Genesis Mission

AMERICAN SCIENCE CLOUD

April 27th, 2026, TREASURE workshop

The American Science Cloud (AmSC)

PRESENTED BY

Debbie Bard

AmSC Product Manager

U.S. DEPARTMENT of ENERGY

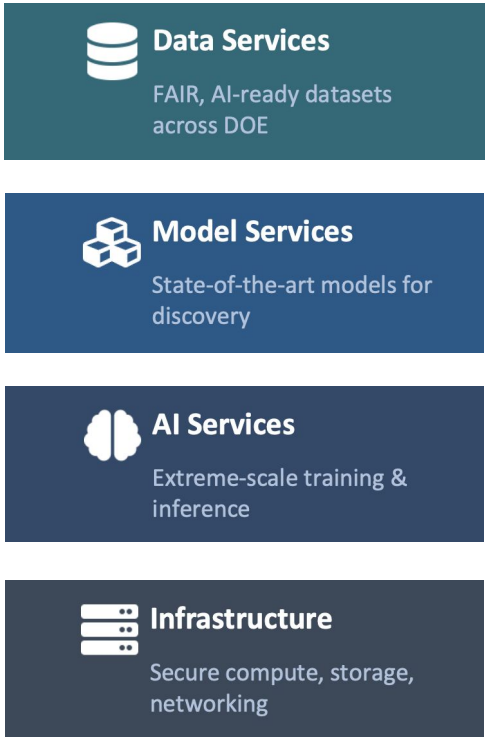
AmSC is a First-of-a-Kind Integrated Platform for Transformative Science

- Deliver a **common fabric** for scientists to build on
- Provide **modular services and abstractions** used to accelerate discovery cycles
- Leverage **science and industry innovations** rapidly as they are integrated into the platform
- Build a science-focused platform through **co-design**

AI-driven scientific discovery in the Genesis Mission

AmSC enables DOE scientists and collaborating teams to:

- Create, access, and integrate world-class AI-ready datasets
- Run scalable model training on contributed compute infrastructure
- Perform large-scale modeling-simulation and AI
- Control instruments and experiments
- Move data efficiently across sites



DOE Labs and Infrastructure Partners

ANL · BNL · FNAL · INL · JLab · LBNL · NETL · NLR · ORNL · PNNL · PPPL · SLAC

AmSC is a cornerstone of the Genesis Mission platform

- **National Science and Technology Challenges:** High-impact scientific efforts to address Genesis Mission priorities, selected in response to the current RFA. These projects may be users of the Genesis Mission platform and therefore AmSC.

- **The Model Consortium (ModCon):** developing tools and frameworks that will be hosted by AmSC infrastructure

- **AmSC is a coalition of Infrastructure Partners (IPs):** DOE labs that contribute and integrate core capabilities into the AmSC platform infrastructure

National S&T Challenges:
Leverage ModCon and AmSC services to accelerate scientific discovery

Model Consortium:
Develop AI models & workflows that are supported by AmSC

AmSC Infrastructure Partners:
Integrate core capabilities into AmSC

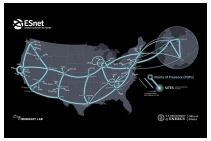
Collaborating Infrastructure Partners in AmSC

Infrastructure Partner(s)	Lab(s)	Key Contacts (POC/PI, Leads)
OLCF	ORNL	Arjun Shankar
ALCF	ANL	Mike Papka
ESnet	LBNL	Inder Monga
NERSC	LBNL	Sudip Dosanjh
HPDF	TJNAF, LBNL	Graham Heyes, Anna Kupresanin
C3	PNNL	Robert Rallo
Scientific User Facilities	LBNL, ANL, TJNAF, BNL, FNAL, SLAC, ORNL	Paolo Calafiura, Nicholas Schwarz

Infrastructure Partner(s)	Lab(s)	Key Contacts (POC/PI, Leads)
DeepLynx	INL	Peter Suyderhoud
DIII-D Digital Twin	GA	Sterling Smith
EDX	NETL	Kelly Rose, Chad Rowan, Jack Searle
FemtoMind	TJNAF, FNAL	Robert Edwards
Fermi Data Platform	FNAL	James Amundson
HAIDIS	TJNAF	Ilya Baldin
SCDF	BNL	Adolfy Hoisie
S3DF	SLAC	Jay Srinivasan
Stellar-AI	PPPL	Shantenu Jha
VEE-ARIES/HERO	NLR	Kristi Potter

AmSC is working with DOE Laboratory Facilities and Infrastructure Partners

Infrastructure and Services



HPDF
High Performance Data Facility
A Partnership of Jefferson Lab & Berkeley Lab

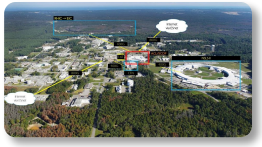
EDX
Energy Data eXchange

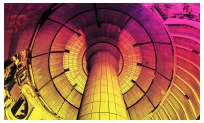
ARIES
by NATIONAL LABORATORY OF THE ROCKIES

