Network for Earth-space Research Education and Innovation with Data (NEREID)

NSF Workshop Report | November 20-22, 2019

Editors:

Catherine B. Cramer | Stephen Miles Uzzo | Daniel Fuka | Katie Naum | Tim Spuck | Peter Tuddenham

Workshop Webpage: https://sites.google.com/view/nereid2019/home

Cover image: word cloud generated from frequency of terms expressed by workshop attendees.

Table of Contents

Authors

We owe our deepest gratitude to all participants for their valuable input throughout the workshop, and presenting their most innovative research and development projects and demonstrating systems under development to identify their most promising ideas for convergence. All workshop participants were active authors for this report:

- Richard Boone, Humboldt State University
- Catherine B. Cramer, Associated Universities Incorporated (AUI)
- Yasmin Catricheo, Associated Universities Incorporated (AUI)
- Robert R. Downs, CIESIN, Columbia University
- Daniel Fuka, Virginia Tech-HI Multi-Climate EcoHydrology Laboratory
- Jose Guridi, Chile Ministry of Science, Technology, Knowledge and Innovation
- Jim Hammerman, TERC
- Sue Ann Heatherly, Green Bank Observatory
- P. Bryan Heidorn, University of Arizona
- Elizabeth Joyner, NASA Langley Research Center
- Randy Kochevar, Education Development Center, Oceans of Data Institute
- Ouida Meier, EarthCube Office
- Margaret Mooney, NOAA Cooperative Institute for Meteorological Satellite Studies
- Katie Naum, Northeast Big Data Innovation Hub, Columbia University
- Shelley E. Olds, University NAVSTAR Consortium
- Leigh Peake, Gulf of Maine Research Institute (GMRI)
- Claire Raftery, National Solar Observatory
- Rebecca Reid, Earth Science Information Partners (ESIP)
- Kartik Sheth, NASA
- Tim Spuck, Associated Universities Incorporated (AUI)
- Gretchen Stahlman, Rutgers University, School of Communication & Information
- Ed Summers, Statistical Analysis System Institute (SAS)
- Laura Trouille, Adler Planetarium
- Peter Tuddenham, College of Exploration
- Stephen Miles Uzzo, New York Hall of Science/AUI
- Tiffany Stone Wolbrecht, Ward Beecher Planetarium, Youngstown State University
- Elena Yulaeva, University of California San Diego, Scripps Institution of Oceanography
- Ilya Zaslavsky, University of California San Diego Supercomputer Center

Acknowledgements

The organizers would like to thank Chaitanya K. Baru, Senior Science Advisor, Office of Integrative Activities at the National Science Foundation for inspiring and supporting this workshop and report. We would also like to thank the following individuals who were unable to attend the workshop but who expressed interest and/or contributed expertise during the organization of the workshop and the compilation of this report:

- Rene Baston, Northeast Big Data Innovation Hub, Columbia University
- Jim Bellingham, Woods Hole Oceanographic Institute
- Virginia Burkett, USGS
- Robert Chen, CIESIN, Columbia University
- Jackie Faherty, American Museum of Natural History
- Vicki Ferrini, Lamont-Doherty Earth Observatory, Columbia University
- Chris German, Woods Hole Oceanographic Institute
- Simon Goring, University of Wisconsin Madison
- Denise Hills, Geological Survey of Alabama
- Josef Kellndorfer, Earth BigData
- Joseph Kerski, ESRI
- Mordecai-Mark Mac Low, American Museum of Natural History
- Jim Madsen, University of Wisconsin River Falls
- Thomas Narock, Goucher College
- Erin Robinson, Earth Science Information Partners (ESIP)
- Patrick Rowley, Charles County (MD) Public Schools
- Mark Serreze, National Snow and Ice Data Center
- Sarah Schoedinger, NOAA
- Jason Williams, Cold Spring Harbor Labs

This effort is supported by the National Science Foundation under Award Number 1939471 to Associated Universities Incorporated. 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.

Executive Summary

This report summarizes the results of a National Science Foundation (NSF) workshop entitled: Network for Earth-space Research Education and Innovation with Data (NEREID) that was inspired by NSF's Convergence Accelerator initiative. The workshop took place at the Green Bank Observatory, Green Bank, West Virginia on November 20 - 22, 2019.

The workshop brought together 28 leading experts in Earth and space data science and education who came from across the hemisphere. Together the participants identified a set of challenges faced by geoscience and space science researchers, practitioners and educators. Those challenges that might be solvable in the near term through convergence research and education in Earth and space science data, as well as those solvable during the next 10 years through transformational approaches to research and education were identified and usage scenarios were detailed and described. Major findings of the workshop are presented here. Specifically, the report intends to:

- Portray the current state of geosciences and space sciences data practice and education as revealed during the workshop;
- Present major challenges and opportunities for convergence in the design of tools and techniques to support scholarly discovery and democratization of data science tools for teachers and lifelong learners, including a timeline of anticipated science and technology innovation that will impact the development of future tools and techniques; and
- Provide recommendations on how both current lines of work and promising future avenues of research and development in academia, government, industry and policy can be exploited to design more effective data science approaches through the convergence of Earth and space data.

Over the course of the workshop, participants made the following observations, relevant to this initiative:

- Data science can provide a transdisciplinary lens on Earth and space sciences that will help identify and actualize areas of convergence;
- General systems theory provides tools and methods to help make connections among Earth and space sciences systems toward convergence of research approaches, while accommodating a diversity of ideas, equity across communities of stakeholders, and inclusion of a broad gamut of perspectives on Earth and space sciences;
- The current Earth and space sciences data resources, while primarily siloed, are well-positioned for accelerated convergence by bringing communities of practice together through team science, policy and education;
- There is growing interest in the kind of interdisciplinarity that the potential convergence of Earth and space data represents;
- Using the socio-technical framing of attitudes, tools and techniques, and social interactions helps outline needs and processes that will lead to a successful convergence effort; and
- In the far horizon, the use of quantum engineering may offer solutions to some of the seemingly insurmountable technical barriers to Earth and space data sensing, imaging and analysis that we face today.

Principal Findings

There are several general, confirmative findings:

- The time is right for the convergence of Earth-space data through bringing together collaborative teams across the Earth, space/astronomy, and data science communities;
- K-14 and community education programs are valuable resources for testing the acceleration of Earth-space sciences convergence, as the data used in each individual program often spans multiple disparate scientific domains;
- Data science can be a catalyst for convergence in our understanding of Earth and space sciences, but communities of users need significant support to use existing tools and develop new ones;
- Resources commonly available to researchers and educators in the space sciences (e.g., planetarium capabilities for modeling) offer a potentially wide-reaching set of resources for the Earth sciences;
- Mature data architectures from the Earth and atmospheric sciences open up significant open-source data sharing capabilities across the space science educational infrastructures;

And a number of findings specific to the workshop topic:

- Diversity, equity, and inclusion are essential characteristics of a convergence accelerator;
- Educating both the current and next generation to embrace complexity and use a variety of data tools, including sophisticated AI-driven data tools, will be essential for accelerating convergence of Earth-space science; and
- Some convergence of Earth and space data is already happening and demonstrates the value of accelerating and scaling innovation and discovery in the short term.

Principal Recommendations

A set of recommendations resulted for addressing the needs of the space and geosciences data research community:

- Catalog existing Earth-space data use cases and bring together larger communities to cross-pollinate and exchange ideas, needs, and aspirations;
- Develop essential support structures for the integration of data sciences into space and geosciences as a crucial first step to prime the research and education communities to align and converge. This includes providing in-depth education in the use of tools and an immediate effort to democratize sophisticated tools for data analysis, cleaning, and visualization and make them widely available across the Earth and space sciences; and
- Bring together teams equally weighted with space- and geo-data-scientists to develop open data specifications and ontologies that mediate information and data exchange across the Earth and space sciences.

1. Workshop Organization and Reporting Structure

This report summarizes the results of a National Science Foundation (NSF) supported gathering entitled "Network for Earth-space Research Education and Innovation with Data (NEREID)", which was a visioning and capacity-building workshop inspired by NSF's Convergence Accelerator (C-Accel) initiative. NEREID is working with academia, industry, education, and research communities to identify efforts toward conducting research, creating tools, and engaging learners and policymakers that can benefit from a convergence effort in Earth and space sciences. Over the span of three days the workshop brought stakeholders together to:

- Explore the challenges of the gathering and use of big data in Earth and space sciences across research communities and within learning settings through a collaborative inquiry process that elicited diverse points of view and enfranchised underrepresented groups;
- Identify and build a convergent, interdisciplinary community of practice;
- Begin to develop and disseminate research-based best practices and curricular resources in teaching and learning with Earth-space data;
- Identify knowledge and policy gaps and develop strategies that will more effectively address those gaps;
- Propose the development of new democratized tools, resources and methods to converge Earth and space data sciences; and
- Define the value and deliverables in terms of tools, data and workforce development of a convergence accelerator in Earth-space data science and develop a timeline for achieving them.

NEREID brought together 28 leading experts in Earth and space data science and education from across the hemisphere to the Green Bank Observatory (GBO) in Green Bank, West Virginia from November 20-22, 2019. During the workshop representatives from industry, education, research and policy-making explored and defined challenges and strategies for the convergence of big data at the intersection of astronomy, ocean, atmospheric, cryospheric, terrestrial and social sciences. The workshop drew out practices and proposal ideas to address issues in research and education, to elicit diverse points of view and enfranchise underrepresented groups, and to identify strategies with the potential to more effectively address knowledge and policy gaps. The workshop produced a set of challenges faced by geoscience and space science researchers, practitioners and educators that might be solvable in the near term through an accelerated program of convergence research and education in Earth and space science data as well as those potentially solvable during the next decade as an outgrowth of transformational approaches to research and education developed through convergence research. Searching for solutions to these near- and longer-term questions thus creates a pathway toward addressing the grand challenges that are writ large across the Earth and sky, affecting not just human health but sustainability for all life on Earth and beyond.

About Green Bank Observatory (GBO):
A world leader in advancing radio astronomy research, innovation, and education, Green Bank Observatory enables leading edge research at radio wavelengths by offering telescope, facility and advanced instrumentation access to the astronomy community as well as to other basic and applied research communities and educators. With radio astronomy as its foundation, the Green Bank Observatory is within the National Quiet Zone, nestled in the mountain ranges and farmland of West Virginia, where radio astronomers are "listening" to the remote whispers of the universe in order to discover answers to our most important and challenging astronomical questions. For more information about Green Bank Observatory visit https://greenbankobservatory.org.

About Associated Universities, Incorporated (AUI)
NEREID is an initiative of Associated Universities, Inc., a non-profit scientific research and education corporation that brings people together to collaborate across organizations, cultures, and borders; and engages people with shared interests in promoting the advancement of science and empowerment of the public through scientific literacy. Through their legacy of creating and supporting large-scale scientific enterprises in fundamental physics and radio astronomy, AUI is leveraging its expertise through NEREID to be a convener, connector and catalyst to accelerate discovery, innovation, and education in Earth and space science data. As the future of humanity depends on the next generation of data-literate public, corporate, scientific, and policymaking sectors, there is no better time than now for NEREID. For more information about AUI visit https://www.aui.edu/.

Figure 1: NEREID participants at the Green Bank Observatory in Green Bank, WV on November 20 - 22, 2019

The purpose of this report is to document the workshop activities and outcomes and to provide a timeline for advancing NSF's priorities by addressing gaps in understanding and processes in the convergence of scientific knowledge across Earth and space data sciences. With the necessary abbreviated timeframe for this effort, the organizers were unable to use a pre-workshop cooperative inquiry process to frame the gathering, so this preparatory work was integrated into the main workshop. We provide background information on how we define convergence and the value of convergence research and education to Earth-space data sciences in Section 2.

The workshop began on Wednesday afternoon with introductions, an orientation, and a "Drake Equation Challenge" to help attendees think about the relationship between celestial, planetary, and Earthbound ideas. (see appendix G). Green Bank Observatory is home to many innovations and transformative ways of thinking. The Drake Equation attempts to identify the number of civilizations in the Milky Way Galaxy with which communication might be possible. The historic meeting that led to the Equation took place in the (now named) Drake Lounge at GBO, the very spot where the NEREID workshop welcome reception took place. (It is important to note that attempts to solve the Drake Equation require an interdisciplinary approach involving astronomers, geoscientists, data scientists, ecologists, biologists, social scientists and others - a reflection of the kind of process that NEREID will require to accelerate convergence.) Participants formed four interdisciplinary teams and spent part of the evening in discussion about the Drake Equation Challenge. Teams were asked to find time throughout the workshop to develop a plan to constrain the variables in the Drake Equation. Each team would then present their plans in Drake Lounge during a Friday evening debrief. This activity set the tone for the workshop from the very start and ended up being a powerful catalyst by engaging participants in a real-time understanding of the convergence process.

The core activity of the workshop was a process of presenting, brainstorming and explicating potential Earth-space data science ideas that would help define the utility of a convergence accelerator. This process took place over a two-day period and consisted of presentations and demonstrations of tools, data and methods in Earth and space sciences, which attendees annotated with various aspects of related knowledge and needs to define the kinds of activities that a convergence accelerator would explore. The presenting and brainstorming activity is described in Section 3. The output of the workshop activities is summarized in Section 4. Section 5 describes an analysis of the structure of relationships among workshop attendees as well as indicators of an emergent community of practice derived from the pre- and post-workshop surveys. Section 6 provides conclusions and next steps. Appendices contain detailed information from the workshop and pre- and post-workshop activities. Artifacts, transcripts and voice recordings of the entire event were used to validate the accuracy of this report and participants were provided with a period for review of the draft report. Additional input was provided by interested parties that were unable to join the meeting and has been integrated into this report as appropriate.

2. General Convergence Opportunities and Challenges

2.1 Historic Context

Since the dawn of humanity, indigenous peoples of the world have observed the Earth and the heavens above, and this Earth-sky connection has been a fundamental element in the development of their culture. In the sky above they perceived information that has helped them to define their culture, religion, agriculture and livestock, spiritual, and even temporal points of view. Observations of the sky have been used to establish calendars, units of measure, the most optimal time to harvest or sow, travel time, orientation, etc. This indigenous worldview connects knowledge of the Earth's surface with the inner layers of the Earth and the known universe. It is a knowledge that has developed over time, and one in which it is impossible to conceive of these pieces individually. This perspective has been compared to systems thinking, which helps to connect everything, from the environment to individual health to business enterprise and ecology (Heke, et al, 2019; Romm, 2015; Leopold, 2007)

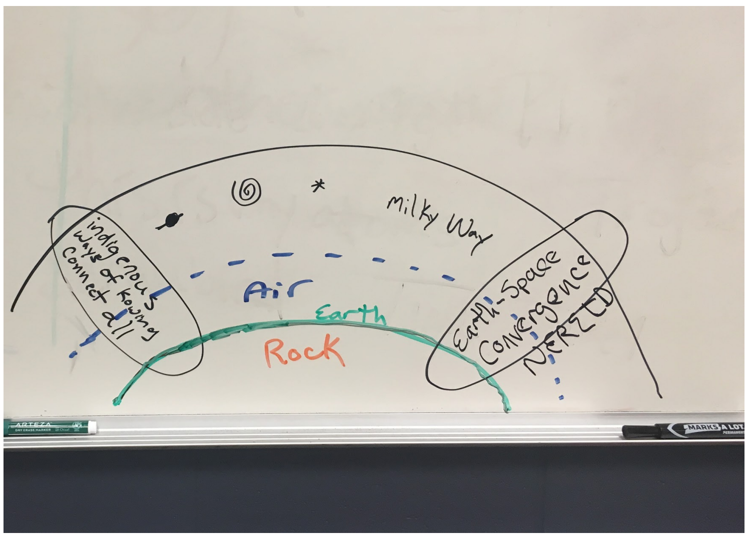


Figure 2: Sketch by Tim Spuck and Yasmin Catricheo Villagran

Through this worldview, indigenous peoples have established connections with everything that surrounds them, including the rivers, plants and animals, and the stars and planets above. This worldview is learned from the environment in which the person grows. As part of the enculturation process, the child begins to learn not only the language and customs but also the assumptions, premises and basic concepts of his/her parents, family and community. Sanchez (2010) states that: "The worldview can be described as lenses, models or maps from the point of view from which people perceive reality". That perception of reality establishes the connections between what exists with what is not even observable.

The "cosmovision" or Earth-space connection of the original people (Grebe, Pacheco & Segura, 1972) has managed to establish worldviews so compelling that many of them persist to this day. One example of this Earth-space relationship is the Inca city of Machu Picchu. Temple exits and entrances and other aspects of the city were constructed based on important observations of the sun and other celestial objects. In North America, the Navajo people used the stars to explain the origin of time and the Earth itself. When Ursa Major was on the horizon during late May and early June, it indicated that it was time to sow. The Maya - known for their very detailed calendars - integrated their myths and beliefs within each of their activities, relating what they read from the stars to their daily tasks. To the south, between Chile and Argentina, are the Mapuche people, who have also based their culture on the observation of the heavens, where the Sun and the Moon have played a fundamental role. In their book "Wenumapu: astronomia y cosmología mapuche" (2014) Canio and Pozo describe how the observation of Venus or Jupiter in its morning phase helped the Mapuches understand that in a short time the day would begin. The Milky Way, called "Wenuleufu" or "the river above", is believed to keep the stories of the ancestors, stories that have been transmitted from generation to generation through the speech of the "Mapudungun", the Mapuche language. Wenuleufu is also an Earthbound river, and at a specific time of the year the river on Earth and the Milky Way align so that one seems to flow into the other, creating an Earth-sky pathway (Kelly, 2018). The Lakota tribes also have an Earth-space connection within their worldview. In the words of John Two-Hawks: "We walk on this earth. Our feet, our bodies and our rhythms are connected to the ground. We must remember this when we reach for the sky" (2014). In Australia, the Ngarinyin people establish that everything that has been created has representation on Earth as in the Universe, implying that both dimensions (Earth-universe) are not conceived individually or separately (Hamacher, 2014).

All of this ancestral knowledge accounts for a great source of information, which has been transmitted in many cases through speech or through Petroglyphs, Pictographs and inlays in *piedas* or caves. This knowledge, or database, is not isolated - it connects the Earth and the universe. Similar to the goals of NEREID, this knowledge establishes a

connection between astronomical data and Earth sciences, in a reliable source, and one that - contrary to what has happened with the knowledge of some people originating in the world - is open and available for whomever wants it.
- *Yasmin Catricheo Villagran, (member of the Mapu Trafün people and workshop participant)*

2.2 How We Define Convergence

The National Research Council (NRC) report on convergence has the subtitle "Facilitating Transdisciplinary Integration of Life Sciences, Physical Sciences, Engineering and beyond". This report provides an introduction and guidance to the topic of convergence. The first sentence in the summary describes convergence as "an approach to problem-solving that cuts across disciplinary boundaries. It integrates knowledge tools and ways of thinking from life and health sciences; physical, mathematical, and computational sciences; engineering disciplines, and beyond to form a comprehensive synthetic framework for tackling scientific and societal challenges that exist at the interfaces of multiple fields" (National Research Council, 2014, p.1).

The NRC report lists the number of common challenges encountered in fostering convergence. They are: establishing an effective organizational culture, structures, and governance; addressing faculty development and promotion needs; creating education and training programs; forming stakeholder partnerships; and obtaining sustainable funding. To meet these challenges the report suggests four characteristics of a successful convergence ecosystem:

People Ensuring that leadership of supporting organizations are committed, as well as encouraging leadership from all participants in a project and encouraging communication among all contributors to the convergence project;

Organizations Organizations need to be open to change and work from their strengths to contribute across disciplines. As convergence happens at the frontiers of knowledge and intersection of disciplines, failure is a real possibility and this needs to be accepted by organizations;

Culture Inclusivity, diversity and mutual respect across disciplines that encourage open sharing and the development of capabilities to converse across boundaries and cultures is vital;

Ecosystem Creating new relationships and value chains is an important characteristic of successful convergence projects.

Included in the Council's report is a set of definitions examining the process of research as it evolves from a discipline focus, through multi- and inter- to transdisciplinarity. As the report indicates, transdisciplinarity transcends disciplinary approaches through more comprehensive frameworks including the synthetic paradigms of general systems theory (GST) and sustainability. General systems theory is defined as the transdisciplinary study of the abstract organization of phenomena, independent of their substance, type, or spatial or temporal scales. It investigates both the principles common to complex entities, and the models that can be used to describe them (Heylighen & Joslyn, 1992). As a language of integration and synthesis, systems science and systems theory offer a number of approaches, governance options, methodologies and tools to achieve convergence in the development and application of data science, data literacy and data use across the Earth-space data field. Several of the participants in this workshop have a rich understanding of GST, which allows the team to think about how converging disciplines connect, mutualistically influence one another and ultimately can be considered a single domain.

Convergence through systems thinking is a way to manage and allow differences to inform change and innovation and to create unique partnerships. These partnerships encourage ideas with multiple differing perspectives to institute an imaginative, creative process of inquiry that would not otherwise be possible. This is in contrast to the notion of homogenizing diverse entities into a uniform set of ideas, which can have the paradoxical effect of stifling innovation (Hedley, 2000). Similar to indigenous approaches to cosmovision, systems theory emphasizes the value of diverse inputs to a system, yet expresses them through wholeness. As Ashby indicated, the robustness of a system depends on the requisite variety needed to adapt to the diversity of challenges that its environment produces. In other words, it needs to have a repertoire of responses as nuanced or complex as the environment it is in (1991). This applies equally well to systems of human knowledge and discovery. Handling the intellectual challenges of research in a converged Earth-space domain demands embracing the kind of complexity and adaptivity that comes from diverse, equitable, and inclusive intellectual ecosystems (Boisot & McKelvey, 2011). GST provides an approach to bringing communities together, understanding the normative behaviors of converging cultures, and revealing the challenges and opportunities of isomorphism (Dimaggio & Powell, 1983).

2.3 A Position on Accelerators

The notion of an "innovation accelerator" has pervaded business parlance for some time but it is more recently being applied to team science (Horowitz, et al 2017). Science accelerators must be able to pipeline innovation into practical outcomes in ways that overcome barriers to diversity, equity and inclusion; enhance the intellectual prowess of diverse team members through mentoring and accessible tools and resources; provide on-demand services for collaboration and data access (in this case Web- or other electronic platform-based rapid, semantic data access and Earth-space ontologies); leverage participatory and inclusive design principles; and develop communities of practice and collaborative network processes to bring exogenous knowledge to participants while maintaining trust (Pauwels, et al, 2016; van Harmelen et al, 2012; Helbing & Balietti, 2011; Ziemer, 2009). But accelerators are not universally applicable to innovations that require a deeper and wider catchment of research and development, and the success of accelerators is not well documented (Konczal, 2012; and Dempwolf, Auer & D'Ippolito, 2014). For those reasons, we differentiate innovations that demand incremental innovation on extant tools, platforms, data and technologies from those innovations that require transformational approaches to research and development for which requisite knowledge may be incomplete. The purpose of this bifurcation of Earth-space data science innovation is to allow for the trajectory of the incremental kinds of innovation on the short horizon, toward a target of long horizon innovation. Thus, accelerator activities become part of a bigger picture moving toward addressing aspects of grand challenges that may be out of reach of an accelerator capacity in the short term. We provide a timeline below (Section 4.0) based on participant feedback to demonstrate how such a process would work.

2.4 Why Accelerate Convergence?

As with many areas of science, we have reached a nexus at which the sheer volume of data streaming from Earth and sky observations requires a transformation in the kind of tools, techniques, and organizing systems used to manage these data, and to create a new generation of analytical and visualization tools useful for the data-driven nature of most scholarly inquiry and discovery (Hey et al, 2009). As the pace of innovation and the amount of data being gathered increases, so must our capacity for sensemaking and analysis. Some sobering statistics support the urgency of accelerating convergence in data sciences: the number of currently active researchers in all science domains exceeds the total number of researchers ever alive; over 50 million scholarly science papers have been published so far and currently are being published at a rate of 2.5 million per year. Some areas of science produce more than 40,000 papers a month (Jinha, 2010, Ware & Mabe, 2015, Parashar, 2009). This "Age of Big Data" (Berman, et al 2018) warrants that we create not only novel tools but a new kind of rapid convergence culture to develop and share tools and techniques for sensemaking and analysis among disciplines, and to accelerate the pace of innovation. This is particularly true in what might be called "Big Science" disciplines (Hallonsten, 2016), and nowhere is this more evident than in geoscience and astronomy data (Bundy, 2007; Zhang & Zhao, 2015).

Within the Earth and space sciences today, remote sensing alone accounts for a significant portion (roughly 10 petabytes) of all the data that currently exists (Chi, et al, 2016), and in 2015 the largest astronomy data center provided approximately 100 petabytes of storage for astronomy data. In the near future, the Large Synoptic Survey Telescope and the Square Kilometre Array will produce 7.3 petabytes and 1 exabyte of data per year, respectively (Stephens, et al. 2015). However, while geosciences and astronomy separately produce vast amounts of data on a daily basis, these data have synergistic qualities. They are both highly spatial, yet contain significant levels of detail that benefit from high sensitivity, high resolution, layered detection and sophisticated visualization techniques, and are ripe for the use of artificial intelligence approaches to improve detection and characterization of both target phenomena and making new discoveries. There is also a well-established body of work on how planetary and Earth sciences inform each other (Lovelock & Giffin, 1969; Lovelock & Margulis, 1974; Kreidberg, et al, 2014; Doyle, et al, 2019). Today, big data in Earth and space science are not only used to get you from point A to point B, or to predict your local weather; they are used to provide early warnings about earthquakes and hurricanes, track near-Earth asteroids, discover and characterize planets around distant stars and unlock the secrets of dark matter and gravity waves (Sellars, et al, 2013; Belehaki, et al, 2016; Tompkins, Cain & Becker, 2019; Elliott, Walters & Wright, 2016). Whether we are looking out into space or back at Earth, the data are often organized and analyzed in similar ways with similar tools, but do not necessarily intersect in ways that will be effective as these data continue to scale.

Here on Earth, the most important and increasingly urgent problems that humanity faces are complex, having many causes and effects across social, biological, environmental, physical and technological systems and, as such, new approaches are needed. Bringing together widely disparate data sources with new tools is becoming a routine interdisciplinary way to understand our world and our communities, including using land management, navigation, ecosystems ecology, socioeconomic, political, defense and security applications (Macauley, 2006). It has become evident that bringing together a wide variety of data from Earth and space systems together with socio-environmental data has the potential to provide a pathway to innovation that can address some of humanity's most pernicious and intractable problems such as climate change, poverty, access to clean water and food supplies, and eradication of disease (Ketter, et al, 2015; Desouza, & Jacob, 2017).

