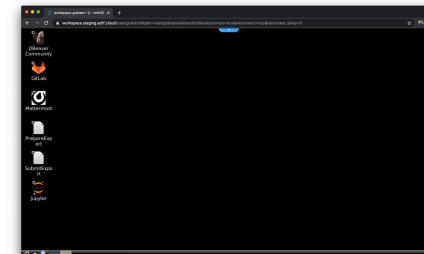
Safe Outputs

Nothing can be exported from the ADRF until it has gone through a disclosure review process.

COMPUTATION

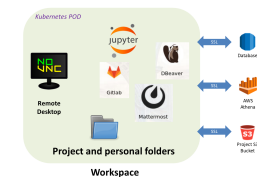
Remote Desktop Environment

The ADRF provides a remote desktop environment that is accessible from a web browser, which means no additional software is required.



Analysis Tools

A number of commonly used analysis tools are available for use, including Jupyter Lab, R Studio, and Stata, as well as Postgres and Athena for database storage and querying.



Customizable workspace

Workspaces can be customized. Projects which may require additional analytical tools, software libraries, or computing and storage are available.

An example project request.

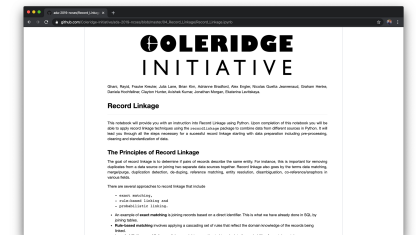
COLLABORATION

Shared Project Folder

Each project workspace includes a shared folder and database schema so that researchers can easily share work with one another inside their workspace.

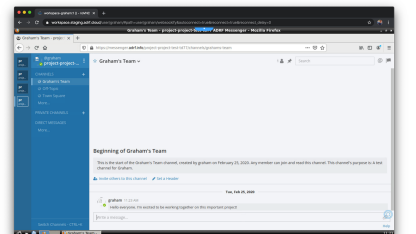
Project Templates

The ADRF has a number of Jupyter notebooks available which have many of the fundamental steps of common analysis patterns already written. Topics include databases, record linkage, text analysis, machine learning, and more.



Internal Messaging

The ADRF also provides the messaging app Mattermost (open-source Slack) so that researchers can communicate securely while inside the ADRF.



The Foundation: Stewardship



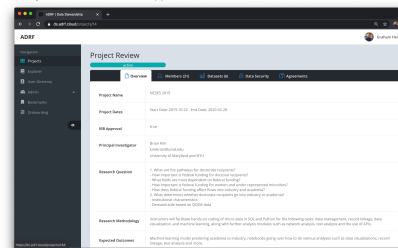
Data Stewardship in the ADRF

Managing, Monitoring, and Measuring the way datasets are used in research projects.

MANAGING

Research Project Request Workflow

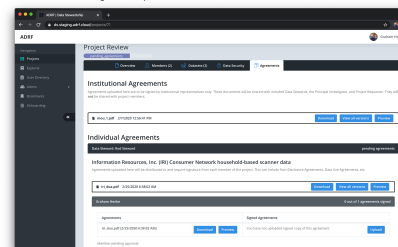
Researchers can request access to datasets via the project request workflow. The appropriate Data Stewards are notified when their datasets are requested. They can then review and approve access.



An example project request.

Data Use Agreements

After a project has been approved, Data Stewards can attach project agreements to the project, such as Data Use Agreements or NDAs, for the researchers to sign and upload.

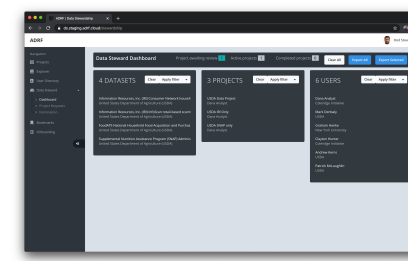


Agreements may be required between institutions and at the individual researcher level.

MONITORING

Projects and Dataset Access

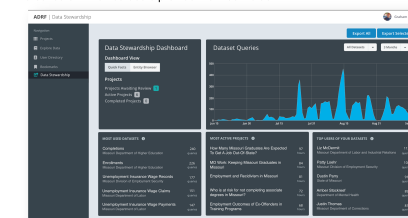
The Data Stewardship Dashboard gives Data Stewards a high-level overview as to which of their datasets researchers have access to as a part of their projects.



Datasets, Projects, and Users in the Data Stewardship Dashboard.

Dataset Usage

The Data Stewardship Dashboard will also provide high-level metrics on which datasets are most frequently used and by whom so that Data Providers can get a sense of which datasets provide the most value.

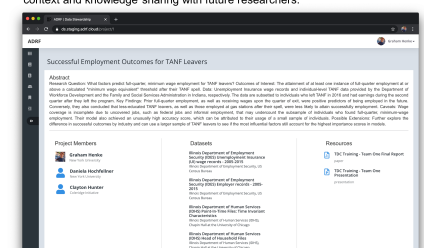


Statistics on datasets, queries, projects, datasets, and users.

MEASURING

Research Papers and Presentations

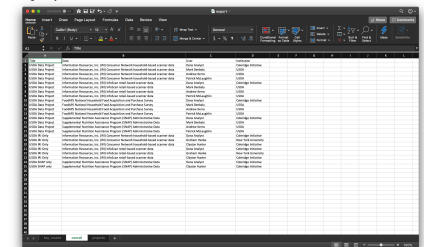
Once a research project has finished, data products such as papers or presentations can be uploaded and associated with the datasets used, providing context and knowledge sharing with future researchers.



An example of a completed project.

Export Metrics

Dataset usage information can be exported from the application so that Data Stewards can generate reports to share amongst key stakeholders at their agency.



