aws

AWS for Research

Jessica Steed

Account Executive

Hatem Siyala

Solutions Architect

Why Use AWS for Research?

Simple for Researchers
Reducing user friction by providing powerful yet simple to access services

Security Controls & Roles
Intuitive to manage access to data

Budget & Cost Management
Cost transparency and simple estimation

Long-term Investment
We're here to stay with stable long-term technology

Open Innovation
Constantly defining new technologies and committed to providing tools to enable research

Culture & People
Shared mindset as we are constantly publishing, teaching, and engaging with academia

AWS Accelerates Research

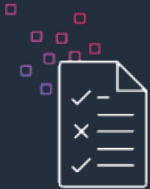### Science, not Servers
Compute when you need it at any scale

Collaboration
Access datasets that span institutions and borders

Research Data Management
Storage, secure access, and management

Share & Reproduce Research
A common platform for reproducing scientific analyses

State-of-the-Art Analytics
Use data science methods in your research

Security
A collection of tools to protect data and privacy

Enterprise IT **and** Research

Enterprise IT cares about…

Stability & Operational Excellence

Cost Optimization

Security & Compliance

Research cares about…

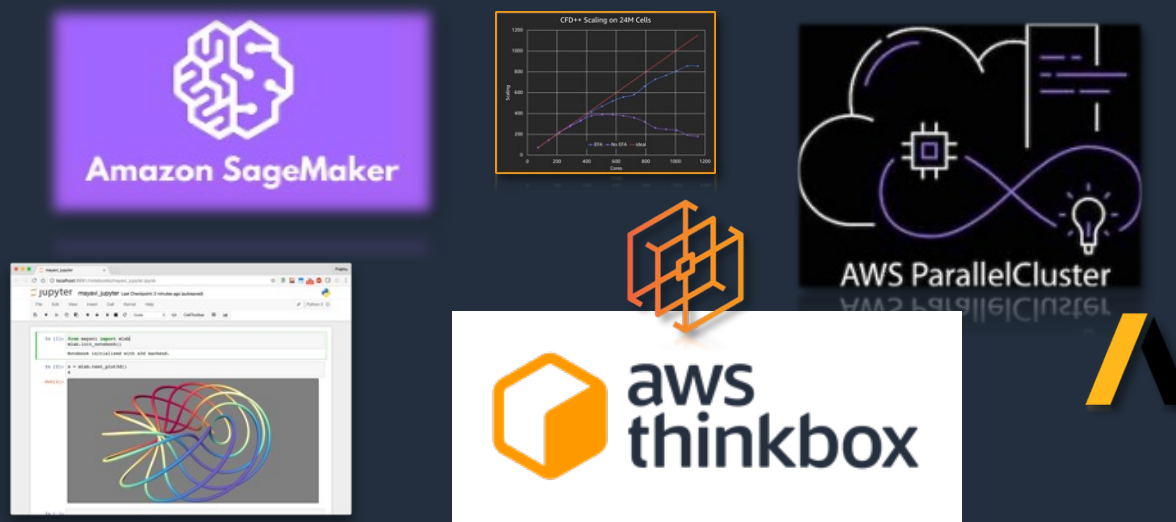**New Ideas & Innovation**

Speed & Agility

Value & Results

Research is not one size fits all
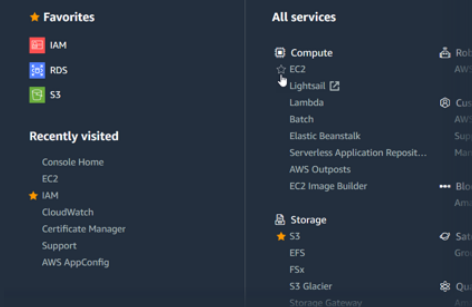
Researcher Workspaces

Research Gateway

RONIN

AWS Research & Engineering Studio

Solutions & Guidance

Amazon SageMaker

AWS ParallelCluster

aws thinkbox

Native AWS

AWS Console, CLI, SDKs, etc.

