



11-2018

# Data-Driven Collective Impact: Driving Social Change as a Community

Natalie Evans Harris

*BrightHive*

Follow this and additional works at: [https://repository.upenn.edu/admindata\\_conferences\\_presentations\\_2018](https://repository.upenn.edu/admindata_conferences_presentations_2018)

---

Evans Harris, Natalie, "Data-Driven Collective Impact: Driving Social Change as a Community" (2018). *2018 ADRF Network Research Conference Presentations*. 17.

[https://repository.upenn.edu/admindata\\_conferences\\_presentations\\_2018/17](https://repository.upenn.edu/admindata_conferences_presentations_2018/17)

This paper is posted at ScholarlyCommons. [https://repository.upenn.edu/admindata\\_conferences\\_presentations\\_2018/17](https://repository.upenn.edu/admindata_conferences_presentations_2018/17)

For more information, please contact [repository@pobox.upenn.edu](mailto:repository@pobox.upenn.edu).

---

# Data-Driven Collective Impact: Driving Social Change as a Community

## **Abstract**

The last decade (or more) has experienced a transformation of data to an action-oriented asset that can draw insights necessary to describe, detect, predict, and evaluate factors to help our communities and the individuals in them to thrive. We've also witnessed threats to these opportunities in the forms of breaches, misinformation, and other erosions of trust that make access to and use of data much more complicated. As a community, it is imperative to take an interdisciplinary approach to data use grounded in public-private collaboration and focused on building trust with the communities we seek to serve.

# Data-Driven Collective Impact: Driving Social Change as a Community

---

**Natalie Evans Harris**

Co-Founder and COO, **BrightHive**

Co-Ideator, **Community-driven Principles for Ethical Data Practices (CPEDS)**

Former Senior Policy Advisor to US Chief Technology Officer, **Obama Administration**

---

# Who Am I?

*I believe that through our **collective** power, data can transform the human experience*

## Technology

Co-Founder and COO,  
**BrightHive**



## People

Co-Ideator, **Global Data Ethics Project (GDEP)**



## Capacity-building

Former Senior Policy Advisor to  
US Chief Technology Officer,  
**Obama Administration**



[www.NatalieEvansHarris.com](http://www.NatalieEvansHarris.com)

@QuietStormnat

LinkedIn: [nevansharris](https://www.linkedin.com/in/nevansharris)



# Why Digital Transformation? The World is Changing

## Citizen-centered impact focuses on:

- **The Journey:** An **end-to-end digital experience** developed from the customer's point of view, accessible anywhere, anytime, and from any device.
- **Protect Access:** A **unique, uniform digital ID** that grants agencies access to the appropriate data and services.
- **Cross-Agency Collaboration:** Mechanisms that allow agencies to **share data** across the state enterprise.

