Group #6: Community of Interest on the Future of Scientific Methodologies

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

|  |
| --- |
| Participants |
| James Fairbanks | Benjamin Gyori |
| Logan Ward | Nobuo |
| 6-Ben Shneiderman | John Bachman |
| Mark Hereld | Ewa Deelman |
| Amber | Heather |
| b | \*fac- Nami |
| Line Pouchard | Sonia Sachs |
| Ben Shneiderman | Josh Greenberg |
| joshua | Benjamin Gyori |
| Stuart Gluck | Jini Ramprakash |
| Shanan Peters |

Contents

[1 Day One - November 2, 2020 3](#_Toc55994164)

[1.1 Discussion: Breakout 1 - Define the Scope of the Problem. 3](#_Toc55994165)

[1.2 Discussion: Breakout 2 - Implications of this Problem. 7](#_Toc55994166)

[1.3 Brainstorm: Day 1 Reflections 9](#_Toc55994167)

[2 Day Two - November 5, 2020 9](#_Toc55994168)

[2.1 Brainstorm: Breakout 3 - Signposts 9](#_Toc55994169)

[2.2 Brainstorm: Breakout 4 - Signpost Plausibility 15](#_Toc55994170)

[3 Day Three - November 10, 2020 22](#_Toc55994171)

[3.1 Brainstorm: Breakout 5 - Pitfalls and Roadblocks 22](#_Toc55994172)

[3.2 Brainstorm: Breakout 6 - Keys to Success 22](#_Toc55994173)

[Appendix 22](#_Toc55994174)

[Live chat 22](#_Toc55994175)

1 Day One - November 2, 2020

1.1 Breakout 1 - Define the Scope of the Problem.

Participants: 0

**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?
	+ (6) Will publications be really the vehicle for knowledge sharing? It is not clear that this will be the best way to share information. Maybe a different kind of organization will be beneficial? (#31)
		- (1) Every day, a lot paper are posted on the arxiv, accepted in journals etc. On the main issues is the inability to process all the documents as well as charting out the full connections that existing in the literature. Can future AI agents become the ultimate tool help navigating the vast sea of information that appears every day ? (#7)
		- Publications are written by humans for humans. That interaction will have to be different if literature is meant for AI agents. (#33)
			* why? our AI should be able to handle pubs. certainly other modes of delivery of content to AI agents should be possible, but ideally they would be able to assimilate pubs, raw data streams/archives, models, etc. and do more with it. (#40)
		- Another concept: "Humans are narrative creatures." Will AI need to meet us where we are, listen to talks, debates on Twitter? (#34)
		- I think we can also think of publications more broadly, to not just be text, but actual artifact, maybe annotated in some fashion (#60)
	+ (5) How can we make the body of existing scientific knowledge, and the new knowledge being generated accessible to machine agents to allow them to be participants in scientific discovery? (Ben) (#3)
		- I agree. Curating the training data (i.e., which facts or concepts from what literature) could be one of the tasks left to human researchers. I can already see in materials engineering and chemistry that even simple facts (e.g., X is a supercomputer with T\_c Y) are clear cut. (#27)
		- +1 on this. Can we use AI to replace/augment critical domain expertise that is being lost or that is difficult to reproduce due to natural progression of disciplines (e.g., alpha taxonomy) (#18)
	+ (4) We need a model of AI science that can progress in complexity and evolve. Perhaps AI Research assistant to AI lab manager to AI collaborator, to AI principal investigator? (#17)
		- i like that! (#20)
		- (1) agreed. this is a great way to organize the problem. start with basic training by assimilating existing knowledge (literature) to point of being able to answer questions, see relationships, reason over it, then move on to being able to contribute to the generation of datasets/models for an existing hypothesis, then finally on to suggesting new hypotheses/experiments. (#47)
	+ (3) In terms of interface: doing science or knowledge synthesis or inference across the body of science is a big aspect/goal. do we imagine this requiring some centralized repository of knowledge? of data? who controls this? is there a distributed model? (#32)
	+ (3) Should we consider whether we are limiting AI by limiting it's ability to communicate "understanding" back to humans? (#35)
	+ (1) Description of the topic: Characterizing the relationship with AI: With increasing sophistication, digital agents will represent a competitive advantage in conducting research, just as human collaborators do today. (#1)
		- (1) It's interesting to think about what "competitive advantage" means here. It suggests that the incentive structures in science will determine the types of AI that are developed to support it. What is success? More publications? Better predictions? Lower cost? (#36)
			* Problem: how to leverage AI and related notions to benefit the research process itself. It appears that the focus at least leans toward the personal relationship between the researcher and some number of AI affordances. (#2)
				+ (1) AI can potentially provide a lot of automation of tedious tasks that are done by scientists today. How can we build trust in that automation process. (#4)

I think delegating tedious tasks to the AI is spot on for how they will be used primarily. Trust is indeed a core issue (#22)

totally agree, and will note that trust is a real issue even when humans are involved, so the bar is unattainably(?) high (#38)

