

UC Berkeley

UC Berkeley Previously Published Works

Title

Understanding the State of Scientific AI: User Needs and Design Implications

Permalink

<https://escholarship.org/uc/item/8d32b7dw>

Authors

Cohoon, Hannah

Paine, Drew

Publication Date

2026-10-16

Copyright Information

This work is made available under the terms of a Creative Commons Attribution License, available at

<https://creativecommons.org/licenses/by/4.0/>



Understanding the State of Scientific AI: User Needs and Design Implications

Study overviews & high level insights from:

- HPDF AI User Needs Study
- BERIL Co-Scientist Study

Oct 16, 2026

Johanna (Hannah) Cohoon, hcohoon@lbl.gov

Drew Paine, pained@lbl.gov

Product & Partnership Experience Team

Table of Contents

[Summary of methods and results from both work streams](#)

[Supplement: Full Takeaways from Both Studies](#)

[Takeaways from Interviews with BERIL Co-Scientist Users](#)

[Takeaways from the AI Needs User Study](#)

Summary of methods and results from both work streams

Two Work Streams Investigating AI Use & Needs

1

AI user needs in a modern data facility

.....

Identifying use cases for AI tooling based on current behaviors self-reported in June-Aug 2026

2

AI Co-Scientist experiences during a May 2026 workshop

.....

Understanding users' experiences with an AI co-scientist and what those imply about UX needs

Work stream 1: Studying **AI user needs**

AI Needs Study Approach: Goals & Methods

Goals

- Identify opportunities to integrate AI into a data facility's UX
- Understand which roles are best suited to take advantage of a given opportunity
- Understand stakeholders' expectations for agentic AI in a modern data facility
- Understand how users determine the utility of AI tools
- Recognize concerns users have about AI tools

Methods



- 18 hour long, semi-structured interviews during summer 2026
- Qualitative thematic analysis
- 10 interviews have been fully analyzed, all have been reviewed multiple times

These results are primarily based on the 10 fully analyzed interviews, with little expectation that overall findings change as analysis is completed

Interviews focused on how users leverage AI now, particularly in ways beyond writing code

AI Needs Participants Summary: Demographics

18 participants

INSTITUTION

14 from national labs
4 from universities

SENIORITY (AS DEFINED BY THE PARTICIPANT)

7 early career
7 mid career
1 senior

Roles⁺

Design Engineer Liaison

Engineer

Faculty Scientist / Associate Professor

Group lead

Lab manager

Postdoctoral Researcher

Principal Engagement Manager

Professor (3)

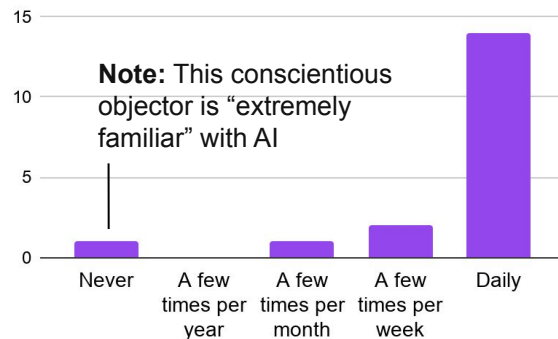
Researcher (7)

Software Engineer (3)

Common disciplines:

Biology, CS, Materials Science,
Environmental Science, Physics

Use of AI



+Some participants listed multiple roles and disciplines; these are all reflected here

AI Needs Participants Summary: Spectrum of AI endorsement

Most participants found use cases they're interested in, while maintaining some uncertainty about how best to use AI and when it is appropriate

P14: Actively campaigns against AI, believing it to be harmful in many ways and irresponsibly advertised and adopted.

P7: Uses AI all the time, pushes people to communicate in repos so that they can leverage agents better

P11: Creates personalized apps to automate as many workflows as possible. Highly invested in sharing AI strategies.

Work stream 2: The KBase **BERIL Co-Scientist** study

Additional results:

J. Cohoon et al., "Toward Calibrated Trust: What Working Scientists Need from an AI Co-Scientist", Jun. 08, 2026, Zenodo. doi: 10.5281/zenodo.20595967.

BERIL Co-Scientist Study Approach: Goals & Methods

Context

- BERIL is an agentic co-scientist with access to the KBase lakehouse (beril.kbase.us; beril-doe.org).
- The agent takes an RQ, produces a research plan, carries that out
- Researchers used BERIL to explore their own RQs during a one-day workshop in early May 2026

Goals

- Understand how users move from idea to analysis and their expectations for publishing outputs
- Recognize concerns users have about the Co-Scientist
- Examine users' trust and relationship to the Co-Scientist

Methods

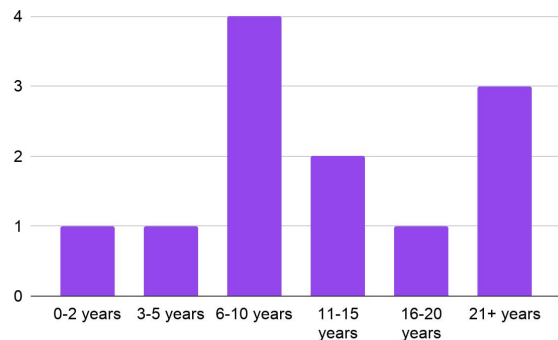
- 12 interviews with 11 participants:
 - Six short interviews during the workshop
 - Six extended interviews after the workshop
 - Qualitative thematic analysis
- Surveys before, during, after workshop

BERIL Co-Scientist Participant Summary: Demographics

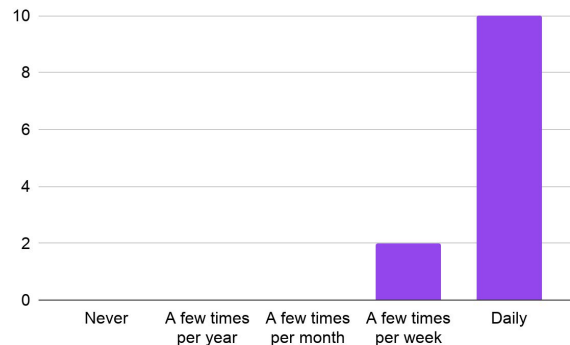
38 workshop
participants

12 interviews w/ 11
participants

Years Professional Experience



Use of AI



BERIL interviewees are more senior than AI User Needs interviewees, but have similar use patterns

User Roles and Their Relationships to Use Cases

User Roles

Modern data facilities have multiple types of users to consider with respect to AI tooling

