Knology

Northeast Big Data Innovation Hub

Mid-Point Evaluation

July 31, 2020

Nezam Ardalan, Rebecca Joy Norlander, Uduak Grace Thomas,
Kathryn Nock, Joanna Laursen Brucker, & John Voiklis

Recommended Citation

Ardalan, N., Norlander, R. J., Thomas, U. G., Nock, K., Brucker J. L., & Voiklis, J. (2020). Northeast Big Data Innovation Hub: Mid-Point Evaluation. Knology Publication #NSF.159.560.01. New York: Knology.

Date of Publication

July 31, 2020

Updated November 13, 2020

Prepared for

Florence Hudson
Executive Director
Northeast Big Data Innovation Hub,
Columbia University
475 Riverside Dr., Room 320-1
New York, NY 10115

This material is based upon work supported by the National Science Foundation under Grant No. (NSF #1946615, #1946558, and #1947000). Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

Knology produces practical social science for a better world.

tel: (347) 766-3399
40 Exchange Pl. Suite 1403
New York, NY 10005

tel: (442) 222-8814
3630 Ocean Ranch Blvd.
Oceanside, CA 92056

Knology Publication #

NSF.159.560.01

Executive Summary

Knology is leading independent evaluation of the outcomes of various efforts led by the Northeast Big Data Innovation Hub, referred to as the "Northeast Hub" or the "Hub" in this report. Specifically, the evaluation explores the impacts of programmatic activities aimed at increasing stakeholder engagement in big data science, developing shared resources and services, and data science education and workforce development. In order to assess these impacts, the evaluation design relies on stakeholder interviews conducted approximately a year apart, allowing evaluators to measure change over time. This is the first of two reports, written following the first round of stakeholder interviews.

In each round of interviews, we are gathering data from two distinct groups of stakeholders to understand their varying perspectives: **Project participants** are data science researchers and practitioners whose role was explicitly described in the grant proposal to the National Science Foundation (NSF), the funding mechanism supporting the Hub; and **Community participants,** who were not listed in the NSF grant proposal, but have engaged with the Hub in a variety of ways and to varying degrees.

Knology is evaluating data from these interviews through the lens of five specific measures that are designed to measure the Hub's progress towards its strategic goals. Generally speaking, these measures assess the Hub's ability to reach various audiences, provide a forum for data science-focused partnerships and collaborations, the benefits of being affiliated with the Hub, the degree to which its activities support the field's projects and goals, and how its work is helping to shape the field. Although the evaluation was structured to collect data from two participant groups, the distinctions between Project and Community participants were largely unimportant to the analysis, as the group type did not correlate to level of engagement with the Hub or the meaningfulness of interaction. As such, these distinctions aren't highlighted in the responses that follow.

In terms of its **Reach / Engagement**, the interviews indicated that participants have various levels of connection and participation with the Hub. We found evidence that the Hub has convened stakeholders across multiple sectors including academia, industry, nonprofit, and government. It provides a forum for data science researchers and practitioners to discuss topics of interest and forge collaborations, some of which have gone on to receive funding. This ties into the Hub's efforts to support the **Integration / Interconnectedness** of the community. We found evidence that the Northeast Hub effectively facilitates cross-sector networking and collaboration. This is crucial for breaking down research silos within academia as well as forging stronger partnerships between academia and industry. And the Hub's efforts are translating into projects with potential applications to societal challenges.

Simultaneously, the interviews also indicated that some members of the Hub feel less connected to its efforts than they would like. The two most frequently mentioned reasons were a lack of awareness about the breadth and depth of the Hub's work, and tight schedules that leave little time for additional activities. These were highlighted not as criticisms of the Hub's work but as a call for the entity to build upon the work that it is already doing. This is important to note because there is evidence that the Hub's efforts have

an **Impact** on its affiliated stakeholders, with potential for even greater impact. Several participants described a legitimizing effect of the Hub on the data science field. While some of these effects were intangible, there was a clear sense that mentioning the Hub and its support from the NSF had helped open doors for participants because it lent an air of credibility to their efforts. Similarly, the **Scale** of the Hub's contributions to the data science field have at least indirectly helped to accelerate the maturation of the field, according to some participants, more quickly and decisively than would otherwise have occurred. Affiliation with the Hub has also contributed to the shaping of data science university programs as well as helped to propel some research forward at a faster clip.

On the topic of **Representativeness**, the evidence suggests that the Hub's activities do align well with those of its stakeholders. Several said they felt that they had shared interests with the Hub and that its activities were relevant to their respective goals. In fact, participants hoped that being aligned with the Hub would give their projects greater visibility. Furthermore, they noted that being part of a united entity carries more weight when starting a conversation with potential partners in other sectors. As one interviewee stated, *"[If] the Hub represents all institutions in the region it's much more credible, you can get people to show up who would not otherwise show up to these conversations."*

Table of Contents

Introduction

The mission of the Northeast Big Data Innovation Hub ("Northeast Hub" or "Hub") is to build and strengthen partnerships across industry, academia, nonprofits, and government to address societal and scientific challenges, spur economic development, and accelerate innovation in the national big data ecosystem. Within that context, the Hub was designed to be a community convener, collaboration hub, and catalyst for data science innovation in the Northeast Region. The Hub in its strategic goals has chosen to focus on translational data science, aiming to equip community leaders with access and knowledge regarding data science so that these leaders can develop new and informed solutions to social change.

In June 2019, the Northeast Hub received its second round of NSF funding (#1916585) along with Hubs in the South, Midwest and West. The first round of funding was under the Big Data Spokes program and is only referred to within this report so far as interviewees commented upon it. Under this second round of funding, in addition to supporting core activities of the Northeast, South, Midwest and West Hubs, NSF also required the Hubs to build a strategic plan[1], create seed funding opportunities, and design a National Coordination Committee (NCC) across the four Hubs.

Knology is leading independent evaluation of the outcomes of various efforts led by the Northeast Hub. The evaluation is designed to assess the progress the Hub has made in relation to its four strategic goals:

1. Build collaborations to address real-world challenges through translational data science approaches;
2. Foster innovation and scale endeavors that reflect regional interests and align with national priorities related to data science;
3. Support and promote representative community engagement / impact across all Hub activities; and
4. Increase data science capacity and talent, emphasizing underserved communities.