Space science is beginning to inform earthbound needs in biomedicine and engineering innovation. Reed (2011) and Hughes (2007) report on the emerging field of astronomical medicine, in which the same kind of astronomy image analysis tools and astronomical algorithms are used to identify, characterize and find biomarkers for cancer cells. A number of similar astronomical technologies and techniques readily transfer to engineering and science problems in other sectors. The charge-coupled device, for instance, which allowed the digital revolution in imaging, was originally designed for use in astronomy (Rosenberg et al, 2014). The adoption of the Flexible Image Transport System (FITS) image data format enabled convergence of astronomy into a multi-wavelength era and was the precursor to multi-messenger. This format allows researchers to examine data from an x-ray satellite and a radio telescope and overlay them (Greisen, 2003).

However, it is abundantly evident that to bridge the silos of Earth and space data science at scale requires a network of innovators prepared to exchange and build the ideas that will serve as a catalyst to converge Earth and space data research communities. They must also work with policymakers and tool developers to identify needs both now and in the future and to create scalable infrastructures to accelerate discovery and innovation at the frontiers of human knowledge. At a recent EarthCube Annual Meeting, a group of attendees were surveyed about NEREID and the convergence of Earth and space data, with the results that EarthCube members:

- Agreed that Earth and space science are intimately connected, and there are needs for integration of tools and methods;
- Expressed an interest in learning about the broader space science research community, and how space science applies to their work;
  Indicated that support, development and sustainability of Earth and space systems education was important and new ways to engage lifelong learners and citizen scientists in Earth and space data research are needed;
- Emphasized the need to focus on what can be learned about social systems and human-occupied geographical regions through new techniques;
- Wanted to learn how NEREID can inform the development of time-series visualization of spatial datasets and can help model the natural habitat of the Earth to help with analysis of data; and
- Agreed that NEREID can help disseminate information about tools to community members and policymakers, and collaborate with online tools and data repositories.

2.5 A Socio-Technical framing for Effective Transdisciplinary Work

In creating a framework to facilitate conversations across the Earth-space and data science domains, NEREID uses the three cross-cutting categories identified by Edelson (2003) into which scientists' activities fit: (1) attitudes, (2) tools and techniques, and (3) social interactions. Here we discuss this framework as it applies to the challenges and opportunities of converging Earth-space data.

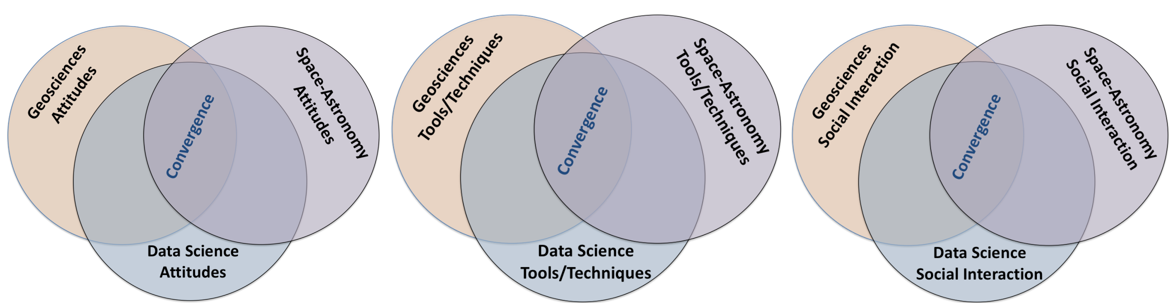


Figure 3: Applying Edelson's Framework to Earth-space convergence

Attitudes

Scientific practice is characterized by attitudes of curiosity, uncertainty and commitment (Edelson 2003). Time is an essential component for real-world science (Bencze 2000; Anderson 2002; Chinn and Malhotra 2002; Rahm, Miller et al. 2003; Robinson 2004; Zion, Slezak et al 2004). Scientists will often spend years investigating a single question. This effort requires strong personal and professional commitment. Further, scientists hold a respect for evidence and do not rush to judgment; they think critically, maintain an attitude of honesty, objectivity, and a willingness to change their minds when confronted with opposing evidence (Kozlow & Nay, 1976). Intuition, or as some might describe it, a "hunch," is also an important attitude in the practice of science. Intuition is grounded in the accumulation of attitudes and beliefs that are derived from individual and cultural experience, and is often used to provide a conceptual framework from which new ideas, research questions, and research findings emerge (Wilder, 1967; Beveridge, 1950). Closely linked to intuition is imagination. Imagination not only leads us to the discovery of new facts, but it is often the birthplace of ideas for new projects and initiatives (Beveridge, 1950). Spuck (2017) found that US astronomers consider the attitudes described above to be either of much importance or extreme importance in the practice of astronomy. As an example, astronomers lead many other disciplines of science in the sharing of data. (Henneken, 2015).

But attitudes across disciplines are another matter. Cultural differences and differing attitudes across qualitative and quantitative data types and disciplinary divides create challenges in integrating and converging data (Chayes, 2013; Miller & Mansilla, 2004; Green, et al, 2007). These cultural challenges relate to developing a diverse, inclusive, and sustainable community of practice that builds interdisciplinarity, supports mastery of the tools and techniques of data science, and evolves them to advance data-driven sciences in effective and efficient ways (Brown, et al, 2015). Evidently there has been no lack of understanding among the geosciences of the importance of advancing tools and infrastructure for sharing data. Perceived barriers are more related to a gap between attitudes about the importance of data sharing and how difficult it is to accomplish in practice (Cutcher-Gershenfeld, et al, 2016). While the disparate cultures between data science and geoscience is an important factor, there is a lack of incentives and support, particularly among geosciences, to leverage advanced data sciences, obtain appropriate training, and work toward a truly interdisciplinary relationship between the two (Cutcher-Gershenfeld, et al, 2016). Thus, the gap between data science and geosciences has more to do with the ways data science and geoscience support excursions into each other's domains. In addition, both astronomy and geosciences have struggled to explain to various stakeholders, including practitioners in other science disciplines, the value that these domains bring to social systems and human-scale challenges (Manduca & Kastens, 2012; Omenn, 2006), which is an impediment to the kind of increased support needed to move Earth-space data sciences forward.

Tools and Techniques

"The practice of science in any modern discipline includes a set of tools and techniques that have been developed and refined over the history of the field" (Edelson, 2003). In addition to attributes such as intuition and imagination, advancements in technology throughout human history have propelled changes in the way geoscientists, space scientists, astronomers, and data scientists engage with their science or come to know new knowledge. There are also similarities in research techniques. For example, geoscientists and space scientists often find their studies involve sampling large populations (e.g., stars, galaxies, fossils, ice, various geologic features, etc.) over time in an effort to understand the historical narrative of our planet and the universe, and to provide a glimpse of the future.

The tools of space scientists and geoscientists have changed significantly over the past 50 years. However, whether one is looking out into space or back at Earth, there are many similarities that present opportunities for convergence. Today, astronomers and other space scientists have access to space-based telescopes, large ground-based telescopes, robotic telescope networks, high-resolution spectrometers, gravity wave detectors, etc. Geoscientists today have access to orbiting satellites, ground/ice penetrating radar, airborne gravimeters and magnetometers, and global and real-time kinematic positioning systems. What these instruments have in common is that they produce large amounts of data that require innovative processing, metatagging, storage, acquisition, and analysis solutions for both the research and education communities. Considering that US astronomers spend on average 70% of their time using computational tools, and regularly use data archives in their research (Spuck 2017), data tools play an increasingly critical role in their work. While data solutions may in some cases be unique to specific disciplines, the data solutions that work in astronomy and space sciences to a large degree work similarly in geosciences, and vice versa. However, major obstacles to moving forward include:

- Large scale data collections for metadata, federation and data distribution lack effective means for publishing at scale and performing useful analyses in a reasonable amount of time. This capacity is needed to provide useful on-demand data services for active research, particularly in climate sciences (Schnase et al, 2015)
- Breakthroughs in intelligent systems are needed in order to provide effective new forms of knowledge representation, data collection and adaptive sampling, data and information integration, interconnection and modeling (system of systems), machine learning and AI, and intelligent user interfaces (Gil, et al. 2018).
- Advanced software solutions, providing the capacity to analyze diverse data sets across disciplines, are necessary.
- Standards for semantic stores and query languages have difficulty encompassing the wide gamut of parameters needed for advanced spatially-enabled semantic data management. The sheer volume of Resource Description Framework (RDF) content renders query execution unstable; semantic heterogeneity problems remain in big spatial data sharing; there is a persistent lack of temporal components; and scalability issues and crude visualization tools will need to be overcome for the semantic web to be useful for spatial reasoning tasks in geosciences as well as space science (Patroumpas et al, 2014; Zhang, et al 2017).

Social Interaction

Social interaction is another key feature of authentic scientific practice. Social interaction among scientists includes the same mix of cooperation, competition, agreement, and argumentation that accompanies all human social activity (Edelson, 2003). However, one of the primary characteristics of research science is collaborative work (Eason, 2004). These social activities can take place either face-to-face, virtually, or through various forms of written and visual communication. A recent survey indicates that U.S. astronomers, on average, collaborate in significant ways with 6 to 10 colleagues per week (Spuck 2017). These collaborations are not limited to other astronomers. A recent survey of U.S. astronomers indicates that 19% of astronomers collaborate with individuals in computer/information sciences, 13% collaborate with electrical/computer engineers, 10% collaborate with aerospace/aeronautical/astronautical engineers, and 9% collaborate with Earth/atmospheric/ocean scientists (ibid). Further, collaboration does not just take place domestically. Nearly 24% of physical scientists engage with others internationally as part of their work (National Science Board, 2012).

This collaboration often results in joint research publications. Most research papers today are written by collaborating authors; there are very few single author papers (Frogel, 2010). More than 70% of US astronomers are considered an author or co-author on 40 or more research publications (Spuck 2017). Geoscientists, astronomers, space scientists and computing scientists also interact socially at various conferences and other meetings. Nearly 43% of U.S. astronomers report attending 1 or 2 professional meetings within the past year, and nearly 29% report attending 3 or 4 professional meetings in this same time period. Additionally, U.S. astronomers report, on average, making slightly more than 4 formal presentations per year (ibid). U.S. astronomers also spend significant time mentoring others, and astronomers indicate that on average they engage in education and public outreach activities somewhere between 4 and 9 times per year (Spuck 2017).

Geosciences needs an equally effective social infrastructure that can enfranchise a broad constituency and provide the kind of resources, training and support to sustain engagement, and converge a wide variety of data into advancing useful knowledge. As Ribes & Bowker (2008) indicate, attention paid to cultivating interaction at "boundary objects" (such as shared infrastructure, interdisciplinary workgroups, meetings, etc.) can set the tone for better collaboration and communication. We must acknowledge, however, that geoscience and data science use differing epistemologies for their respective investigative fields. As Kastens et al. (2009) indicate, the geoscience community of practice is concerned with situating and visualizing knowledge in very large temporal and spatial scales, are theoretically driven, and use keen observational skills to inform their thinking. By contrast, data sciences are driven by the detection, storage and processing of abstractions for both observable and non-observable phenomena, patterns and statistics of those data, and a focus on the machine manipulation of those data to create knowledge. As data science is increasingly used in the pursuit of geoscience and astronomy, differences between these approaches must be resolved through understanding the utility and value of data-driven tools and techniques, and by encouraging strong partnerships between data scientists, geoscientists and space scientists.

Importantly, because astronomical data resides at such high resolution and at such a wide gamut of data types, as increasingly detailed data about celestial objects are obtained, advanced geosciences will accordingly become of increasing utility in characterizing planetary sciences. As indicated in a recent NSF report from the Workshop on Information and Intelligent Systems for Geoscience (NSF, 2015), advancing geosciences in the twenty-first century requires a synthesis with intelligent systems. The report indicates that the geosciences need these advanced computational and data analytical approaches because such data are typically characterized as "uncertain, intermittent, sparse, multi-resolution, and multi-scale. Geoscience processes and objects often have amorphous spatio-temporal boundaries. The lack of ground truth makes model evaluation, testing, and comparison difficult" (p.7). The report enumerates breakthroughs needed to further geosciences, including: better knowledge representation, new collection capabilities, better model integration and interconnection, machine learning, and intelligent user interface development. Finally, the workshop report recommended: (1) interdisciplinary community building through sustained multi-year collaborations; (2) educate and build awareness of computer scientists and geoscientists in each other's fields; and, (3) establish direct partnerships among intelligent systems and geoscience researchers. Thus, aligning these needs with space science becomes a significant guidepost for thinking about increasing knowledge about and the convergence of Earth, planetary and space sciences.

Convergence, Quantum Engineering in Earth-space Sciences and the Long Horizon

New tools and techniques are on the horizon. Vast improvements in our ability to sense and gain knowledge about our world and the universe reveal a level of complexity that was unimaginable only a generation ago. But as the past fifty years of research and development in remote sensing and astronomy demonstrate, many of the phenomena that contribute to the complexity of Earth and space systems remain undetected and poorly understood. Over the past decade there have been increasing calls to work toward a deeper understanding of the limits of photon detection in support of overcoming noise limitations of extant sensing and imaging technologies, and rethinking the very nature of sensing and imaging as we know it. We believe the next great frontier in Earth and space sciences will be explored through a quantum lens. However, while quantum engineering has advanced at the micro scale, advancing quantum engineering at the meso- and macro-scales will require breakthroughs that are only in the earliest stages.

Detection and sensing in Earth and space sciences is limited by a number of factors, both environmental and engineering: scattering, absorptive and thermal atmospheric disturbances of electromagnetic waves and electrostatic and electromagnetic disturbances and phenomena (Uzzo, 1994), as well as sensitivity, resolution and noise limitations of the sensing and signal path. Current technologies to manage these factors, such as the use of adaptive optics, multiple and over sampling, interferometry, reference subtraction, and a variety of other techniques have their own inherent limitations and tradeoffs in terms of expense and the complexity of sampling and processing systems (Morgan, et al, 2018; Yue, et al, 2016; Mennesson, et al, 2016; Kenter, Kraft, Gauron & Amato, 2017; Baker, et al, 2016; and Jorden, et al. 2016). In particular, the need for increased dependence on interferometry as indicated in NASA's Astrophysics Roadmap (Kouveliotou, et al 2014) constitutes a significant barrier to developing smaller, more adaptive and easier to deploy instruments that would enable next generation space science and remote sensing to be less costly and allow more widespread research, educational and commercial applications to be achieved.

For over a decade there has been interest in the potential for quantum optics to both provide breakthroughs in resolution and noise limits and to reveal astronomical objects and phenomena that cannot be detected through classical systems (Dravins, 2008). Its use in interferometry has been proposed in array communication (Gottesman, Jennewein & Croke, 2012), and quantum interferometry is already revealing the structure of gravity waves and plans to synoptically map them are on the horizon (Kouveliotou, et al 2014).

A focus on quantum and hybrid sensing is needed to advance currently limited areas of sensing, imaging and ranging, both classical and quantum (Altmann, et al. 2018). Importantly, much of current research in quantum sensing depends on coherent light, but as fields of study such as quantum information and metrology mature, the gamut of questions about the utility of quantum light is broadening and, increasingly, addressing how quantum and classical light theories actually inform each other and benefit the disciplines as a whole. Indeed, some of the most important questions about sensing come from incoherent light in the classical realm, and the lack of attention to the role of quantum imaging in a wide variety of needs in characterizing celestial and terrestrial sensing needs to be addressed.

Promising efforts currently underway lie in a number of weakly connected areas of quantum research, including, but not limited to: quantum coherence, capacitance coupling, light squeezing and other non-classical approaches to advancing theory and applications for detection and sensing (Moreau et al. 2017; Siwen et al, 2016 Siwen et al, 2018; Zhang, et al 2017; Young, Sarovar, & Léonard, 2018; Dwyer, 2014; Unternährer, et al, 2018). But much of this research remains in the realm of theory or consists of many small-scale experiments that are not well coordinated. There is also a need to assess the relevance of efforts to apply quantum mechanical solutions to computational and data communication systems and metrology that point to new approaches to detection and signal processing systems for other applications (Hosten, et al, 2016: Genovese, 2016). For example, recent interest in entanglement is salient

(Konrad & Forbes, 2019). Comparative testing of ghost imaging and classical laser radar imaging for remote sensing has been conducted (Hardy & Shapiro, 2013). In this case, while performance of ghost imaging LIDAR and classical imaging laser radar systems were shown to be comparable, the potential of ghost imaging to achieve sub-shot noise performance makes it attractive for aggressive research to provide improved practical hybrid imaging systems in the near term, and sub-Poissonian quantum detection and imaging systems in the mid-term (Brida, Genovese & Berchera, 2010; Erkmen, 2012). More recently, Lemos, et al. (2014) demonstrated experimentally the use of quantum interference in the lab to create an image through detection of idler photons (see Fig. 4). Because the illumination photon is not detected, it can allow different wavelengths to be used for illumination and detection and allow entanglement to be maintained. Future experimentation in this domain could have applications in environmental studies and even remote sensing applications. Doriese, et al, (2017) are developing lab-based quantum x-ray and gamma ray sensing systems, essential spectra for detection of celestial phenomena. Mount, et al (2017) describe using microwave detection, essential for experiments to characterize quantum particles in dark matter. A Keck Institute for Space Studies report (Erkmen, Shapiro & Schwab, 2012) indicates the potential use of weak-value techniques in sensing of exoplanets. Finally, the work of Tsang, Nair & Lu. (2016), uses quantum metrology for measuring the separation of stars, which overcomes Rayleigh criterion, the problem in classical optics that obscures discrimination of proximal or overlapping objects in a star field.

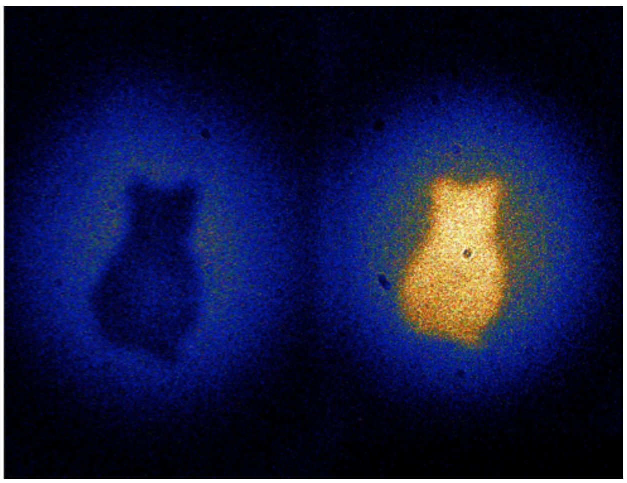


Fig. 4 Image created using constructive and destructive interference (Lemos, et al. 2014).

It is evident from these examples that areas of relevant research vary widely in the approach, research goals, and outcomes. Using an integrative process that works directly and deeply with the broad gamut of communities of research practice to converge approaches on sensing in meso- and macro-scales will lead to the kinds of cross-pollination needed to accelerate discovery. Further, because most imaging applications in astronomical and remote sensing detection and imaging depend on incoherent light and other electromagnetic spectra, we believe that important breakthroughs in overcoming the barriers to classical imaging approaches will come at the intersection of classical and quantum light, optics, and theory and experimentation, thus tapping into the brain trust of both quantum and quantum-inspired classical communities of research, and are required to address mid-horizon questions. Indeed, every area of quantum engineering can benefit from advancing research integration of quantum sensing. Fundamentally, we are at a nexus at which quantum sensing, photonics, and processing research reveals that photons have much more information to provide us about nature and ourselves. While we are not proposing that this problem can be addressed effectively in a near-term convergence accelerator, consideration of the fit between classical and quantum optics can inform near-term questions and work.

3. Workshop Methods, Approaches and Results

3.1 Process

To arrive at the goals of the workshop, the organizers devised a multi-step process that included collaborative group activities intended to draw out individual and group responses as well as general consensus. The intention of the organizers was to use this process in order to arrive at a multi-layered and multi-scale view of the current state and future needs of the Earth-space science community, specifically in regards to data use and education and learning opportunities. This process is described in detail below. Figure 5 shows the steps in the process.

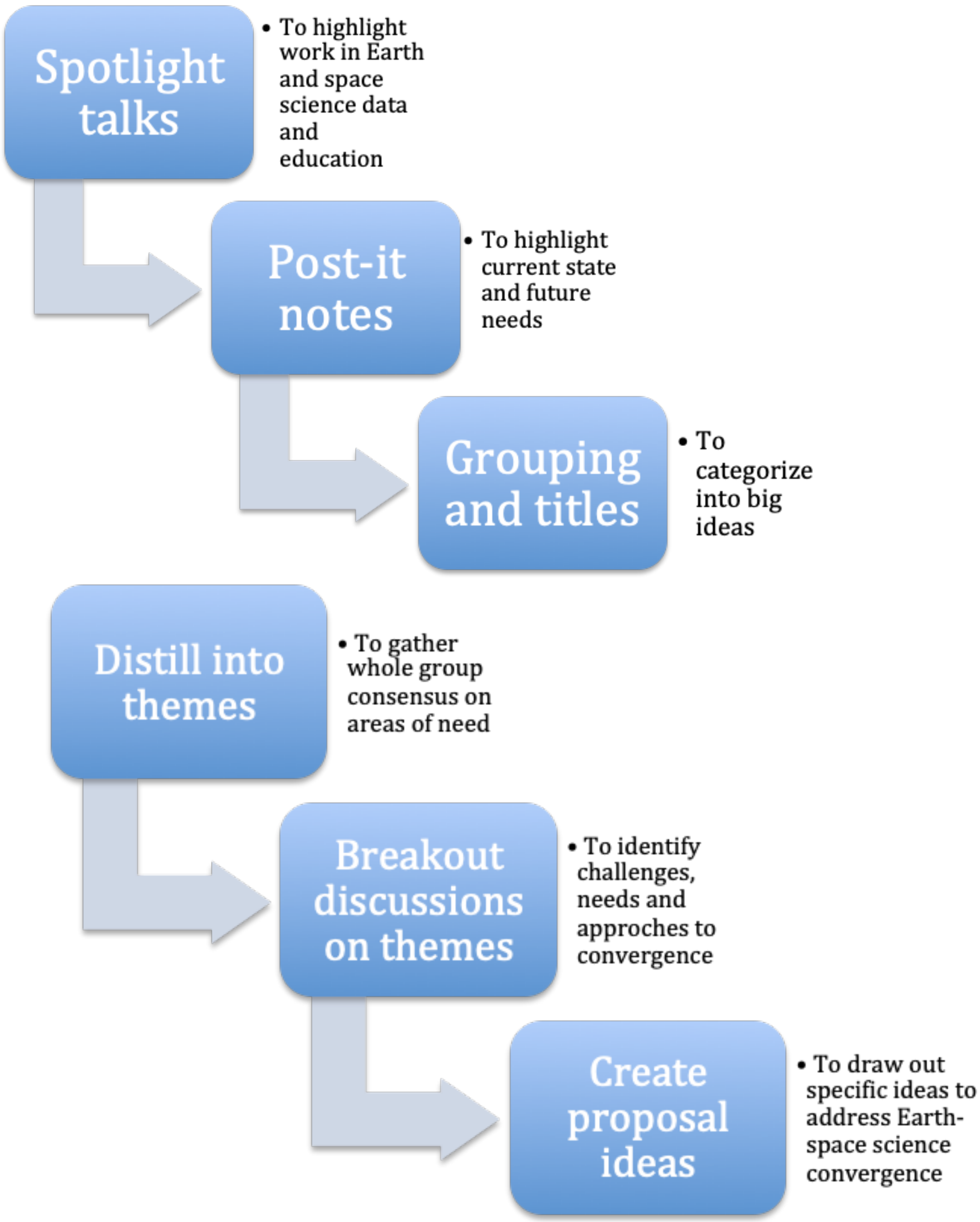


Figure 5: The discrete steps taken to bring the group focus along a trajectory, from individual research interests to collaborative consensus and development of specific project proposals to address Earth-space science convergence with data and education goals.

3.2 Spotlight talks

The bulk of Day One of the workshop was focused on a series of brief talks, called "Spotlights". The purpose of these talks was to provide an opportunity for participants to share their work and highlight the Earth and/or space science resources they use (e.g. tools, programs, databases) as well as challenges they have encountered. During the talks, the other participants were asked to write down specifics on Post-it notes - whether names of tools, challenges to be addressed or ideas for new resources - and then put the Post-its up on the conference room wall. The purpose of this simultaneous activity - listening to the Spotlight talks and commenting in real time - was to elicit on-the-fly ideas and reactions, thereby capturing and creating a multi-variable map of the current landscape of Earth-space science research and education. At the end of the day, participants began to group Post-its with similar topics together, until discrete groupings began to emerge. These discrete groups were then named. It became obvious that within several groups there were Post-its that described resources or tools that do not yet exist but that should. Participants decided to pull out these Post-its and create sub-groups of needs for each discrete grouping. Thus, at the end of Day One, the workshop participants produced a map of the current landscape of Earth and space science data research and education, as well as gaps that need to be filled. A brief description of the Spotlight talks and Post-it topic groups follow.

