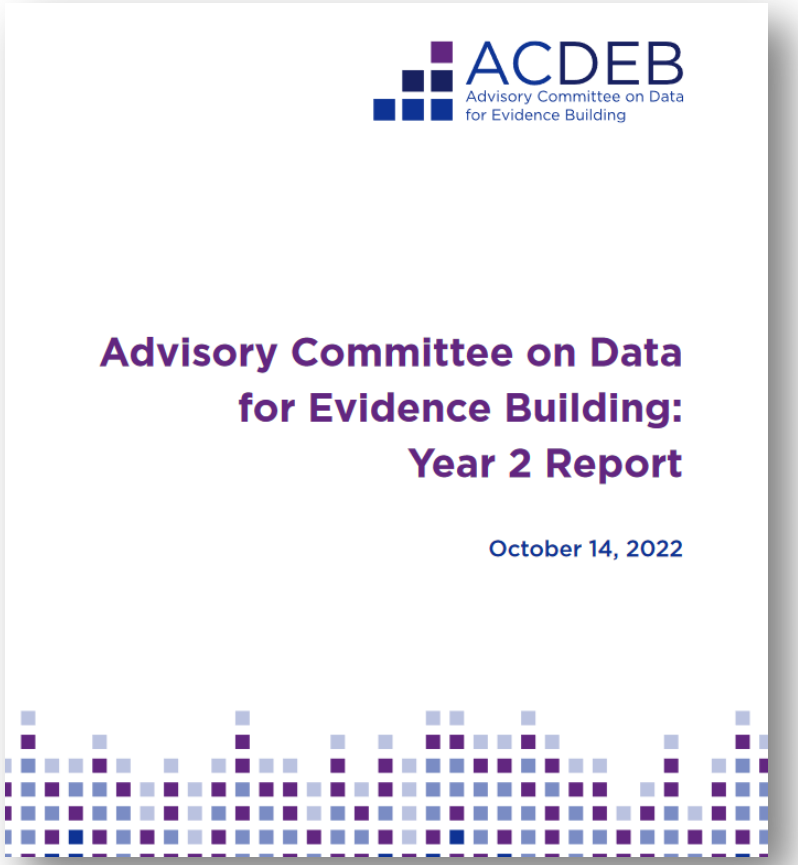
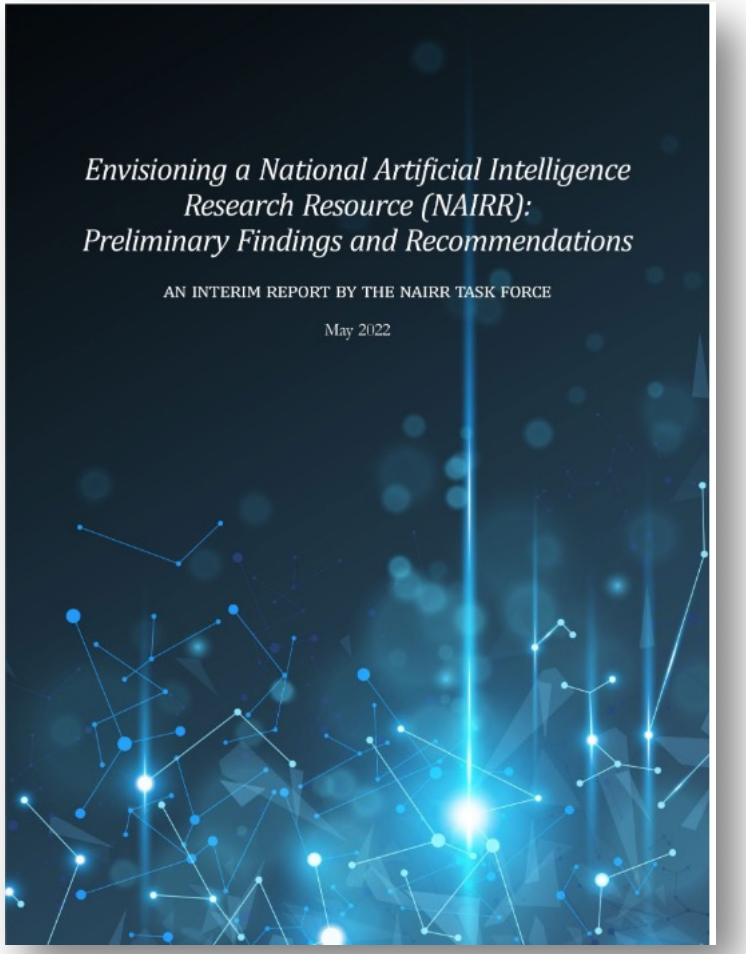
Making change happen