Users in different roles assume different responsibilities and have different needs

The AI Needs and BERIL Co-Scientist studies focused on:

PROJECT MEMBERS

PIs/PROJECT ADMINS

DATA STEWARDS

Results

Note: Opportunities identified are not direct requests, but are insights derived from the research

Opportunities for AI to Assist with Data & Compute Tasks; Relevant Roles

Data & compute workflows

FOR PIs/PROJECT ADMINS

FOR PROJECT MEMBERS

FOR DATA STEWARDS

- Support data discovery, cleaning, analysis, visualization leveraging project context (e.g., allocation requests, project summaries, meeting notes)
- Format data and maintain metadata within Community Zone collections and individual projects
- Help users request resources and launch HPC environments
- Enable assistants such as GitHub Copilot to access facility data and tools from environments like VS Code
- Summarize datasets and other public resources (contents and structure)

Examples

PROJECT MEMBERS

PIs/PROJECT ADMINS

Participant 1: "So if there was just a free-form text field and I said, 'Hey, this is what I'm looking for...I need the date-ordered buy-offs for all of these units,' and then...it would create the query for me that gives me the data I need, or at least it starts the query for me and I can tweak it....that would help me because I don't understand the structure of the database and how to form the query."

PROJECT MEMBERS

PIs/PROJECT ADMINS

Participant 9: "I have a loop, an agentical loop to design a physics analysis....And I want the agent to create an analysis strategy for a specific signature I want to look for."

DATA STEWARDS

Participant 18: "We were collecting photographs for 1,000+ sites. And so I ended up using AI to help set up workflows to seamlessly download the images and label them based on the site numbers and organize them."

Opportunities for AI to Assist with Research Tasks; Relevant Roles

Research assistance

FOR PIs/PROJECT ADMINS

FOR PROJECT MEMBERS

- Summarize recent project work, leveraging Zoom assets and facility logs, acting as a memory aid and monitoring tool
- Connect lit review agents to project context and project data to find more relevant references
- Leverage logs and project context to support documentation and reporting (e.g., reproducibility artifacts, progress reports)
- Create study plans

Examples

PROJECT MEMBERS

PIs/PROJECT ADMINS

Hannah: "Once the meeting is over, then what happens to those notes?"

Participant 16: "We usually take them and put them in our agent memories so they're aware of these things."

PROJECT MEMBERS

PIs/PROJECT ADMINS

Participant 7 shows a prompt in a Github issue: *Send five parallel high effort Edison Scientific queries based on topics and current cited literature in the project proposal.* "And so looking back at one of those, it's opening this pull request....It's chosen its own prompts to send to Edison Scientific, which will then run agents for itself...You get a set of references, non-hallucinated, includes peer-reviewed ones, and then kind of this overview and, and different reasoning pieces there.

Implementation Lessons from BERIL

LESSON

Key steps and available actions must be visible; they should not blend in with more trivial tasks

Frequent permissions requests from the Co-Scientist were fatiguing to users and they stopped paying attention to what was being asked of them. Some users were concerned they missed more important questions in the flood of repetitive permissions questions

LESSON

Humans work iteratively on research projects; agents must support this

Agents should check in with users at key decision points, providing artifacts like scripts, datasets, and plots for inspection, even if a research plan was approved first (with an option to bypass these checkpoints for confident users). This provides opportunity to redirect analyses.

LESSON

When discovering data, people need to be able to explore a dataset before using it

Participant 11 gives an example of why this is necessary: “If you look at the group comparison... of things that had [a particular] enzyme and things that don't, the question would be, ‘Am I data limited before I do that, right?’... Do I only have like two genomes that are [negative]? I couldn't do a statistical test if I didn't know that, right?...Visually looking at it might be helpful.”

LESSON

Users do not want to have to redirect the Co-Scientist and ask it to redo work

Rework is a useful metric for the facility to track—when a human has to intervene and direct the AI to try attempt a task again in a new way, this is a signal of a poor user experience

Opportunities for AI to Assist with Research, Data, & Compute Tasks

Enabling features

- Provide easy access to AI logs for agent monitoring (live and post hoc), but allow monitoring windows to be minimized or navigated away from to reduce unnecessary distraction
- Store or provide access to project context (e.g., Zoom assets, personal notes, grant documents, code repositories)
- Support text and voice interaction with agents
- Offer a skill and prompt library with reusable agent instructions and workflows to reduce redundant efforts, promote standardization, and provide access to common tools
 - Include benchmarks to set performance expectations and offer ways to contact other skill users for mutual support
- Add users to an onboarding project when they sign up, showing them how AI tooling is integrated with the facility

LESSON FROM BERIL

Users need signals/support to help them track the AI's process

Users consistently requested **checklists** and other means of tracking what the Co-Scientist was doing

Expectations: Requests for Tailored Experiences and HPC Access

When discussing a hypothetical data facility:

- **Should offer personalized experiences**
 - Reminders of current tasks
 - Agents approach tasks the same way project members would
 - Write in their voice
 - Ability to share skills with project members
 - Access to domain specific tooling
- **Should not compromise their IP—need ability to keep skills and data private from other projects**
- **Should simplify HPC access**
 - Need increased resources for use, improved ability to run experiments in parallel
 - Want to reduce the need to monitor agent or be involved in handshakes
- **Support data discovery**
- **General preference for multiple agents or well managed memory, believing this to lead to better outcomes**



Access to Github logs, Zoom artifacts, and other project context would improve personalization and could reduce iteration

LESSON FROM **BERIL** + **AI NEEDS**

Research data and project context can be key assets that draw users to engage with a co-scientist/science agent rather than a generic LLM

BERIL users compared the Co-Scientist to their usual LLM of choice: When one user couldn't access the research data they needed through the lakehouse, they wondered why they would use the Co-Scientist at all instead of their usual Claude

AI User Needs study shows that the data that matters is not just scientific data, but project context too.

An open question

How should a facility support some of those use cases?

Is the facility the centralizer, bringing in project context and preferred agents to one place?

Is the facility also a data provider, allowing users to give their external agents access to logs and data?

Does it do both? Or will the facility not connect project context at all?

