Version 1.0
JUNE 2020

WEST BIG DATA INNOVATION HUB

2020-2025 Strategic Plan

a 'living document'

National Science Foundation Grants 1916573, 1916481, 1915774
UC Berkeley, UC San Diego, and University of Washington
Comments encouraged to: info@westbigdatahub.org

FIVE YEARS AGO…

The West Big Data Innovation Hub began with the community.

Through an open-to-all workshop in Salt Lake City, Utah, a researcher from Arizona stood with a data science educator from Washington, identifying opportunities to grow their programs across state boundaries. A small business owner from a rural town brainstormed with colleagues from city government, building bridges that would lead to data sharing, pilot projects, and a new career path.

In our first years, we have focused on data science activities, programs, and initiatives that inspire cross-sector collaboration and exemplify the need for multi-disciplinary approaches that 'translate' across boundaries.

We catalyze and scale a diverse set of joint efforts — whether developing data analysis and tools to support access to safe drinking water, to better understand disease through all 20,000 human proteins, or to facilitate new insights in transportation safety. In addition to these translational initiatives that bridge research, education, and practice, the West Hub aims to support cross-cutting activities that produce frameworks and resources useful to multiple areas of inquiry and practice, from data sharing and cloud computing to responsible data science.

Our coalition provides a creative and inclusive home — an affiliation that sparks meaningful connections and enables valuable work to positively impact science and society.

Pivotal Moments

We began drafting this 'living document' before the current COVID-19 global health crisis and the vast societal upheaval surrounding racial equity and justice in our communities. These pivotal moments are infused with layers of context, difficult questions, and an urgent need for open and collaborative efforts to move forward; these moments are also fundamentally entwined with the potential of the data science community.

This crucial juncture must be an opportunity for us to ***listen*** to each other, to ***acknowledge*** disparities, biases, and the devastating impact of systemic racism, to ***strengthen*** our compassion and resolve, and to ***commit*** to actions we can take towards a more just future. The goals we share here — raising awareness, catalyzing cross-sector collaboration, developing translational data science projects, and supporting education, training, and workforce development — require us to look within ourselves, learn, and work together.

As we continue to ask how we as individuals, teams, neighborhoods, and organizations can better serve as allies, contributing to long overdue systemic change — holding ourselves accountable in the journey ahead — we appreciate and welcome your ideas, insights, and continued partnership.

- The West Hub Leadership Team

VISION + MISSION.

The West Big Data Innovation Hub is a
COMMUNITY for CATALYZING and SCALING data science for societal needs.

Whether working towards the future of data-informed healthcare or tackling the challenges in disaster recovery decision-making, our diverse and growing team of stakeholders envisions a community empowered to contribute to a thriving regional, national, and global innovation ecosystem.

Our Mission is to BUILD and STRENGTHEN strategic partnerships across academia, industry, nonprofits, and government — connecting RESEARCH, EDUCATION, and PRACTICE to harness the data revolution.

GOALS

RAISE AWARENESS
of regional opportunities, inspiring new contributions and commitments

CATALYZE
team formation across academia, industry, nonprofits, and government

DEVELOP + ENABLE
translational data science (TDS) pilot projects in our thematic areas

SUPPORT
data science education, training, and workforce development

STRATEGIES

Creatively leverage and strengthen communication channels to engage stakeholders, expand our reach, and increase Hub participation

- *Ensure our leadership team's presence at synergistic local, regional, national, and global events increases the visibility of Hub activities, programs, and initiatives.*
- *Prioritize paths such as our new science writer position to create community-driven, consistent content that showcases Hub activities and achievements through persistent websites, newsletters, social media, and features in partner communication channels.*
- *Support open-to-all, online remote participation and recorded content that is easily discoverable.*

Reinforce and amplify opportunities for novel and productive community collaboration in our thematic areas

- *Partner with and build upon related initiatives to reach new participants and to maintain a diversity and variety of activities.*
- *Spark new commitments and contributions through timely Hub-led efforts (e.g. public-facing program launches or summits).*

Provide a welcoming environment for participants from all backgrounds to get involved in data science responding to societal needs

- *Facilitate and encourage the 'translation' and reframing of key questions, actions, and takeaways for multiple audiences.*
- *Provide partner-focused 'onramps' to engagement via activities and resources available in-person and online.*

Facilitate effective resource sharing across boundaries, with activities that connect our diverse community

- *Initiate partnerships with data stewards, resource creators, policy makers, and other collaborators to surface and enhance data or data-enabled contributions.*
- *Encourage a culture of responsible data sharing and highlight emerging best practices.*

Emphasize the value of interoperability, community-developed standards, and multidisciplinary teams

- *Capture inspiring real-world stories from our community, including paths for engagement, actionable recommendations, and ideas for potential future work.*
- *Continue to support activities such as the All Hub Infrastructure working group to promote open-access resources from partners across all sectors via monthly webinars and biannual in-person networking events.*
- *Provide a neutral venue for community dialogue and team formation through in-person convenings such as our All Hands Meeting, Challenge launches, and workshops, and online channels.*
- *Develop a community code of conduct with stakeholder input, and mechanisms for reporting issues.*

