

INTERNET<sup>2</sup>  
2017 global  
SUMMIT

**CINC UP: NSF BIG DATA INNOVATION HUBS**

**April 26, 2017**



## CINC UP: NSF BIG DATA INNOVATION HUBS

### AGENDA

- Welcome, Florence Hudson, Internet2
- Big Data Regional Innovation Hubs & Spokes, Fen Zhao, NSF
- Northeast Regional Big Data Innovation Hubs & Spokes, René Bastón, Northeast Big Data Hub

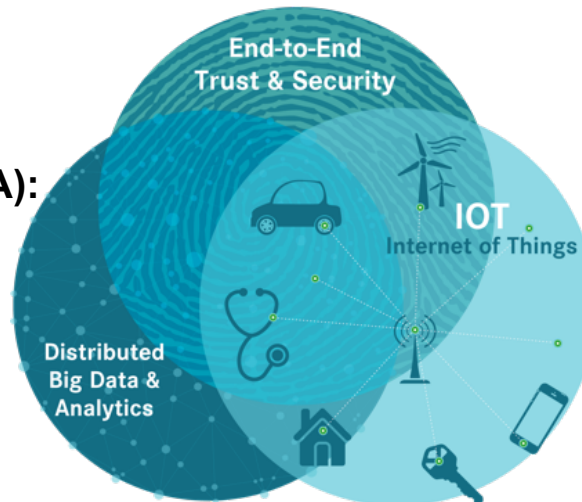
## Internet2 Collaborative Innovation Community focus areas align with the NSF BD Hubs.

### E2E Trust & Security (E2ET&S):

- TIPSS for IoT – Trust, Identity, Privacy, Protection, Safety, Security
- NSF EAGER Cybersecurity Transition to Practice (TTP) Acceleration
- SDP (Software Defined Perimeter), Network Segmentation for IoT

### Distributed Big Data & Analytics (DBDA):

- Health & Life Sciences / Genomics
- Smart Campuses & Cities
- NSF Big Data Hub Collaboration



### Internet of Things (IoT):

- IoT Sandbox
- Smart Campuses & Cities
- Smart Grid Testbed

## **NSF BIG DATA HUBS OVERVIEW**

**FEN ZHAO**

National Science Foundation



# BIG DATA REGIONAL INNOVATION HUBS & SPOKES

Program Briefing

Fen Zhao

April 17, 2017

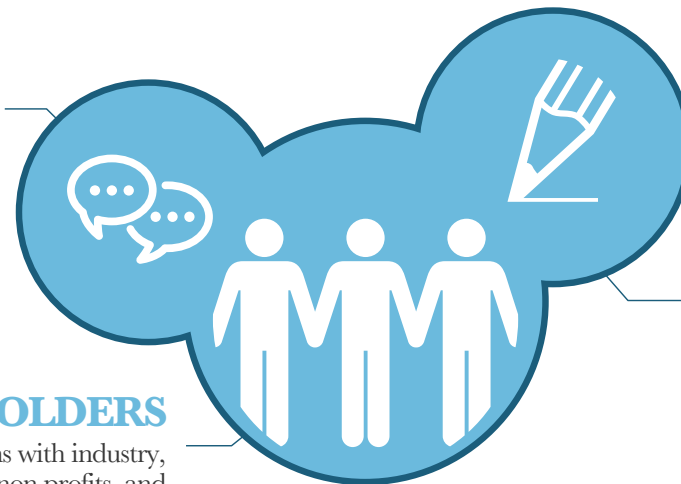
# KEY TAKEAWAYS

## 01 THE PROGRAM

Brings together domain scientists, computer scientists, and end users to use data to solve challenges

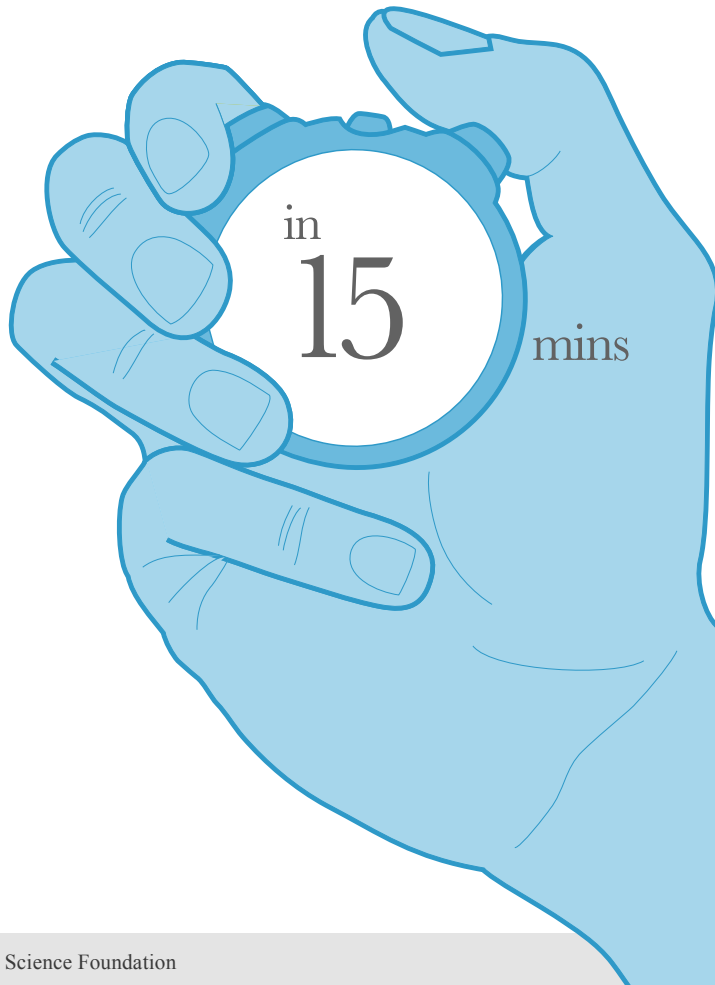
## 02 THE STAKEHOLDERS

Encourages collaborations with industry, state & local governments, non profits, and others that are not typical NSF participants