Authors and Titles
(slides can be found here: https://drive.google.com/drive/folders/1jcy3yI9KjLvTtQlOHhcgdBzSpx34BPJW)

Round 1:

- Daniel Fuka: *Mediation of Heterogeneity: Managing HPC systems, data, and models across scientific domains to accelerate discovery*
- Elena Yulaeva: *Global Forest Link: Engaging Youth Worldwide in Collaborative Environmental Analysis and Decision Making*
- Shelley Olds: *Making Geodesy Accessible and Relevant: Workforce development, educational resources, & data exploration tools at UNAVCO*

Round 2:

- Robert Downs: *Enabling Diverse Learners to Use Integrated Earth Science Data*
- Bryan Heidorn: *ASTROLABE and other CyVerse Projects:Navigating Astronomical and Ecology Data through Advanced Cyberinfrastructure*
- Randy Kochevar: *The Oceans of Data Institute: Preparing individuals to succeed in a data-intensive world*
- Ed Summers: *Insight Without Sight: Non-visual access to data visualization and quantitative analysis*

Round 3:

- Laura Trouille: *Tales from the Zooniverse: Unlocking Data through People-Powered Research*
- Margaret Mooney*: Cooperative Institute for Meteorological Satellite Studies*
- Leigh Peake: *GMRI: Leveraging local and global data sets in linked informal and formal learning experiences*
- Elizabeth Joyner: *MyNASAData*

Round 4:

- Kartik Sheth: *NASA Capacity Building Program: Building Skills to Use Earth Observations*
- Sue Ann Heatherly: *NEREID and the NSF INCLUDES program*
- Tiffany Stone Wolbrecht: *Data to Dome: The Planetarium Pipeline*
- Jim Hammerman: *Earth and Space Science Data Education for K-12 and the Community: Building on What We Know About Data and Statistics Learning*
- Gretchen Stahlman: *Discovering Dark Data Through the Lens of Literature*

3.3 Post-Its

From these short talks (10 minutes or less each), 198 Post-its were placed on the wall. After participants grouped the Post-its, there emerged several topics. (The word cloud on the title page of this report was derived from the fully transcribed text of the Post-its.) These were then further divided into those describing the current state, and those describing needs to be met, or gaps to be filled. Below are the topics and a few high-level insights gleaned from each.

3.3.1 Current (mapping the current landscape in Earth-space science data and education)

Education and Learning

- Importance for students of having personal connection to data
- Data literacy
- Role of data stories
- Models for getting K-12 students engaged with data through working with grad students or researchers
- Embodied experiential grasp of data gathering tools
- Over 30 Earth and space science data analysis tools were mentioned - including “direct human observation”

Accessibility

- Embodied experiential grasp of data gathering
- Sonification
- Kinesthetic data vs. data only

Human-Data Interactions

- Applications to human health and safety
- Using Earth observations for UN sustainable development goals
- Human impact on the water cycle
- Data stored all around us in nature and man-made devices -- how to tap in?

Data Sharing/Access

- FAIR data management
- Too much data

Community Empowerment

- Grand challenges e.g. the first photo of a black hole - can only be achieved through collaboration
- Empower the user community
- What does a community of practice look like in NEREID context?

Partners

- Partners are resources

Funding needs

- Use the NEREID network to identify potential funding sources

3.3.2 Needs (gaps to be filled through convergence of Earth-space science data and education)

Education and Learning

- Many data analysis tools are complex to use, thereby posing challenges to teachers and students. Tools need to be simpler to use.
- Need for projects that connect remote Earth observation data and ground observations for middle and high school students
- Availability of data and tools is problematic - teachers need a “best practices” guide for using data in the classroom
- How do you teach Earth-space science with data? Need for curriculum, searchable collection of resources for integrated exploration of Earth-space science, e.g. teaching about volcanoes on Earth through using data from volcanic activity across the solar system

Tools

- A searchable catalogue/ontology of available tools that has reviews or recommender system, age-appropriate
- Data analysis tool that uses relationships instead of boxes and columns
- Involve other disciplines such as art or music

Accessibility
- Democratize access to data tools
- Lower barriers to data access
- Accessibility in data management should be included from ground up
- Access to data analysis for vision and neuro-diverse users

Inclusivity/Diversity
- Inclusivity on the project team is a critical resource
- Need greater representation from indigenous communities. They bring an understanding of Earth-space connection.
- How do we make sure indigenous voices (and other minorities) are in the conversation

Data Sharing/Access
- How do we encourage researchers to want to share their data? Make them interested in sharing

Data Convergence
- Remote sensing creates massive data sets
- Remote sensing convergence
- Trajectory of increases in resolution of spatial data – how far should it go? How much do we need?
- Phenotype-environmental-genome-multiscale effects, what about interscale effects

3.4 Breakout Groups

The resulting map of the current landscape and indicators of gaps to be filled were reviewed by the group. The groups were asked to frame their thoughts in the context of attitudes of practice, tools and techniques, and social interactions. Through group discussion, new ideas on how to approach the next steps emerged. A collaborative decision was made to identify the key areas across the topics, and then explore them in depth. This exploration would then reveal potential research projects that would advance Earth-space science convergence through the use of data and create relevant connections to education. The key areas or categories that the group identified are Education and Learning; Diversity, Inclusion and Accessibility; and Data Access, Convergence and Interaction. These became three breakout groups.

The breakout group discussion process included:
- Discussing the current status of the category as it relates to Earth-Space-Data sciences.
- Discussing the needs as they relate to Earth-Space-Data sciences convergence and the NEREID framework, looking at the information contained in the Post-It notes, and adding additional ideas as necessary.
- Listing the needs and providing further information on the timeframe for addressing these needs, e.g.:
    - We believe we could address this need now (over the next few years) with the tools and resources that currently exist.
    - We believe that in order to address this need we will need to develop additional tools and resources, but we know what types of tools and resources that would be needed.
    - We believe that in order to address this need we will need a transformation in the domains, major change, or large initiative across the domains.

After an hour of deep discussion, each breakout group reported out on their findings, presented in rough notes below.

3.4.1 DATA ACCESS, CONVERGENCE AND INTERACTION

Group members: Robert Downs, José Guridi, Bryan Heidorn, Gretchen Stahlman, Stephen Uzzo, Ilya Zaslavsky

The group began by discussing specific needs, including:

- The need for incentives for researchers and those in the private sector to participate in Earth-space data science education. NEREID could advocate for making education a more natural part of research workflows and help develop curricula and resources that can serve not just lifelong learners, but also career scientists, engineers, and user communities. The need to democratize tools and algorithms. This will only happen by making the research tools accessible, including simplified interfaces and analytical instruments and data cleaning. It is also necessary to support tools now that lend themselves to convergence, such as image processing and try to improve those that don't.
- The need to unify Earth-space communities. The group warned not to duplicate efforts to reinvent the wheel. Look for best practices and systems that share information easily. Maybe NEREID could model itself after ESIP because ESIP brings together different communities in very well-knit ways.
- The need to use data and education to explore possibilities for supporting life on other planets.

The group talked about how astronomy transfers to other areas and also looking at multidisciplinarity, more broadly, the idea that you can incentivize people to adopt and learn about science by bringing together areas that are of particular concern such as health and disaster preparedness. These are good hooks into enfranchising more diverse participation in using the data because people are actually concerned about getting people to hospitals and escape routes for floods. This idea of using the diversity of different forms of data as a way into enfranchising people who may feel isolated from computational approaches to solving problems was discussed.

Use cases for teachers in different grade bands were also part of the discussion particularly focusing on K-6. High school is an important place to think about integrating data science into different domains. Data science should be introduced across communities of researchers because different domains use data very differently. They very often use different tools and don't think about how to apply them to problems in other domains. The focus needs to be on the curriculum, but also the kinds of tools and data being used and look for what aspects of data gathering, cleaning and analysis could be simplified. Simplifying interfaces to datasets, focusing more on patterns and coming up with ways that people can find themselves in patterns to make it more relevant will also help. If making the tools really conducive to convergence with Earth and space science is a goal, then we need to be explicit in developing new tools or applying existing ones to different kinds of problems.

A couple of very specific cases of convergence came up. The idea that we do remote sensing on Earth, but we also do remote sensing on other planets and how do we compare those data and are the data shared among those communities? We should look into why there are distinct divisions in NASA for remote sensing and planetary exploration for figuring out how well they share data, and if they don't, figuring out how to improve that. But also thinking about can you federate those data? Can they be federated? And looking at the way researchers draw conclusions or patterns. Landers gather ground truth data on Mars, just like we do on Earth (such as soil sampling and chemical analysis with robotics on the Mars surface). How can we take some of these tools that are used particularly for remote sensing in these two different scenarios and figure out how they work together and seeing if that's something that in theory could be addressed in the short term and that's a really potentially fruitful area for conversion, and also looking at whether the fact that we do remote sensing and ground truth on other planets affects our perception of the data and conclusions we draw. (Are we biased because we know the data are from Mars, as opposed to Earth). Cross-fertilization of these two areas could result in richer science and better tools and reasoning. Similarly, data structures and patterns that are used for one domain can be compared and contrasted in other domains with the potential to make discoveries. For example, network science is specifically about seeking patterns of how things connect across many domains. You can find patterns in the connectivity within brains that compare favorably to protein interactions, and router connections on the Internet (small worlds). Finding ways to compare and contrast patterns in data across Earth and space data may similarly be fruitful. These are things NEREID could be doing in the near term.

Because Jose Guridi (Advisor to the Chilean Ministry of Science) was in the group this led to a discussion about the Data Observatory in Chile. While they are working exclusively in the astronomy domain, right now with ALMA (Atacama Large Millimeter/submillimeter Array) data, within the next few years they plan to look very deeply at climate data and climate change and are thinking about how the tools they are developing for astronomy can apply to climate data and other domains. Jose commented that some of the tools that were originally developed for astronomy are being used for everything from biomedicine to other kinds of earth science and social science. NEREID could be aggregating some of those things. Here are some links to some of the kinds of work that could be aggregated in the near term:

- https://science.sciencemag.org/content/331/6018/696
- https://www.iau.org/public/themes/astronomy_in_everyday_life/
- https://link.springer.com/article/10.1007/BF03178440
- https://arxiv.org/abs/astro-ph/0611400
- https://www.cfa.harvard.edu/COMPLETE/astromed/

In the more far off term is thinking about the dramatic rise in identifying and gaining knowledge about exoplanets through determining the makeup of the atmosphere, whether they are Earthlike, etc. Gathering remote sensing, planetary science and exoplanet data researchers together into a single research endeavor might take time, but in the near term seeing how well those communities talk to one another. We may find that it's not as well as we think and NEREID could be finding ways to improve that, but also looking at the outcomes of the increasing numbers of meetings and collaborations outside NASA. NSF should be made aware of the need for this and maybe that would be something to talk about for the next steps in a convergence accelerator.

There are likely education barriers to this kind of unification of planetary sciences and remote sensing that NEREID would envision. But we need more in-depth information about what is actually going on and the trends. Maybe this is a scientometric problem. We should be doing science mapping of these domains to see if they are converging or not (scientometricians were invited but unable to attend the meeting because of a scheduling conflict). Funders might struggle with understanding the value of astrobiology. The Search for Life exhibition at the New York Hall of Science was brought up as an example of the convergence of geoscience and astronomy. The exhibition compared extreme environments on Earth with those on other planets (such as the Atacama Desert and the surface of Mars, deep ocean trenches on Earth to Europa, etc.) Funders for an ExoBio Lab did not grasp the value of astrobiology. Yet what we're talking about NEREID doing is astrobiology with a mirror (remote sensing).

3.4.2 DIVERSITY, INCLUSION AND ACCESSIBILITY

Group members: Katie Naum, Claire Raftery, Ed Summers, Sue Ann Heatherly, Tim Spuck

Why diversity, inclusion and accessibility? The breakout group collectively affirmed that engaging equitably with a diverse and inclusive set of stakeholders "makes the science better," by creating an environment with more perspectives and richer, more creative insights. Non-representative, exclusionary, and inaccessible practices in science do humanity a disservice by harming all people - not only those from underrepresented backgrounds.

The group agreed that Earth and space science topics are widely prevalent in K-12 education, and may serve as an accessible "gateway" science that speaks to stakeholders of many backgrounds. From this perspective, tools and resources for Earth and space science must be inclusively developed, with the needs of minority groups first and foremost rather than as an afterthought.

The group agreed that however the problem being addressed is defined, two questions must be answered from the outset:

1. What populations must be included?
2. Across what dimensions should we frame the diversity, inclusion and accessibility efforts that respond to the needs of these populations?

The group identified the following possible answers to the above questions:

1. Populations for Outreach

1. Diverse genders
2. Racial minority populations, especially those which are underrepresented
3. Indigenous populations

4. People with Disabilities
    a. Neurodiversity
    b. Blind and visually impaired
    c. Deaf and hard of hearing
    d. Physical disabilities
    e. “Invisible” disabilities, such as
        i. Mental health
        ii. Learning disabilities
        iii. ADHD
5. First-generation students
6. LGBTQIA+
7. Populations of low socioeconomic status
8. Late stage onset disability

2. Dimensions/Frameworks for Needs

1. Accessibility - how are information and resources conveyed to stakeholders? How might accessibility be improved to meet existing needs?
   Accessible geospatial data resources are extremely limited for the blind and visually impaired community.
2. Inclusion - how might community support structures support the inclusion of outreach populations?
    a. Ensuring all students have access to representative mentors and are empowered to seek and receive support.
    b. Participation and leadership by outreach populations in Earth and space science must be meaningful - not tokenism. A group member shared an anecdote of an African American professor asked to serve as co-PI on many research proposals - not for his experience, but for his skin color.

The group then identified a set of representative needs for diversity, inclusion and accessibility. These included:

1. Ensuring the next generation of big data explorers in Earth and space sciences is fully representative of the population of the US.
2. Bringing together stakeholders to build pathways/opportunities for funding for minority-serving organizations, industry recruiters, domain scientists and others working with diverse populations.
3. The building of genuine partnerships. Underrepresented communities don’t want others to “fix their problems” on their behalf. The slogan *“Nothing about us without us”* was shared to emphasize that work on diversity, inclusion and accessibility must go beyond well-intended gesturing. External stakeholders should be mindful of the role they serve and the value they provide, rather than making this work “about them.”
4. The creation of a “virtuous” cycle - shifting attitudes on diversity, inclusion and accessibility leads to creating and highlighting diverse role models, fostering inclusive interactions, and generating new tools and resources to support DIA. This leads to a further shifting of attitudes, refreshing the cycle.
5. Standards for accessibility through opportunities for usability research are needed.
6. Funding to support these efforts is needed and could be used to create the resources that are truly accessible, e.g., an educator of students with blindness writing a funding proposal should consider what tools and resources might enable their students to become the next gen of blind data scientists. These might include:
    a. A geosciences camp for the blind and visually impaired community
    b. idata
    c. NGSS

Possible approaches to fostering diversity, inclusion and accessibility:

- Reframing expertise: in an academic setting, “expertise” is typically framed as academic knowledge, minimizing the contributions of end users and others who may not be fluent in this area. Stakeholders who possess other kinds of expertise must be recognized and participate at appropriate levels, creating a more equitable stakeholder map. These might include
    - End users with insight and context into the challenges being addressed
    - Researchers and others with domain knowledge
    - Tool developers and practitioners skilled in tools and techniques
    - Educators and communicators
    - Policy experts
- Equipping stakeholders with a tool kit to get past the preliminary steps of inclusion and accessibility.
    - Informal educators: What fundamentals do you need to consider/respond to, to support inclusion of...
- Harnessing existing resources:
    - Leverage community-led programs such as Northeast Big Data Hub workshops. The Northeast Big Data Hub is organizing a series of “Train the Trainer” workshops for data science, designed to provide skill sets to faculty and others interested in serving as trainers who take curricula back to their home institutions and host training workshops of their own.
    - What best practices might we learn from tech organizations that have diversity and inclusion programs?
    - Support and partner with minority serving organizations: Lesbians who Tech, Black Girls Code
    - Existing resources that are available or could be adapted such as quorum coding environment. Hour of code, code.org has accessibility. IDATA on astronomy side. Using the virtuous cycle model to level up, use these resources to have an experience (informal science perhaps), present findings, changes your attitude and identity, social interaction to interact with mentors, whose attitudes shift. Next level build on skills, have summer camp on Python, work towards authentic research, attend a conference. Timeline 2-3 via Need 3 and 4.
- Obtaining funding: collaborators work together to create and submit proposals to convene stakeholders and develop resources.
- Establishing a matrix of stakeholders: This table would provide a diagrammatic representation of intersecting needs and potential solutions among the many intersections of stakeholders involved in Earth-space sciences.
    - The columns would be the populations for outreach as defined above.
    - The rows might include:
        - Big Tech Tool Providers - SAS, Matlab, Google Earth, SPSS, IBM, ESRI, Microsoft, Apple, Python, Google Drive, Jupyter notebook
        - Industry D&I recruiters who need to recruit competent diverse employees
            - Some of these may hail from the same companies as above
            - In addition, draw on professional associations from large industry - SAS, Redhat, Grace Hopper, disability-IN, LBGG in Tech, SWE, Women in Data Science, Lesbians who Tech etc.
        - Formal Educators, K-12
        - Informal Educators
        - Higher Ed Faculty
        - Higher Ed Disability Support Services
        - Policy makers - especially: those focused on education policy; AAAS; Department of Education; Office of Special Education; NSF influencers; OSTP
        - NSF Large Facilities and other Federal agencies - NASA, NOAA, NSF, USG

3.4.3 EDUCATION AND LEARNING

Group members: Rich Boone, Daniel Fuka, Jim Hammerman, Randy Kochevar, Shelley Olds, Becky Reid, Peter Tuddenham, Tiffany Stone Wolbrecht, Elena Yulaeva, and Sue Ann Heatherly

The group discussed how convergence is necessary to answer questions we cannot answer otherwise. Earth and space data is uniquely positioned for convergence because in both fields, the sheer amount of data that has been and continues to be collected is staggering, and the convergence of these domains will exponentially expand research horizons – the power of tapping into Big Data. Earth-space data also gives rise to valuable learning research opportunities, such as exploring how people take in and process data - what happens when people use tools such as GIS? Research is also needed on the learning of how to use Earth-space data and visualizations and integrating them into teaching practices.

Earth-space data creates opportunities for personal connection. Linking Earth-space data to local data opens an avenue for people to use data to connect to their own world – either by collecting their own data or by using data to explore their own community – and potentially expanding outward to include the distant stars. This is also an excellent way for community engagement and empowerment, when people are allowed to collect and use their own data.

The group identified several key needs in formal education:

- Integrating Earth Space Data into NGSS
- Curriculum development and teacher training
- Identifying Data Competencies in NGSS
- Developing data science gateway tools for preK-14
- Developing a path to data literacy with Earth-space data

Engaging other sectors is seen as key. Partnering with the business community would be a way to understand what knowledge, skills, and abilities they are looking for, in order to ensure pathways to success for K-14 learners. This would also require a review of business-academic partnership best practices. Finally, a general but urgent need is a workable description of what is unique about Earth-space data. Unique qualities include: touchable, sensory, experiential, immersive, the unknown, the “wow” factor.

3.5 Project Proposals

After brief report-outs on the breakouts, all participants were then invited to further distill their thinking into specific project ideas and present them as research project proposals, building on the concepts developed throughout the workshop. The resulting set of proposed science and technology “sweet spots” and corresponding research projects demonstrates the expertise, innovativeness, and proposal writing skills of those participants who accepted the challenge.

Seventeen research proposals were presented, creating a comprehensive look at a range of approaches to Earth-space data science convergence. The titles and authors of the proposed projects are as follows (See **Appendix E: Promising Research Projects** for complete project proposals):

- Cramer, Catherine B.: *Network of Earth and Astronomy Domain Data (NEADD)*
- Downs, Robert: *Collaboratively Improving Capabilities to Use, Analyze, and Visualize Earth and Space Science Data*
- Fuka, Daniel: *DomeBroker*
- Guridi, Jose: *AstroEarth Technology Transfer*
- Hammerman, Jim; Peake, Leigh*: Understanding How Middle School Students Make Sense of Multiple Data Domains in the Context of Climate Investigations*
- Heatherly, Sue Ann: *Convergent Teacher Professional Development*
- Kochevar, Randy; Olds Shelley E.; Yulaeva, Elena: *Solar Impact Collectors On Earth*

- Raftery, Claire; Summers, Ed: *Inclusion and Accessibility Conference*
- Reid, Rebecca; Boone, Richard: *Integrating Earth Space Data into NGSS*
- Spuck, Tim: *Earth-Space Accessible Analysis Tool*
- Spuck, Tim: *Exploring Connections between Life on Earth and Celestial Events*
- Spuck, Tim: *Exploring Human Interaction with Earth-Space Data*
- Spuck, Tim: *Light Pollution Reduction through Analysis of Earth-space and Social Data*
- Stahlman, Gretchen; Heidorn, Bryan: *Resurrecting Data Across Disciplines*
- Uzzo, S.: *Planetary Open Data System (PODS)*
- Wolbrecht, Tiffany Stone: *Exploring Nature to Promote Data Literacy*
- Zaslavsky, Ilya; Tuddenham, Peter: *Open Knowledge Network for Ocean Education (OKNOE)*

In addition, as a result of this activity, two collaborative research proposals are now being developed to submit to NSF. These collaborations are a direct outgrowth of the process outlined above.

Fig 6. Workshop participant teams developing project proposals

4. Timeline

The breakout groups (Data Access, Conversion and Interaction; Diversity, Inclusion and Accessibility; and Education and Learning) discussed perceived needs in each area and how the needs would be filled through NEREID projects. These potential research projects would advance Earth-space science convergence through the use of data gathering tools and methods, and create relevant connections to education. The proposed project ideas stemming from the breakouts were loosely thought of in terms of when they could reasonably be implemented, as follows:

Short term (2-4 years)

- Matrix of stakeholders: intersecting needs and solutions among stakeholders
- Comparing NSF, NASA divisions, etc. with regard to: sharing data and identifying areas for improvement, data federation, structures and patterns, and tool sharing for remote sensing
- Exploring how well geoscience and astro-science communities communicate and collaborate
- Develop prototypes for Earth-space data analysis tools.
- Science mapping of these domains to see if they are converging or not
- Learning research on how Earth-space data and visualizations are used
- Integrating Earth-space Data into NGSS
- Identifying Data Competencies in NGSS
- Curriculum development and development of teacher training programs
- Developing a path to data literacy with Earth-space data
- Workable description of what is unique about Earth-space data.
- Establish/strengthen partnerships with the business community
- Bringing together stakeholders to build pathways for funding for minority-serving institutions
- Harnessing existing resources
- Leverage community-led programs
- Best practices from tech organizations that have diversity and inclusion programs
- Support and partner with minority serving organizations
- Existing resources that are available or could be adapted e.g. quorum coding environment
- Obtain funding

Medium Term (5-10 years)

- Looking very deeply at climate data and climate change and how the tools they are developing for astronomy can apply to climate data and other domains.
- Aggregating tools that were originally developed for astronomy that are being used for everything from biomedicine to other kinds of earth science and social science.
- Integrating Earth Space Data into NGSS
- Curriculum development and teacher training
- Identifying Data Competencies in NGSS
- Developing a path to data literacy with Earth-space data
- Linking Earth-space data to local data and expanding outward to include the distant stars.
- Community engagement and empowerment
- Developing data science gateway tools for preK-14
- Ensure pathways to success for K-14 learners.
- Building genuine partnerships
- Create "virtuous" cycle new tools and resources to support DIA.
- Standards for accessibility through opportunities for usability research
- Funding to support these efforts
- Equipping stakeholders with a tool kit for preliminary steps of inclusion and accessibility.

Long term (10 years +)

- Create true collaborative structure between geoscience and space/astro-science communities
- Ensure pathways to success for K-12 learners
- Ensure the next generation of big data explorers in Earth and space science is fully representative of the population of the USA
- Reframing expertise
- Quantum and hybrid sensing will begin to be realized with the possibility of transformation of tools and techniques and new epistemologies

5. Evaluative activities and analyses

Evaluative activities were conducted to help understand the value of the workshop and the activities around it, and to seek to understand the community of participants and the value of NEREID as a convergence accelerator for Earth-space data. The degree to which the workshop and associated activities foment cross-disciplinary collaborations and inform the development of new ideas, methodologies, applications, research plans, and educational approaches will help gauge its potential for success. We engaged in two distinct activities with separate objectives: a characterization of connections among participants and the kinds of institutions they collaborate with, and feedback on the value of both the workshop and that of an Earth-space data community of practice.

5.1 Network Mapping

As part of the NEREID workshop we wanted to understand the evolution of the convergence process, looking at the changing structure of the network of practitioners, identifying the most effective collaborations and partnerships, and looking at how they change over time. To achieve this goal we deployed a brief survey to obtain a snapshot of those relationships as they existed before the workshop commenced and then mapped the results (see fig X). These data will be used as a baseline to understand and explicate the evolution of the structure of the convergence of Earth-space data research and education going forward; how and when partnerships form; the roles and relationships among the communities of researchers and educators; and resultant intellectual growth through revealing key characteristics of sustainable partnerships and determining relative resiliency of these relationships. It will also provide a dimension of understanding of the network formation, which will inform the effectiveness of a community of practice as it evolves and the degree to which the primary goal of creating public-private, multi-sector partnerships and consortia to build an interdisciplinary community of practice in Earth-space data science research and learning is being met.

Fig. 7 Preliminary network graph based on pre-workshop relationships among participants and the institutions with which they collaborate. Note the clustering around particular nodes, indicating their importance as brokers in an otherwise relatively heterogeneous set of connections.

A preliminary look at the results of the workshop pre-surveys indicate that levels of collaboration vary widely in terms of frequency and directedness. When considering relationships with institutions outside of their own (current or recent relationships), about 68% considered these relationships important to influencing their work, and roughly 58% of these are important enough to partner with them on planning strategically for future work. Also, as Fig X illustrates, there are a few highly connected nodes that are of significant value to the community, and have at least some function as brokers of information and resources. We plan to contextualize this formative work with outcome products, such as co-authorship, co-PI networks and other products that will help provide a richer look at the structure of these relationships and some indications as to how and why they have formed.

5.2 Community of Practice

Participants in the NEREID workshop also filled out an exit survey, offered to provide an opportunity to identify their overall perceived value of workshop. This post-survey also served to seek indications of the potential for a community of practice (CoP) to develop at the intersection of Earth and space data sciences, one which will clearly respond to the needs, challenges and applications for development of a convergence accelerator proposal as well as to assess the value of converging Earth and space data sciences to the field. The survey was developed using the framework developed by Wenger, Trayner, & de Laat (2011), who identify 5 value cycles:

- Immediate Value of convergence activities themselves;
- Potential Value of the interactions in terms of skills, social capital, and resources;
- Applied Value of convergence process' effect on members and how they generalize their knowledge;
- Realized Value of recognition of the value of different abilities and expertise toward both explicit and implicit goals of the group;
- Reframing Value of redefining success, trying new approaches, obsolescing old structures, and identifying new performance metrics to reflect these definitions.

The results of this survey indicate that while a few participants questioned the value of convergence of Earth and space data, the majority expressed significant interest in converging Earth and Space data sciences and that it is of value to their interests and the future of their work. Further, they indicated that they left the workshop with knowledge of resources that they believe will be useful to them, as demonstrated by extremely high marks in this category. They indicated that the workshop itself will be responsible for significant future collaborations and discoveries and will increase the cross-scientific domain convergence research per the desires of the NSF Convergence Accelerator program. The majority of participants significantly increased their understanding of convergence research and NSF's interest in convergence accelerators. They also look forward to more clarity from NSF in the next round of convergence accelerator solicitations.

