AWS State, Local, and Education Learning Days

Boston, MA

aws Learning Days
State, Local, and Education

Research

10:15am – 11:15am

100 level

Transforming Your Research with AWS:

AWS helps universities solve complex research challenges with scalable cloud computing and research credits.

11:30am – 12:30pm

100 level

High Performance Computing (HPC) on AWS:

Cloud bursting enables researchers to dynamically scale HPC workloads between on-premises and cloud environments

1:30pm – 3:00pm

200 level

Workshop: Research on AWS:

Hands-on introduction to AWS cloud services for scientific workflows, from basics to advanced research computing.

3:15pm – 4:15pm

200 level

Building a robust research environment:

Secure research environments help institutions meet global compliance requirements and protect sensitive data for funding agencies.

Secure Research Environment (SRE)

Accelerated Compliance & Security Framework for Regulated Organizations

Education | Healthcare | Research | State & Local Government

Brian McCarthy
Solutions Architect
Amazon Web Services

Doug Morand
Solutions Architect
Amazon Web Services

Agenda

- Secure Research Trends & Challenges
- How AWS Helps Research
- SRE Solution Overview
- SRE + SAS Message to Customers
- SRE Technical Design & Framework
- Benefits of SRE Solution
- Implementation Options
- Call to Action
- Q & A

Trends we have observed

Research is strategic for growth, recruiting and knowledge

New and different types of users need large scale computation

Sustaining on-premise while keeping pace with technology

Increasing grant requirements around secure research

How we work with researchers

Grants – Early Engagement
- Identify grants with cloud identifiable spend.

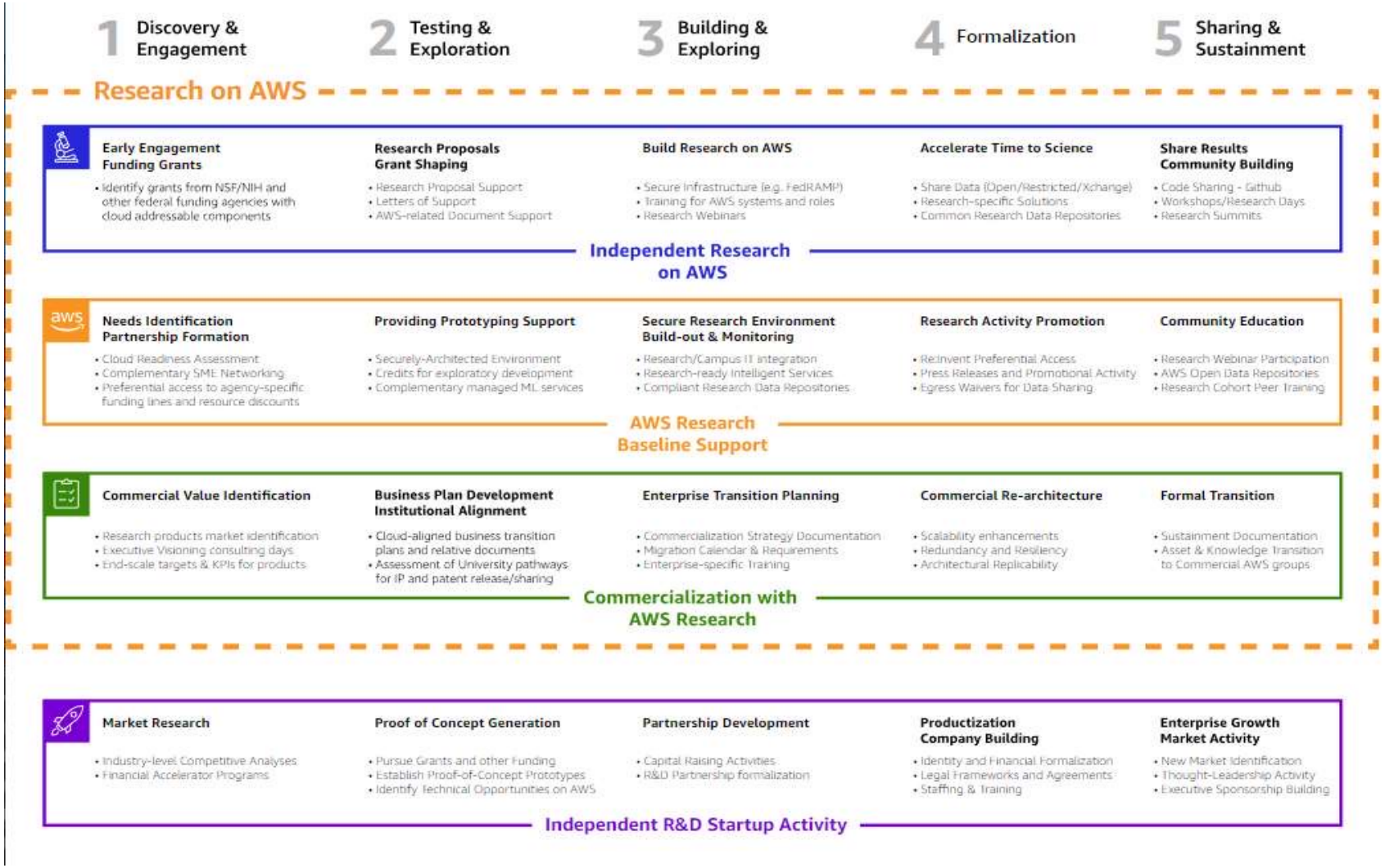
Share Results – Community Building
- AWS Summits
- Case studies, blogposts
- Workshops
- GitHub – Share cod

Research Proposals
- Research proposal support
- Letters of support
- AWS related document support

Accelerate time to science
- Share data – Open Data AWS, Data Exchange
- Research specific solutions
- Data egress waiver

Build Research on AWS
- Simplified access – SSO
- Secure infrastructure
- Get trained on AWS
- Immersion days
- Partner solutions
- AWS cloud credits for research

AWS research support services



Secure Research Environment (SRE)

Secure Research Challenges

- Multi-Framework Compliance & Regulations
- Complex Research Data Handling
- Continuous Security
- IT & Resource Centralization
- Network & Access Management
- Cost & Budget Tracking
- Manual Administration
- Operational Challenges
- Scalability & Elasticity

Secure Research Environment

Designed to help institutions efficiently meet NIST 800-171 Rev 2 requirements through automated security controls.

Built on LZA to provide:

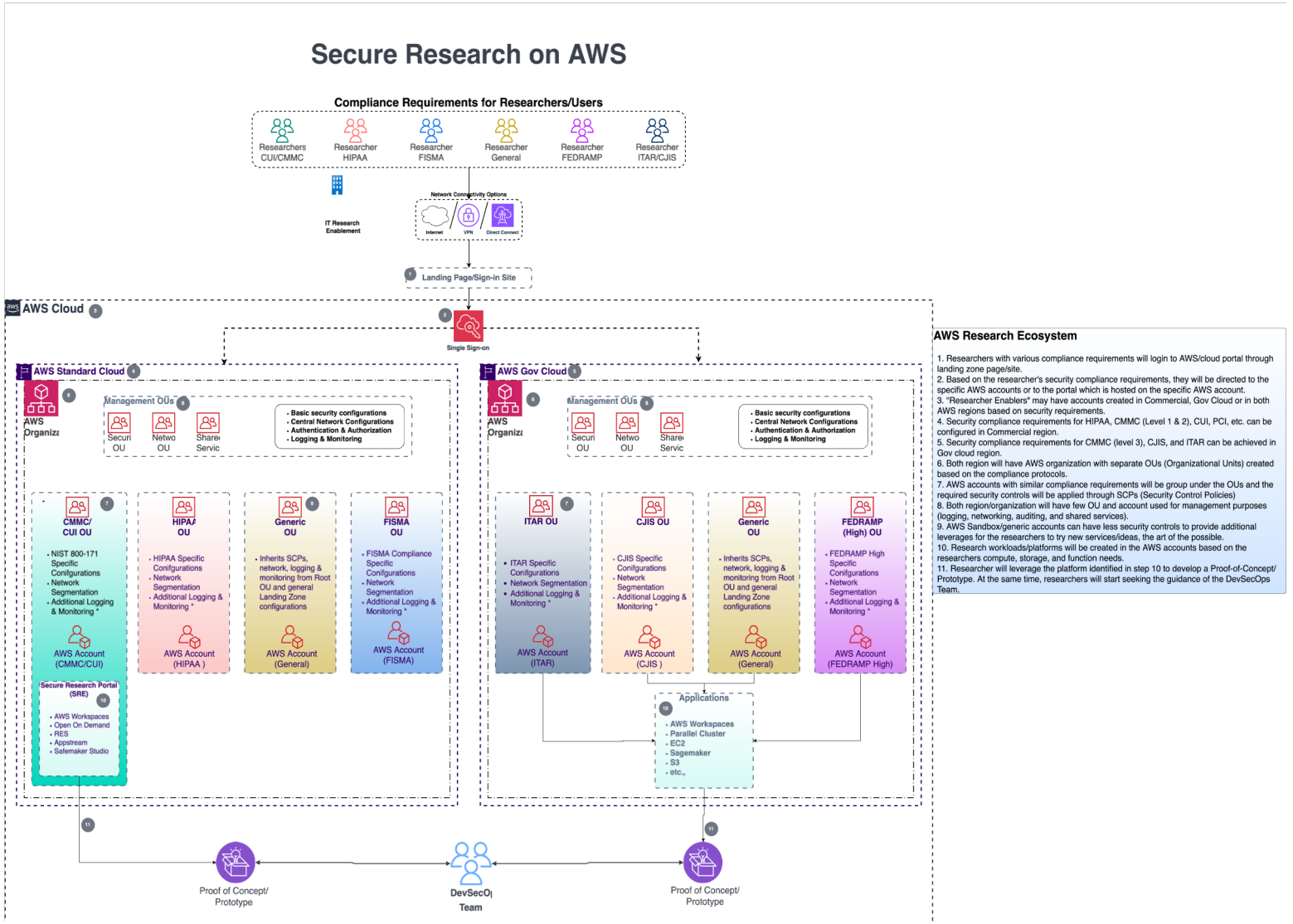
- Comprehensive NIST 800-171 compliance foundation
- Secure CUI data handling capabilities
- Multi-account security architecture

