Biological and Environmental Research Advisory Committee
(BERAC) Meeting Minutes
October 13-14, 2022
Remote Access Meeting

BERAC Members Present

Bruce Hungate, Chair
Caroline Ajo-Franklin
Cris Argueso
Sarah Assman
Bruno Basso
Sen Chiao
Leo Donner
Matthew Fields
Robert Fischetti
Ann Fridlind
Ramon Gonzalez
Jorge Gonzalez-Cruz
Randi Johnson
Kerstin Kleese van Dam
Sonia Kreidenweis
Xiaohong Liu
Maureen McCann
Gerald Meehl
Gloria Muday
Dev Niyogi
Himadri Pakrasi
Kristala Jones Prather
Patrick Reed
Gemma Reguera
Jeremy Schmutz
Daniel Segrè
Karen Seto
Matthew Shupe
Huimin Zhao

Guest Speakers

Asmeret Berhe
John Hill
Taipo Schneider

BERAC Speakers

Ann Fridlind
Maureen McCann
Patrick Reed

Other Contributors

Tom Arrison

Designated Federal Officer

Tristram West

Others

Todd Anderson, Biological Systems Science Division Director, Office of Biological and Environmental Research (BER), Department of Energy Office of Science (DOE SC)
Gary Geernaert, Acting Associate Director, BER, DOE SC
Dan Stover, Earth and Environmental Systems Sciences Acting Division Director, BER, DOE SC
Holly Holt, Science Writer, Oak Ridge Institute for Science and Education (ORISE)

Approximately 246 others were in attendance during the course of the meeting.

All presentations are posted to the BERAC website: https://science.osti.gov/ber/berac/Meetings

Thursday, October 13, 2022

Introduction and BERAC Roundtable

BERAC Chair, **Bruce Hungate,** called the meeting to order at 11:30 a.m. Eastern Time, welcomed attendees, and compared the enthusiasm and connection among BERAC members in their goal to advance science to that of a soccer team at kick-off. Members were encouraged to listen, learn, and engage in discussion as a teammates.

Three out of four BERAC meetings will be conducted via Zoom to reduce travel emissions. Both in-person and virtual meetings offer valuable feedback. The next meeting will be in person in the spring of 2023.

BERAC Roundtable, Dr. Bruce Hungate, BERAC Chair

Hungate discussed broad abiotic and biotic connections influencing the global carbon problem, with systems sometimes reacting in surprising ways. Meta analyses are necessary to unravel these complex interactions.

Committee members were invited to share BER-related thoughts. Discussion traversed several topics, including collaboration to advance science; international competitiveness; significant data and knowledge gaps and the need to integrate information across scales; increasing urgency to reduce human impacts on climate; increased cost-effectiveness of research pipelines; the role of AI/ ML in research; opportunities in biomanufacturing; workforce needs; and the need to imbue diversity, equity, and inclusion (DEI) into all aspects of science. Points are given below.

- The National Academy of Sciences (NAS) has launched a committee to foster U.S.-China scientific engagement to identify mutually beneficial opportunities for collaboration and cooperation related to climate change. This engagement includes the Chinese Academy of Sciences. Also, a new Climate Security roundtable directed by Congress has the goal to help the U.S. Intelligence Community understand climate impacts relevant to national security.
- Passage of the Inflation Reduction Act (IRA) is providing significant funding to DOE to study non-fossil fuel energy sources, along with money for the National Oceanic and Atmospheric Administration (NOAA) and other agencies to study the climate change problem.
- The National Science Foundation (NSF) is funding Regional Innovation Engines through the new Technology, Innovation, and Partnerships (TIP) Directorate.
- Planning for a second Multidisciplinary drifting Observatory for the Study of Artic Climate (MOSAiC) conference is underway.
- The Center for Biofilm Engineering at Montana State University recently acquired state-of-the-art equipment to visualize live biosystems at increased resolution temporally and spatially.
- The Advanced Photon Source (APS) will be shut down on April 17, 2023 for 12 months. Coordination with other light sources, including international partners, will mitigate impact and streamline access to other facilities. An announcement about beamtime access is expected.
- Within the next year, the cost of sequencing will drop by a factor of three while throughput will increase by 3-5 times. From a historical perspective, about 20 years ago, the cost to sequence one billion base pairs was $1M. Next year, the cost will be $1-2 for the same amount of data. Solving difficult global problems in natural systems, carbon cycling, carbon sequestration, and modifying agricultural systems to improve sustainability reducing inputs and greenhouse gas outputs requires the ability to synthesize and integrate data across all areas of biology. It will be important to develop ways to integrate across BER and DOE to make sense of these large quantities of data.
- A member cochaired a workshop last august on Artificial Intelligence and Machine Learning for BioEnergy Research (AMBER) addressing challenges and opportunities to apply AI to bioresearch and biodesign. The report is due in December 2022.
- AI/ML have a role to play in addressing challenges in assembling information across scales.