In terms of Community of Practice Value cycles, workshop activities and post-workshop survey responses indicate evidence of emerging CoP value cycles. The *Immediate Value* of convergence activities was evident in both the proposal ideas and survey results. The *Potential Value* of the interactions in terms of skills, social capital, and resources was evident in the proposal ideas and what were considered the medium-term activities to converge Earth-space data from their proposal ideas, as well as their responses in the survey. The *Applied Value* of convergence process' effect on members and how they generalize their knowledge across disciplines was revealed particularly as they worked together in teams and expressed who they would collaborate with on project ideas. The *Realized Value* of recognition of the value of different abilities and expertise toward both explicit and implicit goals of the group was detected through a high degree of emphasis on diversity, equity and inclusion, which pervaded the conversation on topics throughout the workshop. The *Reframing Value* of redefining success, trying new approaches, obsolescing old structures, and identifying new performance metrics to reflect these definitions was prevalent in conversations around reframing convergence through general systems theory, and interest in the future of quantum imaging. We are in the process of analyzing audio transcripts of the workshop, which will result in a much deeper analysis of CoP and other ideas that emerged in the rich conversations, presentations, and other workshop activities.

6. Conclusions and Next Steps

I think the next [21st] century will be the century of complexity.
- Stephen Hawking, 2000

As stated at the beginning of this report, the National Research Council lists a number of common challenges encountered in fostering convergence, such as:

- Effective organizational cultures, structures and governance;
- Addressing faculty development and promotion needs;
- Creating education and training programs;
- Forming stakeholder partnerships; and
- Obtaining sustainable funding.

While the NEREID workshop revealed the need for answering many of these challenges, it also served to underscore how the focus of NEREID on Earth and space science data provides a unique opportunity for convergence. The qualities and affordances of working with data cuts across domains, making accessible mutualistic tools a ready possibility. Bringing Earth and space science practice and education together through the use of data tools and structures thus appears to be not just possible but inevitable, as together they are poised to deepen understanding of the relationship among phenomena found on Earth, and throughout the Universe. But most importantly, embracing complexity at the intersection of Earth and space data will better prepare us to approach the complexity of humanity's most challenging problems. While it is true that geoscience, space sciences and data science disciplines have some needs that differ, the results of this workshop indicate that there are numerous areas in which convergence across the disciplines can provide fertile ground for discovery and accelerate innovations that are either not possible or impractical when attempted within disciplinary silos.

Specific outcomes from the NEREID workshop indicate the appropriateness of this topic for a Convergence Accelerator. These are:

- The use and application of data science provides a transdisciplinary lens on Earth and space sciences that will help identify and actualize areas of convergence.
- General systems theory provides tools and methods to help make connections among Earth and space science systems toward convergence of research approaches, while accommodating the diversity of ideas, equity across communities of stakeholders, and the inclusion of a broad gamut of perspectives on Earth and space sciences.
- The current Earth and space sciences data landscape, while primarily siloed, is well-positioned for accelerating convergence by bringing communities of practice together through team science, policy and education.
- There is growing interest in the kind of convergence culture that Earth and space data represent.
- Using the socio-technical framing of attitudes, tools and techniques, and social interactions helps outline needs and processes that will lead to a successful convergence effort.
- In the far horizon, the use of quantum engineering may offer solutions to some of the seemingly insurmountable technical barriers to Earth and space data sensing, imaging and analysis that we face today.

We believe that overcoming technical challenges requires a scalable, plug-and-play architecture that can accommodate a wide variety of data types, algorithms, models and visualization and analytical approaches. The synergies across Earth-space sciences and our struggles with common big data problems presents an opportunity to converge efforts to maximize impact on future investment. Because data types, storage and processing demands are constantly changing and growing, the focus of a convergence effort will be on open-source, cloud-based data storage and computing assets to minimize cost and maximize computational power and flexibility, and to leverage new and emerging tools and techniques.

As we look to the future and building the next generation of tools, there are great opportunities for space scientists, geoscientists, and data scientists to come together to explore common solutions to big problems in instrumentation. In the coming decades, geoscientists and space scientists will continue to develop tools that allow us to explore more deeply phenomena here on Earth and in space, and in doing so space scientists, astronomers, and geoscientists will pursue similar goals. These new instruments will require increased sensitivity, lower noise from the electronics, larger detection areas, faster processing, shorter time interval between observations, follow-up capabilities, etc. Converging scientists and engineers across the Earth-space sciences can lead to a greater return on investment. Further, it is likely that quantum sensing and communication will accelerate our ability to collect and analyze vast amounts of data across these disciplines. Bringing together Earth-space scientists, engineers, and educators today to explore this new frontier collaboratively holds great promise for the future.

In addition to instrumentation, if we are to inspire the next generation of Earth-space scientists and maintain a healthy workforce pipeline it is necessary to develop appropriate research-based curricula and training tools and resources, and to do it now. Former National Earth Science Teachers Association President Ardis Herrold (2019) argues that climate change and natural disasters are topics of strong interest among high school students. Astronomy is also considered a “gateway science,” and can be a motivator for various STEM learning and career opportunities (Blandford, 2014; Heatherly et al., 2011; Palmer, 2011; Fisher & Fisher, 2004). This interest can be a powerful tool to motivate young learners in the art of applied interdisciplinary data science and prepare the next generation to grapple with both emergent and established challenges facing humanity.

While virtually every student in the US takes an Earth & Space Science course at some point in their K-12 education, the learning is, for the most part, modularized. Instruction is not delivered in an integrated way, but rather as 6 weeks of geology, 6 weeks of atmospheric science, 6 weeks of astronomy, etc. The convergence of Earth-space sciences provides a rich opportunity to promote learning in an integrated and connected way. Further, due to strong interest in Earth-space sciences by young learners as well as the growing data-centric nature of these disciplines, convergence of these disciplines along with data science can be an effective tool in building the next generation of big-data explorers. To date, there is a void of integrated Earth-space and data science teaching and learning tools and resources, and research-based best practices, an urgent need to be addressed.

GBO’s motto is “The Universe is Whispering to Us.” In the context of NEREID, we might add: “. . . and the Earth is whispering back.” This gathering not only sparked great interest, involvement and commitment from participants, but also clearly highlighted the need to bring together an even larger group from the Earth and space science data communities. This next step is key to developing immediate cross-discipline partnerships in which more tools and data from each sector can demonstrate practical applications and pathways for collaboration, and that together can develop and disseminate interdisciplinary approaches to teaching and learning. Indeed, the time for NEREID is now.

References

Altmann, Y., McLaughlin, S., Padgett, M. J., Goyal, V. K., Hero, A. O., & Faccio, D. (2018). Quantum-inspired computational imaging. Science, 361(6403), eaat2298.

Anderson, R. D. (2002). Reforming science teaching: what the research says about inquiry. Journal of Science Teacher Education, 13(1), 1-12.

Ashby, W. R. (1991). Requisite variety and its implications for the control of complex systems. In Facets of systems science (pp. 405-417). Springer, Boston, MA.

Baker, I., Maxey, C., Hipwood, L., & Barnes, K. (2016). Leonardo infrared sensors for astronomy: present and future. In High Energy, Optical, and Infrared Detectors for Astronomy VII (Vol. 9915, p. 991505). International Society for Optics and Photonics.

Belehaki, A., James, S., Hapgood, M., Ventouras, S., Galkin, I., Lembesis, A., Tsagouri, I., Charisi, A., Spogli, L., Berdermann, J. & Häggström, I. (2016). The ESPAS e-infrastructure: Access to data from near-Earth space. Advances in Space Research, 58(7), 1177-1200.

Bencze, J. L. (2000). Procedural apprenticeship in school science: Constructivist enabling of connoisseurship. Science Education, 84(6), 727-739.

Berman, F., Rutenbar, R., Hailpern, B., Christensen, H., Davidson, S., Estrin, D., Franklin, M., Martonosi, M., Raghavan, P., Stodden, V. & Szalay, A. S. (2018). Realizing the potential of data science. Communications of the ACM, 61(4), 67-72.

Beveridge, W. I. B. (1950). The art of scientific investigation. Vintage Books-Random House Inc., New York.

Blandford, R. (2014). the Committee for a Decadal Survey of Astronomy and Astrophysics 2010, "New Worlds. New Horizons in Astronomy & Astrophysics", National Research Council, National Academy Press, ISBN-10: 0-309-15799-4.

Boisot, M., & McKelvey, B. (2011). Complexity and organization-environment relations: Revisiting Ashby's law of requisite variety. The Sage handbook of complexity and management, 279-298.

Brida, G., Genovese, M., & Berchera, I. R. (2010). Experimental realization of sub-shot-noise quantum imaging. Nature Photonics, 4(4), 227.

Brown, R. R., Deletic, A., and Wong, T. H. (2015). Interdisciplinarity: How to catalyse collaboration.

Bundy, A. (2007). Computational thinking is pervasive. Journal of Scientific and Practical Computing, 1(2), 67-69.

Canio. M, Pozo. G. (2014). "Wenumapu: Astronomia y Cosmologia Mapuche" Chile: OchoLibros.

Chayes, F. (2013, October). Attitudes toward data in the hard and medium-hard sciences. In Data for Science and Technology: Proceedings of the Seventh International CODATA Conference, Kyoto, Japan, 8-11 October 1980 (p. 116). Elsevier.

Chi, M., Plaza, A., Benediktsson, J. A., Sun, Z., Shen, J., & Zhu, Y. (2016). Big data for remote sensing: Challenges and opportunities. Proceedings of the IEEE, 104(11), 2207-2219.

Chinn, C. A., & Malhotra, B. A. (2002). Epistemologically authentic inquiry in schools: A theoretical framework for evaluating inquiry tasks. Science Education, 86(2), 175-218.

Cutcher-Gershenfeld, J., Baker, K. S., Berente, N., Carter, D. R., DeChurch, L. A., Flint, C. C., Gershenfeld, G., Haberman, M., King, J. L., Kirkpatrick, C., Knight, E., Lawrence, B., Lewis, S., Lenhardt, W. C., Lopez, P., Mayernik, M. S., McElroy, C., Mittleman, B., Nichol, V., Nolan, M., Shin, N., Thompson, C. A., Winter, S. and Zaslavsky, I. (2016). Build it, but will they come? A geoscience cyberinfrastructure baseline analysis. Data Science Journal, 15(8), London: Ubiquity Press. 1–14. DOI: http://dx.doi.org/10.5334/dsj-2016-008

Dempwolf, C. S., Auer, J., & D'Ippolito, M. (2014). Innovation accelerators: Defining characteristics among startup assistance organizations. Small Business Administration, 1-44.

Desouza, K. C., & Jacob, B. (2017). Big data in the public sector: Lessons for practitioners and scholars. Administration & Society, 49(7), 1043-1064.

Dimaggio, P. J., & Powell, W. W. (1983). The iron cage revisited: Institutional isomorphism and collective rationality in organizational fields. American Sociological Review, 48(2), 147–160.

Doriese, W. B., Abbamonte, P., Alpert, B. K., Bennett, D. A., Denison, E. V., Fang, Y., … & Hays-Wehle, J. P. (2017). A practical superconducting-microcalorimeter X-ray spectrometer for beamline and laboratory science. Review of Scientific Instruments, 88(5), 053108.

Doyle, A. E., Young, E. D., Klein, B., Zuckerman, B., & Schlichting, H. E. (2019). Oxygen fugacities of extrasolar rocks: Evidence for an Earth-like geochemistry of exoplanets. Science, 366(6463), 356-359.

Dravins, D. (2008). Photonic astronomy and quantum optics. In High time resolution astrophysics (pp. 95-132). Springer, Dordrecht.

Dwyer, S Squeezing quantum noise: Physics Today 67, 11, 72 (2014); doi: 10.1063/PT.3.2596
Eason, G. (2004). Bridging the gap from educator to scientist. Journal of College Science Teaching, 34(3), 26-29.
Edelson, D. C. (2003). Realizing authentic science learning through the adaptation of scientific practice. In B. J. Fraser, & K. G. Tobin (Eds.), International handbook of science education (pp. 317-331). Dordrecht: Kluwer Academic Publishers.
Elliott, J. R., Walters, R. J., & Wright, T. J. (2016). The role of space-based observation in understanding and responding to active tectonics and Earthquakes. Nature communications, 7, 13844.
Erkmen, B. I. (2012). Computational ghost imaging for remote sensing. JOSA A, 29(5), 782-789.
Erkmen, B. I., Shapiro, J. H., & Schwab, K. (2012). Quantum Communication, Sensing and Measurement in Space.
Fisher, C. G., & Fisher, K. E. (2004). Encouraging upward thinking with planetary imagery: science as motivator and enrichment for the adult basic skills learner. In AGU Fall Meeting Abstracts.
Frogel, J. A. (2010). Astronomy's greatest hits: The 100 most cited papers in each year of the first decade of the 21st century (2000–2009). Publications of the Astronomical Society of the Pacific, 122(896), 1214.
Genovese, M. (2016). Real applications of quantum imaging. Journal of Optics, 18(7), 073002.
Gil, Y., Pierce, S. A., Babaie, H., Banerjee, A., Borne, K., Bust, G., Cheatham, M., Ebert-Uphoff, I., Gomes, C., Hill, M. and Horel, J. (2018). Intelligent systems for geosciences: an essential research agenda. Communications of the ACM, 62(1). New York: Association of Computing Machinery. 76-84.
Gottesman, D., Jennewein, T., & Croke, S. (2012). Longer-baseline telescopes using quantum repeaters. Physical review letters, 109(7), 070503.
Grebe, M. E., Pacheco, S., & Segura, J. (1972). Cosmovisión mapuche. Cuadernos de la realidad nacional, 14, 46-73.
Green, T. J., Karvounarakis, G., Taylor, N. E., Biton, O., Ives, Z. G., & Tannen, V. (2007). ORCHESTRA: facilitating collaborative data sharing.
Greisen, E. W. (2003). FITS: A remarkable achievement in information exchange. In Information Handling in Astronomy-Historical Vistas (pp. 71-87). Springer, Dordrecht.
Hallonsten, O. (2016). Big Science Transformed. Basingstoke: Palgrave Macmillan.
Hamacher, D. (2014) Stories from the sky: astronomy in Indigenous knowledge. The Conversation (weblog) https://theconversation.com/stories-from-the-sky-astronomy-in-indigenous-knowledge-33140 . Accessed 11/27/19
Hardy, N. D., & Shapiro, J. H. (2013). Computational ghost imaging versus imaging laser radar for three-dimensional imaging. Physical Review A, 87(2), 023820.
Heatherly, S. A., Rosen, R., McLaughlin, M., & Lorimer, D. (2011, January). Current Results and Future Directions of the Pulsar Search Collaboratory. In Bulletin of the American Astronomical Society (Vol. 43, p. 14506).
Hedley, R. A. (2000). Convergence in natural, social, and technical systems: A critique. Current Science, 592-601.
Heke, I., Rees, D., Swinburn, B., Waititi, R. T., & Stewart, A. (2019). Systems Thinking and indigenous systems: native contributions to obesity prevention. AlterNative: An International Journal of Indigenous Peoples, 15(1), 22-30.
Helbing, D., & Balietti, S. (2011). How to create an innovation accelerator. The European Physical Journal Special Topics, 195(1), 101-136.
Henneken, E. (2015). Unlocking and sharing data in astronomy. Bulletin of the Association for Information Science and Technology, 41(4), 40-43.
Herrold, A. (2019). Personal Interview. May 23, 2019.
Hey, Tony, Stewart Tansley, Kristin Tolle (Eds.). (2009). Jim Gray on eScience: A Transformed Scientific Method. Redmond, WA: Microsoft Research.
Heylighen, F. and Joslyn, C. (1992): "What is Systems Theory?". in: F. Heylighen, C. Joslyn and V. Turchin (editors): Principia Cybernetica Web (Principia Cybernetica, B
Horowitz, C., Shameer, K., Gabrilove, J., Atreja, A., Shepard, P., Goytia, C., ... & Galvez, M. (2017). Accelerators: sparking innovation and transdisciplinary team science in disparities research. International journal of environmental research and public health, 14(3), 225.
Hosten, O., Engelsen, N. J., Krishnakumar, R., & Kasevich, M. A. (2016). Measurement noise 100 times lower than the quantum-projection limit using entangled atoms. Nature, 529(7587), 505.
Hughes, S. (2007). Opportunities for the transfer of astronomical technology to medicine. Australasian Physics & Engineering Sciences in Medicine, 30(4), 292.
Jinha, Arif. (2010). Article 50 million: An estimate of the number of scholarly articles in existence. Learned Publishing. 23. 258-263. 10.1087/20100308.

Jorden, P., Jerram, P., Jordan, D., Pratlong, J., & Robbins, M. (2016, August). e2v CCD and CMOS sensors and systems designed for astronomical applications. In High Energy, Optical, and Infrared Detectors for Astronomy VII (Vol. 9915, p. 991504). International Society for Optics and Photonics.

Kastens, K. A., Manduca, C. A., Cervato, C., Frodeman, R., Goodwin, C., Liben, L. S., & Titus, S. (2009). How geoscientists think and learn. Eos, Transactions American Geophysical Union, 90(31). Washington, DC: American Geophysical Union. 265-266.

Kelly, S. (2018). Articulating Indigenous Rights Amidst Territorial Fragmentation: Small Hydropower Conflicts in the Puelwillimapu, Southern Chile (Doctoral dissertation, The University of Arizona).

Kenter, A., Kraft, R., Gauron, T., & Amato, S. (2017, August). Advancing the technology of monolithic CMOS detectors for use as x-ray imaging spectrometers. In UV, X-Ray, and Gamma-Ray Space Instrumentation for Astronomy XX (Vol. 10397, p. 1039703). International Society for Optics and Photonics.

Ketter, W., Peters, M., Collins, J., & Gupta, A. (2015). Competitive benchmarking: an IS research approach to address wicked problems with big data and analytics.

Konczal, J. (2012). Evaluating the effects of accelerators? Not so fast. Forbes 2012 [cited Aug 8, 2012] from http://www.forbes.com/sites/kauffman/2012/08/08/evaluating-the-effects-of-accelerators-not-so-fast/.

Konrad, T. & Forbes, A (2019) Quantum mechanics and classical light, Contemporary Physics, DOI: 10.1080/00107514.2019.1580433

Kouveliotou, C., Agol, E., Batalha, N., Bean, J., Bentz, M., Cornish, N... & Hartmann, D. (2014). Enduring quests-daring visions (NASA astrophysics in the next three decades). arXiv preprint arXiv:1401.3741.

Kozlow, M. J., & Nay, M. A. (1976). An approach to measuring scientific attitudes. Science Education, 60(2), 147-172.

Kreidberg, L., Bean, J. L., Désert, J. M., Benneke, B., Deming, D., Stevenson, K. B., ... & Homeier, D. (2014). Clouds in the atmosphere of the super-Earth exoplanet GJ 1214b. Nature, 505(7481), 69.

Lemos, G. B., Borish, V., Cole, G. D., Ramelow, S., Lapkiewicz, R., & Zeilinger, A. (2014). Quantum imaging with undetected photons. Nature, 512(7515), 409.

Leopold, A. (2007). Historical ecology: integrated thinking at multiple temporal and spatial scales. The world system and the earth system: global socioenvironmental change and sustainability since the Neolithic.

Lovelock, J.E.; Giffin, C.E. (1969). "Planetary Atmospheres: Compositional and other changes associated with the presence of Life". Advances in the Astronautical Sciences. 25: 179–193. ISBN 978-0-87703-028-7.

Lovelock, J.E.; Margulis, L. (1974). "Atmospheric homeostasis by and for the biosphere: the Gaia hypothesis". Tellus. Series A. Stockholm: International Meteorological Institute. 26 (1–2): 2–10.

Macauley, M. K. (2006). The value of information: Measuring the contribution of space-derived Earth science data to resource management. Space Policy, 22(4), 274-282.

Manduca, C.A., and Kastens, K.A., (2012). Geoscience and geoscientists: Uniquely equipped to study Earth, in Kastens, K.A., and Manduca, C.A., eds., Earth and Mind II: A Synthesis of Research on Thinking and Learning in the Geosciences: Geological Society of America Special Paper 486, Boulder, CO: Geological Society of America. 1–12. doi:10.1130/2012.2486(01).

Mennesson, B., Gaudi, S., Seager, S., Cahoy, K., Domagal-Goldman, S., Feinberg, L., ... & Tamura, M. (2016, August). The Habitable Exoplanet (HabEx) Imaging Mission: preliminary science drivers and technical requirements. In Space Telescopes and Instrumentation 2016: Optical, Infrared, and Millimeter Wave (Vol. 9904, p. 99040L). International Society for Optics and Photonics.

Miller, M., & Mansilla, V. B. (2004). Thinking across perspectives and disciplines. Boston, MA: Harvard Graduate School of Education.

Moreau, P. A., Sabines-Chesterking, J., Whittaker, R., Joshi, S. K., Birchall, P. M., McMillan, A., … & Matthews, J. C. (2017). Demonstrating an absolute quantum advantage in direct absorption measurement. , Scientific reports, 7(1), 6256. New York. : Springer Nature

Morgan, R., Warfield, K., Kuan, G., Stahl, H. P. P., Mennesson, B., Balasubramanian, B., … & Shaklan, S. (2018). Technology maturity for the habitable-zone exoplanet imaging observatory (HabEx) concept. In Space Telescopes and Instrumentation 2018: Optical, Infrared, and Millimeter Wave (Vol. 10698, p. 106980S). International Society for Optics and Photonics.

Mount, B. J., Preece, R. M., Hoff, M. D., Edberg, T. K., Rossiter, P., Kaboth, A., …. & Pagac, A. (2017). LUX-ZEPLIN (LZ) technical design report (No. FERMILAB-TM-2653-AE-E-PPD).

National Research Council. (2014). Convergence: facilitating transdisciplinary integration of life sciences, physical sciences, engineering, and beyond. National Academies Press.

National Science Board. (2012). Science and engineering indicators 2012. (No. NSB 12-01). Arlington, VA: National Science Foundation.

NSF (2015). Final Report on the 2015 National Science Foundation Workshop Report from the Workshop on Information and Intelligent Systems for Geosciences." Yolanda Gil and Suzanne A. Pierce (Eds). Washington, DC: National Science Foundation.

Omenn, G. S. (2006). Grand challenges and great opportunities in science, technology, and public policy. Science, 314(5806), 1696-1704.

Palmer, B. (2011). Sharing the Universe — Why Astronomy Outreach Matters, Astronomy Magazine, August 2011, http://www.astronomy.com/community/astronomyfoundation/2011/08/sharing-the-universe---why-astronomy-outreach-matters. Accessed 12/1/19.

Parashar, M. (2009). Transformation of Science Through Cyberinfrastructure: Keynote Address. Open Grid Forum Banff, Alberta, Canada, October 14. http://www.ogf.org/OGF27/materials/1816/parashar-summit-09.pdf (accessed 12/1/2019).

Patroumpas, K., Giannopoulos, G., & Athanasiou, S. (2014). Towards geospatial semantic data management: strengths, weaknesses, and challenges ahead. In Proceedings of the 22nd ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems. New York: Association of Computing Machinery. 301-310.

Pauwels, C., Clarysse, B., Wright, M., & Van Hove, J. (2016). Understanding a new generation incubation model: The accelerator. Technovation, 50, 13-24.

Rahm, J., Miller, H. C., Hartley, L., & Moore, J. C. (2003). The value of an emergent notion of authenticity: Examples from two student/teacher-scientist partnership programs. Journal of Research in Science Teaching, 40(8), 737-756.

Reed, S. (2011). Is there an astronomer in the house? Science. Vol. 331, Issue 6018, pp. 696-697

Ribes, D. and Bowker, G. C. (2008). Organizing for multidisciplinary collaboration: The case of the geosciences network, in G.M. Olson, A. Zimmerman and Nathan Bos (eds), Scientific Collaboration on the Internet (Cambridge: MIT Press): 311-30.

Robinson, G. (2004). Replicating a successful authentic science research program: An interview with dr. robert pavlica. Journal of Secondary Gifted Education, 15(4), 148-154.

Romm, N. R. (2015). Reviewing the transformative paradigm: A critical systemic and relational (Indigenous) lens. Systemic Practice and Action Research, 28(5), 411-427.

Rosenberg, M., Baldon, G., Russo, P., & Christensen, L. L. (2014). Astronomy in everyday life. Communicating Astronomy to the Public Journal, 14, 30-36.

russels), URL: http://pespmc1.vub.ac.be/SYSTHEOR.html. Accessed, 11/26/19.

Sanchez, D. (2010). El concepto de la cosmovision. Seminario Teologico Centroamericano. KAIRÓS.

Schnase, J. L., Lee, T. J., Mattmann, C. A., Lynnes, C. S., Cinquini, L., Ramirez, P. M., Hart, A.F., Williams, D.N., Waliser, D., Rinsland, P. and Webster, W. P. (2016). Big data challenges in climate science: Improving the next-generation cyberinfrastructure. Washington, DC: National Aeronautics and Space Administration. https://ntrs.nasa.gov/archive/nasa/casi.ntrs.nasa.gov/20160002946.pdf. Accessed 3/3/19.

Sellars, S., Nguyen, P., Chu, W., Gao, X., Hsu, K. L., & Sorooshian, S. (2013). Computational Earth science: Big data transformed into insight. Eos, Transactions American Geophysical Union, 94(32), 277-278.

Siwen Bi, Xuling Lin, Song Yang, Zhiqiang Wu, "Technology study of quantum remote sensing imaging," Proc. SPIE 9755, Quantum Sensing and Nano Electronics and Photonics XIII, 97552J (13 February 2016); doi: 10.1117/12.2212379

Siwen Bi, Youyong Liu, Zhiyong Lu, Wenqi Ma, Na Li, "Infrared quantum remote sensing detection imaging technology that breakthrough quantum noise limit," Proc. SPIE 10765, Infrared Remote Sensing and Instrumentation XXVI, 1076506 (18 September 2018); doi: 10.1117/12.2318262

Spuck, T. S. (2017). What do Astronomers Do: A Survey of US Astronomers' Attitudes, Tools and Techniques, and Social Interactions Engaged in through their Practice of Science. West Virginia University.