Storage and data solutions to meet the specific researcher needs

Researcher Technical Requirements & Capability

Simpler/Common Requirements

Complex/Niche Requirements
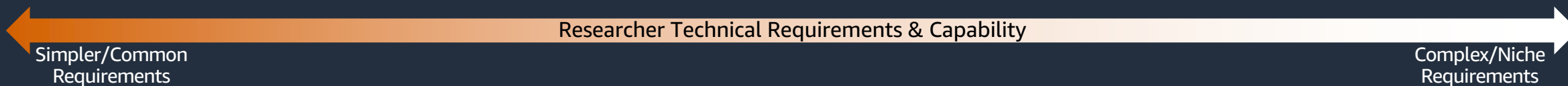
A Cloud-Based Model Promotes Collaboration

Traditional Approach:

Move data to computing resources
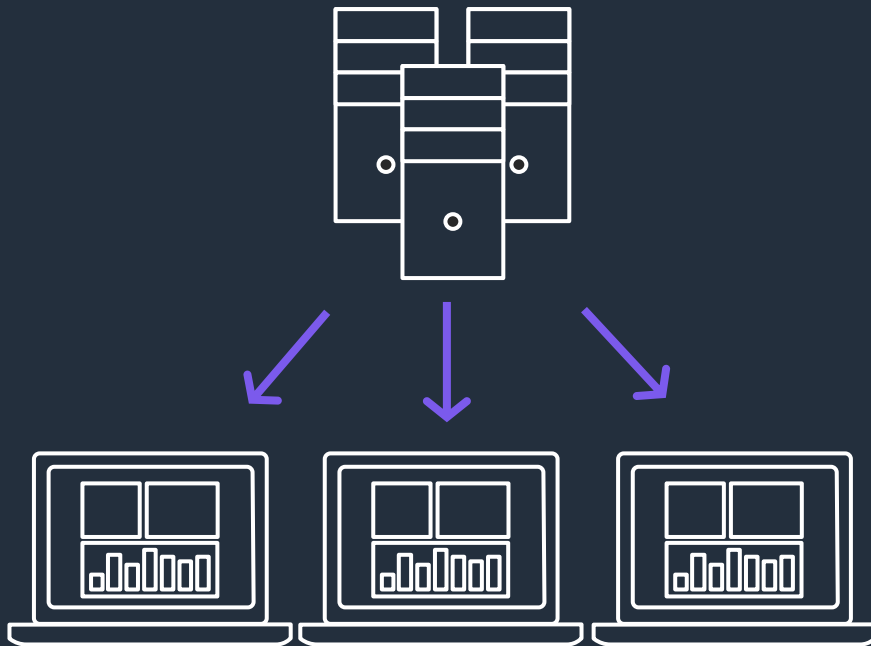
Cloud Approach:

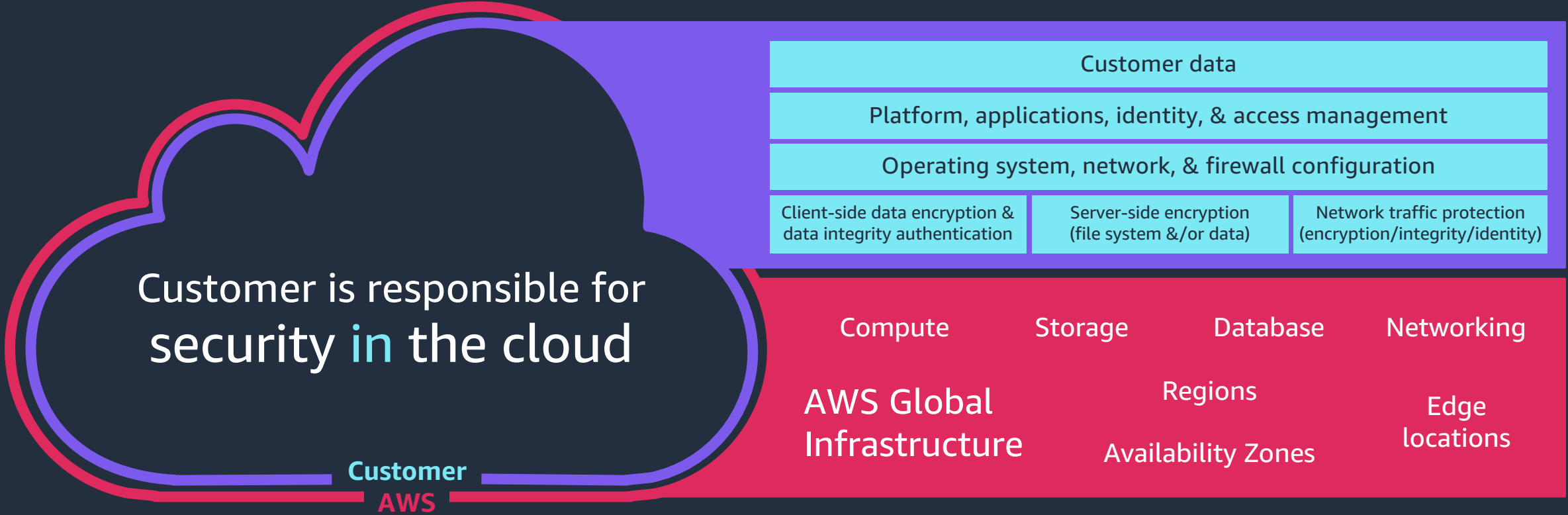Move computing resources to data

Amazon Athena

Amazon EC2

Amazon EMR
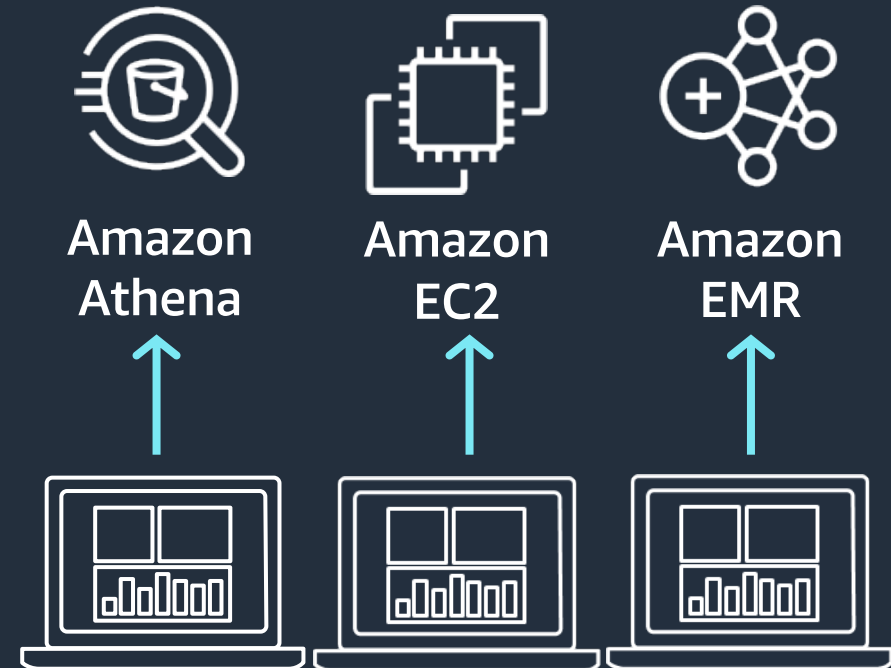
Secure Research

Shared Security Responsibility Model

Customer is responsible for
security **in** the cloud

Customer data

Platform, applications, identity, & access management

Operating system, network, & firewall configuration

Client-side data encryption & data integrity authentication	Server-side encryption (file system &/or data)	Network traffic protection (encryption/integrity/identity)