* + (1) The term "scientist" will potentially undergo significant change as we open up access to science and the scientific method broadly in the future. (#5)
		- (1) One concept a colleague of mine challenged me with awhile back is whether an AI should be listed as an author. I still think "No" but in 30 years that could be different. How do we recognize which AI agents were used and how? (Logan) (#12)
			* I've been thinking about this too. I think it depends on how we define them as agents. At the moment we use them as tools and cite them as such, but for me the boundary would come once the AI agent starts generating new ideas. (#13)
			* Interesting concept. How much contribution from an AI signifies enough to make it an "author"? What if it was a bot that only did your editing - would that be a "ghost writer"? (#14)
	+ Identifying whether AI agents are something developed by a single group or maintained by a community is another issue (#6)
	+ (2) The term agent is going to evolve significantly over time as AI becomes more and more powerful and mature. Right now I think of agents as some sort of tool, but in the future I could envision them evolving to play many different roles. (#8)
		- One role already envisioned is AI guiding which experiments to run. If that work requires more than a few humans, who makes the call on whether to listen to the AI or not? (#16)
	+ (2) Given the asymmetry between human and machine strengths/capabilities, in what ways will their roles be complementary when they collaborate? How can a machine agent be more than just a tool that the human scientist calls upon? (Ben) (#9)
	+ Problem: developing the sensors, data management, actuators, etc. that will allow digital agents to effectively perform essential scientific functions, whether in a lab or in a SC environment, is perhaps an even bigger challenge than the agents themselves. (#10)
		- another way of thinking about this is making lab equipment accessible/controllable via APIs so that machine agents without physical sensors and actuators can use them (Ben) (#25)
	+ (1) Problem: bridging giant gap between observational/experimental data required to train effective AI in scientific domains and models used to understand natural systems (#11)
	+ How will theory-driven models, data-driven models (e.g. ML) and new experimental methods (allowing questions to be answered empiricially, no prediction necessary) be used jointly for more efficient and effective scientific discovery? (#15)
	+ (1) Do we envision a future in which autonomous agents are capable of taking into account contextual factors (e.g., practical relevance, fairness, cost, actionability) when designing and executing scientific studies? Or do we envision that humans will always be in the loop for guiding these decisions? (#21)
	+ Does this topic resonate with the much more science fictiony slash singularity notion of the augmented mind? The things that researchers do to push the intellectual envelope are integrated with the tools that they use. To the extent that the tools can be natural and immersive, they might be thought of as extensions of the individual. Natural and immersive -- might require careful personalized tuning of the underlying generic AI tech to the individual in a way that is analogous to the "personalized medicine" that is sure to become popular. (#19)
		- Maybe the idea of "autonomous" agents could or ought to be skewed toward the idea of "collective" in the Minsky Society of Mind sense. The nominal autonomous agent might perform big independent tasks for one like "do my taxes" or "field my trivial email". On the other hand, a more tightly integrated and finer grain autonomy could provide for a nice model. (#29)
	+ (1) Historically, automation is accepted in a field when it clearly produces better results for less cost. In science, where discovery is based on understanding complex processes, how will this be quantified? (#26)
	+ (Logan, capturing Josh's comment) Is there a way we can break out AI by which role it replaces? AI Technician vs AI Lab Manager, vs AI PI? (#30)
	+ if in 2050 you had a chip in your brain that allowed you to read every publication and understand every code, would you still want an AI assistant? (#37)
		- Yes. Because ultimately this is a huge-dimension problem and "individual creativity/domain expertise" is at some level a specific "projection" of this into your own context-rich reference frame. one "projection" is not sufficient (#39)
	+ Test comment (#59)
	+ Visual user interfaces empower creativity for scientists at every stage of their career. Tools to see the dynamic growth or decline of topics, understand the impact of a paper, person, group, or topic. Then tools to help develop new ideas, plan research, write papers, and disseminate results to targeted users. For example, I would like to send a new paper to anyone who has cited or downloaded my previous papers on a topic. More powerful tools would enable me to view the uptake of my work in different communities and reveal its impact by business practitioners and patent filings, as well as theory innovations. (#71)
* Who would develop it (basic research to advanced deployment)?
	+ (5) I imagine a "bottom up" kind of development process. Unique AIs geared for a certain task will be created by communities that, as they gain acceptance, start to be developed by a broader communities (#45)
		- Specifics before generics seems almost necessary when creating digital agents...? (#48)
		- Yes, I agree--it is likely that general AI capabilities and infrastructure will be applied and refined by people working in specific scientific domains. (#49)
		- I agree. I think that we already have these to some degrees. e.g. the gmail AI that autocompletes my emails for me. (#52)
	+ (2) This is a cross-domain project. It requires domain scientists partnering with AI and data and computing infrastructure. Different fields will have different requirements and will require some custom considerations. John's point about experiments vs. no experiments is a good one. Additional complications come up if the field has data collected from humans (ethics and HSR concerns) or if the implications of the field could impact people directly (economics maybe or public health) (#42)
		- There would be a lot of ethical considerations in this domain. (#55)
	+ (1) Developing AI agents would currently be a research program in of itself. I think that the future when we might use them is still many years away. If successful, I would expect this to have significant commercial applications. (#24)
	+ One issue to work through in this space is how domain specific a digital agent would be versus trying to create a generalized model of a digital agent that could work across domains. (#41)
	+ Would be developed at least initially by a mix of national governments and universities, with specific domain applications moving into the commercial space (e.g., imagine a powerful AI for drug discovery developed and deployed by a specific drug company for competitive advantage). As is currently the case for a lot of complex ML models currently, access to large volumes of data and computing power would determine which actors would be driving the development. (#44)
	+ (1) Advanced deployment could be an area for national laboratories or universities. The kinds of hardware to train the AI and the non-trivial computing requirements for using them demand access to resources beyond what an individual lab could have, especially with moving more compute to the cloud (#50)
	+ Is an AI agent going to be a computer program or a specific device that you might buy just like we buy our laptops and phones these days? (#54)
	+ (1) Will there be cases where the AI is sensitive to bias (either due to limited data or high impact on society) that will require humans to curate training data by committee? (#57)
* Who would use it and what skills would they need to use it effectively?
	+ (1) Many important points here. We should not expect domain scientists to also become experts in ML or AI. Usability, accessibility, etc. will be key. (#23)
	+ Ideally domain scientists are the users and they need nothing more than their domain expertise operating in a conversation-style interaction. (#61)
	+ I think that this is true in the long run, but I think that the early adopters are going to be the people who develop these tools and agents for themselves. I think that the long-term agent probably won't require any special skills, but I suspect that this would become the reality once a commercial solution has been developed. (#62)
	+ In the future, I would expect that all scientists would be, out of necessity, familiar with these AI tools and the separation between "domain scientist" and AI expert will be blurred. (#67)
	+ We would need to make sure that AI tools and technologies are broadly available, so that we don't create an even greater disparity between scientists. (#68)
* When would it be expected to be in production use (N years in the future)?
	+ I would hope that a general purpose AI research assistant constrained to the world of computation and modeling and knowledge and data (i.e. not physical actuated) would be 5 years. targeted applications in robotic cloud labs and etc. could also be accessible in those times. (#28)
		- I think that the more general purpose agents would be at least 10+ years in the future. (#64)
	+ The NIH is just kicking off a new effort to make biomedical sciences at all scale "AI ready" by defining the computable experimental procedures, knowledge and meta-data structures, sharing protocols, etc. i believe they think in 5-10 years they can have a large proportion of NIH funded research using machine readable data and knowledge standards. making synthesis across experiments and etc. possible. (#51)
	+ Maybe there is no IT, but rather a they. If development efforts include fine grain AI and the ideas required to connect them, one might expect continuous deployment of production-grade capabilities as they emerge from the ecosystem of AI agent tech. (#53)
		- +1 (#65)
	+ See comment 17 (green panel): a progression mimicking to some extent the progression of functionality/capability between grad student -> lab manager -> PI would be good. For grad student, hopefully not far in future, for PI, we are many. years out (#63)
* Where, and how widely, would it be deployed?
	+ (1) The competitive advantage of having a digital agent in the scientific research process will potentially create barriers to the kind of openness that is seen in today's community projects (#43)
		- (1) However, having a model where the knowledge is open even if the tools are not can be beneficial to the entire process. (#46)
	+ I'd hope for ubiquitous deployment in some sense, the idea being to maximize impact. (#56)
	+ These days you'd expect the first big general solutions to this to come out of one of the private research shops like deepmind or openai or ai2 or etc. There's no doubt that there will be both open and commercial tools. big +1 to the importance of keeping the knowledge open. But how can that be incentivized? (#58)
		- I don't know how to incentivize keeping the knowledge public, but I think its very important. (#66)
	+ We need to make to understand implications of AI on things like energy and climate. (#70)
* What is the setup time and/or process for using it?
	+ AI is already being used in science, but the parameters for applying AI in science to obtain a valid result, or the definition of a valid result are not clear. (#69)