Stephens, Z. D., Lee, S. Y., Faghri, F., Campbell, R. H., Zhai, C., Efron, M. J., Iyer, R., Schatz, M.C., Sinha, S. & Robinson, G. E. (2015). Big data: astronomical or genomical?. PLoS biology, 13(7), e1002195.

Tompkins, J., Cain, S., & Becker, D. (2019). Near Earth space object detection using parallax as multi-hypothesis test criterion. Optics express, 27(4), 5403-5419.

Tsang, M., Nair, R., & Lu, X. M. (2016). Quantum theory of superresolution for two incoherent optical point sources. Physical Review X, 6(3), 031033.

Two-Hawks, J. (2014) Of Earth and Sky. Weblog entry https://www.johntwohawks.com/blog/of-earth-and-sky. Accessed 11/27/19.

Unternährer, M., Bessire, B., Gasparini, L., Perenzoni, M., & Stefanov, A. (2018). Super-resolution quantum imaging at the Heisenberg limit. Optica, 5(9), 1150-1154.

Uzzo, S. (1994). Light and Imaging in Extreme Environments. In Growth and Environment: Challenging Extreme Frontiers Proceedings, 2nd International Design for Extreme Environments Assembly. Montreal, PQ: McGill University, Centre for Northern Studies.

van Harmelen, F., Kampis, G., Börner, K., Van den Besselaar, P., Schultes, E., Goble, C., ... & Hayes, C. (2012). Theoretical and technological building blocks for an innovation accelerator. The European Physical Journal Special Topics, 214(1), 183-214.

Ware, M., & Mabe, M. (2015). The STM report: An overview of scientific and scholarly journal publishing. The Hague: International Association of Scientific, Technical and Medical Publishers.

Wenger, E., Trayner, B., & De Laat, M. (2011). Promoting and assessing value creation in communities and networks: A conceptual framework. The Open Universiteit Nederland.

Wenger, E., Trayner, B., & De Laat, M. (2011). Promoting and assessing value creation in communities and networks: A conceptual framework.

Wilder, R. L. (1967). The role of intuition. Science, 156(3775), 605-610.

Young, S. M., Sarovar, M., & Léonard, F. (2018). Fundamental limits to single-photon detection determined by quantum coherence and backaction. Physical Review A, 97(3), 033836.

Yue, L., Shen, H., Li, J., Yuan, Q., Zhang, H., & Zhang, L. (2016). Image super-resolution: The techniques, applications, and future. Signal Processing, 128, 389-408.

Zhang, Y. and Zhao, Y., 2015. Astronomy in the Big Data Era. Data Science Journal, 14, p.11. DOI: http://doi.org/10.5334/dsj-2015-011.

Zhang, Y., Wu, Y., Wang, X., Fossum, E. R., Kumar, R., Liu, J., … & Yu, S. Q. (2017). Non-avalanche single photon detection without carrier transit-time delay through quantum capacitive coupling. Optics express, 25(22), 26508-26518.

Ziemer, N. (2009). Collaborative networks as innovation accelerators. In IIE Annual Conference. Proceedings (p. 391). Institute of Industrial and Systems Engineers (IISE).

Zion, M., Slezak, M., Shapira, D., Link, E., Bashan, N., Brumer, M., . . . Valanides, N. (2004). Dynamic, open inquiry in biology learning. Science Education, 88(5), 728-753.

Appendix A: Biographies of Participants

RICHARD BOONE

Dr. Richard Boone is Professor of Ecology in the Departments of Forestry & Wildland Resources and Biological Sciences at Humboldt State University (HSU). Prior to joining HSU, Boone was Professor of Ecosystem Ecology in the Department of Biology and Wildlife and the Institute of Arctic Biology at the University of Alaska Fairbanks. He received his B.A. in Biology from Oberlin College, M.S. in Forest Ecology from Oregon State University, and Ph.D. in Forestry from the University of Massachusetts, Amherst. Boone was appointed as an Aldo Leopold Leadership Fellow in 2001 and a National Academies Education Fellow in the Life Sciences in 2007. He served as Program Director of the Integrative Graduate Education and Research Traineeship (IGERT) Program September 2011-May 2015; Program Director of the National Science Foundation Research Traineeship Program May 2014-May 2015, and Acting Deputy Director, Division of Graduate Education July 2014-May 2015 at NSF. Boone's professional passions are STEM education, ecosystem ecology, and climate change science.

YASMIN CATRICHEO

Yasmin Catricheo is currently working at the Office of Education and Public Engagement at Associated Universities Inc. She is a physics educator from Chile. Of Mapuche origin, she is passionate about the teaching of science and more recently has focused in the area of astronomy. In her professional training program, she has participated in courses taught by astronomy faculty at the University of Concepción, and has earned a Master's degree in Education from the University of Bío-Bío. Yasmín is also a member of the indigenous group "Mapu Trafun", and she works closely with the Mapuche community to recover the culture and communicate the message of the Mapuche Worldview. In 2018 was selected to enjoy the Astronomy in Chile Educator Ambassador Program (ACEAP) founded by FSF.

CATHERINE CRAMER

Catherine Cramer works at the intersection of data-driven science and learning, specifically as it pertains to the understanding of complexity and its application to data and network sciences, with a focus on underrepresented communities. For over 20 years she has developed tools and programs for the teaching and learning of complex network and data science, centering on identifying, creating, sustaining and growing productive and innovative collaborations and partnerships between research, industry and academia. She worked with the centers for Ocean Science Education Excellence (COSEE) and the Ocean founders of the Network Literacy and Network Science in Education movements. She remains active in both, most recently organizing the 8th annual Network Science in Education symposium at the University of Vermont as part of the 2019 International School and Conference on Network Science, and is on the Board of the Network Science Society. She is co-editor and co-author of the Springer volume Network Science in Education, published in October 2018. She is currently co-leading the data literacy efforts at the Northeast Big Data Innovation Hub, located at the Columbia University Data Science Institute, as well as a Social Network Analysis of the Hub itself. She is also Principal and Director of the Woods Hole Institute.

ROBERT R. DOWNS

Dr. Robert R. Downs is a senior staff associate officer of research and serves as the senior digital archivist and acting head of cyberinfrastructure and informatics research and development at CIESIN, the Center for International Earth Science Information Network, a research and data center of the Earth Institute of Columbia University. His research focuses on the development, management, use, evaluation, and value of information systems, focusing on data facilities. He is an elected member of the CoreTrustSeal Standards and Certification Board, a member of the Editorial Board of the CODATA Data Science Journal, and a co-chair of the Research Data Alliance (RDA) Repository Platforms for Research Data (RPRD) Interest Group. Dr. Downs has been designated as a Senior Member of the Association for Computing Machinery (ACM) and is a member of the American Geophysical Union (AGU) and the International Association for Social Sciences Information Services and Technology (IASSIST).

DANIEL R. FUKA

Daniel is a cross-disciplinary scientist working on the Big Island of Hawai'i for Virginia Tech University. He is passionate about mentoring researchers young and old about the scientific opportunities that exist when bridging many disparate sciences. Daniel is an active member of the EarthCube and ESIP communities, a key collaborator on the EC Architecture BCube brokering project, and is a Co-PI on the NSF BALTO data brokering project. He has been active in the NSF RCN on Intelligent Systems for Geosciences (IS-GEO) creating collaborations to better our understanding of Earth systems through applications of intelligent information systems. Outside of his research scientist day job, Daniel enjoys the challenges and rewards of teaching and mentoring students who may start their college education under-prepared, and he has been using data architectures in unique ways to bring less prepared students up to speed quickly while not compromising the education of their better-prepared peers.

JOSÉ A. GURIDI

José currently serves as an advisor to the ministry of science, technology, knowledge and innovation as part of the Future team, where he leads the development of the artificial intelligence policy of Chile. Previously, he worked at the Ministry of Economy where he participated in the creation of the Data Observatory. He is an industrial engineer and Msc. in industrial and systems engineering from Pontificia Universidad Católica de Chile.

JIM HAMMERMAN

Jim Hammerman, Ed.D., directs the STEM Education Evaluation Center at TERC, in Cambridge, MA. He has over 35 years' experience teaching, developing, researching, and evaluating STEM education innovations in both formal and informal education settings. Dr. Hammerman has been PI, Co-PI, senior researcher and evaluator on NSF-funded projects focused on math, science, computing, and data and statistics, including *Educating About Statistical Issues in Large Scientific Data Sets*, *Innovators Developing Accessible Tools for Astronomy (IDATA)*, the *Visualizing Statistical Relationships (VISOR)* project, *Levels of Conceptual Understanding in Statistics (LOCUS)*, the Adler Planetarium's *Leveraging Citizen for Informal STEM Learning* project focusing on galaxy classification, among others. He served on Advisory Boards for the Oceans of Data project, and the Gulf of Maine Research Institute's Strategic Planning Education Committee. He has an interest in broad and equitable access to educational innovations, and inquiry oriented data analysis and exploration software.

BRYAN HEIDORN

Bryan Heidorn arrived at the University of Arizona 10 years ago and was the founding director of the School of Information. Prior to that he served two years as a program director in the NSF Division of Biological Infrastructure. He was on the faculty of the University of Illinois for 10 years with the final rank of Associate Professor. Bryan serves on the Organization for Tropical Studies Board of Directors and various other committees. He served on the Board of Trustees for the JRS Biodiversity Foundation for seven years as a chair of the grants committee, vice president and president. JRS has an activist board that fosters research and infrastructure for the dissemination of biodiversity information particularly in Africa. His research interests include information retrieval, text processing and digital libraries use in context. His work includes a redefinition of the book into electronic format with higher functionality for different user groups. His primary area of application is biological informatics including scientific communication, automatic markup of unstructured text for enhanced information retrieval, information extraction and information structuring from a variety of biological information resources ranging from images of natural history museum specimens to natural language description of location and taxonomic literature.

ELIZABETH JOYNER

Elizabeth Joyner serves as Project Lead for My NASA Data at NASA Langley Research Center's Science Directorate. Elizabeth taught for over 10 years in both public and private institutions, as well as worked in informal settings such as the American Geosciences Institute, Virginia and South Carolina Space Grants, SC Sea Grant, NOAA/NSF's Center for Ocean Sciences Education Excellence (COSEE-SE) and others. Her focus is leveraging NASA earth systems data resources and bridging these data to support NGSS phenomena-based teaching and learning.

RANDY KOCHEVAR

Randy Kochevar trained as a deep-sea biologist at the University of California, Santa Barbara, studying the physiology of hydrothermal vent and hydrocarbon seep animals. The discovery of similar animal communities in the Monterey Submarine Canyon brought Randy up the coast, initially to the Monterey Bay Aquarium Research Institute (MBARI), and then to the Monterey Bay Aquarium, where he was involved in exhibit development, website development and print publishing, and media relations. From 2008-2015 Randy worked at Stanford University's Hopkins Marine Station, collaborating with Dr. Barbara Block on electronic tagging of marine apex predators including tunas, billfish, and sharks. In 2015 Randy was invited to take over as director of the Oceans of Data Institute at the Education Development Center (EDC), where he oversees a portfolio of projects to build data literacy skills in K-16 students, and to help build pathways for data science and analytics careers.

MARGARET MOONEY

Margaret Mooney is the Education and Public Outreach lead at CIMSS. As a former National Weather Service meteorologist, Margaret has ample experience working with formal and informal audiences, leveraging early career experience with a degree in public policy to promote weather and climate education as avenues towards stewardship and sustainability. Margaret is active in several education communities, including NOAA Climate Stewards, NASA Initiatives in Climate Education, the GOES-R Education Proving Ground, the Earth Science Information Partners, and the Wisconsin Initiative for Climate Change Impacts.

KATIE NAUM

Katie Naum is the operations manager for the National Science Foundation's Northeast Big Data Innovation Hub, and a science writer for the National Center for Supercomputing Applications. Katie also consults on writing, editing, and grant writing projects, specializing in topics related to science & technology as well as creative writing. She holds a degree in sustainable development from Columbia University.

SHELLEY OLDS

Olds has been leading curriculum development projects, developing science education products, and teaching science and technology to educators for close to 20 years. She has a Masters of Education in instructional systems development, a B.S. in Earth Science / Geophysics, and is currently a PhD Candidate investigating human dimensions of natural hazards. Currently Shelley is a Science Education Specialist for UNAVCO's Education and Community Engagement program, creating data-rich, place-based Earth science curricula using high-precision GPS and other data types for secondary-level and undergraduate courses, designing science-focused museum exhibits, evaluating educational resources, and leading the professional development Earth science programs for K-12. In her spare time, she likes to photograph landscapes and wildlife and hiking in the mountain of Colorado.

LEIGH PEAKE

Leigh Peake is Chief Education Officer at the Gulf of Maine Research Institute (GMRI) a marine research laboratory based in Portland Maine's bustling working waterfront. GMRI integrates fundamental marine research, data modeling, and forecasting, with substantive work across the seafood supply chain and with learners of all ages. GMRI's extensive education programming engages more than 70% of the state's middle-school student cohort annually as well as more than 200 middle-school educators. Our programs use a Preparation for Future Learning framework to deliberately link informal and formal learning experiences. Our work focuses on leveraging local and global data sets to enable investigations of changing ecosystems across Maine as a context for supporting student work with data and modeling.

BECKY REID

A lifelong educator, Becky Reid holds California teaching credentials in both biology and earth science, along with an M.S. in Education. She has taught science and basic computer coding in the high school classroom and as an informal educator with youth serving organizations. Becky has been involved with the Federation of Earth Science Information Partners (ESIP) since the summer of 2009 when she was invited to attend the Summer meeting in Santa Barbara. She has been volunteering with the ESIP Education Committee in various capacities ever since, serving as the Education Committee chair in 2013 and 2019. With other members of the committee, she has been involved in numerous professional development workshops for educators at ESIP meetings. She currently teaches at Cuesta Community College in San Luis Obispo, CA.

TIM SPUCK

Tim is Director of Education & Public Engagement at AUI where he manages the development and implementation of innovative STEM Education efforts. He currently serves as PI on Network for Earth-space Research Education and Innovation with Data (NEREID). Innovators Developing Accessible Tools for Astronomy (IDATA), the Astronomy in Chile Educator Ambassadors Program (ACEAP), and Big Astronomy in Chile through Dome+. Tim also leads ongoing collaborations with the National Society of Black Physicists, Organization of American States, Cortes-Solari Foundation, International Astronomical Union, and others. Prior to his role with AUI he taught astronomy and earth sciences at the high school and university levels and served as a K–12 Science Coordinator. Tim has received the Albert Einstein Distinguished Educator Fellowship, American Institute of Aeronautics & Astronautics Educator Achievement Award, Tandy Technology Scholars Award, and the Pennsylvania Christa McAuliffe Fellowship. He holds a doctorate in Curriculum & Instruction from West Virginia University, and a master's degree in Science Education from Clarion University of PA. Tim served as lead editor for the Best Practices in STEM Education: Innovative Approaches from Einstein Fellow Alumni, 2nd edition; the 1st edition was awarded the 2014 Peter Lang Publishing Book of the Year.

GRETCHEN STAHLMAN

Gretchen Stahlman is an Assistant Professor of Library & Information Science at Rutgers University's School of Communication & Information. Gretchen's research areas include: long-term management of science data; scholarly communication paradigms; and the social, technical and institutional infrastructures that mediate data curation. Most recently, Gretchen's research has focused on identifying and characterizing data associated with journal articles published over the past several decades in astronomy. Gretchen's Ph.D. work was conducted at the University of Arizona's School of Information. She holds a Master of Science degree in Library Science from Clarion University of Pennsylvania, and she previously worked as a documentation specialist for the Atacama Large Millimeter/submillimeter Array telescope project.

ED SUMMERS

Ed Summers is a blind software engineer and an accessibility specialist. He has a B.S. in Computer Science and 25 years of professional experience as a software developer and a development manager. Ed's personal mission is to enable people with disabilities to realize their full potential in the classroom and the 21st century knowledge economy. He fulfills that mission as a leader in the software industry and disability-related not-for-profit organizations. Ed is Director of Accessibility at SAS, the market leader in analytics software and services. The SAS accessibility team enables people of all abilities to access the power of analytics. Ed also serves on the Advisory Council for the North Carolina Museum of Natural Sciences and the North Carolina Board of Education Advisory Council for the Governor Morehead School for the Blind.

LAURA TROUILLE

Laura Trouille is Vice President of Citizen Science at the Adler Planetarium, Research Associate at Northwestern University, and Co-PI for Zooniverse, leading the Adler's 7-person Zooniverse web development team, three postdoctoral research fellows, and the Adler's Teen Programs team. The Adler-Zooniverse team works closely with Oxford and the University of Minnesota on strategic planning as well as building and maintaining the Zooniverse platform together. While earning her Ph.D. in astrophysics in 2010, she also completed the Center for the Integration of Research, Teaching and Learning's (CIRTL) Delta certificate for STEM education research. As a postdoctoral research fellow within the Center for Interdisciplinary Exploration and Research in Astrophysics (CIERA) at Northwestern University, she continued her supermassive black hole research as well as co-led the Computational Thinking in STEM project, bringing computational thinking and modeling curricular materials to high school science and math teachers.

PETER D. TUDDENHAM

Peter D. Tuddenham is working for the future of 12 grandchildren. He is President of the College of Exploration (TCOE), a not-for-profit global online learning community he co-founded with Tina Bishop in Virginia USA in 1991. He is Past-President of the International Society for the Systems Sciences and a trustee of the American Society for Cybernetics. He is a co-founder and director of the European Marine Science Educators Association based in Belgium. He is Managing Director of CoExploration, Limited in Dorset, England. As President of TCOE he leads several initiatives to create and support organizations, networks and learning experiences for all ages that build whole systems approaches and appreciations to issues such as global ocean literacy and global systems literacy. TCOE conducts program and project evaluations using developmental approaches. Current activities include encouraging transdisciplinary approaches for business, government, education and the broader civil society to address complex environmental and economic issues. The focus is on building collaborative actions to encourage innovation, change and transformations. He is a Member of the Institution of Royal Engineers. He sails a 22-foot BayRaider Expedition in Barnegat Bay, New Jersey.

STEPHEN UZZO

As Chief Scientist for the New York Hall of Science, Stephen Uzzo develops and leads large-scale initiatives to research and integrate cutting edge science into teaching and learning. He currently develops initiatives to build communities of practice in complexity, data-driven science and engineering, and improve science, technology, engineering and math (STEM) literacy of the public. His background includes over 20 years experience in the research of connected systems and teaching and learning in STEM; and prior to that, 10 years in video and computer graphics systems engineering. Dr. Uzzo's research interests include the coupling of complex human and natural systems, complex networks, smart cities, and the impact of big data on communities of need. He holds a terminal degree in network theory and environmental studies and serves on a number of institutional and advisory boards related to his interests. His work also includes developing, studying and teaching graduate programs in STEM learning. Having never lived very far from the ocean in New York and California, Dr. Uzzo has been a lifelong advocate for marine conservation.

TIFFANY STONE WOLBRECHT

Tiffany is the Planetarium Lecturer at YSU's Ward Beecher Planetarium where she coordinates and leads programming, impacting over 20,000 members of her local community annually. In 2018, Tiffany was selected as one of 8 astronomy educators across the US by the Astronomy in Chile Educator's Ambassadors Program, or ACEAP, to tour NSF-supported astronomy facilities in Chile and receive extensive training about the observatories and their telescopes, instruments, science, and data. From this experience, Tiffany developed a live planetarium show kit Eyes on Chile Skies, distributing it freely in the planetarium community. Tiffany is also serves as web content and show distribution lead on an NSF Advancing Informal STEM Learning (AISL) project developing a bilingual planetarium show entitled Big Astronomy: People, Places, Discoveries about observatories in Chile, the diverse careers needed to run them, and the incredible discoveries they uncover, along with a robust collection of live events with observatory staff, educational activities, and ongoing science experiments hosted on a web portal. Tiffany also serves on the American Astronomical Society's Committee for the Status of Women in Astronomy (CSWA) as well as the Great Lakes Planetarium Association (GLPA)'s Executive Committee. Tiffany obtained her Master's of Education in Secondary Instruction at Edinboro University of Pennsylvania after earning her Bachelor's of Science in Physics at Penn State University, The Behrend College.

ELENA YULAEVA

Dr. Yulaeva is a climate researcher and an educator. She received her Ph.D. in atmospheric sciences from the University of Washington, and has been a researcher at UCSD. For the last 20 years she has been engaged in climate changes analysis and prediction and in analysis of climate change impacts on global communities. Elena is the Executive Director of Community Commons, a San Diego based non-profit organization that develops innovative solutions to support informed, healthy, equitable, and sustainable communities. In this capacity, she has organized and led educational STEM programs at several levels, from K-12 to university students. She is the founder and co-director of the award-winning Global Forest Link Project (globalforestlink.com) that engages youth in the online collaborative analysis of local forest health and helps them explore impacts of environmental change.

ILYA ZASLAVSKY

Ilya Zaslavsky is director of Spatial Information Systems Lab at the San Diego Supercomputer Center, University of California San Diego. His research focuses on distributed information management systems and spatial and temporal data integration. Zaslavsky received his Ph.D. from the University of Washington (1995), and earlier a Ph.D. equivalent from the Russian Academy of Sciences (1990). He has been leading design and technical development in several large cyberinfrastructure projects supported by the U.S. National Science Foundation.

Appendix B: Workshop Agenda

Wednesday Nov 20: Arrival

5:00 Arrival at GBO, registration and pre-workshop surveys *Residence Hall*
5:30 Dinner *Drake Lounge*
7:00 Reception/Orientation/Spotlight sign-up *Drake Lounge*
8:00 Drake Equation Challenge kick-off

Thursday Nov 21/Day One: Terrain

8:00 Breakfast *Cafeteria*
9:00 Welcome and introductions *SkyLab, Science Center*
9:30 Why NEREID? Tim Spuck
9:45 Charge for the meeting Steve Uzzo
10:00 Agenda Review Catherine Cramer
10:15 Spotlights Round 1
- Daniel Fuka
- Elena Yulaeva
- Shelley Olds

10:50 Coffee Break
11:05 Spotlights Round 2
- Robert R. Downs
- Bryan Heidorn
- Randy Kochevar
- Ed Summers

12:00 Lunch *Cafeteria*
12:45 Tour of GBO *Zone 1 and Science Center*
2:30 Spotlights Round 3 (via teleconference) *SkyLab*
- Laura Trouille
- Margaret Mooney
- Leigh Peake
- Elizabeth Joyner

3:15 Coffee Break
3:30 Spotlights Round 4
- Kartik Sheth (via teleconference)
- Sue Ann Heatherly
- Tiffany Wolbrecht
- Jim Hammerman
- Gretchen Stahlman

4:00 Building the current landscape (binning exercise)
5:00 Dinner *Cafeteria*
6:30 Evening activity: radio telescope demonstration *40-foot*
8:00 Social *Drake lounge*

Friday Nov 22/Day Two: Blast Off

8:00 Breakfast *Cafeteria*
9:00 Welcome back, charge for the day *SkyLab*
9:10 Talk: Chile Data Observatory Jose Guridi
9:30 Exploring online-accessible astronomy databases Tim Spuck
9:50 Exploring SuAVE/Data Discovery Studio Ilya Zaslavsky
10:45 Coffee Break
11:00 Earth-space convergence needs brainstorm (three groups)
12:00 Lunch *Cafeteria*
1:00 Convergence brainstorm small groups (continued)
2:15 Small group report-outs
2:30 2^{nd}-best-idea write-ups - individual or group
3:15 *Coffee Break*
3:30 Report-outs
4:00 Expectations for report: timeline, process; wrap up
5:00 Dinner *Cafeteria*
7:00 Drake Challenge presentations *Drake Lounge*
8:00 Social *Drake Lounge*

Saturday Nov 23

Breakfast
Departure

Appendix C: Spotlight Talks Abstracts

Downs, Robert R.: ***Enabling Diverse Learners to Use Integrated Earth Science Data***

Scientific data products and services offered by the NASA Socioeconomic Data and Applications Center (SEDAC) provide diverse opportunities for learning and science. The data products and services available from SEDAC are open and free to use. These resources include web accessible data, services, and maps that are available for various platforms and devices and can be used by diverse audiences to improve their understanding of human interactions with the environment. The data and maps offered by SEDAC are applicable to the physical, social, and health sciences, as well as other disciplines. SEDAC resources are appropriate for both formal and informal learning opportunities and can be used in classrooms, libraries, laboratories, and homes. Many of the resources available from SEDAC are integrated data products and services that are produced by combining geospatial data and social science data. The hands-on data analysis and visualization capabilities of SEDAC products and services provide opportunities to explore and study the relationships between variables within and among represented locations, globally. SEDAC data have been used in over 4,000 published journal articles, conference papers, and proceedings that report on scientific studies within various disciplines, and can be used by experts and novices across many grade levels. In addition, SEDAC offers over 2,000 maps that also are free to use and many of these maps have been republished in textbooks and course materials that are used in K-12 classrooms and at the undergraduate and graduate levels, internationally.

Fuka, Daniel: ***Mediation of Heterogeneity: Managing HPC systems, data, and models across scientific domains to accelerate discovery***

The challenges of data interoperability, within and across scientific domains in the earth and space sciences, are well documented. Researchers often must learn distinct discovery and access methods for each data set and repository with which they interact, both for analytical scientific exploration, as well as for running domain-specific models. The cost of data access, preprocessing, and conversion—i.e., "data wrangling"—can exceed 80% of the effort required for a new scientific discovery, both in research as well as in educating the public. Therefore, enhancing interoperability is high on the list of the National Science Foundation's (NSF's) goals. While standardization, when possible, is useful in any scientific endeavor, standards do not exist until a critical mass works in, as well as agree to, the overarching goals and best practices of these cross-domain studies. As such, work has been underway for at least a decade on means to achieve data interoperability without seeking the impossible, i.e., universal agreement on standard data formats and encodings.

In single-domain communities with widely agreed standards, brokering is generally of little value because everyone talks the same data language. In contrast, endeavors that span multiple domains, especially where domain-specific methods and data formats are the norms (e.g., scientists speak different data languages), brokering offers important advantages to those who need cross-domain information. The problem of data interoperability can exhibit a great degree of complexity: if the goal is to pair M source-data types with N usage contexts, the brute-force approach is a problem of the order MxN (this arises if users are on their own when matching unfamiliar data sets to their preferred analysis tools). In contrast, a well-designed brokering approach (built using modular middleware that employs an appropriate data abstraction) can simplify the problem to order N+M, requiring a single software interface for each type of source data and each usage context, all mapped to a common data model. Recent research has shown that the time and effort savings to an end-user can be substantial, though automated brokering requires a high level of description for the objects being paired, requiring more upfront time to develop the data model.