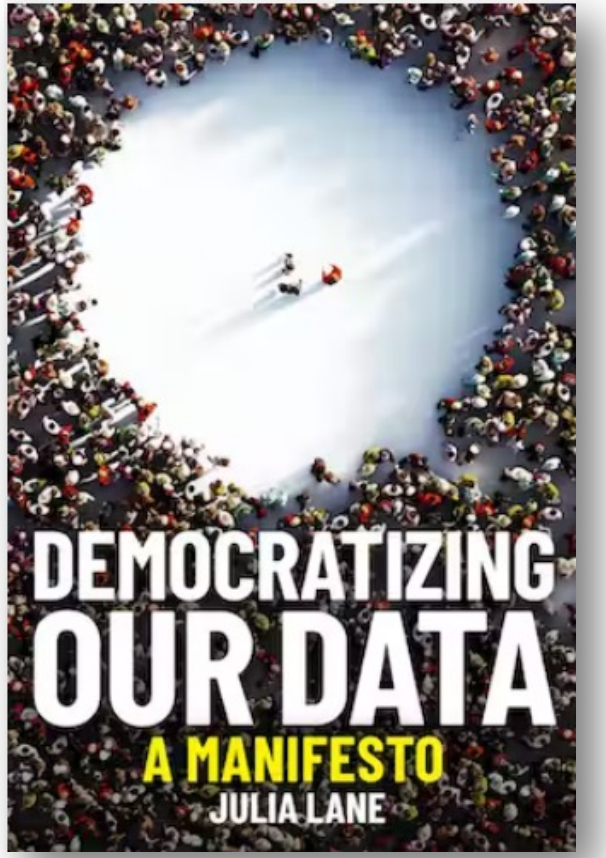
Julia Lane

New York University

Context


ACDEB
Advisory Committee on Data for Evidence Building
Advisory Committee on Data for Evidence Building: Year 2 Report
October 14, 2022


Envisioning a National Artificial Intelligence Research Resource (NAIRR): Preliminary Findings and Recommendations
AN INTERIM REPORT BY THE NAIRR TASK FORCE
May 2022


DEMOCRATIZING OUR DATA
A MANIFESTO
JULIA LANE

TL;DR

1. Massive increase in interest in data and evidence – risk and utility should be separately measured but jointly determined
2. Theory of change should be clearly described – build in evaluation from the beginning
3. High value use cases should form the basis for initial investments – establish innovation sandboxes/testbeds
4. Rethink current data infrastructures

TL;DR

1. Massive increase in interest in data and evidence – risk and utility should be separately measured but jointly determined
2. Theory of change should be clearly described – build in evaluation from the beginning
3. High value use cases should form the basis for initial investments – establish innovation sandboxes/testbeds
4. Rethink current data infrastructures

Regulation Content: Risk and Utility Framework

Recommendation 1.6. OMB should adopt a risk-utility framework as the basis for standards on sensitivity levels, access tiers, and risk evaluations as part of the regulation on expanding secure access to CIPSEA data assets.

Risk and utility should be separately measured but jointly determined. Evidence on data use should be used to inform the measurement of value, models from the public and private sectors should be used to measure risk, and the Five Safes framework should be applied to develop combined risk-utility metrics that are open and transparent. Key considerations on these aspects are described below.

Utility. Because there is limited information on data use, agencies have historically relied on a variety of methods, including standing advisory committees of expert data users, data user surveys, and literature reviews of citations, to find out how their data have been used. In addition, the lack of an automated method to search for and discover what data sets are used in empirical research leads to fundamental reproducibility challenges, threatening the utility of these data for research. The Committee discussed examples of how evidence on data use can inform the measurement of value and, by extension, be used to increase value, including ACDEB's use cases, the Democratizing Data project (also known as "rich context"), measures of conservation value in the forestry service, and the automation of research workflows in Federal Statistical Research Data Centers and elsewhere. For more information, see Appendix B. ACDEB Use Cases and the Supplemental Information—Other Models and Examples posted with this report.

Risk. The Committee explored well-tested frameworks and tools to measure and mitigate risk and examined tools to assist organizations with risk assessments. Risk assessments provide a key method to evaluate the information a data set contains while weighing and evaluating the value and benefits against potential privacy risks associated with a release. For more information, see the Supplemental Information—Other Models and Examples posted with this report.

Combining risk and utility. The Five Safes framework provides an appropriate approach to operationalizing value and risk metrics. Box 3 provides high-level considerations of this framework.

Statistical agencies can demonstrate value to data providers and users by enabling safe projects, people, and data.

- ***Safe projects.*** Statistical agencies can identify high-value projects that advance the goal of evidence-based policymaking by working closely with federal programmatic agencies and state and local government agencies. Utility measures can be identified by the agency, researchers, and stakeholders and validated through automated measures of data usage.
- ***Safe people.*** Agencies can increase the size of the skilled workforce available to produce evidence through standards for accrediting safe researchers in a transparent and accountable manner. Learning agendas can be expanded to include more hands-on training in privacy and confidentiality.
- ***Safe data.*** Agencies can increase access to safe data—both to safe confidential microdata and to safe summary data.

Statistical agencies can mitigate risks using restricted data by implementing the Five Safes framework, particularly involving safe people, settings, data, and outputs.