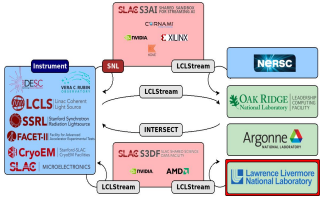
DeepLynx nexus

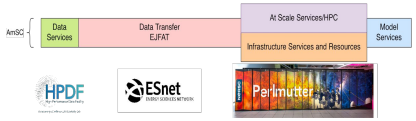
aws Azure Google Cloud

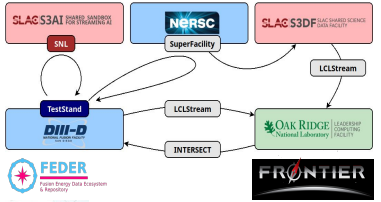
Scientific Facilities and Workflows

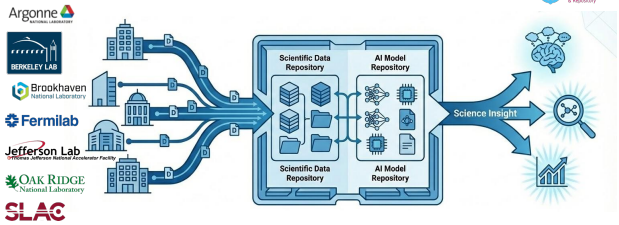


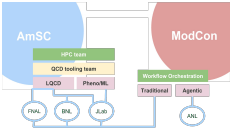












AmSC Organization Chart

Science Council

Chair:Andreas Kronfeld (FNAL)
Kevin Yager (BNL)
Michael Begel (BNL)
Kjiersten Fagnan (LBNL)
Oliver Gutsche (FNAL)
Joel England (SLAC)
Todd Satagota (JLAB)
Jana B. Thayer (SLAC)
Neeraj Kumar (PNNL)

Industry Council

AWS
AMD
Cisco
Dell
Google
HPE
Microsoft
NVIDIA
Nokia
Cornelis

IP Council

Infrastructure Partner PoCs

AmSC Project Leadership Office

DIRECTOR: **Gina Tourassi** (ORNL)

DEPUTY: **Arjun Shankar** (ORNL)
DEPUTY: **Inder Monga** (LBNL)

CTO: **Sarp Oral** (ORNL)
Product Manager: **Deborah Bard** (LBNL)
PMO: **Denise Hoomes** (ORNL)
CISO: **Ryan Adamson** (ORNL)
Integration Architect: **Zach Mayes** (ORNL)

ModCon CoDesign Liaisons

Wahid Bhimji (LBNL)
Venkat Vishwanath (ANL)
Feiyi Wang (ORNL)

IRI Liaisons

Deborah Bard (LBNL)
Thomas Uram (ANL)

OPERATIONS

L1: **Ashley Barker** (ORNL)

Virtual Organization
Veronica Vergara (ORNL)
Adam Slagell (LBNL)

Intelligent Operations
Ed Balas (LBNL)
Eric Pershey (ANL)

User Engagement
Haritha Siddabathuni Som (ANL)
Chris Fuson (ORNL)

INTERFACES AND SERVICES

L1: **Mike Brim** (ORNL)

AmSC Interfaces
Taylor Childers (ANL)
John MacAuley (LBNL)

Infrastructure Services
Zach Mayes (ORNL)
Shane Canon (LBNL)
P. Shyamshankar (ANL)

Data Services
Ilya Baldin (TJNAF)
Rajesh Kalyanam (ORNL)

AI SERVICES

L1: **Thomas Uram** (ANL)

Model Services
John Gounley (ORNL)
Huihuo Zheng (ANL)

At-Scale Services
Murali Emani (ANL)
Steven Farrell (LBNL)

Intelligent Interfaces
Wahid Bhimji (LBNL)
Shreyas Cholia (LBNL)

AmSC PARTNER INTEGRATION

L1: **Chin Guok** (LBNL)

IRI Integration
John MacAuley (LBNL)
Paul Rich (ANL)

CSP Integration
Roger Cass (PNNL)
Jack Deslippe (LBNL)

Industry & Gov IP Integration
David Martin (LBNL)
Rebecca Hartman-Baker (LBNL)
Ben Mintz (ORNL)

The American Science Cloud (AmSC)

A Unified Cloud Platform for Transformational AI and Science Across DOE
A cornerstone of the Genesis Mission

Mission

Deliver a seamless integration of DOE science instruments, computing, experimental, and networking capabilities, AI, data, modeling and simulation software tools in a single advanced programming interface – the **AmSC API**

