Group #2: Community of Interest on the Future of Scientific Methodologies

|  |  |
| --- | --- |
| Date | November 2, 2020 |

|  |  |
| --- | --- |
| Participants | |
| Ravi Madduri | Xi Yang |
| Valerie Taylor | Nathan Urban |
| Deep Medhi | Bruce |
| Nami Ishihara | Kjiersten |
| Mike (ANL) | Rachana Ananthakrishnan |
| Daniel | Prasanna Balaprakash |
| Rich Carlson | Chin Guok |
| Josh Greenberg | Richard Gerber |
| Carlos X. Soto |

Contents

[1 Day One - November 2, 2020 3](#_Toc55484438)

[1.1 Discussion: Breakout 1 - Define the Scope of the Problem. 3](#_Toc55484439)

[1.2 Discussion: Breakout 2 - Implications of this Problem. 13](#_Toc55484440)

[1.3 Brainstorm: Day 1 Reflections 16](#_Toc55484441)

[2 Day Two - November 5, 2020 16](#_Toc55484442)

[2.1 Brainstorm: Breakout 3 - Signposts 16](#_Toc55484443)

[2.2 Brainstorm: Breakout 4 - Signpost Plausibility 19](#_Toc55484444)

1 Day One - November 2, 2020

1.1 Breakout 1 - Define the Scope of the Problem.

**The following participants have not been active:**  
Richard Gerber

**Question or instruction for the discussion:**  
Breakout 1 - Define the Scope of the Problem.  
The purpose of this session is to lay the foundation for the next 5 sessions. That is, each breakout group will define a key piece of technology, a new device, or methodology that would have an impact on how the labs/scientists operate. The details should include answers to the questions below.



**Sticky points:**

 Top Takeaways (5 points per participant)

