

# BIG DATA REGIONAL INNOVATION SPOKES

Advancing Big Data Applications

**Fen Zhao**  
Program Coordinator  
CISE/CNS

# THE BD SPOKES TEAM

Many thanks goes out to...

## **CISE Team:**

- Chaitan Baru
- Silvia Spengler
- Ken Whang
- Deborah Shands
- Joanna Chan
- Stephanie Rodriguez
- Gabriel Perez-Giz
- Elena Zheleva

## **Co-program Officers:**

- Earnestine Psalmonds (EHR)
- Cheryl Eavey (SBE)
- Lin He (MPS/CHE)
- Seta Bogosyan (OISE/CISE)

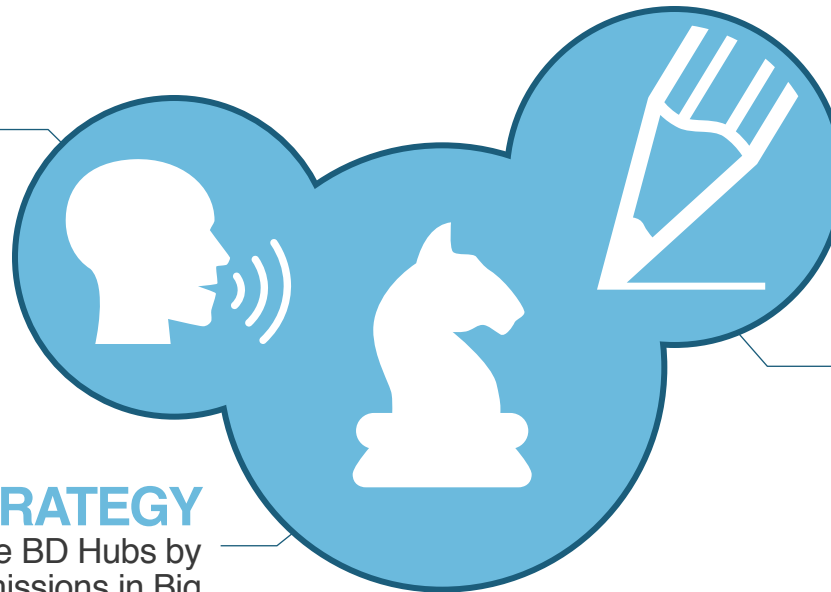


# WHAT IS THE BD SPOKES PROGRAM?

An Agenda for The Discussion Today

## 01 THE HISTORY

BD Spokes is the second phase of the BD Hubs program



## 02 THE STRATEGY

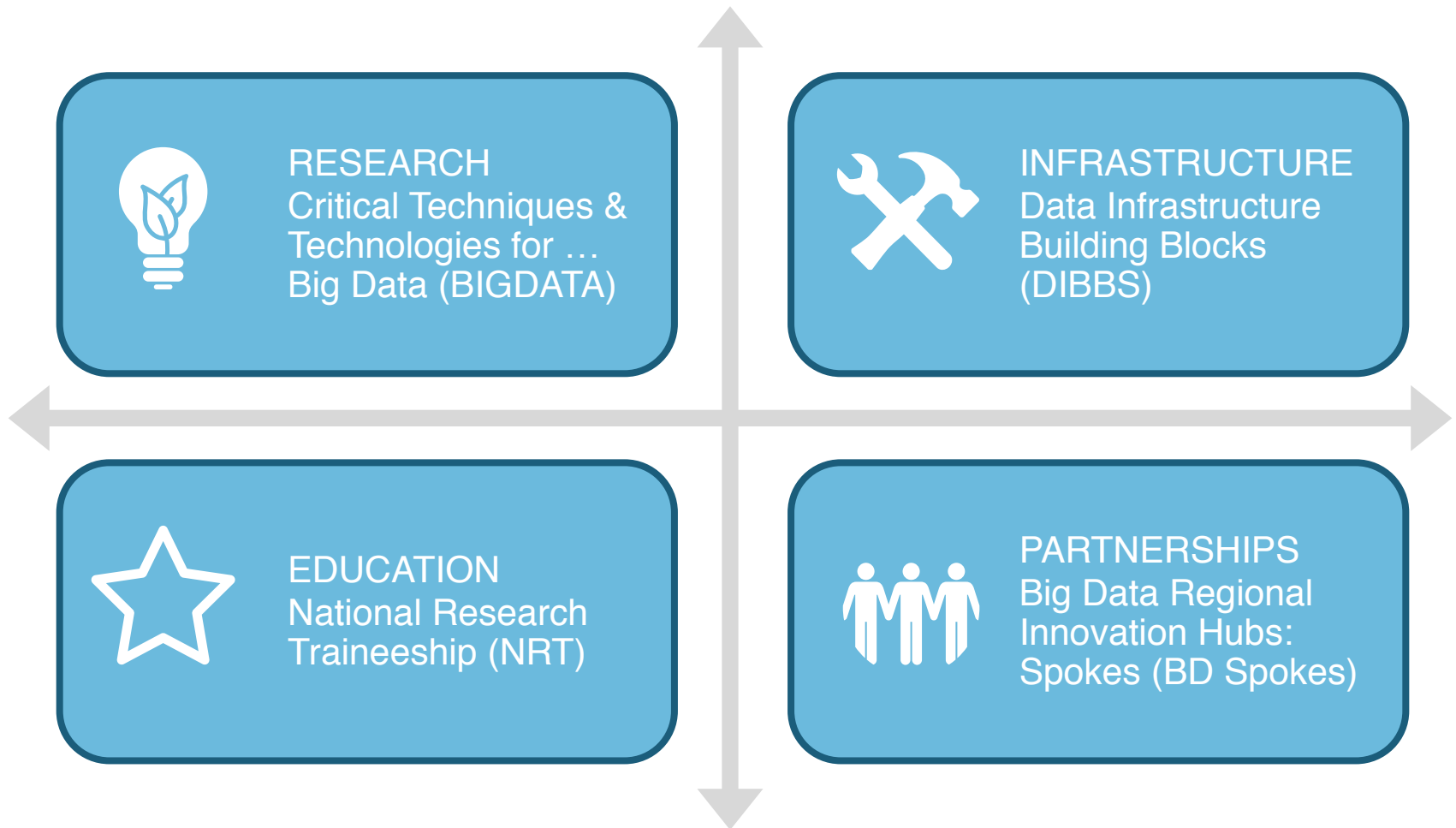
BD Spokes enable the BD Hubs by supporting targeted missions in Big Data application areas

## 03 THE CALL

The second NSF Spokes (17-546) solicitation with full proposal due 9/18/2017.