- ***Safe people.*** Agencies can institutionalize and operationalize the concept of safe people by creating open and transparent access policies that specify who can read or edit data, for what purposes, and how derived data products may be shared. Indemnification could be expanded to include analysts accessing confidential summary tabulations, as has been implemented by the Midwest Collaborative.
- ***Safe settings.*** Agencies can enable more modes of data access through secure remote environments and privacy-preserving technologies.
- ***Safe data.*** Agencies can offer tools and technical assistance to produce safe data, including hashing algorithms and guidance on producing synthetic data.
- ***Safe outputs.*** Agencies can support production of safe outputs, including confidential summary tabulations, synthetic data, traditional statistical disclosure limitation techniques, and evolving methods such as differential privacy.

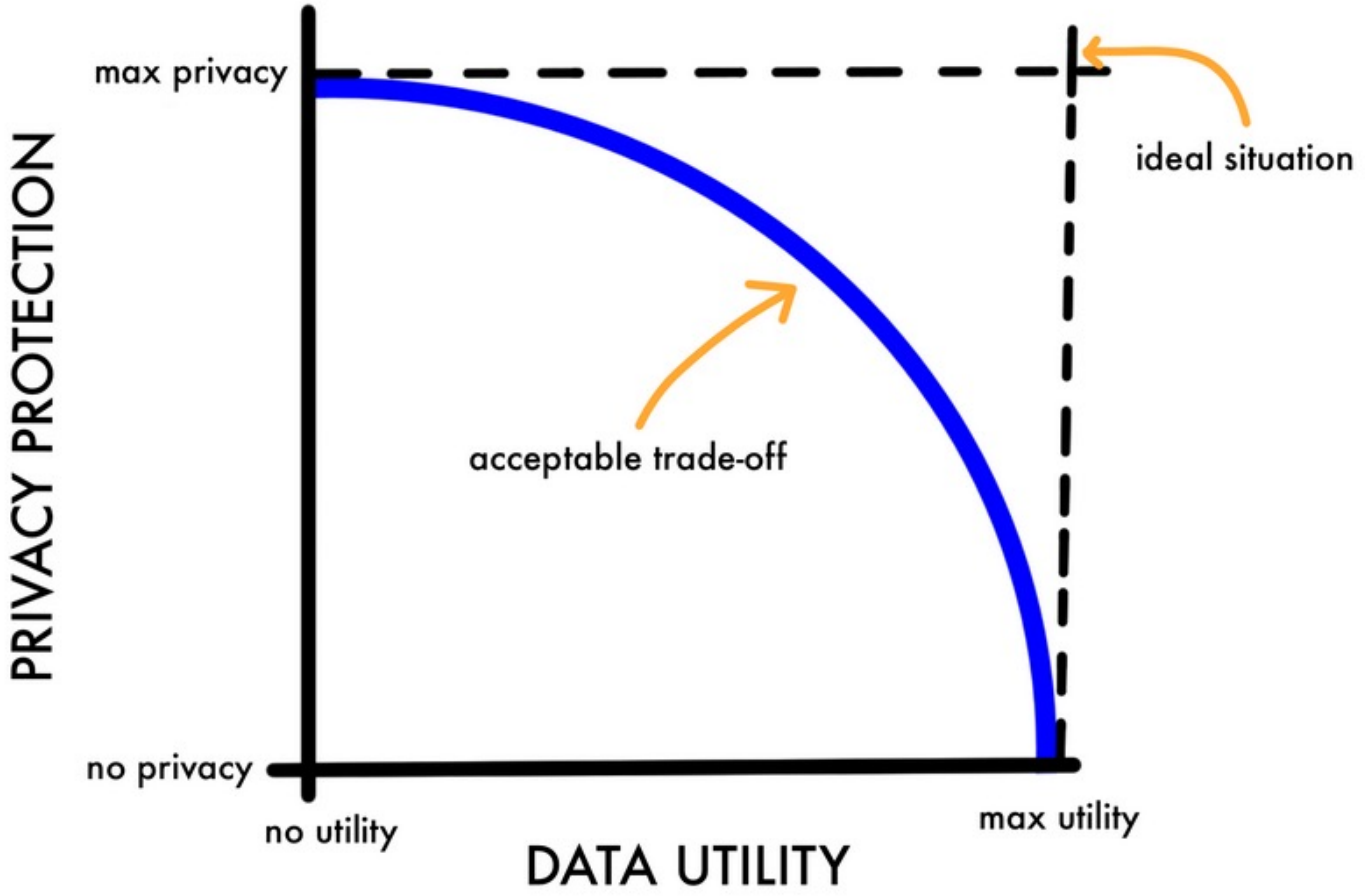
Regulation Content and Implementation Strategy: Tiered Access in Practice

During its second year, the Committee engaged with a wide variety of experts, including from the ICSP and other entities inside and outside the federal government, to understand (1) the need for better data access and (2) options to make data more accessible. OMB should use these insights to inform guidance on tiered access as part of the regulation on expanding secure access to CIPSEA data assets.