Dr. Daniel Fuka brings to his spotlight a multi-decadal review of cross science data and model mediation, starting with a review of how successful model and data mediation can take bleeding edge shock and materials computational fluid dynamics research from DOE to accelerate health initiatives in the biomedical sciences; create novel weather risk products for the industries that perform electrical transmission and wholesale-sale of electricity and natural gas in interstate commerce from combined weather forecasts from many different countries weather forecasting agencies; and accelerate scientific discoveries by injecting mediation metadata into Long-Tail data from within the earth and space sciences.

Hammerman, Jim: ***Earth and Space Science Data Education for K-12 and the Community: Building on What We Know About Data and Statistics Learning***

Educating citizens to participate in community decision-making built on understanding of what Big Data, including Earth and space science data, can tell us must begin in K-12 schools, and include lifelong learning opportunities as well. Not everyone needs to become a data scientist, but we want to provide enough basic understanding about working with data so that everyone can make sense of data and its implications, and to create a wide enough pipeline to build a community of data scientists. Fortunately, we already know some things about how people learn about data and statistics and about how working with "big data" changes that. This talk describes some of those prior findings and resources and suggests new directions for learning research that address the special features of Earth and space science data (e.g., spatial and time series relationships). It also describes the importance of a community to curate data, points to issues in ethics and privacy around data, and makes links to the data component of computational thinking.

Heatherly, Sue Ann: ***NEREID and the NSF INCLUDES program***

The Green Bank Observatory is lead PI on an NSF INCLUDES Alliance called the First2 Network. The Network is based in West Virginia and aims to double the percentage of STEM students who graduate with a STEM degree. Our target audience is first generation students who have declared a STEM major. Our focus is K- the first two years of college. One of the well-studied best practices to increase retention in STEM pathways is to engage students early in the undergraduate career, in authentic participation in research communities. This is easy for large universities to accomplish but there is not equal access across smaller predominantly undergraduate institutions. NEREID, with a goal of converging access to big data across Earth and Space Sciences could open research opportunities for all students. Understanding how to build in mentoring would be key.

Heidorn, Bryan: ***ASTROLABE and other CyVerse Projects: Navigating Astronomical and Ecology Data through Advanced Cyberinfrastructure***

I will talk about the CyVerse Cyberinfrastructure system and three projects running on it including the, The HDR Phenotype from Environment and Genome (PEG), James Webb Space Telescope NIRCam working group, the Event Horizon Telescope. CyVerse provides life scientists with powerful computational infrastructure to handle huge datasets and complex analyses, thus enabling data-driven discovery. Our extensible platforms provide data storage, bioinformatics tools, image analyses, cloud services, APIs, and more. Recently CyVerse has begun supporting other sciences. CyVerse is funded by the National Science Foundation's Directorate for Biological Sciences. We are a dynamic virtual organization led by the University of Arizona to fulfill a broad mission that spans our partner institutions: Texas Advanced Computing Center, and Cold Spring Harbor Laboratory.

Kochevar, Randy: ***The Oceans of Data Institute: Preparing individuals to succeed in a data-intensive world***

Established in 2013, the Oceans of Data Institute is dedicated to transforming education to help people succeed in school, work, and life, in a data-intensive world. This is accomplished through two major strands of work; one focused on integrating data literacy skills into K-16 education; the other on building pathways to careers in data science and analytics. ODI's work is grounded in occupational profiles in data-intensive careers, both at the mid-level and at the professional level. Data literacy education projects currently serve students and teachers from grade 3 through undergraduate, and focus on a variety of topics including Earth science, life science, social science, and urban mobility. In these projects, which are funded through a combination of federal (typically NSF) and non-federal sources, we conduct education research by designing novel interfaces for accessing and analyzing data; building curricula to support student learning; creating professional development programs for educators; or a combination of these activities. Our work tends to be multi-disciplinary and collaborative, and typically involves subject matter experts and data providers from beyond our walls. We look forward to engaging with the NEREID community to explore new directions and opportunities for partnerships.

Olds, Shelley: ***Making Geodesy Accessible and Relevant: Workforce development, educational resources, data exploration tools at UNAVCO***

Geodesy is the study of Earth's shape, gravity field, and rotation. Geodetic research defines the terrestrial reference frame; and quantifies changes in the properties of Earth's surface and subsurface, ice sheets and glaciers, and oceans and atmosphere. Geodesy's broader benefits include help with preparedness and mitigation of hazards; and foundational support for space-based operations, navigation, communications, surveying, resource management, and national security. UNAVCO operates the National Earth Science Geodetic Facility, under an award from the National Science Foundation (NSF), that includes support from both NSF and National Aeronautics and Space Administration (NASA). Through our Geodetic Infrastructure and Geodetic Data Services Programs, UNAVCO operates and supports geodetic networks, geophysical and meteorological instruments, a free and open data archive, software tools for data access and processing, cyberinfrastructure management, technological developments, technical support, and geophysical training.

The UNAVCO Education and Community Engagement (ECE) program includes four areas: community engagement, education, outreach, and workforce development. Geodetic applications and data are used to understand many societal issues and can be incorporated into almost any topic being taught in geosciences. Resources and data exploration tools make geodesy accessible to both educators and students by providing content that can be leveraged and incorporated into geoscience courses and investigations. Through the workforce development programs, UNAVCO provides multiple career and mentoring resources for both students and faculty through intentional preparation and talent development of populations historically underrepresented in the geoscience workforce.

Peake, Leigh: ***Gulf of Maine Research Institute***

I will present a concrete instance of GMRI's work to leverage local and global data sets in linked informal and formal learning experiences where students investigate the connections between warming temperatures in the Gulf of Maine, shifts in the range of the state's iconic crustacean (Homarus americanus), and the emergence of an invader "from away" (the black sea bass). This example from practice will highlight the challenges of putting Earth data into students' hands in ways that are meaningful, pedagogically sound, and practical to support and sustain. It will also illustrate the profitability of climate change as a topic to motivate student work with data--these stories of change over both space and time pose authentic, purpose-driven contexts for student encounters with essential concepts such as trend, variability, and uncertainty. The example will also highlight central challenges for experience designers such as 1) finding the "sweet spot" of a compelling narrative with a strong climate signal for which local and global data sets are accessible; and 2) the absence of data tools accessible for young learners that still have the power to deliver multiple representations of data.

Sheth, Kartik: ***NASA Capacity Building Program: Building Skills to Use Earth Observations***

These are different kinds of programs aimed at getting users to use NASA Earth observations data for applied science purposes: ARSET, DEVELOP, GLOBE and SERVIR.

Stahlman, Gretchen: ***Exploring the Long Tail of Astronomy Data***

This spotlight talk highlights two projects that aim to support data curation and scholarly communication efforts in several domains. The first project - "Exploring the Long Tail of Astronomy Data" – implemented a mixed-methods approach to 1) capture insights into the research practices and data infrastructures of astronomers; and 2) identify indicators of uncurated astronomical data associated with scholarly articles. The second project – "Mining Georeferences from Biodiversity Literature" - was an exploratory study conducted in Summer 2018 through the LEADS-4-NDP fellowship program, which provided recommendations on methods and workflows for parsing geographic references in the text of Biodiversity Heritage Library (BHL) collections. For each project, I discuss related work, present an overview and status, and highlight possible outcomes and implications of this research.

Summers, Ed: ***Insight Without Sight: Non-visual access to data visualization and quantitative analysis***

Data visualization is inherently visual. Does that mean people with visual impairments or blindness (VIB) are out of luck? Not anymore. Join Ed Summers, Director of Accessibility at SAS, for a demonstration of technology that enables students and professionals with VIB to access the same quantitative insights as their sighted peers.

Trouille, Laura: ***Tales from the Zooniverse: Unlocking Data through People-Powered Research***

Processing our increasingly large datasets poses a bottleneck for producing real scientific outcomes and citizen science - engaging the public in research - has proven a creative solution, particularly when coupled with automated machine learning efforts. Zooniverse.org is the world's largest platform for online citizen science, with over 2 million registered volunteers participating in over 100 active projects; tagging animals in wildlife images, discovering exoplanets, transcribing artist's notebooks, tracking resistance to antibiotics, and much more (see https://www.zooniverse.org/projects for a full list). Since the launch of the first project, Galaxy Zoo, in 2007, Zooniverse projects have led to over 150 peer-reviewed publications (see zooniverse.org/publications). With the launch of the Zooniverse 'Project Builder' (zooniverse.org/lab) in 2015, anyone can build their own Zooniverse crowdsourced research project for free using a simple, browser-based interface. In this talk I'll provide an overview of Zooniverse, the Project Builder, and briefly mention our education efforts through classroom.zooniverse.org.

Villagran, Yasmin Catricheo ***Mapuche culture and its Cosmovision of the Universe***

From an indigenous perspective, this talk will explore how indigenous peoples from around the world have used connections between Earth and sky that have helped them to define their culture, religion, agriculture and livestock, spiritual, and even temporal points of view. Observations of the sky have been used to establish calendars, units of measure, the most optimal time to harvest or sow, travel time, orientation, etc. This indigenous worldview connects knowledge of the Earth's surface with the inner layers of the Earth and the known universe. It is a knowledge that has developed over time, and one in which it is impossible to conceive of these pieces individually. This perspective has been compared to systems thinking, which helps to connect everything, from the environment to individual health to business enterprise and ecology.

Wohlbrecht, Tiffany Stone: ***Data to Dome: The Planetarium Pipeline***

Data to Dome is an international initiative within the planetarium community to get data sets on planetarium domes quickly and easily, enabling planetariums to continue being a valuable source of the latest information in science for their communities. This presentation will provide an overview and vision of this initiative, current examples of its implementation, challenges it faces, and how it might be utilized in NSF's vision for the Convergence Accelerator program.

Yulaeva, Elena: ***Global Forest Link: Engaging Youth Worldwide in Collaborative Environmental Analysis and Decision Making***

Global Forest Link (globalforestlink.com, GFL), an award-winning environmental program, engages youth worldwide in active protection of the environment, with the goal of creating a more sustainable future. Initially developed as an educational extension of the WRI's Global Forest Watch platform, the Global Forest Link project has brought together over 2000 students from 100+ schools and youth groups, and 7 countries. Its key objective is to nurture a new generation of 'world stewards,' who are skilled in modern earth observation technologies, data collection and analysis, evidence-based education, international collaboration, environmental advocacy, and journalism. GFL teaches youth to explore key environmental change issues by integrating space imagery and local data, understand the environmental and socio-economic consequences, and to communicate their findings to peers and communities.
GFL innovations include:

- Bringing new elements to established citizen science and environmental science learning practices and approaches;
- Teaching youth the 21st century skills;
  Introducing youth to modern software tools for collaborative exploration of forest data and data science, such as SuAVE (Survey Analysis via Visual Exploration), which lets students from around the world to view, map, and analyze collections of forest images and videos collected by their peers, and link them to the GFW platform
- Engaging youth in social learning and data-driven journalism.

We will discuss the strategies of cultivating the youth ability to create a compelling story based on the integrative analysis of information.

Zaslavsky, Ilya: *Exploring Data and Data Tools*

Data Discovery Studio is an online platform for data discovery, which indexes 1.69+ million datasets and other resources from more than 40 geoscience data repositories. Dataset metadata is automatically augmented using text analytics and several geoscience ontologies and gazetteers, to enable more powerful search across data sources and domains. The metadata records are accessible via standards-based APIs, and are also exported in schema.org for indexing by Google. Users can edit metadata descriptions, track provenance, organize the discovered datasets into collections, and launch Jupyter notebooks for any dataset or a collection of datasets. In the context of one of C-Accel Track A projects (KONQUER), DDStudio is used to federate data discovery across geoscience and biomedical data repositories, to better understand environmental stressors leading to diseases such as Valley fever.

SuAVE is an online platform for visual exploratory analysis of surveys and image collections, with multiple applications in the Earth sciences and other domains. Users can publish their surveys and image galleries, analyze them visually, statistically and cartogaphically, annotate distribution patterns and individual items, and share their findings. In addition, SuAVE serves as a gateway into advanced data science tools, as users can launch Jupyter notebooks residing on multiple Jupyter hubs to perform image processing, machine learning, and other operations, bringing the results back into the exploratory analysis environment. The platform has been used to teach several UCSD undergraduate classes on research methods and data science, and in multiple research projects.

Appendix D: Promising Research Projects

All participants were invited to distill their thinking into specific project ideas and present them as research project proposals, building on the concepts developed throughout the workshop. The resulting set of proposed science and technology "sweet spots" and corresponding research projects demonstrates the expertise, innovativeness, and proposal writing skills of those participants who accepted the challenge. Seventeen research proposals were presented, creating a comprehensive look at a range of approaches to Earth-space data science convergence. With permission of the intellectual property owners, some of these project proposals are reproduced here. Contact information is provided for each author and we highly encourage you to contact the original authors if you are interested in collaborating on or funding one or more of these projects.

Cramer, Catherine B.: ***Network of Earth and Astronomy Domain Data (NEADD)***
Addressing the urgent issues facing our world today will require collaboration and communication on all fronts. Key among these needed collaborators are the domains of Earth and astronomical science. Increasingly, these currently largely separate science domains are converging on the methods and tools for sensing, gathering and organizing enormous data sets, moving toward pulling geoscience remote sensing, planetary science and exoplanet data researchers together into a single research endeavor, much needed to help solve these enormous challenges such as climate change.

Patterns that are revealed in one are increasingly being seen in the other, offering insight and inspiration. Much needed is a way to quickly see these connections in order to better understand and use them. The Network of Earth and Astronomy Domain Data (NEADD) is a substantial next step to answering that need. Tools originally developed for astronomy are now being used for everything from biomedicine to social science. NEADD will allow researchers on both fronts to quickly see connections and subsequently look deeply into tools being developed for astronomy and how they may be applied to studying climate data, for example. NEADD will reveal connections between tools according to how they are being used, which will reveal capabilities that may be applied to other domains. NEADD will also reveal the extent to which the domain communities are collaborating and communicating with each other, through its Social Network layer, measuring the increasing number of cross-domain research communities such as EarthCube and ESIP.

NEADD will also be useful as a tool for education. Students will be able to literally see the connections between domains and choose tools to use for their own data-gathering projects, or invent new ones. Teachers will be able to look at the Tools layer to see which tools are most appropriate for their classroom and compare them across domains. In addition NEADD will provide the information for science mapping of these domains in order to measure convergence.

Intellectual Merit This effort will make visible the kinds of research, data and tools being used by the fields of both Earth and astronomical science as well as the relationships between them, which in turn will help advance a convergence of use in order to address the wicked problems facing our world. The fact that most of the important discoveries in contemporary science are emerging from data-driven approaches indicates that skills and knowledge in these approaches must play a significant role in 21st century STEM learning, as well as in addressing issues of equity.

Broader Impacts: NEADD will be used to engage and advance the skills of early career scientists, post-docs, grad students, undergrads and – most importantly - the next generation of explorers. It will offer an even playing field open to all who are curious about where the convergence of knowledge of Earth and space is leading, and how it is best suited to solving our most urgent problems.

Downs, Robert: ***Collaboratively Improving Capabilities to Use, Analyze, and Visualize Earth and Space Science Data***

Summary

The use, analysis, and visualization of Earth and space science data offers opportunities for diverse members of society to understand our world. Earth and space science data can be analyzed by students to learn about how the actions of society influence the environment that they inhabit. The use of such data also enables scientific research on the Earth and space sciences. Using such data also facilitates decision-making by planners and policy makers. Collaborative development of easy-to-use interfaces offers opportunities to facilitate analysis and visualization of Earth science data for learning, research, and practice. The proposed project will engage experts in Earth, space, and computer science to work collaboratively across disciplines to investigate opportunities for improving the capabilities for various segments of society to use, analyze, and visualize Earth and space science data.

Intellectual Merit

The proposed project will advance knowledge on synergies that can be achieved by developing interfaces to use, analyze, and visualize data from both the Earth and space sciences. Similarities and differences in the approaches used to analyze different types of data will inform understanding on how people from diverse fields can learn, conduct research, and engage in practical uses of Earth and space science data.

Broader Impacts

The proposed project will identify synergies among the capabilities needed to utilize both Earth and space science data. The proposed project also will develop interfaces to improve capabilities for students, researchers, and practitioners to use, analyze, and visualize data representing observations from both the Earth and space sciences.

Fuka, Daniel: *DomeBroker*

Summary

While the museum and school planetariums give us amazing glimpses of the Universe and Pink Floyd laser shows, they're also commonly becoming a means of educating the public about current earth science discoveries in addition to the space science learning per their original objectives. What is needed is a reliable means to discover, access, and locally archive bleeding-edge earth and space science research that non-traditional educators in the planetariums can incorporate into their shows as scientists make and publish new discoveries.
Dome Broker, or DBroker, proposes a core open-source interface to discover, access, and make locally accessible, geoscience data sources for displaying current global and locally relevant geoscience discoveries at Planetariums. DBroker joins methods and software from EC Building Blocks, to create a next-generation community education. Leveraging a large installed base within the Earth and Atmospheric Sciences, and its open-source flexibility, DBroker extends Hyrax—OPeNDAP's data server with capabilities for distributed brokering, to become both a multi-scientific domain data broker as well as an auto-syncing data repository that can cache and make locally available data necessary for planetarium education programs communicating current discoveries to the public.

Intellectual Merit:

Towards NSF's, NASA's, and USGS's community educational goals, this exchange will offer "data-as-service" capabilities that match diverse data sources with established applied community educational resources in local and portable planetariums.

Broader Impacts:

Education's greatest challenges require community educators to quickly assimilate cross-discipline research. DBroker outcomes will accelerate knowledge acquisition for both traditional (licensed K-16) and non-traditional (planetarium staff and operators) educators from all levels, simultaneously fostering the expertise needed for its evolution. Ultimately, this will yield a better public understanding of the Earth and Space Systems, globally and locally. With this pipeline from brokering to the extension of developing DomeCasting capabilities, it opens up the world for remote planetarium community educators to reach communities with portable domes to fill entire planetarium day series off-site, and for geoscientists educators to dynamically interface alongside the space sciences, giving the teaching medium/environments twofers.

Guridi, Jose: ***AstroEarth Technology Transfer***

Summary

Convergence between different areas like astronomy and earth sciences can take place not only in the analysis of common spaces, but also in the technology and technique used across disciplines. For example, there are many experiences of astronomy transferring knowledge to very different disciplines like medicine and other examples (There are some examples in here: https://science.sciencemag.org/content/331/6018/696; https://www.cfa.harvard.edu/COMPLETE/astromed/; https://www.iau.org/public/themes/astronomy_in_everyday_life/). Given this diagnosis, we propose to build a collaboration to explore and identify technological convergence between astronomy and earth sciences. This projects would need to group a large network of researchers to study tools used in astronomy and earth science during a year and come up with areas of application. The project will need an executive group that will work on the convergence of technologies and will relate with academic institutions, government organizations, researchers and industry using data science tools in earth and astronomy fields. A first step could be to collaborate with the Chilean Ministry of Science is working on a Climate Change Observatory with oceans and cryosphere data and will probably collaborate with the Data Observatory, enabling an ecosystem with astronomical and climate change data which can be a natural testbed to test technology and knowledge transfer.

Stakeholders: Data Observatory, Chilean Government (Ministry of Science), institutions and researchers related to astronomy and earth sciences, industry interested in these kind of data or tools, and astronomical observatories and earth science institutes.

Intellectual Merit

Even though there are cases of technological and knowledge transfer between astronomy and earth science and other areas, the general rule is that they are isolated cases and there is no systematic effort looking for commonalities and transference. There would be great merit in making a systematic convergence effort of technologies and techniques.

Broader Impacts

If there is a group permanently looking for convergence areas between two fields, they could develop a methodology to increase technology and knowledge transfer between science, industry and society.

Heatherly, Sue Ann: ***Convergent Teacher Professional Development***

Summary

Who:

a. With formal teachers who teach earth and space sciences- diversity of perspectives– urban, rural, underrepresented
b. With ESRI, other tool providers/developers
c. With researchers in the field- including graduate students
d. With National Lab representation – EG: NOAA, NRAO, NSO,

What: This group will

a. understand potential relevant research questions of today
b. understand existing tools and barriers to use- rapid feedback and fixes…
c. converge and integrate these questions into cross disciplinary challenges – what they? Ideas mentioned here are water
d. develop scope and sequence of curriculum that centers these challenges through the lens of local, relevance and community, that requires students to conduct research using data, report it and share it
e. create a community of practice spanning the stakeholders listed above.
f. Pilot the curriculum with teachers and students.

Intellectual Merit

I have a very willing group of teachers who teach ninth grade ESS, and the curriculum, I'm sure does not generate the next generation of big data explorers—this proposal would transform the ESS experience for students and develop big data literacy, as well as research and communication skills

Broader Impacts

understanding between domain professionals, faculty and teachers, resource that would be broadly disseminated.

Kochevar, Randy; Olds, Shelley E.; Yulaeva, Elena: ***Solar Impact Collectors On Earth***

Summary

We propose to develop an inquiry-based curriculum in which students will simultaneously collect data from solar radiation using two different techniques; one based on observing the sun directly using XXXX, and the other detecting voltage signals in Earth-based antennas. Through the proposed learning experience, different groups of students will use different techniques to observe solar coronal activity; one through direct observation of the sun, and the other through a ground-based antenna. The culminating experience will involve combining and analyzing the data from each experiment, looking for correlation between observed solar phenomena and measurements of Earth-based antennas. In order to ensure the success of this project, we propose to collaborate with members of the astronomy community with expertise in solar observation and radio astronomy, as well as data science education.

Intellectual Merit:

The transition from individual, student-collected data to broader, more complex datasets is thought to be a key component in mastering data literacy skills. By combining self-collected data with data collected simultaneously by peers, students take an initial step towards this transition, as well as learning fundamental data life cycle skills; beginning with data acquisition and storage, visualization and analysis. They will also learn about fundamental statistical concepts including variance and correlation.

Broader Impacts:

Data literacy skills are a fundamental part of inquiry-based science learning, as laid out in the Next Generation Science Standards (NGSS). Furthermore, as our society becomes increasingly data-focused, the ability to work effectively with data is becoming critical to Success In high educational pursuits and in the job market. THrough sharing data results with teams across the country and World, students are contributing to collaborative knowledge sharing and …

Madsen, J.: ***Connections Between Earth and Space Science Related to IceCube***
(sent electronically before the workshop)

The IceCube Neutrino Observatory (IceCube) at the South Pole has a wide science reach. The primary purpose is to explore the high-energy universe using neutrinos, almost invisible cosmic messengers created in extreme astrophysical environments and engines. The telescope consists of over 5000 light sensors that have been placed a cubic kilometer of ice, at depths between 1450 and 2450 meters below the surface.
The grid of light sensors, most 125 meters apart horizontally and 17 meters apart vertically, record light given off by electrically charged particles created when neutrinos interact. Cosmic rays, high energy electrically charged particles from outer space, create a shower of secondary particles when they collide with an atom in the earth's atmosphere. This creates neutrinos and another elementary particle called a muon. Muons are electrically charged and the ones with sufficient energy can reach IceCube. We see about 2800 muons per second, or 100 billion per year.
The cosmic rays also create neutrinos, but they are invisible unless they interact which is highly improbable but not impossible. Like the lottery, if enough people play someone always wins. About 10 times per hour we detect an event that was caused by a neutrino interaction. About once or twice a month, we see a very high energy neutrino event from a neutrino that must have originated from outside the Earth's atmosphere. It was from an extreme environment like those around massive black holes or from an extreme event like mergers of black holes or neutron stars or a hypernova.

The above examples are good ways to talk about probabilities. What is the difference between impossible and improbable, for example. The data analysis illustrates two big approaches often encountered in the era of big data. The large sample allows one to look for small deviations. For example, the moon and sun block some of the cosmic rays so there is a shadow which is seen as a deficit of events in the direction of the moon and the sun. The moon shadow is constant in size and can be used to test how precise and accurate our angular reconstruction of the data is. The sun shadow changes measurable based on how magnetically active the sun is, which varies over a 22-year cycle. So we can learn about the sun from the measurements at the South Pole.
The 2800 muons per second rate depends on the season because the temperature and hence density of the atmosphere changes. So we can measure the atmospheric temperature over Antarctica while looking for high energy particles. Another aspect is that we need to know the optical properties of the ice where the light sensors are located. The ice is the clearest material ever measured but still has some dust layers that change the absorption and scattering lengths as a function of depth. In fact, the dust layer are tipped about 50 meters across the 1000 meters that are instrumented. The ice properties also are about 10% different in the direction the glacier is moving compared to the perpendicular direction. We need to characterize this to reconstruct our data to determine the enegy and direction of the detected particles. But understanding how these variations in ice properties occurred tells about past climate including possibly the wind speed.

The ice may also hold isotypes created from past supernovae that can be recovered from ice cores to study supernova rates. The ice cores are crucial for learning about CO_2 levels and temperatures.
There are a couple of immediate project that students could do to collect data. The first one, DECO (https://wipac.wisc.edu/deco/home) , uses cell phones to look for cosmic rays. It uses some data but the app is free and pretty straight forward to install and use. There ae a number of simple projects that could be done to investigate how to shield (block) cosmic rays for example. It is also a nice project for connecting students at different locations since cosmic ray rates vary with latitude and elevation.
Another project is Cosmicwatch (http://cosmicwatch.lns.mit.edu/about). This self-contained cosmic ray detector costs about $100. It is easy to use and again there are multiple projects that can be done including looking at the cosmic ray rate through out the year to see if it changes with the seasons. Like with DECO, comparison of results from other locations or from different altitudes or environments such as in abasement of tall building or on the top floor. The technology is similar to what is used in medical imaging so that opens up ways to talk about that connection.
Hope this helpful!