What We Deliver

✓ Unified APIs for data, model, and workflow access
✓ Secure, federated identity (Virtual Organization model)
✓ Seamless interoperability between DOE facilities and CSPs

Reduce "time to insight" from months to days through composable AI-driven workflows

First year sprint to develop the Genesis Mission Platform MVP

3-MONTHS
FEDERATED ACCESS

✓ Unified login across multiple DOE HPC sites; first operational **REST API and Python client** enabling authentication, dataset discovery, and workflow execution

6-MONTHS
SCIENCE TEAM ONBOARDING

✓ Two to three research teams execute **end-to-end workflows** (data search → transfer → model fine-tuning) on **Frontier**, **Perlmutter**, or **Aurora** using the AmSC Data Catalog

9-MONTHS
INTELLIGENT WORKFLOW MILESTONE

✓ First early **agentic workflow framework** autonomously discovers data, trains, fine-tunes, and performs **reasoning-based inference** across federated DOE and cloud resources

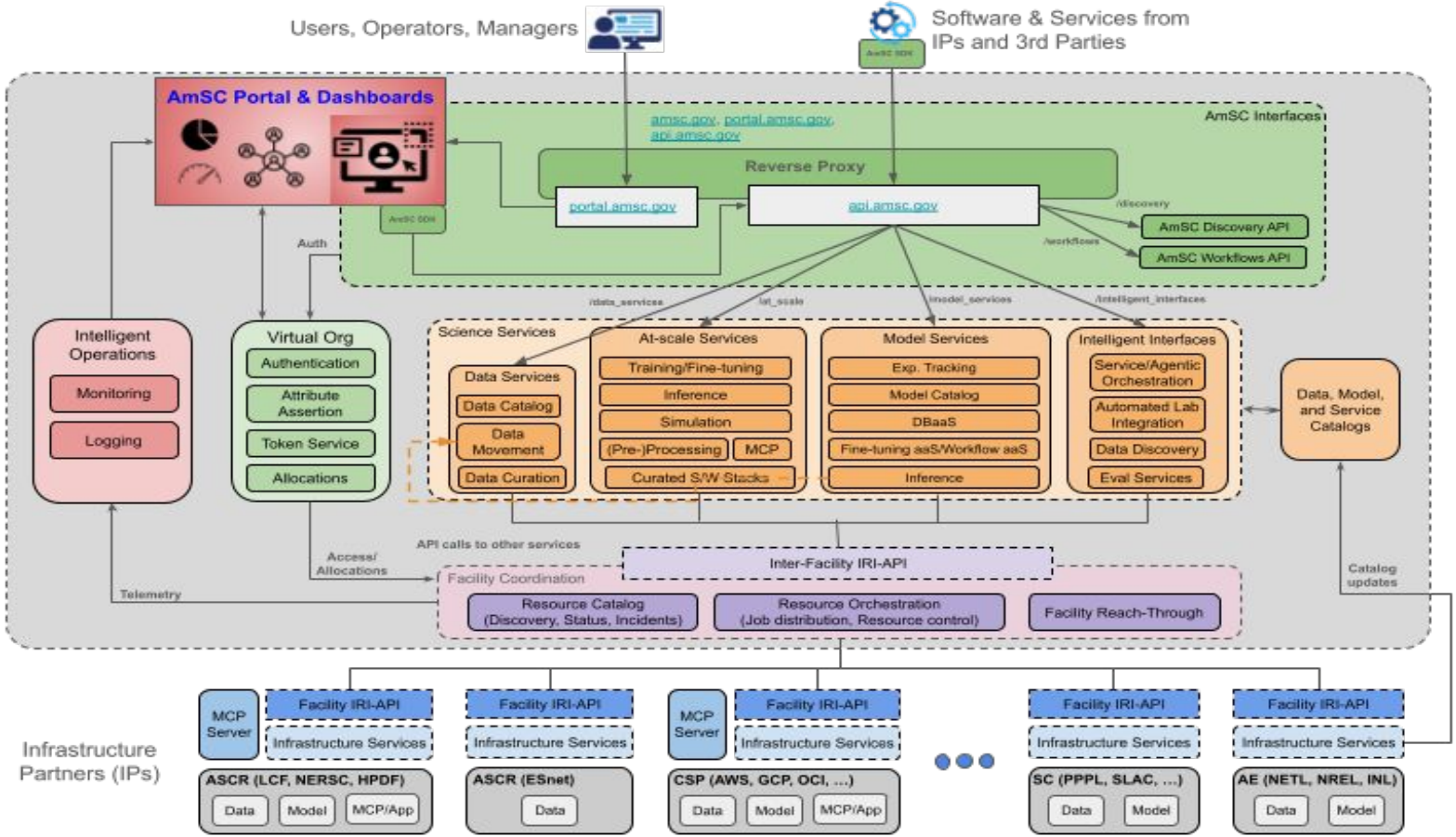
12-MONTHS
MINIMUM VIABLE PRODUCT

✓ Link IPs and CSPs into a **unified pre-production environment** to support AI-ready data services, agentic frameworks, large-scale model training/inference and advanced simulation. Establish a **foundation for future expansion** to additional partners and services.