- A member discussed two aspects of AI that offer great potential. When data is scarce, there is potential to adapt information from related areas — for example, using natural language processing models to search journal articles — to prime models. Secondly, AI can make digital twin models faster so they require less time and compute power and therefore use less carbon. A recent effort replacing solver models with ML models achieved similar accuracy.
- There are opportunities to incorporate AI/ ML into synthetic biology pipelines to improve international competitiveness. However, there is a lack of training and funding for computational scientists. Focus is needed to build training and funding systems.
- Workforce support is needed from graduate and postdoctoral levels through early- to mid-career levels. Many new graduates want to develop translatable skills because they are already thinking about leaving science. It is vital BER create stable career paths for junior scientists. There are unexplored partnership opportunities for DOE and other agencies to reframe training to include management and soft skills. PIs also have an obligation to support student career interests on an individual basis.
- Rice University will soon inaugurate its first immigrant, first African American, first Civil Engineer to be its 8th president, a direct reflection of how the face of science and higher education is changing. Reginald DesRoches brings a strong scientific vision for research, with keen interest putting forth new programs on energy and the environment.
- A BERAC member's role as an editor for a scientific society's flagship journal has highlighted how DEI must be prioritized through data-driven policies.
- Attention was called to a new land sampling model for precision agriculture and optimal utilization of marginal croplands. Another member called for decarbonization of fertilizer which is currently generated through an energetically intensive process and represents approximately 1/3 of the greenhouse gas emissions generated by agricultural activities.
- BERAC members discussed research to improve crop defenses against climate change and pathogens. Research touched on studies in *Arabidopsis*; and a new rice mutant showing improved drought tolerance.
- New research points to the role of circadian clocks in cyanobacteria photosynthesis and nitrogen fixation.
- A member is considering how to address a lack of standardization for studies and data related to microbial networks and recently published a perspective article on generating Findable, Accessible, Interoperable, and Reusable (FAIR) information in this area.
- A member discussed studies of the atmospheric microbiome and its interactions with aerosol physics and meteorology.
- Several members highlighted important interactions between urban environments and climate change, with discussion addressing urban heat islands; wind; an algorithmic approach to compare atmospheric data from urban and suburban areas, demonstrating the impacts of air quality, pollution, and heat waves on local people; studies in coastal environments; and a meta analysis revealing locations where cities create excess rainfall. Major challenges are presented in the need to integrate human dynamics, climate dynamics, and urban systems with regional systems to create tools for climate-ready cities and protect vulnerable populations. The Intergovernmental Panel on Climate Change's (IPCC's) Assessment Report-7 (AR-7) offers perspectives on urban environments. Urban impacts must be significantly reduced to meet the global warming target of two degrees. Our global population will increase by 2.5B people in cities and towns in 30 years. Our ability

to mitigate urban impacts and develop green and blue infrastructure will determine our success. Knowledge and opportunities are expanding through use of urban digital twins and the DOE's Urban IFL initiative. There are opportunities for multi-agency collaboraitons.