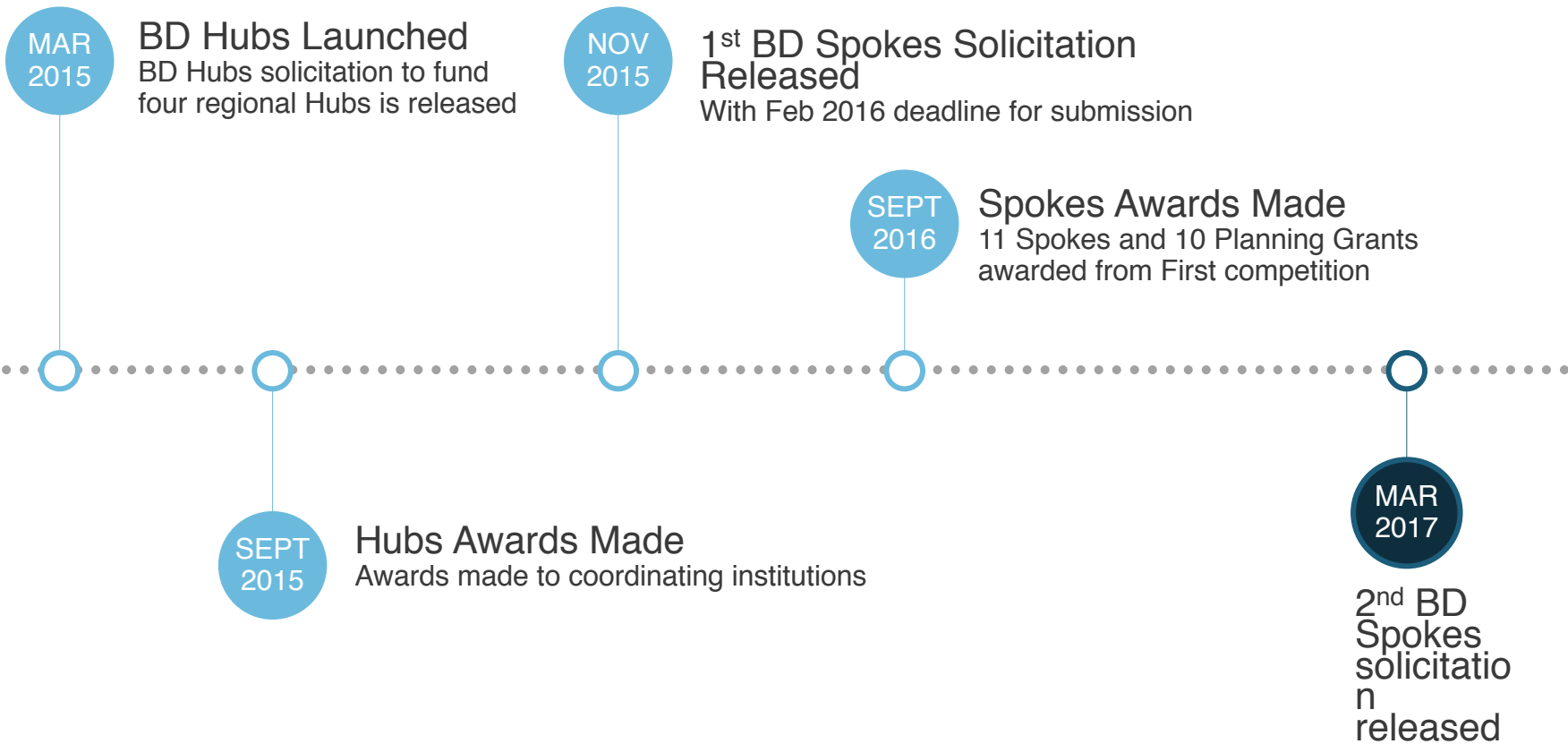
# THE NSF BIG DATA PORTFOLIO OF PROGRAMS

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



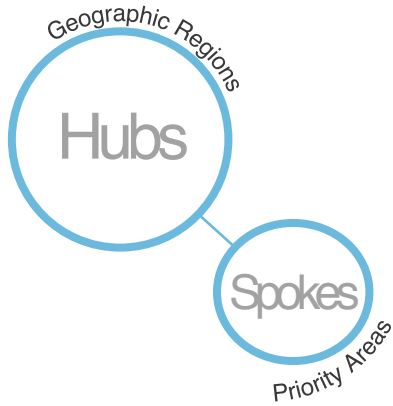
# THE HISTORY BEHIND BD SPOKES

BD Spokes is the second phase of a long term NSF agenda for Big Data Partnerships