* What is the problem, issue, technology, device, methodology?
  + Technologies that will democratize science. What may be the future of experimental facilities and discovery science if technologies that were formerly "big science" and "big facilities" become ubiquitous.
    - clarified data access and ownership... maybe akin to net neutrality (#20)
    - Another prerequisite for democratizing big science (and one relevant to computing) is the need for autonomous control. i.e. giving access to all scientists without having to specialize in the facility as well. (#21)
    - how does open source model looks like for facilities (#22)
    - AI could help by separating how something is done from what is the problem (#40)
  + Technologies that will change how we collaborate. What may be the impacts on science and scientists of being able to have interpersonal interactions in a virtual environment comparable to interacting in person.
    - Distant co-presence in physical facilities (#17)
    - What does home work environment need to look like? (#25)
      * Is it a home environment, or a different sort of office? (#28)
      * What would a local "telepresence facility" look like that would let you totally immerse in a remote space/facility/meeting? (#32)
  + New energy production, delivery and storage methods. How may fundamental changes in the energy sector drive change in experimental facilities.
    - I have little to no expertise here ... but the more we want to compute, access, and store data ... the more power we'll need to enable that. The end of oil will demand other energy sources... solar isn't sufficient without better batteries. Do we just assume necessity will drive access to more energy sources or... do we assume that the future represents energy access only for the privileged? (#23)
      * I think energy production, delivery and storage are keys for the future ideas to come to fruition (#42)
  + Scientist with each other
    - Teleconferencing with a holographic feature so that we'll feel like in the same room (#27)
    - Immersive realtime XR (#8)
      * Is there a meaningful difference b/w VR and AR for scientific use cases? (#12)
      * interaction with digital twins of physical apparatus/instruments (#15)
        + We probably need standards for interacting with digital twins (#38)
        + +1. General framing could be symmetrical standards for interaction with other humans and non-human AI (collaboration) and interaction with models/instruments/field sites (involving automation/digital twinning) (#41)
      * Interactive publication of data and results (akin to distill.pub) (#10)
        + Do you think this will still be an issue in 30 years? (#37)
      * Combine advances in VR with controlling physical lab resources (#29)
        + alleviates the need for wetlab/bench scientists to be in the lab to get their work done (#35)
    - Truly high-fidelity shared whiteboarding (#24)
      * this feels very short-term/tactical (#34)
  + Scientist with the public
    - It may be useful to start with defining what democratization of science. One definition would be: creating institutions and practices that fully incorporate principles of accessibility, transparency, and accountability. It means considering the societal outcomes of research at least as attentively as the scientific and technological outputs. It means insisting that in addition to being rigorous, science be popular, relevant, and participatory. This is the link where I got this from: https://issues.org/p\_guston-3/ (#14)
      * I like this: both how do we bring science to everyone, and everyone to science (#39)
    - Democratization - is this similar to commoditization - is commoditization what's needed to enable democratization? (#16)
    - for democratization is how to make technology accessible to everyone (easy, reducing the overhead, move from expert-driven to automated facilities) (#18)
      * One topic we focused on in the DOE 5G workshop was "democratize the radio edge for science". What that means was to provide open platform for edge computing, radio frequency and sensor innovations so that scientific user facilities can leverage a common infrastructure and integrate diverse edge technologies and use cases. (#31)
    - democratization means easier access so that the size of your institution doesn't become a factor (barrier) for a researcher's to use a facility (#19)
      * I like this concept ... convince Congress that just because we're DOE funded, doesn't mean we shouldn't have access to NSF or NIH resources if the projects make sense! (#30)
  + Scientists with science
    - Experimental "breadcrumbing" that would allow asynchronous collaboration across users of a facility (#33)
    - Robotic-assisted telepresence for collaborative work (#6)
      * what would this look like? (#36)
        + Some control and feedback interface would be needed at user's end, with a robotic operator or team of robots to assist in physical tasks. Likely AI assistance would be needed, particularly if a more intent-driven approach is desired (e.g. issuing commands like "collect these samples and get them ready for X") (#57)
    - Intent driven automated seamless (workflow) execution (e.g., Star Trek-ish computer queries) (#13)
      * Removing the "technology wizard" requirement. (#26)
      * Results orientated and not process focused (#43)
    - Workspace of the future: at home and in the office (#137)
      * Features and interaction tools: telepresence (holographic, AR/VR), remote operation (robotic), AI assistance (e.g. multi-user moderation) (#138)
* Who would develop it (basic research to advanced deployment)?
  + Technologies that will democratize science. What may be the future of experimental facilities and discovery science if technologies that were formerly "big science" and "big facilities" become ubiquitous.
    - Funding agencies in the US and elsewhere will have a role in funding activities that would foster democratizing science (#47)
      * need more $ to foster a culture of risk-taking for scientific applications (#50)
    - open sourcing the expertise (code, data, workflows, etc) (#48)
  + Technologies that will change how we collaborate. What may be the impacts on science and scientists of being able to have interpersonal interactions in a virtual environment comparable to interacting in person.
  + New energy production, delivery and storage methods. How may fundamental changes in the energy sector drive change in experimental facilities.
  + Scientist with each other
    - Learn from what folks have done already to identify what works and what doesn't... (#52)
  + Scientist with the public
    - If goal is to move specific expertise to a wide audience, need buy-in from the experts: share their knowledge, work on autonomous systems, devote their time/money to this purpose (#49)
    - Data (#54)
    - Need to educate taxpayers to demand access to the data (#55)
    - Technologies and policies that would enable to compute on private data (#58)
    - We need to be able handle deep fakes and disinformation (#59)
    - there is a lot of interesting work in industry for holographic telepresence and augmented reality - opportunities for partnership to apply this to science? (#44)
  + Scientists with science
    - Very likely industry will lead here. But DOE should get in early to help establish requirements and use cases (#45)
      * I agree - question is how to get in to help shape requirements through use cases (#46)
        + One unique motivator for DOE-specific collaboration and work requirements is the large scale scientific user facilities (#51)
      * If DOE needs are unique (e.g. science collaboration has needs that are different from more general collaboration), then DOE needs to be part of development (#53)
    - If we want to build a highly integrated end-to-end solution, all the DOE facilities involved will need to collaborate on a unified solution. (#56)
* Who would use it and what skills would they need to use it effectively?
  + Technologies that will democratize science. What may be the future of experimental facilities and discovery science if technologies that were formerly "big science" and "big facilities" become ubiquitous.
    - Data systems infrastructure (#60)
    - Hopefully everyone! Which means we need better science education (#61)