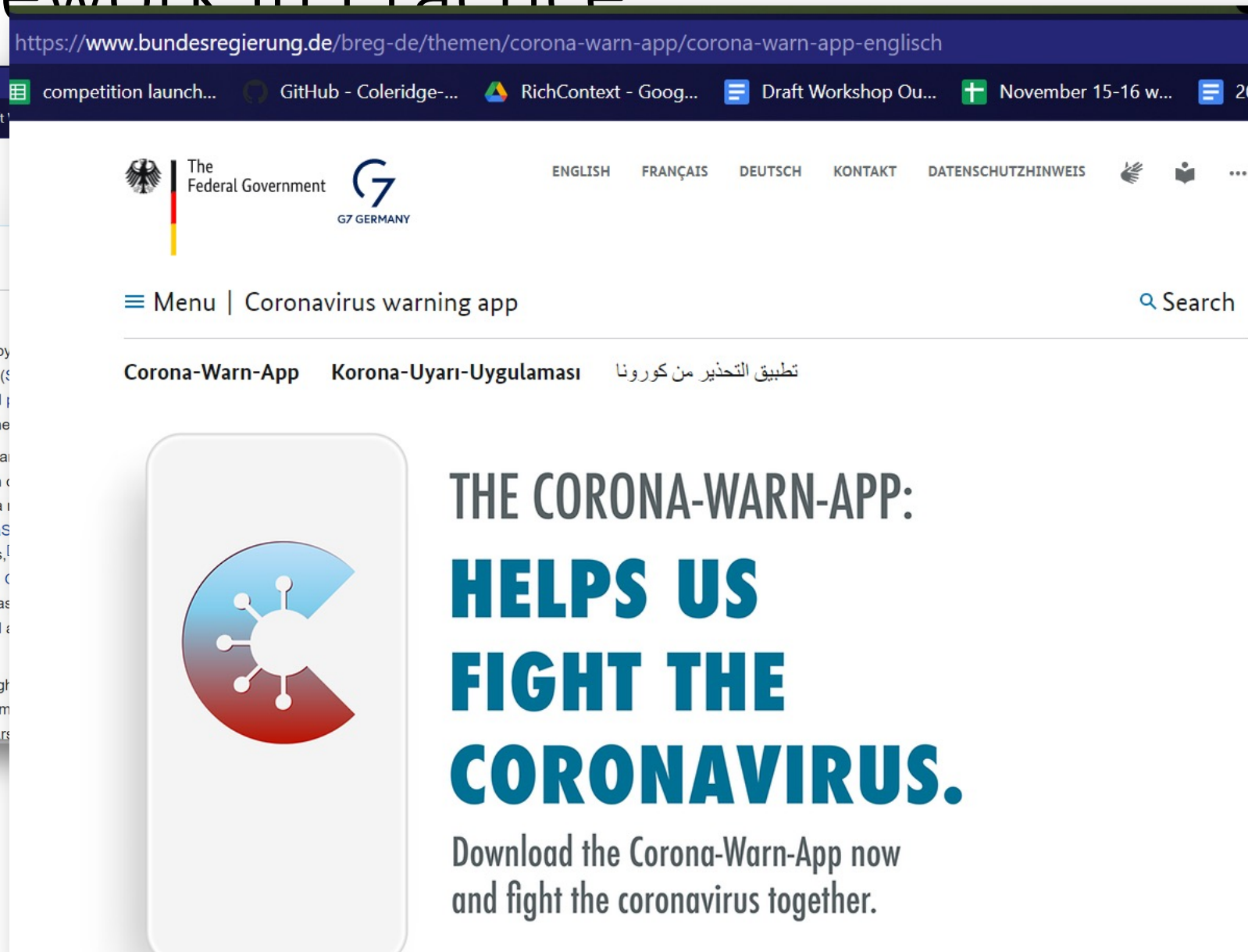
These experts expressed many needs and challenges related to better data access, including the following:

- ***Needs.*** Needs include the ability to access timely, relevant data with sufficient granularity to inform policy development and decisionmaking at all levels of government and in the private sector. This often requires bringing together data from multiple sources—federal, state, local, and private—for purposes ranging from monitoring the local or regional impacts of economic, health, or environmental conditions to highly technical analytical analyses, including those using artificial intelligence approaches.

Risk-Utility Framework



Risk-Utility Framework in Practice



TL;DR

1. Massive increase in interest in data and evidence – risk and utility should be separately measured but jointly determined
2. Theory of change should be clearly described – build in evaluation from the beginning
3. High value use cases should form the basis for initial investments – establish innovation sandboxes
4. Rethink current data infrastructures

ACDEB
Advisory Committee on Data for Evidence Building

Advisory Committee on Data for Evidence Building: Year 2 Report

October 14, 2022

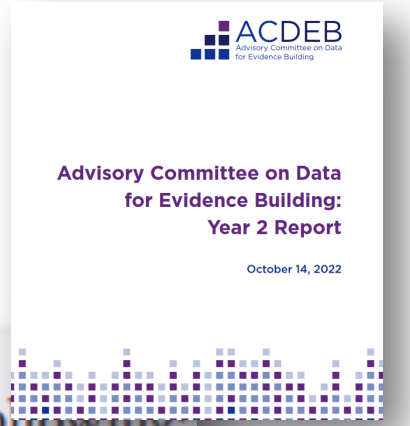
Box 8. NSDS Theory of Change

The NSDS theory of change can be broken into the five discrete categories shown in Figure A and described below.