Specifically, the evaluation will assess progress along five measures:

- **Reach / Engagement** measures the number and variety of communities affected by Hub activities and the number and variety of partnerships that the Hub works with;
- **Integration** or **Interconnectedness** measures the extent to which Hub partnerships amplify and cross-pollinate Hub-related activities to create economies of scale;
- **Impact** measures the perceived added benefit for stakeholders related to collaboration or integration with the Hubs, and the degree to which those who volunteer their time on Hub activities are satisfied with those efforts;

[1] The strategic plan for the Northeast Hub is available at this link *http://nebigdatahub.org/wp-content/uploads/2020/06/NEBDHub-Strategic-Plan-6.1.2020.pdf*

- **Representativeness** is the extent to which Hub activities align with and support the goals of partner and community projects; and
- **Scale** is the extent and depth of change attributable to the activities, projects, and initiatives.

About this Report

We recognize that the Hub is engaged in many different aspects of growing and supporting the data science community. In designing this evaluation, Knology seeks to capture the breadth and depth of the Hub's current activities, measure progress against goals to date, as well as identify ways that the Hub might seek to expand its impact during the following years of grant funding. As such, Knology conducted a first round of interviews with 22 data science researchers and practitioners across a range of institutions and sectors to understand the scope of the Hub's efforts and how those efforts have produced meaningful results for the data science community. The interviews provided insight into how the Hub works with the community to develop and leverage data science approaches to catalyze action from knowledge and insight, and enable the development of new solutions to societal problems. As a result, the analysis does not expressly comment on the NCC or Seed Funding directly unless raised by the interviewees.

To capture further progress made along the five measurement areas, Knology plans to conduct a second round of interviews with the same group of participants. The evaluation design originally allowed for a year between rounds of interviews, but the exact timeline remains pending.

Background

The Northeast Big Data Innovation Hub was initially conceived to help support data science knowledge and application, serving as a way to coalesce the efforts of researchers and practitioners working with large datasets. The Northeast region includes collaboration between project partners working primarily in Connecticut, Maine, Massachusetts, New Hampshire, New Jersey, New York, Pennsylvania, Rhode Island and Vermont. To address local issues in collaboration with cross-sector communities across the Northeast, the Hub initially structured work within four application areas and four cross-cutting themes:

Application Areas

- Education
- Health
- Rural / Urban Spectrum
- Science

Cross-Cutting Themes

- Data Literacy
- Data Sharing

- Responsible Data
- Privacy & Security

In the NSF proposal, each application area, as well as each cross-cutting theme, was designed to have a Volunteer lead from the community. Additionally, twelve projects that focused on these specific application areas and themes were highlighted as part of the proposal. Each project had a specific team named in the proposal.

Proposal Projects

- EmPAWRed Exposome Data Exchange Pilot: Reducing Barriers to Data Access and Interoperability: *Health, Rural/Urban, Data Sharing, Privacy & Security, & Responsible Data Science.*
- Smart Cities Data Exchange: *Health, Rural/Urban, Science, Data Sharing, Privacy & Security, & Responsible Data Science.*
- Seamless Licensing Agreements Collaboration: *Data Sharing, Privacy & Security, & Responsible Data Science.*
- Responsible Data Science in Government: *Rural/Urban, Science, Data Sharing, Privacy & Security, & Responsible Data Science.*
- Building Capacity with The Carpentries: *Data Literacy, & Responsible Data Science.*
- Data Science for All: *Data Literacy & Responsible Data Science.*
- Management Training: Sandra Will, Will Squared LLC: *Data Literacy, & Responsible Data Science.*
- Collaborative Resource and Understanding eXchange: *Data Literacy & Responsible Data Science.*
- Learning and Career Development Initiative: *Education & Responsible Data Science.*
- Outreach on Educational Data Mining: *Education, Data Literacy, & Responsible Data Science.*
- Cybersecurity Risk Initiative*: Privacy & Security, & Responsible Data Science.*
- Computational Steering: *Science & Responsible Data Sharing.*

The Hub, housed at Columbia University, is a collaborative entity guided by many voices in order to sustain impact across such a diversity of societal issues. The project team (PI, Co-PIs, Executive Director, Project Manager) has been supported by a Steering Committee, Advisory Board, Seed Fund Steering Committee, Project Activity Leaders and the NCC.

Methods

Participants

Knology worked with the Executive Team at the Hub to identify individuals who could speak to the impact that the Hub has had to date as well as opportunities for success moving forward. Ultimately, the Executive Team and the evaluators created a list of individuals representing people named in the 2018 proposal to NSF either as project team members for the application areas, cross-cutting themes, or projects (Project Participants), as well as individuals whose engagement, while not originally in the grant, has been critical or evolved since the funding was awarded in 2019 (Community Participants). The evaluation design

recognizes that projects evolve over the course of implementation, and that participant levels of engagement vary over time.

A total of 22 participants agreed to do interviews, fairly evenly split between the two Participant groups:

- **Project participants** (*n* = 12) were defined as individuals named in the NSF proposal as members of a team working on an application area, cross-cutting theme, or project;
- **Community participants** (*n* = 10) were defined as individuals who have had some connection with the Hub which was either undefined in the proposal or has evolved since.

The Hub's Executive Team connected the interview candidates with Knology evaluators, who subsequently scheduled the interviews.

Among the Project and Community Participants, some were also part of the various support committees outlined in the previous section. Of the 22 participants, 6 were either on the Hub's Steering Committee, the Advisory Board, or a co-PI. To preserve anonymity, we refer to all six of these individuals as members of the Hub "Leadership Team" when we note their responses in the analysis.

Instrument

The Interview protocol was designed to assess progress on the five measures guiding the evaluation, listed in the introduction. It consisted of eight primary questions with secondary follow-up questions or probes as needed. Interviewers adapted the protocol to each interviewee based on their personal history and connection to the Hub. Interview protocols for Project versus Community participants varied only slightly in how they framed similar questions. See the Appendix to this report for both interview protocols.

Three Knology researchers participated in data collection, rotating through roles. Two Knology staff attended each interview, one to conduct the interview and another to take notes. All interviews took place over Zoom and were recorded with verbal consent. Each interview lasted for approximately 30 minutes.

Analysis

Analysis of the interview data was performed by the same researchers who conducted the interviews. We used a grounded theory approach to see what themes surfaced in the conversations organically. We chose this approach because the participants had such diverse relationships and experiences with the Hub, and grounded theory allows for the emergence of unanticipated and disparate findings (as opposed to, say, developing a coding scheme at the outset of analysis). We then used the five measures as a way to interpret and provide structure to the analysis, as is evident from the headings in the Results section of this report.

While the groupings of Project and Community participants were helpful for the interviewee selection process, data from both groups were largely analyzed together. Due to the broad criteria for inclusion in either group we did not anticipate meaningful analytic categories to

result from the interviews. However, we used designations when knowing more about interviewees' roles provided additional insight into their comments.