Key Features:

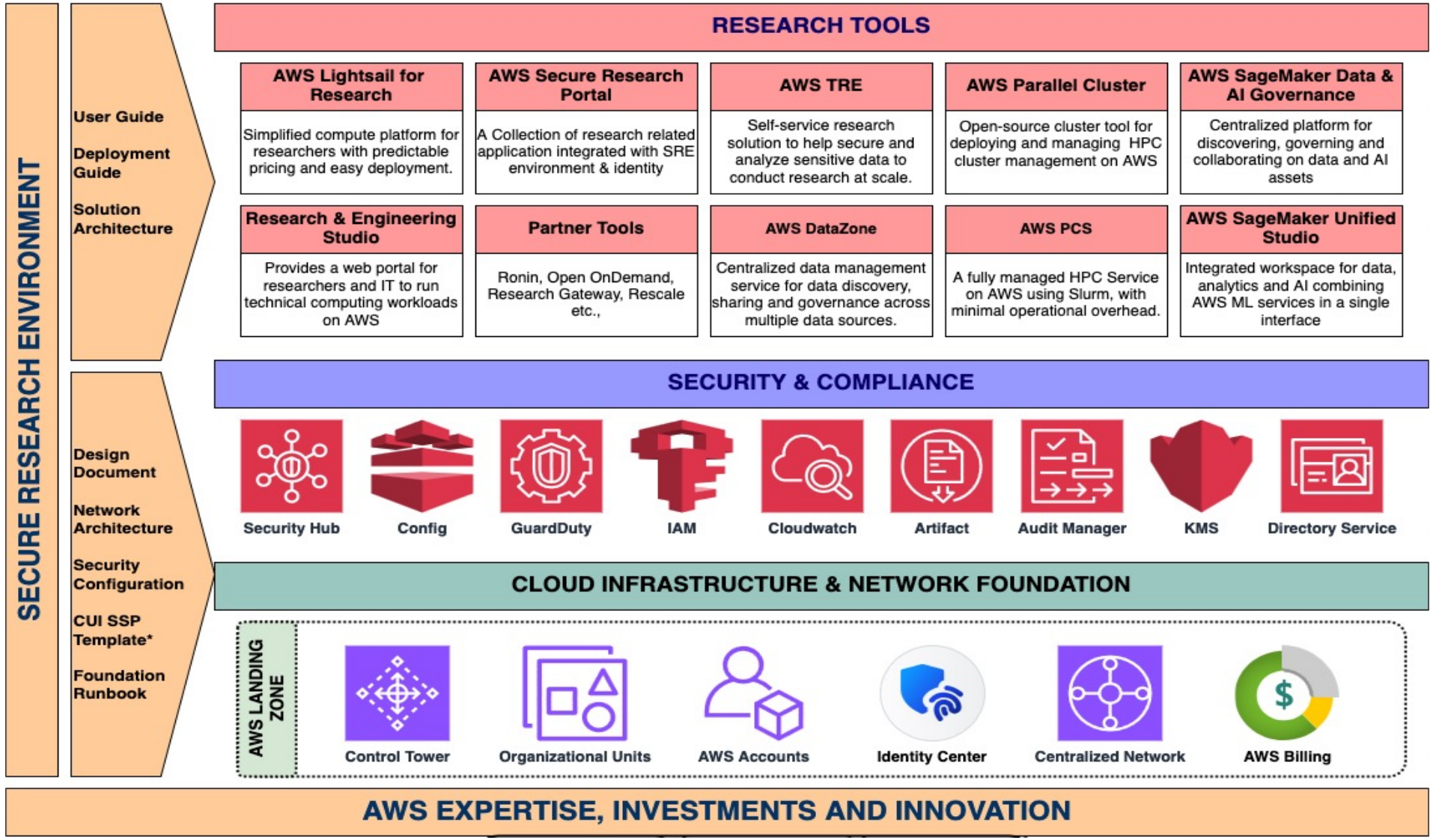
- Pre-configured security controls
- Automated compliance monitoring
- Integrated identity management
- Infrastructure-as-code deployment

Benefits:

- Rapid path to NIH mandate compliance
- Automated security guardrails
- Continuous compliance validation
- Streamlined CUI protection