\*Deloitte Insights, Delivering the Digital State

## Technical Barriers

**Data Silos** Within and Across Verticals

Lack of Internal **Analytics Capacity**

Lack of **Standard** Metrics and Algorithms



## Cultural Barriers

**Policies** Limiting Data Interoperability

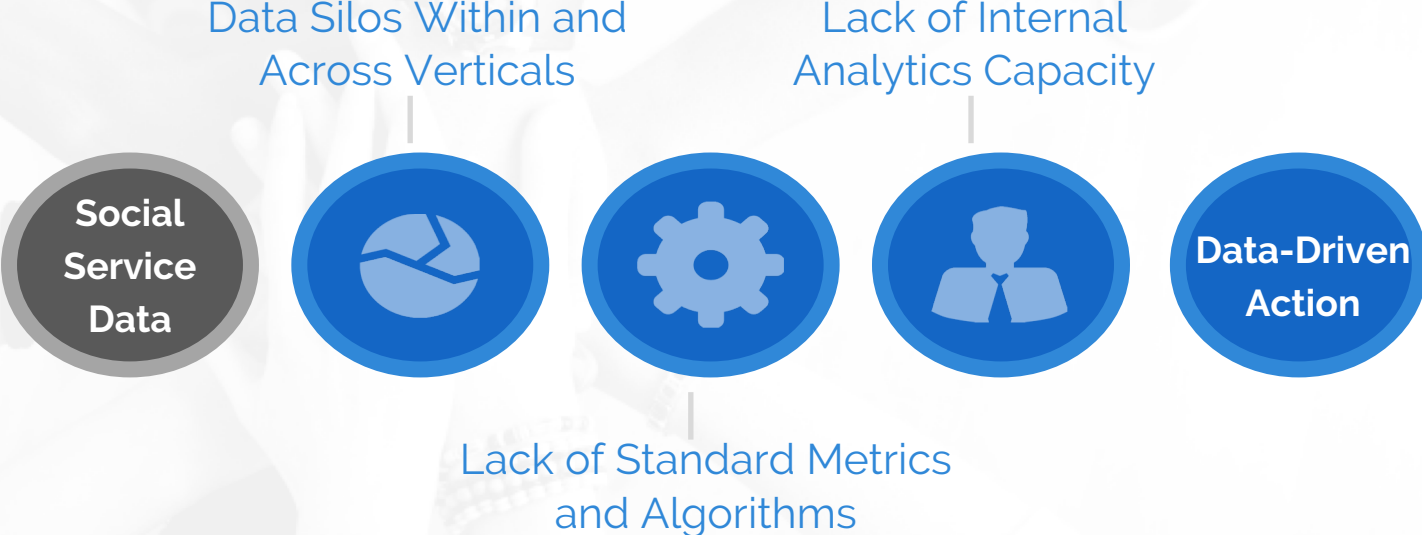
Lack of **Talent** Pipeline

Lack of Dedicated **Funding**

Limited Pathways for **Sustainment**

# Challenges to using data for social good

# Current Data Infrastructure for Identifying, Improving, and Investing in Impact



# Future Infrastructure must be Open and Collaborative





## Balancing Individual Rights With Social Impact

Building Trust requires:

Empowering **People** to speak on data use,  
establishing **Processes** to clarify practices,  
and leveraging **Technology** to hold us accountable

***We must define the "common good" and begin again to shape a common future.***

*Rep. Barbara Charline Jordan of Texas  
1976 Democratic National Convention Keynote Address*



# Global Data Ethics Project

## By Data Scientists, For Data Scientists:

Increases responsiveness to the needs and concerns of data scientists

Better captures the diverse spectrum of interests across the data science community

May facilitate adoption of the code of ethics

**Bloomberg**<sup>®</sup>



\*\*Formerly Community-driven Principles for Ethical Data Sharing (CPEDS)

# As a Community, What's Important Is...

Informed and purposeful **consent**

**Protect** anonymous data subjects

Foster **diversity**

Clearly established **provenance**

**Communicate** responsibly and accessibly

Guarantee the **security** of data, subjects, and algorithms

**Transparency** as the default

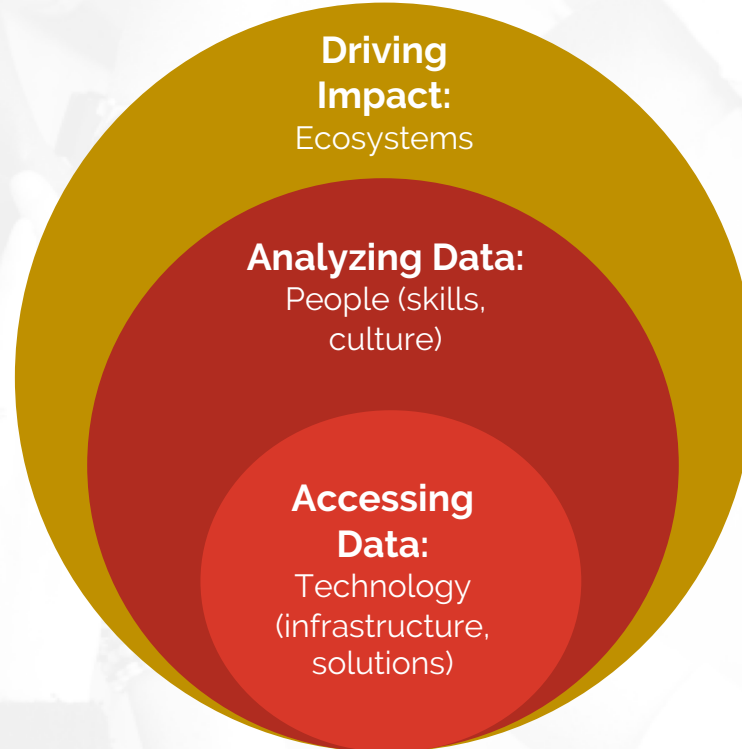
Acknowledge and mitigate unfair **bias**