Throughout this report we refer to individuals using the singular "they" to protect identity.

Results

The Hub Demonstrates Reach and Engagement

Reach / **Engagement** measures the number and variety of communities affected by Hub activities and the number and variety of partnerships the Hub works with. We understand this to mean how the Hub occupies a central role as a **convener** that has engaged stakeholders through **events** and **communication** efforts.

Our analysis resulted in the following key findings related to Reach / Engagement:

- The Hub acts as convener for a wide variety of stakeholders, resulting in tangible outputs;
- Members of the data science community have benefitted from past events and recognize the Hub's outreach efforts, but acknowledge room for improvement; and
- Stakeholders are eager for additional ways to engage with the Hub but not always sure how to do so.

Although both Community and Project participants had varying levels of engagement with the Hub, the majority of all interviewees described the Hub's role as a convener as key to building collaborations and fostering innovation. Evidence suggests that the Hub has effectively connected stakeholders across research, academic, nonprofit, industry, and – to a lesser extent – governmental sectors. One result of this convening function has been the development of collaborative grant proposals, some of which are aimed at advancing the inclusivity of the data science community in the Northeast region. For example, one interviewee had recently submitted a proposal to work with teachers in Title I schools to bring data science education to students in under-resourced areas. The interviewee indicated various members of the Hub that were integral to the proposal project, saying, *"These relationships are very much alive, it's a network that's really important to the work we've been doing."* A Leadership Team member reiterated this idea, saying, *"I've benefitted from the ecosystem that has emerged around [the Hub]."*

Furthermore, interviewees described the Hub as having an *"umbrella role"* that unites data science activities across the Northeast region. When asked what might be lost if the Hub didn't exist, one participant emphasized that this convening function is vital. *"If the Hub didn't exist, or if it went away, then the [loss of the] opportunity to have a regional convener for activities around big data and the opportunity to work with that convener would be unfortunate."*

Interviewees, especially Project Participants who had been involved with the Hub for many years, spoke positively about Hub-organized events. One described a Hub-related convening as *"one of the best events I attended,"* and many others expressed a desire for increased involvement in Hub activities. Interviewees felt that regular events in different Northeast locations would help keep the Hub *"top of mind,"* and increase engagement though they acknowledged the prohibitive expense of in-person gatherings. One participant felt that having a talk series over Zoom could bolster the Hub's public presence. This participant also noted that now would be an appropriate time to begin virtual meetings or workshops given

the current restrictions on gatherings due to COVID-19. Several other interviewees also noted the potential benefit of online convenings while simultaneously recognizing the general feeling of saturation with video conferencing. For future Hub-related events, one interviewee suggested that the Hub could provide clearer, actionable follow-up.

One respondent commended recent Hub efforts related to website redesign and others appreciated the Hub's community outreach efforts via its newsletters. However, the interview data suggests that the broader data science community may be unaware of all the Hub's activities. Some attributed their lack of awareness to the general chaos surrounding COVID-19. Participants mentioned feeling out-of-touch or only knowing about annual events, while others were aware of more frequent Hub-related activity.

The varying degrees of awareness may be related to their level of engagement with the Hub, or with the Executive Team directly. The level of connection also seemed to vary with geographic location. Generally speaking, interviewees that were based in New York City had more frequent engagement with the Hub and knew more about its efforts. This suggests that the Hub may have a stronger presence there than in some other Northeast areas. Some interviewees noted that a communications and outreach strategy aimed at reaching participants in other parts of the Northeast could help broaden engagement with the Hub. Along these lines, another participant suggested that the Hub could emphasize Spanish-speaking workshops or events as one way to be more inclusive with outreach efforts.

Overall, participants wanted the Northeast Hub to have more frequent activity, whether through online activities or through events and publications. One participant stated that it would be helpful for the Hub, *"To get back on everybody's radar on a more regular basis so … that opportunities and engagement areas are more fully known."* Considering other opportunities for deeper engagement, one participant mentioned utilizing volunteer-led working groups to help the community more fully engage with the Hub and to capitalize on areas of shared interest. One interviewee felt that the Hub could play a bigger role in supporting efforts to educate members of the public about big data or artificial intelligence. They noted that this could include support for learning happening in public libraries or community centers which are attended by non-academic audiences.

Another frequently mentioned obstacle to participating in Hub-related activities and events was time. This was true for both Project and Community participants, who said that although they saw the Hub as a valuable entity, finding time in their schedules to attend events and other activities was often a real hurdle. Yet even those who had the lowest levels of current engagement seemed to desire to re-engage, suggesting that they found value in the Hub and that time limitations were not insurmountable. However, they were occasionally unsure about how to become more involved, how proactive to be, or what stage in a research project was optimal for pursuing the Hub's support.

Interviewees had a number of suggestions to increase overall reach and engagement. One saw the need to *"bring in new blood,"* noting that there has been limited infusion of new participants since the Hub's founding. Others had specific suggestions for getting more participants involved. This included things like sharing newsletters spotlighting opportunities across the field (such as hackathons or creating open source datasets), and using the Northeast Hub website to offer clear invitations to action or specific types of involvement.

They acknowledged having to look across many sources for opportunities in data science and felt the Hub as a centralized communicator and resource would be extremely valuable for the field.

The Hub as a Unique Force for Integration and Interconnection

Integration or **Interconnectedness** measures the extent to which Hub partnerships amplify and cross-pollinate Hub related activities to create economies of scale. We understand this to mean ways that the Hub promotes cross-sector collaboration and works to reduce disciplinary silos. It also includes how the Hub connects academic research with real-world application. Interviewees, aware that much of this connecting function falls on the Hub's Executive Team, offered strategies to help amplify efforts while reducing burden.

Our analysis resulted in the following key findings related to Integration and Interconnectedness:

- Networking efforts of the Executive Team are recognized and appreciated by stakeholders;
- The Hub has successfully facilitated collaboration within academia as well as between academia and other sectors; and
- The Hub connects data science research with opportunities for meaningful application, a function that is particularly valuable for data science students.

By far the most frequently mentioned benefit of the Hub was related to networking, as a way of generating new ideas and pushing the boundaries of the data science field. Most participants said they had benefited from being in the Hub network or that they expected to benefit in the future.