Dataset usage metrics can be exported in common format like Excel and CSV.

The Foundation: Training

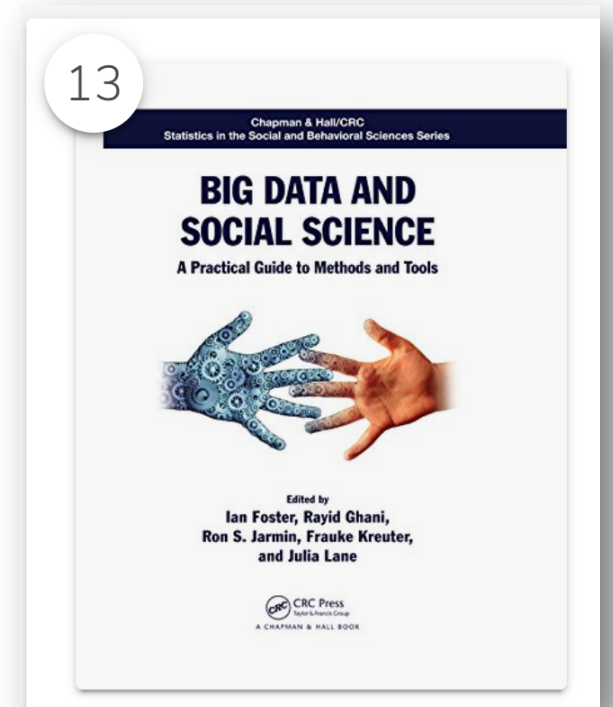
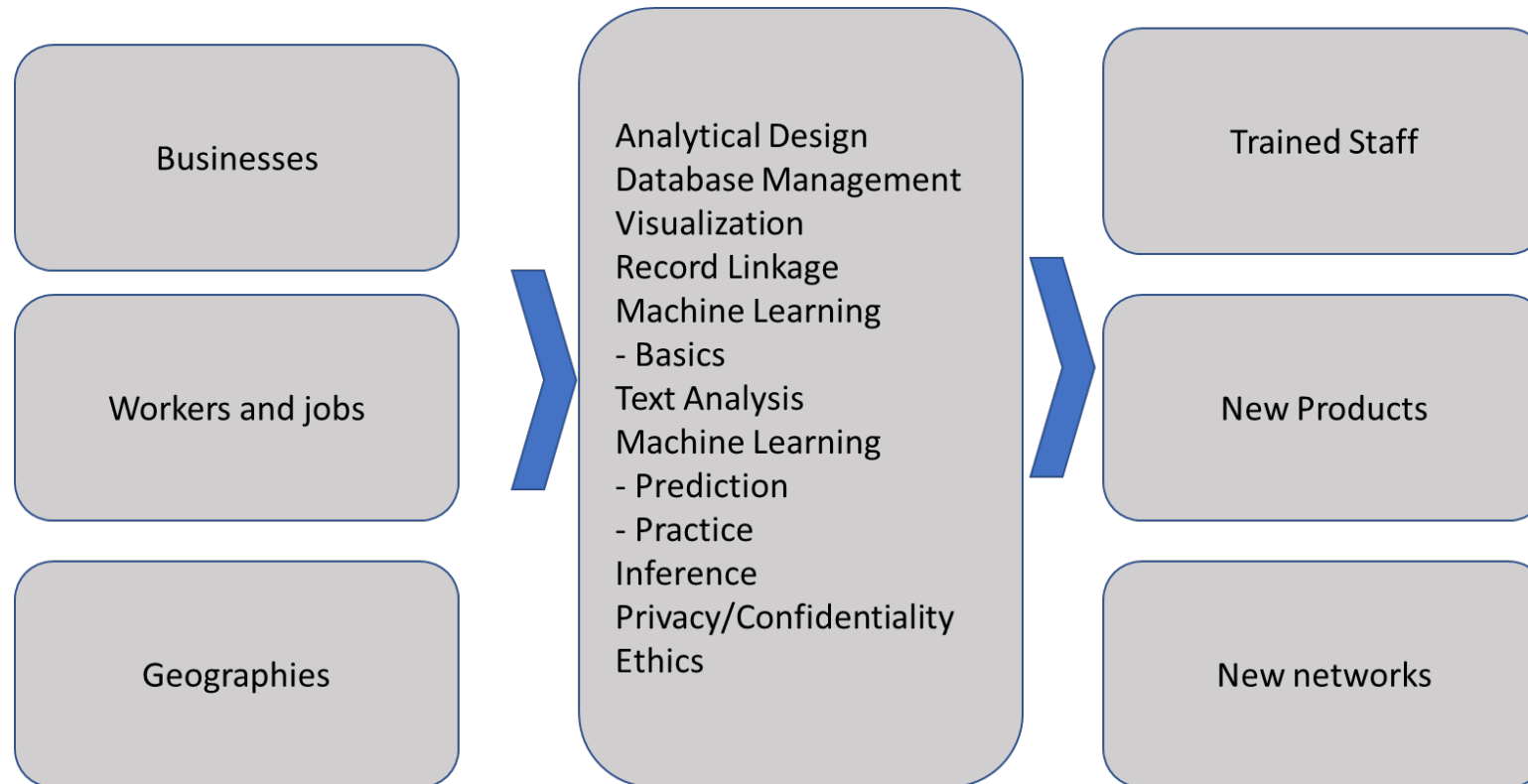




Photo by [Camylla Battani](#) on [Unsplash](#)

Q&A (use chat)





Before we go

- Stats DC cancelled
- May webinar
- Podcasts
- Give us (and each other) a shout!
- Thank you for the work you do every day.

You know where to find us!



Maureen, Kristin & the SPR/BrightHive TA Team

Maureen_Sarver@spra.com

Kristin_Wolff@spra.com