## 03 PARTICIPATION

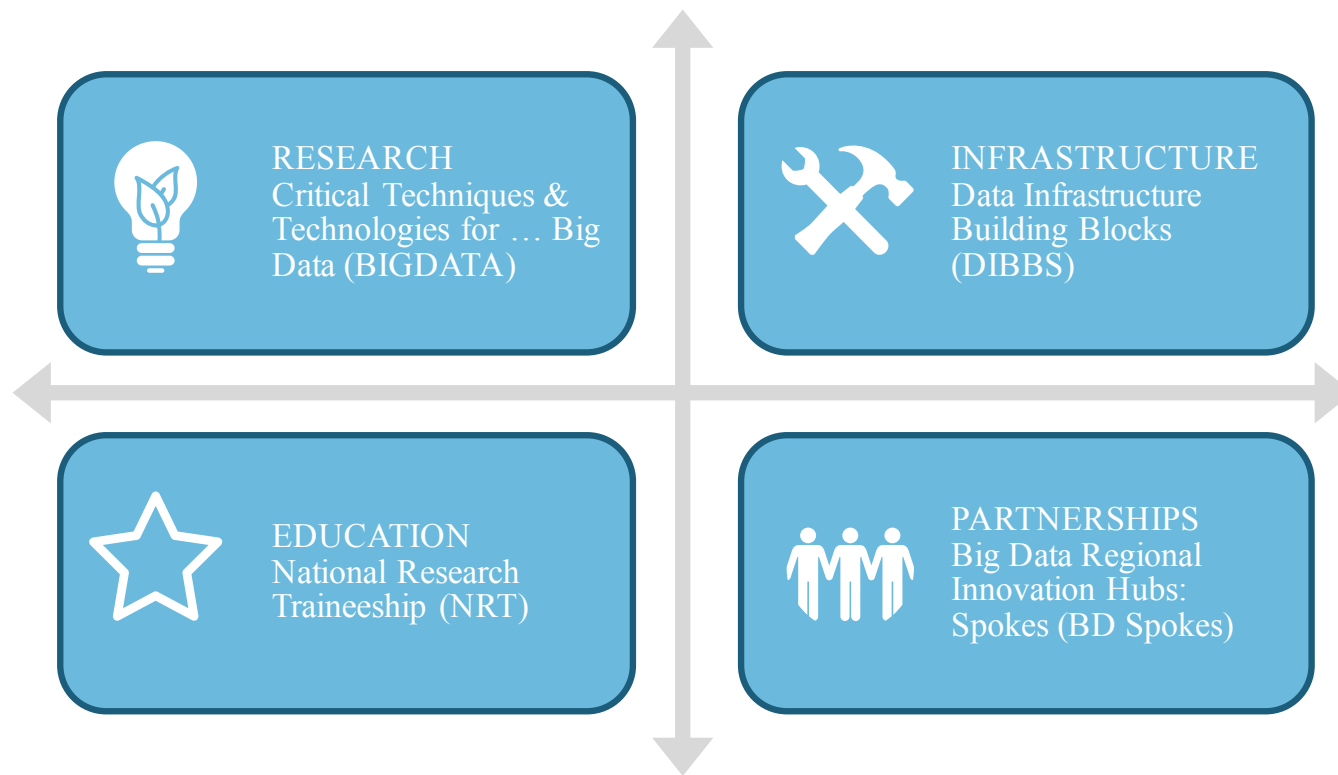
Opportunity the I2 community to participate!



vision of the BDHubs program  
activities of funded Hubs  
spokes awarded  
opportunities for participation

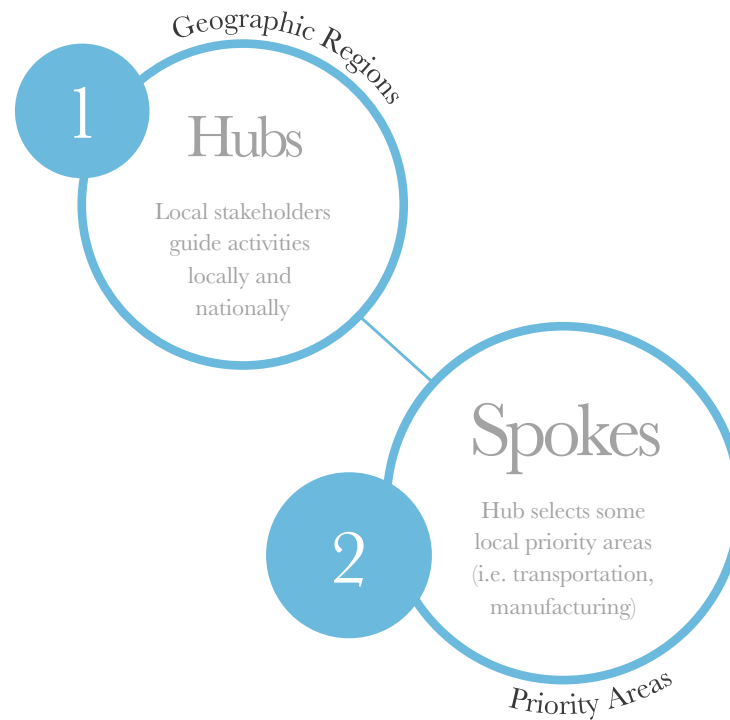
# WITHIN THE BIG DATA PORTFOLIO OF PROGRAMS

Within the broader portfolio, BD Hubs and BD Spokes  
focuses on building partnerships around Big Data