- A member called for a geoengineering workshop to address deep knowledge gaps in understanding atmospheric particles and climate change, especially across scales. These gaps limit geoengineering approaches to address global warming. Other significant knowledge gaps across scales persist for megadroughts and sea ice.
- The European Union's Destination Earth is a groundbreaking model enabling monitoring climate change impacts and testing interventions through digital twins. The U.S. is urged to develop comparable capabilities.
- Basic science in biomanufacturing must be translated to energy-efficient product and industry-oriented applications to realize a profitable bioeconomy that also addresses climate change. Example studies in this area were presented.
- The White House issued an executive order on Advancing Biotechnology and Biomanufacturing Innovation for a Sustainable, Safe and Secure American Bioeconomy.
- A NAS workshop on Successes and Challenges in Biomanufacturing is planned for October 2022, bringing international voices to the table.
- A recent major breakthrough in biomaterials found mutating a single gene can cause micron-sized bacteria to self-assemble to produce centimeter-scale materials.
- DOE must step up to the challenge of recycling rare earth elements; 80% of the resource market is currently held by China. There is a potential role for biomanufacturing in this area.

DOE Introductory Remarks, Dr. Asmeret Berhe, DOE Office of Science, Director

BER and other SC offices, guided by their advisory committees, support the DOE's scientific mission and broader research communities as well as the nation's success.

The bipartisan Creating Helpful Incentives to Produce Semiconductors for America (CHIPS) and Science Act tops SC's agenda, with authorization levels that can significantly enable many goals, including to drive frontiers of science across disciplines and to develop a strong 21st century Science, Technology, Engineering, and Math (STEM) workforce rooted in DEI principles.

It is important DOE reaches historically underrepresented and minoritized communities where they are. Several DOE initiatives have been launched and will be expanded, including Reaching a New Energy Sciences Workforce (RENEW); Funding to Accelerate Inclusive Research (FAIR); and Accelerate Innovations in Emerging Technologies (ACCELERATE). Peer Plans promoting inclusive and equitable research are a new requirement for all DOE solicitations and funding opportunity announcements, including those from national laboratories. Proposals for conferences must have an established code of conduct, including an anti-harassment policy and must submit a recruitment and accessibility plan for speakers and attendees, including recruiting from historically underrepresented groups. SC is providing webinars and resources to assist with plan development and review. SC welcomes community input.

Another priority is to ensure the community, policymakers, and citizens know of SC and why it matters. Publicly funded scientists have an obligation to ensure work is truly serving the public.

Behre is committed to ensuring that SC's capabilities are recognized and exercised in achieving a net zero CO_2 emissions economy by 2050. BER has responded to new legislation and executive orders related to this effort, including the delivery of 40% of overall benefits of certain federal investments to disadvantaged communities. BER supports this requirement with the Urban Integrated Field Labs (Urban IFLs). BER is also participating in the SC's Earthshot initiative to deliver transformative clean energy and climate solutions within a decade.

BERAC's International Benchmarking Report and the recent Committee of Visitors (COV) Review of the Biological Systems Science Division are important contributions given the incredibly competitive global landscape.

BERAC will receive a new charge to review existing and anticipated needs and capabilities in data management and supporting infrastructure.

Discussion

One individual sought Behre's thoughts on outreach to underrepresented communities at larger, established universities where individuals may also face barriers. **Berhe** clarified DEI efforts are not limited to Historically Black Colleges and Universities (HBCUs) or Minority Serving Institutions (MSIs), but seek to broaden participation from people that have not historically engaged at different institutions. Broadening participation is not limited to one type of institution. There is also focus on institutions that have not historically received proportionate public funding. Responsible stewardship of public resources means serving everyone.

A second comment expressed concern regarding the role of basic science in the Earthshots given their emphasis on applied research. **Berhe** emphasized cross-cutting efforts ensuring incorporation of basic science in the Earthshots. Deployment of this initiative is possible because of basic science supported decades ago.

In response to a question regarding engagement opportunities to expand DEI in Earthshot research for both researchers and impacted communities, **Behre** pointed to the Urban IFLs. Extended conversations are vital to understanding community needs and reaching compromises between science goals and community requirements.

Behre addressed an inquiry regarding funding provided through the IRA for climate change mitigation. The $1.55B to DOE will support ongoing projects and facility infrastructure; no funding was allocated to BER through the IRA. However, this funding will complete or significantly advance a number of SC projects, freeing budget resources for allocation to other things. More support may be delivered through elevated appropriations authorized by the CHIPS and Science Act.