Figure A. The Theory of Change

Inputs. Resources at the disposal of the project, including statistical and program staff, compute facilities, existing research, and budget.

Activities. Actions taken or work performed to convert inputs into outputs, specifically, the NSDS core functions—coordination, communication, research and development (R&D), and data standardization.



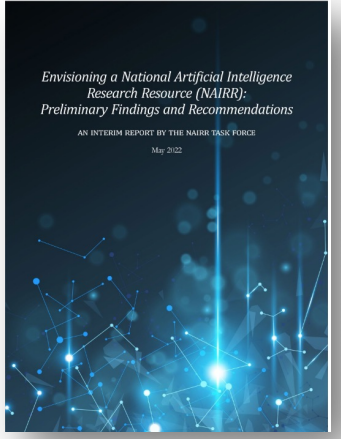
Outputs. The tangible goods and services that the project activities produce, including the following:

- Producing value for the American public while protecting privacy and confidentiality;
- Coordinating and supporting evidence-building efforts that cut across entities;
- Facilitating linkage of, secure access to, and analysis of nonpublic data;
- Providing capacity-building services for data users, data providers, and related communities of practice;
- Communicating the value and use of data for evidence building and how data are protected;
- Facilitating R&D and the adoption of practices and methods that enhance privacy and confidentiality and improve record linkage; and
- Fostering and promoting data standardization to enable more efficient and high-quality linkage, access, and analysis.

Outcomes. Results likely to be achieved once the beneficiary population uses the project outputs, including new evidence and products for decisionmaking, lower costs and higher quality evidence, greater transparency and accountability, communities of practice, and better collaboration across levels of government.

Impacts. The results achieved indicating whether project goals were met, including better decision-making; more timely, actionable, and policy-oriented research; policies that are more responsive to local conditions; and more effective local interventions.

In much of the discussion about evidence-based policymaking and the NSDS, the theory of change has been cut short, where the focus on inputs and activities stops with the endpoint of interest exclusively on outputs. The Committee encourages the OMB Director to work with agencies, the ICSP, the Evaluation Officer Council, the Chief Data Officer Council, and others to emphasize the value of short-term, intermediate, and long-term outcomes in the NSDS theory of change in relationship to the broader evidence ecosystem.



Recommendation 3-15: NAIRR evaluation methods, including definition of metrics and indicators of success for the NAIRR, should be grounded in established best practices.

Successful program design requires a clear understanding of the activity's overall near- and long-term goals and a conceptual model for how progress toward achieving them will be measured. Such a so-called logic model defines the underlying assumptions of how the program would work, its desired outcomes and impacts, the activities and outputs that would support these outcomes, the resources or inputs needed to achieve the outcomes, how all these elements work together to achieve desired impacts, and the metrics that can be used to characterize each.[24] The entities responsible for evaluating the NAIRR should adopt this approach to assessing its performance, similarly informing the evolution of the NAIRR over time.

TL;DR

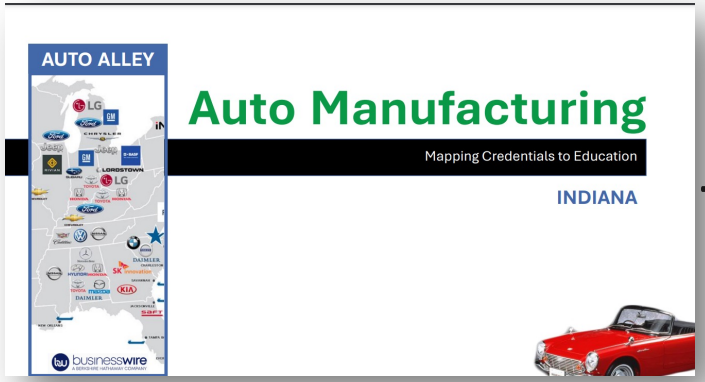
1. Massive increase in interest in data and evidence – risk and utility should be separately measured but jointly determined
2. Theory of change should be clearly described – build in evaluation from the beginning
3. High value use cases should form the basis for initial investments – establish innovation sandboxes/testbeds
4. Rethink current data infrastructures

ACDEB
Advisory Committee on Data for Evidence Building

Advisory Committee on Data for Evidence Building:

Year 2 Report Supplemental Information

October 14, 2022



f Va

We define success as our students enrolling in a postsecondary program, completing a credential, and securing employment that sets them up for economic mobility in Chicago

Our three key success metrics

Postsecondary enrollment	Percentage of students who **enroll directly in a certificate, two-year or four-year program after high school**
Postsecondary completion	Percentage of students who **earn a postsecondary credential** (certificate, two-year degree or four-year degree) **within 150 percent time**
Employment and economic mobility	Metric to be developed between CPS and CCC