# 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:

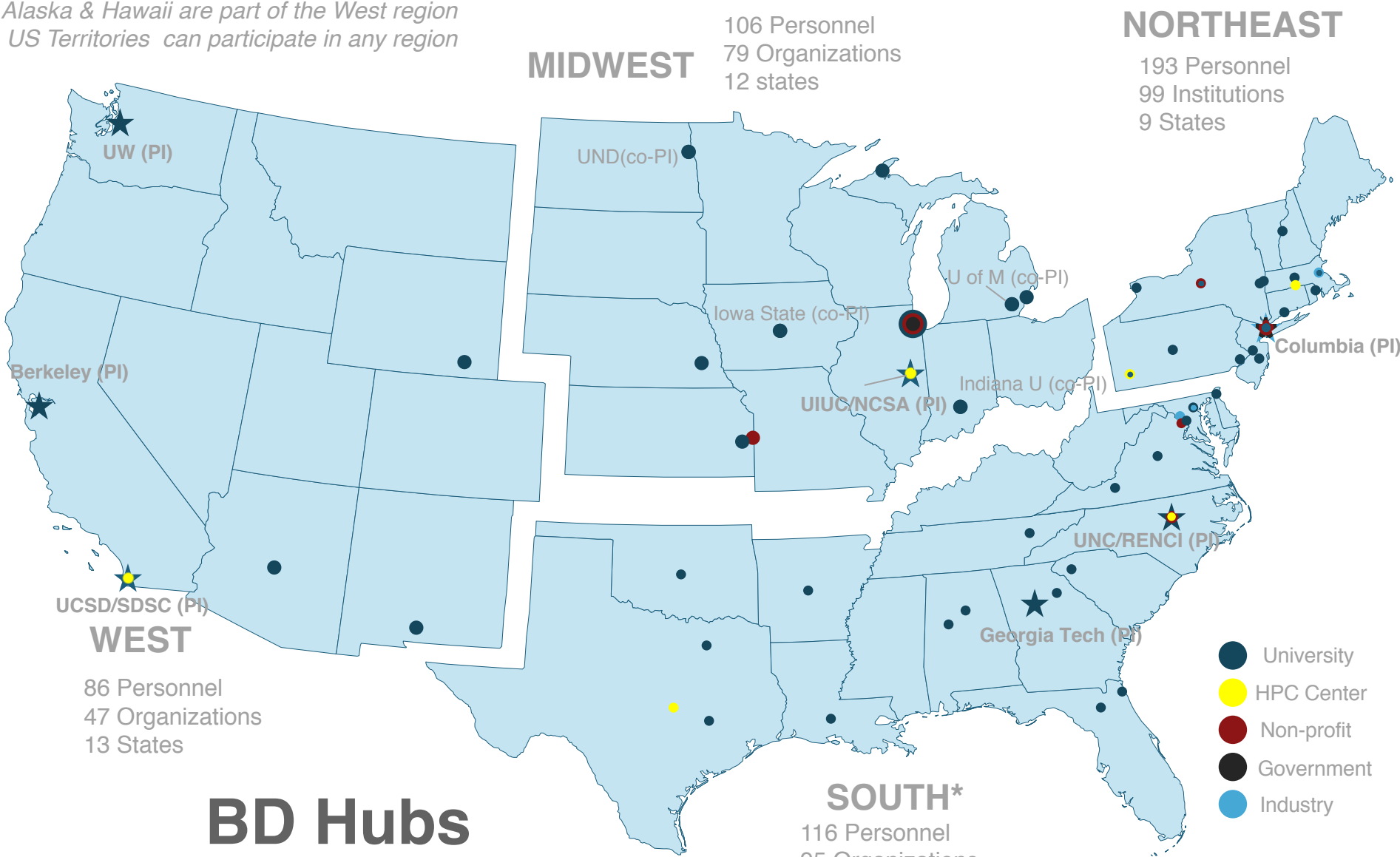
- Improve **data sharing**,
- **Automate** the big data lifecycle, and
- Use data to address societal **grand challenges**