**Respect** relevant tensions

**Exercise Our Ethical Imagination!!**

<https://www.datafordemocracy.org/project/global-data-ethics-project>

# Public-Private Collaboration





# Technology - Civic Data Infrastructure

Collaborated with AISP, MetroLab and Casey Foundation to answer this question:

What are the **key considerations in building and sustaining civic data infrastructures** and the various technology approaches that may be helpful in overcoming challenges in data integration?

A digital version of this report with resource web links can be viewed and downloaded as a PDF here: <http://bit.ly/2xaQW2L>

# Pain Points

## Lack of Agility

- **Dependence on vendor** to run inquiries - Procurement processes should require data be interoperable and accessible
- Support linking new streams

## Lack of Standards

- Lack of **data descriptions** the further up the data lifecycle

## Lack of Government Alignment

- Shared **procurement** to reduce costs and increase interoperability
- Improved **governance** to streamline legal and political challenges
- Reduce **duplication** of efforts

# Technology for Civic Data Integration

The purpose of this report is to **describe key considerations in building and sustaining IDS** and the various technology approaches that may be helpful in overcoming challenges in data integration.

**Consideration 0:** Staffing Expertise  
Data Collection

**Consideration 1:** Data Management  
Data Storage

**Consideration 2:** Security and Privacy  
Data Linking

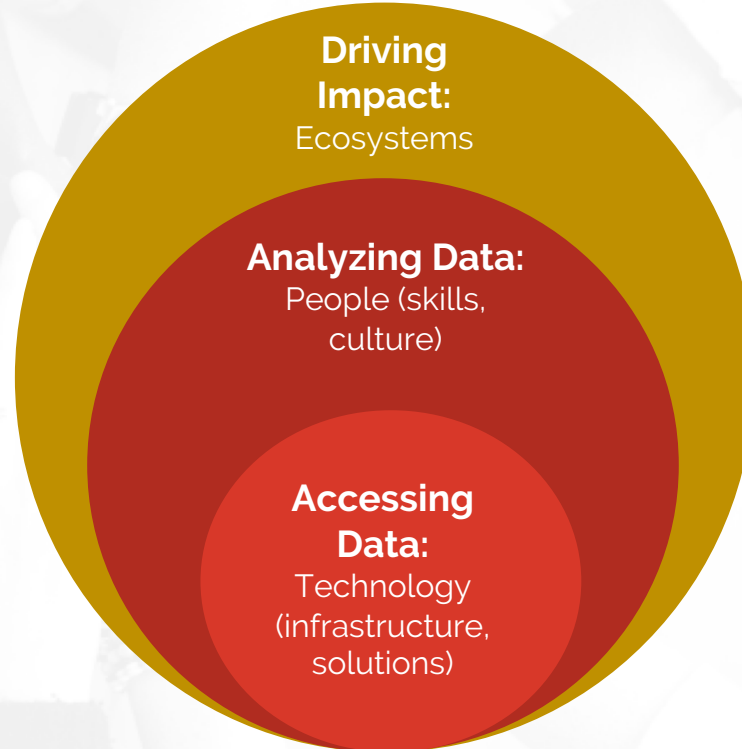
**Consideration 3:**

**Consideration 4:**

**Consideration 5:**

**Consideration 6:** Data Access and Dissemination

# Public-Private Collaboration



# People Make Data Meaningful



Building public sector capacity to **turn data into valuable insights** that enable informed action