Judging Utility

Users primarily judged AI utility based on how much more quickly they get through a task

- Participants appreciate speed and ability to do tasks that were previously impossible or extremely difficult (because of scale or lack of knowledge)
- Some questioning of when AI is most appropriate—participants had no clear cut guidelines to reference
- Benchmarks needed to help users judge which approaches are most effective



A data facility can learn from aggregate use patterns, leveraging logs to provide performance benchmarks and recommend skills and agents for certain tasks

Concerns & Constraints on AI Use

Users must be able to opt out of AI tools, interfacing with the data facility in more traditional ways

Constraint - Cost & Limited Tokens

- Token access is limited and costly; facility can offer value by allocating tokens
- Users will need to be able to switch their agents when token limits are reached

Constraint - Security & Regulation

- Institutional security policies limit AI use (but users will push boundaries!)
- Need clarity on which policies impact users, especially for cross-facility workflows
- Need for thoughtful regulation developed with users

Constraint - Awareness

- Lack of knowledge of best practices and ideal use cases; need for information sharing among users

Concern - Critical Thinking

- Need for useful friction to slow down work, encouraging critical thought and knowledge (not just data) production

Concern - Privacy & Hesitance to Share

- Users hesitate to show unrefined AI outputs with peers, partly because it's seen as unhelpful—what outputs and logs should project members share with one another?
- Users want to be sure their valuable data or AI skills are not being illicitly shared with others

Concern - Expertise

- Students/junior colleagues do not always have the necessary knowledge to accurately validate AI output
- References, benchmarks, and contextually relevant information is important to help users validate AI outputs
- Monitoring tools can help mentors see output quality

Concern - Environment

- Concern around environmental impact (but not usually enough to truly limit AI use)
- Careful messaging needed around environmental impact
- Design choices should minimize water and energy usage

Follow up Research Opportunities

- 1 People like summaries and accelerated work, but need to be intellectually engaged. How does a data facility introduce the right amount of friction in the right places?
- 2 What policies do users want to see around AI in a data facility?
- 3 Do users want the facility to centralize their existing resources (data, code, project documents and notes) or should it be a source of information that can be leveraged by external tools?
- 4 What AI logs/outputs get shared within teams and what should be personal?
- 5 What are the key steps in mental models for data discovery, workflow execution, script writing and editing, etc.? How do mental models differ among user archetypes?

Supplement: Full Takeaways from Both Studies



Takeaways from Interviews with BERIL Co-Scientist Users

Lessons for a modern data facility

June 2026

HPDF Product & Partnership Experience Team

Context

BERIL is an agentic co-scientist with access to the KBase lakehouse (beril.kbase.us; beril-doe.org)

The agent takes a research question (RQ), produces a research plan, carries that out

38 BER researchers used BERIL to explore their own RQs during a one-day workshop

We gathered logs, conducted surveys, and held interviews with participants. These slides report on findings from interviews, as does [this report on Zenodo](#)

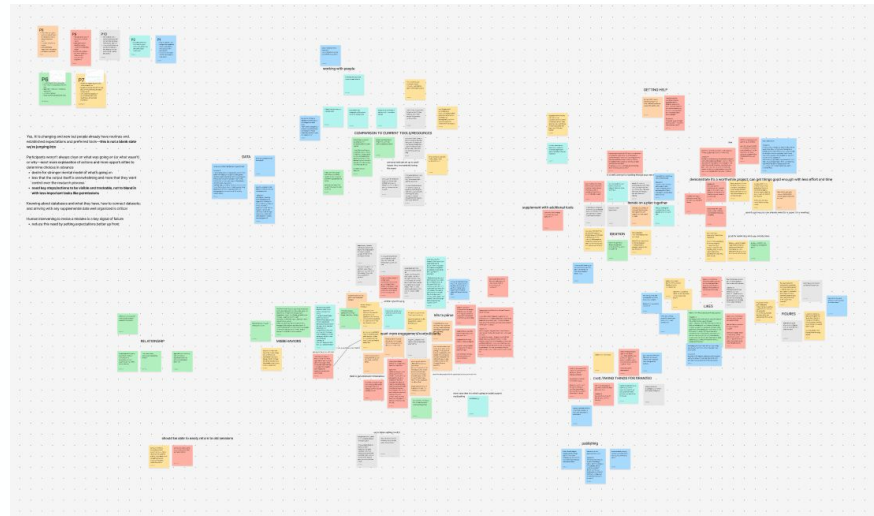
An HPDF User Experience Researcher planned the surveys, conducted and analyzed interviews, and contributed substantially to report writing

Research Approach

12 interviews with 11 participants:

- Six short interviews during the workshop
- Six extended interviews after the workshop

Qualitative thematic analysis to identify key takeaways



A screenshot from early in the thematic analysis process

Summary: All Study Takeaways

- 1 When working with unfamiliar data, users need to be able to explore it before incorporating it into an analysis
- 2 Users have a more iterative process than the Co-Scientist; this misalignment felt like poor usability
- 3 Users need signals/support to help them track the AI's process
- 4 Key steps and available actions must be visible; they should not blend in with more trivial tasks
- 5 Data is a key asset that draws users to engage with a co-scientist/science agent rather than a generic LLM
- 6 Users want to avoid rework

Summary: All Implications for a Modern Data Facility

- Agents should provide humans the opportunity to explore & tinker with data before proceeding
- Agents should check in with users at key decision points, providing artifacts (e.g., scripts, datasets) for inspection, even if a research plan was approved first (with an option to bypass these checkpoints for confident users)
- Need to know: What are the key steps in mental models for data discovery, workflow execution, script writing and editing, etc.? How do mental models differ among user archetypes?
- It should be easy for users to distinguish between an agent's high impact science/strategy questions and more procedural questions
- It should be easy for users to configure their agents' permissions so that they do not need explicit permission for each individual action
- An value proposition for a modern data facility is strongly affirmed: Enable data discovery and use through improved federation, breaking down silos.
- A promising marketing angle: Users don't have to bring data to an LLM, they can work with it in situ

Key Takeaways

Takeaway 1

When working with unfamiliar data, users need to be able to explore it before incorporating it into an analysis

Users need to explore data to be able to judge if it can be used to answer their research question

Users wanted to be “getting a description of what's in each table and looking at a few sample rows from each table” (Participant 2)



In workflows where agents help users discover and work with data, agents should provide humans the opportunity to explore and tinker with data before proceeding

Participant 11 gives an example of why this is necessary:

“If you look at the group comparison... of things that had [a particular] enzyme and things that don't, the question would be, ‘Am I data limited before I do that, right?’... Do I only have like two genomes that are [negative]? I couldn't do a statistical test if I didn't know that, right? So I'd want to— so visually looking at it might be helpful.”