Compute   Storage   Database   Networking

AWS Global Infrastructure   Regions   Edge locations   Availability Zones

Customer
AWS

AWS is responsible for
security **of** the cloud

Inherit Global Security & Compliance Controls

Certifications/Attestations		
DoD CMMC	🌐 🇺🇸	✓
FedRAMP	🇺🇸	✓
FISMA	🇺🇸	✓
HITRUST	🇺🇸	✓
ISO 9001	🌐	✓
ISO 27001	🌐	✓
ISO 27017	🌐	✓
ISO 27018	🌐	✓

Laws/Regulations/Privacy		
FERPA	🇺🇸	✓
GDPR	🇪🇺	✓
GLBA	🇺🇸	✓
HIPAA	🇺🇸	✓
IRS 1075	🇺🇸	✓
ITAR	🇺🇸	✓
PIPEDA [Canada]	🇨🇦	✓

Alignments/Frameworks		
CIS (Center for Internet Security)	🌐	✓
CJIS (US FBI)	🇺🇸	✓
CSA (Cloud Security Alliance)	🌐	✓
FFIEC	🇺🇸	✓
FIPS	🇺🇸	✓
NIST	🇺🇸	✓
PHR	🇺🇸	✓

🌐 = industry or global standard

AWS Solutions for Research

Introducing Amazon Lightsail for Research

- **For users with limited experience running large computation**

  - Ex: Researchers currently doing all analysis on their laptops and looking to quickly spin up an environment to test ideas

  - No need to configure or manage cloud infrastructure components

- **Access to applications through browser based, simple create wizard**

  - Pick from a fixed set of compute options matched to pre-installed application

  - Run scripts, analyze data, and install additional packages

  - Drag and drop web interface to upload datasets or download results

- **All-inclusive pricing**

  - Includes compute, storage, and data transfer only for duration

  - Data can persist after power down

  - Budgeting tools to manage costs, power down idle research computers, etc.

Amazon Lightsail for Research

12

AWS ParallelCluster

© 2024, Amazon Web Services, Inc. or its affiliates.

13

AWS ParallelCluster

```yaml
 1 ▼   HeadNode:
 2          InstanceType: c5n.large
 3 ▼   Ssh:
 4          KeyName: KP – PDX
 5 ▼   LocalStorage:
 6 ▼       RootVolume:
 7              VolumeType: gp3
 8 ▼   Networking:
 9          SubnetId: subnet – 04e5 d7521c24f6a7e
10 ▼   Scheduling:
11          Scheduler: slurm
12 ▼   SlurmQueues:
13 ▼       –Name: on–demand – queue
14      AllocationStrategy: lowest – price
15 ▼   ComputeResources:
16 ▼       –Name: on – demand – queue – cr – 0
17 ▼   Instances:
18 ▼       –InstanceType: c5n.large
19      MinCount: 1
20      MaxCount: 40
```

SageMaker

Features

Features	
What	AWS service
Use cases	Development of machine learning (ML) models
Configuration	Configurable
Setup persona	Research IT / Partner / ProServe
Setup skills	Advanced
User skills	Advanced
Compliance	Depends on Setup

Pricing & Billing

Pricing & Billing	
Egress traffic	Per service usage
Pricing	Metered service
Billing	As per services used