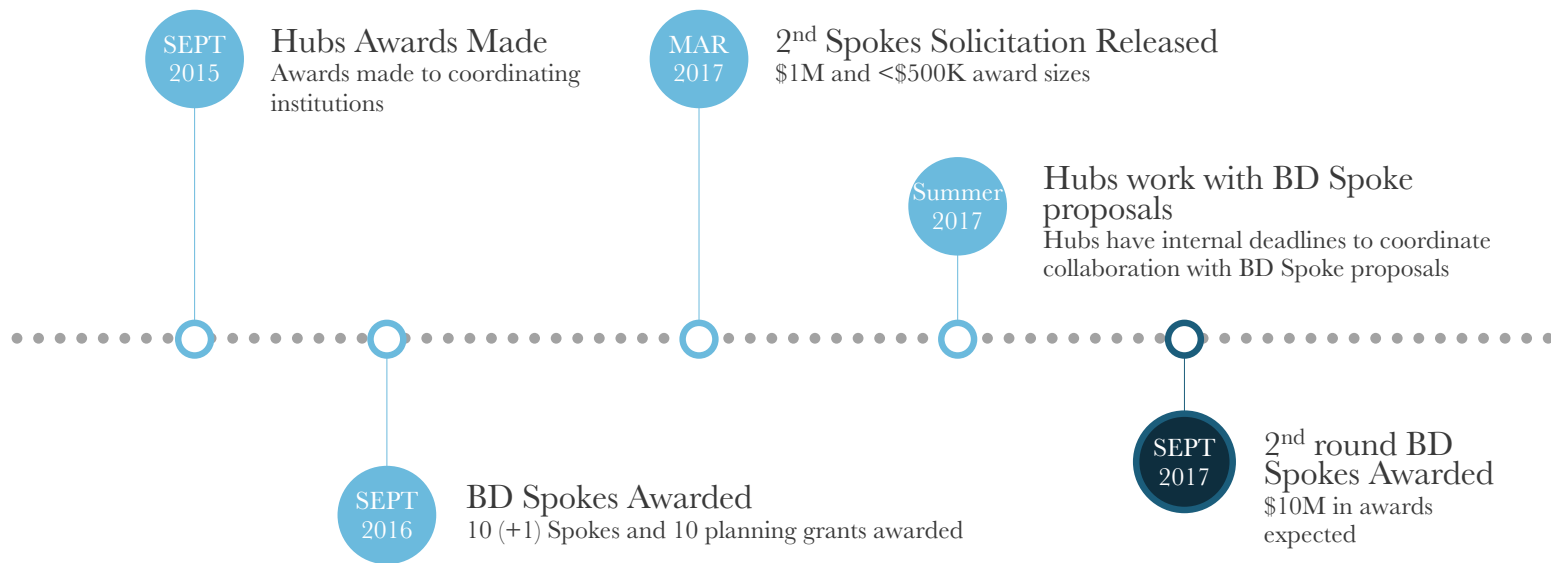
# WHAT IS THE BDHUBS NETWORK?

“Hub and Spoke”– A Nation-Wide Network for Data Innovation

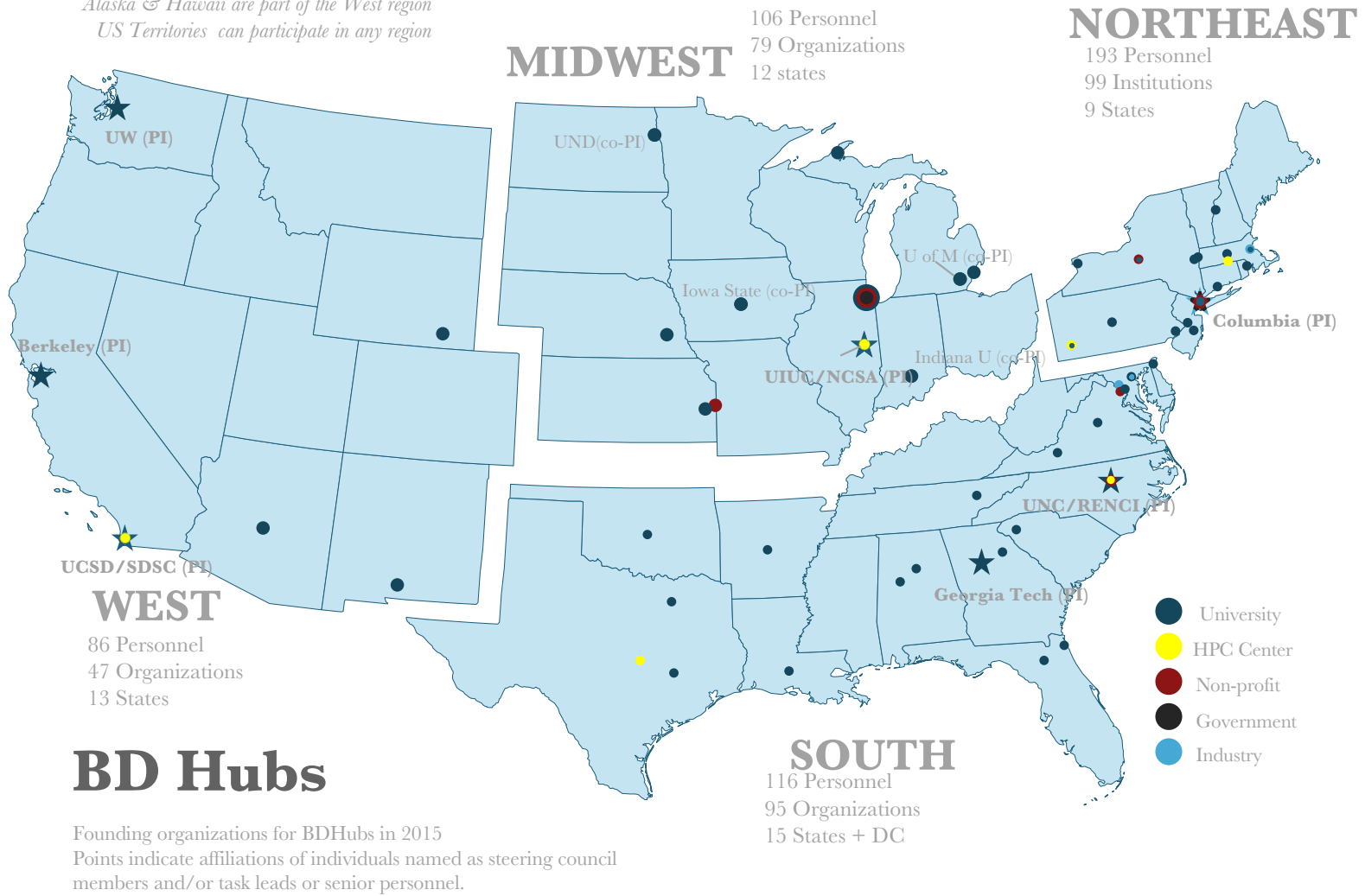


# THE BD HUBS & SPOKES TIMELINE

Currently



Alaska & Hawaii are part of the West region  
 US Territories can participate in any region





# Big Data Hubs and Spokes Ecosystem

Each BD Hub works with multi-institutional and multi-sector teams, or BD Spokes, to organize and coordinate projects to use data to address regional priorities.