Democratizing Data Challenge: Illinois partnership with Coleridge Initiative

- Visualizations for local decision makers on unemployment spell behavior by education categories of local claimants
  - Impact of unemployment spell duration on reemployment probability by education levels
  - Equity groups- low wage workers, Black and Hispanic/Latinx workers, women, and individuals living in rural areas
- Timely measures of reemployment activity
  - Program performance
    - Quarterly UI Wage Records
  - Evaluate alternative data sources for timely allocation of local resources
    - Illinois Monthly UI Wage Records, Directory of New Hires
- Applied data analytics training
  - Equity challenges in claimant reemployment by level of education/training
    - Extension of 3-month training in 2021 funded by USDOL/ETA; 120 participants, 30 states
  - Partnership among MidWest, Eastern, and Southern Regional Collaboratives in cooperation with local universities
  - Audience: staff of employment services and workforce agencies

Expanding on the Multi-State Postsecondary Report Dashboard to Include Workforce Credentials and Students Without Degrees

Ohio Education Research Center, John Glenn College of Public Affairs

Unemployment to Reemploym… Portal: Common Data Model

…al dynamic of unempl…
… individual claimant
claims … weekly flows
…imant measures
…k when a claimant files a new
…ants who share the same ben…
…city
…fferent phases of rollout

Up Next:

Products and Practice: Facilitated Conversations and Priority Setting

Breakout 1	**Unemployment to Reemployment: New Measures and Confidential Summary Tabs**
Breakout 2	**Postsecondary Transitions to Employment: Building Common Data Models and Products**
Breakout 3	**Training Practitioners: Human Services**
Breakout 4	**Training Practitioners: Work-based Learning**
Breakout 5	**The Elements of State-led Governance**

CHAPIN HALL
AT THE UNIVERSITY OF CHICAGO

Newsroom Contact Us

ABOUT US IMPACT AREAS OUR WORK EXPERTS CAREERS SUPPORT US

Report

Illinois Childcare Workers Experienced Employment Interruptions During Early Months of the COVID-19 Pandemic

2022

By Robert Goerge, Emily Wiegand, David McQuown

https://kystats.ky.gov/Reports/Tableau/2022_MSPSR

Multi-State Postsecondary Report

COLERIDGE INITIATIVE

Employment by Institution | Employment by Demographic | Employment by Major | Technical Notes

Filter Report by State: Kentucky Postsecondary Completers

This dashboard includes three sections: Employment Outcomes by Major Group, Credential, and High School Origin. Each of these sections are filterable by Institution and Years Post Completion. Wage information can be filtered by percentile group using the wage filter. Employment outcomes looks at Kentucky Postsecondary Completers across the 2007 through 2017 academic years (AY) followed to post completion employment in Indiana, Kentucky, Ohio, and Tennessee. A double asterisk in a data table represents redacted values, blank data represents no data available.

Filter the Dashboard by Institution ▼: All

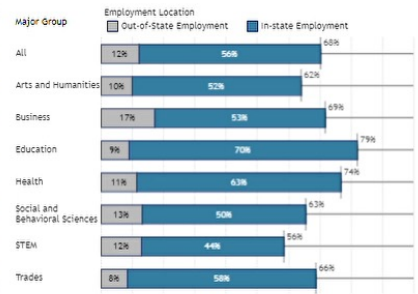
Filter the Dashboard by Years Post Completion ▼: 3-Years Post Completion

Filter Wages by Percentile ▼: Median Wage

Employment Outcomes by Major Group

This section provides employment outcomes for each major group. Major groups are categories that each contain several majors. Find more information in the technical notes.

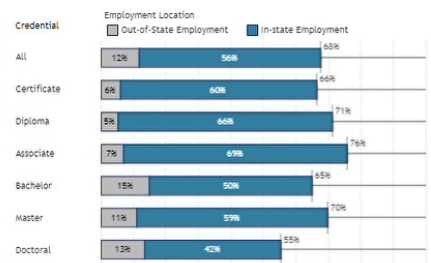
Major Group	Completer Count	In-State Median Wage	Out-of-State Median Wage
All	309,720	$40,601	$43,744
Arts and Humanities	50,251	$29,522	$31,837
Business	53,200	$42,097	$48,806
Education	38,290	$48,704	$42,228
Health	66,226	$44,068	$50,885
Social and Behavioral Sciences	45,003	$35,215	$38,460
STEM	34,603	$45,628	$51,778
Trades	22,147	$36,035	$39,598

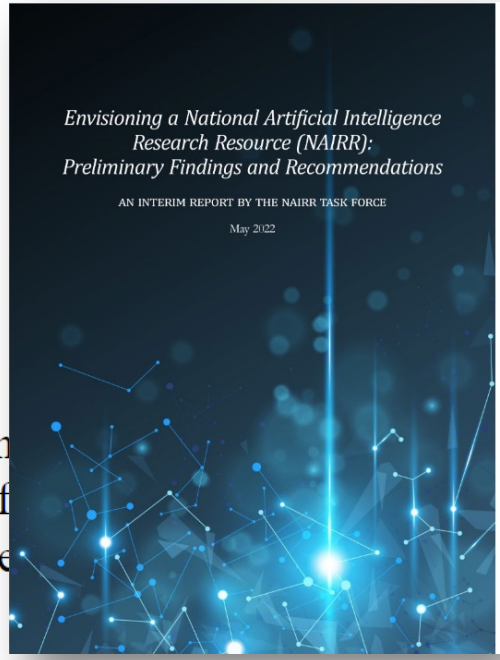


Employment Outcomes by Credential

This section provides employment outcomes for each credential level.