Acknowledging the importance of responsibly stewarding public funds, an individual asked about potential for DOE, NSF, and the National Aeronautics Space Administration (NASA) to coordinate via a NAS-level joint report, similar to the NASA/ NOAA/ and United States Department of Agriculture (USDA) decadal surveys from space. **Behre** suggested BERAC could make such a recommendation for DOE support.

The level of destruction wrought by recent weather events in Puerto Rico and Florida demonstrates that climate has been manifesting in different ways and repeatedly, with vulnerable communities most affected. What is SC is doing to help vulnerable communities avoid damage in the near future and accelerate preparation? **Berhe** acknowledged communities most affected by climate are the least represented among those doing science and developing recommendations to policy makers. Simply put, SC is not hearing from or including these communities to the

degree it should. When stewarding basic research, SC's role is to make sure the community that is addressing issues is as diverse as can be. Underrepresentation is a dire issue.

News from BER, Dr. Gary Geernaert, Acting Associate Director

[Presentation posted]

Discussion

An individual offered congratulations for the realization of the Urban IFLs. This is a transformational DEI effort. All of BERAC is urged to support and engage in this program. **Geernaert** added many agencies are represented among BERAC members. All are needed to advance this initiative's goal. BER will be reaching out on all fronts.

Another individual commented advance notice of the need for partnerships in forming Urban IFLs would have been beneficial. How can outreach be done when there is not yet a funding commitment? **Geernaert** suggested dedicating time for communications during BERAC meetings or townhalls at scientific meetings. With respect to the Urban IFLs, waiting for an appropriation was a limiting factor, but this communications point is well received. The Urban IFL initiative will serve as a template for future integrative activities.

Hungate dismissed the meeting at 2:03 p.m. for a break and reconvened at 2:15 p.m.

Update on the Earth and Environmental Systems Science Division (EESSD), Dr. Dan Stover, Acting Division Director

[Presentation posted]

Discussion

When asked about active coordination with other agency efforts for Urban IFLs, **Stover** explained this effort is at the opening stage with awards just made. Many awardees have plans to engage with various activities and groups. There are ongoing discussions with NOAA, NASA, and the U.S. Forest Service on levering resources. Baltimore's IFL, for example, could leverage NSF's Long-Term Ecological Research (LTER) program. There are many interagency collaboration opportunities. **Geernaert** added BER briefed the U.S. Global Change Research Program and the White House's Urban Heat Working Group on the IFLs with the intent to solicit other agencies for collaboration. The National Institutes of Health (NIH) expressed great interest due to existing agency investments in urban regions. More agency engagement will be sought following the Urban IFL Principal Investigator (PI) meeting in November 2022. A BERAC member commented on alignment among interests from Urban IFLs, NIH, and NASA.

Update on the Biological Systems Science Division (BSSD), Dr. Todd Anderson, Division Director

[Presentation posted]

Discussion

When asked about using sorghum as a study model for biosafety and biosecurity, **Anderson** explained the goal is to create a template addressing an agnostic biothreat to guide DOE capabilities in response to a national emergency. This is an initial Biopreparedness Research Virtual Environment (BRaVE) effort, and future steps will likely focus on using broader portfolio capabilities, including the user facilities.

An individual sought comments on the implementation status of Behre's vision in BSSD. **Anderson** remarked bigger changes in FY23 revolve around FAIR and ACCELERATE efforts which do not include BSSD. However, BSSD is enacting comparable broadening participation this year. The Earthshots in which BSSD involved also include broadening participation components.

BERAC Subcommittee on International Benchmarking, Dr. Maureen McCann, National Renewable Energy Laboratory; and Dr. Pat Reed, Cornell University

[Presentation posted]

Vote and Discussion

BERAC voted unanimously to approve the U.S. Scientific Leadership Addressing Energy, Ecosystems, Climate, and Sustainable Prosperity Report. No one present abstained or voted against the report. **Hungate** requested that discussion focus on key report aspects, rather than a voting stance. Comments were complimentary of the effort to produce the report and emphasized the need for BER to use the report to advise DOE decision-making. Discussion covered BER's role in improving U.S. competitiveness through funding, collaborations, and shifting from plans to action; the impact of BER visibility on DEI, talent retention, and funding; and advising DOE on which issues to champion. The following summaries reflect discussion points.