**\$12M Awarded in FY2016-17:**

**11 Spokes Grants** (\$1M over 3 years); **10 Planning Grants** (\$100K for 1 year)

Themes:

- Improve **data sharing**,
- **Automate** the big data lifecycle, and
- Use data to address societal **grand challenges**

Midwest midwestbigdatahub.org	Northeast nebigdatahub.org	South southbigdatahub.org	West westbigdatahub.org
Neuroscience Network	Licensing Model and Ecosystem for Data Sharing	Medical Informatics	Reproducibility in Scientific Computation
Digital Agriculture	Data-Driven Education	Smart Grids	MetroInsight
Integrative Materials Design	Observational Health Research	Big Data for Environmental Sustainability	Policing in the Western US
Bridge Health	Big Data Literacy	Rare Disease Observatory	Poteogenomics Applications of Genetic Data
Cyberinfrastructure for Sensor Originated Big Data	Big Data Cyber Attack Awareness		
Networked Resilience of Communities	Privacy and Security in Big Data		
	Energy Cycle Innovation		

Spoke Grant
Planning Grant



# HUB ACTIVITIES

Hubs ideate and coordinate Spokes, but also host a variety of activities for the community



Microsoft awards  
Hubs \$3M in cloud  
computing credits

Massive regional  
All-Hands with  
hundred of attendees

Early career  
researcher  
programs  
with CCC

3 years  
sociotechnical  
study of Hubs



# The strategy behind BD SPOKES

BD Spokes are  
not your typical R&D project  
nor are they mini Hubs

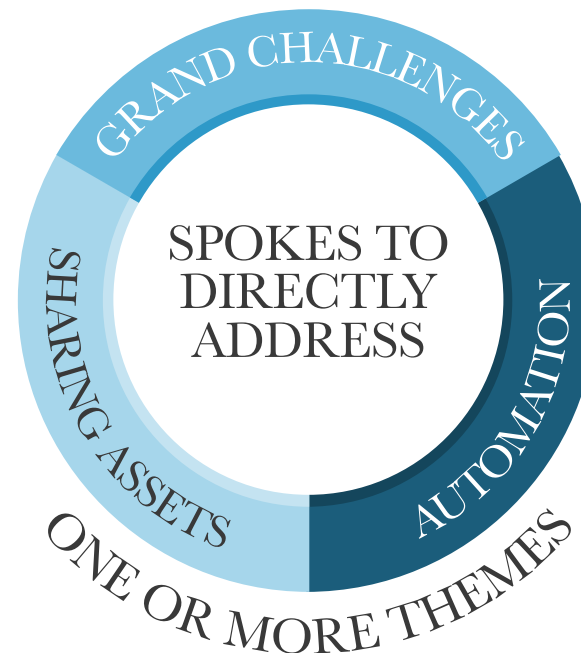
# MISSION DRIVEN SPOKES

BD Spokes proposals must articulate a clear focus within a specific Big Data topic or application area, while highlighting their Big Data Innovation theme.

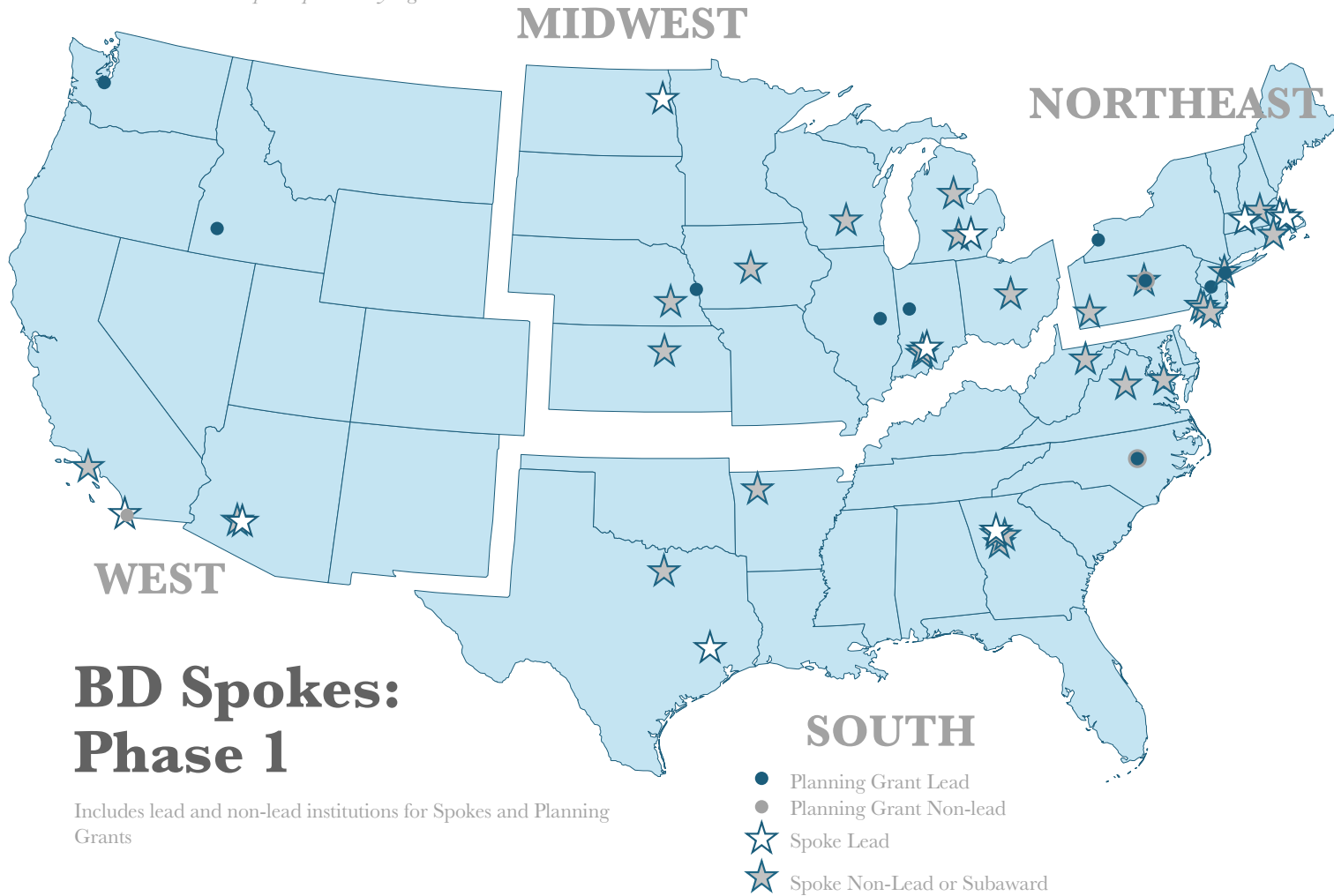
All BD Spokes must have clearly defined mission statements with goals and corresponding metrics of success.