1.2 Breakout 2 - Implications of this Problem.

Participants: 0

**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?
	+ Infrastructure for capturing and sharing (with data and publications) rich context information, knowledge embeddings, etc. in machine readable formats. (#1)
		- Currently in most experimental fields, it is very hard to compare across labs because simple differences in equipment, procedures, and methods can seriously effect the data. Even in the same lab, if a postdoc leaves or a piece of equipment is updated the data may be hard to compare to the same experiments last year. (#5)
		- We need infrastructure to support aggregating and distributing large and rapidly growing datasets that span disciplines (ideally, some of these will be AI-constructed/curated) and provide the ability to run models calibrated with/assessed over these data. (datasets in the broadest sense, from field deployed instruments, lab machines, to human-collected/processed samples/observational data) (#2)
	+ for both computational and experimental research, we need to make resources (incl. lab equipment but also computational resources) accessible via APIs for intelligent machines to be able to directly interact with and control them (#4)
	+ we need a new paradigm for disseminating scientific results that are accessible in a useful way to both humans and machines (#9)
	+ We need new mediums for building computational models that isn't code. code is a terrible medium for modeling. (#13)
		- (Shannon) interactions need to be conversational, not code. Domain scientists are not trained to be sufficiently adept at programming. (#32)
	+ We will need a way that humans are able to trust the results from a digital agent - enough confidence (with a margin of error) to believe in the outcome produced by it. (#14)
		- I'm not sure I understand how this differs from using a computer program to produce a result at the moment? (#16)
	+ There may be a number of natural augmenting technologies, notably hardware to facilitate human-AI interaction, robotic hardware to facilitate interaction between agent and world. (#15)
	+ new tools that automatically process existing and future literature such as graphs , equations, text as extraction tools are needed in order to develop the AI agent. That is, agent without tools is not an agent. (#20)
	+ We need to make to understand implications of AI on things like energy and climate. (#34)
* Who is/will develop this companion technology/service?
	+ Research teams at labs and universities. Tech companies. Start ups? (#6)
		- Certainly. There may be nothing in particular that is special about the relationship between existing innovation institutions and the ancillary technologies and services that might support the AI agentry we are discussing. (#24)
	+ This needs to be a three-way partnership between domain scientists, computer scientists, and computing/networking infrastructure technologists. (#7)
		- yes. and incentives need to be there to make this happen. large center grants (NIH is pursuing this method) or DARPA style programs that include cross-domain and integration components and that explicitly incorporate workforce education wherever possible. (#17)
			* How do we incentivise industry? (#22)
				+ provide them with well-trained students that they can hire ready-to-go? (#27)
	+ And in this connection there is also the issue of open access discussed previously, intellectual property, democratization of these new capabilities. (#26)
	+ partnership with journals (#29)
		- S. Peters says: "have fun with that one" :-p (#33)
* What skills/knowledge does the end user require?
	+ In an ideal world a digital agent would fulfill a very specific need intuitively but while we are working to perfect this paradigm end users will still need to have ways to learn about the digital agent and how to get their specific need fulfilled by it/them. (#3)
		- The technology itself could contribute to making this ideal more of a reality. (#11)
		- +1 (#21)
	+ The ability to articulate well-posed problems/hypotheses. The AI tech assistant then serves-up the right answers and/or data/model components and generates results and tests of these hypotheses (#12)
	+ Perhaps an AI agent doesn't need to "explain" every finding or discovery it makes. As long as it points the research in a useful direction, that is a valuable contribution. But perhaps the agent DOES need to explain itself. to be able to teach users how to use it. Seems like we value scientists that are able to teach, so why not our AI scientists. (#25)
* What are the training/support requirements?
	+ Let's hope that this issue is no more onerous that current learning curves. We can do at least as good as we've already done in other tech adoption areas: common interface idioms to manage expectations, better documentation and help facilities (natural language, voice, etc.). (#8)
		- let's require that AI agents be capable of doing all the required training themselves (posted with more depth in red-box) (#28)
		- Such agents can be used for training, e.g. new students (#30)
	+ Significant training will be required to teach researchers. I don't think its an interface issue, but rather a cultural one in that researchers will need to move away from doing a large set of tasks themselves. (#10)
		- Ideally most of the training would be focused on how to pose problems that are well-suited to the AI tech. It is not uncommon in many domains for emphasis to be on "brick makers" (to use joshua's must-read cartoon example) rather than edifice builders. We need more of the latter and in an AI-powered "futureworld" this requires the right training and approach. (#18)
			* Joshua's link: https://massivesci.com/articles/chaos-in-the-brickyard-comic-matteo-farinella/ (#23)
	+ For a human scientist to productively interact with a machine "scientist", it has to be clear what the machine can and cannot do (at all or reliably), where it can be trusted and where its contributions need to be questioned. (#19)
		- How different would these requirements be from those that are required of human-human collaboration? (#31)
	+ There may unintended consequences from, for example, being able to hire "AI grad students". We need to think through the implications in regards to their training to become scientists and also the functioning of the research group. (#35)

