

Q1: Is there an IRI API defined already?

**A:**

The IRI API is a work in progress and has been under development for almost two years. A formal group is working on its definition. We are working within AmSC, there is an IRI integration group, focusing on interacting with existing IRI products. The Infrastructure Services level is providing input into the design discussion so the IRI API can meet the needs of AmSC. You can currently find draft documentation for the [facility status API](#) (information about resources and their availability at different facilities). They are being deployed at many facilities. There is also ongoing work in data, authentication, running, and monitoring jobs.

---

Q2: When will the websites referenced in the American Science Cloud detail architecture be stood up and be publicly visible?

**A:**

The websites, including portals, dashboards, and API websites, are under active development. The goal is to deploy the first versions in early January, with potential incorporation into demos shortly thereafter. However, these websites may not be publicly *usable* initially, as they will be proof-of-concept for early users.

---

Q3: Will users request access to the entire Science Cloud or individual subsets of services?

**A:**

Users without an account at a DOE computing facility can apply for an AmSC account or use existing accounts from federated providers to access the American Science Cloud. Access will be granted to resources within the facilities that are allocated to their accounts.

---

Q4: How is "AI-ready" data defined? What is the plan to access domain specific knowledge?

**A:**

AI-ready data refers to cleaned, validated, and machine-readable data with sufficient metadata for model training. It does not imply a single format, but rather modular standards tailored to domain-specific applications. Seed/Flashlight teams provide domain semantics for which the ModCon DBS group will build adapters (validation tools). AmSC will provide the catalog and infrastructure so the curated datasets can be accessed consistently.

---

Q5: What is the difference between a seed team and a flashlight team?

**A:**

Flashlight and SEEDs refer to the same projects. SEED projects are renamed as Flashlight projects.

---

Q6: How does data governance fit into the strategy, especially for sensitive data?

**A:**

Different enclaves will be supported for varying sensitivities of data, integrated into catalog functionalities, user accounts, resource access, and data movement. Appropriate protections will be implemented for sensitive data movement and usage.

---

Q7: Will the American Science Cloud provide access to commercial Electronic Design Automation (EDA) tools commonly deployed with CSPs?

**A:**

The American Science Cloud lays the foundation for facilities and infrastructure partners to connect to the ecosystem. It does not target specific tools or datasets but supports integration based on partner preferences.

---

Q8: Will the Science Cloud support integration of open and proprietary datasets for ModCon use cases?

**A:**

Yes, the Science Cloud will support secure compute and various data security enclaves, allowing data providers to control access and movement securely.

---

Q9: How are priorities decided when targets shift quickly? Who sets the priorities?

**A:**

AmSC has a very well-defined scope and deliverables for the next year. Priorities are guided by headquarters and co-dependencies with broader AI strategies. The team embraces agility and focuses on building lasting foundational capabilities.

---

Q10: When will the first round of Science Cloud demos occur, and what are the use cases?

**A:**

The first demonstrations are planned for end of March, with use cases to be determined based on science teams' needs, and technical capabilities being developed by AmSC. We expect to have a better understanding of the use cases in the months leading up to the demo.

---

Q11: Building agentic workflows and curating datasets in an agentic way requires access to foundation reasoning LLM models. Will AmSC provide access to state-of-the-art models to gain expertise with the tools?

**A:**

Yes, state-of-the-art models will be made available through AmSC, through multiple interfaces as noted in the presentation, supporting both small-scale and large-scale inference campaigns integrated with agentic flows. The 9-month demo targets agentic flow so we are working towards that goal.

---

Q12: Will the Science Cloud architecture handle various data types, including open, sensitive, and proprietary data?

**A:**

Yes, secure workflows will be supported through existing capabilities, spanning open to sensitive data classifications, within the 12-month period. Additional development work is being planned to continue support beyond the initial 12-month period.

---

Q13: How will the Science Cloud address data security for HIPAA datasets?

**A:**

Resources within the Science Cloud will be rated for different classifications. Access will be granted based on what the account has access to in terms of resource and what the account is authorized to access.

---

Q14: What is the relationship between model teams and seed teams?

**A:**

Model teams and seed teams (or seedling teams) are essentially the same, focusing on foundational development and AI-enabled workflows.

---

Q15: Will users be able to (or have to) send jobs to specific compute platforms or will the jobs be split up across multiple platforms?

**A:**

The architecture supports launching jobs on the appropriate facility. It uses a 'facility coordinator' to manage where jobs are launched. It hides the details of different facilities and automatically picks the best one if the job doesn't specify a facility. Users can also target a specific facility. The architecture does not currently support splitting jobs across multiple facilities simultaneously.

---

Q16: Will there be a Science Cloud BoF or get-together at SC'25?

**A:**

Yes, there will be an internal AmSC developers meeting at SC, but no public BoF due to timing constraints.

---

Q17: Is the Science Cloud funding stable, and what is the long-term outlook?

**A:**

Funding is secured for one year, with efforts focused on building lasting foundational capabilities to advocate for future funding. We are optimistic in our outlook and aggressive in our approach to what we deliver.

---

Q18: Will the Science Cloud support ML for science use cases not based on AI agents or foundational models?

**A:**

Yes. The approach with AmSC is to enable and accelerate AI based science across all model teams and in general across DOE. The infrastructure supports hybrid use cases combining simulations and existing data to generate training data for models in an overall simulation+AI approach to solving the problems. The infrastructure will support flexible execution of all types of codes across facilities.

---

Q19: Is ModCon supporting model teams working on science not covered by seed teams?

**A:**

Yes, ModCon supports a broad range of scientific model development.

---

Q20: How can the playing field be leveled for labs with varying access to foundational LLMs?

**A:**

Standardization of practices and procedures across labs is needed, and discussions are ongoing to address these disparities. There is a team across the labs and within HQ that is working on these long-term challenges. We recognize the need to have thoughtful conversations and come to decisions that harmonize the ecosystem.