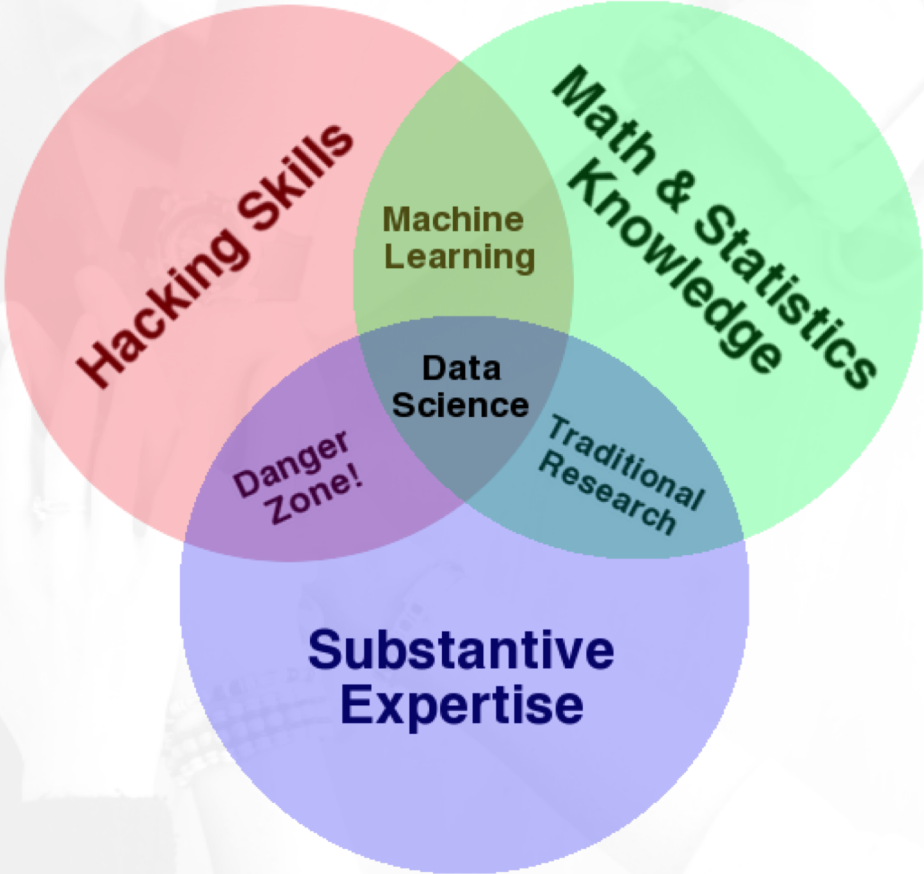
A background image showing several hands of different skin tones reaching towards the center, holding a globe. The image is faded and serves as a backdrop for the text.

# Data Literacy is the Ability To Read, Create, and Communicate Data as Meaningful Information

## Data literacy program:

- must be agile and adaptive
- must view data literacy as a continuum
- must empower people

# Critical Skills



# The Critical Skills for the Big Data Future Look Like the Skills of the Little Data Past

- \* Ability to connect people and organizations
- \* Ability to build and lead cross-sector coalitions
- \* Ability to communicate clearly and sell the big vision
- \* P-values: Patience, Persistence, Perseverance
- \* Seeing the opportunities through the risks
- \* Openness benefits all



# The Mythical Data Scientist





\*Data is a  
Diverse  
Team Sport

# Ethical Culture

**Self sovereignty and informed consent:** Empowers individuals to control their own data and determine its uses.

**Cooperation:** Promotes collaboration between people and institutions

**Transparency & Openness:** The origins and ownership are clear and workings are intelligible to non experts; information defaults to being open and free.