      * Do facilities need to have a stronger community engagement and outreach arm to better understand what matters to the public? (#66)
        + Both outreach and also a roll in education. (#70)
      * Out of scope for this breakout, but DOE should invest heavily in K-12 science education (particularly middle school age), and especially reach out further to underrepresented groups at this early age. (#68)
    - Domain scientists, facility users (we have to make it easy to use with minimum skill) (#65)
    - citizens will have easier access (not just limited to scientists) (#67)
    - Everyone! It might be useful to think about our education system and workforce development that would enable critical thinking so any one with basic education can ask the questions (#71)
      * +1 (#72)
    - Reducing the barriers to entry (#74)
  + Technologies that will change how we collaborate. What may be the impacts on science and scientists of being able to have interpersonal interactions in a virtual environment comparable to interacting in person.
  + New energy production, delivery and storage methods. How may fundamental changes in the energy sector drive change in experimental facilities.
    - Everyone (#73)
  + Scientist with each other
  + Scientist with the public
  + Scientists with science
    - Will be completely pervasive (like asking who would use a smartphone or laptop computer today) (#64)
    - Ideally we would want minimal specialized skills to use new tools effectively (the Star Trek approach) (#62)
      * +1 (#63)
      * yes - so the systems are intuitive to non-experts... like someone was saying previously, you don't need to know what's underneath (#69)
* When would it be expected to be in production use (N years in the future)?
  + Technologies that will democratize science. What may be the future of experimental facilities and discovery science if technologies that were formerly "big science" and "big facilities" become ubiquitous.
    - 5 years (#76)
    - Education component can be immediate. Universal access could be 10-20 years away (#78)
      * Disagree - we need to know what we're educating people on before we can design a curriculum (#79)
  + Technologies that will change how we collaborate. What may be the impacts on science and scientists of being able to have interpersonal interactions in a virtual environment comparable to interacting in person.
    - 5 years (#75)
    - Expect regular rollout of new technologies (particularly from industry providers). Major fundamental changes every ~10 years (#77)
      * e.g. paradigm shifts to heavy video/mobile use (current decade), possibly AR/VR usage (next decade), seamless AI and robotic assistants (~20 years) (#81)
    - Plausible for early adoption in specific cases within 5 years, definitely there within 10 years (#80)
  + New energy production, delivery and storage methods. How may fundamental changes in the energy sector drive change in experimental facilities.
  + Scientist with each other
    - This will come with a culture change that encourages collaboration...which is likely on a decadal time scale without some sort of serious shift akin to COVID influencing the sharing of data amongst scientists (#85)
  + Scientist with the public
  + Scientists with science
* Where, and how widely, would it be deployed?
  + Technologies that will democratize science. What may be the future of experimental facilities and discovery science if technologies that were formerly "big science" and "big facilities" become ubiquitous.
  + Technologies that will change how we collaborate. What may be the impacts on science and scientists of being able to have interpersonal interactions in a virtual environment comparable to interacting in person.
  + New energy production, delivery and storage methods. How may fundamental changes in the energy sector drive change in experimental facilities.
  + Scientist with each other
    - at the end-user, at the edge (facilities), at the compute centers (it will be distributed/federated) (#88)
    - New technology to enable scientific collaboration is available now - but is tricky to use (look at this meeting as an exemplar). There are so many - how do we collect data and information from the scientists to identify what works and what doesn't. (#89)
    - Deployed on every ones' computing device (#90)
    - Within traditional workspace as well as at home (#91)
  + Scientist with the public
    - Are orgs like DOE allowed to "advertise" what they do to the public? (#92)
    - open source data, experiments, models, hypothesis, reports, end-results (#93)
    - The more people the better! Try to reach everyone. (I'm picturing Gary Oldman) Goal is to both drive support for our research (everyone) and educate the next generation of scientists (youth, especially URGs). (#94)
      * EVERYONE! (#135)
    - Can we run effective data challenges that encourage the use of DOE or other data resources to address challenges? (#95)
    - Public databases maintained by federal agencies and other groups (#96)
    - Data and results publications (#97)
      * Translation of the publications into digestible formats that get folks excited or help them relate to what was done. Focus on meaningful impact. (#98)
    - Publication organizations, arxiv (#99)
    - How can we cut through all the noise and alternative facts so scientists become a trusted resource in our society? (#100)
  + Scientists with science
    - Within the DOE facilities and science collaboration groups (#86)
    - With smaller research groups (e.g. colleges, etc.) that otherwise cannot access big science, including under-represented groups. (#87)
    - Scientists will no longer need to be physically present at facilities to do work - so the process can be done at home :) (#101)
    - every step in science investigation (literature survey, related work, sharing knowledge) (#103)
* What is the setup time and/or process for using it?
  + Technologies that will democratize science. What may be the future of experimental facilities and discovery science if technologies that were formerly "big science" and "big facilities" become ubiquitous.
    - As quick turn around as possible. Think about "boot my personal quantum computer" vs "reserve my time slot in next month at the XYZ computer center". (#111)
      * quick usually translates to having a surplus of resources - can we get away from thinking about metrics like utilization? (#124)
  + Technologies that will change how we collaborate. What may be the impacts on science and scientists of being able to have interpersonal interactions in a virtual environment comparable to interacting in person.
    - The realistic and friction-free experience is the key. Future technologies should provide the truly immersive environment that eliminate all the current overhead. (#127)
      * Can we identify some of the current overhead challenges? (#130)
        + Like networking and processing capacity, data and infrastructure decoupling, layers of policy barriers between data and users etc. (#134)
  + New energy production, delivery and storage methods. How may fundamental changes in the energy sector drive change in experimental facilities.
  + Scientist with each other
    - Starts early with academic training... though hopefully the tools are intuitive enough (Star Trek style) to make this just a natural part of how we collaborate. Like how we started using the Internet to share ideas early and create forums for discussion. (#105)
      * Agreed. Also the incentive structure should be in place (#112)
        + What kind of incentives could we suggest? (#121)
    - Education and lining up incentives are important (#108)
    - cheap, fast, and accessible compute and data movement infrastructure (#118)
      * I'll add usable to this list! (#122)
  + Scientist with the public
    - Engaging with the public may require some specialized training and tools (#102)
      * basic idea of communicating complex ideas to general public will still be needed. (#119)
    - Educating the public will first require educating scientists on how to do outreach (#104)
      * Clone Anthony Fauci (#113)
        + +1 (#114)
        + +1 (#115)
        + +1 (#116)
        + More seriously - find the cadre of science communicators that can establish trust with the public. Or make this a career path that's highly valued. We need the public to trust scientists. (#117)