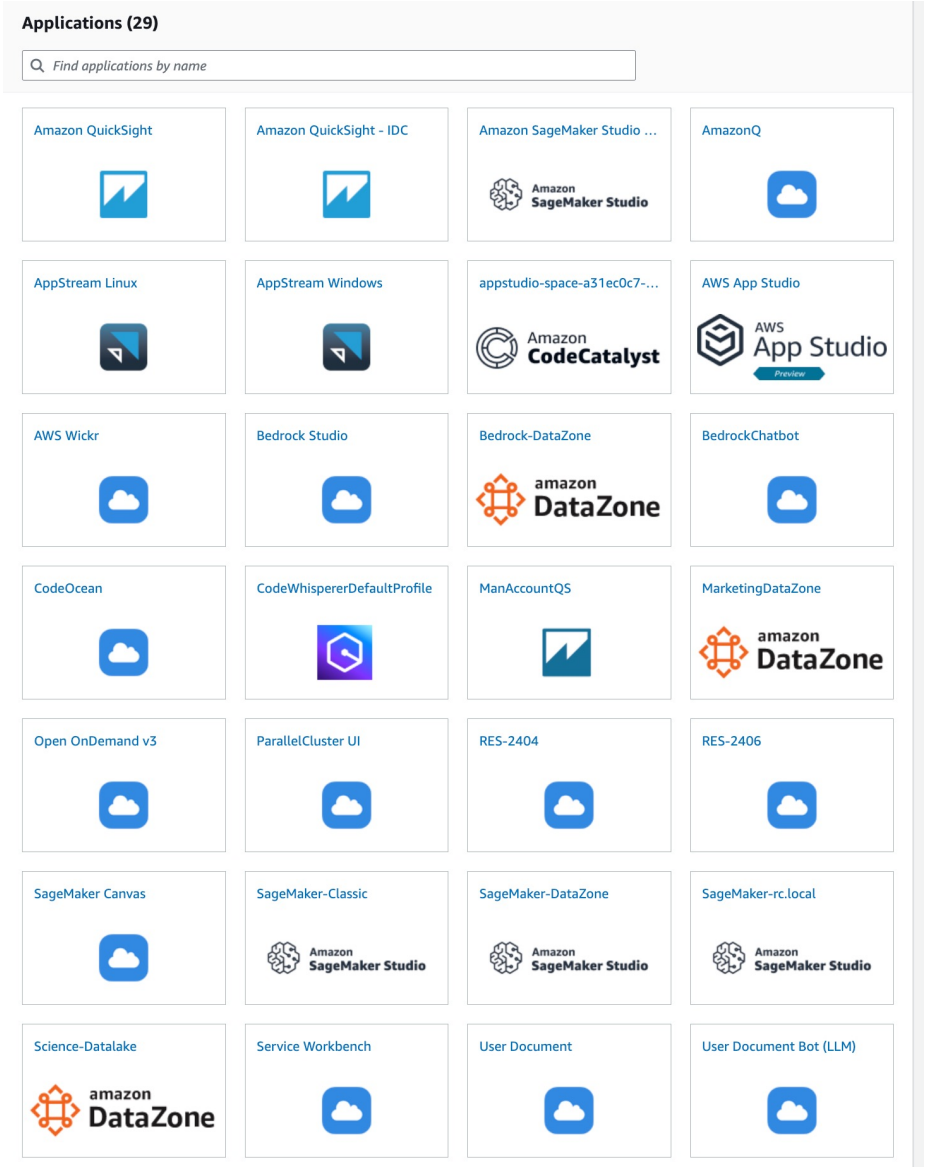
Secure Research Environment Framework



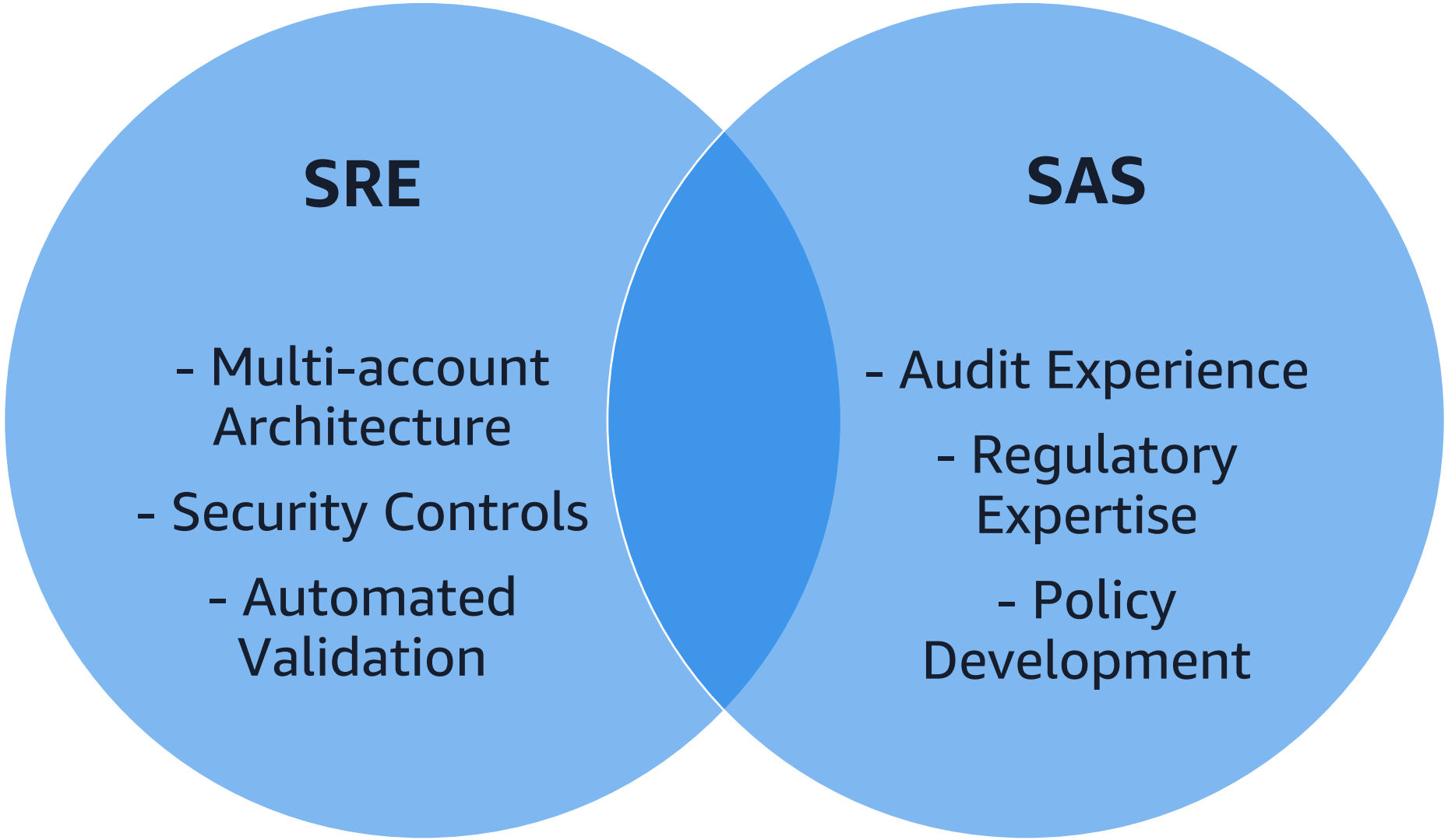
Secure Research Portal

- The Secure Research Portal (SRP), integrated into the Secure Research Ecosystem (SRE), it’s a landing page that simplifies AWS service access for researchers and departments, bypassing the complexity of the AWS console.

- Aimed at all disciplines that requires cluster and end-user computing and storage, such as weather, genomics, geospatial and other researcher domains. The SRP offers a user-friendly interface ensuring federal compliance, including NIST-800 standards which are inherited from SRE.

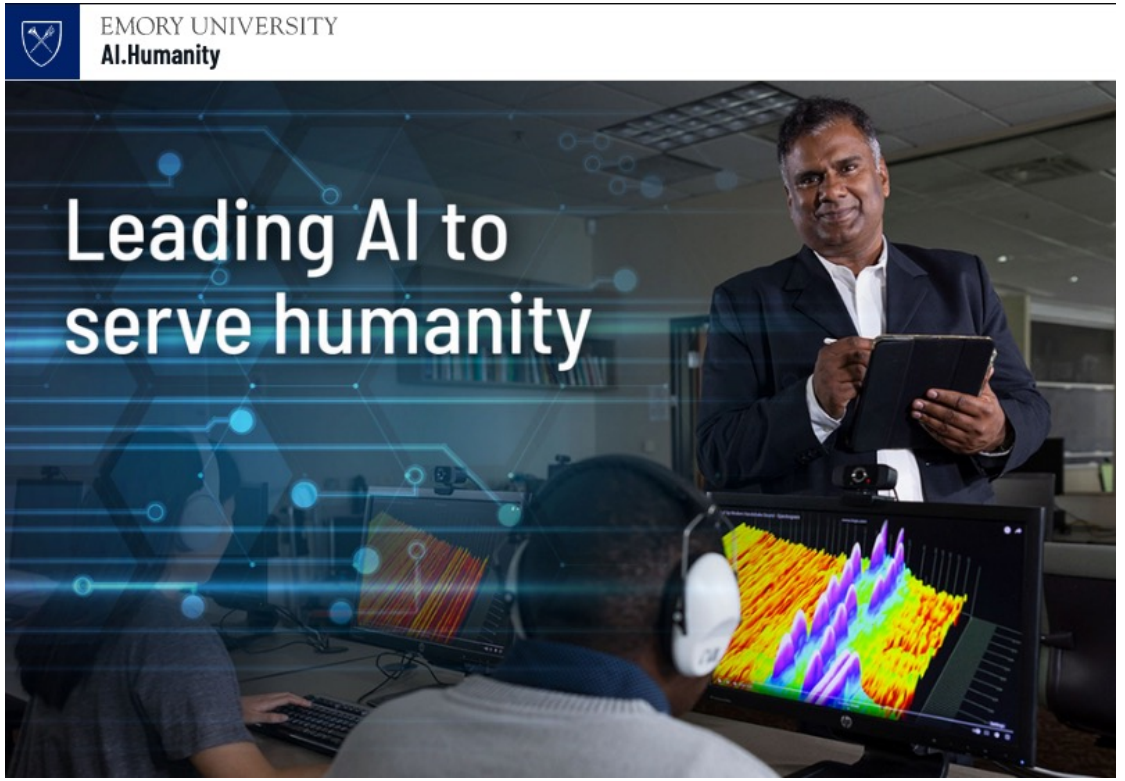


SRE + SAS: Compliance Synergy



Emory AI Humanity Initiative

- Launched in 2022 to explore the societal impacts of artificial intelligence (AI)
- Recruit 60 new faculty who infuse AI-related research and inquiry into disciplines at Emory
- AWS-powered HPC cluster with additional resources around GPU, storage, data analytics, and security (i.e. HIPAA & NIST800-171)
- Enable Emory to become a leading advocate for ethical use of AI in educational and research spaces



https://aws.amazon.com/blogs/publicsector/emory-university-supports-ai-humanity-initiative-with-high-performance-computing-on-aws/

Q&A

Call to Action

Review the recent AWS blog post “Complying with updated NIH Genomic Data Sharing policies on AWS” for a deeper dive on the NIH GDS policy update and AWS can help

Reach out to your AWS account team to discuss your compliance

Reach out to the AWS Solutions Architects on this presentation for further consultation.

- Brian McCarthy (btmccar@amazon.com)
- Doug Morand (domorand@amazon.com)

Thank you!