Support the cross-pollination and co-design of data-enabled approaches, methods, infrastructure, and tools that can lower barriers to community engagement and promote reuse / reproducibility

- *Ask 'who is not at the table?' and identify potential barriers to community engagement when designing and executing any activity.*
- *Map the 'lifecycle' of the outputs and outcomes of activities, programs, and initiatives to determine ways we can promote reuse.*

Support a 'virtuous cycle' of use-inspired learning and improvement

- *Encourage collaboration across stakeholder groups to build a cohesive set of use cases, objectives, and avenues for project development and shared, transferable learning.*
- *Facilitate a 'safe space' for testing assumptions and building new analysis, products/services, and partnerships addressing societal needs.*

Structure interactions that highlight real-world contexts, ethics, and the need for multidisciplinary problem-solving

- *Capture examples of how data science approaches, tools, and analysis have multi-faceted impacts, both positive and negative.*
- *Encourage holistic project design frameworks and transparent, community-driven measures of progress*
- *Design activities to help solidify a 'fluency' across project stakeholders and a culture of open innovation.*
- *Incorporate techniques for integrating different community voices into project design and evaluation.*

Identify and grow sustainable mechanisms and models for collaboration in the translational data science lifecycle

- *Periodically assess the sustainability of our activities, programs, and initiatives, including the potential for generating revenue.*
- *Compare different models of collaboration and engagement across our portfolio and with other organizations, including the other hubs in the national network.*

Meet people where they are

- *Organize open-to-all stakeholder meetings and design multiple onramps for initiatives, appreciating the wide spectrum of interests, commitment levels, and constraints.*
- *Collect and review participation statistics, to evolve our portfolio of community events and activities, ensuring representation from across the region, different sectors, and underrepresented groups.*

Broaden access and participation from all potential contributors, noting how diversity, equity, inclusion, and a sense of belonging are core to all Hub efforts

- *Continue to support early career participation and training-focused activities.*
- *Build upon our existing Steering Committee and recruit diverse candidates that reflect the broad range of Hub stakeholders and target audiences.*
- *Balance representation amongst event speakers and other highlighted Hub constituents.*

Identify, create, or modify activities, programs, and initiatives that promote open access educational content, tools, and data that are findable, accessible, interoperable, and reusable (FAIR)

- *Work regionally and with the national network of hubs and other partners to augment local efforts.*
- *Offer training, workshops, and reusable curricula that expand access or bridge gaps – for example, building upon momentum from Data Science for Social Good, Data Science Corps, FAIR data awareness and stewardship, and data access in the public sector.*

ADDITIONAL CROSS-CUTTING STRATEGIES

Align incentives through human-centered community discussion, inspirational data storytelling, and iterative co-design

- *Encourage active listening, empathy, and an environment where stakeholders can share their perspectives*

Strengthen our network of ambassadors

- *Enable scaling of our community through a culture of empowerment and growth, developing a dedicated network of formal and informal Hub ambassadors*

IMPACT MEASURES

Reach of activities, programs, or initiatives: stakeholder engagement metrics
Extent and depth of change attributable to Hubs efforts (e.g., new commitments and contributions surfacing due to Hub activities such as 'Calls for Commitments')
Diversity of our activities, programs and initiatives, noting target audience and actual audience, amount of effort and resources spent, number and variety of new partners or beneficiaries
Stakeholder perception and usage of generated content, resources, infrastructure, community-convening activities, and project acceleration frameworks/pathways
Integration of activities, programs, and initiatives (the extent to which the Hub's programming amplifies and cross-pollinates)

WEST BIG DATA INNOVATION HUB

West Hub Activities, Programs, and Initiatives connecting to our 2020-2025 GOALS and THEMATIC AREAS

GOALS (see additional table for related stratgies and objectives)

RAISE AWARENESS	CATALYZE	DEVELOP + ENABLE	SUPPORT
of regional opportunities, inspiring new contributions and commitments	team formation across academia, industry, nonprofits, and government	translational data science (TDS) pilot projects in our thematic areas	data science education, training, and workforce development

INITIAL THEMATIC AREAS

Metro Data Science	Natural Resources + Hazards	Health + Medicine	Data-Enabled Discovery + Learning

Cross Cutting Themes include Infrastructure and Cloud Computing, Data Storytelling, and Responsible Data Science

2020 2021 2022 2023 2024 2025

All Hands Meetings (since 2016)
Signature open-to-all convening
RAISE AWARENESS | CATALYZE TEAMS

Transportation Data + Driver Video Privacy Challenges (since 2017)
Community engagement and translational research, responsible data science
DEVELOP PROJECTS | CATALYZE TEAMS | RAISE AWARENESS | EDUCATE/TRAIN

Data Carpentries (since 2017; every-other-year)
Train-the-Trainer program, tribal + MSI community workshops
EDUCATE/TRAIN | CATALYZE TEAMS

Water Data Collaborations (since 2018)
Series of open data projects, content, and partnerships
State/local government, industry, foundation financial support
DEVELOP PROJECTS | RAISE AWARENESS | CATALYZE TEAMS | EDUCATE/TRAIN

ABOUT THIS RESOURCE

This framework for viewing how West Hub Activities, Programs, and Initiatives (West Hub 'API') connect to our goals and thematic areas is being developed as an interactive, online resource for community members to search, filter, and provide project information over time. We welcome all interested individuals and groups contributing to data science for societal needs to collaborate and help design and shape this resource to be useful. In particular, we aim to integrate user-friendly ways for activities, programs, and initiatives to acknowledge and credit collaborators, indicate open questions/needs, and highlight ways to participate over time.