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: 0

**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)

* (7) 27. The key precursor technology/ies will be intelligent systems, that, like steps on a ladder, 1) can incorporate both causal knowledge of the domain and data-driven reasoning in generating predictions and explanations (see overlap with breakout #8), 2) can integrate knowledge from disparate domains of science to generate novel causal hypotheses, 3) can design, execute, and interpret relevant experiments for testing those hypotheses, 4) can report findings from these experiments/analyses in both machine and human-comprehensible formats and 5) can, when faced with alternative projects/hypotheses, differentiate among specific research directions in terms of their novelty, or applied or theoretical value.
	+ Comments
	+ +1. I would also include the ability to generate and run dynamical models of systems capable of making predictions and prioritizing new data collection efforts. (#42)
	+ I agree. Arguably that could be considered as a combination of #1 (incorporating causal/mechanistic knowledge) and #3 design/execute experiments. (#47)
	+ I would like to add reproducibility to this, the ability of the intelligent system to be able to reproduce a scientific process (#52)
	+ i think these are good steps for the "AI principal investigator" or whatever phase, but i would argue that there are important steps before this that will be valuable by themselves but will be precursors to these precursors. (#54)
	+ Certainly. Though it's worth noting that 1 and 2 are happening now and have many proofs of concept. (#55)
* (4) 39. AI can also have an important role in reproducibility and provide tools that follow and record the exploration that a scientist does, create a record that can be shared with others and test the reproducibility.
* (3) 13. A lot of scientific understanding comes from the ability to discriminate between hypotheses that could be true from hypotheses that don't make sense. Post simulation/experiment V&V and a priori validation of ideas is a key signpost for AI systems that \*understand\* the scientific domain expertise.
	+ Comments
	+ Just getting to the point where we have a system that could potentially incorporate and/or do a simulation/experiment AND use data from many different sources (literature, databases/data archives/live streams) and then validate/refine would be an achievement. (#79)
* (3) 18. The ability to connect small-scale agents together would be a telling signpost: precursor to facility with integrating the technology, building new and larger intelligent functions from smaller ones.
* (3) 41. Development of these technologies could be supported by new infrastructure for persistent execution of such agents -- including live modification, continuous testing. Perhaps this points to a need for dedicated facility resources for such live development.
* (3) 34. Machine should evolve from being used as tools that are given specific instructions to actually being able to contribute to science creatively, in which case they could be considered "co-authors" in some sense
	+ Comments
	+ I strongly concur. As a scientist, I'm really looking for AI technical assistants that can serve as "domain expert consultants". That is, an AI assistant that can take very basic prompts from me (a picture of an organism, a data set/result, a model) and then provide substantive input/feedback. (#40)
* (3) 32. DOE should have an ethical framework for how to conduct AI enabled science FAIR, IRB, engineering disasters, military ethics doctrines for autonomous weapon systems all provide guidance. The military academy scholarship on autonomous weapon systems looks at history and the laws of war. Ethics in science have a significantly shorter history and AI augmented scientists will face new questions about responsibility and restraint.
* (2) 11. A critical bottleneck to address is making sense of the flood of new knowledge being produced in science and engineering, machines can meaningfully help with that
* (2) 6. If AI are going to contribute to science in the future through knowledge absorption and dissemination then we need to start building the agents that are ingesting all this knowledge
* (2) 4. Precursor technologies must include ability to automate menial/repetitive tasks at scale with high quality across all domains of knowledge.
	+ Comments
	+ As a signpost, is it necessary to have system(s) that can do this over all domains of knowledge? or is a proof of concept in one domain sufficient? (#36)
* (2) 7. There are large conceptual signposts/stages. I like James' self-driving car analogy, automating the menial tasks that currently take most of the time of grad students. Then automating higher level functions until we have a true collaborator or leader.
	+ Comments
	+ But i think we need to distinguish different threads. automating the menial tasks of experimental biology is very different from automating the menial tasks of computational biology. One requires sensors and actuators that will have to advance significantly before generally applicable. (#14)
* (2) 8. A system with multi-agents that enlarge the capabilities of a scientist (or a program manager) to do his/her work is likely to have emergent behavior. Some of it could be good, but some may have unintended consequences. What has been done about simulating/modeling such multi-agent system for AI agents? I know there is a lot of work on this in economics and in transportation/cities, but not aware of recent work on multi-agent AI systems.
* (1) 2. For automating a program manager: it seems that the role of the PM is to make value judgements about the kinds of work to invest in and that is based on the distributions of outcomes for the project and how it matches the values of the funding organization. How could that be automated beyond the detail oriented tasks of estimating risk and keeping track of budgets?
	+ Comments
	+ That is part of it i think, but its also about steering the field in a particular area towards specific goals or valuable avenues. obviously the model differs by agency, but PMs have alot of influence on where the field goes, not just calling balls and strike. (#43)
	+ Identifying promising directions that are not sufficiently covered may very well be automatable and would help to reduce the biases and conservativism that exist in program management that can slow progress in a field. (#45)
	+ I can certainly imagine in the very near future venture capital companies using models to calculate risk/benefit associated with different investments. If this is conceivable then an AI program manager (or assistant to a PM) is not far off. (#51)
* (1) 3. Building on the idea that digital agents evolve we should identify the agents that are already assisting scientists today to see what form those take and how those could evolve to the next "level"
	+ Comments
	+ This is going to be a pretty short, idiosyncratic, and domain/problem-specific list... (#53)
* (1) 5. We should make a matrix of sophistication of task x power of capability to track both aspects of analogizing Scientific AI agents to self driving vehicles and replacing/augmenting researchers as a function of career stage.
* 9. In the context of peer-review system, typically papers are evaluated by few people. We do have technologies in social media, and perhaps AI agents+wisdom of the crowd can potentially improve the peer-review system
	+ Comments
	+ And if AI start writing papers, would review by humans be the right approach to their evaluation or do we need to build AI that are evaluating these scientific publications? (#16)
	+ I would prefer to start with AI tech that can assess novelty of contribution and/or measure in some way the value of a paper beyond the usual metrics (citations, etc.). (#46)
* 10. Another thought is: what formal methods can be used to design AI agents?
* 12. Based on reading teh comments in MSPheres, as I didnot attend, what resonated with me in the discussions is that the discovery and dissemination of knowledge would be profoundly different from what it is today, facilitated by AI, and also aimed at being "read" by AI as well as by humans. The dissemination would necessitate a different infrastructure that enables this knowledge to be stored and accessed at all levels and across disciplines. Thinking about it now in writing, what people formulated and discussed on MOnday is the emergence of Mr. Data, the old Star Trek character, facilitated by AI.
	+ Comments
	+ One way to think about this (perhaps colored by my bio-geo systems oriented view) is a that key step of knowledge dissemination should be incorporation into very large dynamic models of systems that are constantly being revised and improved (#23)
* (2) 15. We should also look at the failings of scientists (local failings) and the failings of science (global or domain failings) and consider signposts when AI has "fixed" these.
	+ Comments