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



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



# THE BD SPOKES SOLICITATION (17-546)

Details and Specifications



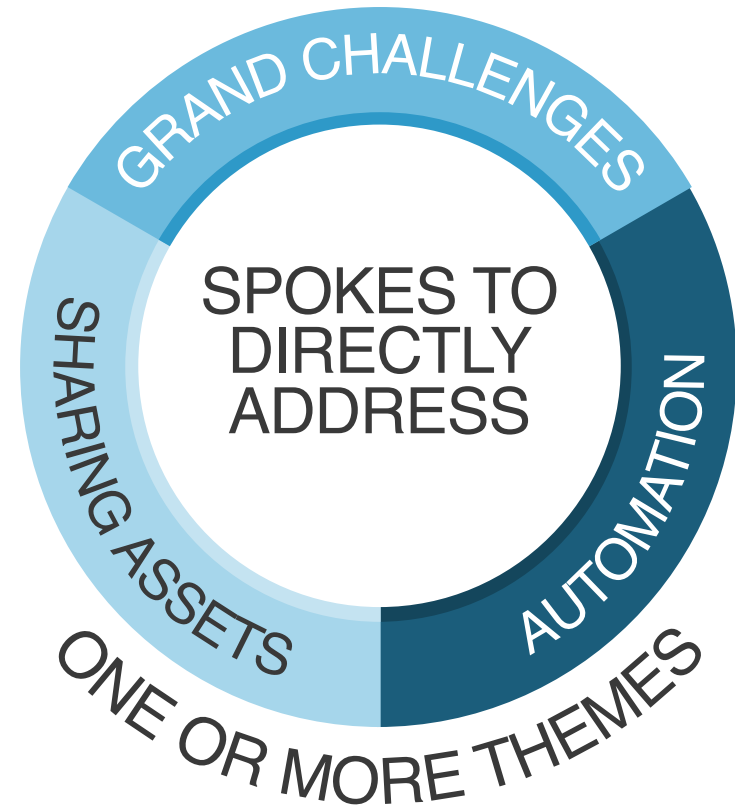
# MISSION DRIVEN SPOKES

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





Accelerating progress towards societal

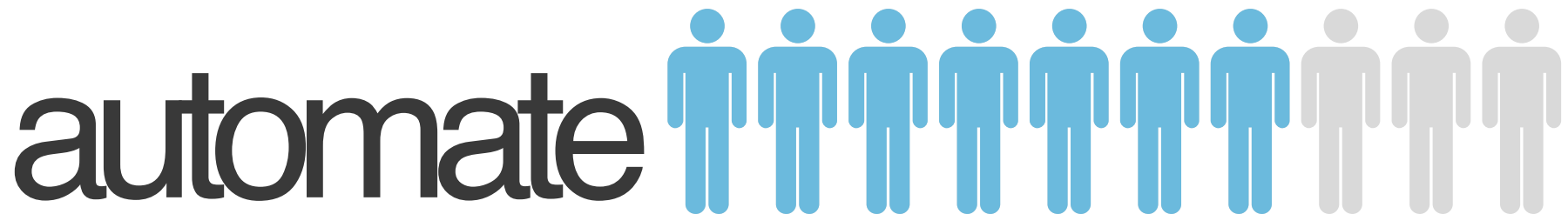
# grand challenges

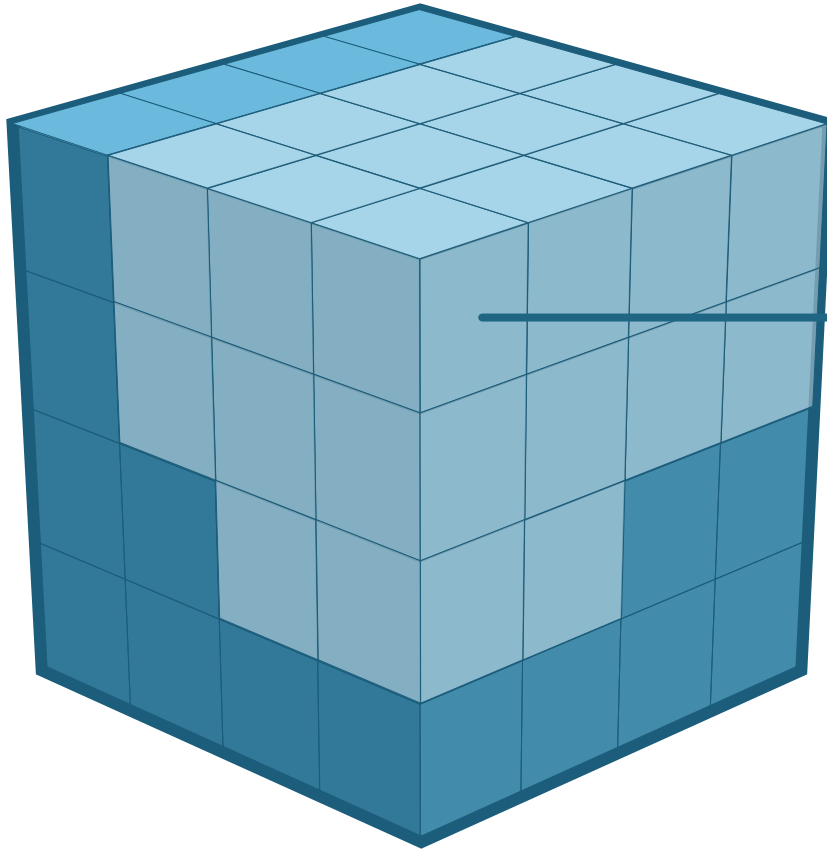
relevant to regional and national priority areas.

Due to the pervasiveness of Big Data in virtually all national priority areas, the BD Spokes have the opportunity to bring rapid change in application areas, by facilitating the creation of interdisciplinary and multidisciplinary data-intensive teams.