Brian McCarthy
Solutions Architect
Amazon Web Services
btmccar@amazon.com

Doug Morand
Solutions Architect
Amazon Web Services
domorand@amazon.com

Please complete the survey for this session

Research Track

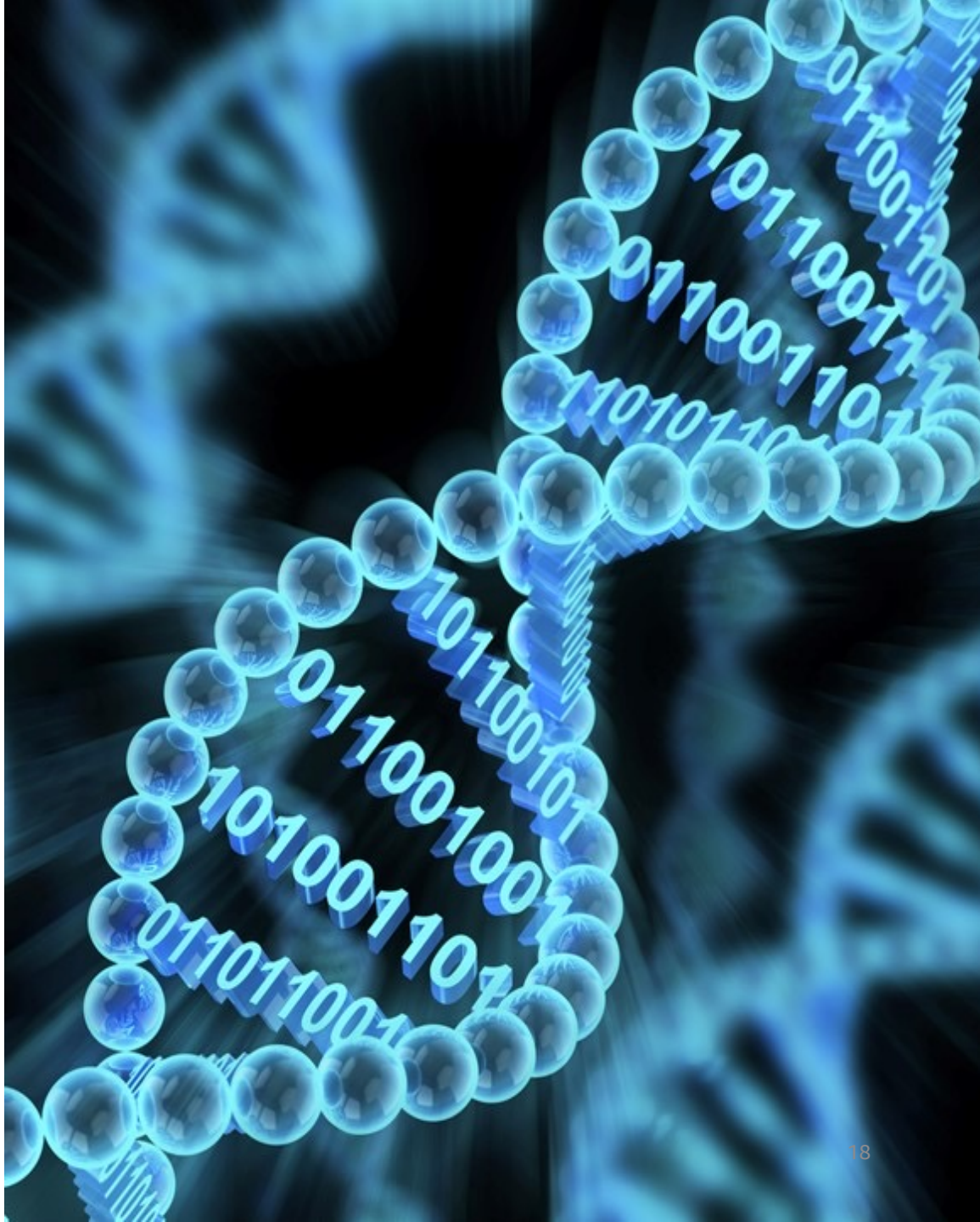
Secure Research Environment (SRE)

aws Learning Days
State, Local, and Education

What is the NIH?

- National Institute of Health (NIH) is part of the U.S. Department of Health and Human Services and is the nation's largest biomedical research agency
- It is comprised of 27 Institutes and Centers, each with a specific research agenda, often focusing on particular diseases or body systems (e.g. National Cancer Institute (NCI), National Human Genome Research Institute (NHGRI) and Center for Informational Technology (CIT)
- NIH had a total Budget of ~$47.3B in 2024; it is the largest public funder of biomedical research in the world

NIH National Institutes of Health
Turning Discovery Into Health



NIH Data Policy Update

What is changing?

National Institute of Health (NIH) has provided detailed updates for two practices under the NIH Genomic Data Sharing (GDS) Policy:

- Modernizing security standards provided in the "NIH Security Best Practices for Controlled-Access Data Subject to the NIH Genomic Data Sharing (GDS) Policy"
- Establishing minimum expectations for access to controlled-access data by developers

The update outlines how researchers can access and manage certain controlled data[1]:

- Approved U.S. and non-U.S. users[2] will be required to attest to NIH that their institution and any third-party system or Cloud Service Providers involved in data analysis or storage comply with **NIST SP 800-171** or an equivalent IT security standard.
- Adherence to the new standard will be included in new or renewed Data Use Certifications or similar agreements.

When does it take effect?

The updated policy takes effect on **January 25, 2025**

Who does this apply to?

User – U.S. and Non-U.S. **NIST SP 800-171** or an equivalent IT security standard
Host - NIST SP 800-53 Moderate
Developer - submit a Developer Use Statement to the NIH Developer Data Access Committee

Notes:
[1]https://nexus.od.nih.gov/all/2024/09/23/updates-to-data-management-and-access-practices-under-the-nih-genomic-data-sharing-policy/
[2]Non-U.S. users of controlled access data that are unable to align to the NIST SP 800-171 are permitted to use the ISO/IEC 27001/27002 instead.
Additional Resources: NIH Security Best Practices for Users of Genomic Controlled Access Data Day 1

Compliance Obligations for Universities

Compliance is a requirement

- With NIH GDS policy update, compliance is a key requirement for higher ed institutions
  - Universities conducting federally funded research must demonstrate compliance to maintain funding eligibility.
  - On Prem systems are subject to CMMC assessments (when applicable) 18 requirements are often evaluated in person by the auditor.
  - Specific clauses like DFARS 252.204-7012, DFARS 252.204-7019, DFARS 252.204-7020 and DFARS 252.204.7021 in defense contracts mandate compliance.
- Dept. of Ed does not explicitly mandate NIST 800-171 for all programs, Higher Ed institutions that receive funding or work on federal projects involving CUI from the Dept. of Education are expected to comply with NIST 800-171 requirements.
- Some examples that the Dept. of Ed deems to be CUI:
  - Federal Student Aid Data
  - Data related to Research Programs
  - FERPA governed data, Federal Work-Study and other Title IV Program Data
- Criminal Justice Information System (CJIS)- 13 Policy areas, 2024 Update MFA implementation.
- HIPAA- January 2025 Health and Human Services proposing new changes to the HIPAA Security Rule

Consequences of non-compliance

- Loss of access to controlled data sets
- Implications on future funding for research
- If an institution repeatedly fails to meet these requirements, it could face broader restrictions on accessing NIH controlled data across multiple research teams
- Violations of NIST 800-171 controls can result in severe consequences beyond fines, including federal lawsuits, potential loss of government contracts, and significant damage to an institution's reputation and credibility in handling sensitive information
- Department of Justice is the enforcement arm of the DoD- False Claims Act fines in the millions
- CJIS audits performed by the FBI non-compliance can be loss of access to the CJIS data
- HIPAA enforcement is done by Office of Civil Rights (OCR)- data breach can result in large fines based on negligence

Overview NIST 800-171

- NIST 800-171 is a key framework that provides guidance on protecting sensitive information, specifically **Controlled Unclassified Information (CUI)**, when it is processed, stored, or transmitted in non-federal systems. Its relevance to HCLS environments stems from the NIH GDS guidance which closely aligns with this framework and becomes effective Jan 25, 2025 for federally funded research and projects that require handling of CUI.
- The primary purpose is to safeguard CUI and ensure that non-federal entities (e.g., universities, public and private organizations including research orgs) can demonstrably secure sensitive data associated with federal contracts. This applies to organizations receiving federal funding, handling federal contracts, or working on projects involving CUI.
- NIH GDS and NIST 800-171 Common Goals: Data Security, Confidentiality, Compliance.
- Areas of Alignment: Access Control, Audit and Accountability, Data Security and Encryption, and Incident Response.

Overview NIST 800-171

- What is Controlled Unclassified Information (CUI)?
  - CUI is information the federal government considers sensitive but not classified. Examples include:
    - Research findings funded by a federal agency (NASA, NIH, DoD, Dept. of Education, Dept. of Energy, Dept. of Commerce, etc.).
    - Data related to infrastructure, healthcare, or export-controlled technologies.
    - For HCLS:
      - Clinical trials and study data, PII.
      - Sensitive student or institutional data shared with federal entities.
      - Genomics Data, Patient, and Healthcare Data.
      - Biodefense and public health data.

Key Components of NIST 800-171R2

The framework includes 110 security controls 320 objectives organized into 14 families that cover various aspects of information security. These include:

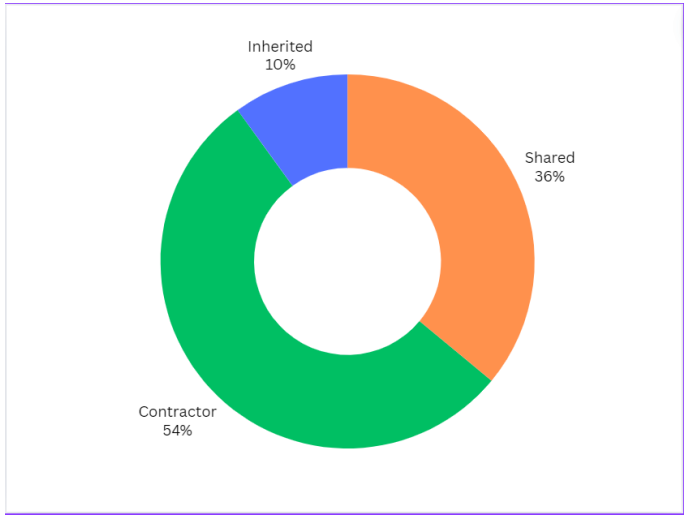
1. **Access Control** (e.g.: Restricting access to only authorized users / enabling MFA)
2. **Awareness and Training** (e.g.: Regular documented training for staff and researchers on security risks and policies)
3. **Audit and Accountability** (e.g.: Monitoring and logging system and user activity)
4. **Configuration Management** (e.g.: Ensuring secure configurations of systems and software)
5. **Incident Response** (e.g.: Detecting, reporting, and mitigating security incidents)
6. **System and Communications Protection** (e.g.: Protecting data in transit and at rest)
7. **Identification and Authentication** (e.g.: Ensure the identity of users and systems accessing resources)
8. **Maintenance** (e.g. Control tooling, techniques, and processes used to maintain systems securely)
9. **Media Protection** (e.g.: Safeguard sensitive data on physical and digital media)
10. **Personnel Security** (e.g. Ensure personnel handling sensitive information are trustworthy)
11. **Physical Protection** (e.g.: Prevents unauthorized physical access to systems and facilities)
12. **Risk Assessment** (e.g.: Identifies and evaluates potential threats, risks, and gaps in security posture.)
13. **Security Assessment** (e.g.: Monitors and evaluates the implementation of security controls)
14. **System and Information Integrity** (e.g.: Detects and addresses vulnerabilities and errors in a timely manner)