"*It's a place to meet people... There are times and places where I've run into people who I wouldn't otherwise run into, [and] that has led to ideas that would not otherwise have occurred. I think that is the principle outcome of participating in the Hub.*"

A similar comment confirmed the idea that the existence of the Hub serves to break down barriers that normally exist in the field of data science: *"There are a lot of really smart people in the Hub who I wouldn't naturally bump into without something like the Hub, so I think that's a real benefit."*

Evidence suggests that direct relationships with either current or past members of the Executive Team are largely responsible for the Hub's identity as a connecting force. Interviewees unambiguously praised both previous and newer members of the Executive Team and their ability to facilitate integration:

- *[Names] are a "superstar team!"*
- *"[Name] does a hell of a lot of work…they're awesome. If [the Hub] lost them they'd be in big trouble. They're a lynchpin of the organization."*
- *"[Name] is probably the best person to do [networking] because they are definitely a connector."*

- *"[Name] was really good at reaching out and saying 'Hey you're looking at this issue involving data and [another researcher] who is in health care is also looking at data. Maybe you two should talk?'"*
- *"Somebody like [Name]…to be such a supporter and really championing our work and connecting us with people, that's invaluable."*

Interviewees commented on the unique position of the Hub to facilitate connection between academic researchers who tend to operate in disciplinary silos. One interviewee who works extensively with research networks called this the *"cross-fertilization of research activity."* This function is particularly pertinent because collaboration across different institutions can be challenging, due to self-interest (in particular when it comes to recognition for research and procuring funding).

Those interviewed for this evaluation clearly identified the Hub as a body able to connect researchers within and across institutions of higher education. For example, a member of the Leadership team commented that, "*the Hub has the opportunity to bring together people who are working in data science in different research fields together in a way that probably wouldn't happen without something like the Hub. I think that is an irreplaceable role."* Another participant noted a similar idea, saying, *"The Hub enables, in a very non-threatening environment, academia to come together for a greater purpose then themselves.*" Ultimately, this interconnectedness can lead to a greater capacity that is not possible when institutions work separately. As one participant noted, *"I like to talk about the Hub as taking on challenges that are above and beyond what any single institution can do alone."*

Additionally, several participants specifically highlighted the value of having a central entity that integrates diverse stakeholders from across different sectors in ways that can profoundly shape the field of data science. In the words of one interviewee,

"If the Hub didn't exist, how would we have those kinds of conversations? Who would bring together all the key players that are involved in decision making for data science, especially for data science education and training? Who would bring those people together if we did not have the Northeast Big Data Hub?"

Participants also felt that the Hub could do more to broker relationships between academia and other sectors like industry, government, and nonprofits to address real-life challenges facing society. One participant viewed the Hub as primarily an academic entity but saw its potential for impact as a connector with industry. Another agreed, noting that not all academics are interested in pursuing partnerships that focus heavily on real-world engagement, but those who do, tend to be *"self-sorting."* Two mentioned the use of Research-Practice Partnerships (or RPPs) as a possibility to advance cross-sector integration.

Interviewees felt that focusing on problem-solving (and the cross-sector collaboration that makes it possible) has particular value for data science students who will ideally emerge from their educational programs ready to tackle societal challenges. One interviewee said that they became aware of the Hub through an internship opportunity while in graduate school and now focuses on using technology for social justice. They attributed their career choice at least in part to the real-world experience provided by the internship. They also highlighted a need for a continued focus on connecting research to its various applications, indicating that societal challenges have *"put a spotlight on big data, both positives and*

negatives of how it can impact our world. We're starting to see an important need for the Northeast Big Data Hub to be at the center of the conversation."

The Impact of the Hub on Stakeholders

Impact measures the perceived added benefit for stakeholders related to collaboration or integration with the Hubs, and the degree to which those who volunteer their time on Hub activities are satisfied with those efforts. We understand this to mean how individual stakeholders rank their own satisfaction based on engagement with the Hub, as well as what overall impact on the data science field can be attributed to the existence of the Hub. Our main findings related to impact include:

- The existence of the Hub has a legitimizing function on the emerging field of data science; and
- Independent of their level of personal satisfaction with the Hub, interviewees presented thoughtful feedback on how the Hub can prioritize its efforts in an attempt to increase its impact.

Various interviewees highlighted the role the Hub plays in legitimizing their work, particularly with institutional colleagues or administrations. They felt that the Hub's name and reputation carried weight, as did its association with NSF. A few mentioned specific instances where they thought this legitimization occurred, such as using the Hub's name when creating a new graduate program at the university where they worked, how mentioning the Hub's support gave credibility to their proposals, or that being affiliated with the Hub had improved their chances of receiving funding or resulted in funded projects. One participant described how being affiliated with the Hub simplifies the task of justifying their work, saying, *"It's much easier for me. That's why I usual defer and say, we are working with the National Science Foundation Big Data Hubs."*

Some participants felt that the Hub legitimizes the data science field as a whole. Data science was described by many interviewees as an emerging field, one that is still taking shape, a process that felt to one like "*building the plane while it's flying*." Interviewees spoke about the Northeast Hub and other regional Hubs' visionary and defining role in this field-building process, expressing opinions such as, *"The Hubs are playing a major role in the emergence of data science as a discipline."* Data, and big data in particular, are complicated to work with for many reasons – issues of privacy, compatibility, and ownership all come into play. The Hub, as a legitimizer, can work across these issues and help smoothen the process. One interviewee summed up this thought by saying,

"Having the Hub at the table at those conversations makes it possible. It's very hard for an individual PI to approach a city or some entity that's gathering data, so you need a trusted intermediary that makes those conversations possible."

Impact can also be gauged in terms of how satisfied stakeholders are with their relationship with the Hub. The level of satisfaction across participants varied widely. On the positive side, one interviewee (who has had a close and long-term relationship with the Executive Team) said, *"Out of ten, maybe a ten…just because they've been such*

champions of our work, we wouldn't be where we are now." Another was dissatisfied, citing a lack of clarity about how to engage with the Hub and provide value. Other responses fell across the spectrum, depending on personal experience and individual context.

Virtually all interviewees shared suggestions for the Hub to increase its impact moving forward. While many of these ideas relate to the other four measurement areas (and are discussed elsewhere in this Results chapter), we thought it would be helpful to present them as a single list in this section, as all are intended to increase impact in some way.

Specific ideas suggested by participants include:

- Workshops to educate the public about big data and data science topics;
- Events that bridge academia and industry;
- Greater web presence to spotlight current data science projects and promote broader awareness of data science;
- Regularly scheduled newsletters and virtual meet-ups to keep the field abreast of current research trends, opportunities, and funding possibilities;
- Support for data science standards and defining of common terminology used by the field;
- Access to big data sets by the public and specific communities who can benefit;
- Serving as a clearinghouse or resource archive for data science; and
- Outreach efforts across the Northeast region.