**Decentralization:** Ownership, production, and control are distributed and driven by a community; default to open source.

**Flexibility:** Easy for users to modify, adapt, improve, or inspect its core; Individuals and institutions may freely choose to use it or give it up

**Redundancy:** More than one solution to every data and technology problem. No monopolies or “one platform to rule them all”

**Efficiency:** Minimizes new resource requirements and personnel costs to realize impacts



# Ethical Practices - Data Science Oath

**I recognize that data science** has material consequences for individuals and society, so no matter what project or role I pursue, I will use my skills for their well-being.

**I will consider the privacy, dignity and fair treatment** of individuals when selecting the data permissible in a given application, putting those considerations above the model's performance.

**I have a responsibility to bring data transparency, accuracy and access** to consumers, including making them aware of how their personal data is being used.

**I will act deliberately to ensure the security of data** and promote clear processes and accountability for security in my organization.

**I will invest my time and promote the use of resources in my organization** to monitor and test data models for any unintended social harm that the modeling may cause.

# Ethical Design Checklist

- Have we listed how this technology can be attacked or abused?
- Have we tested our training data to ensure that it is fair and representative?
- Have we studied and understood possible sources of bias in our data?
- Does our team reflect diversity of opinions, backgrounds, and kinds of thought?
- What kind of user consent do we need to collect and use the data?
- Do we have a mechanism for gathering consent from users?
- Have we explained clearly what users are consenting to?
- Do we have a mechanism for redress, if people are harmed by the results?
- Can we shut this software down in production if it is behaving badly?
- Have we tested for fairness with respect to different user groups?
- Have we tested for disparate error rates among different user groups?
- Do we test and monitor for model drift to ensure our software remains fair over time?
- Do we have a plan to protect and secure user data?

# Real World Example - National Head Start Association (NHSA)

- Serves over 1 million children, 1600 grantee programs w/ over 200K staff across the United States.
- NHSA programs, however, vary in the populations and communities they serve, the data and technological capacities of their staff, and the data and outcomes that they collect.

*Striving to not only identify and align **meaningful early childhood education outcomes**, but also to create a culture of collaboration and collective capacity building across their network...*

# Data-Driven Collective Impact

## Developing and Deploying a New Data and Outcomes Measurement Infrastructure in Early Childhood Education



NATIONAL HEAD START ASSOCIATION

4 pilot participants, Head Start program sites from across the county, identified data infrastructure challenges and helped us scope **open-source data solutions** to build a shared data infrastructure

### Ethical Challenges:

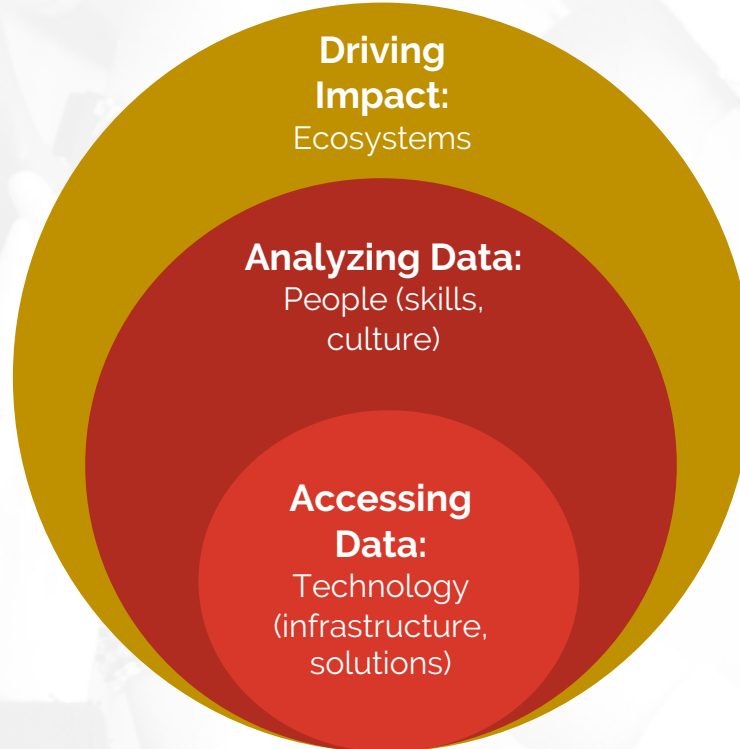
1. Guarantee that the sharing of information within and across Head Start programs is limited to the minimal
2. Ensure algorithms predicting risks for children, their families, and the staff serving them are transparent and unbiased