	+ Scientists are notoriously bad at recording the provenance of how they get to discoveries. because the discovery process is messy and non-linear. AI assistants that can help to monitor, structure, and publish these would be hugely valuable. in particular, this would solve the problem of success/failure publication bias, because the failed ideas or experiments along the nonlinear path would be catalogued and published. (#19)
	+ Two such failings relate to other comments above: (1) failure of the current approach to scientific publishing and (2) failing to critically evaluate and make sense of the flood of information being published in an unbiased way (#20)
	+ The field of science has huge broad challenges with the scope and scale of knowledge and the accessibility of publications, which could be solved by AI. (#22)
	+ Reproducibility, transparency (of model code and methods just for one example), comparability of experiments/models/results, communication across disciplines, etc etc. (#33)
	+ We could identify which of these problems are likely to be solved and in one order and how and these would be critical signposts toward the AI Program Manager. (#35)
* (2) 24. Since data will be central to all things AI, DOE & Lab policies will have to change to allow free exchange of data while preserving the attribution of that data.
* (2) 25. Did we talk about ethics and trust in AI?
	+ Comments
	+ [Jini] This topic is very very interesting to me. (#28)
	+ Who decides where to draw the line between what is possible and what \*should\* be possible / allowed? (#31)
* 17. I'd like to think more about how we want the scientific process and teams to evolve once such technologies start to become available. If we do that, then we have a chance to steer the technologies to achieve the outcomes that we might want.
* 21. In the beginning the content of wikipedia was not fully trust worthy, and it took years for people to start believing on the contents. The challenge is how design AI agents that can be trusted and not biased.
	+ Comments
	+ I really believe that it problematic to let this play a key role up front in any effort. Scientists are skeptical distrusting creatures. "Trust" is going to be a factor in every single effort of any kind, human or machine run. Make it transparent, easily understood and accessible and that's that. (#29)
* (2) 26. I keep going back to the idea of AI GRAs etc.. I think one thing that we are missing is that GRAs do bring with them expertise, sometimes through work in industry or through training in a different discipline, or by having different educational training. So, when we think about AI GRAs, AI can not only automate our tedious tasks, but I would like to have AI tools that provide me with expertise in areas needed for my work (potentially in a different domain). So, potentially AI can advance inter-disciplinary research.
* (2) 30. Explainability is a precursor technology that is in the works now but needs to be supported by better tools, methods.
* 37. in the context of human learning, education is not homogeneous across the board. It depends on social instances, geographic locations etc. Can AI agents be the next generation of teachers?
* 38. One metric to watch when developing the technology will be the level to which it becomes autonomous.
* 44. One task program managers could get a lot of help with is the work of matching reviewer expertise to proposals in a review panel. Very time consuming...
* (1) 48. Our focus now will be to arrive at a vision that is more or less unified.
	+ Comments
	+ Is the vision, what we want at the end or the process to get there? The vision is the end result, the process is the signposts. (#56)
	+ JE: our vision is that "automation should replace the large majority of menial tasks, so that all scientists at every level of career can focus on the creative parts of science. At the higher level, we want an AI that can act as a collaborator. AI should also help in reproduction and dissemination." (#57)
	+ JB: are we utopian vs dystopian? What if we don't care about the proceses of science and just want results? The role of the human is substantially different. (#58)
	+ BG: menial tasks might not be the place to start -- perhaps looking at problem with fresh perspective could provide a better approach (#59)
	+ SP: really wants a domain expert collaborator -- not so much an agent that simply automates a well understood (perhaps complex) task (#60)
	+ BG: why are we trying to use AI to automate broken processes, we should fix the root causes of the problems with publication, instrument compatibility. (#61)
	+ ED: we underestemate GRAs we collaborate with/use currently. they bring interesting and useful experience. (#63)
	+ SP/JE: scientific expertise is projecting expertise from high dimensional to low dimensional (#65)
	+ HG: Utopia vs dystopia we should take responsibility over the outcome. We need to build in constraints (#67)
	+ HG: take responsibility for trying to shape the process -- perhaps focusing more on the path than a statement about having an all powerful AI at the end (#68)
	+ LP: we have a more fleshed out and full basis for a vision, but not yet a single pithy statement of what it is (#69)
	+ LP: suggests we move on to talking about SIGNPOSTS here (I am opting to keep this in the current comment block) (#70)
	+ JB: think about accountability, relates to explainability; think about accessibility, fairness (#71)
	+ HG: e.g. biases in current AI such as racism (#72)
	+ JF: let's not forget root cause of things like AI bias is driven by business model (#73)
	+ MH+: facilities might be a useful area where DOE could participate. existence of those facilities could be useful signposts. (#76)
	+ MH: roles of facilities -- R&D, scaling, persistence, collaboration (#78)
* (1) 62. what it means to communicate about science, not just automation of menial tasks, an interaction with an AI system that can generate hypothesis, against whom I can play my ideas
* (1) 66. utopian vs dystopian vision: we need to take responsibility for shaping it, how do we achieve something that would be better?
	+ Comments
	+ DOE labs have the ability to build AI technology that isn't financially incentivized towards unfairness (#74)
	+ DOE has a facilities thinking mindset which can help shape the approaches in a way that industry can't (#75)
* (1) 77. SIGNPOST: DOE has an agreement on policies for how to study AI systems at the facility level.
* 64. discovery into the context of multi-disciplinary
* 1. We already have a number of AI technologies out there but many of them have not been applied in the science arena. So, more exploration is needed there.
* 80. ONE LIST OF SIGNPOST FODDER
	+ Comments
	+ The ability to connect small-scale agents together would be a telling signpost: precursor to facility with integrating the technology, building new and larger intelligent functions from smaller ones. (#81)
	+ The ability to connect small-scale agents together would be a telling signpost: precursor to facility with integrating the technology, building new and larger intelligent functions from smaller ones. (#82)
	+ The ability to connect small-scale agents together would be a telling signpost: precursor to facility with integrating the technology, building new and larger intelligent functions from smaller ones. (#83)
	+ The ability to connect small-scale agents together would be a telling signpost: precursor to facility with integrating the technology, building new and larger intelligent functions from smaller ones. (#84)
	+ The ability to connect small-scale agents together would be a telling signpost: precursor to facility with integrating the technology, building new and larger intelligent functions from smaller ones. (#85)
	+ The key precursor technology/ies will be intelligent systems, that, like steps on a ladder, 1) can incorporate both causal knowledge of the domain and data-driven reasoning in generating predictions and explanations (see overlap with breakout #8), 2) can integrate knowledge from disparate domains of science to generate novel causal hypotheses, 3) can design, execute, and interpret relevant experiments for testing those hypotheses, 4) can report findings from these experiments/analyses in both machine and human-comprehensible formats and 5) can, when faced with alternative projects/hypotheses, differentiate among specific research directions in terms of their novelty, or applied or theoretical value. (#86)
		- Certainly. Though it's worth noting that 1 and 2 are happening now and have many proofs of concept. (#100)
		- +1. I would also include the ability to generate and run dynamical models of systems capable of making predictions and prioritizing new data collection efforts. (#87)
		- I agree. Arguably that could be considered as a combination of #1 (incorporating causal/mechanistic knowledge) and #3 design/execute experiments. (#92)
		- I would like to add reproducibility to this, the ability of the intelligent system to be able to reproduce a scientific process (#97)
		- i think these are good steps for the "AI principal investigator" or whatever phase, but i would argue that there are important steps before this that will be valuable by themselves but will be precursors to these precursors. (#99)