# SPOKES MAJOR THEMES

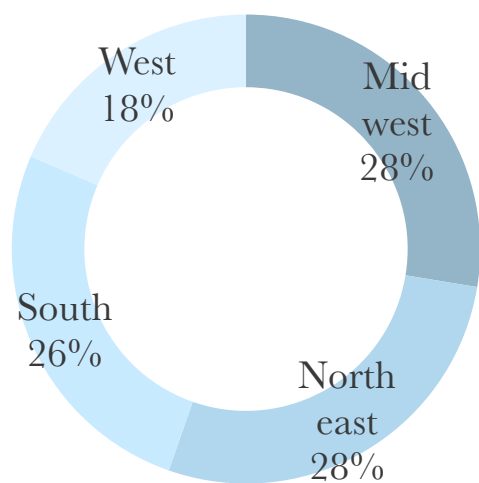
Three different ways of slicing  
the Big Data Innovation  
problem



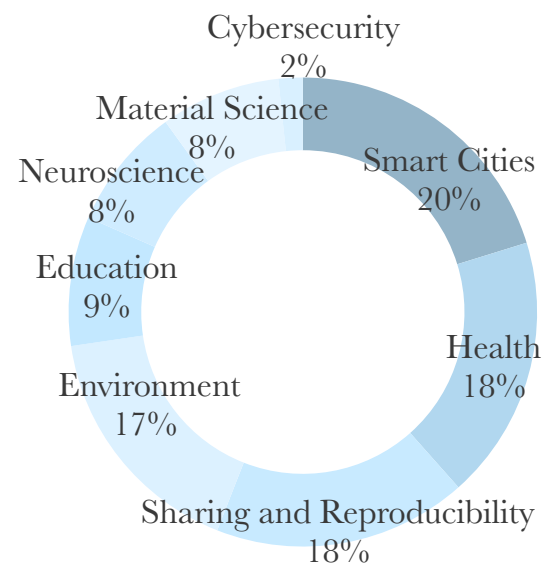
*Alaska & Hawaii are part of the West region  
US Territories can participate in any region*



Percent funding per region



Percent funding per topic area

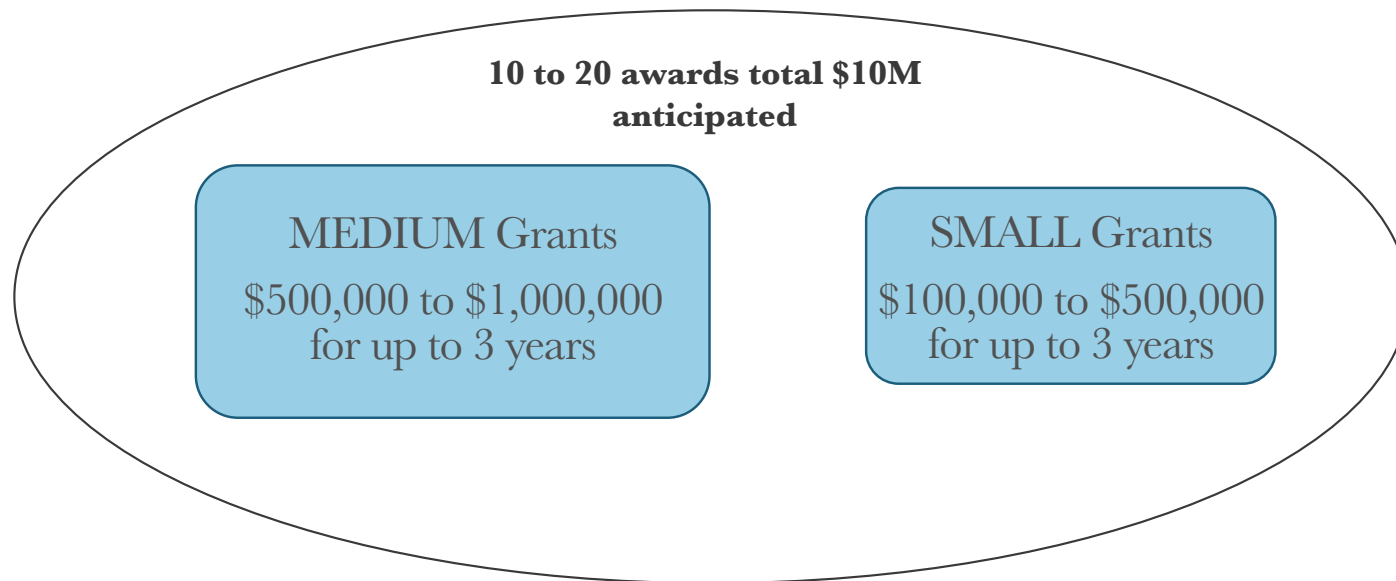


Total Spokes ~\$12M in first round



## Big Spokes Solicitation (17-546)

Same 3 Themes: Data Sharing, Automation, Grand challenges



**Proposal submission deadline: 9/18/2017**  
**No Letters of Intent**

# TANGIBLE OUTCOMES

Medium proposals are expected to articulate tangible outcomes:

- **Explicit results** from data-enabled or data-facilitated inquiry in a scientific or engineering field or other domain area, including social science and education research; must explicitly define the field- or domain-specific questions addressed and which specific researchers would be conducting the research to answer these questions
- A **prototype** or proof of concept for a technology platform, data product, data standards, or other data infrastructure
- An innovative **education or workforce development program** with a plan for evaluating the effectiveness of the program. The program could be virtual-, classroom-, or workplace-based and should engage underrepresented groups.