LONG-TERM VISION

Accelerate discovery cycles by building the world's most integrated scientific platform

Scalable architecture incorporates multiple DOE/non-DOE providers



Example usage: the AmSC platform can be used for data discovery and model training

1. Register & Discover Data
- Register datasets in Data Catalog
- Search for similar relevant datasets

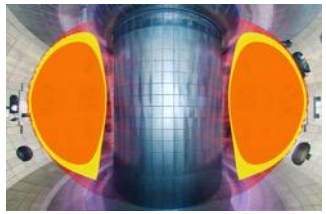
2. Find Base Model
- Search Model Catalog for appropriate model

3. Train Model
- Move data to appropriate site
- Train model with selected data
- Run on AmSC infrastructure

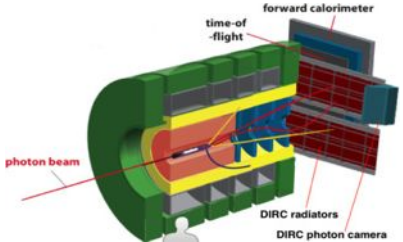
Examples of science workflows that fit this pattern:

- Foundation model development for network planning and decision-support.

- Surrogate model development for experiment facility digital twin.



- Extracting physics information from experiment or simulation data.



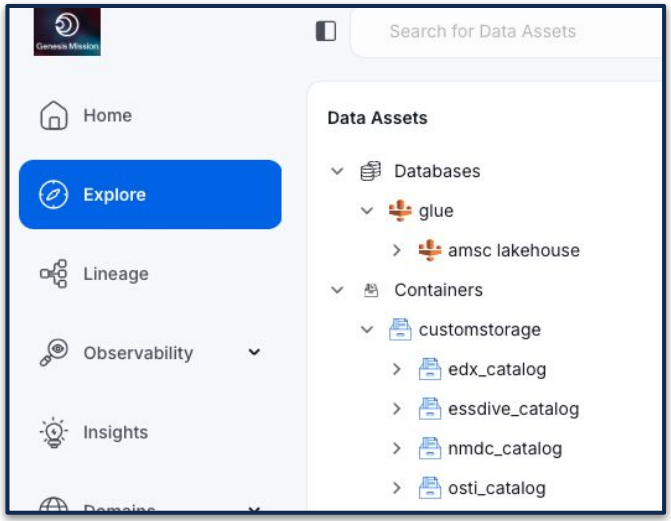
How does AmSC enable this workflow?

1. Register & Discover Data
- Register datasets in Data Catalog
- Search for similar relevant datasets

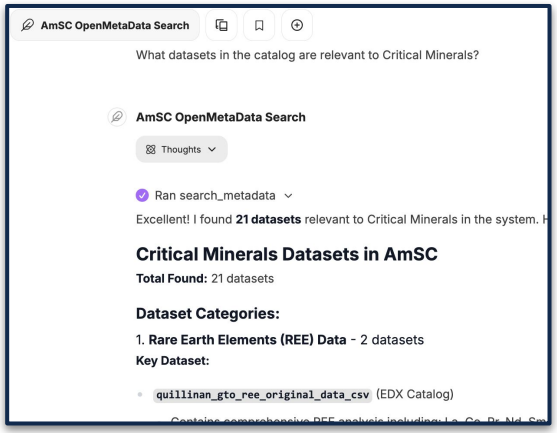
2. Find Base Model
- Search Model Catalog for appropriate model

3. Train Model
- Move data to appropriate site
- Train model with selected data
- Run on AmSC infrastructure

User can upload, inspect and query data tables in AmSC Lakehouse and search across IP data catalogs.



User can ask an agent to find similar datasets



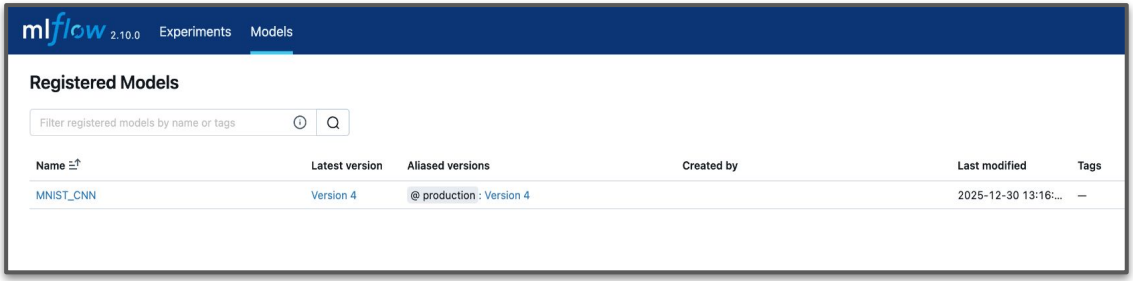
How does AmSC enable this workflow?