Completely agree. Science communicators of the past have been invaluable, but rare. We need to normalize training in science-to-public literacy. (#125)

* + - Educating/training scientists about the importance of public engagement, incentives? (#123)
      * +1 - translation of the scientific work to impact on the every day lives of folks - even if they might not see that impact immediately. Hate to say this, but we need a marketing branch for science. (#126)
    - Why science is exciting and supercool so that the next generation becomes interested. (#131)
    - We need a "clippy for science" that pops up across articles and gives folks guidance on what's being reported in the articles (#133)
      * could potentially be achieved if we worked on a science communication AI-bot (#136)
  + Scientists with science
    - Experimental facilities will have to be adapted to support new interaction modalities (possibly a lengthy and costly process), and new ones will have to be designed with this in mind (#106)
      * Scaling facilities and serve large number of scientists (breaking geographic broaders) (#132)
      * Completely agree. Facilities will have to focus on the interfaces to their resources... what will the analogs to the mouse and keyboard of the future look like (#109)
      * Other considerations also needed, particularly data security and validity (#120)
    - There will be a trade off between ease of use and complexity of the system, but this progress in the systems' ability can be incremental. (#107)
      * +1 (#129)
    - Investment still needed in fundamental research to develop smart tools. 10-20 years on the complex end. (#110)
    - invert the model of scientists coming to facilities. We need facilities to come to scientists (#128)
* Summary and Notes
  + Technologies that will democratize science. What may be the future of experimental facilities and discovery science if technologies that were formerly "big science" and "big facilities" become ubiquitous.
    - Scientists or public interaction with data (#82)
    - Scientists interaction with each other (#83)
    - Scientist interaction with workflows (whole process of science) (#84)
  + Technologies that will change how we collaborate. What may be the impacts on science and scientists of being able to have interpersonal interactions in a virtual environment comparable to interacting in person.
  + New energy production, delivery and storage methods. How may fundamental changes in the energy sector drive change in experimental facilities.
  + Scientist with each other
    - Adding a thought from the group 9 discussion - Incubators - different structures for bringing scientists together to determine "what is the right problem to solve?" (#272)
    - "Design thinking for Science" (#273)
    - A career path for interdisciplinary folks to guide interdisciplinary teams - how to we minimize getting teams up and running when they come from different disciplines? (#274)
    - How do we bridge rather than destroy silos? (#275)
  + Scientist with the public
  + Scientists with science
* Group 9 - What is the problem, issue, technology, device, methodology
  + Technologies that will democratize science. What may be the future of experimental facilities and discovery science if technologies that were formerly "big science" and "big facilities" become ubiquitous.
  + Technologies that will change how we collaborate. What may be the impacts on science and scientists of being able to have interpersonal interactions in a virtual environment comparable to interacting in person.
  + New energy production, delivery and storage methods. How may fundamental changes in the energy sector drive change in experimental facilities.
  + Scientist with each other
  + Scientist with the public
  + Scientists with science
  + Contributions not referencing a 'prompt'
  + Interdisciplinary collaboration: breaking down disciplinary silos (#139)
  + Flexible/agile collaboration: move to a less predetermined/structured so researchers are involved as best fit, lesser barriers for integration with data/tools/resources, better legal/org structure to support that (#140)
  + How do you support collaboration on a global scale? language and time zone barriers (#141)
  + Reducing the start-up time needed for productive, interdisciplinary collaborations. (#142)
  + How to start from problem being solved and work backwards to teams that are needed (including new capabilities or bridging facilitators), rather than starting from existing capabilities and declaring they apply to the problem (#143)
  + issues with domain-specific knowledge and bridging those silos and shortening the time to understanding (#144)
  + Interagency collaborations: breaking down the "ownership" problem (#145)
  + How to facilitate disruptions in collaborations, to move to thinking out of the box? (#146)
    - Cultural/disciplinary barriers to accepting disruptive change (#172)
      * both at the global scientist level but also with the culture of discipline (#173)
      * This is more about getting people to engage many different disciplines in the brainstorming process to think out of the box. (#174)
  + Massive data stores: how to get access to them (from other projects, labs, agencies, industries), and how to actually get them onto DOE hardware for compute (#175)
  + DOE laboratory structure hinders cross laboratory collaboration (#176)
  + Solve data "translation" problem - relevance, usefulness beyond first imagined. (#177)
  + quality control in citizen science (data, understanding) (#178)
  + Need for specific "mediator" staff who are good at bridging disciplines? (#179)
  + Supporting the dynamic aspect of collaborations, that change over time. (#180)
  + Need mechanisms to engage citizen science as appropriate. (#181)
  + Integrated with cyber/trust work and its application to ensure broad collaborations scale well across organizations/security structures (#182)
  + simplifying access to resources, current barriers require many forms and reviews (#183)
  + How to bring new staff into an interdisciplinary culture, rather than bringing them up within a discipline initially and then expecting them to branch out after ways of thinking are set? (#184)
  + expansion of userbase at computing facilities, support higher-level tools (think scratch for grade school students access to programming) (#185)
    - expand to domains that are not traditionally steeped in HPC (#201)
    - increasing accessibility to other areas/disciplines (#202)
    - High-level "modeling programming language" for integrating models and data? (#203)
  + allowing for failures in collaborations without penalty or judgement (#204)
  + How do we define success? What are the metrics for success? (#205)
* Group 9 - Who would develop it?
  + Technologies that will democratize science. What may be the future of experimental facilities and discovery science if technologies that were formerly "big science" and "big facilities" become ubiquitous.
  + Technologies that will change how we collaborate. What may be the impacts on science and scientists of being able to have interpersonal interactions in a virtual environment comparable to interacting in person.
  + New energy production, delivery and storage methods. How may fundamental changes in the energy sector drive change in experimental facilities.
  + Scientist with each other
  + Scientist with the public
  + Scientists with science
  + Contributions not referencing a 'prompt'
  + It is important for the users to be involved in the development. The issue is that the users can change over time. Need to allow for the dynamic changes with the users. (#206)
  + Lab invests in sustainable products and brings in skill set to develop and maintain stacks to support collaboration (#207)
* Group 9 - Who would used it and what skills?
  + Technologies that will democratize science. What may be the future of experimental facilities and discovery science if technologies that were formerly "big science" and "big facilities" become ubiquitous.
  + Technologies that will change how we collaborate. What may be the impacts on science and scientists of being able to have interpersonal interactions in a virtual environment comparable to interacting in person.
  + New energy production, delivery and storage methods. How may fundamental changes in the energy sector drive change in experimental facilities.
  + Scientist with each other
  + Scientist with the public
  + Scientists with science
  + Contributions not referencing a 'prompt'
  + Can be used across different science domains as well as facilities. (#208)
  + Also citizens can be engaged in the collaborations. (#209)
  + Ensure adoption by researchers across domains, past the HPC community. Needs expertise from UX researchers and HCI to build such platforms to reduce time to adoption and use. (#210)
    - not topics DOE traditionally supports but really really needed for the future - accessibility (#229)
  + Increasingly stakeholders or consumers of science as active users, not just DOE scientists (#230)
* Group 9 - When would it be in production?
  + Technologies that will democratize science. What may be the future of experimental facilities and discovery science if technologies that were formerly "big science" and "big facilities" become ubiquitous.
  + Technologies that will change how we collaborate. What may be the impacts on science and scientists of being able to have interpersonal interactions in a virtual environment comparable to interacting in person.
  + New energy production, delivery and storage methods. How may fundamental changes in the energy sector drive change in experimental facilities.
  + Scientist with each other
  + Scientist with the public
  + Scientists with science
  + Contributions not referencing a 'prompt'
  + This is hard to determine. (#231)
  + This is a cultural change, may be slow (#232)
    - success accelerates change/shift (#243)
  + As you have successes, people may be open to come around to change. (#244)
  + DOE "incubator" support for teaming in different ways (#245)
  + Providing details about how a group achieved success. (#246)
* Group 9 - Where, and how widely, would it be deployed?
  + Technologies that will democratize science. What may be the future of experimental facilities and discovery science if technologies that were formerly "big science" and "big facilities" become ubiquitous.
  + Technologies that will change how we collaborate. What may be the impacts on science and scientists of being able to have interpersonal interactions in a virtual environment comparable to interacting in person.
  + New energy production, delivery and storage methods. How may fundamental changes in the energy sector drive change in experimental facilities.
  + Scientist with each other
  + Scientist with the public
  + Scientists with science
  + Contributions not referencing a 'prompt'
  + The collaborations can be global. (#247)
  + It is important to facilitate international collaborations and data sharing. (#248)
  + Also, can engage citizen science, which means some education is needed for broad audience. (#249)
    - or technology needs to intervene to aid in the process (#258)
* Group 9 - What is the setup time?
  + Technologies that will democratize science. What may be the future of experimental facilities and discovery science if technologies that were formerly "big science" and "big facilities" become ubiquitous.
  + Technologies that will change how we collaborate. What may be the impacts on science and scientists of being able to have interpersonal interactions in a virtual environment comparable to interacting in person.
  + New energy production, delivery and storage methods. How may fundamental changes in the energy sector drive change in experimental facilities.
  + Scientist with each other
  + Scientist with the public
  + Scientists with science
  + Contributions not referencing a 'prompt'
  + shorten time to collaboration startup (#259)
  + access to both data and resources (#260)
  + More integrative computational/software frameworks for easing collaborative science (rather than making more complex burden) (#261)
    - imbed collaboration in the frameworks (#271)

1.2 Breakout 2 - Implications of this Problem.

**The following participants have not been active:**  
Mike (ANL), Prasanna Balaprakash, Rich Carlson, Richard Gerber, Josh Greenberg

**Question or instruction for the discussion:**  
Breakout 2 - Implications of this Problem.  
Each group will now develop a list of issues and implications for the issue/technology/community they settled on. There are lots of implications for how a technology can be used, or further developed.