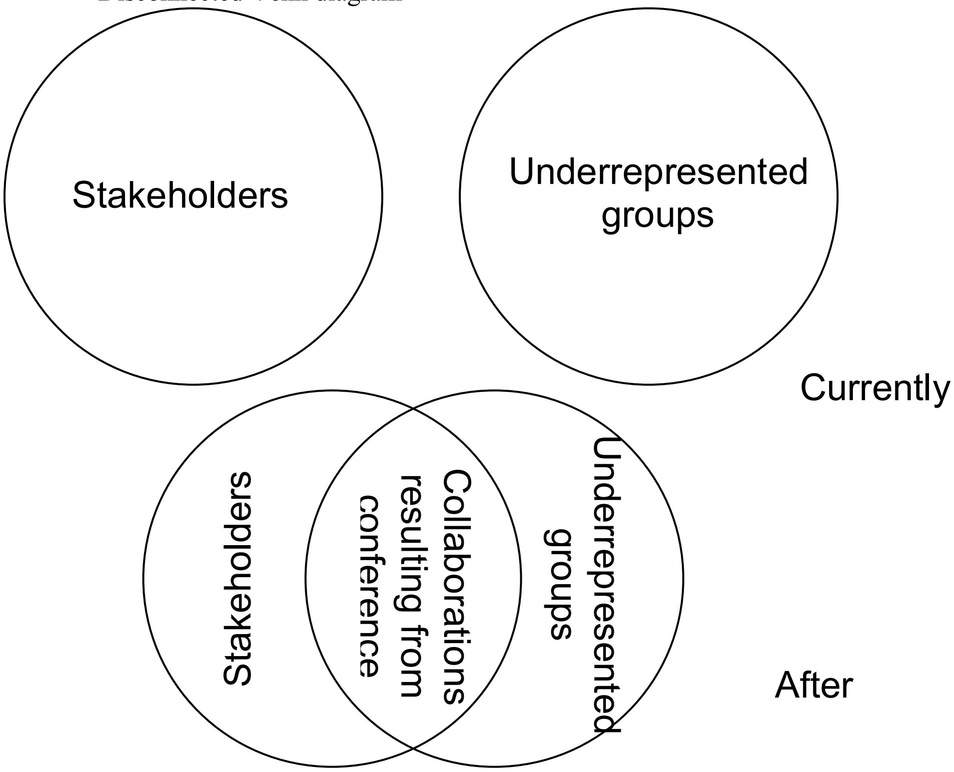
Raftery, Claire; Summers, Ed: *Inclusion and Accessibility Conference*

Summary

- Describe the obvious fact that NSF is interested in pursuing DIA activities. We believe that inclusion of underrepresented populations makes the science better.
- Meager progress has been made, low impact with various populations. Support with studies about populations and rate of inclusion in STEM, data on pay gaps, employment of people with disabilities
- Theory why meger progress has been made. What is the problem?
    - Systems thinking approach has not been applied, it has been catch as catch can (isolated, unconnected attempts). Cohesive framework has not been established.
    - “Nothing about us without us” hasn’t been applied. People with disabilities haven’t been included in an explicit, sustained way.
    - Stakeholders (industry and academic institutions) have a role to play often don't have specific action, understanding or access for specific underrepresented populations. (Columns = underrepresented populations)
- We believe Earth-space science represents a convergence zone within science that's big enough but not too big. This is included throughout K-12 education and is foundational to early education, and is a commonly recognized gateway science. Computational sciences are a critical component of Earth Space science, and will become more so moving forward. We believe that a systems approach to DIA within the convergence zone of earth space science is applicable to many other fields of many other fields of science. With the majority of NSF’s Large Facilities residing in this convergence arena, there is added incentive to make this

area the gold standard for DIA. In this way, NSF becomes the exemplar of DIA across other federal agencies, setting the standards for e.g. NOAA and NASA.

- In order to accomplish this, we propose the following concept as the framework for a systems approach to DIA.
- Disconnected Venn diagram



We propose an inclusion matrix the represents stakeholders as rows and underrepresented populations as columns. The full and equal participation of any particular underrep population is facilitated or not by the actions of the stakeholders in the cells throughout a particular column. Each cell requires further input from the stakeholder and the specific underrepresented population. To that end, the first step in implementing the framework approach is to convene a meeting with representation from all stakeholders and underrepresented populations. The objective of this conference is to conduct a gap analysis regarding the needs of that particular population within the field of that stakeholder. For example, how can informal educators include youth with visual impairments or blindness in a planetarium shows about Earth and Space science topics? An outcome of the conference would be to facilitate both groups jointly designing a proposal to the NSF to fulfill needs identified using the resources of the stakeholder. This might include developing tools, online training modules for providers of informal education, detailing best practices in engaging and supporting underrepresented students, building engagement programs and providing exposure and opportunities for students. In our example above, the informal science educator would work with the National Federation of the Blind and individuals with lived experience to create manipulatives (tactile resources) that would represent the data that is being delivered visually. [Need to emphasize is to identify needs and resources needed to fulfill needs, not create the resources themselves]

Intellectual merit:

- SIG/access (special interest groups for accessibility). Accessibility research will provide opportunities for researchers to engage with researchers and underrepresented populations to conduct and test hypotheses.
- Research agenda - plan for researchers to move the field forward.

Broader impacts:

- Measure impacts on individuals in underrepresented group in their identities and attitudes towards science, and seeing themselves as scientists.
- This conference will facilitate knock on impacts for long term benefits to underrepresented groups.

Reid, Rebecca; Boone, Richard: ***Integrating Earth Space Data into NGSS***

Summary

Identify data science educational resources that already exist: formal and informal; curriculum, programs, activities, online tools, etc. and connect these identified resources to NGSS Crosscutting Concepts— 1) Patterns, 2) Cause and Effect, 3) Scale, Proportion and Quantity, 4) Systems & system models, 5) Energy & Matter, 6) Structure & Function, 7) Stability & Change.

Intellectual Merit

Many educational resources and teaching tools already exist, but they are not utilized in science classrooms with an already full set of concepts to "cover", and teachers without a data science background don't know how to integrate data science into their science classrooms.

Data science skills are increasingly important in multiple careers and arguably should be a core learning competency for all high school graduates. Although data skills are frequently identified through NGSS, data science skills with large data sets in particular are rarely, if at all, addressed. The proposed project would identify the NGSS that specifically address data literacy skills, gaps in skills that focus specifically on obtaining information from large data sets, and opportunities to use existing tools and large land-space data sets to foster those skills. The project will include identify approaches in undergraduate data science courses that could be applied with modification at the high school level plus what is known about data-science pedagogy.

Broader Impacts

We need to prepare students at all levels to interact with data in earth and space science, both in post secondary schools and in daily life. Encountering data science in their earth and space science classrooms will integrate data into science as a common practice. The ability to obtain useful information from large data sets will become increasingly important in STEM and non-STEM fields. Current and emerging tools can be applied at various grade levels to visualize, analyze, and manipulate large data sets. The development of data science competencies among students at the primary and secondary school level offers the potential for community-based projects involving evaluation of large data sets of relevance to a community. Current K-12 teachers do not have time to develop an integration of earth and space data science skills, knowledge and behaviors into their classes. This idea would give them the tools to prepare students to encounter data science in post secondary education.

Spuck, Tim: ***Earth-Space Accessible Analysis Tool***

Summary

Develop a common data analysis tool that works across geospatial/astronomy data. The tool would be accessible to BVI/DHH. This would be built on (or modeled after) the existing Afterglow Access software being developed via the Innovators Developing Accessible Tools for Astronomy project. The tool may be built in a way that individual features could be turned on and off, reducing confusion for new users.
(Collaborators would include AUI, SAS, TERC, UNC-Chapel Hill, National Federation for the Blind or similar, DHH group(s), astronomy and geoscience researchers and educators.)

Intellectual Merit:

Common tool that can be used across disciplines, reduces barriers to use, and increases potential science output.

Broader Impacts:

Tool is accessible and can be used to diversify the STEM pipeline.

- May be able to be used in medicine and other areas.
- Both astronomy and geosciences benefit because training for educators, learners, and even researchers can use common tutorials and training resources.

Spuck, Tim: *Exploring Connections between Life on Earth and Celestial Events*

Summary

Survey of Celestial Event Impact on Nocturnal Animals. Over thousands of years indigenous people have recognized the Earth-sky connection, understanding the Earth-sky collectively, not separately. There are a number of examples of convergence in Earth-space research yield significant results. It has been demonstrated that dung beetles use the sun, moon, and the Milky Way for navigation. In addition, research indicates whales may use the stars for navigation. Increasing amounts of data available on migratory patterns and animal movement present unique research opportunities. There is an increasing interest in the use of orbiting satellites for commercial advertisements. For example, the Russian company StartRocket is working on launching satellite ads that would be viewed in the night sky as early as 2021. It is important that we understand the impact this activity might have on animal behavior and ecosystems. A deeper understanding of the impact naturally occurring celestial events have on animal behavior and ecosystems will help us better predict the impact of these orbiting satellites and work to mitigate potential problems. This project will use nocturnal animal tracking data, Earth surface topography and vegetation data, and astronomical data to explore correlations between animal behavior and know celestial events.
(Collaborators would include indigenous knowledge holders, ecologists, sociologists, animal behaviorists, geoscientists and astronomers).

Intellectual Merit

- Changes in migratory patterns and ecosystems can be disruptive to human activity in communities around the globe. A better understanding of how life on earth is impacted by activity in the sky can help us better prepare for and mitigate potential problems.
- Findings can inform the use and development of tools and best practices for commercialization of space.

Broader Impacts:

This project presents an excellent opportunity for individuals of all ages from across the globe to engage in the research effort via citizen science. Improving STEM literacy for all can not only inform individual career decision in young learners, but it can help create a society that understands and values the STEM enterprise. Partnering with indigenous communities on the research presents an opportunity to have meaningful engagement with traditionally underrepresented groups.

Spuck, Tim: Exploring Human Interaction with Earth-Space Data

Summary

Explore user multisensory interactions across earth-space data. What are the benefits of exploration of earth-space data through sight, touch, and sound? Are their patterns that are more easily identified in images through sight, touch, sound or some combination. How does this differ across BVI, DHH, and sighted users? How does this differ between musicians and non-musicians?
(Collaborators would include AUI, TERC, National Federation for the Blind or similar, DHH group(s), astronomy and geoscience researchers and educators, and learners/users.)

Intellectual Merit

Understanding human interaction with data can lead to improvements or new ways of analyzing data, potentially leading to new discoveries in Earth-space sciences.

Broader Impacts

This improved understanding can lead to the construction of tools that increase data accessibility in Earth-space science, and diversification of the STEM pipeline. Findings can inform the use and development of tools and techniques outside the Earth-space sciences convergence.

Spuck, Tim: *Light Pollution Reduction through Analysis of Earth-space and Social Data*

Summary

Improving Astronomy and Human Health: Identification of Best Practices in Light Pollution Reduction - The Socioeconomic Data and Applications Center (SEDAC) is a powerful tool to explore population densities and other socioeconomic variables. As population density increases so does the human footprint. Light pollution has been known to have a negative impact on observations of the night sky, as well as a negative impact on human health and a variety of life on Earth. We propose using data from SEDAC, satellite image data, and Globe at Night data to identify areas of high population density and low levels of light pollution. These targets would then be further studied to identify best practices that could potentially be replicated elsewhere, reducing light pollution and the impact on energy consumption, astronomy, human health, and animal well-being.

(Collaborators include AUI, CIESIN, astronomical observatories, ecologists, etc.)

Intellectual Merit:

The reduction of light pollution increases the opportunity for scientific discovery at astronomical observatories around the world. In addition, if more effective lighting practices can decrease energy consumption and improve human health, a variety of disciplines can benefit from the knowledge gained via this project.

Broader Impacts:

Obviously, an improvement in human health has broad societal benefits. However, there are indirect benefits as well. Currently, 80% of the US population can not see the Milky Way from their own “backyard”. Astronomy is a gateway science that can inspire young learners to consider a variety of STEM careers. A reduction in light pollution makes the wonders of the night sky more accessible to all, increasing the opportunities for astronomy education and outreach. In addition, the findings in this study could significantly inform a similar effort to decrease RFI and increase opportunities in radio astronomy.

Stahlman, Gretchen; Heidorn, Bryan: *Resurrecting Data Across Disciplines*

Summary

This project brings together an advanced computing environment and large-scale curation of valuable raw and processed data currently stored on physical media such as magnetic tapes and floppy disks. The project will develop protocols and workflows and initiate a national collaboration of institutions to test processes on pilot datasets, as well as to prepare for ongoing ingest and use of the data beyond the duration of the funding period. Reuse of older data is critical to enable time-domain research, particularly in astronomy and ecology. The strategic collaboration proposed here will broadly facilitate open scholarly communication by sharing access to rare digital media readers and cloud storage for open access to and analysis of historical data, including linking to associated journal publications and integrating with other data. This project also proposes to develop training programs for librarians and educators about data curation and incorporating newly resurrected data into education initiatives. These training programs will be leveraged as human capital, by providing hands-on experiences working with data and in collaboration with researchers helping to make their data discoverable.

Intellectual Merit

Studies across disciplines have demonstrated the potential for new findings from archival and legacy data, as well as the citation advantages of making data associated with publications available and discoverable. Furthermore, a number of current reports urge exploration of new paradigms for locating and managing data and software across time and place, and in support of research transparency and open science. Curating older data – including data stored on obsolete media such as Exabyte tapes – is a socio-technical and data science challenge. A tipping point is quickly approaching after which it will be problematic or impossible to adequately curate a large number of potentially-valuable datasets without the expertise of the original data creators, many of whom are approaching or past retirement (particularly in the case of data stored on media for which media readers are increasingly difficult to obtain). This problem complements another tipping point – in which new instruments such as the Large Synoptic Survey Telescope will universally change the way science is conducted, through massive data output as a community service and educational resource for the public.

Broader Impacts

Focused attention, innovation and expertise in this area can: reduce barriers to enable data creators and librarians to easily share data and metadata in trusted, long-lived repositories and cyberinfrastructure; determine and work towards a precise economic framework for sustainable data sharing; establish strategic partnerships and inter-agency agreements for development of cyber- and social infrastructures; characterize, prioritize and target highest-value data and metadata to enable grand challenge research across disciplines; and preserve publicly-funded knowledge and efficiency of science.

Possible collaborators:
- Rutgers University, LIS Department
- University of Arizona iSchool
- Center for Digital Society & Data Studies
- Astrolabe
- CyVerse
- CODATA
- EDI
- Data Observatory
- Publishers
- Libraries
- Educators

Uzzo, Stephen: ***Planetary Open Data System (PODS)***

Collaborators

NASA (JPL, Goddard, Ames), NASA DAACs, EarthCube, ESA, NOAA, AGU, optical and radio telescope research organizations, BD Hubs, along with research labs, universities, scinetometricians, and corporate tool developers (like IBM) etc.

Project Summary

There are many entities across the globe advancing various aspects of earth and space data sciences (sensor webs and remote sensing about earth, planetary sciences, exoplanet exploration and discovery, etc.), but finding ways to federate these many disparate data are elusive. The potential for the many kinds of data to inform each other and create rich ontologies, make extensive use of emerging artificial intelligence algorithms in seeking patterns and defining new area for research and discovery is great. The goal of this proposed project is to define, catalog and circumscribe exiting efforts in earth and space science data, characterize them, and envision and evolve a new generation of tools, resources, and epistemologies for making connections dynamically across domains of geoscience and space science knowledge.

The project would begin with a World Earth-Space Summit (WESS) that would be a highly inclusive convening of the many efforts across the globe and negotiate open data agreements throughout the earth space data domains of knowledge that would ultimately result in a Planetary Open Data System (PODS). Because of the massive scale of this undertaking it would have to consist of a network of resources, rely extensively on federating many disparate data systems through next-generation deep learning-driven semantic ontologies, so would have to include a wide gamut of sectors in computational and data sciences. Cooperation from the private sector will be essential to supporting and sustaining this effort. Through this process, the ultimate goal will be to develop a large-scale cyberinfrastucture or planetary data observatory that will allow both existing and new data streams to be federated, and through machine learning push patterns of interest in synoptic data to research teams to explore further and reason about.

Because of the innovative forward-looking nature of project a significant investment in learning research, curriculum, training and resource development would be an essential aspect of the project from the outset.

Intellectual Merit

Advancing data in earth-space sciences through a series of large-scale cooperative efforts will result in accelerating innovation across many sectors and advance many domains of knowledge. It will also spur the development of next-generation tools, resources and methods, that will accelerate discovery in earth and space data.

Broader Impacts

The kind of cyberinfrastructure and AI driven tools envisioned under this proposal will have implications for many distal and orthogonal field of science, including biomedicine, social sciences, and even models for sociopolitical and business systems. Further, because education and learning are deeply embedded in the project throughout all phases of its development, they would be immediately available for a wide variety of formal and informal learning settings from early learners through mid-career executives, teachers, managers, and others to use, remix, and develop to advance emerging areas of planetary and geosciences.

Wohlbrecht, Tiffany Stone: *Exploring Nature to Promote Data Literacy*

Summary

Developing Informal Education Resources for E-S data: GScouts, BSA, museums, and planetariums. Incorporating data collection and analysis projects into existing Earth Science and Astronomy badges. Educating museum and planetarium professionals on how to incorporate large Earth Science data in their work.
Developing a pipeline from Earth Science data to planetariums to immerse audiences in real, relatable data in real-time.
Connecting nature to data- other lifeforms use natural data to inform their lives and make decisions. This concept could be utilized to develop educational projects and materials.
Collaborators: Astronomical Society of the Pacific, International Planetarium Society, Adler Planetarium, NEREID Network

Intellectual Merit

Understanding data and how it can be used to understand your environment; nature has data for you to discover and learn from; can this be connected to climate change awareness?

Broader Impacts

Creating a data literate society. People who understand Data Ethics; a data literate society capable of tackling the issues around who owns data, how data is collected and used.

Zaslavsky, Ilya and Tuddenham, Peter: *Open Knowledge Network for Ocean Education (OKNOE)*

Project description

This project will accelerate integration across science education and particularly ocean science, Earth-space data, and knowledge representation and networking, in order to build a new generation of tools and knowledge structures that would allow teachers and students to navigate, explore and understand the body of knowledge about ocean and ocean processes in new and exciting ways. In particular, we will assemble and formalize knowledge representation of ocean science curricula, to make it possible for educators to create personal pathways through knowledge concepts for students at different levels and locations, integrating worked examples, both raw and cleaned datasets, simulations, models, and expert contributions into a personalized ocean knowledge dashboard. For every concept in the curriculum, users will be able to find related concepts - both those that are foundational to understanding the concept in question, and those that would derive from it, find materials illustrating this concept and testing its understanding, and also contribute additional materials and concepts, thus expanding the knowledge graph.

The project will involve a collaboration between oceanographers, professional educators, evaluators, computer scientists, and visualization and ontology researchers, as well as several companies, working together with a group of dedicated classroom teachers on field-testing the proposed tools and learning pathways. It will include a knowledge graph and associated software development, curriculum knowledge organization, and iterative testing of the software in years 2 and 3 of the project, with the goal of making it widely adopted and self-sustained by the end of the project period.

Intellectual merit. This project will attempt to resolve a critical and long-standing problem of efficient engagement of broad groups of students into personal exploration of ocean related materials, including ocean data and models, creating a novel science education gateway into data science tools focused on ocean exploration, while integrating data and knowledge from neighboring domains such as biology, atmospheric science, fisheries, climate change research, and coastal processes. Currently, such an integration is performed on an ad hoc basis in research, and is not formally present in education. We will propose and evaluate new curriculum approaches for students to achieve a higher level of competency in ocean sciences and, more generally, in science classes. A particular focus will be on creating a convergent curriculum that will help students address critical questions at the intersection of several domains.

Broader impacts. The project will engage several hundred high-school students from 2 schools, and 5 science teachers. In addition, the project will create a resource that can be used by students and educators around the world. The Decade of Ocean Science for Sustainable Development is 2021-2030. This will support the research and education for Ocean Literacy as part of the decade. Students will be able to relate their research on the dashboard with Sustainable Development Goals, a key UN initiative focused, in part, on improving ocean health.

Appendix E: Post-Workshop Survey Results

Module 1: Achieving Workshop Objectives

Please reflect on the 2-day workshop you have just completed and indicate your level of agreement with the following statements, and please comment on the answers to the following questions.

Question	Mean	Min	StDev	N
I believe convergence of Earth and space data may be useful to addressing some of humanity's greatest challenges	4.54	3	0.66	13
I learned about valuable Earth-space data resources that I didn't know about before the workshop	4.85	4	0.38	13
My understanding of challenges and opportunities to the convergence of Earth-space data practice and learning improved through this workshop	4.38	1	1.12	13
The workshop helped me understand the value of convergence research	3.85	1	1.14	13

Module 1: Answer

The workshop has been tremendously successful in bringing together researchers with diverse backgrounds and having them converge on a common agenda, challenges, and exploration of potential solutions.
I'd suggest we should explore use cases where integration of space and earth data would be most useful. Also, Earth measurements and instruments are more diverse than those used for space data. Methodological convergence would be more successful if there is some commonality in data types

Module 1: Answer

Throughout the workshop we circled around and provided examples of convergence and its application to solving Big Data questions. As a whole, I have a better appreciation of the need for convergence, the progress that is being done within this area, and the educational opportunities to bring this to the classroom.

Module 1: Answer

Talks were very informative. Would have liked to see more representation from Astronomy

Module 1: Answer

The NEREID workshop improved my understanding on the convergence of Earth and space science data.

Module 1: Answer.

I'm new to the idea of convergence research, and could use a more basic primer to it. I did read the NSF solicitation but didn't have time to see the types of projects funded.

Module 1: Answer.

I found it valuable to have these discussions with a diverse group of people with very different backgrounds.

Module 1: Answer.

The workshop enhanced my understanding of the valuable and diverse work already being done to converge earth and space data and stimulated many new ideas about possibilities for the future.
I learned that convergence is already happening naturally, but strategic planning around shared goals will make the convergence much more valuable to all on earth.

Module 1: Answer.

The workshop demonstrated the synergy between Earth and Space data, analysis tools and research. Those valuable resources and techniques would complement each other if integrated into rigorous education curricula and integrated in diverse existing courses.

Module 1: Answer.

We spent a long time getting to know one another's work and learning about potential resources. There may be potential in bringing together these data in ways that move beyond what each offers, and we discussed possible challenges and opportunities in doing so. But it's not entirely clear whether and how that will play out, and we didn't really get much opportunity to generate ideas for what convergence would provide, nor what tools, resources, infrastructure, knowledge, methods, etc. would be needed to really take advantage of that -- hints of possibilities, but without clarity about the intended nature of the convergence, nothing resembling a plan. There's an opportunity to continue that discussion and really figure out what the next steps would be, and what the most exciting opportunities for convergence would be, but I didn't feel that we ended up with clarity about how that would proceed either.

Module 1: Answer.

I want to preface my comments by saying that I appreciated the opportunity to participate in the workshop. I met a number of interesting people, many of whom I am sure I will collaborate with in the future. Seeing the Observatory was an amazing experience, and an unexpected treat. I know how much work it takes to put these things together – especially under impossibly tight deadlines – and my comments below should in no way be taken as personal criticism of any of the organizers, all of whom I enjoyed working with, and would gladly do so again.

I felt that the concept of "convergence research" was not clearly articulated before or during the workshop, and the small amount of time we spent actually discussing possibilities for convergence practice were insufficient to improve my understanding. I think that the experience might have been more effective had the need and vision been clearly articulated from the beginning, and the workshop focused on developing goals and strategies to fulfill and expand that vision.

Module 1: Answer.

The only reason why I did not strongly agree with the second phrase is because I already understood the value of convergence because of my work with the Data Observatory. Nevertheless, it was a great workshop to see what different people is doing with earth and space data and to force this community to start thinking in a more interdisciplinary and convergent way.

Module 1: Answer.

I think one very compelling area of convergence is the set of activities that are required to collect, clean, organize, host, and share big data sets. The private sector uses the term "dataops" to refer to these activities:
https://en.wikipedia.org/wiki/DataOps

Module 1: Answer.

The workshop was extremely stimulating, though it was evident that convergence requires much more than just tools and applications. While we all got along well and shared ideas and perspectives enthusiastically (and I came away with several new collaborations), without having a common language (along with related disciplinary frameworks and shared priorities), real communication took extra effort, and I am not convinced that we produced a coherent path forward towards convergence with data from disparate domains.

Module 1: Answer.

Convergence of Earth-space data is a big idea and requires a shift in thinking. Two days provided an opportunity to get a good start, but more time is needed for this group to discuss the issues around convergence with Earth-space data. I do think something that was especially helpful, and I heard it from other participants was engagement in the Drake Equation Challenge. This provided an opportunity to explore and experience this convergence in a real-world situation. It was also great to hear about all of the amazing resources that are out there that I did not know about. Someone needs to create a library of these resources, and then we can really start exploring ways to use these resources to advance convergence of Earth-space data, tools, attitudes of practice, and social interactions.

Module 2: Rationale for Future Work

Please reflect on the long-term big data literacy initiative as discussed in the workshop and indicate your level of agreement with the following statements.

Question	Mean	Min	StDev	N
Being part of this initiative is an exciting opportunity for me.	4.54	3	0.78	13
Being part of this initiative is an important professional opportunity for me.	4.00	2	0.91	13
I believe I have a valuable contribution to this initiative moving forward.	4.38	4	0.51	13
I believe this initiative will positively impact my institution / organization.	4.08	3	0.64	13
I feel inspired by this initiative.	4.54	3	0.66	13
I plan to invest my time and energy to make this initiative a success.	4.00	2	0.91	13
This initiative is relevant to the success of my work in other areas.	3.77	3	0.83	13

Module 2: Answer

Do you believe that the convergence of Earth and space data is important to data research and education? If so, why. If not, what ranks as more important?

I believe that Earth and space data are interesting to a large number of people, and important sources of information for understanding our world and what we need to do to keep it and the people living here healthy, happy, growing and supported, and to do so in equitable ways. I believe that it's vital for people to learn to work with such data, including dealing with the complexities related to time-series and spatially correlated data, and that this is a difficult task. For those who are not going to be directly working with analyzing data, it's also vital for people to learn to be critical consumers of research built on such data, and decision-making processes that build on that research. I think it's important for people to understand both how large data sets are just extensions of smaller and more local data sets that are more easily grasped, but also how large data sets and the combining of related large data sets collected for different purposes can go well beyond those smaller data sets and allow for understanding emergent patterns, larger systems and interactions, and smaller/ quieter signals amidst all kinds of variability. At the same time, there are special issues in dealing with these larger data sets so that the findings make sense, and that issues of privacy, ethics/ use of data, ownership of data, are addressed.