2.2 Breakout 4 - Signpost Plausibility

**The following participants have not been active:**
Amber, b, Stuart Gluck, Josh Greenberg, Benjamin Gyori, Sonia Sachs, 6-Ben Shneiderman, Ben Shneiderman, Logan Ward

**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)

* 48. The key precursor technology/ies will be intelligent systems, that, like steps on a ladder, 1) can incorporate both causal knowledge of the domain and data-driven reasoning in generating predictions and explanations (see overlap with breakout #8), 2) can integrate knowledge from disparate domains of science to generate novel causal hypotheses, 3) can design, execute, and interpret relevant experiments for testing those hypotheses, 4) can report findings from these experiments/analyses in both machine and human-comprehensible formats and 5) can, when faced with alternative projects/hypotheses, differentiate among specific research directions in terms of their novelty, or applied or theoretical value.
	+ Comments
	+ These ideas map onto the scientific career stages (something we discussed before as it relates to AI) quite nicely, i.e., research assistant, grad student, postdoc, PI, program manager (#51)
	+ I'm not sure we should call this a ladder because I don't think the order is linear. e.g. I think that automating aspects of experiments is already done now, while 1) requires quite a bit more sophistication. (#52)
		- I interpret "design" experiments in 3) as a huge achievement which makes it more complicated than 1 or 2 (#61)
			* I think we should separate design for the rest there. (#64)
		- agreed. ordering here doesn't quite align, but the notion of increasing levels of complexity very appealing (#54)
	+ #3 and #4 are both being worked on today but #1, #2, & #5 are very far away. (#57)
		- There is quite a bit of non-trivial reasoning required for "reporting findings". It's not just dumping some tables and plots, rather, crafting a narrative that turns results into high-level findings. (#67)
		- #1 is making good progress now too. for example, by automating the reading of scientific papers to extract specific causal relations relevant to a model (#59)
* 49. AI can also have an important role in reproducibility and provide tools that follow and record the exploration that a scientist does, create a record that can be shared with others and test the reproducibility.
	+ Comments
	+ Probably this is different by field, but in my field things are already highly reproducible -- we put quite some effort into that. Can we clarify that this is an issue in some fields but not necessarily all? (#56)
	+ integrating failure models into reproducibility, it would become easier to learn from the failed experiments if reproducibility for AI was already in place (#93)
	+ Additionally, reproducibility (especially in fields where it's not easily achieved) becomes a goal, the steps towards achieving it involve transparency. Transparency for AI models, leading to the discussion on Explainability (#95)
* 50. A lot of scientific understanding comes from the ability to discriminate between hypotheses that could be true from hypotheses that don't make sense. Post simulation/experiment V&V and a priori validation of ideas is a key signpost for AI systems that \*understand\* the scientific domain expertise.
* 1. Our focus now will be to arrive at a vision that is more or less unified.
	+ Comments
	+ Is the vision, what we want at the end or the process to get there? The vision is the end result, the process is the signposts. (#2)
	+ JE: our vision is that "automation should replace the large majority of menial tasks, so that all scientists at every level of career can focus on the creative parts of science. At the higher level, we want an AI that can act as a collaborator. AI should also help in reproduction and dissemination." (#3)
	+ JB: are we utopian vs dystopian? What if we don't care about the proceses of science and just want results? The role of the human is substantially different. (#4)
	+ BG: menial tasks might not be the place to start -- perhaps looking at problem with fresh perspective could provide a better approach (#5)
	+ SP: really wants a domain expert collaborator -- not so much an agent that simply automates a well understood (perhaps complex) task (#6)
	+ BG: why are we trying to use AI to automate broken processes, we should fix the root causes of the problems with publication, instrument compatibility. (#7)
	+ ED: we underestemate GRAs we collaborate with/use currently. they bring interesting and useful experience. (#8)
	+ SP/JE: scientific expertise is projecting expertise from high dimensional to low dimensional (#9)
	+ HG: Utopia vs dystopia we should take responsibility over the outcome. We need to build in constraints (#10)
	+ HG: take responsibility for trying to shape the process -- perhaps focusing more on the path than a statement about having an all powerful AI at the end (#11)
	+ LP: we have a more fleshed out and full basis for a vision, but not yet a single pithy statement of what it is (#12)
	+ LP: suggests we move on to talking about SIGNPOSTS here (I am opting to keep this in the current comment block) (#13)
	+ JB: think about accountability, relates to explainability; think about accessibility, fairness (#14)
	+ HG: e.g. biases in current AI such as racism (#15)
	+ JF: let's not forget root cause of things like AI bias is driven by business model (#16)
	+ MH+: facilities might be a useful area where DOE could participate. existence of those facilities could be useful signposts. (#17)
	+ MH: roles of facilities -- R&D, scaling, persistence, collaboration (#18)
* 19. The DOE AI Townhalls had a lot of examples of people working on AI for science.
* 20. Communities in formal methods in software and automated theorem provers in mathematics are developing technologies that will be necessary to automate scientific understanding. That infrastructure is required well before the adoption of AI (beyond ML) in the scientific enterprise.
* 21. We need facilities for actively studying the AI models and techniques that are open to the scientific community. Compared to Big Tech, the DOE is a very open place for scientists to do fundamental work on understanding these methods. We need a strong theoretical understanding of data driven processes that go into building Scientific Assistants before we can deploy them in a reliable, robust, safe, application.
* 22. I'll mention a few things from DARPA programs:
Synergistic Discovery and Design (SD2) is developing the data, meta-data, computational, and ML infrastructure to insert AI into the design-test-build-learn cycle for super complex experimental fields like synthetic biology and materials science/chemistry. It includes work in machine-readable experimnetal protocols that can (in some cases) connect to robotic equipment in "cloud labs" to automatically conduct new experiments and iterate.

The NIH is using SD2 infrastructure as the starting point for their new AIBLE initiative (AI for biomedical excellence or something) which is intended to create the data and knowledge infrastructure necessary to make biological data, from molecular to social scales, machine readable and accessible to not just today's algorithms but to the algorithms of the future. This is a 7 year initiative i think.