International standing: Some individuals expressed that U.S. competitiveness is limited and may decline in the future as a result of increasing collaborative international research. Also, an individual highlighted a NASA report concluding that limited data and technology access limit innovation. However, there are serious risk-related considerations that can and should limit that access.

Investments in Research: There was a general sentiment that investments in BER and much of the U.S.'s scientific enterprise are not commensurate with the challenges faced by society. Funding for science has "flatlined" such that scientific investments are not keeping pace with economic growth. Several individuals stated that science investments should be scaled with the economy, while others suggested scaling to the size of the problem. DOE's funding is disconnected in terms of the scale of the problems versus Congressional funding, however funding is always impacted by political conditions. Funding opportunities for blue sky research, working with industry through public-private partnerships, and large team modalities for ambitious projects require additional consideration. One individual pointed out the level of funding and investments may differ based on whether BER's objective is to compete or to coordinate with others as the whole world rises. BER needs to continue horizon scanning to effectively inform DOE on what to strategically champion for competitiveness. On the other hand, consideration of how the U.S. can have world-wide impact through international collaborations is vital, and is facilitated through the creation of relationships among project managers across agencies.

COVID-19 Pandemic and Impact on the National Laboratory Workforce: The pandemic dramatically changed workforce culture, with greater impacts on younger researchers and women. In addition, the pandemic created a disconnect between decision-makers and those affected by the decisions. National laboratories are experiencing the same workforce retention challenges faced many U.S. employers. National labs cannot necessarily offer better pay, but they can bring a

sense of belonging as well as feeling valued and respected to employees. Maintaining employee engagement is critical to improving retention.

Lack of BER Visibility: BER is relatively invisible in the community. Some noted BER's scientific breakthroughs are visible, but not the association of the breakthrough with BER. This lack of visibility impacts talent retention, funding, and competitiveness. A second aspect of limited visibility is on the smaller scale of recognizing individual contributions. So much of BER's research is collaborative that individuals are somewhat lost in the crowd, with negative consequences for younger or junior researchers seeking to establish themselves. This unintended consequence of collaborative work can again affect recruitment and retention. Limited venues for funding blue sky research at national laboratories versus in academia also impact talent attraction and retention. Thus, it is important for BER to improve its visibility and program awareness by developing cohesive communication and marketing strategies internally and externally. Conveying successes will inform potential talent, decision-makers, and the public that BER is doing critical work that cannot be done elsewhere. Marketing should consider targets and use all tools, including social media.

DEI: The importance of DEI was threaded throughout discussion. There was a general consensus that workforce statistics do not show progress desired by current thinking. It is important to understand DEI across the spectrum, from feedback to high-level efforts. Furthermore, DOE must move beyond DEI plans to conducting specific actions. It is also important to be sensitive to those in the workforce who have the least access and are most vulnerable. One individual commented the U.S. may be ahead of the global curve in defining what is needed for DEI, but focus is primarily on representation, which is only a first step. In actuality, the U.S.'s performance is poorer than or similar to that of other countries in studied areas, such as the gender gap. Not all can sit at the table when there are barriers to entry.

Hungate dismissed the meeting for the day at 4:23 p.m.

Friday, October 14, 2022

Hungate convened the meeting at 11:00 a.m.

BSSD COV Report: BSSD Response, Dr. Todd Anderson, Division Director

[Presentation posted]

Discussion

A committee member called for a more structured process with the capacity to engage broader research communities in peer review than workshops for vetting new DOE capabilities and large-scale facilities. **Anderson** appreciated these remarks.

Conversation segued to whether the Small Business Innovation Research (SBIR) program coordinates "reverse exhibitions" at conferences whereby DOE technologies could be presented to industry. **Anderson** responded SBIR does hold meetings of funded projects where Program Managers are invited to view portfolio activities. BER has considered inviting BER-related SBIR

awardees to PI meetings to promote programmatic cross pollination. However, restricted in-person attendance presents challenges.