I guess I believe that deepening people's understanding of data in all its forms is really an essential feature of what it means to be a responsible citizen nowadays, and that Earth and space data are an important part of that mix. But I'm not sure that I know yet about the role of convergence of these specific types of data.

What do you believe is the most important outcome of a multi-sector collaboration among academia, industry, and government in the convergence of Earth and space data?

The workshop didn't really address the role of industry in this collaboration, and did include the importance of educational actors who are not part of academia. That said, and given the limitations about convergence expressed above, I think a multi-sector collaboration can build on the different resources that these different actors bring to build tools, methods, and infrastructure for sharing and working with and educating about data.

Module 2: Answer

Do you believe that the convergence of Earth and space data is important to data research and education? If so, why. If not, what ranks as more important?

I am not convinced that this convergence is necessary to support either research or education. In the example ideas that we developed on Friday afternoon, I did not hear any ideas in either research or education that were so compelling and important that I felt like, "Yes! We must to this now!" Rather, I was left with the impression that people were struggling to find ways of combining Earth and space data in ways that made sense. This may very well have been partially because we were working with hastily-formed teams under a ridiculously tight time frame, to do a job that would typically involve weeks or months of conversation.

As far as what ranks as more important – right now, we are struggling just to develop a clear understanding of how best to teach and learn with data. The rapid integration of machine learning and AI technology is going to be changing the game as we go. And then we have the issue of how to fit all of this into existing education structures. Do we change the way we teach biology, chemistry, physics, Earth science, across the board? Do we replace "Statistics" with "Data science?" Or maybe just convince folks to ditch Calculus in favor of something students are more likely to actually need? When viewed through this lens, the question of, "How about if we combine two different scientific disciplines and start teaching about that new, converged field?" seems far from low-hanging fruit.

What do you believe is the most important outcome of a multi-sector collaboration among academia, industry, and government in the convergence of Earth and space data?

As worded, this sounds like a solution in search of a problem. So I'm going to turn it around as a thought experiment: Climate change is an existential problem for the human species, and life on Earth in general. There is mounting evidence that we are causing the next major extinction event on our planet, and if we are to continue to exist beyond the fleeting moments since Homo sapiens emerged (in geological time), we must rapidly extend our habitable zone beyond that of our native home. This will involve a new level of understanding of those elements of Earth that are critical to human existence; and of the galactic neighborhood in which we might be able to extend our footprint. In order to gain this understanding, we must establish a unique collaboration among academia, industry, and government to understand Earth and space in a completely new way – as citizens of a larger galaxy rather than as earthlings.

Module 2: Answer

Do you believe that the convergence of Earth and space data is important to data research and education? If so, why. If not, what ranks as more important?

I believe it is important because of two reasons. First, earth and space data are relevant to tackle societal problems like climate change and to answer fundamental questions for humanity such as if we could inhabit other planets. Furthermore, I think that they are not only related on research topics, but also in techniques (e.g. algorithms they use, analysis methods, etc.) that could benefit very much of convergence. Second, I think they are relevant for education both because of the data techniques they use, but also because they can inspire children and youngsters into science.

What do you believe is the most important outcome of a multi-sector collaboration among academia, industry, and government in the convergence of Earth and space data?

I think there could be a lot of different important outcomes. The Data Observatory is an example of these kind of collaboration and, if it incorporates climate change data, it will precisely an example of convergence that aims to generate benefits for society, research, industry, etc.

Module 2: Answer

Do you believe that the convergence of Earth and space data is important to data research and education? If so, why. If not, what ranks as more important?

I believe it is important because of two reasons. First, earth and space data are relevant to tackle societal problems like climate change and to answer fundamental questions for humanity such as if we could inhabit other planets. Furthermore, I think that they are not only related on research topics, but also in techniques (e.g. algorithms they use, analysis methods, etc.) that could benefit very much of convergence. Second, I think they are relevant for education both because of the data techniques they use, but also because they can inspire children and youngsters into science.

What do you believe is the most important outcome of a multi-sector collaboration among academia, industry, and government in the convergence of Earth and space data?

I think there could be a lot of different important outcomes. The Data Observatory is an example of these kind of collaboration and, if it incorporates climate change data, it will precisely an example of convergence that aims to generate benefits for society, research, industry, etc. Furthermore, convergence among disciplines and sectors can foster creative thinking and enable new ways of educating, conducting research and generate economic and social development.

Module 2: Answer

Do you believe that the convergence of Earth and space data is important to data research and education? If so, why. If not, what ranks as more important?

Yes, because earth and space sciences clearly share common challenges that need to be solved and researchers/practitioners in both fields can benefit from solutions.
What do you believe is the most important outcome of a multi-sector collaboration among academia, industry, and government in the convergence of Earth and space data?
Greater efficiencies in dataops.

Module 2: Answer

Do you believe that the convergence of Earth and space data is important to data research and education? If so, why. If not, what ranks as more important?
Convergence of Earth and space data is important to data research and education is important since it builds on existing data, tools and curricula that can be implemented across various disciplines
What do you believe is the most important outcome of a multi-sector collaboration among academia, industry, and government in the convergence of Earth and space data?
The ability to support the Framework for Science education, in particular: 3D learning (Disciplinary Core Ideas + Crosscutting Concepts+ Science and Engineering Practices) and Coherence (building and applying ideas across time)

Module 2: Answer

Do you believe that the convergence of Earth and space data is important to data research and education? If so, why. If not, what ranks as more important?
Convergence would be useful where justified. Can't assess relative importance of different approaches outside use cases. Demonstrating methodological consistency across space and earth data analysis would be very useful for education. There may be a few research cases where methodological cross-pollination would be useful, too.
What do you believe is the most important outcome of a multi-sector collaboration among academia, industry, and government in the convergence of Earth and space data?
Compendium of convergence use cases, success stories, technologies and ongoing projects.

Module 2: Answer

Do you believe that the convergence of Earth and space data is important to data research and education? If so, why. If not, what ranks as more important?
We are facing complex problems that only convergence of data from Earth systems science and space will need to solve in order to retain a viable living environment on Earth. From an efficiency standpoint, it makes more sense to solve the data storage, handling, analysis etc together than separately. The data will be shared back and forth between these disciplines and many others, making sure the pathways between are clear will help with solving of problems that require convergence of multiple disciplines and Big data strategies.
From an educational perspective, we will need to teach both basic data literacy and the new techniques of big data analysis to our upcoming workforce (k – graduate school).
What do you believe is the most important outcome of a multi-sector collaboration among academia, industry, and government in the convergence of Earth and space data?
Borrowing from ESIP, the outcome of multi-sector collaboration is to "leverage their collective expertise and technical capacity to address common challenges related to Earth science data." Each sector brings unique strengths to the conversation and applications resulting from convergence of Earth and space data – research into what it means, stream-lined solutions borrowed from other disciplines reapplied to E-S realms, the development of global data-sharing policies, democratization of data and education to all people, etc. As we learn more about what it takes to have a livable environment on newly discovered planets of our galaxy, this knowledge can be applied to our understanding of Earth's environment, perhaps in surprising ways.

Module 2: Answer

Do you believe that the convergence of Earth and space data is important to data research and education? If so, why. If not, what ranks as more important?
Yes, we have lived in the silos for far too long. There are too many important questions about human survival that could be helped by a convergence approach

What do you believe is the most important outcome of a multi-sector collaboration among academia, industry, and government in the convergence of Earth and space data?
Everyone can benefit from each other's research, try new ideas and tackle complex problems

Module 2: Answer

Do you believe that the convergence of Earth and space data is important to data research and education? If so, why. If not, what ranks as more important?

Yes. The convergence of Earth and space science data is valuable for improving understanding of our planet and its relationship with other planets and planetary systems.

What do you believe is the most important outcome of a multi-sector collaboration among academia, industry, and government in the convergence of Earth and space data?

Multi-sector collaboration on the convergence of Earth and space science data has the potential to improve our understanding of Earth and space systems and offers the potential to improve education and research on Earth systems, planetary systems, and the relationships between Earth and planetary systems.

Module 2: Answer

Do you believe that the convergence of Earth and space data is important to data research and education? If so, why. If not, what ranks as more important?

I can clearly see the value within K-16 education. I also have noted in the past the techniques astronomers use to sample, acquire store and reduce data are applicable to other disciplines.

What do you believe is the most important outcome of a multi-sector collaboration among academia, industry, and government in the convergence of Earth and space data?

Removal of access barriers to data, tools and shared curriula

Module 2: Answer

Do you believe that the convergence of Earth and space data is important to data research and education? If so, why. If not, what ranks as more important?

Yes

What do you believe is the most important outcome of a multi-sector collaboration among academia, industry, and government in the convergence of Earth and space data?

Data standardizations and pipelines to educational institutions, both formal and informal.

Module 2: Answer

Do you believe that the convergence of Earth and space data is important to data research and education? If so, why. If not, what ranks as more important?

Yes to both. For data research, this convergence supports better understanding of both (all) data sets. The mingling of good minds and processes is beneficial to both (all) disciplines.

For education, this convergence supports the Next Generation Science Standards (NGSS) standards of crosscutting concepts (concepts that appear in all scientific fields) and scientific practices (such as asking questions, using models, analyzing data) that all scientists engage in. It has the potential to provide examples for teachers to use when implementing NGSS in classrooms and informal education settings.

What do you believe is the most important outcome of a multi-sector collaboration among academia, industry, and government in the convergence of Earth and space data?

A more scientifically aware and educated human population, allowing them to make personal and organizational choices leading to sustaining human life on this planet.

Module 2: Answer

Do you believe that the convergence of Earth and space data is important to data research and education? If so, why. If not, what ranks as more important?

Yes, but it is important to recognize the distinctions between research and education. Convergence for research seems to be the first big challenge, followed closely by convergence for education with research data. Also, researchers need strong incentives to participate in data literacy initiatives (for example, the first priority of academics is to publish research in the important journals of one's own field, and participating in curriculum-focused initiatives such as NEREID would likely need to fall into the category of service).

What do you believe is the most important outcome of a multi-sector collaboration among academia, industry, and government in the convergence of Earth and space data?

Development of shared infrastructures

Module 2: Answer

Do you believe that the convergence of Earth and space data is important to data research and education? If so, why. If not, what ranks as more important?

I do believe that convergence of Earth and space data is important to data research and education. Because of the similarities in data parameters and qualities there is a natural opportunity for convergence in Earth-space sciences with data. In addition, leaners are very interested in Earth-space science topics providing an opportunity to capitalize on this interest and build the next generation of big data explorers, and a data literate society. Finally, Major challenges facing the planet naturally sit in the Earth-space sciences domain, including, climate change, energy production, water resources, origins and evolution of life, etc.

What do you believe is the most important outcome of a multi-sector collaboration among academia, industry, and government in the convergence of Earth and space data?

Collaboration between academia, industry, and government in the convergence of Earth-space data is very important. It is government that establishes policies and programs that can influence education and workforce development nationally and internationally. In addition, government funds can often be used for high risk/high yield opportunities, advancing practice and technology in collaboration with academia and industry, to a point where industry can pick it up and refine it to achieve maximum impact. Academia of course plays a key role in getting these innovative practices and tools integrated into their education and training programs to ensure the workforce is prepared to take advantage of these innovations.

Module 3: Developing a Community of Practice

Please reflect on the long-term big data literacy initiative as discussed in the workshop and indicate your level of agreement with the following statements. What value do you feel you would bring to a long-term Earth-space data science convergence initiative? (Note: The word value here can be interpreted broadly, to include skills, relationships, and resources)

Module 3: Answer

I have a solid understanding of the Next Generation Science Standards and would like to be part of using this convergence to benefit science teachers looking for ways to improve their teaching.

The work of NEREID can be part of ESIP-Education's work to promote teacher professional development through workshops and educator collaborations.

I believe that my experience with both curriculum development and teacher professional development may be an asset to the group.

Module 3: Answer

Education / influencing public opinion

Module 3: Answer

My primary potential contributions are:

1. a focus on understanding how novices (K-12 students; community members) learn about data and statistics and how that's similar and different with "Big Data" and with Earth and space science data;
2. an ability to create curriculum, professional development materials and methods, and help to design software tools to support inquiry and learning about data, including substantial prior work with educationally oriented data and statistics software such as Fathom/ TinkerPlots/ CODAP;
3. curiosity and creativity about data, the sciences, as well as systems and communities of divergent actors and how they can produce synergistic results; and
4. a concern about ethics, social justice, and privacy, and how those can be protected through education and policy, especially as the power of Big Data (and associated AI and machine learning) has been developing beyond the existing policy/ regulatory environment that was not designed with its capacities in mind.

Module 3: Answer

I felt that both Jim Hammerman and myself brought a unique (and similar) perspective to the work, and either or both of us could make a valuable contribution to shaping the educational goals and objectives of it moving forward. The ultimate value of this work to me and my field, however, will depend entirely upon what directions the initiative takes, and how its goals are prioritized. I don't feel that I have a clear enough sense of this right now to comment one way or the other about its value as an opportunity.

Module 3: Answer

I bring critical insights and relationships that can improve the diversity, inclusiveness and accessibility of earth-space science.

Module 3: Answer

I bring a perspective on complexity and a background in socio-environmental issues

Module 3: Answer

I think I can contribute with relationships and knowledge about innovative initiatives like the Data Observatory. I've been researching the astronomy ecosystem in Chile and how it can produce spillovers into other domains (culture, infrastructure, talent, etc.) and it has a lot to do with convergence with other sciences and data.

Module 3: Answer

We will expand the Global Forest Link (a successful STEAM learning program) model that combines citizen-based framework for ground-truth data collection and scholarly analysis and remote sensing satellite data to include additional data science learning modules based on existing Earth-Space data collections and analysis tools.

Module 3: Answer

As a pedagogical expert, I have spent my career translating scientific and data concepts to educational environments. The convergence project presents a challenge to stretch beyond the current educational practices to develop innovative approaches to teaching about big data to solve E-S problems.

Module 3: Answer

Resources and exploratory tools that can be used for space and earth data discovery and exploration; expertise developing infrastructure for cross-domain research; visual analysis tools that can be used in space and earth science education and research, and a visual gateway into [spatial] data science tools in Jupyter

Module 3: Answer

I can contribute to a long-term Earth-space science data initiative by applying what we have learned through experiences in Earth systems informatics to the study Earth-space science data.

Module 3: Answer

My networks, my skills as a education developer, my long-standing interaction with researchers

Module 3: Answer

Librarianship perspective (long-term management and assessment of information); Experience with cyberinfrastructure and long tail data

Module 3: Answer

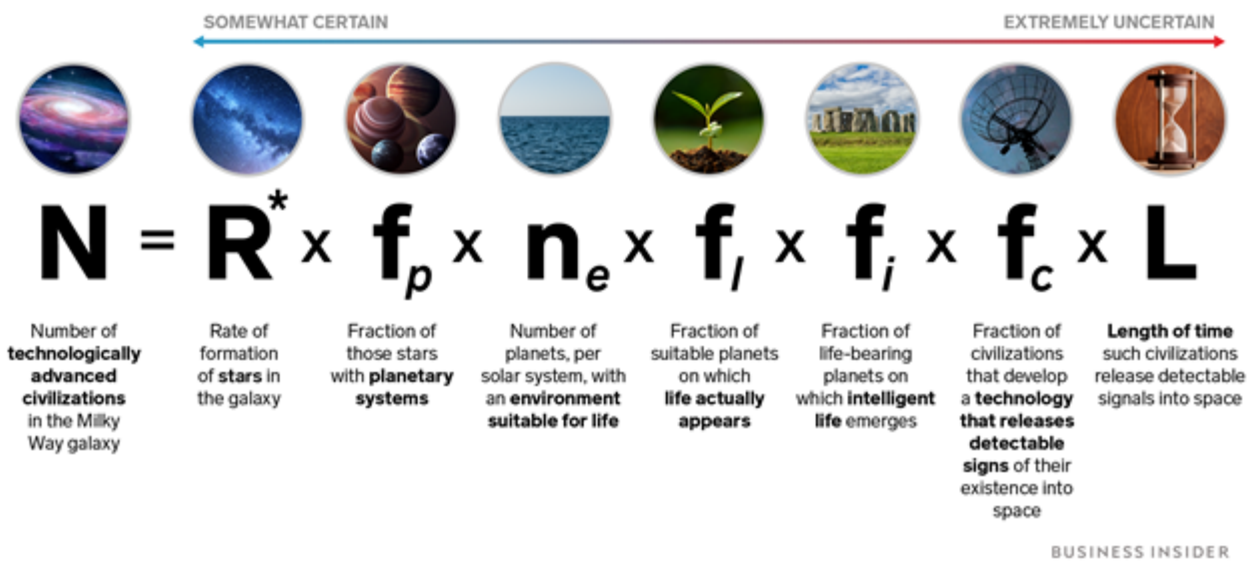
I believe I can bring my knowledge of earth and space/astronomy education and research to the convergence initiative. In addition, through my work, I can connect others in the geosciences, space/astronomy sciences, and data science to the convergence initiative, and demonstrate the power this convergence has to advance disciplinary and interdisciplinary research

Appendix F: Drake Equation Challenge

The Drake Equation Challenge – NEREID 2019

What do we need to know about to discover life in space? How can we estimate the number of technological civilizations that might exist among the stars? While working as a radio astronomer at the National Radio Astronomy Observatory in Green Bank, West Virginia, Dr. Frank Drake conceived an approach to bound the terms involved in estimating the number of technological civilizations that may exist in our galaxy. The Drake Equation, as it has become known, was first presented by Drake in 1961 and identifies specific factors thought to play a role in the development of such civilizations. Although there is no unique solution to this equation, it is a generally accepted tool used by the scientific community to examine these factors.

-- Frank Drake, 1961



Your Team Challenge

Within your team, think about N above that we are trying to solve. Consider the diverse sources of data you know about today, and ones that may be available in the future. Develop your plan to solve for N or at least better constrain N.

Your team will be asked to present your plan to the group Friday evening.

Orange Team Slides: https://drive.google.com/drive/folders/1jcy3yI9KjLvTtQlOHhcgdBzSpx34BPJW

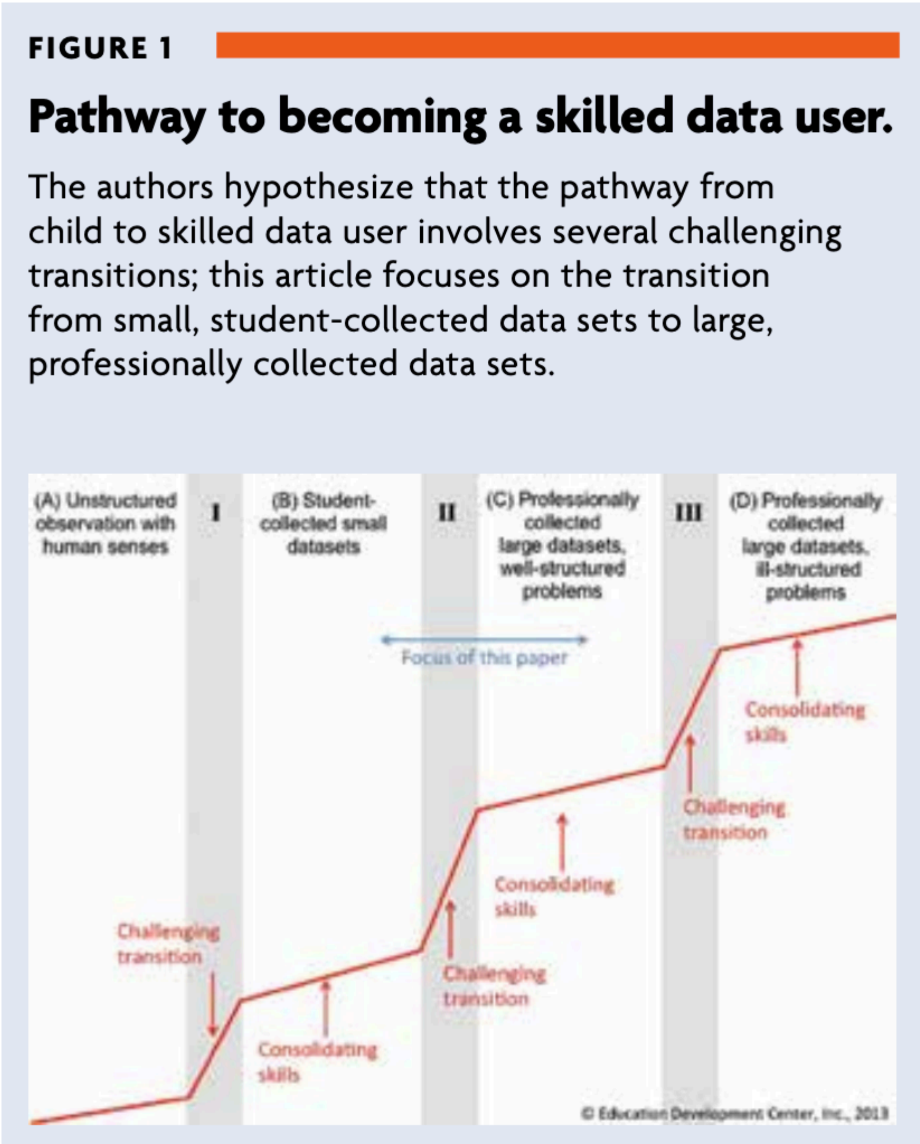
Green Team notes:

- How to measure data in forms we do not currently have
- Carbon based life - go back in time
- Factor n(e) - what is habitable is much bigger
- Factor f(l) is really interesting right now e.g. water bear factors
- Get rid of factors!

- When you lose the need for a form of communication the form of communication stops. However without the written word we would no longer know about constellations.
- Pollution as indicator of factor f(c) - detection of heavy elements
- Combination of elements as proxy for civilization at a certain phase
- Worlds with no climate swings could be hugely important to L - or is the opposite true??
- Engage other thinker - science fiction writers, artists

Red Team notes:
- Focus was mainly on FI, FE, and FC
- Belief that we need to include a social factor (fS) because intelligence does not evolve into communication technology without a social system.
- How often does the cognitive niche and niche construction happen?
- What are the pressures to develop complex language and communication conventions
- There is no incentive to develop detectable signal technologies without complex language and the cognitive niche.
- Intelligence does not by itself mandate sophisticated language (ex. Dolphins, Whales, etc.)

Appendix H: Resources Vetted by Participants



Kastens, K.A., Krumhansl, R., and Baker, I. 2015. Thinking big: Transitioning your students from working with small student collected data sets towards ''big data.'' Science Teacher, 82:25– 31.

Astronomy Data Archives: Ease of use 1=simple, 2=moderate, 3=difficult

SkyView (1) - https://skyview.gsfc.nasa.gov/current/cgi/titlepage.pl) SkyView is a Virtual Observatory on the Net generating images of any part of the sky at wavelengths in all regimes from Radio to Gamma-Ray. NOTE: There is a non-Astronomer interface.

Science & Data Center for Astrophysics & Planetary Sciences (2) – https://www.ipac.caltech.edu/projects/category/archives-and-data-science IPAC operates three of NASA's six Astrophysics Data Centers: IRSA, the NASA Exoplanet Archive, and NED. IPAC also serves the community by supporting the Keck Observatory Archive, and the ZTF time domain survey. IPAC data science activities seek to develop tools and techniques to facilitate archive exploration and scientific discovery in the era of big data.

SIMBAD Astronomical Database (2) - http://simbad.u-strasbg.fr/simbad/ The purpose of Simbad is to provide information on astronomical objects of interest which have been studied in scientific articles. Simbad is a dynamic database, updated every working day. It provides the bibliography, as well as available basic information such as the nature of the object, its coordinates, magnitudes, proper motions and parallax, velocity/redshift, size, spectral or morphological type, and the multitude of names (identifiers) given in the literature. The CDS team also performs cross-identifications based on the compatibility of several parameters, in the limit of a reasonably good astrometry.

Sloan Digitized Sky Survey (1) - http://www.sdss3.org/ Massive spectroscopic surveys of the distant universe, the Milky Way Galaxy and extra planetary system.

NASA/IPAC Extragalactic Database (1) - https://ned.ipac.caltech.edu/ Database of objects outside the

Spitzer Space Telescope Archive (2) - https://sha.ipac.caltech.edu/applications/Spitzer/SHA/

From Joseph Kerskii https://spatialreserves.wordpress.com

An example we could point to ... https://www.google.com/amp/s/phys.org/news/2017-04-scientists-dung-beetles-milky.amp
https://www.forbes.com/sites/alexknapp/2011/04/20/do-whales-use-the-stars-to-navigate/#428d1fb09152

ESIP: https://www.esipfed.org

Your phone can be a cosmic ray detector. https://crayfis.io/about

SuAVE: Survey Analysis via Visual Exploration; http://suave.sdsc.edu/

Earth Cube Data Discovery Studio http://datadiscoverystudio.org/geoportal/#searchPanel
NSF's BALTO - Earth Cube's Brokered Alignment of Long-Tailed Observations https://www.earthcube.org/group/brokered-alignment-long-tail-observations-balto

OPeNDAP & OPeNDAP's Hyrax https://www.opendap.org/

Sonifier https://afterglow.skynet.unc.edu/login

Kubernetes (K8s) open-source system for automating deployment, scaling, and management of containerized applications. https://kubernetes.io/

Container management Kubernetes (K8s) is an open-source system for automating deployment, scaling, and management of containerized applications. It groups containers that make up an application into logical units for easy management and discovery. Kubernetes builds upon 15 years of experience of running production workloads at Google, combined with best-of-breed ideas and practices from the community.
https://kubernetes.io/

http://www.data2dome.org/

LOCUS Level of Conceptual Understanding of Statistics - assessment tool for statistical learning

YouTube VR360 Live

My NASA Data

Zooniverse Largest on-line citizen science program in the world https://www.zooniverse.org/lab/

Accelerator SAS extension for Chrome browser allows non-visual interaction with data

Oceans of data case study re best practices: http://oceansofdata.org/our-work/visualizing-oceans-data-ocean-tracks-%E2%80%93-case-study

CyVerse.org Platform for sharing data

SEDAC

SEDAC Hazards and Population Mapper (mobile app)

From UNAVCO https://www.unavco.org/software/visualization/GPS-Velocity-Viewer/GPS-Velocity-Viewer.html

From Concord Consortium CODAP, SAGE Modeler Data cleaning/management and localization

Compelling narratives for climate change education have a strong local component. We need a tool/process/institution to acquire both global and local data sets, clean them, and allow them to be localized to an appropriate scale of granularity for the narrative.