**Sticky points:**

 Top Takeaways (5 points per participant)

* What other/companion technologies, services, software/hardware must also be developed and deployed?
  + Scientist with public (#1)
    - We need tools and services that can help us identify fact from fiction (#17)
    - We need to create space (not necessarily physical) where highly-trained scientists, engineers, facilitators can collaborate on a single problem (grand challenge) (#18)
    - Need for tools to help educate and excite a broad audience about the collaboration. (#30)
      * It would be nice if some of the tools that scientists use make their way into broader dissemination... design the tools for the masses - make things that high school kids can use (#35)
      * +1. Lots of different areas here: new ways for public to interact with science (visualization, AR/VR, etc.), ways for public to discuss (social media, etc.), ways for public to self-educate (MOOCs, tours, etc.) (#37)
        + what might this look like in 30 years? do we have systems that learn preferences for how we view the world and organize things accordingly? can we fight the bias of a fixed mindset where folks stick with what's comfortable? (#44)
  + Scientist with each other (#2)
    - Platform to extract some "knowledge" from data in other disciplines to determine potential for use in other domains (#20)
      * generate context and meta information such that scientist from some other discipline and see potential for use of the data (#54)
      * An "oracle" that could look at data from different disciplines and find relationship between them. (#63)
    - High-level "programming language" for specifying how models, data are integrated to solve problems, make predictions ... abstract away details of data sources, coupling ... propagate uncertainties / errors (#21)
    - Tools to aid in communication (#23)
      * +1 e.g. if scientists wants to bring their tool to the scientist-masses, what tools let them educate others? Can we go beyond manuals since we know nobody RTFMs? Virtual guides, smart guides, etc. (#50)
        + +1 to no one reading the manuals... folks attention spans are shorter and shorter these days (#60)
    - Technology and process methods that lower barrier of entry - less redtape to contribute and collaborate (#38)
      * +1 And would add to redtape, also technological advances to make using a quantum computer the same as a spinning up a AWS instance or even using google to do a search (#45)
    - Collaboration tools - remote whiteboards - ways to build on creative work. Training and practice using these tools in low-risk situations (#70)
  + Scientist with science (#3)
    - Way to focus and gather around a problem, without predisposed choices driven by organization/defined structure (#19)
      * +1 (#25)
    - Intelligence (e.g., AI) to understand the ask and coordinate the retrieval of the data and scheduling of resource to process it. (#26)
    - Automation of routine work/tasks - to reduce cognitive load and keep the scientist focused on the science (#29)
    - Tools that can automatically detect and alert us to biases (#40)
    - How to bring big data to compute when the data may "belong" to other organizations / agencies / governments / corporations. (#42)
      * Or, computationally, how to interact with petabytes/exabytes of data that is stored on non-DOE systems, but needs DOE leadership computation? (#46)
        + +1 do we have petabytes/exabytes of data that are well-described enough to facilitate reuse? or do we need "knowledge extraction" here as well? (#59)
    - Optimal resource allocation: big integrative collaborations will require decisions about how to allocate compute time, experimental facility time, etc. How to objectively quantify this throughout a whole project, as opposed to giving each subteam an arbitrary allocation. (#58)
  + Cross cutting ideas (#13)
    - the need for technologies (knowledge extraction, etc.) that reduce the time for effective collaborations. (#22)
      * can we define "knowledge extraction"? (#34)
        + Being able to extract sufficient abstractions of the data so another researcher can see how to use it... Extracting context for data and analysis. (#52)
        + Context is critical. (#53)
      * Translation tools to aid in collaboration - common language for scientific context (#57)
    - Intuitive ways to get assurance about data/results and integrity (imagine a browser certified/secure symbol for other things) (#24)
      * Yes - this may also need to be something that's interactive - let the user explore the rationale for something being "secure" or "trusted" (#32)
    - Investment in user-centric solutions: invest in people and resources with relevant skill sets (#43)
      * Gamification of doing science (addressing a scientific grand challenge) (#56)
    - Rethinking of organization and culture in the research complex e.g. alignment around problem areas, risk/reward ratio changed, foster more collaboration (rather than compete), etc. (#62)
  + "you don't know what you don't know" - help defining relationships between the silos (#64)
* Who is/will develop this companion technology/service?
  + Scientist with public (#4)
    - Important to have a broad audience engaged in the development, to insure that the technologies appeal to the public (#36)
      * This is where techniques from "user-centered design" can be adapted for a scientific pursuit (#47)
  + Scientist with each other (#5)
    - Technologies for virtual immersion is primarily in the commercial space. (#69)
    - how can we experiment with what's being developed in the commercial space? (#72)
  + Scientist with science (#6)
    - Generic technologies (e.g, natural language processing, etc) will be developed in the commercial space, but integration of this technologies to support science would be a DOE effort. (#67)
      * Yes, and I'd argue that industry won't be able to adapt these technologies to scientific data (#73)
  + Cross cutting ideas (#14)
    - VR-based Video game creators (#71)
    - Those who understand that scientific data must be involved in all of these areas. Would be great to collaborate with technologists. (#74)
* What skills/knowledge does the end user require?
  + Scientist with public (#7)
    - Low barrier to understand: should be able to comprehend results and some real world effects (#33)
  + Scientist with each other (#8)
    - GPT3 for science (#78)
  + Scientist with science (#9)
  + Cross cutting ideas (#15)
    - Writers - those who can express scientific ideas with limited jargon (#75)
      * https://www.amazon.com/Thing-Explainer-Complicated-Stuff-Simple/dp/0544668251 (#76)
      * Up Goer 5! (#77)
    - Involve decision makers / end users in problem and solution formulation. (#48)
    - "Project architects" or "system architects" who see interdisciplinary / grand challenge big picture, end-to-end solution of a problem, focus on interfaces between components/teams (#49)
    - "Computational architects" who can design high-level modeling / programming / workflow abstractions, data-model coupling, to bring large integrated projects together in an easy-to-specify way. (#51)
    - Integrate treatment of uncertainty (and decision/application consequences of uncertainty) in a more meaningful way, from the ground up, tightly into projects. (#55)
* What are the training/support requirements?
  + Scientist with public (#10)
  + Scientist with each other (#11)
    - Need for professional "interdisciplinary" scientists who can mediate/translate between disciplines, lower barriers to cross-discipline collaboration (#31)
      * Workforce development: how to avoid early career scientists getting trapped in disciplinary thinking due to initial technical focus (#39)
  + Scientist with science (#12)
  + Cross cutting ideas (#16)
    - Problem-oriented project formulation: start with problem to solve, work backwards to determine personnel needs ... rather than starting with disciplinary capabilities and working forward to what scientists think is needed (#27)
    - Early engagement of stakeholders / decision makers / users (#28)
    - Need to identify the criteria for success, which may evolve with the collaboration (#41)
  + "Incubators": can the DOE facilitate pilot interdisciplinary demonstrator projects that are scoped, designed, staffed, and organized differently than usual? (#61)
    - Really like the idea of incorporating no context switching in the incubators - when you're part of the incubator, that's all you do (#65)
      * with a virtual world where we don't have to travel to another place to work with others all day every day, this becomes more of a possibility (lots of folks couldn't travel and live elsewhere for months at a time) (#66)
    - Demonstrate tolerance of fail fast/early, and as more is learnt on the problem pivot as needed (#68)