Takeaways 2 & 3

Users have a more iterative process than the Co-Scientist; this misalignment felt like poor usability

Users need signals/support to help them track the AI's process

Users wanted to inspect data and outputs, working iteratively, but the Co-Scientist worked linearly

Users consistently requested checklists and other means of tracking what the Co-Scientist was doing

Aligning to users' established mental models can improve the user experience:

- Helps users feel more in control and able to trust the Co-Scientist
- Helps users avoid bad paths and potential rework



AI agents should check in with users at key decision points, providing artifacts like scripts, datasets, and plots for inspection, even if a research plan was approved first (with an option to bypass these checkpoints for confident users)



A data facility needs to know: What are the key steps in mental models for data discovery, workflow execution, script writing and editing, etc.? How do mental models differ among user archetypes?

Takeaway 4

Key steps and available actions must be visible; they should not blend in with more trivial tasks

Frequent permissions requests from the Co-Scientist were fatiguing to users and they stopped paying attention to what was being asked of them.

Some users were concerned they missed more important questions in the flood of repetitive permissions questions



It should be easy for users to distinguish between an agent's high impact science/strategy questions and more procedural questions



It should be easy for users to configure their agents' permissions so that they can work "dangerously" if they wish, *i.e.*, without explicit permission for each individual action

Takeaway 5

Research data is a key asset that draws users to engage with a co-scientist/science agent rather than a generic LLM

BERIL users emphasized that its power comes from connecting multiple data sources in an environment where they can easily perform analysis

Data need to be “good quality, reliable, uniformly processed” (Participant 11) and users need to understand how it is organized to take advantage



Reaffirms one of a modern data facility's value propositions: Enable data discovery and use through improved federation, breaking down silos.



Indicates a promising marketing angle: Users don't have to bring data to an LLM, they can work with it in situ

BERIL users compared the Co-Scientist to their usual LLM of choice:

When one user couldn't access the data they needed through the lakehouse, they wondered why they would use the Co-Scientist at all instead of their usual Claude

Takeaway 6

Users want to avoid rework

Users considered it a negative thing to have to redirect the Co-Scientist and ask it to redo work



Rework is a useful metric for a data facility with AI to track—when a human has to intervene and direct the AI to try attempt a task again in a new way, this is a signal of a poor user experience



Supplemental Information

Additional methodological information can be found in:

J. Cohoon et al., "Toward Calibrated Trust: What Working Scientists Need from an AI Co-Scientist", Jun. 08, 2026, Zenodo. doi: 10.5281/zenodo.20595967.

20 min Interview Protocol (During Workshop)

1. What phase of your work are you at right now? *(Have you produced a research plan yet?)*
2. What routines have you established while working today? *(What strategies have been helpful? Have you noticed the Co-Scientist repeatedly behaving in a particular way—how do you respond to that?)*
3. What do you think about the way the Co-Scientist engages with you?
 - a. Is it asking you the right questions at the right time? Why do you say that?
4. What's surprised you about your interactions with the Co-Scientist today?
5. What else should I know about?

60 min Interview Protocol (After Workshop)

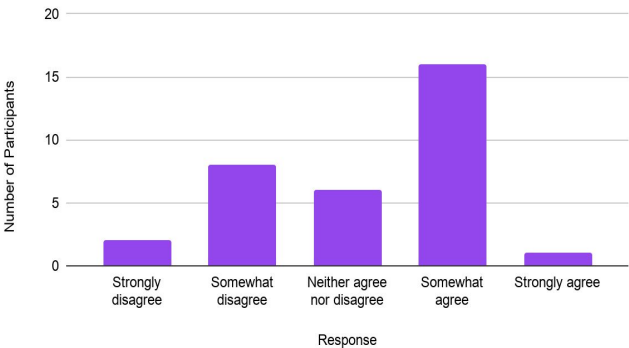
1. To start, a quick question: How technical would you characterize yourself as being?
2. Okay and can you share your general impressions from the KBase Co-Scientist event?
3. You told us in your earlier survey that these are the steps you usually take to carry out a study: **xxx**
Thinking of your experience with the Co-Scientist, is this the approach you used?
 - a. What was different?
 - b. What steps did the Co-Scientist carry out that would usually involve another human?
4. Can you share an example from the workshop of when you had to change your strategy or approach to a task?
5. Were you involved in the research process in the ways you'd like to be?
 - a. Why do you say that?
6. We asked in the surveys about whether the Co-Scientist was a collaborator, a supervisor, a thought partner....but maybe those aren't the words you would use. How would you characterize your relationship to the Co-Scientist?
7. Comparing your answers from the first survey to the last one, you've shifted some in how much you think you will use AI to _____. Can you tell me some about your thinking there?
8. In the survey, your responses suggested that you found it [easy/difficult] to work with the Co-Scientist. Tell me some more about that.
9. You described yourself as feeling _____ in one of the surveys during the workshop. Why were you feeling that way?
10. What do you think the Co-Scientist's major limitations are?
11. Do you think you produced something like a toy project, or something more publishable? How come?
12. Based on your experience at the workshop, what expectations do you have for how publishing will look in the future?
 - a. What about scientific collaboration—do you foresee any changes in how we team up to do research?
13. Are there any other notable experiences you'd like to share with us from the workshop?
14. As you know, we're going to be writing a meta paper about the Co-Scientist and the event. Based on your own experience, what do you think the abstract should say?
15. Is there anything else you'd like to say?

Interviewees

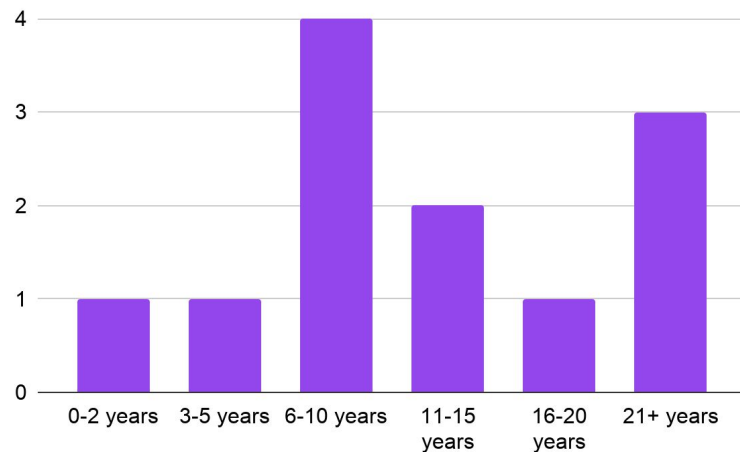
Person ID	Frequency of AI use	Post-event agreement with “I am confident in the AI Co-Scientist”
P1*	Daily (reported in interview)	(no response)
P2*#	Daily	Strongly disagree
P3*	Daily	Somewhat agree
P4*	Daily	Somewhat disagree
P5*	A few times per week	Somewhat agree
P6#	Daily	Neither agree nor disagree
P7#	Daily	Somewhat agree
P8#	Daily	Somewhat agree
P9#	A few times per week	Somewhat agree
P10#	Daily	Strongly disagree
P11*	Daily (reported in interview)	Somewhat agree
*participated in extended interview #participated in brief interview		