Meeting these compliance objectives requires extensive expertise, capabilities, and a robust governance infrastructure and can be a significant lift for customers to address on their own

Compliance requirements for the 14 family of controls can be addressed with AWS services and solutions and with TAM, AWS ProServ, or partner engagements

Governance and Regulatory Requirements- Domains

1. Access Control (AC)
2. Awareness & Training (AT)
3. Audit & Accountability (AU)
4. Configuration Management (CM)
5. Identification & Authentication (IA)
6. Incident Response (IR)
7. Maintenance (MA)
8. Media Protection (MP)
9. Personnel Security (PS)
10. Physical Protection (PE)*
11. Risk Assessment (RA)
12. Security Assessment (CA)
13. System & Communication Protection (SC)
14. System & Information Integrity (SI)



Key Components of NIST 800-171R3

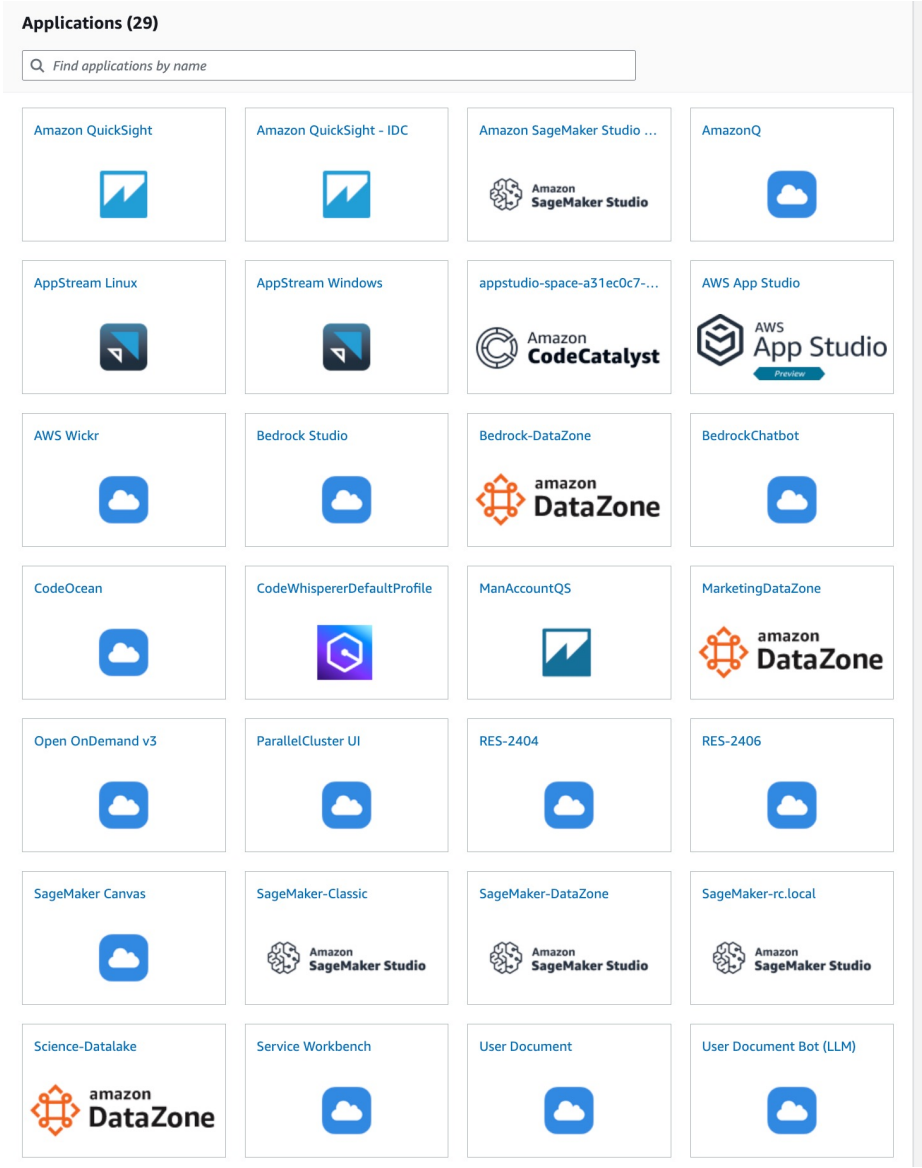
The framework includes 109 security controls 445 objectives organized into 14 families that cover various aspects of information security. These include:

1. **Access Control** (e.g.: Restricting access to only authorized users / enabling MFA)
2. **Awareness and Training** (e.g.: Regular documented training for staff and researchers on security risks and policies)
3. **Audit and Accountability** (e.g.: Monitoring and logging system and user activity)
4. **Configuration Management** (e.g.: Ensuring secure configurations of systems and software)
5. **Incident Response** (e.g.: Detecting, reporting, and mitigating security incidents)
6. **System and Communications Protection** (e.g.: Protecting data in transit and at rest)
7. **Identification and Authentication** (e.g.: Ensure the identity of users and systems accessing resources)
8. **Maintenance** (e.g. Control tooling, techniques, and processes used to maintain systems securely)
9. **Media Protection** (e.g.: Safeguard sensitive data on physical and digital media)
10. **Personnel Security** (e.g. Ensure personnel handling sensitive information are trustworthy)
11. **Physical Protection** (e.g.: Prevents unauthorized physical access to systems and facilities)
12. **Risk Assessment** (e.g.: Identifies and evaluates potential threats, risks, and gaps in security posture.)
13. **Security Assessment** (e.g.: Monitors and evaluates the implementation of security controls)
14. **System and Information Integrity** (e.g.: Detects and addresses vulnerabilities and errors in a timely manner)
15. **Planning (e.g.: Policy, procedures, rules of behavior)**
16. **System and Services Acquisition (e.g.: SDLC, external systems, etc.)**
17. **Supply Chain Risk Management (e.g.: SCRM policy/procedures)**

Secure Research Portal

The Secure Research Portal (SRP), integrated into the Secure Research Ecosystem (SRE), it's a landing page that simplifies AWS service access for researchers and departments, bypassing the complexity of the AWS console.

Aimed at all disciplines that requires cluster and end-user computing and storage, such as weather, genomics, geospatial and other researcher domains. The SRP offers a user-friendly interface ensuring federal compliance, including NIST-800 standards which are inherited from SRE.



Key Components of NIST 800-171

Control Family	AWS Services
	IAM, VPC, SSO, MFA, NACLs
	AWS Training, Security Hub, Trusted Advisor
	CloudTrail, CloudWatch, Config
	Config, Systems Manager, CloudFormation
	IAM, Cognito, Secrets Manager, MFA
	GuardDuty, Security Hub, CloudTrail, SNS
	Systems Manager, CloudFormation
	S3, KMS, Backup
	IAM + Identity Provider Integrations
	AWS Data Center Security
	Trusted Advisor, Inspector, Well-Architected
	Security Hub, Audit Manager, CloudWatch
	KMS, ACM, VPC, Direct Connect, VPN

Controlled-access repositories implementing NIH Security Best Practices (1/2)

Repository	Access System	URL
Database of Genotypes and Phenotypes (dbGaP)	dbGaP Access System	https://www.ncbi.nlm.nih.gov/gap/
BioData Catalyst	dbGaP Access System	https://biodatacatalyst.nhlbi.nih.gov/
The NHGRI Genomic Data Science Analysis, Visualization, and Informatics Lab-Space (AnVIL)	dbGaP Access System	https://anvilproject.org/
National Cancer Institute (NCI) Genomic Data Commons	dbGaP Access System	https://gdc.cancer.gov/
Cancer Data Service (CDS)-Trusted Partner	dbGaP Access System	https://dataservice.datacommons.cancer.gov/#/home
Kids First Data Resource	dbGaP Access System	https://kidsfirstdrc.org/resources/
INvestigation of Co-occurring conditions across the Lifespan to Understand Down syndromE (INCLUDE) Data Hub	dbGaP Access System	https://portal.includedcc.org/login?redirect_path=/dashboard
Restricted Portion of Sequence Read Archive (SRA)	dbGaP Access System	https://www.ncbi.nlm.nih.gov/sra
National Institute of Mental Health Data Archive (NDA)	NDA Access System	https://nda.nih.gov/
NDA: National Institute on Alcohol Abuse and Alcoholism Data Archive (NIAAADA)	NDA Access System	https://nda.nih.gov/niaaa/
Adolescent Brain Cognitive Development Study (ABCD)	NDA Access System	https://abcdstudy.org/scientists/data-sharing/