In the prototyping stage, the West Hub leadership team is leveraging the framework and additional mechanisms to gain insights about potential gaps and opportunities, aiming to support a balanced and diverse portfolio as our community engagement and capacity both evolve.

For example, not visually elaborated above, but to be included in the interactive online resource are: more detail about Infrastructure and Cloud Computing initiatives (all goals), Data Storytelling (raise awareness, catalyze team formation, educate/train), Responsible Data Science (all goals, with a focus on ethics, privacy, and security), and programs such as the National Data Science Education Workshops (educate/train, raise awareness), FAIR data workshops (educate/train), Exposome Pilot, and Community Calls-for-Collaborative-Projects (all goals), as well as partnerships with U.S. Census Bureau (raise awareness, educate/train), COVID Info Commons (all goals), Women in Data Science (educate/train), and Data Science for Social Good (educate/train, develop projects).

WEST BIG DATA INNOVATION HUB

GOALS + STRATEGIC OBJECTIVES

The West Hub's vision and mission drive a set of four overarching goals depicted on Page 4. Collectively, we aim to:

- RAISE AWARENESS of regional opportunities, inspiring new contributions and commitments;
- CATALYZE team formation across academia, industry, nonprofits, and government;
- DEVELOP and ENABLE translational data science pilot projects in our thematic areas; and
- SUPPORT data science education, training, and workforce development.

As described in further detail on Pages 15-18, we note a set of 3 strategic objectives for each goal, and initial measures of impact corresponding to each goal. Additional cross-cutting strategies shared on Page 4 emphasize the importance of enabling a strong network of ambassadors and aligning incentives through human-centered community discussion, inspirational data storytelling, and iterative co-design.

In addition to West Hub goals and strategic objectives, we anticipate synergy with national goals and strategic objectives set forth through the development of a national coordination effort across all four regional innovation hubs.

The objective of the National Coordination Committee is to support a diverse portfolio of strategically driven activities at the national scale, as well as compliment regional activities with national resources to create accomplishments that would not have occurred without our collective contributions. The National Coordination Committee role may include:

- Understanding each Hub's priority areas / activities to strategically suggest new community engagement and avenues for scaling (e.g., involvement of constituents, groups, or initiatives from each of the Hubs),
- Helping the Big Data Hubs respond to national-scale NSF opportunities, and
- Proposing strategic non-NSF national-scale opportunities.

PORTFOLIO MANAGEMENT + DECISION MAKING

With a seemingly boundless and dynamic ecosystem of opportunities that could potentially connect to our regional and national goals, this Strategic Plan serves to support the effective management of the West Hub's portfolio of activities, programs, and initiatives – guiding impact-focused reflection and informing Hub decision making. Regular review of the Hub portfolio on a semi-annual cadence, noting how, when, and why any activities are augmented or reduced is critical to allow for a healthy "lifecycle" of activities and an appropriately nimble process for resource allocation, as allowable by the Hub's Cooperative Agreement or other funding guidance, and as aligned with existing scopes of work.

Particularly for decision points with significant impacts on limited resources, the Hub leadership and senior personnel will employ both our draft measures of success as well as a *Spectrum of Agreement* decision making process to facilitate substantive discussion that supports collaborative and informed action. The *Spectrum of Agreement* framework we have "piloted" for internal decision making starting in December 2019 provides a scale as follows:

1. I do not support this action/decision or recommendation
2. I need revisions to support this action/decision or recommendation.
3. I can support this with additional information.
4. I am mostly satisfied with this and will support it.
5. I strongly support this action/decision or recommendation.

As a guiding principle, we will aim to converge on actions, decisions, and recommendations where all parties indicate a "3, 4, or 5" level of support, as inspired by the facilitated decision making process used for a multi-stakeholder State of California Technology Working Group. Notably, while this process is being used to support streamlined organizational decision making for the Hub with our Leadership Team and senior personnel, the framework can help at a broader scale for small groups and other cases for the Hub community.

SELECTION PROCESSES

In addition to the strategic growth and "pruning" of the Hub portfolio of activities, we recognize that another form of decision making occurs through competitive selection processes where community demand is greater than available capacity. Examples include applications for Participant Support or selected roles, seed funding for collaborative projects, and merit-based recognition.

Intellectual merit and the potential for impact guide these selection processes, with efforts to broadly recruit participation supporting diversity across sectors, technical focus, institutions/affiliations, roles and responsibilities, demographics as allowable by law, and the level of prior engagement with the topic and/or the project entities.