Automating Scientific Knowledge Extraction (ASKE) is developing the tools to extract and synthesize knowledge wherever it lives (publications, source code, etc.) and create, maintain, augment, and compare computational models without writing code. Plus using those models to conduct procedures that lead to insight and meta-AI tools to learn over model structures and explore model spaces within domain restraints. obviously focused on fields where knowledge is mostly synthesized and applied in the form of computational models.
	+ Comments
	+ Yes. It seems like a lot of what has come up here is really how do we integrate a number of parallel/complementary efforts into a facility-type, data- and AI infrastructure. (while of course solving the significant challenges of scaling some of the AI tech, meeting computational demands, addressing access and fairness, etc) (#41)
* 23. ONE LIST OF SIGNPOST FODDER
	+ Comments
	+ The ability to connect small-scale agents together would be a telling signpost: precursor to facility with integrating the technology, building new and larger intelligent functions from smaller ones. (#24)
	+ The ability to connect small-scale agents together would be a telling signpost: precursor to facility with integrating the technology, building new and larger intelligent functions from smaller ones. (#25)
	+ The ability to connect small-scale agents together would be a telling signpost: precursor to facility with integrating the technology, building new and larger intelligent functions from smaller ones. (#26)
	+ The ability to connect small-scale agents together would be a telling signpost: precursor to facility with integrating the technology, building new and larger intelligent functions from smaller ones. (#27)
	+ The ability to connect small-scale agents together would be a telling signpost: precursor to facility with integrating the technology, building new and larger intelligent functions from smaller ones. (#28)
	+ The key precursor technology/ies will be intelligent systems, that, like steps on a ladder, 1) can incorporate both causal knowledge of the domain and data-driven reasoning in generating predictions and explanations (see overlap with breakout #8), 2) can integrate knowledge from disparate domains of science to generate novel causal hypotheses, 3) can design, execute, and interpret relevant experiments for testing those hypotheses, 4) can report findings from these experiments/analyses in both machine and human-comprehensible formats and 5) can, when faced with alternative projects/hypotheses, differentiate among specific research directions in terms of their novelty, or applied or theoretical value. (#29)
		- +1. I would also include the ability to generate and run dynamical models of systems capable of making predictions and prioritizing new data collection efforts. (#30)
		- I agree. Arguably that could be considered as a combination of #1 (incorporating causal/mechanistic knowledge) and #3 design/execute experiments. (#31)
		- I would like to add reproducibility to this, the ability of the intelligent system to be able to reproduce a scientific process (#32)
		- i think these are good steps for the "AI principal investigator" or whatever phase, but i would argue that there are important steps before this that will be valuable by themselves but will be precursors to these precursors. (#33)
		- Certainly. Though it's worth noting that 1 and 2 are happening now and have many proofs of concept. (#34)
* 35. The curation of open data from all federally funded experiments must be in place within 10 years in order to provide ample time for technology that requires accurate access to the corpus to develop in 20 years to hit 2050 target. The availability of the data will create an ecosystem of new research programs.
	+ Comments
	+ +1. Directly related to our "facilities" discussion. (#36)
	+ +1 (#38)
	+ I think this needs to be more nuanced. My experiment makes huge amounts (like google scale) of data. So we can trivially make it open but its absolutely not useful to anyone. (#42)
* 37. The DOE should use its incentives towards the user facilities and influence policies that encourage the facilities to coordinate further data acquisition, collection, and curation so that these data can be exploited with AI across instruments and experiments, and inter-operate more easily with computational data.
	+ Comments
	+ Also directly related to our facilities discussion (#39)
* 40. There is a significant transition in workforce skill that will be required to get over the challenges of the next 5-10 years. We need to put emphasis and resources NOW behind efforts to train our graduate students in all disciplines on the principles of computational science and data science and AI. One way to do this is through real encouragement of collaboration between bio/physical/social departments and computer science departments. this is nothing new to DOE of course (Ian Foster started the computation institute to do this 20 some odd years ago), but it would be worthwhile to re-emphasize this and update the model to reflect the current and future potential of AI.
	+ Comments
	+ There is already significant amount of conversation taking place among DOE user facilities to see how they can coordinate the data deluge that they are collectively facing. Individual facility policies and DOE wide policies are part of that discussion. However, their bias towards action means they are looking at specific use-cases and how they're enabling the existing problems with that integrated interaction. (#45)
	+ I think that we should engage universities in this. In fact, a number of leading universities have already taken big steps in this direction, by adding new courses, new cross-discplinary divisions, etc. How can the DOE engage with these efforts or do something complementary. (#47)
* 43. Programs should be initiated setting guidelines and exploring infrastructure support for attribution and accessibility of models and data, e.g. for proprietary models that are built principally from data contributed by others (see e.g. the controversy around a handful of Sloan Kettering AI researchers spinning out a startup company to read pathology slides based on contributions of many individual pathologists over many years). Government funders could set guidelines on the use of data generated by public research funding.
* 44. The idea of creating small-scale AI agents came up in several places -- this is a very plausible early step in developing the ideas, and in developing tool chains, and in perhaps developing AI agents as components should such Gaggle-of-Agents approach prove useful. I'd say moving in those directions are already happening, but perhaps not with the focus we've discussed. Concrete progress could be seen in 5 year time frame with significant development by 10 years.
	+ Comments
	+ as for when this precursor (small scale components) would be NEEDED, it should be an early achievement (5 to 10 at most) in order to provide critical feedback to the development timeline (#55)
* 46. IBM Watson seems to be an example of precursor and scientific community can profit from such technology. In the next 5 years, it seems plausible that national labs can provide such technology for a variety of key domain experts
	+ Comments
	+ This does fit well into several parts of our discussion about capabilities and facilities (#62)
* 53. Support for hypothesis generation, management, evaluation (a la signpost #50) does seem plausible and at a guess might be further out than 10 years before it is a routine component of scientific workflows.
* 58. Demonstrating the contribution of an AI method/agent on one concrete scientific problem in one domain is important and can serve as a template for generalization
* 60. AI in support of reproducibility should be in development early with wide-spread use within 20 years -- as a goal.
* 63. Notes
	+ Comments
	+ MH: Looks like bits and pieces of the signposts we have listed are already being worked on with a variation is when those will be available. (#65)
	+ MH: Which is the hardest of our signposts? (#66)
	+ JB: #1 is the topic of a whole other breakout at the unconference (#68)
	+ JB: #1 is also being actively worked on in the national labs, academia, and industry today (#69)
	+ JE: New program called "Machine common sense" where the goal is to create an AI that has the common sense of a 2-year old (#70)
	+ JE: The examples in causal reasoning is very specific but what we want is some general reasoning that can get us to an interesting conclusion. (#71)
	+ LP (as notetaker) : perhaps figuring the right questions to ask for an AI system would be one step towards plausibility of a generelized approach for bullet 1 (#72)
	+ JE/JB: Is this that the specific is needed before the generic is available? We need to do the specifics but we almost need to do the generic in parallel. (#73)
	+ SP: An AI system that constructs, hypotheses, and then make a better prediction than what humans can do today (#74)
	+ LP: One of the things needed is to work towards integrating data from disparate sources especially making data from failed experiments available. This leads to tremendous re-inventing the wheel. (#75)
		- MH: This could fall nicely within the reproducibility context. (#76)
		- LP: Lot of the times though this is a goal but may never be achieved. Instead making the research transparent could assist in creating these AI assistants of the future. (#77)
		- SP: There's a social aspect to get over because of the competitive aspect of sharing failures. (#78)
		- LP: This could be done by an AI agent that could keep it anonymous and share across labs. (#79)
		- BG: Perhaps the suggestion is to change the incentive structure to incentivize people to share all data and findings and not just the successful ones (#80)
		- JB: Even if DOE is unable to change incentives, then making the data tools available to be able to share across and adding a positive value to a negative experiment can make a difference. (#81)
	+ BG: Specific vs general - in the next 5 years we focus on how AI can solve the specifics and then over the next 5 years after that we can generalize to other domains where we have less of an opportunity to make specifics (#82)
	+ JE: The key infrastructure pieces that need to be changed include (1) computational (2) metadata, etc. (#88)
	+ NS: What happens when we cannot understand AI? Maybe the hypothesis doesn't make sense to us but it makes complete sense to it? Are we going to call this a failure? (#89)
		- JF: Not just that the AI is explainable but that the AI is creating the explanation of the phenomena. (#90)
		- JE: You could have an AI that crunches a lot of data and humans can go in and understand what the AI did in that context and that's a more near-term signpost. (#91)
		- NS: When looking at fundamental science that is very complicated for human understanding if an AI comes along and comes up with a hypothesis that cannot be explained, then how do we proceed? (#92)
			* Analogy to cyber security investigation. An intrusion detection system provides an alert to the security team. An explanation of the model is a description of the math that the system used to determine the presence of the cyber attack. When the DOJ goes to prosecute the cyber criminals, they need an explanation not of the model, but of the attack. They need to determine who used what exploits and code to attack which resources and construct a narrative explaining the structure and execution of the attack. That narrative is an explanation of the phenomenon instead of an explanation of the model. (#96)
	+ JB: If the machine is producing the decisions being made (like in cell biology) then what role do the humans play in this scenario - are we just the people who maintain the decision engine? (#94)
	+ HG: What isn't plausible to me is allowing the 'agent' the agency to take decisions about experiments or the scientific direction of the field. This would imply a huge amount of trust, which I don't see a path towards. (#98)
	+ Attempt at a "timeline of signposts": 4-ish years. AI-system reads entire body of scholarly works and generates its own curated version of WikiScholar, which becomes most widely-used technical reference system on internet for students and researchers. 5-ish years. 32 papers published in Science, Nature, PNAS in 1 year use critical datasets curated from literature by AI technical assistant. AI assistant spurs key new data collection efforts. 6-ish years. AI-system integrated into bio- and climate- modeling system at national facilities; which retrieves, curates and updates models with instrument/literature data in near real time. 12-ish years. AI-system constructs, parameterizes, and tests with data streams a novel Earth systems model that proves to better predict some dynamical responses to carbon forcing than current human state-of-art models. 20+ years. Nobel-prize winning scientists gets in protracted scientific argument with an AI agent about the plausibility of her hypothesis. She is perceived to win, which sets back AI research. AI agent proven right 12 years later by novel instrumentation that yields new measurements. (#99)
* 83. The key precursor technology/ies will be intelligent systems, that, like steps on a ladder, 1) can incorporate both causal knowledge of the domain and data-driven reasoning in generating predictions and explanations (see overlap with breakout #8), 2) can integrate knowledge from disparate domains of science to generate novel causal hypotheses, 3) can design, execute, and interpret relevant experiments for testing those hypotheses, 4) can report findings from these experiments/analyses in both machine and human-comprehensible formats and 5) can, when faced with alternative projects/hypotheses, differentiate among specific research directions in terms of their novelty, or applied or theoretical value.
* 84. The key precursor technology/ies will be intelligent systems, that, like steps on a ladder, 1) can incorporate both causal knowledge of the domain and data-driven reasoning in generating predictions and explanations (see overlap with breakout #8), 2) can integrate knowledge from disparate domains of science to generate novel causal hypotheses, 3) can design, execute, and interpret relevant experiments for testing those hypotheses, 4) can report findings from these experiments/analyses in both machine and human-comprehensible formats and 5) can, when faced with alternative projects/hypotheses, differentiate among specific research directions in terms of their novelty, or applied or theoretical value.
* 85. The key precursor technology/ies will be intelligent systems, that, like steps on a ladder, 1) can incorporate both causal knowledge of the domain and data-driven reasoning in generating predictions and explanations (see overlap with breakout #8), 2) can integrate knowledge from disparate domains of science to generate novel causal hypotheses, 3) can design, execute, and interpret relevant experiments for testing those hypotheses, 4) can report findings from these experiments/analyses in both machine and human-comprehensible formats and 5) can, when faced with alternative projects/hypotheses, differentiate among specific research directions in terms of their novelty, or applied or theoretical value.
* 86. The key precursor technology/ies will be intelligent systems, that, like steps on a ladder, 1) can incorporate both causal knowledge of the domain and data-driven reasoning in generating predictions and explanations (see overlap with breakout #8), 2) can integrate knowledge from disparate domains of science to generate novel causal hypotheses, 3) can design, execute, and interpret relevant experiments for testing those hypotheses, 4) can report findings from these experiments/analyses in both machine and human-comprehensible formats and 5) can, when faced with alternative projects/hypotheses, differentiate among specific research directions in terms of their novelty, or applied or theoretical value.
* 87. The key precursor technology/ies will be intelligent systems, that, like steps on a ladder, 1) can incorporate both causal knowledge of the domain and data-driven reasoning in generating predictions and explanations (see overlap with breakout #8), 2) can integrate knowledge from disparate domains of science to generate novel causal hypotheses, 3) can design, execute, and interpret relevant experiments for testing those hypotheses, 4) can report findings from these experiments/analyses in both machine and human-comprehensible formats and 5) can, when faced with alternative projects/hypotheses, differentiate among specific research directions in terms of their novelty, or applied or theoretical value.
* 97. It seems like one of the less plausible signposts is 5) above (automated program management/scientific direction choosing) because the knowledge required to make those decisions lies outside of the realm of what is accessible and the utility functions are political.
* 100. Recycle bin
	+ Comments
	+ Attempt at a "timeline of signposts": 4-ish years. AI-system reads entire body of scholarly works and generates its own curated version of WikiScholar, which becomes most widely-used technical reference system on internet for students and researchers. 5-ish years. 32 papers published in Science, Nature, PNAS in 1 year use critical datasets curated from literature by AI technical assistant. AI assistant spurs key new data collection efforts. 6-ish years. AI-system integrated into bio- and climate- modeling system at national facilities; which retrieves, curates and updates models with instrument/literature data in near real time. 12-ish years. AI-system constructs, parameterizes, and tests with data streams a novel Earth systems model that proves to better predict some dynamical responses to carbon forcing than current human state-of-art models. 20+ years. Nobel-prize winning scientists gets in protracted scientific argument with an AI agent about the plausibility of her hypothesis. She is perceived to win, which sets back AI research. AI agent proven right 12 years later by novel instrumentation that yields new measurements. (#101)