The Hub is Aligned with the Data Science Community

Representativeness is the extent to which Hub activities align with and support the goals of partner and community projects. We understand this to mean whether and in what ways stakeholders working across the data science community feel supported by the Hub. Our findings related to this measure indicate that:

- Certain stakeholders feel supported by the Hub;
- The Hub is considered an important way to increase the visibility of its members' activities; and
- Interviewees indicated that having a unified voice for data science in the Northeast Region is valuable for advancing the field.

Perhaps unsurprisingly, those interviewees who had longer and closer relationships with the Hub or the Executive Team felt the most supported. When asked to describe their relationship with the Hub, one such interviewee, who continues to work with the Hub, acknowledged, *"It's hard to keep it concise, the Hub has been part of my life for the last six years."* The interviewee went on to explain how they and other stakeholders in the Hub came together due to their shared interest in data literacy, and then applied for and received a planning grant. In this and similar cases, our analysis showed a high degree of alignment between Hub activities and the goals of individual stakeholders.

Those without this deeper level of engagement most often thought about the Hub's support in terms of promoting awareness or visibility on an individual and collective basis. They acknowledged working on a wide range of data science projects and initiatives, and saw a role for the Hub to leverage its network to help make those activities more visible.

Collectively, interviewees understood alignment to mean supporting the field to speak together with one voice, especially in conversations between academia and industry. They felt that this was especially important as it relates to the process of making large datasets available for research. One participant commented on this need, indicating there was a general *"lack of appreciation for the complexity of the process leading up to a uniform data set."* They suggested that the Hub could align with those who are already working on this issue, including translational and applied research scientists on the academic side as well as thought leaders from industry. Interviewees thought that doing this would create a feedback loop, increasing the authority of the Hub as spokesperson. *"[If] the Hub represents all institutions in the region it's much more credible, you can get people to show up who would not otherwise shown up to these conversations."*

The Hub is Positioned for Long-Term Scale

Scale is the extent and depth of change attributable to the Hub's activities, projects, and initiatives. In our analysis, we understand scale to be best represented by the ways the Hub is contributing to accelerating and amplifying the work of data science researchers. Key findings include:

- The Hub's efforts to date have helped advance the field more quickly and decisively than would otherwise have occurred;
- The Hub has been instrumental in promoting broader awareness of certain aspects of data science – such as data ethics – and can provide additional value by amplifying the work of its stakeholders.

Overall, interviewees felt that NSF's investment in the Hub(s) has paid off by helping the field establish a stronger sense of identity. *"I think that data science has gone, during the lifetime of the Big Data Hubs, from a 'What are you doing?' to a 'Oh yes, this is one of the most important things we could be doing*.'" Several stakeholders acknowledged that the field is maturing quickly, indicated in part by the growing number of programs in universities, and attributed some of this acceleration to the work of the Hub.

They acknowledged that many projects and initiatives could have happened without the Hub's involvement, but that the Hub's efforts have helped speed up some of the work. As one participant said,

"It's hard to pinpoint something and say that would never have happened…but there are lots of places where things have moved along faster and come to fruition in ways that they would not otherwise have if the Hub were not in place."

Another participant emphasized the Hub's role as a catalyst for certain types of research projects noting that, *"The Big Data Hubs have been an astoundingly good and successful way of getting science to turn on a dime, so to speak."*

We also found evidence suggesting that Hub efforts have helped some researchers think differently about the data science field, and they are sharing that knowledge with others in the community. Following one Hub in-person meeting, a university-based researcher commented that the meeting *"actually put me in a whole new path of research I hadn't been in, which is ethics and data science, which has proven to be very important."* The researcher went on to describe a project where they had doctoral students try to explain the ethics of their work. This exercise led them to realize that the students lacked formal training or exposure to the ethics component of their work. The interviewee mentioned the *"scalable value"* of taking what they learned at the Hub event, and working it more systematically into university practice.

A small number of participants noted that the Hub helps to disseminate and / or publicize their work, and that without the Hub, the data science community would lose a space to highlight projects and research that haven't been published or shared broadly. Interviewees felt that the Hub's capacity for amplifying the efforts of stakeholders would result in scalable impact. One interviewee noted that this could be done through co-branding efforts, and taking advantage of other established campaigns that already have traction (such as Cybersecurity Awareness Month). Two participants who are currently involved with research networks noted the potential for these and other networks to amplify or promote further awareness of the Hub and its activities. They described it as a *"two-way street"* where they can provide the Hub with input from the data science communities they work with, while promoting the Hub's messaging across their networks.

Discussion & Recommendations

Overall, responses from both the Community and Project participants interviewed for the evaluation suggest that the Hub is making good progress towards its stated goals. These data suggest that Hub occupies a central role, one that positions it to engage stakeholders from various sectors of the data science field. And it is able to do so in a way that brokers deeper and richer collaborations than any individual institution would be able to do. In large part due to its role as a convening entity, respondents felt that the Hub would be well suited to address some current gaps in the field primarily around gathering and sharing large datasets. They also observed that there a few current areas of activity where the Hub could do more to maximize its efforts.

Build collaborations to address real-world challenges through translational data science approaches

There is a significant amount of evidence from the responses in support of the Hub's role in enabling collaboration, often between individuals who might not otherwise have connected. Participants mentioned conversations during Hub-organized events or named specific members of the Hub's Executive Team as crucial to facilitating interactions. This web of connections made through the Hub has resulted in funded projects as well as fostered collaboration across disciplines, institutions, and sectors. One important benefit of brokering relationships between academia and other sectors like industry, government, and nonprofits is the opportunity to leverage these partnerships to address real-life societal challenges. Several participants mentioned the possibility of using research-practice partnerships as a way to accomplish this.

Given the importance of its role as a connector, several participants interviewed for this evaluation wanted the Hub to have a much stronger presence across the Northeast region, voicing a desire for clarity around its efforts to engage stakeholders. This suggests that there is room for the Hub to be a bit more strategic in its communication approach and perhaps leverage a broader range of outreach mechanisms to inform the community about planned events and activities. We offer some suggestions for this purpose in the recommendations section below.

Foster innovation and scale endeavors that reflect regional interests and align with national priorities related to data science

While we cannot really speak to regional interests or national priorities as this was not the focus of the evaluation, we can speak to innovation and scale. In general, interviewees felt that the Hubs as a whole have contributed, at least in part, to the rapid maturation of the data science field. Although the exact particulars of how the Hubs have accelerated the field are a little unclear, some participants were able to mention specific instances where their affiliation was of benefit. For example, a participant mentioned that their affiliation with the Hub was supportive in creating a data science program at their university. Several participants also described the Hub as a catalyst for research projects. Related to this,

various interviewees highlighted the Hub's role in legitimizing their work to institutions and funders.