Our next version of the Strategic Plan will include samples of the selection processes currently employed, to encourage discussion of best practices. For example, at an overview level:

- The review of Water Data Challenge entries and allocation of awards (with funding from external sponsors) has consisted of a categorized rubric process with several volunteer judges from the community selected by the initiative's multi-stakeholder team. Beyond the standard judging process, a "People's Choice Award" is often provided by open voting from all in-person participants at an event.
- The review for hiring a seed-funded West Hub Science Writer that will support regional seed-funded and core-award-funded projects consisted of a rubric process that included a Steering Committee member, a Leadership Team member, and a community stakeholder outside of the Hub senior personnel.
- The review for our Data Carpentry Train-the-Trainer cohort included "matchmaking" across teams from similar geographic areas to support a diversity of locations and sectors. For example, we notified the applicants from a given state to encourage collaboration and support a "hybrid" team where each institution could participate, albeit with less people per institution, given limited seating.
- In our first phase, participant support applications were encouraged from early career community members, with support to anyone in our 13 states providing written details of the anticipated impact; as demand increases for support, we will formalize the resource allocation process with a rubric and multi-stakeholder review.

Finally, we note that despite finite resources and logistical constraints, the West Hub aims to intentionally design avenues for varying levels of engagement, to broaden and scale the impact of a set of opportunities whenever possible. Consequently, the planning process for any activity considers ways to sustainably increase access to content for those unable to join any limited-capacity opportunity.

LINKAGE OF STRUCTURE TO GOALS

The 'West Hub Activities, Programs, and Initiatives' resource shown on Page 5 serves to methodically link West Hub activities, programs, and initiatives to the objectives and goals they address – to spark discussion and strengthen a culture of continuous improvement.

Throughout the coming year, the Hub leadership will be working with an evaluation team to more comprehensively inventory and assess our activities, programs, and initiatives, noting how our current and future efforts have a range of governance and operational structures that influence progress towards objectives and goals.

For example, the organizational governance and operational structure noted in the Management + Staffing section – in particular, the in-kind personnel through partnerships – sets a foundation for deeper reflection about our Hub collaborative framework and the mechanisms, norms, and measurable outputs contributing to our goals.

Notably, in the context of example vignettes included below, the links between structures, goals, and objectives of the West Hub and our collaborators emerge:

- **The California Water Data Challenge**, an effort now in its third year of being led by the West Hub with the California Governor's Office of Planning and Research, aims to facilitate team-based data projects and tool development. The Challenge attracts both new and returning participants and mentors, having been cited as an "extraordinary" piece of community engagement critical to innovation, feedback, and implementation of the first-in-the-nation 2016 legislation for more accessible water and ecological data (AB1755).

  As collaborators from the Water Foundation have noted, the Challenge results helped "make the case" in 2019 for a historic statewide Safe and Affordable Drinking Water Fund of **$1.3 Billion** over the next decade. The synergy between the Hub efforts and the wide-reaching water data collaborations across more than eight state agencies can be seen in jointly-generated video content (supported by State staff) as well as the formation of an entirely new, State-endorsed nonprofit 501(c)(3) organization, the California Water Data Consortium. With the Consortium hiring its first President and CEO beginning in Summer 2020, the Hub anticipates additional opportunities to link evolving community structures towards shared goals.

- **The National Transportation Data Challenge** (NTDC) and the **Driver Video Privacy Challenge** represent a progression of multi-sector collaborations leveraging Federally Funded Research and Development Centers, industry experts, and joint efforts through a Federal Interagency Agreement. In contrast to the water data efforts mentioned above, the structure of the transportation challenges began with national scaffolding, drawing inspiration from the global Vision Zero movement and the Transportation Research Board's flagship Second Strategic Highway Research Program (SHRP2) Naturalistic Driving Study (NDS) across the states of Washington, Florida, Indiana, New York, North Carolina, and Pennsylvania. With more than 35 million vehicle miles and unprecedented levels of in-cabin video data at stake for potential research insights, the West Hub led the launch of these Challenges in close collaboration with the network of Hubs across the nation, cross-sector partners, NSF, and the U.S. Department of Transportation. The 6-month NTDC series of community problem-solving

sessions, roundtables, and technology demonstrations produced new data sharing commitments and data analysis focused on transportation safety. The West Hub is building upon the NTDC results in subsequent partnership with the Northeast Hub and an advisory team, framing avenues for de-identifying facial video data while addressing issues related to privacy, ethics, bias, and security. The structure of the activities leverages the backdrop of an estimated $500 Billion annual economic cost of U.S. traffic crashes, 2 petabytes of SHRP2 data (**1.2 million hours of video**), seed funding, and the authority of 15 U.S.C. 3719 to stimulate proof-of-concept research and surface privacy-preserving video analytics techniques. Creative structuring of 'Category Chairs' and topical ambassadors combined with student leaders help support a broad sense of belonging and a growing set of community contributions.