## AREAS OF EMPHASIS FY 17

Some NSF priority areas include



NEUROSCIENCE



REPLICABILITY &  
REPRODUCIBILITY  
IN DATA SCIENCE



DATA DRIVEN  
RESEARCH IN  
CHEMISTRY



EDUCATION



DATA INTENSIVE  
RESEARCH IN THE  
SOCIAL,  
BEHAVIORAL, &  
ECONOMIC  
SCIENCES



DATA ANALYTICS  
FOR SECURITY



## LIMITED COMPETITION

An organization may participate as submitting organization (or, in the case of collaborative proposals, as the lead organization) for at most one proposal responsive to this solicitation.

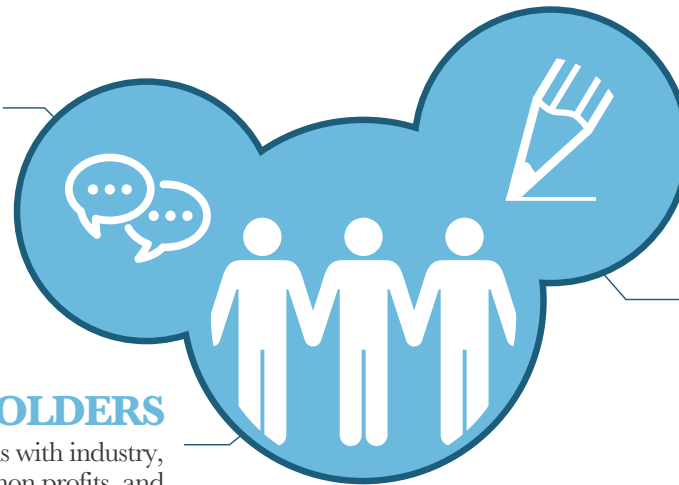
No limit is placed on an institution submitting a non-lead proposal in a collaborative proposal.

Proposal submissions are limited to maintain a balanced geographic representation of the Regional Hubs and Spokes program and to increase diversity of participating institutions

# KEY TAKEAWAYS

## 01 THE PROGRAM

Brings together domain scientists, computer scientists, and end users to use data to solve challenges



## 02 THE STAKEHOLDERS

Encourages collaborations with industry, state & local governments, non profits, and others that are not typical NSF participants

## 03 PARTICIPATION

Opportunity the I2 community to participate!

# FOR FURTHER QUESTIONS CONTACT

Fen Zhao, [fzhao@nsf.gov](mailto:fzhao@nsf.gov) 703 292 7344



NSF Headquarters, Arlington VA

## **SPOKE AWARDS**

**RENÉ BASTÓN**

Northeast Big Data Hub



**NORTHEAST**  
**BIG DATA**  
INNOVATION HUB

---

René Bastón, Executive Director

Kathy McKeown, PI

*Coordinated by Columbia University*

# NE Regional Big Data Innovation Hub

---

Provide frameworks for public-private, multi-sector collaborations to address high-priority challenges with data-driven solutions



CA, CT, IL, MA, MD, ME, NJ, NY, NH, PA, RI, TX, VA, WA  
...D.C., and QUÉBEC!

# Not Just About Technology





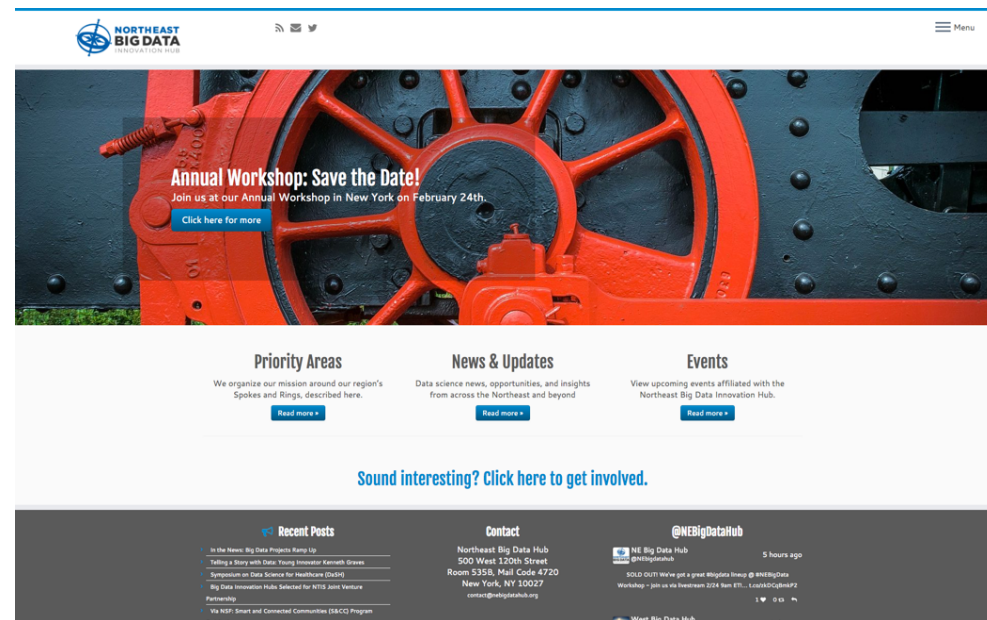
# Year 1 - By the Numbers



## Community/Communications:

- Events: Convened or participated in 18 events; over 2700 people
- Website Relaunch: Blog, Newsletter, Twitter; ~4600 visitors/month

Altogether, 10s thousands reached



# Year 1 - By the Numbers



## Funding and In-Kind Contributions

- \$4.5 million

## 9 New Projects

- 3 Spoke Projects
- 4 Planning Projects
- Innovator Internships
- BD-Map

# Events



Organizational Change: Building Data  
Enabling Industry and Government to Work Together  
SCHEDULE, AND A WORKING GROUP OF OFFICIALS, AT A  
MISSION, D.C.