1. Register & Discover Data
- Register datasets in Data Catalog
- Search for similar relevant datasets

2. Find Base Model
- Search Model Catalog for appropriate model

3. Train Model
- Move data to appropriate site
- Train model with selected data
- Run on AmSC infrastructure

User registers their model, and searches for other models using industry-standard tools, such as MLFlow or ClearML.



How does AmSC enable this workflow?

1. Register & Discover Data
- Register datasets in Data Catalog
- Search for similar relevant datasets

2. Find Ba
- Search Mo
appropriate model

User deploys model training via interactive web interface, API or agent.

3. Train Model
- Move data to appropriate site
- Train model with selected data
- Run on AmSC infrastructure

Under the hood AmSC orchestrates the infrastructure necessary to run workflow seamlessly across multiple AmSC sites:

- **VO provides access tokens**
- **Resource Orchestration identifies available resources across all AmSC sites (ASCR, IP and CSP)**
- **Data Movement co-locates data with model**
- **Model Services provides experiment tracking framework, including databases**
- **At-Scale Services provides performant software across all sites**
- **IRI API launches jobs**
- **Intelligent Operations tracks job execution for provenance**

Example usage vignette: AmSC platform can be used for inference workflows

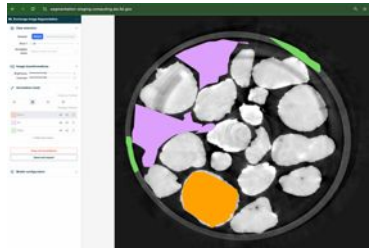
1. Data produced at experiment facility
- Moved via API to computing facility
- Data registered in Data Catalog

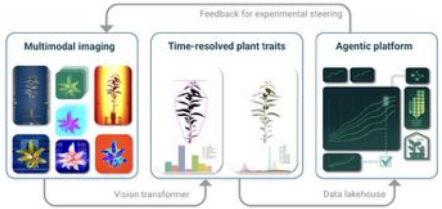
2. Data processed with AI-enhanced analysis
- Access AmSC Inference service
- May launch simulations via IRI-API

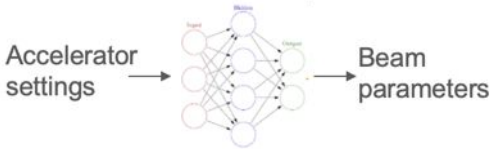
3. Results inform next steps
- AmSC supports Agentic workflows drive autonomous experiments

Examples of workflows across facilities using AmSC to accelerate experimental science:

- Image analysis, used to give lightsource or microscope users feedback on experiment progress in real time.
- Agent-driven feedback for steering autonomous lab experiments.
- Build framework of AI-powered real-time feedback for experiment operators.







How does AmSC enable this workflow?

1. Data produced at experiment facility

- Moved via API to computing facility
- Data registered in Data Catalog

Experiment site supports AmSC API endpoints, used to:

- **Register raw dataset with AmSC Data Catalog and set appropriate access permissions**

2. Data processed with AI-enhanced analysis

- Access AmSC Inference service
- May launch simulations via IRI-API

3. Results inform next steps

- AmSC supports Agentic workflows drive autonomous experiments

How does AmSC enable this workflow?

1. Data produced at experiment facility
- Moved via API to computing facility
- Data registered in Data Catalog

2. Data processed with AI-enhanced analysis
- Access AmSC Inference service
- May launch simulations via IRI-API

3. Results inform next steps
- AmSC supports Agentic workflows drive autonomous experiments

- **Intelligent Operations (with VO) identifies an appropriate available AmSC site accessible to experiment**
- **Data Movement transfers data to AmSC site**
- **Inference Services applies model to data, via user-specified framework**

At-Scale services enables performant large-scale simulations on HPC, based on inference results

How does AmSC enable this workflow?

1. Data produced at experiment facility
- Moved via API to computing facility
- Data registered in Data Catalog

2. Data processed with AI-enhanced analysis
- Access AmSC Inference service
- May launch simulations via IRI-API

3. Results inform next steps
- AmSC supports Agentic workflows drive autonomous experiments

- **Data Catalog registers addition to analyzed dataset, sets appropriate access permissions**
  - **Dataset may be made available to wider community immediately**
- **Data Movement returns results to user for local analysis/viz**
- **OR results are evaluated by API to perform next step in automated workflow**
- **OR agent analyzes results and determines next step in workflow**

Integrating the DOE complex to accelerate scientific campaigns

AmSC is a pathway for scientists to build on the DOE complex's integrated capabilities: access facilities, data, software, and tools – to accelerate components of their workflows.