Source: *https://sharing.nih.gov/accessing-data/NIH-security-best-practices*

Controlled-access repositories implementing NIH Security Best Practices (2/2)

Repository	Access System	URL
The Neuroscience Multi-omic Data Archive Brain/NeMo	NDA Access System	https://nemoarchive.org/data/
The CommonMind Consortium Knowledge Portal	NDA & Synapse Access Systems	https://www.synapse.org/Synapse:syn2759792/wiki/197283
PsychENCODE Knowledge Portal	NDA & Synapse Access Systems	https://www.synapse.org/Synapse:syn4921369/wiki/235539
National Institute on Aging (NIA) Genetics of Alzheimer's Disease Data Storage Site (NIAGADS)	NIAGADS Access System	https://www.niagads.org/home
Accelerating Medicines Partnership® Parkinson's Disease (AMP® PD)	AMP® PD Access System	https://amp-pd.org/
Parkinson's Disease Biomarkers Program Data Management Resource (PDBP DMR)	PDBP DMR Access System	https://pdbp.ninds.nih.gov/
PEGS: Personalized Environment and Genes Study	PEGS Access System	https://www.niehs.nih.gov/research/atniehs/labs/crb/studies/pegs
NIMH Repository and Genomics Resources (NRGR)	NRGRAccess System	https://www.nimhgenetics.org/
NIDCR FaceBase	FaceBase Access System	https://www.facebase.org/

Source: *https://sharing.nih.gov/accessing-data/NIH-security-best-practices*

AWS Services for NIST 800-171 Compliance

AWS Identity and Access Management
Securely manage access to AWS services

AWS Security Hub
Automate AWS security checks and centralize security alerts

Amazon GuardDuty
Protect your AWS accounts with intelligent threat detection

Amazon Inspector
Automated and continual vulnerability management at scale

AWS Identity Center
Centrally manage workforce access to multiple AWS accounts and applications

Amazon CloudWatch
Observe and monitor resources and applications on AWS, on premises, and on other clouds

AWS Config
Assess, audit, and evaluate configurations of your resources

AWS CloudTrail
Track user activity and API calls

AWS Organizations
Policy-based management for multi-account strategy on AWS

AWS Key Management Service (KMS)
Create and control keys used to encrypt or digitally sign your data.

AWS Research and Engineering Studio (RES)
Self-service portal designed for scientists and engineers to securely access and manage their workspaces

AWS security, identity, and compliance solutions

Identity & access management

AWS Identity & Access Management (IAM)

AWS Single Sign-On

AWS Organizations

AWS Directory Service

Amazon Cognito

AWS Resource Access Manager

Detection

AWS Security Hub

Amazon GuardDuty

Amazon Inspector

Amazon CloudWatch

AWS Config

AWS CloudTrail

VPC Flow Logs

AWS IoT Device Defender

Infrastructure protection

AWS Firewall Manager

AWS Network Firewall

AWS Shield

AWS WAF – Web application firewall

Amazon Virtual Private Cloud (VPC)

AWS PrivateLink

AWS Systems Manager

Data protection

Amazon Macie

AWS Key Management Service (KMS)