1.3 Day 1 Reflections

Participants: 0

**Brainstorm question or instruction:**  
Day 1 Reflections  
This area is for the Moderator to note key discussion points to summarize what was accomplished in Day one. Remember that day one is focused on Identifying a new technology or methodology and identifying the implications and possible consequences of it. The moderator can populate this individually at the end of the day or request input from the group here.



2 Day Two - November 5, 2020

2.1 Breakout 3 - Signposts

Participants: 2

**Brainstorm question or instruction:**  
Breakout 3 - Signposts  
What we are looking for is technology or social trends that would give us clues that we are on the right track. o How would precursor technologies/services be identified? o What are the precursor technologies/services? o Is there a rank order for when specific technologies/services need to be available? o What DOE or Lab policies need to be in place now, in 5 years? o What facilities need to be in place now, in 5 years?



**Sticky points:**

 Top Takeaways (5 points per participant)

* 1. Synthesis of topics from discussion on Day 1: https://docs.google.com/document/d/1tFDBhqUhDC5LThvDABglzwUod0gnia7WR2Chj4KHW7U/edit#
* 2. Unconstrained collaboration: Overwhelming adoption of open-access publication across scientific disciplines (papers, data, code, etc)
* 3. Immersive interactions: Commoditized novice-level access to VR/AR (or future equivalent) equipment (i.e. as common as smartphones today). OR elimination of need for specialized equipment
* 4. Topic 1 - Scientists’ interaction with Science - Simplicity of use (e.g., Star Trek)
  + Comments
  + How would precursor technologies/services be identified? Sharing of to data/code/provenance is commonplace (because of mandate, reward structure, cultural) Across scientific domains (currently only common in some areas) Technologies that can understand intent of the (science) query and render it into actionable items (e.g., retrieve the right dataset, schedule compute time, run the right algorithms, etc) High-level descriptive/declarative "integrated modeling language" for scientists to specify how simulations, data, statistics, ML, etc. gets combined? (Spend less manual labor on hand-coupling of scripts, design of ensembles, building of statistical models / surrogates to integrate data.) Not just a "software coupling" type of system, but allow for the specification of statistical assumptions / probabilistic models, quantitative rules for combining information Expose / make explicit scientific assumptions as part of the "analysis program" ... allow users to challenge, query, alter assumptions and trace changes in assumptions through to new consequences Make analysis workflow transparent, traceable, interrogatable, changeable (and have a way of quickly understanding effects of "what-if" changes in assumptions) Easy for new collaborators to add information sources by specifying their nature, location, how they integrate with others Integrate into automated AI decision loop to plan simulation/experimental campaigns within resource/budget constraints? Users specify problem, AI figures out how to plan it, and even "compile" it into an efficient workflow Explainable/Interpretable AI Early signpost: commonplace adoption of AI-based analysis and inference tools in mission-critical settings (currently much hesitancy due to model opaqueness) Later: conversational AI agents that can explain their rationale and inferences in natural language and other modalities This will help scientists understand the consequences of an AI decision AI and robotic research assistants Data and domain-knowledge driven AI hypothesis engines Robotic experiment automation AI data analysts Cross-experiment, cross-lab data synthesis More information-fusion like approaches as the initial design concept and designing what sources of information are needed to solve problems ... as opposed to taking existing piles of simulations and data (or obvious new simulation and experimental campaigns as extensions) and figuring out how to integrate/synthesize them after the fact E.g., if we are using ML/AI to integrate multiple data streams into a prediction or decision, think about what kind of simulations/data would best inform the AI, rather than what experiments we think are most interesting Role for optimized design of scientific experiments? (Using AI to help determine the best design/method to support a scientific experiment, e.g., get an intelligent estimate of the result before taking up 100K cores to compute the result) Routine use of sci fi-like assistive technologies VR/AR, etc. NLP (#5)
  + What are the precursor technologies/services? Web-based computing Jupyter notebooks technology wise (sharing/transfer) ML more common place and available as a tool . Used for "knowledge extraction" so published data/artifacts have enough context associated for reuse AI/ML for natural language processing Ubiquitous and ultra-speed connectivity (5G) [unconstrained access to information] High powered processors, sensors Data commons and automated policy enforcement Differential access (Here's what I mean to see is that some data may not be made available for everyone, so how the enabling technology allows us to do this) Ultra high-end (capacity, predictive, resilient) networking to connect facilities Access to HPC and other compute resources for domains that don't traditionally use them and are not specifically trained for them (usability barriers) Access to affordable (and right type of) storage for scientific data Massive storage with high-speed networking outside of DOE to bring large non-DOE data sources into DOE compute (#6)
  + Is there a rank order for when specific technologies/services need to be available? Order of technology may not be critical as each technology development may provide incremental utility. Network speed, computing power (especially on edge), User experience enhancements allowing seamless interaction with complex technologies Order may not be as important as long as we understand where the bottlenecks are and work from there Synthesis of complex data, trend analysis, knowledge extraction - Technologies that enable seamless "what-if" analysis (#7)
  + What DOE or Lab policies need to be in place now, in 5 years? Less concern about "mission" space and focus on problem space Investigation into the softer sciences around evaluation and usability DOE does not have/support a way of evaluating this - that needs to change today (metrics and evaluations) Policy that requires sharing of scientific artifacts (code, data, methods) Also, the flexibility of policy change fairly dynamically Policies that promote integration of facilities (e.g., cross facility allocations, common Authentication/Authorization methods) Effect/influence workforce pipeline focus to include skills for interdisciplinary work Simplification of funding models that enable/reward long-term (#8)
  + What facilities need to be in place now, in 5 years? More seamless networking with extremely high-data rate and huge storage (both at the edge and in the core of the network) Automated coordination of resources across facilities (e.g., HPC, HPN, storage) to support complex workflows (#9)
* 10. Topic 2 - Scientists’ interaction with other Scientists - Unconstrained Collaboration  
  (should include Scientists’ interactions with Science - Immersive Interactions)
  + Comments
  + How would precursor technologies/services be identified? Technologies that facilitate interdisciplinary collaborations -- reducing the time to gain appreciation for different disciplines Identification/recruitment of personnel in integrative areas: cross-disciplinary scientists, systems-level architects, software architects (or even programming language designers) for integrated scientific workflows Language for communication between scientists of different disciplines to help bring up speed (#11)
  + What are the precursor technologies/services? Funded demonstration projects of new approach to collaboration: design technologies, integrative approach, and personnel interactions from scratch rather than emerging from existing disciplinary silos and technical capabilities Current technologies in VC, telework (including holographic representation) remote interaction begin to enable scientist to scientist interaction but require a notion of the scientific workflow truly support future technologies Communication technologies that can support true seamless real time language translation (e.g., universal translator) AI assisted ontology translation and data transformation for data modeling and sharing in collaborative workflows Tools that connect, scientists to scientist, data and resources (instruments/compute) and facilitate the collaboration Solutions for data discovery (services for search, access to data, connect to researchers engaged) Explainable AI (#12)
  + Is there a rank order for when specific technologies/services need to be available? (#13)
  + What DOE or Lab policies need to be in place now, in 5 years? Policy for open and shared access to scientific artifacts (code, data, methods, platforms) Break down/reduce barrier: organization policy/constraints on how collaborations can be formed - across lab complex, across agencies, industry(?) Formal recognition of explicit staff roles for integrative science, rather than adding a new hat to disciplinary scientists Allow collaborations that are not predetermined - iterative and evolves as new data emerges Solution-oriented problem formulation: work backward from needed decision or prediction to technologies, models, data sources, info fusion approaches \*and teaming\* needed to solve problem, rather than presupposing capabilities and personnel to be involved and working forward Support for sustaining technology and work products that support such collaborations Have more opportunities for collaborative projects, working across funding agencies Easier data access for scientists outside DOE Labs (#14)
  + What facilities need to be in place now, in 5 years? Mechanisms that allow for people to plug in from a different discipline (ways for one to easily capture the relevant knowledge or information from other disciplines as it relates to a given problem) (Citizen science) last mile network solved Gamification of science (#15)