- **West Hub Supported Data Science Education, Training, and Workforce Development** activities, particularly those structured to engage a diverse cohort of leaders and potential Hub ambassadors, advance common goals both regionally and more broadly. The cadence, structure, and models of community engagement for activities such as Data Carpentry Train-the-Trainer workshops, FAIR (Findable, Accessible, Interoperable, and Reusable) data training, the WiDS (Women in Data Science) Datathon, and the National Data Science Education Workshop series, all aim to leverage economies of scale with key partners, and to spark new connections across our constituents – to meaningfully broaden access and participation from traditionally underserved stakeholders. Emphasis on early career support, programmatic content that is open-to-all, and resources that are generated by the community and for the community aim to "meet people where they are" and support multiple onramps to our Hub portfolio.

Moreover, the four Big Data Innovation Hubs have established strong relationships with each other, laying the foundation for more structured national-scale coordination, including the formation of a National Coordination Committee (NCC). Collectively, the Hubs have committed to coordination in four key areas:

- Communications: Developing and executing a national-scale communications strategy to consistently engage and grow our communities, incorporating stakeholder research, development of core messaging, social media strategy, branding, marketing assets, shared bigdatahubs.org website scoping, and public relations.
- Education and Workforce: Recognizing our mission to support data science education and workforce training, we plan to continue, coordinate, and scale education and workforce activities, including programs and events to boost data science capacity, support data literacy, and emphasize responsible data science, with a focus on underrepresented /underserved populations and in collaboration with other sectors.
- Cyberinfrastructure and Data Access: Coordinating and scaling current national cyberinfrastructure and data access activities, such as the open storage network and the All Hub Infrastructure working group.
- Evaluation and Metrics: We plan to work toward a coordinated evaluation strategy and potentially shared evaluation activities, methods, and metrics, comprising both qualitative success stories and quantitative measures.

MANAGEMENT + STAFFING

Organizational Structure. The West Hub is led in collaboration across three host institutions: the University of California, Berkeley; the University of Washington; and the University of California, San Diego. The structure of three PIs, an Executive Director/CoPI, two Deputy Directors/CoPIs and any additional CoPIs enables coordinated on-campus presence across the host institutions. As we expand upon our activities and the community-building inherent to all lines of effort, we will leverage additional Hub staffing through part-time fellows and project leads.

In addition to our Leadership Team of PIs, CoPIs, and Directors, the West Hub includes a volunteer Steering Committee with leaders from across our community. Our inaugural West Hub Steering Committee, which was the result of an open nominations and election process, includes representation from local government, industry, academia, and nonprofits. The Steering Committee has committed to volunteer their service extending into the second phase of the Hub, with consensus that the committee would ideally stagger terms in order to preserve continuity of knowledge, while welcoming new engagement. The structured rotation of Steering Committee membership also serves to achieve representation from different disciplines, application areas, geographies, and backgrounds, aligned with our intentions for facilitating a Hub environment that is welcoming to all, and appreciative of a diverse set of community contributions.

We will continue to encourage collaboration between our rotating Steering Committee members and our Leadership Team of PIs, Co-PIs, and Directors. In the first phase of the Hub, the PIs, Co-PIs, and Directors served as *ex officio* members of the Steering Committee. Given the increased responsibilities and opportunities for engagement of the Steering Committee, in particular through input on community *Calls-for-Collaborative-Projects* throughout the year, we will work to pair small teams of committee members for specific tasks, cognizant that the Steering Committee role is voluntary and not remunerated. This strategy builds upon our existing Steering Committee-led formation of Subcommittees focused on specific sustainability, communications, and internal coordination opportunities.

Specifically**,** the community-focused seed funding through *Calls-for-Collaborative-Projects* will be managed through a coordinated effort across the three lead institutions, with oversight and engagement from a subset of the Steering Committee as well as stakeholders external to the Hub formal committee structure. For example, the recent competition for a West Hub Science Writer employed a review process that included input beyond the Leadership Team and Steering Committee, with policies for avoiding Conflict of Interest and leveraging a rubric similar to project-specific scoring criteria developed in alignment with Hub strategic goals and objectives. This intentional structuring and integration of different perspectives will aid the Hub in effectively advancing a diverse group of topics, and opportunities, bringing in valuable new community engagement.

The external evaluation team is from one organization, was approved as a subrecipient as of December 2019 by NSF, and is currently in contract with UC Berkeley as a vendor.

West Hub Personnel

June 2020

University of Washington (Collaborative PI Institution)