The following key capabilities of AmSC are aimed at facilitating your end-to-end scientific workflows.

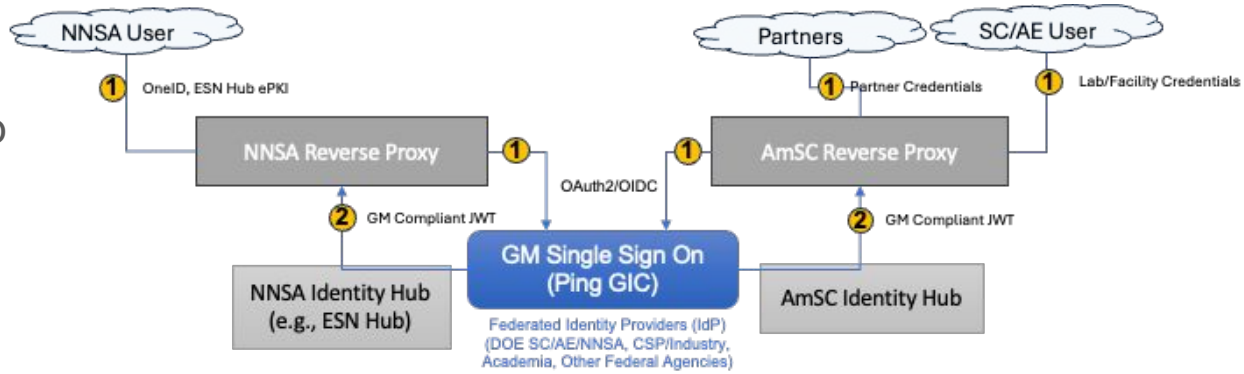
I. Federated ID
II. Data Services/APIs
III. AI Model and Inference Services
IV. Compute Services

I. AmSC Federated Identity

Goal: Provide a unified login experience across multiple DOE and CSP sites through AmSC interfaces

- In this context, federated identity means that a user can authenticate at one trusted DOE site or identity provider and use that same identity to access web-based resources at other DOE facilities—without needing to create separate accounts at each site.
- This is a common source of frustration today

- Will be essential for automated, resilient workflows
- Many of the challenges to implementing FedID are related to **policy and governance** rather than technical obstacles



II. Data Services and APIs

Goal: Users can find, register and move datasets, setting appropriate access controls.

Data Catalog	• Centralized metadata store for **discovering** AmSC datasets • Catalog of catalogs synthesizing dataset registries from IPs and facilities
Data Lakehouse	• Columnar data **storage** supporting highly scalable SQL-like data querying • Hosting solution for AI-ready curated datasets generated from raw data
Data Curation	• **Pipelines** for processing and generating AI-ready data from raw data • Facilitating ingestion into the Data Lakehouse
Data Movement	• **Transferring** registered AmSC datasets from IPs to compute facilities • **Streaming** data transfer from experimental facilities to compute

III. AI Services

Goal: **Users can leverage industry-standard tools to run automated training, fine-tuning and optimization pipelines, including agentic workflows.**

At-Scale services	• Provide curated software stacks across AmSC resources • Benchmark AI applications to validate software stack functionality and track performance (with ModCon)
Model Services	• MLFlow, ClearML and other frameworks for experiment tracking and model catalog • DBaaS for AI
Inference Services	• Serving LLM and non-LLM models hosted at DOE sites and CSPs • Accessible OpenAI-compliant API
Intelligent Interfaces	• Deploying Agentic stack in AmSC - including LLM API access, ChatUI, Agent Builder and MCP Server hosting • Interface to LLM inference services at facilities and commercial providers

IV. Cross-DOE and CSP Compute Access

Goal: Enable automated workflows across all AmSC sites

AmSC services all use a common API to interact with AmSC IP sites, including HPC facilities, DOE lab resources and CSPs. Endpoints include:

- Resource status
- Job management
- Local file management
- Allocation information
- Resource orchestration
- WAN management

operationId	DemoAdapter	NERSC	ALCF	ESnet_West	ESnet_East	ORNL	PNNL	SLAC	BNL
GET /api/v1/facility	PASS	PASS	PASS	PASS	PASS				
GET /api/v1/facility/sites	PASS	PASS	PASS	PASS	PASS				
GET /api/v1/facility/sites/{site_id}	PASS	PASS	PASS	PASS	PASS				
GET /api/v1/status/events	MISSING	MISSING	PASS	MISSING	MISSING				
GET /api/v1/status/events/{event_id}	MISSING	MISSING	PASS	MISSING	MISSING				
GET /api/v1/status/incidents	PASS	PASS	PASS	PASS	PASS				
GET /api/v1/status/incidents/{incident_id}	PASS	PASS	PASS	PASS	PASS				
GET /api/v1/status/resources	PASS	PASS	PASS	PASS	PASS				
GET /api/v1/status/resources/{resource_id}	PASS	PASS	PASS	PASS	PASS				
GET /api/v1/account/capabilities	PASS	PASS	FAIL	PASS	PASS				
GET /api/v1/account/capabilities/{capability_id}	PASS	PASS	FAIL	PASS	PASS				
GET /api/v1/account/projects	FAIL	PASS	PASS	PASS	PASS				
GET /api/v1/account/projects/{project_id}	PASS	PASS	PASS	PASS	PASS				
GET /api/v1/account/projects/{project_id}/project_allocations	PASS	PASS	PASS	PASS	PASS				