Credential	Completer Count	In-State Median Wage	Out-of-State Median Wage
All	309,720	$40,601	$43,744
Certificate	30,471	$23,497	$25,306
Diploma	5,133	$32,619	$34,434
Associate	53,350	$36,646	$39,135
Bachelor	143,920	$38,039	$41,247
Master	62,155	$52,491	$54,599
Doctoral	14,691	$70,197	$66,561





Testbeds

AI testbeds are simulated, live, or blended environments that support research, prototypin development, and testing of AI applications that are robust and trustworthy. The concept of testbed can encompass the environment itself—hardware and software—as well as the data se and frameworks that support evaluation, and the talent needed to manage the resource.

Testbeds can accelerate AI research by providing virtual or physical environments to test, simulate, explore, and develop AI. They can spur innovation in specific areas, provide opportunities to benchmark and check the quality of research, and foster cross-disciplinary collaboration.[53] Expanding efforts to centrally catalog AI testbeds[54] could increase accessibility to researchers and support Federal efforts to assess testbed needs. The NAIRR has an opportunity to act as a hub, cataloging and making available existing testbeds and test sets to accelerate research, increase access to inspiring testbeds, and broaden participation in AI.

Finding 4-12: Testbeds can support quality assessment of AI.

AI researchers have previously evaluated the performance of their new AI technologies on internal or proprietary data, making claims of improved performance difficult to measure or replicate. Standardized competitions can help prevent this occurrence.

Finding 4-13: Testbeds can inspire participation in AI research.

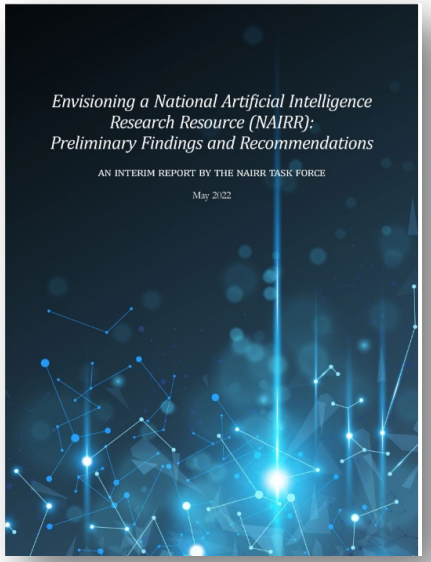
If students from kindergarteners through PhD candidates hear primarily about uninspiring, harmful, or problematic AI use cases (e.g., advertising technology, surveillance, or social media manipulation), there is a risk of deterring a diverse group of minds from considering working in or supporting the field of AI. Larger public works projects to improve safety, education, health, and equity might increase enthusiasm for working in the field. Testbeds such as those created to support the DARPA Autonomous Vehicle Challenge or various public health competitions can be helpful.

TL;DR

1. Massive increase in interest in data and evidence – risk and utility should be separately measured but jointly determined
2. Theory of change should be clearly described – build in evaluation from the beginning
3. High value use cases should form the basis for initial investments – establish innovation sandboxes/testbeds
4. Rethink current data infrastructures

Reconsider Data Collection

The full value of AI is often not realized without high quality, trusted, dense, transparent data