**November 28th webinar - register here!**

[bit.ly/BrightHiveWebinar](https://bit.ly/BrightHiveWebinar)

[learnmore@brighthive.io](mailto:learnmore@brighthive.io)

# Public-Private Collaboration





# Driving Impact Requires Collective Energy



Define Shared  
Needs Assessment  
& Use Case  
Discovery

Collaboratively  
Develop Data  
Standards &  
Interoperable Tools

Provide Technical  
Assistance and Change  
Management Services

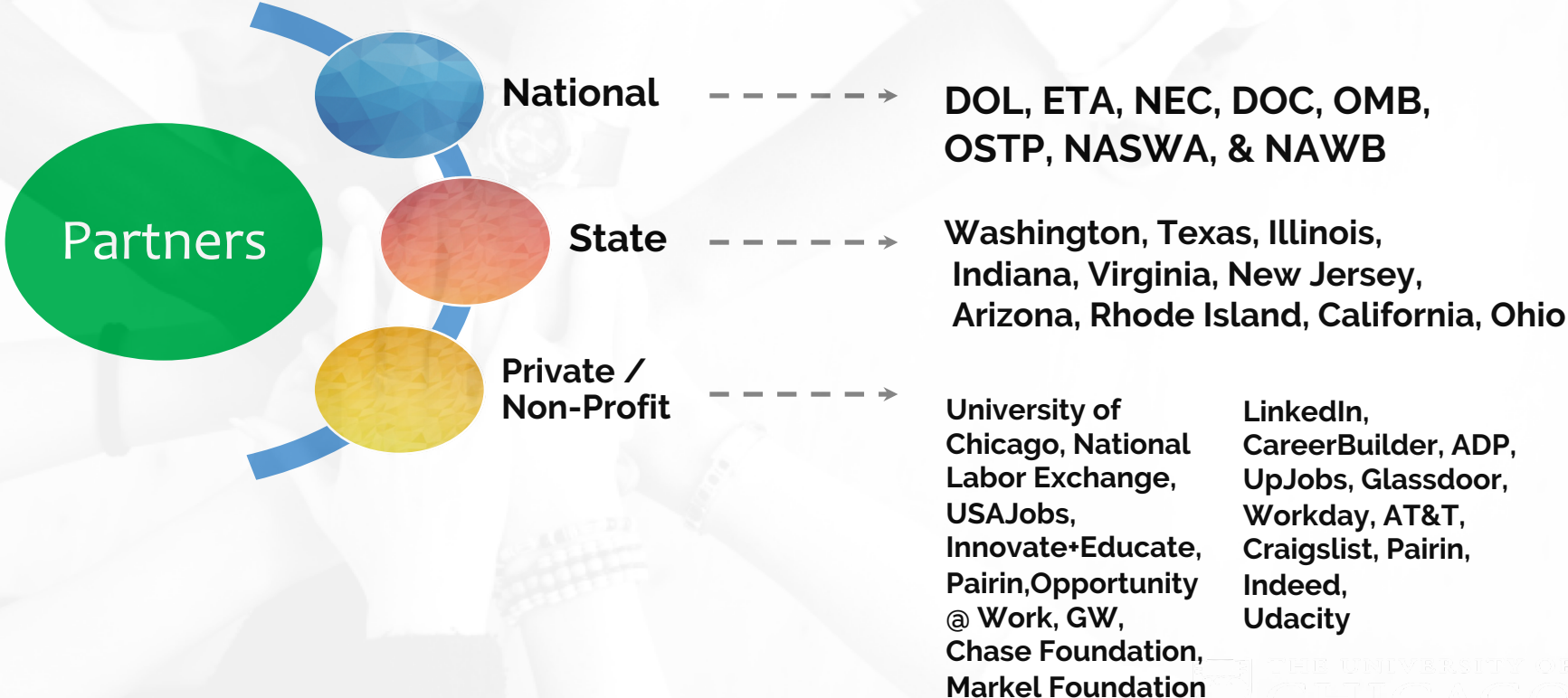


**Building partnerships** to drive development of standards, tools, frameworks, etc to support the community in navigating the data life-cycle responsibly

## The Open Skills Project

Combining survey data with unstructured and semi-structured administrative data to create new, open labor market data for academia, government, and industry

# Who We Are Working With



# DATA AT WORK

Data, tools, and lightweight standards to make a more open and interoperable 21st century workforce data ecosystem.

## Data

Our Open Skills Project publishes standardized job titles, skills, and training outcomes to provide the foundation of a *more open, connected and interoperable* workforce data ecosystem.

[Explore Data](#)

## Tools

Our Training Provider Outcomes Toolkit provides a set of open source tools that to make it easy to collect, connect, analyze, and aggregate data on education and training outcomes.

[Browse Tools](#)