Soon! Work in Progress

Foundation for future expansion

AmSC is designed to be a fabric that people can build on. Beyond the MVP, Science teams will be able to bring their models, connect their datasets and plug in their infrastructure.

Foundation for future expansion

AmSC is designed to be a fabric that people can build on. Beyond the MVP, Science teams will be able to bring their models, connect their datasets and plug in their infrastructure.

Example: Connect a new data source to AmSC, make datasets available to all DOE

- Register with the Data Catalog and support Data Movement API to provide access to AmSC users
- Integrate with AmSC FedID for auth and access control
- Data source may send data and compute tasks to AmSC computing sites via the API

Foundation for future expansion

AmSC is designed to be a fabric that people can build on. Beyond the MVP, Science teams will be able to bring their models, connect their datasets and plug in their infrastructure.

Example: Connect a new data source to AmSC, make datasets available to all DOE

- Register with the Data Catalog and support Data Movement API to provide access to AmSC users
- Integrate with AmSC FedID for auth and access control
- Data source may send data and compute tasks to AmSC computing sites via the API

Example: Science team develops a new tool/capability, useful for all AmSC users

- Register with the Service Catalog & document
- Provide AmSC API endpoints for other AmSC services, interfaces and users to make use of it

Foundation for future expansion

AmSC is designed to be a fabric that people can build on. Beyond the MVP, Science teams will be able to bring their models, connect their datasets and plug in their infrastructure.

Example: Connect a new data source to AmSC, make datasets available to all DOE

- Register with the Data Catalog and support Data Movement API to provide access to AmSC users
- Integrate with AmSC FedID for auth and access control
- Data source may send data and compute tasks to AmSC computing sites via the API

Example: Science team develops a new tool/capability, useful for all AmSC users

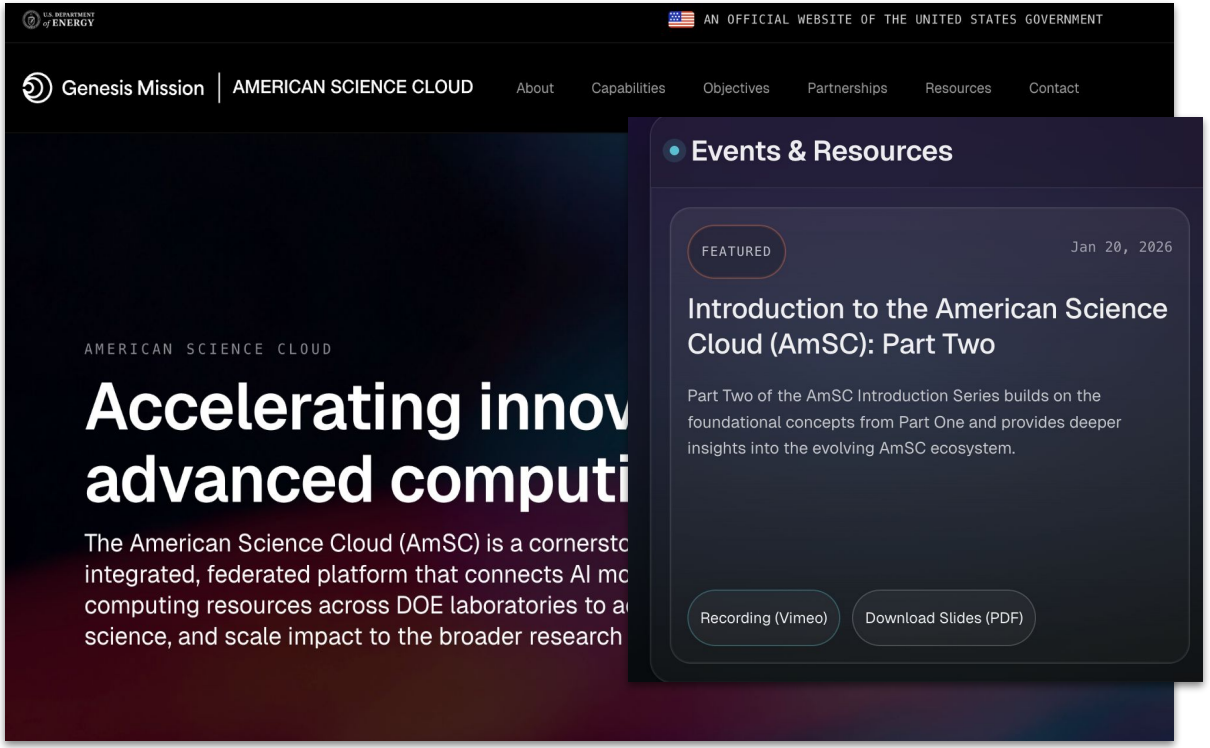
- Register with the Service Catalog & document
- Provide AmSC API endpoints for other AmSC services, interfaces and users to make use of it