A few participants noted the role that the Northeast Hub plays in spotlighting projects and research that had not been published or disseminated broadly. In addition to increasing the visibility of data science projects, this particular role has implications for the Hub's goal to foster collaborative innovation moving forward by providing an alternate – yet highly visible – space for sharing the work of its stakeholders, beyond traditional academic outlets. Moreover, spotlighting projects in this way can lead to new funding opportunities or help collaborations target real-world problems.

Support and promote representative community engagement / impact across all Hub activities

Although the levels of awareness and involvement varied, evidence from this first round of interviews suggests that the Northeast Hub is working to engage members of the data science community across the board. Several participants specifically highlighted the Hub's role in building bridges between data scientists and practitioners within academia as well as between academia and industry. Most felt that the Hub has made progress toward creating a vital and interconnected network that would be lost if the entity were no longer there. And many recognized that there were ample opportunities to bring in stakeholders from other sectors, including more industry players as well as people from non-profits, government, and even the general public. Partnering with libraries who already offer programs for the public was one of the ways participants felt the Hub could engage more deeply with non-academic audiences.

Increase data science capacity and talent, emphasizing underserved communities.

The general consensus from our analysis of the interview responses was that the Hub has an important part to play in increasing data science capacity, access, and opportunity for the broader community of potential stakeholders. Participants recognized the Hub's past efforts to engage the data science community through various activities including in-person and virtual events. This is work that many hope will continue and expand. Some of the concrete suggestions that came out of the analysis include promoting specific opportunities that are designed to increase participation from students, and to better engage with communities that have been traditionally underrepresented in fields related to data science. One concrete way that the Hub could increase students' participation is by connecting them to more internship opportunities. These internships could focus on academic applications of data science or on applications of data science methods to important societal challenges. In terms of general accessibility, one suggestion was to co-sponsor workshops and events in commonly-spoken languages like Spanish with partners who are already engaged in that work.

As noted above, participants felt that the Hub could refine its communication strategy around activities and opportunities to encourage greater involvement from all facets of the community, and to better connect participants in regions of the Northeast outside of New York City. Besides awareness, another frequently mentioned barrier to engagement was

time. In the recommendations section, we suggest ways that the Hub can navigate concerns about time constraints.

Recommendations:

Based upon the interviews, we provide the following recommendations for expanding the impact of the Northeast Hub, understanding that some of these ideas may be best pursued jointly with other regional Hubs. We recognize that some of these recommendations are already occurring to various degrees. However, interviews made it clear that either the community lacks awareness (indicating a need for refining the Hub's communication strategy) or the frequency / depth of these activities should be increased. Wherever possible, acknowledging funding and staffing limits at the Executive team level, we have sought to identify ways that the Hub can leverage its impact through the work of its partners.

- **Community education events**: Co-sponsor / promote workshops, online classes and / or events to educate the public about big data and data science topics, prioritizing venues that are widely accessed by non-academic audiences whenever possible;
- **Bridge academia with real-world applications:** Co-sponsor / promote ways – events, discussion forums, etc. – that allow for academia to find applications across other sectors, such as nonprofit, industry, and government;
- **Web presence:** Utilize the new website and other social media tools to have a stronger web presence to promote broader awareness of data science, making use of existing networks with listservs to increase the reach of communications;
- **Sharing of new developments in the field:** Whether through regularly scheduled newsletters, the website or other platforms, become the distributer of current research trends, opportunities, and funding possibilities;
- **Flexible Scheduling:** Working with partners to vary the times when planned events occur or offering multiple versions of the same events may give more members opportunities to participate that fit with their schedules. Another option might be, where possible, to record events and make these publicly available for participants to view in their own time.
- **Promote diversity, equity and access:** Reach out to stakeholders who already have a focus on inclusivity and promote their efforts. Use internships and convenings as an opportunity to bring younger and newer voices into the field.
- **Contact points:** The Hub has access to a wealth of talent and expertise that can be utilized to a much greater degree. Create clear mechanisms for interested individuals to get in touch with the Hub and guidelines for when and why to reach out, along with specific ways they can help.
- **Volunteer-led working groups:** We concur with the suggestion from participants to clearly establish and promote working groups around different aspects of data science, led by volunteers from the community who are experts in those domains. This creates another opportunity for like-minded individuals to connect and collaborate on topics and issues of interest to them, and can be an additional way to engage new members.

Interviewees emphasized the importance of several larger undertakings that they saw as priorities for advancing the field. While we believe these to be outside the scope and capacity of the Northeast Hub, we note them as important long-term opportunities for

strengthening the field. We encourage the Hub to consider its role in addressing these greater challenges.

- **Clearinghouse:** Become a clearinghouse or resource archive for data science;
- **Standards Designer:** Design and promote standards for data science and coordinator for common terminology used by the field; and
- **Data Repository:** Become a central place for the sharing of or access to big data sets.

Appendices

Appendix A: Interview Protocols

Knology has been hired to help the NE Big Data Innovation Hub assess its impact. To do so, Knology will interview two groups of stakeholders: **project teams** and **community partners**. Two sets of interviews will occur with each group, reflected by the 4 sections shown in the following protocol. We have called these "Round 1" (expected to occur in April / May 2020) and "Round 2" (expected to occur following the NSF site visit in Spring 2021).

The evaluation is designed to explore:

- **Reach / Engagement** measures the number and variety of communities affected by HUB activities and the number and variety of partnerships the HUB works with.
- **Integration** or **Interconnectedness** measures the extent to which HUB partnerships amplify and cross-pollinate HUB related activities to create economies of scale.
- **Impact** measures the perceived added benefit for stakeholders related to collaboration or integration with the Hubs, and the degree to which those who volunteer their time on Hub activities are satisfied with those efforts.
- **Representativeness** is the extent to which Hub activities align with and support the goals of partner and community projects.
- **Scale** is the extent and depth of change attributable to the activities, projects, and initiatives.