Steps in the data lifecycle include: ingestion, validation, curation, quality assessment, anonymization, publication, active data management, and analysis (including information extraction, visualization, and annotation).

Automated (or, semi-automated) techniques are needed in order to keep up with the rapid data rates, large volumes, and immense heterogeneity of Big Data. Automation may also aid the reproducibility of data processing and analysis workflows.





Enabling access to and increasing the use of valuable, available **data assets**, also including international data sets, where relevant

One of the desirable roles for a BD Spoke is as a catalyst for organizing and sharing datasets and related data services among a larger set of stakeholders, across disciplinary areas, within the geographic region, or across the national community.

# BD Spokes Solicitation (17-546)

Same 3 Themes: Data Sharing, Automation, Grand challenges

10 to 20 awards  
total \$10M anticipated

MEDIUM Grants

\$500,000 to \$1,000,000  
for up to 3 years

SMALL Grants

\$100,000 to \$500,000  
for up to 3 years

**Hub's Letter of Collaboration deadline: 6/19/2017**

**Proposal submission deadline: 9/18/2017**

**No Letter 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

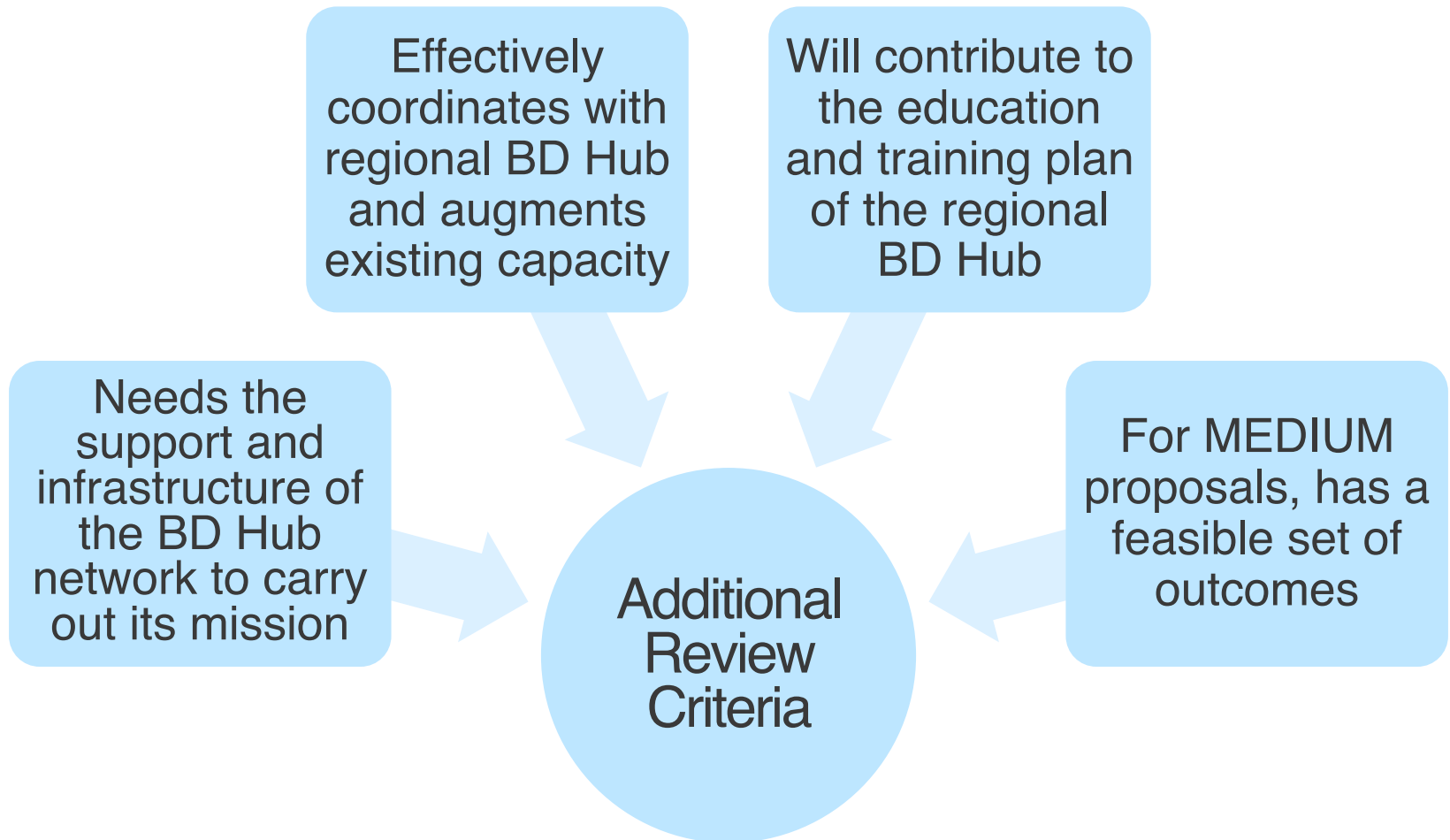


DATA ANALYTICS  
FOR SECURITY



# SOLICITATION SPECIFIC REVIEW CRITERIA

Additional to NSF standard Intellectual Merit and Broader Impacts criteria



# 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

# SUPPORT FROM BD HUBS

The regional BD Hub Steering Committee will provide general guidance to BD Spokes and assist with coordinating with the national BD Hub network, other BD Spokes, and the broader innovation ecosystem.

All proposals to the BD Spoke solicitation must include a letter of collaboration from a BD Hub coordinating institution.

# Letter of Collaboration from BD Hubs

Teams should fill out the form with the Hub that covers **the location of the lead PI's institution**. Submissions will be reviewed on a rolling basis.

*Please note that this is a limited submission and we recommend that PIs request Letters of Collaborations before their institutions decide on which proposal to submit.*

**Letter of Collaboration Request Deadline: June 19th, 2017**

For more info: <https://www.bigdatahubs.io/spokes17/>



## Building the Network

If the topic of a proposed SMALL or MEDIUM project is substantively similar to that of a previously funded BD Spoke, even if the activity will take place in a different BD Hub region, the proposal must justify the need to fund the new activity.