# Pounding the Pavement

---



## Northeast Road Trip

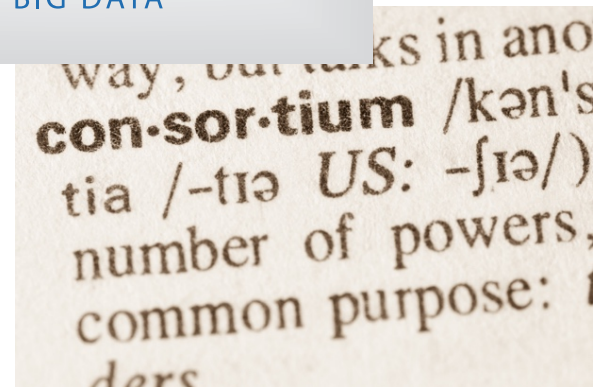
- 9 cities, so far
- Discovering Stakeholder Needs
- Identifying Data Resources in the NE
- Exploring Membership, Consortia, and Partnership / Collaboration Models

# Exploring New Initiatives

## Big Data Map: Assets Platform (BD MAP)

- Big Data Resources
- Exploring collaborations with Hypothes.is; Pittsburgh Super Computing Center; and XSEDE

## Cross Sector Consortia





# Cross-Sector Outreach



# Cross-Sector Outreach



# Cross-Sector Outreach



THE CITY OF NEW YORK  
OFFICE OF THE MAYOR  
NEW YORK, NY 10007



Empire State  
Development



# Cross-Sector Outreach

---



**International Educational Data Mining Society**

# Cross-Hubs Coordination

---

- Cloud Vendors
- DOT / SHRP2 Opportunity
- NTIS Joint Venture Partnership
- NSF Directorate and Federal Agency Outreach
- National Big Data Infrastructure



# Funding / In-Kind

---



## \$4.5 Million in Additional Resources and Awards

- Graduate Student Supplementation Awards (BD-MAP) (\$80,000)
- Computing Community Consortium Early Career Data Scientist Grant (\$300,000)
- Microsoft Azure Cloud Credits (\$750,000)
- Spoke Projects (\$3.3 million)
- National Technical Information Service Joint Venture Partnership

# Spoke Awards

---

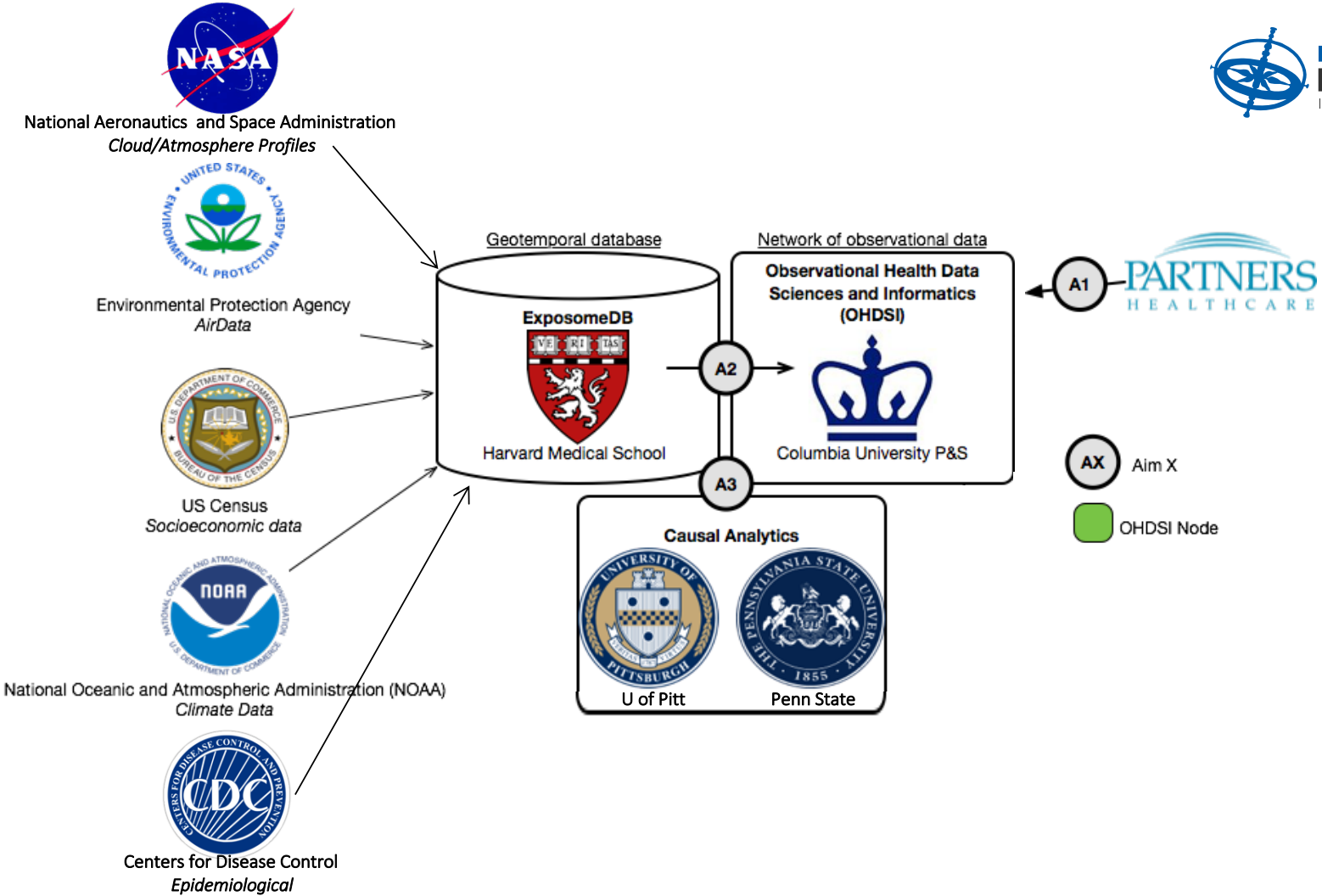


7 NSF Spoke Awards



Integration of Environmental  
Factors and Causal Reasoning  
Approaches for Large-Scale  
Observational Health Research









# A Licensing Model and Ecosystem for Data Sharing





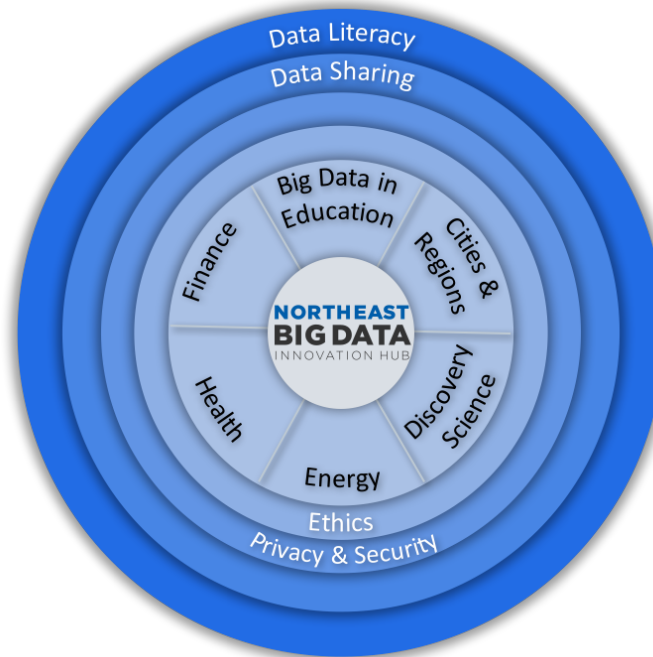
# Grand Challenges for Data-Driven Education