Text enclosed with brackets (i.e., [text]) represent explanatory notes for the interviewer, and the participant will not see or hear these notes. Examples include: [Prompt] or [Follow-up]

Appendix B: Project Team Interview Script: Round 1

Introductory script

Thanks so much for agreeing to speak with us today. My name is [Interviewer's name] and my colleague here is [insert transcriber's name]. As you saw from our initial email, we're with Knology, a social science research organization that has partnered with the Northeast Big Data Hub to help understand their current community engagement efforts - and how the Hub might evolve or continue current efforts to increase the positive impact for the Hub and community in the future. As we said in the email, this is the first of two half-hour conversations, with a follow-up planned for about one year from now.

We'd like you to be as honest as possible during this conversation. While the Hub knows who we are speaking with, we will not identify interviewees in our reporting, so they won't be able to match up people with their specific comments. [name of notetaker] will be taking notes, while the two of us talk. If it's ok with you, we'd also like to record the conversation. We won't share this recording with anyone, it's just to make sure we capture the conversation accurately. May I record?

[If yes, begin recording, if not transcribe only]

Do you have any questions before we begin?

Thanks!

Questions

For our record, can you say your name again and a sentence or two about who you are and the work you do.

[REACH/ENGAGEMENT] Can you tell us a little about your relationship with the Northeast Big Data Hub. Anything you think is important for me to know.

> [Prompt] In what way do you interact? What value do/did you get out of these interactions?
>
> [Prompt] Are you currently working on a project together or have plans to work together?
>
> > [Prompt] How did that work originate (e.g., out of a Spoke project)?

[INTEGRATION/INTERCONNECTEDNESS & REACH/ENGAGEMENT]

Have you made any useful professional connections through the Hub?

How did that connection result in other valuable interactions?

> [Prompt] What value did you get out of this interaction? e.g., shared resources, maintain an ongoing dialogue, built a new collaboration or project, furthered your research?

[IMPACT] How do you feel the Hub supports or might support you or your work?

> [Prompt: Are there any goals that they are helping you to achieve?]

[Prompt – As applicable] Does the Hub support translation or transfer of academic research into a real-life application? [Follow-up] What evidence do you have that this is occurring? [Interviewer asks about specific projects, depending on interviewee and the goal(s) of their project(s).]

[Prompt] What is your level of satisfaction with your relationship working with the Hub? (FH question - are we planning to give them a high, medium, low range, or ask them to just tell us in their words?) Can you tell me more about the reasons for your answer?

[IMPACT] What are the biggest challenges or obstacles, if any, getting in the way of your progress on Hub activities?

[SCALE] What are the biggest opportunities that you see for your work with the Hub in the future?

[Prompt] Can the Hub help mitigate the challenges you just described or support new opportunities? e.g., Specific skill-building? Resources? Collaboration?

[REPRESENTATIVENESS] Thinking about the diversity of work currently happening in your area of expertise, where do you feel the Hub should place its priorities to ensure they are supporting the diversity of programs in the field?

[IMPACT] What might be lost if an entity like the Hub didn't exist?

Conclusion

We are nearing the end of our time together. I'd just like to ask you one final question. Is there anything else that you think is important for me to know that we didn't already cover in today's conversation?

If you think of anything you'd like to add you're more than welcome to email us. Thanks again for your insights.

Appendix C: Community Partner Interview Script: Round 1

Introductory script

Thanks so much for agreeing to speak with us today. My name is [Interviewer's name] and my colleague here is [insert transcriber's name]. As you saw from our invitation email, we're with Knology, a social science research organization that has partnered with the Northeast Big Data Hub to help understand their current community engagement efforts - and who they might work with in the future. As we said in the email, this is the first of two half-hour conversations, with a follow-up planned for about one year from now.

We'd like you to be as honest as possible during this conversation. While the Hub knows who we are speaking with, we will not identify interviewees in our reporting, so they won't be able to match up people with their specific comments. [name of notetaker] will be taking notes, while the two of us talk. If it's ok with you, we'd also like to record the conversation. We won't share this recording with anyone, it's just to make sure we capture the conversation accurately. May I record?

[If yes, begin recording, if not transcribe only]

Do you have any questions before we begin?

Thanks!

Questions

For our record, can you say your name again and a sentence or two about who you are and the communities you work with relation to Big Data.

[REACH/ENGAGEMENT] Can you tell us a little about your relationship with the Northeast Big Data Hub. We realize this might be a new relationship, but we would like to hear how you learned about the Hub and what you've done together so far – if anything.

[REACH] Have you made any useful connections through the Hub?

> [Prompt] Have you leveraged those connections in some way? e.g., engaged in a collaboration or program?

[IMPACT] How do you feel the Hub supports or might support you or your work? [Prompt: Are there any goals that they are helping you to achieve?]

> [Prompt – As applicable] Does the Hub support translation or transfer of academic research into a real-life application? [Follow-up] What evidence do you have that this is occurring? [Interviewer asks about specific projects, depending on interviewee and the goal(s) of their project(s)
>
> [IMPACT] What is your level of satisfaction with your relationship working with the Hub? What do you feel drives that level?

[IMPACT] What are the biggest challenges or obstacles you see for leveraging data in your field?

[SCALE] What are the biggest opportunities for leveraging data in your field that you see in the future?

> [Prompt] Can the Hub help you achieve your goals or support new opportunities? e.g., Specific skill building? Resources?

[REPRESENTATIVENESS] Thinking about the diversity of work currently happening in your area of expertise, where do you feel the Hub should place its priorities to ensure they are supporting the broad diversity of programs in your field or the broader data science ecosystem?

[IMPACT & SCALE] What else could the Hub do to better support your long-term goals or vision in support of data initiatives?

[IMPACT] What might be lost if an entity like the Hub didn't exist?

Conclusion

We are nearing the end of our time together. I'd just like to ask you one final question. Is there anything else that you think is important for me to know that we didn't already cover in today's conversation?

If you think of anything you'd like to add you're more than welcome to email us. Thanks again for your insights.

Appendix D: Response Letter from Northeast Big Data Hub

Letter to Knology in response to initial evaluation report of the Northeast Big Data Innovation Hub

The Northeast Big Data Innovation Hub extends our gratitude to the team at Knology for their work in conducting the first phase of their evaluation of NSF award #1916585. We look forward to keeping the team up-to-date on current and upcoming activities at the Hub, and to the insights they will provide in the second phase of the evaluation, slated for spring 2021. The below summarizes our response to the findings of the evaluation's first phase, including specific actions with timelines being taken by the Hub's Leadership team.

The mission of the Northeast Big Data Innovation Hub is to build and strengthen partnerships across industry, academia, nonprofits, and government to address societal and scientific challenges, spur economic development, and accelerate innovation in the national big data ecosystem. Our collaborations with this community are critical to this work. Consequently, we were delighted to hear confirmation in this report that our efforts provide real value to our community - including effectively facilitating cross-sector networking and collaboration, organizing activities that align well with the priorities of our stakeholders, and supporting our community by amplifying their work.