For example, the proposal can explain how the new activity might expand or augment the previously funded activities or differentiate itself from those other activities.

For proposals that expand or augment previously funded BD Spokes, the project team should coordinate in advance with those awarded BD Spoke teams; a Letter of Collaboration is encouraged

# AWARD INFORMATION & DUE DATES

Each Spoke project will be funded up to a maximum of \$1,000,000 for up to 3 years, subject to the availability of funds.

Each Planning Grant will be funded up to a maximum of \$100,000 for up to 1 year, subject to the availability of funds.

Total anticipated funding amount of \$10M

An estimated 9 Spokes and 10 Planning Grants is anticipated.

Letter of Intent Deadline:

Jan 12, 2016  
(5pm proposer's local time)

Full Proposal Deadline:

Feb 25, 2016  
(5pm proposer's local time)

# ELIGIBILITY INFORMATION

Proposals may be submitted by

- ✓ Universities and Colleges
- ✓ Non-profit, non-academic organizations
- ✓ State and Local Governments

NSF welcomes collaborative proposals that include for-profit organizations and FFRDCs. However, such organizations may only participate as subawardees.

An individual may only serve as the PI or co-PI in at most one submission.

# AWARD INFORMATION & DUE DATES

## 2016

- Each **Spoke** project will be funded up to a maximum of \$1,000,000 for up to 3 years
- Each **Planning** Grant will be funded up to a maximum of \$100,000 for up to 1 year

Total funding amount of \$10M

10 Spokes and 10 Planning Grants awarded

Full Proposal Deadline: Feb 25, 2016

Letters of Intent required

## 2017

- Small Spokes Grants (**SMALL**): total budgets from \$100,001 to \$500,000 for up to 3 years
- Medium Spokes Grants (**MEDIUM**): total budgets from \$500,001 to \$1,000,000 up to 3 years

Anticipated funding amount is \$10M

**10 to 20** awards are anticipated (both SMALL and MEDIUM)

Full Proposal Deadline: **Sept 18, 2017**

Letters of Intent NOT required

Limited Competition

Tangible outcomes (MEDIUM Grants)



# FOR FURTHER QUESTIONS CONTACT

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



NSF Headquarters, Arlington VA

The script from today's webinar will soon be made available on the webinar's event page – at the link shown [here](#) – and the presentation slides will also be posted in the coming days.