All workshop participants

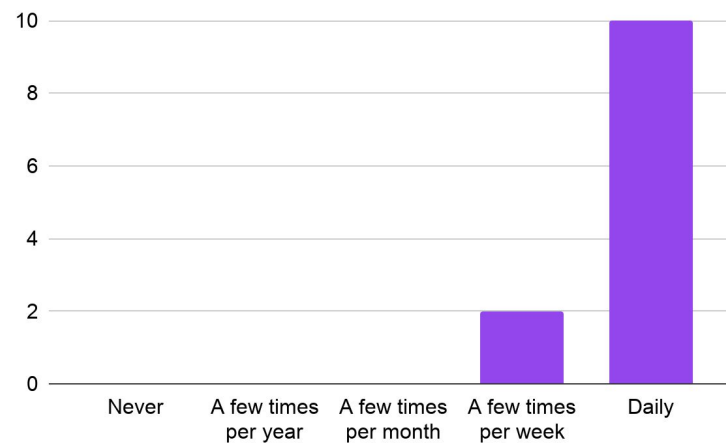
Post-event agreement with "I am confident in the AI co-scientist"



**Co-Scientist Interviewees' Years
of Professional Experience**



Co-Scientist Interviewees' Use of AI



Funding Acknowledgement for Co-Scientist Study

The Biological and Environmental Research Intelligence Layer (BERIL) is funded by the U.S. Department of Energy, Office of Science, Office of Biological and Environmental Research under Award Numbers DE-AC02-05CH11231, DE-AC02-06CH11357, and DE-AC05-00OR22725.

The DOE Systems Biology Knowledgebase (KBase) is funded by the U.S. Department of Energy, Office of Science, Office of Biological and Environmental Research under Award Numbers DE-AC02-05CH11231, DE-AC02-06CH11357, DE-AC05-00OR22725, and DE-SC0012704.

Cohoon's work is supported by the High Performance Data Facility (HPDF) through the U.S. Department of Energy, Office of Science under contracts 89243126CSC000213 (TJNAF) and DE-AC02-05CH11231 (LBNL).



Takeaways from the AI Needs User Study

Early Results

Sept 2026
Product & Partnership Experience Team

Key Takeaways

1

Context is key

Prolific AI users attempt to supply AI tools with as much relevant information as they can to improve outputs and reduce iteration

2

Summaries needed

Users want AI to help them quickly grasp large amounts of information about datasets, code repos, project progress, literature, etc.

3

Tokens are in demand

AI use is frequently constrained by cost or the limited availability of institutional supplied tokens

*! Despite the need for speed, some participants expressed concern about accelerated research, noting that using AI can reduce how much critical thinking they do. **Some friction is useful.***

About the Study & Participants

Study Approach: Goals & Methods

Goals

- Identify opportunities to integrate AI into a data facility's UX and understand which roles are best suited to take advantage of a given opportunity
- Understand stakeholders' expectations for agentic AI in a modern data facility
- Understand how users determine the utility of AI tools
- Recognize concerns users have about AI tools

Methods



- 18 hour long, semi-structured interviews
- Qualitative thematic analysis
- 10 interviews have been fully analyzed, all have been reviewed multiple times

These results are primarily based on the 10 fully analyzed interviews, with little expectation that overall findings change as analysis is completed

Study Approach: Interview Focus

Because people are not good at predicting their behavior, the study focused on **how users leverage AI now.**

Because we already know AI is frequently used when writing code, we sought examples of **use cases other than writing code.**

Opportunities identified are not explicit requests made by users, but use cases we do know to be of value.

Participants Summary: Demographics

18 participants

INSTITUTION

14 from national labs
4 from universities

SENIORITY (AS DEFINED BY THE PARTICIPANT)

7 early career
7 mid career
1 senior

Roles⁺

Design Engineer Liaison
Engineer
Faculty Scientist / Associate Professor
Group lead
Lab manager
Postdoctoral Researcher
Principal Engagement Manager
Professor (3)
Researcher (7)
Software Engineer (3)

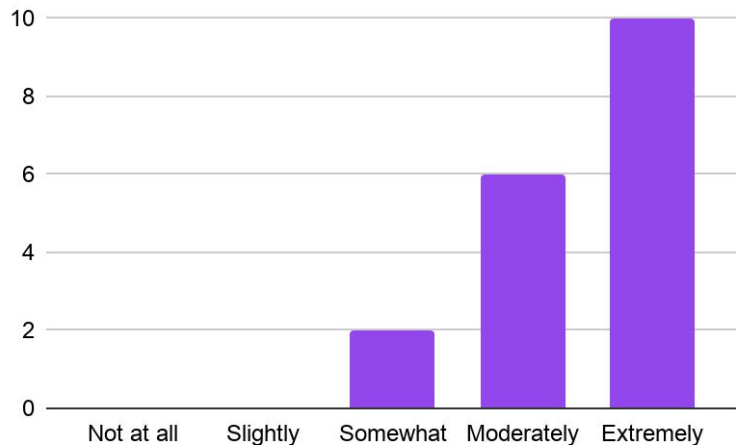
Disciplines⁺

Bioengineering
Computational Biology
Computational Materials Science
Computational Science & Engineering
Computer Science (3)
Earth & Environmental Sciences (2)
Electrical Engineering
Environment Science (2)
Materials Science
Mechanical Engineering
Physics (4)
Social Psychology

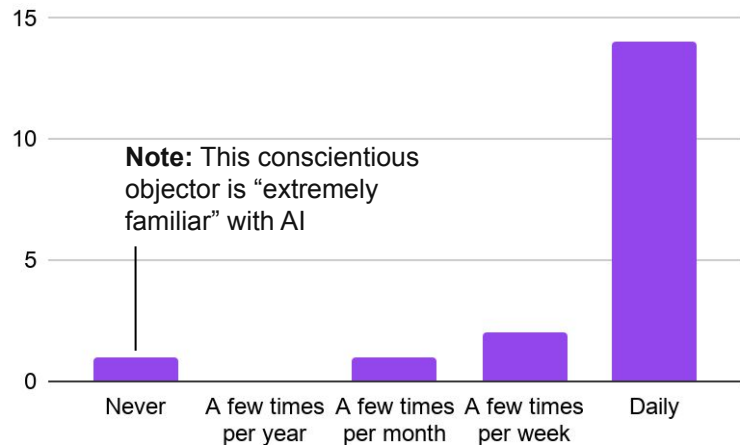
+Some participants listed multiple roles and disciplines; these are all reflected here

Participants Summary: AI Experience

Familiarity with AI



Use of AI



Participants Summary: Spectrum of AI endorsement

Most participants found use cases they're interested in, while maintaining some uncertainty about how best to use AI and when it is appropriate

P14: Actively campaigns against AI, believing it to be harmful in many ways and irresponsibly advertised and adopted.