The Executive Team has discussed Knology's list of recommendations (see Appendix) for enhancing our existing impact, and how we build them into our strategy and operations. Our guiding principles for this work include:

- Be **proactive** in engaging with our community, listening to them, and inviting them to get involved. This includes proactive leveraging of other channels to our constituency, including partnering with Regional Research and Education Networks, and developing the COVID Info Commons Community.
- Be **accessible and inclusive**, reaching individuals and institutions of all kinds, from all parts of our region.
- Be **focused**. Our June 2020 strategic plan identified four focus areas, consolidating from a more widespread set of 8 after research and discussion with community stakeholders. By focusing our efforts, we grow our strengths and provide easier access points for engagement.

Specific actions in alignment with these principles that we have taken, are taking, or plan to take within the next quarter, include:

We will continue to develop community programs and organize events that serve as a bridge between academia and other sectors. These include the Northeast Student Data Corp Founding Committee, designed to bring together leaders from all sectors and students of all backgrounds to develop a program teaching data science fundamentals

475 Riverside Drive, Suite 320, New York, NY 10115 • nebigdatahub.org

to underserved populations. The initial cohort for the committee convened for the first time in October 2020.

Beyond our existing community events, we will reach broader and non-academic audiences by:

- Participating in existing conferences and events popular among the community as well as reaching communities we may not yet know. As a recent example, the Hub submitted a birds of a feather proposal on diversity, equity and inclusion to the Academic Data Science Alliance for October 2020, and offered it as a base for a larger BOF. Leading on this has resulted in combining the Chan Zuckerberg team in the BOF, and Fred Hutchinson Cancer Research Center.
- Organizing new community events. These include monthly COVID Information Commons (CIC) Community webinars, beginning with the CIC launch in July 2020 and continuing with monthly COVID-19 lightning talk webinars with invited COVID RAPID PIs from around the country, which kicked off in September 2020 and are slated to run through Q1 2021. We plan a CIC student paper competition by early 2021 to engage more students in the COVID-19 research community.
- Co-branding existing and new events with collaborator organizations, extending our reach. These include a Connected Healthcare Cybersecurity workshop cosponsored with IEEE, for which an organizing committee was first convened in October 2020. The event is expected to be held in Q1 2021. Other co-branded events include leveraging our partnerships with the regional RENs (Research and Education Networks) such as a webinar with KINBER for Pennsylvania schools and libraries held in September 2020, and a future event we are planning with NYSERNET for education and research communities in New York state.

We will continue to leverage our web presence and share new developments in the field by:

- Creating more content highlighting Northeast data science projects, outcomes, and opportunities, from community voices - e.g., success stories from our community and interviews with COVID-19 researchers, which launched during the summer of 2020.
- Incorporating topics highlighting new developments in the field, e.g. insights from CIC COVID-19 researchers, and new IEEE cybersecurity standards efforts.
- Expanding existing multi-channel communications - web, Twitter, LinkedIn, YouTube, Northeast Hub and COVID Info Commons Community newsletters, and program-specific channels such as Slack - on an enhanced production schedule deployed in August 2020.
- Meeting quarterly to review engagement analytics to inform communications strategy, with the first meeting slated for October 2020.

We are operationalizing our mission statement on diversity, equity and inclusion, created in the summer of 2020, through intentional outreach to and collaboration with

Knology Publication #NSF.159.560.01

underserved communities on an ongoing basis throughout all aspects of our programming. As a recent example, we reached out to 84 new institutions across the Northeast to invite participation in our 2020 seed fund, of which 32 were HBCUs, primarily black institutions, and Hispanic-serving institutions. The outreach list was built to encourage inclusion from a demographic perspective, and also geographically (reaching every state in the Northeast) and institutionally (reaching universities and two- and four-year colleges).

We will develop programs that support underserved communities, such as the Northeast Student Data Corps, in partnership with collaborators from these communities.

As an ongoing commitment to our community, we leverage flexible scheduling by offering multiple time slots and deadlines for Hub events and programs, as staffing and community interest allows, and schedule events to maximize participation from the Hub community. We will continue to leverage live streaming and recording to share our events more broadly via our multiple social media channels including the website, YouTube and Twitter.

We have integrated organization-level touch points for contact and collaboration throughout all aspects of our programming - as an example, see our form to attract new collaborators at nebigdatahub.org/collaborators. We will continue to leverage program-specific touch points, such as community Slack channels, topical webinars, and topical working groups with open calls for community participation. These include ways of contacting the Hub, but also ways for community members to contact each other.

To further our commitment to community participation, we disseminate open calls for community participation in volunteer-led working groups that further the Hub's mission. Current and future volunteer-led working groups include the Hub's Steering Committee, Advisory Board, Seed Fund Steering Committee, Northeast Student Data Corps Founding Committee, Connected Healthcare Cybersecurity Working Group with collaboration with IEEE on responsible data science for clinical IoT and cybersecurity, and Urban to Rural Communities Working Group, all of which are already active or expected to launch in 2020. We are building such opportunities for community participation in all future community projects.

Given the additional recommendations highlighted by the findings, we will explore roles for the Hub as a clearinghouse, standards designer, and/or data connection repository, leveraging existing programs discussed above. Building off our strength in community engagement, we will explore opportunities - particularly through volunteer-led working groups, web presence, and community events - to share resources and data sets and promote the design of standards and common terminology.

NORTHEAST BIG DATA INNOVATION HUB

We will publish Knology's report, including our response, on a dedicated web page, outlining our commitment to our community, our ongoing actions to scale our value to the community, a "fact sheet" on community engagement metrics, and a feedback mechanism for anonymous community input. We expect this page will be launched by December 2020.

We are grateful for the many opportunities we have to serve our community, and consider our work in doing so to be ever-evolving, with our goal to always learn and improve. We thank Knology for their role as evaluator on this project, and look forward to future insights from them and, as always, from our community themselves.

Jeannette Wing, Principal Investigator
Florence Hudson, Executive Director
Katie Naum, Operations Manager

475 Riverside Drive, Suite 320, New York, NY 10115 • nebigdatahub.org

Knology

Behaviors

Biosphere

Culture

Media

Wellness

Systems

Knology.org
info@knology.org

tel: (442) 222-8814
3630 Ocean Ranch Blvd.
Oceanside, CA 92056

tel: (347) 766-3399
40 Exchange Pl. Suite 1403
New York, NY 10005