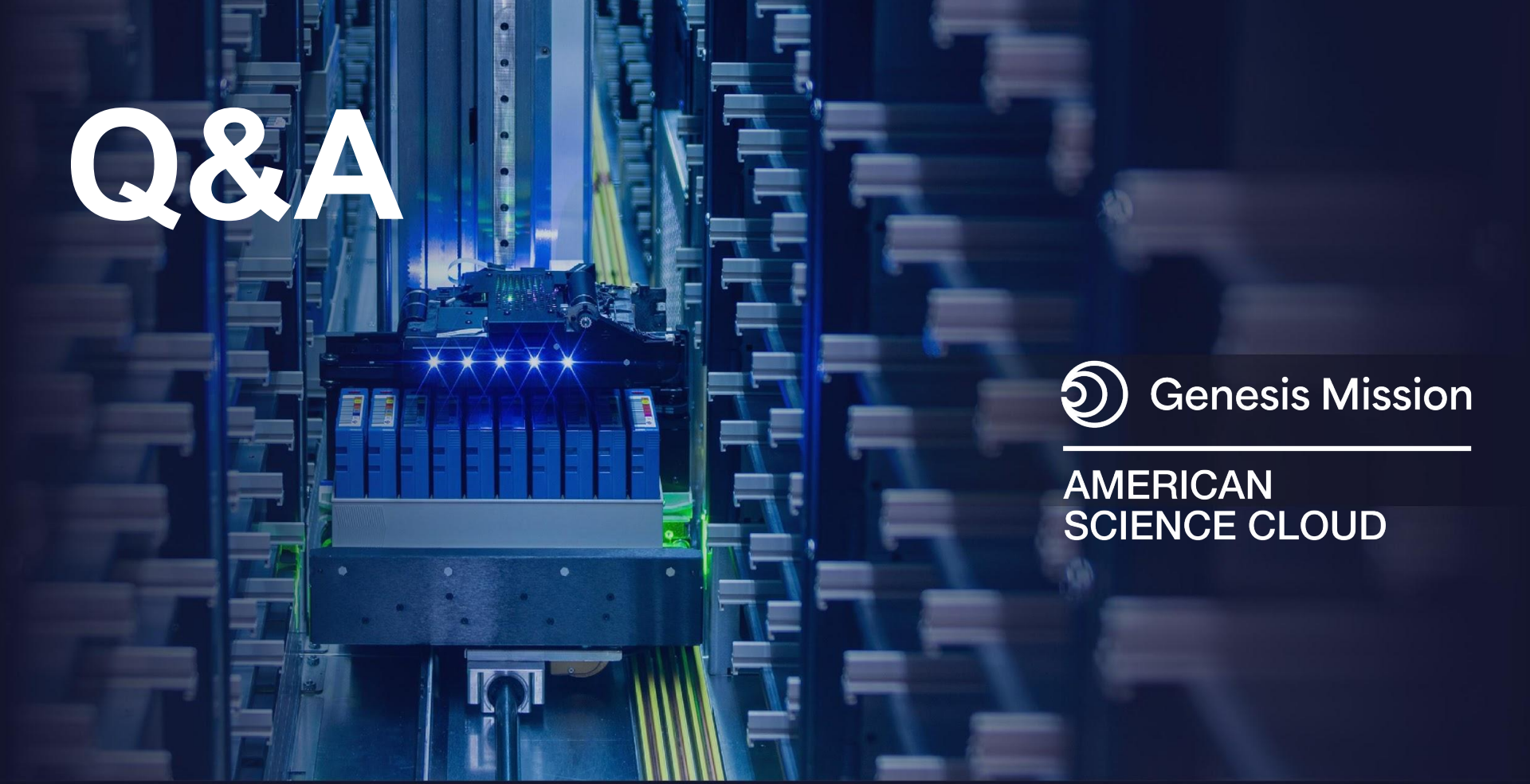
Example: Expose a new compute resource to AmSC users

- Register site with the Resource Catalog; support endpoints to AmSC API.
- Integrate with AmSC FedID for auth and access control
- Deploy AmSC software and services, accessible to all AmSC users/agents via API

Where/how to find more information

- https://amsc.energy.gov/
- Introductory webinars available on our website




Q&A
Genesis Mission
AMERICAN
SCIENCE CLOUD