3 Day Three - November 10, 2020

3.1 Breakout 5 - Pitfalls and Roadblocks

Participants: 1

**Brainstorm question or instruction:**
Breakout 5 - Pitfalls and Roadblocks
Detailed discussions on identifying pitfalls and potential roadblocks. If possible, list in rank ordering. o What could prevent the technology/service/device from being developed (funding, materials, policies, researchers, operations staff, etc.)? o How will progress be measured/evaluated? o How will lack of progress be measured/evaluated? o Who will decide if progress is being made? o What are the consequences of not engaging in this area?

**Sticky points:**

 Top Takeaways (5 points per participant)

* 1. Pitfalls/roadblocks:
- lack of data sharing and reusability
- lack of community standards (incl. for controlling instrumentation, accessing data, accessing literature, etc.)
- incentives and funding not supporting to above explicitly

3.2 Breakout 6 - Keys to Success

Participants: 0

**Brainstorm question or instruction:**
Breakout 6 - Keys to Success
Identify who needs to be engaged, research communities, domain science communities, staff, management. Identify needed skills and knowledge (give examples) o What benefits would society obtain? o What benefits would the science/research community obtain? o What research communities need to be involved? o What domain science communities need to be involved? o What staff and management communities need to be involved? o What kind of management structure is required? o How broadly will this impact society and/or the science community?

**Sticky points:**

 Top Takeaways (5 points per participant)

Appendix

Live chat

**The following messages were exchanged via Live chat.**

* Nov 2, 2020, 19:54 UTC
	+ I think this cartoon is essential for this conversation: (joshua elliott | Nov 2, 2020, 19:54 UTC)
	+ https://matteofarinella.com/Chaos-In-The-Brickyard (joshua elliott | Nov 2, 2020, 19:54 UTC)
	+ sorry, full cartoon here: https://massivesci.com/articles/chaos-in-the-brickyard-comic-matteo-farinella/ (joshua elliott | Nov 2, 2020, 19:55 UTC)
	+ I like it! (Shanan Peters (U. Wisconsin-Madison) | Nov 2, 2020, 20:05 UTC)
	+ Wow - that was enlightening. Definitely useful for our discussion (Jini Ramprakash | Nov 2, 2020, 20:19 UTC)
	+ (shannon) Interactions need to be conversational, not code. Domain scientists are not trained to be sufficiently adept at programming (joshua elliott | Nov 2, 2020, 20:25 UTC)
	+ (ben) there are many domains where the essential knowledge is simply expressed as descriptions in text and difficult to extract and structure and etc. (joshua elliott | Nov 2, 2020, 20:31 UTC)
	+ (mark) you have an AI and yr grad student has an AI and you use your AIs to communicate back and forth and help synthesize and etc. the Superhuman Grad Student? (joshua elliott | Nov 2, 2020, 20:38 UTC)