# What's Next?

# Spokes & Rings

---



- Advancing existing Spoke and Planning Projects
  - Assisting with next NSF solicitation

# Next Steps & New Directions

- Innovator Internships
- BD Map Pilot
- Consortia Models
  - Transportation
  - Cybersecurity Insurance
- Expanded Access to Data
- Cross Hub Communications Plan
- Ad Hoc Workgroups



# Executive Education Pilot

---

- Bridging the Big Data Knowledge Gap
- A short course targeted at Executives
- Goal: Help decision-makers understand how to effectively use data-driven tools for decision support



# Community Engagement



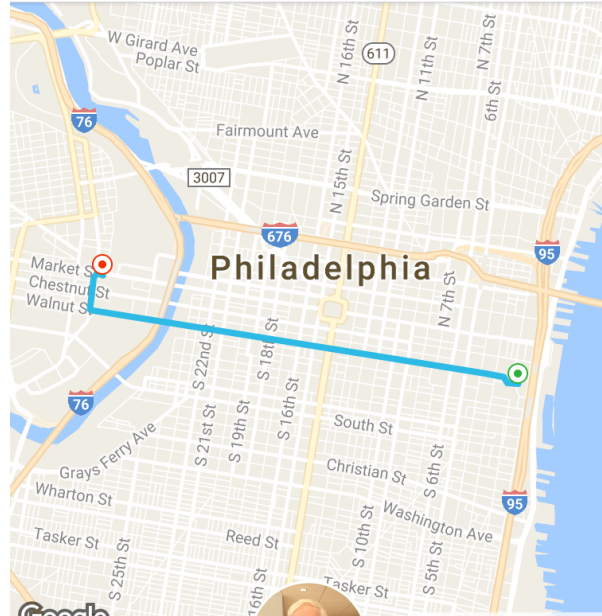
- Steering Committee Workgroups
- Expanding engagement and activities in priority areas
- Exploring new areas in which NE Hub can help
- Expanding number of participants across NE Hub



Search ●●●● LTE 1:11 PM 56%

← SEP 30, 2016, 8:28 AM

- 272 Dock St, Philadelphia, PA 19106, USA
- 3268-3276 Market St, Philadelphia, PA 19104, USA



YOU RATED  
★★★★★



KARL

TRIP TOTAL  
⚡ \$34.41

NEED HELP?

UBERX RECEIPT



# For More Information

---

[nebigdatahub.org](http://nebigdatahub.org)

[rb70@columbia.edu](mailto:rb70@columbia.edu)

[contact@nebigdatahub.org](mailto:contact@nebigdatahub.org)

[@NEBigDataHub](#)

[#BDHubs](#) | [#NEBigData](#)

[tinyurl.com/NEBDHubList](http://tinyurl.com/NEBDHubList)



## Internet2 Engaging with the Big Data Innovation Hubs

- Help with coordinating use of current infrastructure, for current and future needs
  - As a national service provider, we can help coordinate with state/regional networks, international partners
  - Participating in “A Licensing Model and Ecosystem for Data Sharing” Spoke in NE BDHub.
    - September 2016 workshop report: <http://cci.drexel.edu/mrc/news/2016-11-bigdatahubworkshop/>
  - Engaging in Midwest BDHub to support genomics big data transfer
  - Engaging South BDHub “Smart Grids Big Data” spoke, attending TAMU workshop April 18
- Incorporating BDHubs needs in future infrastructure
  - Giving a voice to the BDHubs community for future needs



## CINO Sponsored Schedule of Events at GS17

- **Sunday, April 23, 4:15-5:30pm: CINC UP: CINO Program Advisory Group Meeting (Open), Meeting Room 15**
- **Tuesday, April 25, 8-10am: CINC UP: Collaborative Innovation Community Meeting: IoT, E2ET&S, Smart Campus, Renaissance Ballroom West B**
  - Collaborative Innovation Community & Innovation Working Groups Update: IoT, E2ET&S, DBDA
    - Smart Campus Initiatives Update and invitation to participate
    - Smart Campus: IoT Systems Risk Management Task Force Update. Shodan & Censys.io demonstrations 4/24 & 4/25
    - TIPSS for IoT: ITANA Collaboration and White Paper
  - Smart Campus-themed Cybersecurity Transition to Practice Researcher Presentations
  - IoT Pedagogy
- **Wednesday, April 26, 7:30-8:30am: CINC UP: NSF Big Data Innovation Hubs, Meeting Room 10/11**
  - NSF Big Data Hubs and Spokes Overview by Fen Zhao, NSF, and René Bastón Northeast Big Data Innovation Hub ED
  - How to get involved, connections for researchers, regional networks, and IT
- **Wednesday, April 26, 12:30-5:30pm: CINC UP: Cybersecurity Research Acceleration Transition To Practice (TTP) Workshop and Showcase (NSF #1650445), Meeting Room 8/9**
  - Join us for an interactive discussion to determine how – working together – we can accelerate Transition To Practice (TTP) of cybersecurity research into operational environments. Regional networks, IT, industry, labs, students: everyone is invited.
  - University CIO Perspective on Leveraging Cybersecurity Research
  - 12 Researcher Presentations on Identity & Access Management, Network Security, Smart Grid, Cloud Security & Storage, Data Analytics & Security, and IoT
  - Discussion, Pilot Opportunities, and Feedback, tell us what cybersecurity assets you need
  - Poster Session & Networking at breaks. Additional poster sessions on Monday and Tuesday: breakfast, lunch, breaks.

INTERNET<sup>2</sup>  
2017 **g**lobal  
SUMMIT

**CINC UP: NSF BIG DATA INNOVATION HUBS**

**April 26, 2017**

