Ethics & Opportunities

Natalie Evans Harris

Co-Founder and Head of Strategic Initiatives
BrightHive

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

Former Senior Policy Advisor to U.S. Chief Technology Officer
Obama Administration
Who am I?

I believe that through our **collective** power, data can support transforming the human experience.

<table>
<thead>
<tr>
<th>Accessing Data</th>
<th>Using Data</th>
<th>Governing Data</th>
</tr>
</thead>
<tbody>
<tr>
<td>Co-Founder and COO, BrightHive</td>
<td>Co-Ideator, Global Data Ethics Project (GDEP)</td>
<td>Former Senior Policy Advisor to U.S. Chief Technology Officer, Obama Administration</td>
</tr>
</tbody>
</table>

[Image: Community Principles on Ethical Data Practices]
Challenges to data-driven social impact

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
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 (TX)
1976 Democratic National Convention Keynote Address
Public-private collaboration is critical

- Developing the ecosystems for data-driven social impact
- Investing in infrastructure that uses data to support missions
- Defining Value: People (skills, culture)
- Accessing Data: Technology (infrastructure, solutions)
- Building trust with communities (citizens, technology, academia)

Driving Impact: Ecosystems
Current data infrastructure for identifying, improving, and investing in impact

- Data silos within and across verticals
- Lack of internal analytics capacity
- Lack of standard metrics and algorithms
Future infrastructure must be open and collaborative

- Provide vendor-neutral platform for secure data access and integration
- Provide intuitive views to outcomes and support existing BI tools
- Provide standard outcomes, KPIs, and advanced analytics
- Social Service Data
- Data-Driven Action
Public-private collaboration is critical

- Developing the **ecosystems** for data-driven social impact
- Investing in **infrastructure** that uses data to support missions
- Building **trust** with communities (citizens, technology, academia)

**Driving Impact:**
Ecosystems

**Defining Value:**
People (skills, culture)

**Accessing Data:**
Technology (infrastructure, solutions)
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.

**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
- Foster diversity
- Respect relevant tensions

Exercise our ethical imagination!

https://www.datafordemocracy.org/project/global-data-ethics-project
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.
Ethical culture

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 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?
Data sharing governance

Checklists are great for increasing intentionality of building trust around data use.

But the most impact occurs when you bring these principles together through building your data sharing governance process—the process of controlling the way your data is protected, administered, used, and shared.
Traditional data sharing governance is often limited and under-resourced

**Short-Sighted**
Addresses immediate needs using limited bilateral agreements without a plan for sustainability.

**Isolated**
Often sustained by a single champion, but can fail due to changes in leadership or isolated engagement.

**Imposed**
Imposed on organizations which lack the capacity to participate on equal footing.

**Self-interested**
Lawyers engage to protect interests and eliminate liability at the expense of value to stakeholders.
Data sharing governance should be equitable, sustainable

**Sustainable**
Addresses immediate needs while supporting the potential for long-term growth.

**Coordinated**
Supports shared decision-making for trust-managed data resources and organizational data ownership.

**Empowered**
Levels the playing field for all data sharing members, regardless of size, capacity, or incentives to participate.

**Group-oriented**
Eliminates liability while creating value for communities and individuals who should benefit most from the data sharing.
Data Sharing Governance must be collaborative

To ensure **transparency**, build **trust**, improve **understanding**, and promote a **culture** of collaboration

To detail technical & data specifications, **legal & ethical permissions and use**, designate trustee(s), and establish a governance board

To **leverage existing** legal agreements, knowledge, relationships, and collaborations

To ensure agreed upon changes are **managed through amendments**, rather than re-negotiation
Technical infrastructure is modular, flexible, and secure

- Built to the **data** and **technical specifications** articulated in the Data Trust Member Agreement (DTMA)
- Uses **open source** and **open standards**
- Flexible, **modular stack** integrates with members' existing data systems
- Meets each Data Trust Member's data **security** and **privacy** thresholds

BrightHive DTMA can be found on [GitHub](https://github.com)
Governance board monitors and sustains the data trust

Consists of data trust members who collectively manage, monitor, and sustain the data trust over time.

Ensures data and technical specifications, data access and use permissions, use of pooled resources are consistent with the legal agreement.

Identifies and agrees upon data trust expansions, including new use cases, new members, new pooled resources, and third party users.

Collaboratively amends the DTMA to meet the expanding and evolving needs of the Data Trust and its beneficiaries.
Public-private collaboration is critical

- Developing the **ecosystems** for data-driven social impact
- Investing in **infrastructure** that uses data to support missions
- Building **trust** with communities (citizens, technology, academia)
Driving impact requires collective energy

Define shared needs assessment & use case discovery

Provide technical assistance and change management services

collaboratively develop data standards & interoperable tools
Scalable Impact must have Shared Vision

“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 (TX)
1976 Democratic National Convention Keynote Address
Natalie Evans Harris

natalieevansharris.com

Twitter
@QuietStormnat

LinkedIn
nevansharris

Tea/Wine
natalie@brighthive.io

“If you want to go quickly ... go alone. If you want to go far ... go together.”

African Proverb