2.2 Breakout 4 - Signpost Plausibility

Participants: 1

**Brainstorm question or instruction:**  
Breakout 4 - Signpost Plausibility  
Now that we have the list of signposts, the groups need to consider how plausible they are and what DOE needs to do to either ensure they happen or the implications of them not happening. o Who is actively working on these precursors? o When would these precursor technologies/services be needed? o What active or pending research programs need to be in place now? In 5 years? 10? o What existing or planned facilities need to be in place now? In 5 years? 10? o What software services or capabilities need to be in place now? In 5 years? 10? o How successful has the community been in meeting previous goals?



**Sticky points:**

 Top Takeaways (5 points per participant)

* 1. Topic 1 - Scientists’ interaction with Science - Simplicity of use (e.g., Star Trek)
  + Comments
  + Who is actively working on these precursors? DOE already has some examples of large-scale collaborative grand challenge projects (e.g., some of the new COVID work), which are heading in the right direction in terms of new integrative cross-disciplinary problem-oriented science ... but there are deeper questions of how to formalize, develop workforce, develop tools/tech to make these grand challenge solutions more effective collaboratively, simpler, more routine To facilitate the seamless integration of DOE facilities, DOE will have to build a roadmap of efforts along with corresponding policies to enable this Companies like FaceBook, Google, Apple building out consumer facing devices like Oculus, Apple Glasses etc that are enabling AR and VR Companies building out AI frameworks (PyTorch, TensorFlow etc) High-level domain-specific modeling languages already exist in various open-source and company-sponsored projects (probabilistic programming languages, optimization modeling languages, high-level ML "languages" like Keras ... scientific workflow frameworks NIH is working on creating a data repository for biomedical data that enables discovery, reuse etc that could be used as a model for similar efforts in other scientific domains Similar efforts in NIH to standardize identity, Authentication, Authorization and Access to data Adopt/collaborate to make this cross agency rather than reinvent another silo for "DOE research" Better UX (user experience) enabling technologies (e.g., web-frameworks that industry is building (e.g., Siri), game controllers (PlayStation controller) Explainable/Interpretable AI (XAI) is a very active area of research at present in academia and industry. Many directions - surrogate models, activation visualizations Universities, such as Stanford, UC Berkeley, MIT (#2)
  + When would these precursor technologies/services be needed? NLP is currently available, but it has to be customized (i.e., understand the jargon, etc) for science queries Intelligent scheduling and orchestration of resource (E.g., HPC, HPN, storage) across different administrative boundaries (#3)
  + What active or pending research programs need to be in place now? In 5 years? 10? Organizing the world's scientific data so it is easily searchable, findable, verifiable Formalize what it means to stand the shoulders of big data Development of workforce that has the capabilities (social, technical) for interdisciplinary research Professional (yet simple) usability of emergent technologies like AR/VR needs significant work. Presently limited mostly to games, tech demos, etc. Need more exploration and evaluation of work- and science-driven applications If holographic and robotic telepresence are to have a place in the scientific workplace of the future, in the near future we will need human usability studies for different modalities (10 year) Investment in "integration science" R&D as a programmatic end-goal, not an afterthought: technologies to support information fusion, uncertainty quantification targeted at multi-modal, heterogeneous, biased-source data problems, optimal scientific decision making, interactive "integrative modeling" programming environments, scientific workflow frameworks, ... (AI/ML) methods to tag data to preserve provenance, as well adding labels that can be cross referenced by other scientific disciplines Common data schemas to facilitate collaborative discoveries Investment on usability aspects of tools/services - how do study them, and apply to capabilities developed for researchers (lot of good work is being done in the industry on this) Programs to build out methods for assurances with data/process/provenance Platforms for integrating across facilities (being able to access compute/storage/instruments/services across the lab complex, perhaps beyond) Solving secure access (authN/authZ) across these resources Investment in scientific modeling / analysis / workflow automation ... modeling languages, interactive analysis environments, information fusion paradigms and instantiations as computational workflow architectures, AI for optimal scientific design (#4)
  + What existing or planned facilities need to be in place now? In 5 years? 10? Increased networking, computational, and storage resources Way to combine DOE HPC resources with massive datasets on non-DOE servers Large storage facilities (at or near compute) to meet the growing volume of new science data, as well as existing data that will be re-processed [ongoing] secure/sensitive computing without need for dedicated enclave (not classified work) [5 year] Diverse computing models (e.g., Quantum computing, HPC, local-compute, edge-compute, GPU, etc) to support diversity in computing needs, e.g., AI, etc Program/group focused on building solutions to enable collaborations, ideally cross cutting (across labs / agencies) Facilities that support remote experiments (#5)
  + What software services or capabilities need to be in place now? In 5 years? 10? Resources for storing and managing scientific artifacts (storage, services, personnel) Infrastructure automation and service orchestration software to manage scientific data and workflows end-to-end Data driven, AI assisted infrastructure monitoring and measurement tools Software that can verify the validity and reliability of the data (e.g., Fauci-bot) Interactive inspection of data to deliver automatic trend analysis, anomaly detection, suggestions for downstream analysis and other related data (think about YouTube recommendations) (#6)