A member commented BER's three-year application cycle hampers diversification and engagement of early career researchers. **Anderson** responded BER must reach out to early career researchers; there is a mutual benefit in inviting these individuals to participate in review panels so they can see and gain a better understanding of the characteristics of winning proposals. In return, DOE gains a larger population of reviewer perspectives.

Another individual inquired about reducing disparities between reviewer's comments and scores. Research shows numerical scoring can introduce bias. **Anderson** said BSSD does place a heavy emphasis on comments and works with review panels to elicit critical feedback. Scores can represent an overall level of enthusiasm for proposals, and there is a need for balance.

When asked about issues related to the Joint Gene Institute (JGI), **Anderson** explained the COV was commenting on the range of capabilities JGI brings to the portfolio and ensuring those capabilities appeal to a broad user base and align with BER goals.

Briefing: SC Roundtable on Biopreparedness and Response, Dr. John Hill, Brookhaven National Laboratory

[Presentation posted]

Discussion

One BERAC member highlighted discussion of the need to make data easily available for different purposes. **Hill** agreed. This was a cross-cutting theme of the Roundtable, especially in addressing how to incorporate datasets that require sharing constraints.

When asked to provide lessons learned from the COVID-19 pandemic, **Hill** explained the National Virtual Biotechnology Laboratory (NVBL) stood up programs well, but they were short-lived and leveraged capabilities in place. A recommendation is to stand up experimental and computational capabilities now, so expertise is already working as a team the next time there is a national emergency.

One individual asked if there were discussions on the role of Urban IFLs and how the built environment will impact climate change. **Hill** answered there was not specific discussion on Urban IFLs. However, there was discussion around urban centers, epidemiological models, and complementarity with other agencies like the NIH. Another individual added climate change is very important to understanding disease transmission within and across species; urban models will be important for outcomes.

Conversation pivoted to coordination with other agencies such as the Department of Homeland Security (DHS) and the National Institute of Standards and Technology. Are there concrete plans for new funding programs to support interagency efforts? **Hill** emphasized DOE's role falls in the basic sciences only. The challenge is effectively and proactively relay BER capabilities to other agencies. Other agencies are aware of this roundtable and are eager to see the report, including DHS and the Office of Science and Technology Policy. Hill is unaware of interagency funding goals other than the $50M BRaVE act funding. Various BERAC members suggested exploring potential connections to the NIH Rapid Acceleration of Diagnostics (RADx) initiative and USDA. There are opportunities to work the latter on biopreparedness related to plant pathogens and bioenergy crops. Another individual noted BER just funded a center with BRaVE to examine plant pathogens at Brookhaven National Laboratory. The center plans to execute future work in a distributed, nation-wide fashion.

When asked about leveraging atmospheric science and modeling considering aerial viral transmission, **Hill** commented the NVBL was focused on the fate and transport of pathogens, airflow in buildings, and microbial resistance.

BERAC Science Talk: ***NASA GISS ESM Development from a Cloud Physics Perspective: Strategies and Recent Results*****,** Dr. Ann Fridlind, NASA Goddard Institute for Space Science (NASA GISS)

[Presentation posted]

Discussion

A BERAC member appreciated the value of this research in advancing the modeling of microphysics in climate systems. How does the E3 model move from the details of cloud microphysics to the issues at the macrophysics scale? What are the implications for instrument requirements to inform E4, especially for scales ranging from particles to 10 meters in the context of turbulence and inhomogeneous mixing? **Fridlind** responded that at present, there is a poor macrophysics scheme defining the subgrid variability of cloud microphysics. Current approaches include asking the climate model to reproduce the cloud environment that the Large Eddy Simulation (LES) produces, given boundary LES conditions. This approach has proven powerful for stratiform cloud systems and systems without a lot of mesoscale organization. At the LES scale, the problems of the subgrid scale are reduced, allowing focus on factors like liquid water path and radiative fluxes at the column-level. However, changes have been made to the macrophysics, and the degree to which the LES aligns with observations is known. Forward simulations are comparing results across different data sources. Tuning strategies employed for microphysics have been extended to the macrophysics using emulators. Simulations reveal the LES supposedly has resolved a lot of inhomogeneous mixing, however, much can be learned from comparing LES to observational data. Chamber design will be crucial and requires community input and debate.