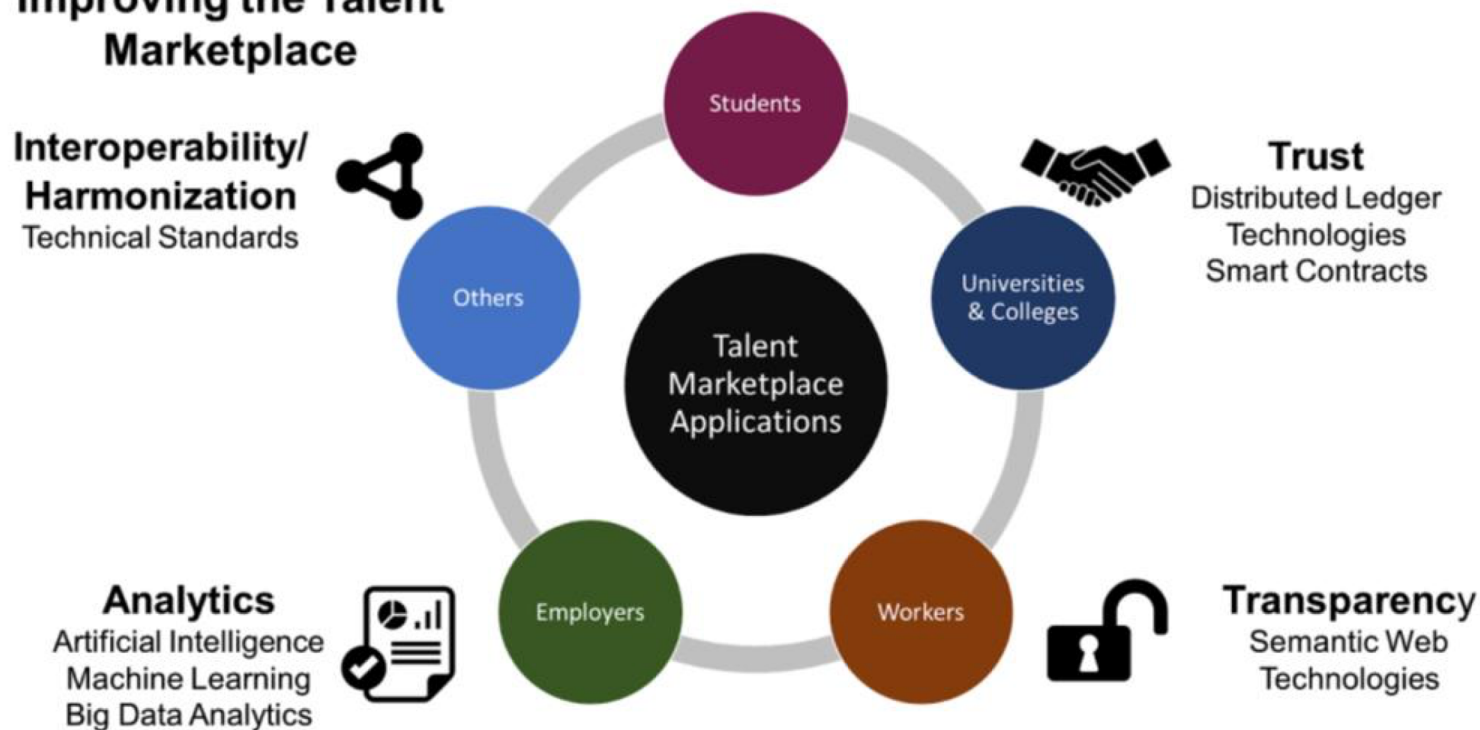
## Working Groups

Our open skills data and open source tools are developed collaboratively through technical working groups staffed by government, industry, academia, and nonprofit institutions.

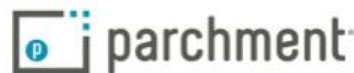
[Get Involved](#)

# T3 Innovation Network

## Improving the Talent Marketplace



# 150+ Participating Organizations





# T3 Network Guiding Principles

**Principle 1:** Focus on High-Impact Stakeholder Use Cases

**Principle 2:** Promote Web 3.0 Convergence

**Principle 3:** Foster Open Collaboration

**Principle 4:** Develop Open Technical Standards and Protocols

**Principle 5:** Utilize Open Competency Frameworks, Taxonomies, and Ontologies

**Principle 6:** Empower Individuals and Enable Self-Sovereign Identity and Data Management

**Principle 7:** Facilitate Open Data Access in Public-Private Data Infrastructure

**Principle 8:** Promote Ethical Practices as well as Equity Considerations

---



# Future of the T3 Innovation Network

1. Promoting and gaining **acceptance of the T3 Guiding Principles**
  1. Exploring and developing **a public-private data and technology infrastructure** that includes:
    - Public-private data standards (WG2)
    - An open and distributed competency data and technology infrastructure (WG3)
    - New architectures and uses of individual-linked data (WG4)
  3. Developing and promoting **participation in high-impact projects** that address the most critical stakeholder use cases (WG1)