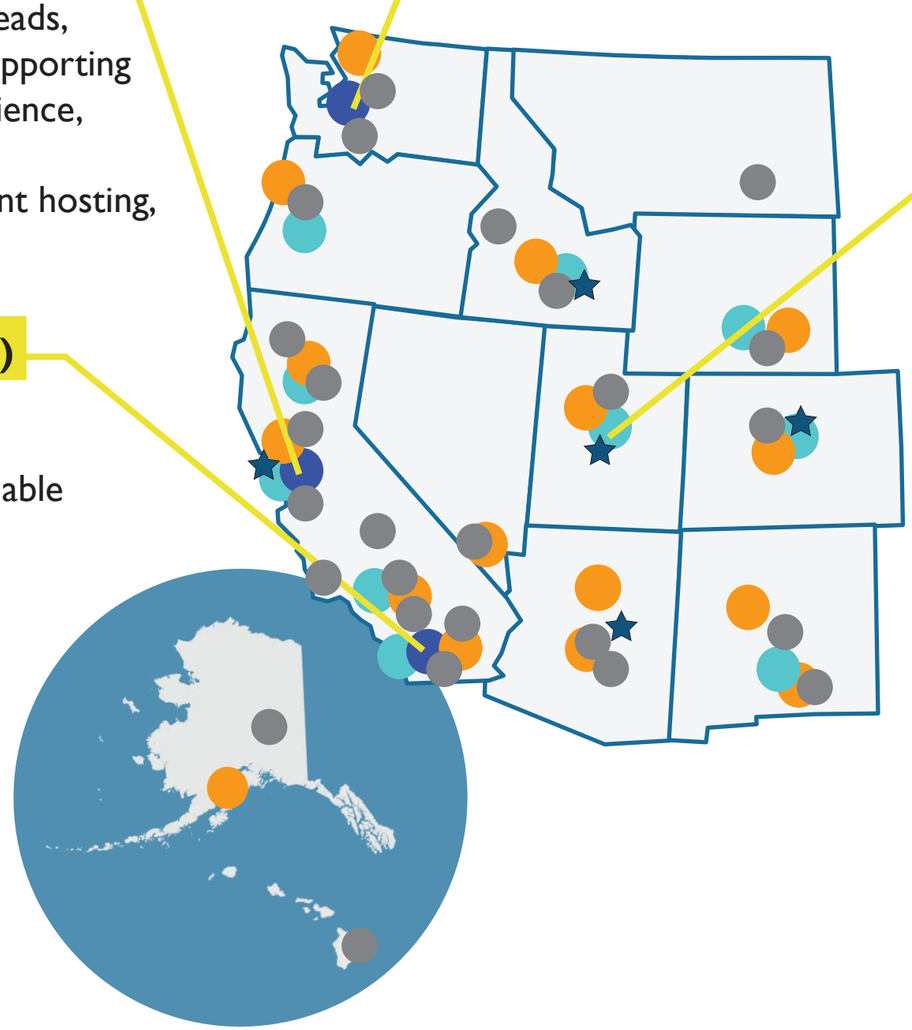
- Deputy Director/CoPI + Principal Investigator + CoPI
- Project Senior Personnel and as available and aligned, personnel supporting key initiatives such as Data Carpentries and DSSG workshops, cloud computing

UC Berkeley (Lead PI Institution)

- Executive Director/CoPI + Principal Investigator
- Part-time Administrative Assistant
- As available and aligned, Senior Fellow project leads, students, and in-kind collaborative personnel supporting key intitiatives such as natural resources + resilience, Census, transportation, and education
- Contracted vendors for graphic facilitation, event hosting, and the independent evaluation team

UC San Diego (Collaborative PI Institution)

- Deputy Director/CoPI + Principal Investigator
- Part-time Science Writer
- Additional Project Senior Personnel and as available and aligned, personnel supporting key initiatives such as cyberinfrastructure, FAIR and Data Carpentries programs, cloud computing



- **Steering Committee (community-elected)**
- **Local personnel for co-hosted activity**
- **Participant Support**
- ★ **All Hands Meeting Program Leads beyond Hub leadership team (close collaboration with local host committee)**

National Coordination Committee community representative

Additional personnel or community members that can be geographically visualized include: trainees, participants not receiving financial support, speakers, facilitators, recipients of grants beyond participant support, recipients of letters of collaboration, applicants, and people subscribed to Hub communications channels, and survey respondents.

WEST BIG DATA INNOVATION HUB

Risk Mitigation. In the event of the loss of key personnel to the project, such as Director(s), Principal Investigators, or senior staff, the relevant project management roles, responsibilities, budget, and supporting collateral materials as appropriate will be transitioned.

As part of our long-term management/staffing sustainability and risk mitigation efforts, we continue to maintain key resources that have been developed thus far in a transparent, open, and online workspace whenever possible. These key resources include but are not limited to works-in-progress such as public-facing online documentation, curated partnerships and 'intellectual infrastructure' (governance and commitments facilitated through the hub), and frameworks/guides that enable community engagement across academia, government, nonprofits, and industry constituents.

Ongoing activities will emphasize continuity of existing operations, capturing hub-facilitated projects, initiatives, shared resources, and progress for broader distribution through existing online platforms (including but not limited to westbigdatahub.org and bigdatahubs.org websites, shared working group online spaces, social media, and open access documentation).

Moreover, the Leadership Team of PIs, CoPIs, and Directors continues to take initiative to meet in person when possible as budget allows (e.g., to hold strategic planning discussions when a critical mass is co-located); these sessions allow the leadership to reflect on the Hub's management, trajectory, and upcoming opportunities.