Another member emphasized the importance of scale interactions in cloud microphysics and invited further comment on the need for instrumentation and possible observational strategies. **Fridlind** expounded on the value of a variety of data collection modalities ranging from airborne- to ground-based campaigns and integration with satellite data. There are many ways to collect data, but emphasis should be placed on improving scientific understanding of observational data and pushing the limits of high-resolution LES. This will provide a path to training climate models for many but not all cloud types.

An individual remarked the climate community agrees clouds are a huge issue, and doing this work is an important recommendation to BER. The world needs the modeling, instrumentation, and other aspects the speaker laid out. **Fridlind** appreciated these remarks.

Hungate dismissed the meeting for a break at 1:04 p.m. and reconvened at 1:20 p.m.

Discussion of New Charge, Dr. Bruce Hungate, BERAC Chair

Hungate presented a new SC Charge to 1) review existing and anticipated data management capabilities and supporting infrastructure relevant to BER; and 2) recommend a strategy for the next generation of data management and analysis within a unified framework. **BERAC** offered reactions, and ensuing discussion explored the following topics.

This subject of data needs and infrastructure is not new; its recurrent nature and the ongoing demand for solutions speaks to the magnitude of the challenge. It will be important to leverage related efforts at DOE SC Office of Advanced Scientific Computing and Research (ASCR) and other agencies while also learning from the shortcomings of previous attempts.

It is critical that BER specifically define the purpose, focus, and aims of the framework. Otherwise, the volume of information to be managed as well as the increasingly velocity with which it will be generated could become overwhelming. Goals may include providing fast access to large quantities of data; establishing mechanisms to access new data while integrating existing data; providing data that is scalable in a way that does not compromise security; and/ or generating hubs for new science collaborations and questions rather than creating another repository. Indeed, there are opportunities to consider how such a facility or framework could address scientific questions that currently cannot be answered by empowering users to integrate information across disciplines and scales without requiring individuals to have universal topical expertise. Findings from the International Benchmarking Report proposing the U.S. take a competitive leadership in key areas, however, may limit data sharing and integration.

The framework's goals will inform its structure, which may be built from scratch and/ or consist of junctures between collaborating entities across agencies with mechanisms to enable data sharing. Industry involvement through cloud services and other computational platforms may be important.

Beyond the framework, processes for ensuring data quality, formatting, and metadata inclusion must be defined. Data itself is becoming more complex as experimental and computational datasets in addition to computational models themselves are generated. However, managing data ontology can easily become an all-consuming activity, so standardization efforts must be targeted. The nature of future data generation and analysis may take unexpected paths as new technology, including machine learning (ML), and artificial intelligence (AI) come to the fore. A nimble approach will be key to adapting to technological advances. Implementation of Findable, Accessible, Interoperable, and Reproducible (FAIR) data principles will be paramount for tracking both data provenance and conferring recognition to contributors. Mechanisms to control data misuse must be considered to encourage researchers to contribute. In addition to community outreach, there must be a coordinated effort to work with journal publishers to encourage data deposition.

DEI considerations are foundational to meeting community needs. BER must seek broad community input through many platforms, including the composition of the Charge subcommittee itself. Infrastructural considerations must take variation in computational capacity at different institutions into account. Indeed, many MSIs and HBCUs lack computational power. Likewise, training in data access, analysis, and deposition must be provided to expand the scope of engaged researchers and institutions. Finally, it is important to consider how these activities will integrate with ongoing efforts to broaden participation in BER science.

Hungate asked members to think about key elements of subcommittee composition required to effectively tackle the charge. It is important to include others that are not on BERAC; the full community must be brought to bear. One year is a short delivery timeframe.