P7: Uses AI all the time, pushes people to communicate in repos so that they can leverage agents better

P11: Creates personalized apps to automate as many workflows as possible. Highly invested in sharing AI strategies.

Summary Results

Opportunities for AI to Assist with Research, Data, & Compute Tasks

Identified based on current AI use

Data & compute workflows

FOR PIs/PROJECT ADMINS

FOR PROJECT MEMBERS

FOR COMMUNITY ZONE DATA STEWARDS

- Support data discovery, cleaning, analysis, visualization leveraging project context (e.g., allocation requests, project summaries, meeting notes)
- Format data and maintain metadata within Community Zone collections and individual projects
- Help users request resources and launch HPC environments
- Enable assistants such as GitHub Copilot to access facility data and tools from environments like VS Code
- Summarize datasets and other public resources (contents and structure)

Research assistance

FOR PIs/PROJECT ADMINS

FOR PROJECT MEMBERS

- Summarize recent project work, leveraging Zoom assets and facility logs, acting as a memory aid and monitoring tool
- Connect lit review agents to project context and project data to find more relevant references
- Leverage logs and project context to support documentation and reporting (e.g., reproducibility artifacts, progress reports)
- Create study plans

Opportunities for AI to Assist with Research, Data, & Compute Tasks

Identified based on current AI use

Enabling features

- Store or provide access to project context (e.g., Zoom assets, personal notes, grant documents, code repositories)
- Support text and voice interaction with agents
- Offer a skill and prompt library with reusable agent instructions and workflows to reduce redundant efforts, promote standardization, and provide access to common tools
 - Include benchmarks to set performance expectations and offer ways to contact other skill users for mutual support
- Provide easy access to AI logs for agent monitoring (live and post hoc), but allow monitoring windows to be minimized or navigated away from to reduce unnecessary distraction
- Add users to an onboarding project when they sign up, showing them how AI tooling is integrated with the facility



Users usually feed AI tools their project documents; facility could reduce the need for that manual effort

Expectations: Requests for Tailored Experiences and HPC Access

When discussing a fictional data facility:

- Should offer personalized experiences
 - Reminders of current tasks
 - Agents approach tasks the same way project members would
 - Write in their voice
 - Ability to share skills with project members
 - Access to domain specific tooling
- Should not compromise their IP—need ability to keep skills and data private from other projects
- Should simplify HPC access
 - Need increased resources for use, improved ability to run experiments in parallel
 - Want to reduce the need to monitor agent or be involved in handshakes
- Support data discovery
- General preference for multiple agents or well managed memory, believing this to lead to better outcomes



Access to Github logs, Zoom artifacts, and other project context would improve personalization and could reduce iteration

An Open Question

How should a data facility support some of those use cases?

Is the facility the centralizer, bringing in project context and preferred agents to one place?

Is the facility also a data provider, allowing users to give their external agents access to logs and data?

Does it do both? Or will the facility not connect project context at all?

Answering requires reflection on the data facility's value propositions

Judging Utility

Users primarily judged AI utility based on how much more quickly they get through a task.

- Participants appreciate speed and ability to do tasks that were previously impossible or extremely difficult (because of scale or lack of knowledge)
- Some questioning of when AI is most appropriate—participants had no clear cut guidelines to reference
- Benchmarks needed to help users judge which approaches are most effective



Facility can learn from aggregate use patterns, leveraging logs to provide performance benchmarks and recommend skills and agents for certain tasks

Concerns & Constraints on AI Use

Users must be able to opt out of AI tools, interfacing with the data facility in more traditional ways

Constraint - Cost & Limited Tokens

- Token access is limited and costly; facility can offer value by allocating tokens
- Users will need to be able to switch their agents when token limits are reached

Constraint - Security & Regulation

- Institutional security policies limit AI use (but users will push boundaries!)
- Need clarity on which policies impact users, especially for cross-facility workflows
- Need for thoughtful regulation developed with users

Constraint - Awareness

- Lack of knowledge of best practices and ideal use cases; need for information sharing among users

Concern - Critical Thinking

- Need for useful friction to slow down work, encouraging critical thought and knowledge (not just data) production

Concern - Privacy & Hesitance to Share

- Users hesitate to show unrefined AI outputs with peers, partly because it's seen as unhelpful—what outputs and logs should project members share with one another?
- Users want to be sure their valuable data or AI skills are not being illicitly shared with others

Concern - Expertise

- Students/junior colleagues do not always have the necessary knowledge to accurately validate AI output
- References, benchmarks, and contextually relevant information is important to help users validate AI outputs
- Monitoring tools can help mentors see output quality

Concern - Environment

- Concern around environmental impact (but not usually enough to truly limit AI use)
- Careful messaging needed around environmental impact
- Design choices should minimize water and energy usage

Follow up Research Opportunities

- 1 People like summaries and accelerated work, but need to be intellectually engaged. How does a data facility introduce the right amount of friction in the right places?
- 2 What policies do users want to see around AI in a data facility?
- 3 Do users want the facility to centralize their existing resources (data, code, project documents and notes) or should it be a source of information that can be leveraged by external tools?
- 4 What AI logs/outputs get shared within teams and what should be personal?