Finally, as part of the management and staffing plan for the hub, we have planned to support national-scale coordination. The National Coordination Committee will be a small, nimble body that allows autonomy of the regional Hubs, while facilitating grassroots coordination among the four regions. Each Hub's Executive Director will serve on the NCC, plus at least one additional member from each region, with each Hub having an equal number of representatives. One member of each Hub's Leadership Team will serve as the NCC's point of contact with the National Science Foundation as assigned on a rotating basis of one year among all of the Hubs.

MEASURES OF IMPACT

With an understanding of the often multi-faceted, interconnected, and dynamic nature of the West Hub's portfolio since its launch, the evaluation effort will provide structured and methodical analysis that stimulates reflection and will inform the Hub in the strategic planning and execution of activities, projects, and initiatives.

The impact-focused questions, measures, and scenarios developed are anticipated to incorporate context from programmatic activities designed to accelerate stakeholder engagement in data science, activities focused on developing socio-technical resources/services, and data science education, training, and workforce development activities.

We anticipate the listed draft impact measures in this section will be designed and discussed with the evaluation team, using these working definitions:

- *Scale*: the extent and depth of change attributable to the activities, projects, and initiatives.
- *Reach/Engagement*: the number and variety of communities affected by Hub activities and the number and variety of partnerships with which the Hub works.
- *Integration or Interconnectedness*: the extent to which Hub partnerships amplify and cross-pollinate Hub activities to create economies of scale.
- *Impact*: the perceived added benefit for stakeholders related to integration with the Hubs.
- *Representativeness*: the extent to which the activities align with the goals of the project and its stakeholders.
- *Responsiveness*: the extent to which the activities address the needs and values of community stakeholders at appropriate or reasonable timescales.

Working collaboratively with the West Hub Leadership Team and active personnel such as our Steering Committee members, the independent evaluation team, Knology, will first identify key measures for documenting *Scale, Engagement, Integration,* and *Impact* through mapping current projects, activities, and initiatives. Second, leveraging existing community surveys and input, the evaluation team will establish best practices for collecting data to address evaluation objectives, including a focus upon *Representativeness* and *Responsiveness* of activities.

For our 4 broad West Hub Goals, draft measures of impact are noted below each of the strategies over the next pages:

GOAL 1: RAISE AWARENESS of regional opportunities + INSPIRE new contributions and commitments

STRATEGIES:

(1.1) Creatively leverage and strengthen communication channels to engage stakeholders, expand our reach, and increase Hub participation

- Ensure our Leadership Team's presence at synergistic local, regional, national, and global events increases the visibility of Hub activities, programs, and initiatives.
- Prioritize paths such as our new science writer position to create community-driven, consistent content that showcases Hub activities and achievements through persistent websites, newsletters, social media, and features in partner communication channels.
- Support open-to-all, online remote participation and recorded content that is easily discoverable.

(1.2) Reinforce and amplify opportunities for novel and productive community collaboration in our thematic areas

• Partner with and build upon related initiatives to reach new participants and to maintain a diversity and variety of activities.

• Spark new commitments and contributions through timely Hub-led efforts (e.g. public-facing program launches or summits).

(1.3) Provide a welcoming environment for participants from all backgrounds to get involved in data science responding to societal needs

• Facilitate and encourage the 'translation' and reframing of key questions, actions, and takeaways for multiple audiences.

• Provide partner-focused 'onramps' to engagement via activities and resources available in-person and online.

MEASURES may include:

- Reach of content (stakeholder engagement metrics)
- Extent and depth of change attributable to Hubs communications/content — for example, the number and extent of new commitments and contributions surfacing due to Hub actions such as "Calls for Commitments"
- Diversity of our activities, programs and initiatives, noting
  - target audience and actual audience,
  - amount of effort and resources spent,
  - number and variety of new partners or beneficiaries
- Stakeholder perception and usage of content and communication channels/environments
- Integration of activities, programs, and initiatives (the extent to which the Hub's programming amplifies and cross-pollinates)

GOAL 2: CATALYZE team formation across academia, industry, nonprofits, and government.

STRATEGIES:

(2.1) Facilitate effective resource sharing across boundaries, with activities that connect our diverse community

• Initiate partnerships with data stewards, resource creators, policy makers, and other collaborators to surface and enhance data or data-enabled contributions.
• Encourage a culture of responsible data sharing and highlight emerging best practices.

(2.2) Emphasize the value of interoperability, community- developed standards, and multidisciplinary teams

• Capture inspiring real-world stories from our community, including paths for engagement, actionable recommendations, and ideas for potential future work.
• Continue to support activities such as the All Hub Infrastructure working group to promote open-access resources from partners across all sectors via monthly webinars and biannual in-person networking events.
• Provide a neutral venue for community dialogue and team formation through in-person convenings such as our All Hands Meeting, Challenge launches, and workshops, and online channels.
• Develop a community code of conduct with stakeholder input, and mechanisms for reporting issues.

(2.3) Support the cross-pollination and co-design of data-enabled approaches, methods, infrastructure, and tools that can lower barriers to community engagement and promote reuse / reproducibility

• Ask 'who is not at the table?' and identify potential barriers to community engagement when designing and executing any activity.
• Map the 'lifecycle' of the outputs and outcomes of activities, programs, and initiatives to determine ways we can promote reuse.