NAS Briefing: Automated Research Workflows for Accelerated Discovery: Closing the Knowledge Discovery Loop, Dr. Taipo Schneider, California Institute of Technology and NASA Jet Propulsion Laboratory (NASA JPL)

[Presentation posted]

Discussion

When invited to describe the process of creating the report, **Schneider** explained efforts began with an initial committee meeting, followed by a workshop which generated a number of use cases. Two years of committee meetings and consultations with the broad scientific community followed. **Tom Arrison** (NAS) agreed the workshop was the main mechanism for gathering information. The committee identified further areas requiring refining. Rather than direct specific agency budgets and initiatives, the purpose was to introduce ideas and facilitate intellectual change regarding the next stages of computing-enabled research. Crafting the message and appropriate recommendations took time as did reviews of reports.

A BERAC member considered the possibility of applying these methods for Earth system modeling. Accessing data sets that can be used in automated workflows is a challenge. What imperatives will a future data management system incorporate to enable optimal use of methods? **Schneider** commented observational data is currently not centralized. Future ensemble simulations will regularly produce data on the 100-PB scale, precluding archive storage and necessitating cloud storage. Ideally, both observational and simulation data will be provided through a central cloud location. However, cloud costs can be large. Generating a strategy for how to achieve these goals would be very helpful and accelerate work.

Another member disagreed, stating a cloud solution is not needed, and DOE cannot afford one. DOE has data centers that routinely handle transfer of large data volumes. The plethora of workflows generated by the community are a barrier to reproducibility. Is there an opportunity for setting quality measures? To what degree will workflows be autonomous? **Schneider** responded the goal is to develop tools that enable workflows to learn from data that builds on cumulative scientific knowledge. Tool development should be open; public accessibility will enable reproducibility. Scientists will remain indispensable to the process, especially in variable selection, making workflows more automated than autonomous.

Discussion, Dr. Bruce Hungate, BERAC Chair

Open discussion among BERAC members addressed several topics.

Several BERAC members discussed the benefits of exploiting *Arabidopsis* to meet DOE's and other agencies' goals. *Arabidopsis'* rapid growth rate and amenability to transformation makes it a tractable model for training students. The plant's low maintenance costs speak to its potential to support workforce diversity. All other plant genomes are annotated from *Arabidopsis,* and there is a niche for *Arabidopsis* to accelerate synthetic biology studies. This is relevant to the idea presented in the International Benchmarking Report of staking a leadership position in plant biosystems design, and the same concept translates to select microbe models. This should be recommended to DOE and BER. However, at present, it is hard to obtain DOE-funded proposals centered on *Arabidopsis* unless research is for proof of principle. Another individual, however, cautioned that while *Arabidopsis* has value, it is important to transform other plant species more closely related to other target crops. There is value in many models. Related to this dialogue is the need for a national plant transformation facility. DOE's focus on infrastructure makes it an ideal sponsor. If funded, such a facility could also provide grant programs to fund researchers' costs to use the facilities, broadening participation. There is currently an active dialogue within BER on the need for such a facility.

DOE is risk averse. Solicitations favor specific topics. Consequently, BER misses the opportunity to include diverse ideas with the potential to yield inventions.

Although BER has some digital twin-type efforts in E3SM with associated integrated assessment and MultiSector Dynamics (MSD) activities, BER needs to host a workshop on digital twins. Such a workshop would be timely as NASA and NAS are also hosting workshops on this topic. BER is well-qualified to undertake a leadership role in this area.

Another BERAC member suggested BER host workshop to determine how it can contribute to science of extremes in climate and weather. Further study of boundary processes and the ability to predict extreme events with less uncertainty is important for advising policy makers.

As part of the new charge, BERAC may consider advising block chain technology to track data ownership and how data is used.

On the subject of pandemic preparedness, some believe computation is all that is needed. However, experimentation is still critical to validate computations, and computations can increase experimental efficiency by using AI and ML to direct experiments. The capacity to synthesize compounds is also important for biopreparedness. Such infrastructure was not available for the COVID-19 pandemic.

While many want in-person meetings, BERAC may consider a hybrid option to be more inclusive of a wider audience.

Public Comment

None.

Hungate adjourned the meeting at 3:18 p.m.

Respectfully submitted on November 11, 2022
Maureen Leavitt and
Holly Holt, PhD
Science Writers, ORISE