  + How successful has the community been in meeting previous goals? NLP and AI developments have advanced at a very rapid pace in the past decade. Robotics has been reinvigorated in past ~20 years by expansion of social robotics R&D and commercialization (e.g. in health care, hospitality settings). These developments have direct implications for robotic telepresence in the future scientific workplace (self-driving laboratories) There are some initial efforts in the orchestration of resources across administrative domains Advances in distributed computing are driven by commercial entities at a faster pace than the research enterprise (#7)
* 8. Topic 2 - Scientists’ interaction with other Scientists - Unconstrained Collaboration  
  (should include Scientists’ interactions with Science - Immersive Interactions)
  + Comments
  + Who is actively working on these precursors? Leverage from Topic 1 above (#9)
  + When would these precursor technologies/services be needed? Substantially improved connectivity, particularly for remote work (needed within 5 - 10 years?), e.g. satellite constellation internet Next generation commercial services for telepresence (e.g. the teleconferencing of 2040) - e.g. holographic presence, physical integration, AI avatars, AI moderation, etc. Expect 2-3 paradigm shifts over next 2 decades Secure communications (e.g., QKD) to ensure information integrity Digital Twin technologies that allow for virtual world objects to interact with real world equipment and standards (like the BlueTooth specification) Services that enable faster turnaround time for InterAgency collaboration, IP discussions, licensing arrangements (#10)
  + What active or pending research programs need to be in place now? In 5 years? 10? Tools and services that support data reuse - automated extraction of information, context addition etc. (tied to ML perhaps) [5 yr] AI processing of data that is not machine readable in traditional sense, while adoption of FAIR data principles gains traction Quantifying bias in AI-driven research [5yr] Support for discovery of useful scientific work across domains (rather than force query where researchers find data without prior knowledge it is needed for specific work) [10yr] Funding for explicit "integration" exercises and low-TRL R&D for technologies needed to support integrative science (this has a policy related function) ... as opposed to funding all the science pieces up front and expect the integrative science and methods to be bolted on later ... seed projects oriented explicitly at integration [5yr] Some of these may look like operations research / optimization / decision theory, statistics, AI, computer science, systems engineering, applied math / model reduction ... other bridge/"glue" disciplines between traditional physical domain sciences Integration as a programmatic activity / area ... what PM supports "integration" as an R&D area and "discipline" in its own right? Scientists who work in this area (integrative research) are not recognized as a critical component of the overall research activity (this is a sociological aspect that needs to be addressed). Quantum networking to support Quantum Information systems (QIS) [10yr] E.g., Quantum sensing, transduction, etc (#11)
  + What existing or planned facilities need to be in place now? In 5 years? 10? New workforce development models: hire or grow new kinds of "bridge" staff (interdisciplinary scientists / "mediators", systems architects, software architects or programming language developers for large integrative science/data frameworks, operations researchers or even science R&D-flavored "industrial engineering", liaisons to end-user communities, ...) [5yr] Mechanisms to fund/support them ... formalize their roles, recognize them as being distinct from domain science / basic science, and valued [5yr] Empower them to make decisions about project direction (as opposed to current "decisions about interdisciplinary projects are made by the most senior disciplinary domain scientist" model) [5yr] An incubator model for conducting research (problem centric approach) to vet ideas for what org structure/resources will be needed to do this at the scale of DOE labs or across agencies. (Not a facility, but don't see a better bucket for this) May need to try out different collaboration models ... different visions More basic than collaboration model: decide what kinds of capabilities and personnel are necessary but don't exist / aren't easily visible in org structures Programmatic examples of deciding who/what needs to be in place to solve a problem before any discussion of teaming or leveraging existing projects, and funding "hole filling" (possibly including hiring in new areas - requires long-term support) rather than expecting to just pull together staff from existing teams Shift in strategic plans, goals and metrics at labs to support more collaborative work, with higher tolerance for risk and thus higher reward. [10yr] Also more support for projects that have explicit mechanisms and R&D intended to facilitate to transition to predictive or applications science useful to end-users outside DOE, decision support Testbed facilities (e.g., dark fiber, Quantum labs, etc) to support QIS research [5 yr] (#12)
  + What software services or capabilities need to be in place now? In 5 years? 10? Federated identity management and policy based authorization for collaborations among groups, institutes and agencies. [5yr] Continued development and deployment of cloud based collaboration platforms [5yr] Software to support QIS [10yr] Capabilities for researchers to make data available to the community, with policy constraints they may need to place. (Storage, tools, integration with their w/f) [5 yr] (#13)
  + How successful has the community been in meeting previous goals? The community has done relatively well from a research perspective. However the track record of transferring the research to production has always been challenging and in most cases does not happen. Solution-oriented activities like uncertainty quantification, decision support / optimal design, combining multiple sources of relevant data (maybe outside of DOE), making use of non-DOE models ... all usually viewed as ancillary or not in DOE mission, but hinder addressing the integrated mission problem (#14)