  3. Convening stakeholders to **review progress, share information, and develop new initiatives**
-

# Individuals Are The Power

We are a people in a quandary about the present. We are a people in search of our future.  
**We are a people in search of a national community.**

We are a people trying not only to **solve the problems of the present**, unemployment, inflation, but we are attempting on a larger scale to fulfill the promise of America.

We are attempting to fulfill our national purpose, to create and sustain **a society in which all of us are equal.**"

*Rep. Barbara Charline Jordan of Texas  
1976 Democratic National Convention Keynote Address*



**Thank You For Your Time!**

**[www.NatalieEvansHarris.com](http://www.NatalieEvansHarris.com)**

**@QuietStormnat**

**LinkedIn: [nevansharris](#)**

**Tea/Wine: [natalie@brighthouse.io](mailto:natalie@brighthouse.io)**

Backup Slides

# Understand Societal Outcomes and Ethical Implications

## Identifying Overlapping Services and Coordinating Service Provision Across Networks of Service Providers



A network of 6 service providers in Richmond, VA wanted to know who their overlapping populations were in order to better coordinate preventative care for superutilizing individuals.

### Ethical Challenges:

1. Guarantee that the sharing of information across service providers is limited to the minimal
2. Ensure algorithms predicting risks for high need population are transparent and unbiased