Results Deep Dive

Motivators & Facilitators for AI Use

- Desire to speed up and scale up, especially when experiencing time crunch
- Enhanced ability to take on new tasks
- Ease of iteration
- Perception that AI offers well curated search results
- Institutional endorsement and support
- Desire to use the tools/approaches peers use
- Cultural acceptance
- Alignment with skills and experience
- Reputational gains given AI hype



Note that the perception of quality curation was driven in part by the small number of results AI offers users—facility could create a similar effect for standard data search by making the first 5 results more prominent and de-emphasizing pagination

AI Use Cases, Real and *Imagined, Reported by Interviewees

- Writing code:
 - Troubleshooting
 - Optimization
 - Parallelization
 - Script writing
 - Planning (tool choices, repo structure)
 - Getting acquainted with new repos
- Maintaining documentation
- Accessing HPC resources
- AI management:
 - Prompt writing
 - Skill writing and maintenance
 - Write logs
 - Orchestrating other agents
 - Validate AI outputs
 - Monitor and report on errors in AI workflows
- Data work:
 - Discovery (query formulation in unfamiliar DBs and discovering unknown DBs)
 - Summarizing
 - Collection
 - Formatting
 - Tidying
 - Integration
 - Analysis
 - Visualization
- Creating apps (for end users or personal workflow automation)
- Literature review
- Understanding a new discipline
- Search (instead of Google)
 - Appreciation for summaries, the perception of curated results, and source citations
 - Identifying potential collaborators
 - For institutional knowledge, policies*
- Brainstorming project ideas
- Writing/editing (papers, reports, proposals)
- Paperwork, e.g. grant proposals, updating organizational docs*
- Writing email
- Memory aid
 - Meeting summaries, transcripts, action items

Strong appreciation for summaries to catch up on previous work/meetings and understanding a new topic or resource

Users interact with agents via desktop apps, browser apps, VS Code extension. They use talk and text.

Examples of AI Use Cases

DATA & COMPUTE
WORKFLOWS

PROJECT MEMBERS

Participant 1: "So if there was just a free-form text field and I said, 'Hey, this is what I'm looking for...I need the date-ordered buy-offs for all of these units,' and then...it would create the query for me that gives me the data I need, or at least it starts the query for me and I can tweak it....that would help me because I don't understand the structure of the database and how to form the query."

DATA & COMPUTE
WORKFLOWS

CZ DATA STEWARDS

Participant 18: "We were collecting photographs for 1,000+ sites. And so I ended up using AI to help set up workflows to seamlessly download the images and label them based on the site numbers and organize them."

DATA & COMPUTE
WORKFLOWS

PROJECT MEMBERS

Participant 9: "I have a loop, an agentical loop to design a physics analysis for ...Large Hadron Collider....And I want the agent to create an analysis strategy for a specific signature I want to look for."

RESEARCH ASSISTANCE

PROJECT MEMBERS

PIs/ADMINS

Hannah: "Once the meeting is over, then what happens to those notes?"

Participant 16: "We usually take them and put them in our agent memories so they're aware of these things."

Expectations: Requests for Tailored Experiences and HPC Access

When asked about their opinions on a data facility “where you can store and discover data, implement workflows, and leverage the [computing facility] of your choice...and you can integrate agents,” participants expressed the following wishes:

- Should offer personalized experiences
 - Reminders of current tasks
 - Agents approach tasks the same way project members would
 - Write in their voice
 - Ability to share skills with project members
 - Access to domain specific tooling
- Should not compromise their IP—need ability to keep skills and data private from other projects
- Should simplify HPC access
 - Need increased resources for use, improved ability to run experiments in parallel
 - Want to reduce the need to monitor agent or be involved in handshakes
- General preference for multiple agents and/or multi-level memory, believing this to lead to better outcomes

Pilot participant:

“I think having some sort of separate agents might be a good idea. We understand that a key feature of agents is some sort of memory management. As I would not want to conflate or mix, my interaction with the administration side of the work with the research side of the work. I think that that's sort of important.”

Quotes on Concerns & Constraints

Constraint - Cost & Limited Tokens

- Token access is limited and costly; data facility can offer value by allocating tokens
- Users will need to be able to switch their agents when token limits are reached

Participant 9, limited by cost: "So I was using mostly Codex. The reason I use Codex because it's unlimited due to the lab subscription. Otherwise I would use Claude with some other models."

Constraint - Security & Regulation

- Institutional security policies limit AI use **(but users will push boundaries!)**
- Need clarity on which policies impact users, especially for cross-facility workflows
- Need for thoughtful regulation developed with users

Participant 7, pushing boundaries: "I'm connecting it with HPC, but that connection with HPC is not smooth at all.... I have to get the agent to pick up a one-time passcode for me within 30 seconds."

Participant 14, on regulation: "You get...chronic statements from both public and private sector that a lack of regulation is necessary for innovation....And if we look historically, that just isn't true at all. Like the thing that makes products safe and trustworthy, that creates the innovation to reduce the harmful externalities is regulation."

Participant 17, pushing boundaries: "...There's a lot of work that isn't on open source repos that policy-wise can't be putting into Claude and whatnot. I'll still like drop technical details...or domain details. And it's like, 'I need a function that, you know, does this standard data analysis thing, give that to me.'"

Note: Some participants expressed security concerns

Participant 14 said, "...Agents should be understood as malware from a security standpoint."

Quotes on Concerns & Constraints (continued)

Constraint - Awareness

- Lack of knowledge of best practices and ideal use cases; need for information sharing among users

Participant 4, limited awareness: "I only learned what agentic AI was last week."

Concern - Expertise

- Students/junior colleagues do not always have the necessary knowledge to accurately validate AI output
- References, benchmarks, and contextually relevant information is important to help users validate AI outputs
- Monitoring tools can help mentors see output quality

Participant 6, students lacking expertise: "...A lot of my students will use it for coding tasks. And that's where I've seen the most issues or the most red flags to me is because they will get to the point where they're like... 'I worked with Claude or ChatGPT'....and I need to present these results....And then I look at them, and I mean, the code that ChatGPT wrote is correct, but it's not valid for whatever problem they were trying to solve."

Concern - Critical Thinking

- Need for useful friction to slow down work, encouraging critical thought and knowledge (not just data) production

Pilot participant, on slowing down work: "...You can be extremely productive. But, when you are presented with that result like half an hour later...it's hard to understand all these results. But in the old way, you spend half a year to get something done. So over that half a year, you actually have like a lot of time thinking about the project."

Quotes on Concerns & Constraints (continued)

Concern - Privacy & Hesitance to Share

- Users hesitate to show unrefined AI outputs and prompts, perceiving it as unnecessary information and too verbose. Summaries and talking points are more thought to be appropriate for sharing; prefer to share refined work
- Tone with AI might be more casual than the professional self they want to project to peers
- Users want to be sure their valuable data or AI skills are not being shared without credit

Counterpoints on Hesitance to Share

- Mentors share prompts and logs with mentees to demonstrate AI usage
- Meeting transcripts are widely shared
- Team members share skills with one another

Concern - Environmental Impact

- Concern around environmental impact (but not usually enough to truly limit AI use)
- Careful messaging needed around environmental impact
- Design choices should minimize water and energy usage

Participant 17, on sharing unrefined outputs: "I don't want to hand my project managers, my teammates, something fuzzy. I see it as my responsibility to give them like high quality information that has been checked over."