MEASURES may include:

- Reach of team-formation efforts (stakeholder engagement metrics)
- Extent and depth of change attributable to Hubs team-formation and resource-sharing efforts — for example, newly identified team members / cohorts and/or the extension of prior investments
- Diversity of our activities, programs and initiatives, noting
  - target audience and actual audience,
  - amount of effort and resources spent,
  - number and variety of new partners or beneficiaries
- Stakeholder perception and usage of resources, infrastructure, and community-convening activities
- Integration of activities, programs, and initiatives (the extent to which the Hub's programming amplifies and cross-pollinates)

GOAL 3: DEVELOP + ENABLE translational data science (TDS) pilot projects in our thematic areas.

STRATEGIES:

(3.1) Support a 'virtuous cycle' of use-inspired learning and improvement

• Encourage collaboration across stakeholder groups to build a cohesive set of use cases, objectives, and avenues for project development and shared, transferable learning.
• Facilitate a 'safe space' for testing assumptions and building new analysis, products/services, and partnerships addressing societal needs.

(3.2) Structure interactions that highlight real-world contexts, ethics, and the need for multidisciplinary problem-solving

• Capture examples of how data science approaches, tools, and analysis have multi-faceted impacts, both positive and negative.
• Encourage holistic project design frameworks and transparent, community-driven measures of progress
• Design activities to help solidify a 'fluency' across project stakeholders and a culture of open innovation.
• Incorporate techniques for integrating different community voices into project design and evaluation.

(3.3) Identify and grow sustainable mechanisms and models for collaboration in the translational data science lifecycle

• Periodically assess the sustainability of our activities, programs, and initiatives, including the potential for generating revenue.
• Compare different models of collaboration and engagement across our portfolio and with other organizations, including the other hubs in the national network.

MEASURES may include:

- Reach of TDS pilot project efforts (stakeholder engagement metrics)
- Extent and depth of change attributable to Hubs TDS pilot project efforts
- Diversity of our TDS pilot project activities, programs and initiatives, noting
  - target audience and actual audience,
  - amount of effort and resources spent,
  - number and variety of new partners or beneficiaries
- Stakeholder perception and usage of resources, infrastructure, community-convening, and project acceleration frameworks/pathways
- Integration of activities, programs, and initiatives (the extent to which the Hub's programming amplifies and cross-pollinates)

GOAL 4: SUPPORT data science education, training, and workforce development.

STRATEGIES:

(4.1) Meet people where they are

• Organize open-to-all stakeholder meetings and design multiple onramps for initiatives, appreciating the wide spectrum of interests, commitment levels, and constraints.
• Collect and review participation statistics, to evolve our portfolio of community events and activities, ensuring representation from across the region, different sectors, and underrepresented groups.

(4.2) Broaden access and participation from all potential contributors, noting how diversity, equity, inclusion, and a sense of belonging are core to all Hub efforts

• Continue to support early career participation and training-focused activities.
• Build upon our existing Steering Committee and recruit diverse candidates that reflect the broad range of Hub stakeholders and target audiences.
• Balance representation amongst event speakers and other highlighted Hub constituents.

(4.3) Identify, create, or modify activities, programs, and initiatives that promote open access educational content, tools, and data that are findable, accessible, interoperable, and reusable (FAIR)

• Work regionally and with the national network of hubs and other partners to augment local efforts.
• Offer training, workshops, and reusable curricula that expand access or bridge gaps — for example, building upon momentum from Data Science for Social Good, Data Science Corps, FAIR data awareness and stewardship, and data access in the public sector.

MEASURES may include:

- Reach of education, training, and workforce development efforts (stakeholder engagement metrics)
- Extent and depth of change attributable to Hubs education, training, and workforce development efforts
- Diversity of education/workforce activities, programs and initiatives, noting
  - target audience and actual audience,
  - amount of effort and resources spent,
  - number and variety of new partners or beneficiaries
- Stakeholder perception and usage of education, training, and workforce development resources, infrastructure, community-convening, and individual or group learning frameworks/pathways
- Integration of activities, programs, and initiatives (the extent to which the Hub's programming amplifies and cross-pollinates)

Notably, in the process of being responsive to community-driven needs and partnerships, we have evolved the West Hub's coordination mechanisms to serve more long-term goals, with less 'one-time' activities and more strategically shaped interaction that consists of a series of connected activities. This type of community engagement offers a more nuanced opportunity for time-dependent evaluation and will play a key role moving forward. Evaluator-led interviews will provide one avenue for reflecting on the progression of roles for community members (e.g., viewer, active attendee, team lead/facilitator) and seek to assess how volunteers gain short and long-term value from their Hub interaction and participation.

Reflecting upon additional mechanisms for gathering both quantitative and qualitative — and both formative and summative — metrics, we continue to capture and document ideas in our existing shared online collaboration space. Moreover, in addition to our regional measures, we will work with the Leadership Teams across the hub network to identify metrics shared across all Hubs to monitor national-scale impacts and strengthen national coordination.