AWS CloudHSM

AWS Certificate Manager

AWS Secrets Manager

AWS VPN

Server-Side Encryption

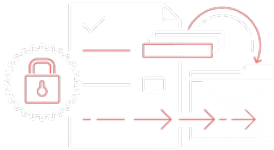
Incident response

Amazon Detective

CloudEndure DR

AWS Config Rules

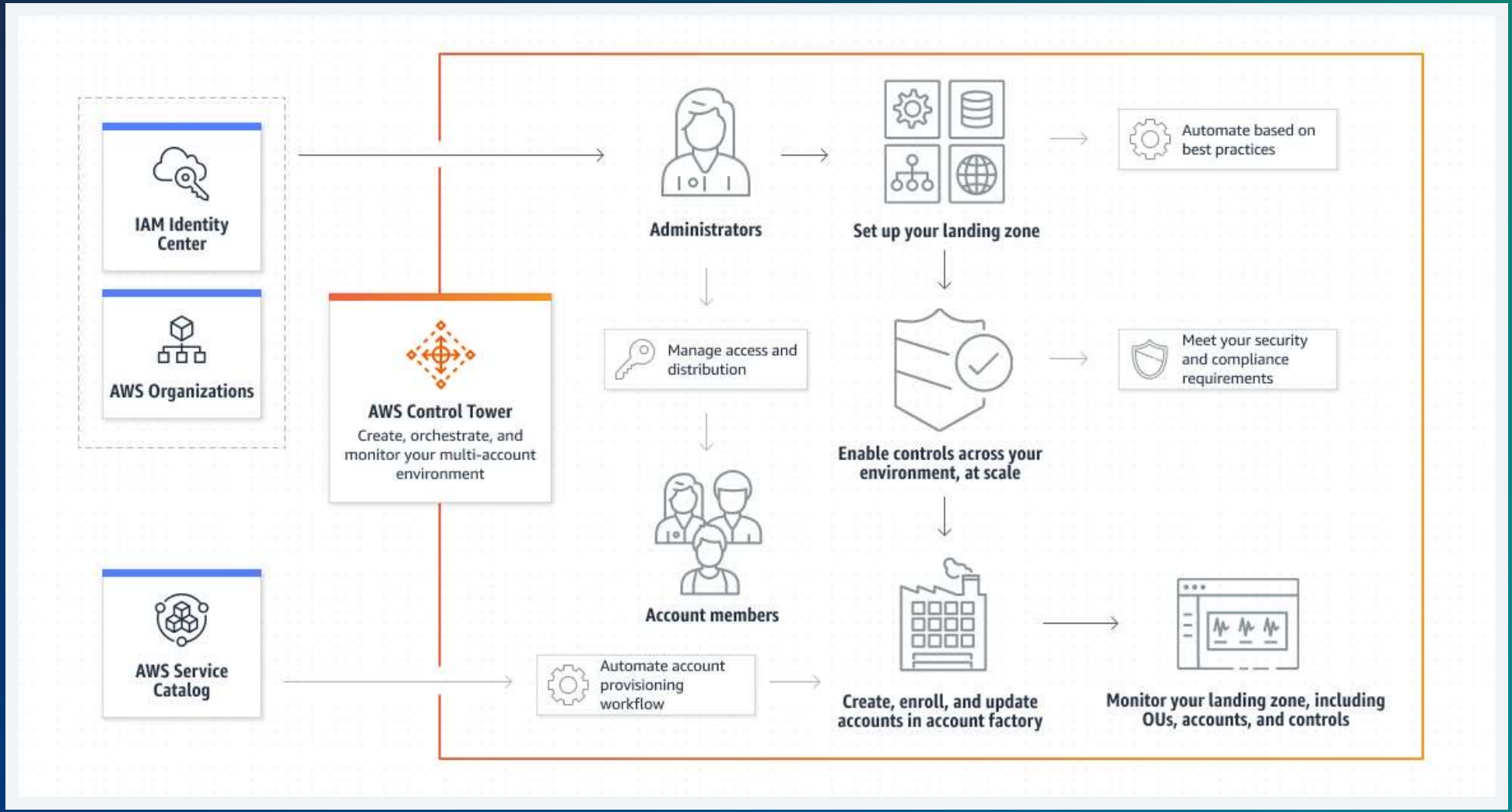
AWS Lambda



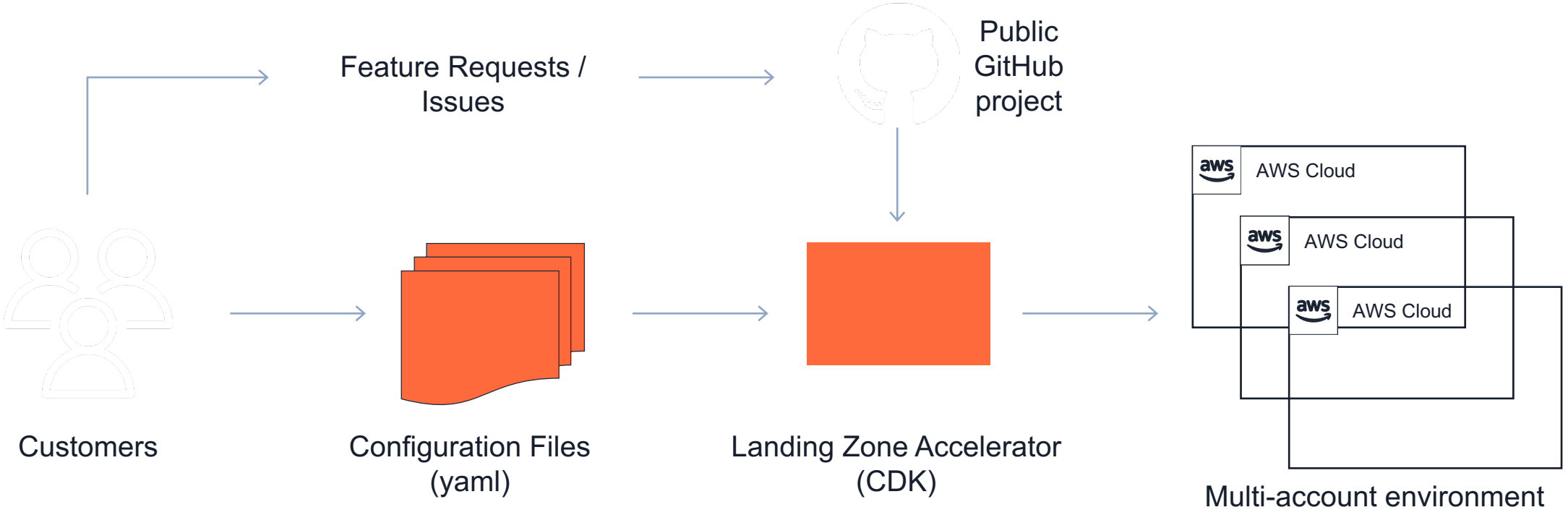
AWS Artifact

AWS Audit Manager

AWS Landing Zone – Generalized Architecture

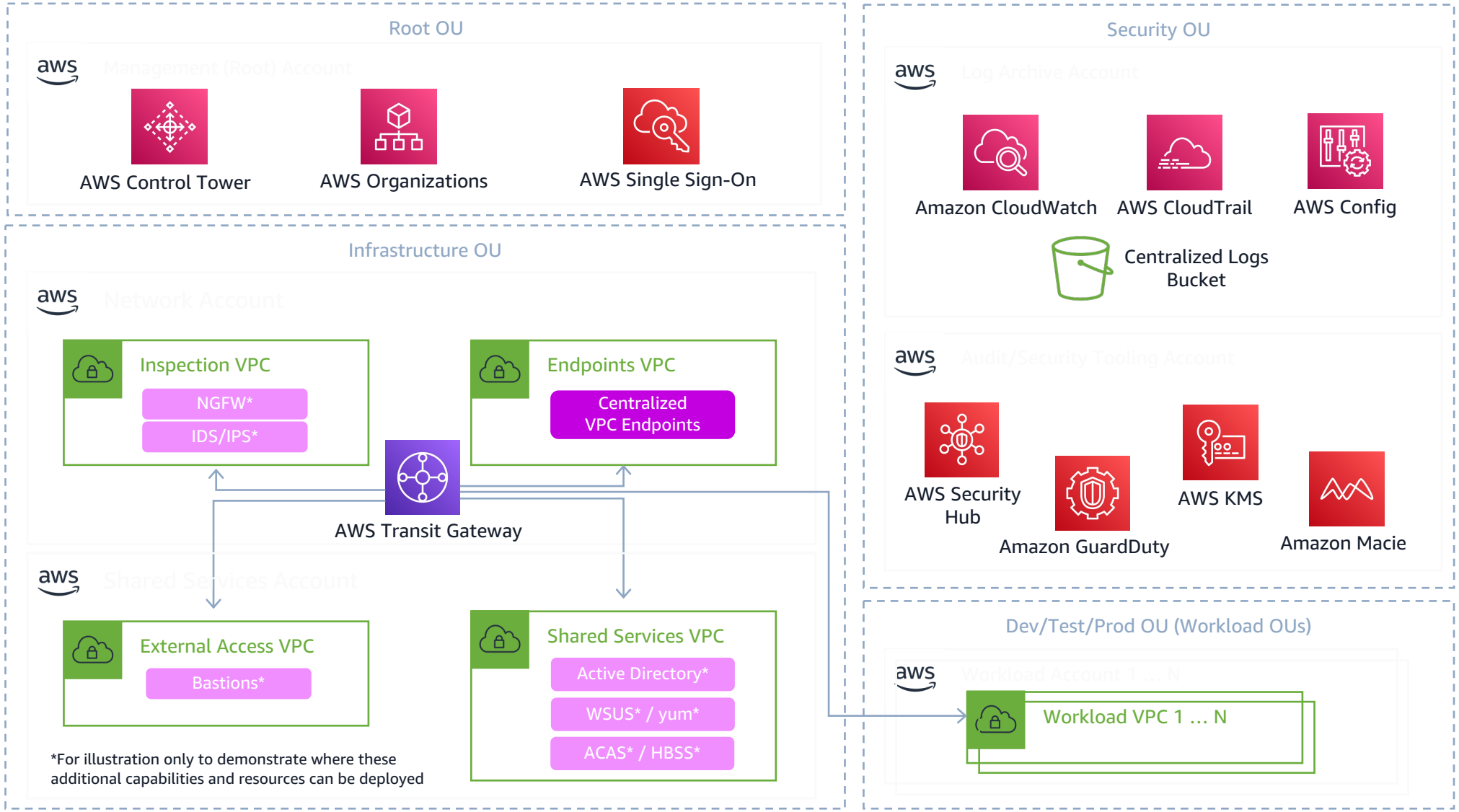


Architecture



Landing Zone Accelerator on AWS

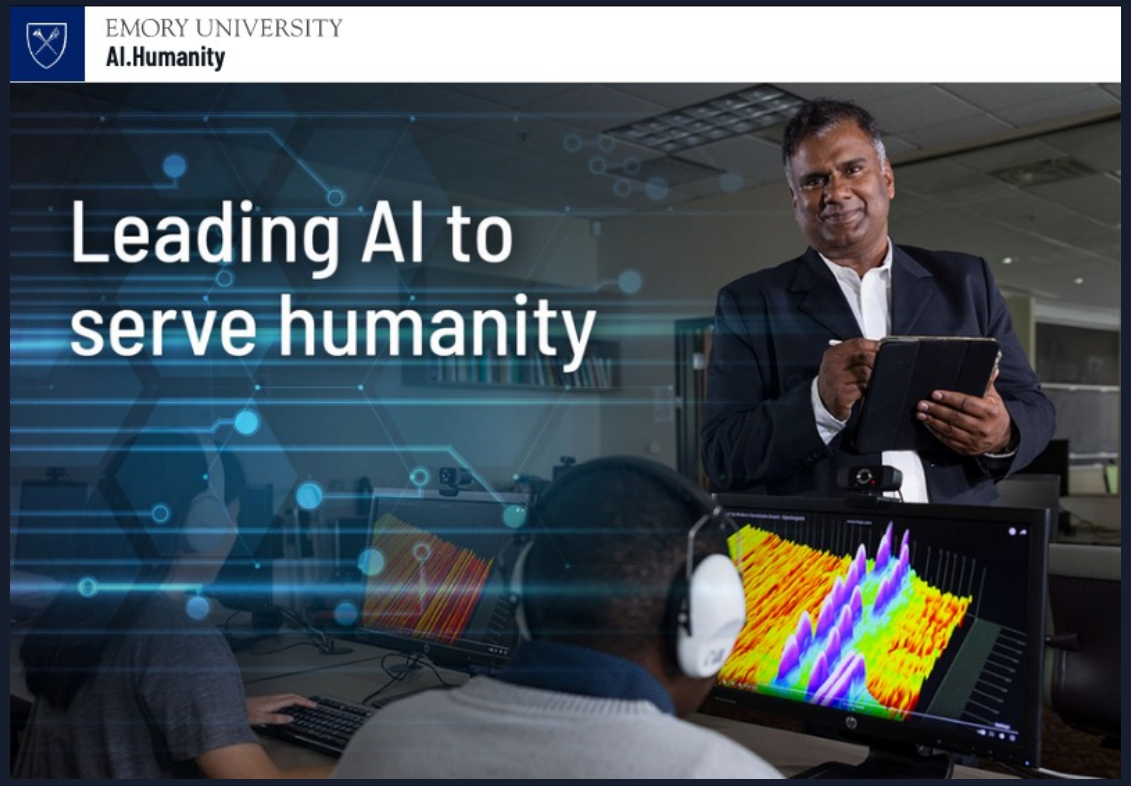
1. Installation Template (AWS CloudFormation)
2. Configuration Files (YAML)
3. AWS CodePipeline
4. AWS CDK



Landing Zone Accelerator

Emory AI Humanity Initiative

- Launched in 2022 to explore the societal impacts of artificial intelligence (AI)
- Recruit 60 new faculty who infuse AI-related research and inquiry into disciplines at Emory
- AWS-powered HPC cluster with additional resources around GPU, storage, data analytics, and security (i.e. HIPAA & NIST800-171)
- Enable Emory to become a leading advocate for ethical use of AI in educational and research spaces



https://aws.amazon.com/blogs/publicsector/emory-university-supports-ai-humanity-initiative-with-high-performance-computing-on-aws/

Children's Hospital of Philadelphia - FPGA

Fastest-Ever Analysis Of 1,000 Genomes

1,000 diverse pediatric genomes were processed into useable data files in two hours and twenty-five minutes

World Record for processing Genomes using FPGAs

Deployed on 1,000 Amazon EC2 F1 instances

One of the largest cohorts for this demographic that has been sequenced to date

Utilized Edico Genome's DRAGEN™ Genome Pipeline

Secure Research Environment

THE NEED

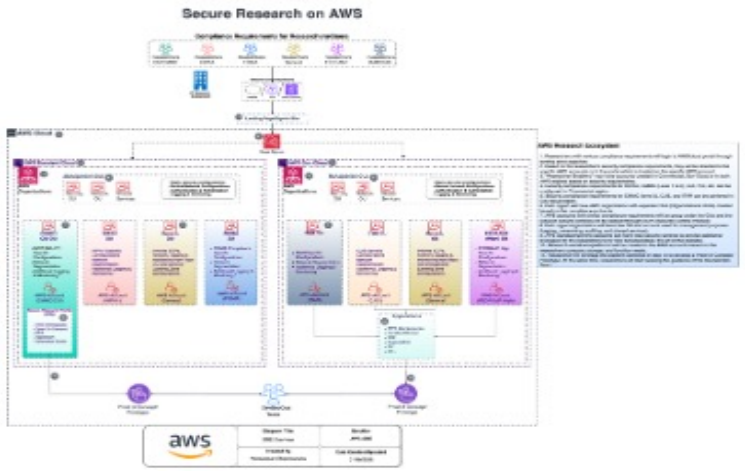
Problem

In today's digital landscape, academic institutions and their researchers, funded by organizations like the Department of Defense (DoD), the National Institutes of Health (NIH), and the National Science Foundation (NSF), face a growing need to establish secure research environments that meet stringent compliance requirements. Initiatives such as the Cybersecurity Maturity Model Certification (CMMC), implemented by the DoD, and the NIH's Information Security Program (ISP), underscore the importance of robust cybersecurity practices and protection of sensitive data.