P16, on sharing AI skills:
"You're now trying to take a secret sauce of sorts, right? I mean, it's the intellectual aspect of it as well....how do you assign credit in the first place?"

Participant 11, first reaction to the idea of a data facility: "I heard like it has— it will bring so negative impact on the environment, and I think that is something we need to concern. But maybe I'm kind of too positive because I don't really know how severe it is."

Effective AI Assistance Depends on Project Context

Users want to cut down on iteration and ensure AI produces content that matches their style.

To reduce interventions, users provide provide the following context:

- Paper and grant reviews
- Github repos (issues, PRs, code, docs)
- Mockups and descriptions of user flows
- Screen recording of workflows (for automation)
- Performance feedback
- Data
- Meeting notes and transcripts



HPDF can help users connect and reuse project context like meeting records, project documents, & development histories



Users must be able to control what is shared with agents & project members

Using Multiple Agents and AI Tools

Daily AI users who are moderately/extremely familiar with AI report using multiple agents or AI tools
(not necessarily in a single workflow)

Choice:

- Users have a feel for which AI tools perform best for a given task (e.g. Copilot for summarizing repos, prompt writing from ChatGPT or Claude)
- To comply with policies, sometimes certain tools must be used for certain tasks (e.g. involving CUI)

Use cases for AI teams:

- Use one agent to orchestrate others, writing prompts, fetching results
- Use AI tools as red-team reviewers of another agent's output

Participant 16: "We have...agents that talk to each other and synthesize a research problem end-to-end and see how they should go about solving that...And as we do that sort of workflow, we normally end up thinking through what are the steps that need to happen, And many of them are actually agentified."

Participant 7: "That has become kind of this bread and butter of the workflows that I do in terms of agentic workflows: going from creating an issue in a GitHub repository tied to a specific project, assigning it to Copilot or Claude coding agent, it making a pull request, grabbing context from other issues and comments that I've made across the repository, having it ping Edison Scientific and get results back."

Tensions Between the Need for Monitoring & Desire for More AI Autonomy

Need for monitoring

- Logs/human-in-the-loop (HITL) needed to ensure quality control and monitor progress and resource consumption
 - Junior colleagues and students may need more monitoring
- Logs help increase trust; sources and reasoning are visible

VS.

Desire to reduce AI monitoring

- A primary value of AI is that it takes over menial tasks, but monitoring *is* tedious
- Log DBs can get large, making monitoring tedious and slow
- HITL can be enough friction to drive a user to put off a task

Log capture is critical but users do not want to babysit AI

- Logs understood as an important data source for improving AI performance
- Logs sometimes shared with mentees to show how to use agents

Participant 9:

"[I] say [to the AI], 'Every time you try something, you should document it, and the tasks are in this task folder.' So I have a folder specific for the task. So what I do, I just create new files in that folder. And this [agent] will automatically look into it and work on the tasks. And after they finish, they will report back to the Markdown file. So I don't need to—I don't manage them very often."

Interview Protocol

Interview Protocol

Warm Up - 3 Min

1. First, tell me a bit about your role and responsibilities.

AI Use - 15 min

We're going to talk about AI tools for most of this conversation. I know there's a lot of different forms of AI, but whatever comes to mind when we're talking is fine to share. You don't need to limit yourself to just LLMs and agentic tools, for instance.

2. *Scientists:* Outside of coding tasks, can you tell me about a time you used AI in your research process?
Data stewards: Outside of coding tasks, can you give me an example of how you've used AI when working with data?
 - a. Walk me through this in detail—how did you interact with the AI and what happened?
 - b. Have you done this multiple times or was it a one off?
 - c. What spurred you to try this in the first place?
 - i. How did you come to choose *[the tool]*?
 - d. How did you assess the results AI gave you?
 - e. What did you like about using AI in that situation?
 - f. Was there anything that felt risky or uncertain about using it? Tell me about that.

If no AI use:

3. Can you tell me why you're not using AI?
 - a. What feels risky about using AI to you? (*probe for what kinds of AI they're thinking of, ask for an example if needed*)
 - b. Aside from those risks, what else has stopped you from using AI?
4. What *does* seem appealing to you about using AI?

Interview Protocol (continued)

Opportunities for AI - 25 min

1. Thinking about the tasks you routinely engage in when working with data, what's something that you imagine could be sped up or offloaded by integrating a new *AI agent*?
 - a. Describe how you imagine working with the agent in that situation.
 - b. How would you know it's working well?
2. Thinking beyond just your current routines, what are some intriguing opportunities you think AI in any form could bring to your kind of work in the future? *[new questions, methods, products]*
 - a. In what ways would this be useful to you?
 - b. What is needed to make that vision come to life?
3. What do you see as the biggest obstacle right now to implementing AI workflows in your kind of work? Tell me about that.
4. What data do we need to be collecting to make effective use of AI in your field? *[If there were a model trained for people like you, what would it be trained on?]*
5. What other resources do you think are important for an AI-enabled future?

Interview Protocol (continued)

AI and Data Facility Interfaces - 10 min

1. Before we wrap up, let's talk about AI and data facilities. Imagine a data facility where you can store and discover data, implement workflows, and leverage the LCF of your choice. This facility has several key interfaces—like for data search and workflow construction—and you can integrate agents into all of them if you wish.
 - a. First, what are your initial reactions to this idea?
 - b. Would you prefer agents from different interfaces to share memory with one another or would you want them to be more independent?
 - i. Tell me about why you say that.
 - ii. What exceptions might there be to your preference here? (*When might it be okay to share/preferable to not share memory*)
 - c. What sorts of personalization would you like to see in a facility like this? (*What should agents know about you, and how should they leverage that?*)

Wrap Up - 3 min

2. Is there anything you haven't gotten to say that you think we should know?
3. Do you have any questions for us?

Thank you again for your time.

Funding Acknowledgement for AI User Needs Study

This material is based upon work supported by the U.S, Department of Energy, Office of Science, Office of Advanced Scientific Computing Research under contracts 89243126CSC000213 (TJNAF) and DE-AC02-05CH11231 (LBNL)