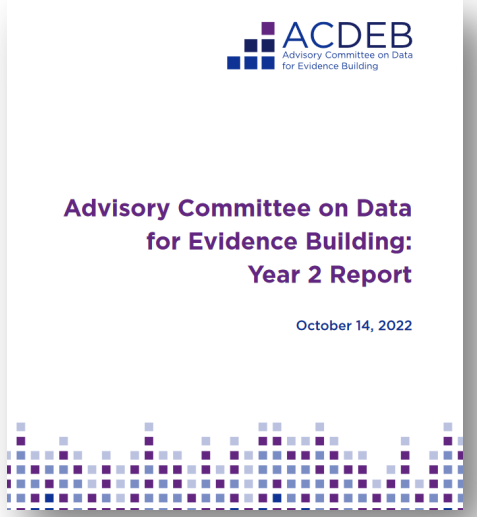


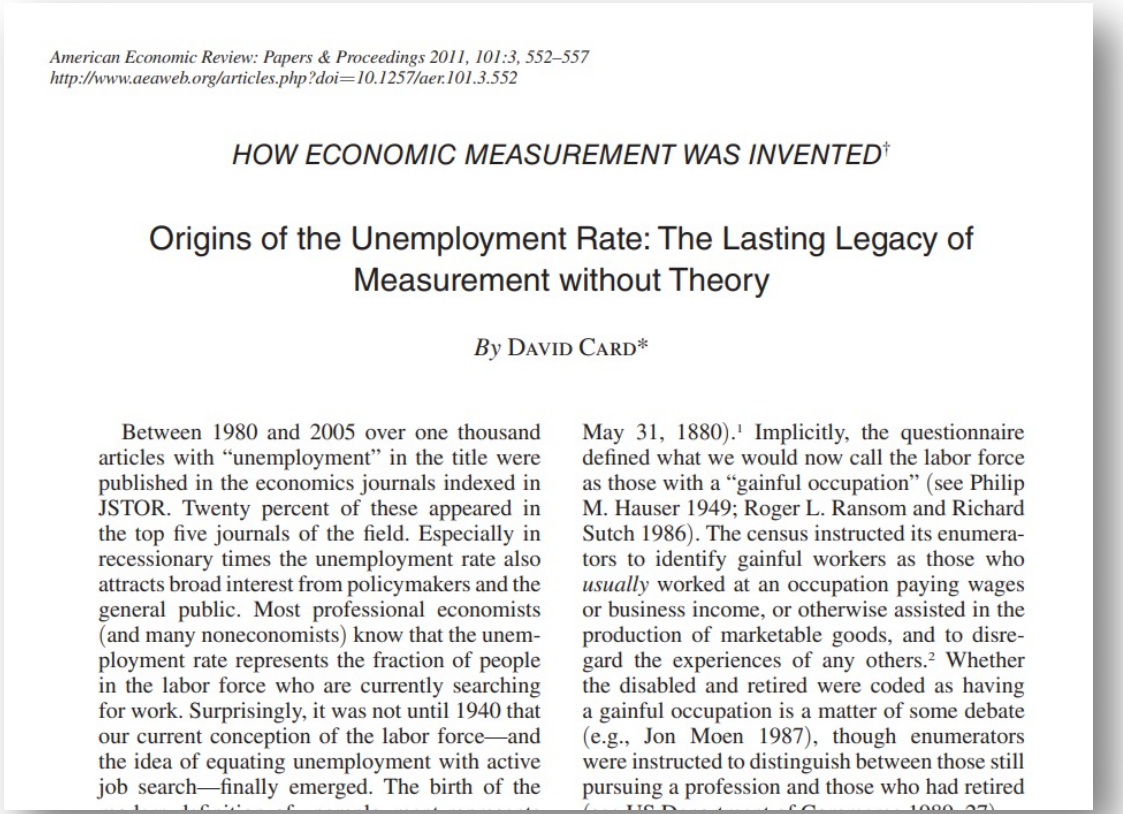
For example

1. In social data, badly trained criminal justice algorithms and lack of transparency can lead to social harm
2. For customer care data, underrepresentation of certain demographic groups in training data leads to false positives and resultant inequitable treatment
3. In natural hazards data, sparse data leads to poor simulations of the impact of hurricane surges
4. For self-driving cars, sparse data on rare events can have significant negative impact

Reconsider construct measurement Employment

- Labor market activity (stability, employer match; coverage)
  - Employment
  - Unemployment
  - Reemployment
- Labor market outcomes
  - Earnings
  - Number of jobs
- Organizational framework
  - Firms
  - Industries
  - Occupations
  - Geography




American Economic Review: Papers & Proceedings 2011, 101:3, 552–557
http://www.aeaweb.org/articles.php?doi=10.1257/aer.101.3.552

HOW ECONOMIC MEASUREMENT WAS INVENTED†

Origins of the Unemployment Rate: The Lasting Legacy of Measurement without Theory

By David Card*

Between 1980 and 2005 over one thousand articles with "unemployment" in the title were published in the economics journals indexed in JSTOR. Twenty percent of these appeared in the top five journals of the field. Especially in recessionary times the unemployment rate also attracts broad interest from policymakers and the general public. Most professional economists (and many noneconomists) know that the unemployment rate represents the fraction of people in the labor force who are currently searching for work. Surprisingly, it was not until 1940 that our current conception of the labor force—and the idea of equating unemployment with active job search—finally emerged. The birth of the

May 31, 1880).[1] Implicitly, the questionnaire defined what we would now call the labor force as those with a "gainful occupation" (see Philip M. Hauser 1949; Roger L. Ransom and Richard Sutch 1986). The census instructed its enumerators to identify gainful workers as those who *usually* worked at an occupation paying wages or business income, or otherwise assisted in the production of marketable goods, and to disregard the experiences of any others.[2] Whether the disabled and retired were coded as having a gainful occupation is a matter of some debate (e.g., Jon Moen 1987), though enumerators were instructed to distinguish between those still pursuing a profession and those who had retired

TL;DR

1. Massive increase in interest in data and evidence – risk and utility should be separately measured but jointly determined
2. Theory of change should be clearly described – build in evaluation from the beginning
3. High value use cases should form the basis for initial investments – establish innovation sandboxes/testbeds
4. Rethink current data infrastructures