Benefits of Solution

The Secure Research Environment on AWS Leverages AWS's advanced infrastructure and compliance capabilities, institutions can establish a comprehensive approach that encompasses data protection, access controls, and cybersecurity measures. This approach enables them to effectively address the specific compliance obligations mandated by funding agencies

ARCHITECTURE DIAGRAM



SOLUTION

Approach

Delivering an AWS Secure Research Environment requires a comprehensive approach that involves various stakeholders and interdependencies. This environment is built upon the collaboration between AWS Utilities, AWS Builders (including CAT, ProServe, Partners, and Customers), Customers' responsibilities, and Vendors (such as EDU Resellers and 3rd Party assessors.

Technology

- ***Infrastructure:*** The initial phase focuses on setting up the core infrastructure of the AWS Secure Research Environment. This involves creating a foundational environment that includes essential services such as Virtual Private Cloud (VPC), Identity and Access Management (IAM), and network security controls. The core infrastructure provides a secure and scalable foundation for the subsequent enclaves or OUs.
- ***Organizational Units:*** OU is created to meet the compliance requirements set forth by regulations such as NIST Special Publication 800-171 and HIPAA. This enclave/OUs implements additional security controls such as data encryption, access restrictions, and continuous monitoring.

TARGET PERSONAS

Organization Type

- Higher Education Institutions
- Nonprofit organization (NPO)
- State and Local Government (SLG)
- Defense Integrated Business Systems (DIBS)

Primary (Decision Makers)

- CIO
- CISO
- CTO

Secondary (Influencers)

- VP of Research
- Director of Research Computing

KEY BUILDERS

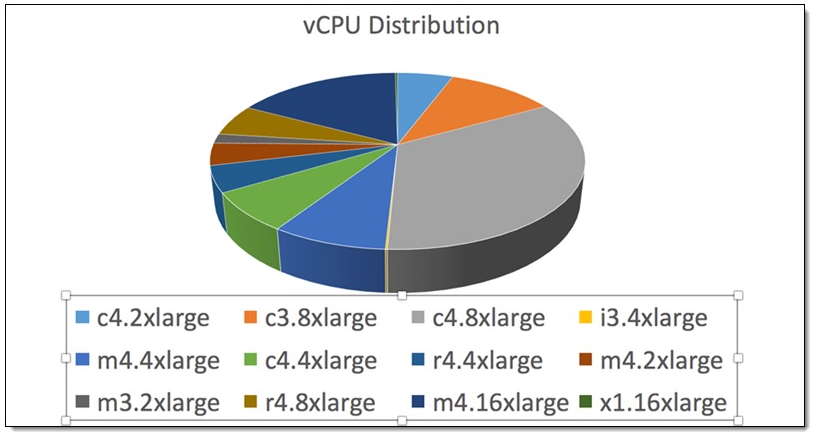
- ProServe (including CAT when needed)
- StackArmor
- Bladestack

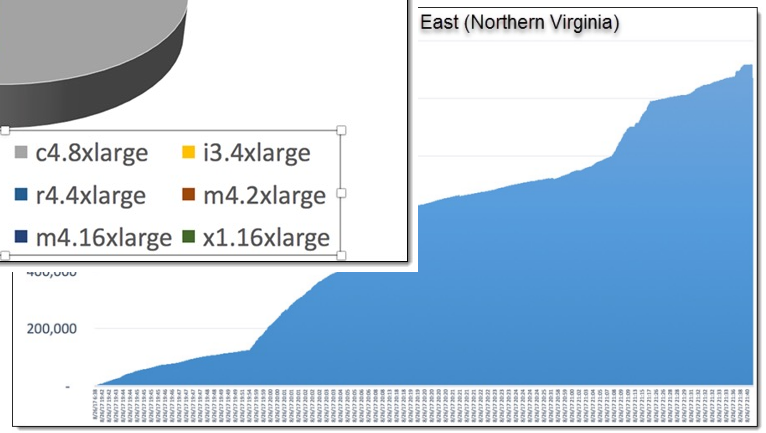
Clemson University - Natural Language Processing

The researchers conducted nearly half a million topic modeling experiments to study how human language is processed by computers.

The 1.1 Million vCPU count usage is comparable to the core count on the largest supercomputers in the world.

CLEMSON UNIVERSITY

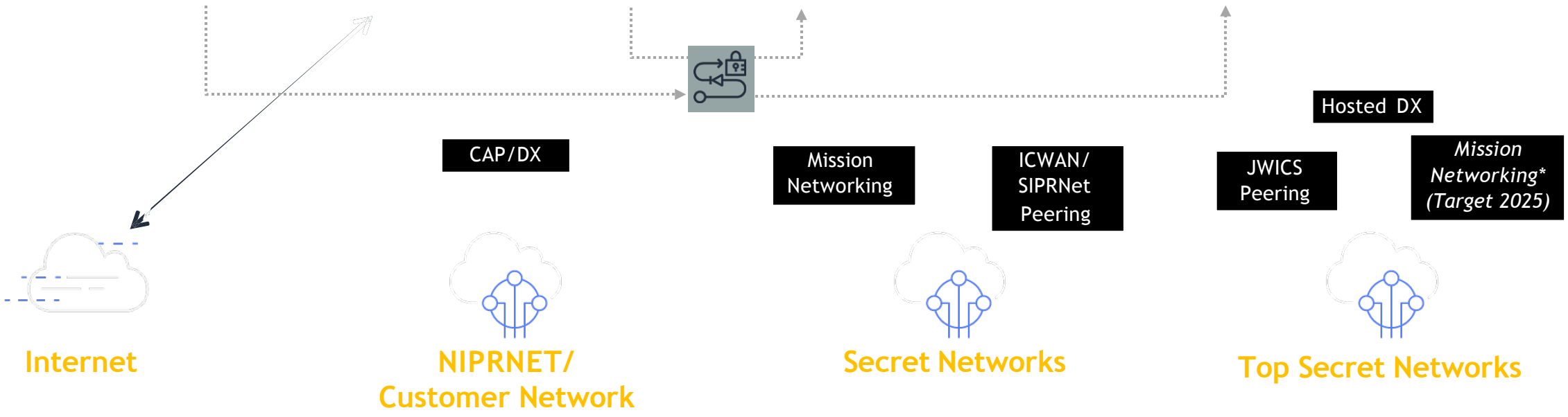


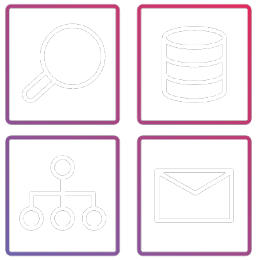


Amy Apon

AWS US Regions Infrastructure

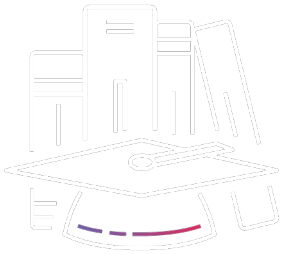
US Intelligence Community (IC) Sponsored





AWS provides the broadest and deepest solutions for research globally

We provide common research tools, cluster technologies, and solutions for security and compliance



We have experts who understand your research needs

The AWS ML stack

Broadest and most complete set of machine learning capabilities

AI SERVICES

Vision
Rekognition

Chatbots
Lex

Business tools
Personalize, Forecast
Fraud Detector
Lookout for Metrics

Search
Kendra

Healthcare
HealthLake
Comprehend Medical
Transcribe Medical

Speech
Polly
Transcribe

Text
Comprehend
Translate
Textract

Contact centers
Contact Lens
Connect Voice ID

Code + DevOps
CodeGuru
DevOps Guru

Industrial
Panorama Appliance and SDK, Monitron, Lookout for Equipment, Lookout for Vision

AMAZON SAGEMAKER

Label data

SAGEMAKER STUDIO IDE

Data collection prep | Store features | Detect bias and explain predictions | Visualize in notebooks | Pick algorithm | Train models faster | Tune parameters | Deploy in production | Manage & monitor

CI/CD

Manage edge devices

ML FRAMEWORKS & INFRASTRUCTURE

TensorFlow, PyTorch, Apache MXNet | Deep learning AMIs & containers | GPUs | Inferentia | Elastic inference | FPGA

aws