Links

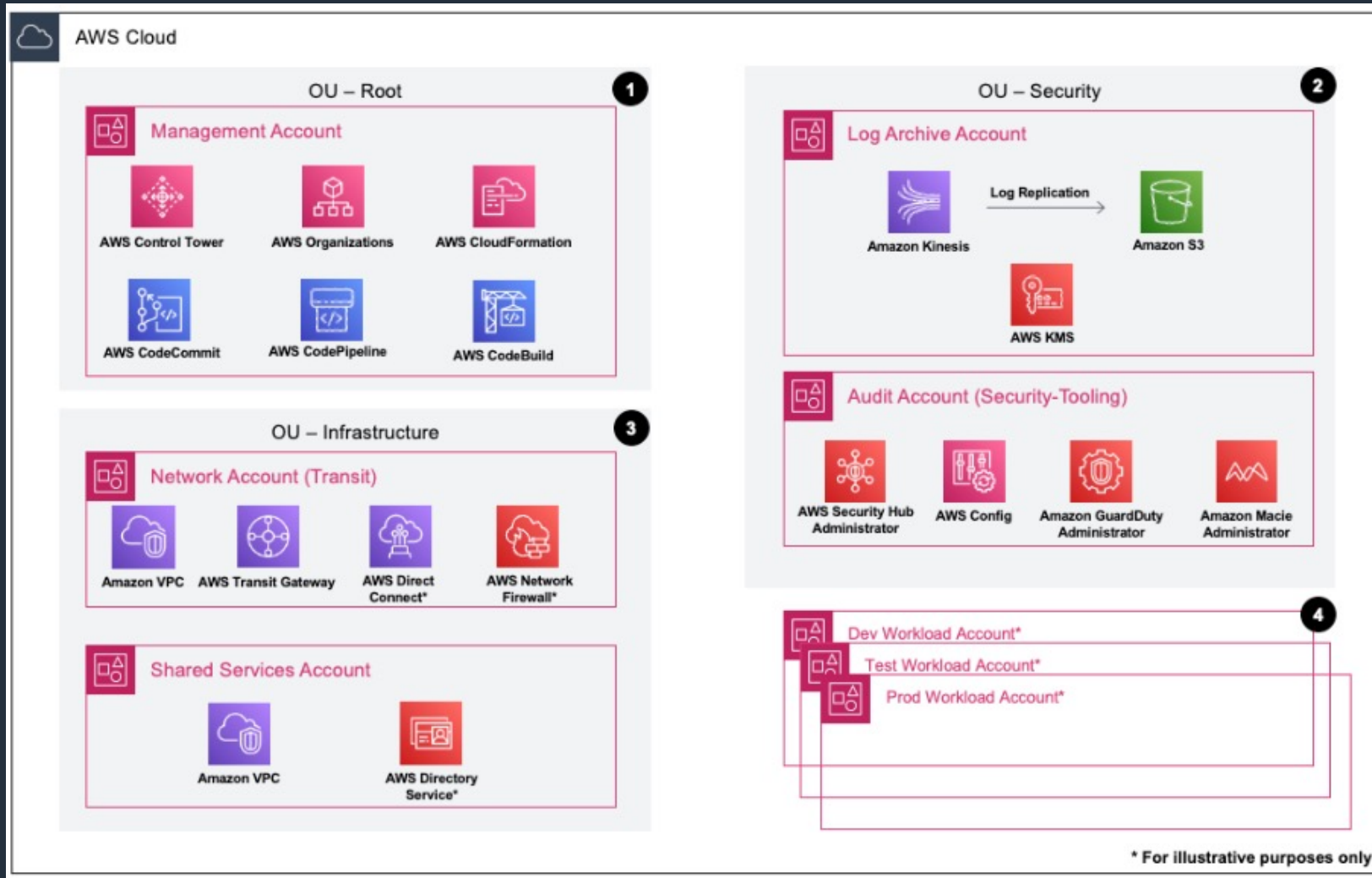| Links |
|---|
| [Public service page](#) |

Solution Data

Solution Data	
Description	Build, train, and deploy ML models for any use case with fully managed infrastructure, tools, and workflows
Key points	• Access, label, and process large amounts of structured data (tabular data) and unstructured data (photo, video, geospatial, and audio) for ML • Reduce training time from hours to minutes with optimized infrastructure. Boost team productivity up to 10 times with purpose-built tools • Enable more people to innovate with ML through a choice of tools—IDEs for data scientists and no-code interface for business analysts • Automate and standardize MLOps practices and governance across your organization to support transparency and auditability
Category	Research tools

Other Research Resources
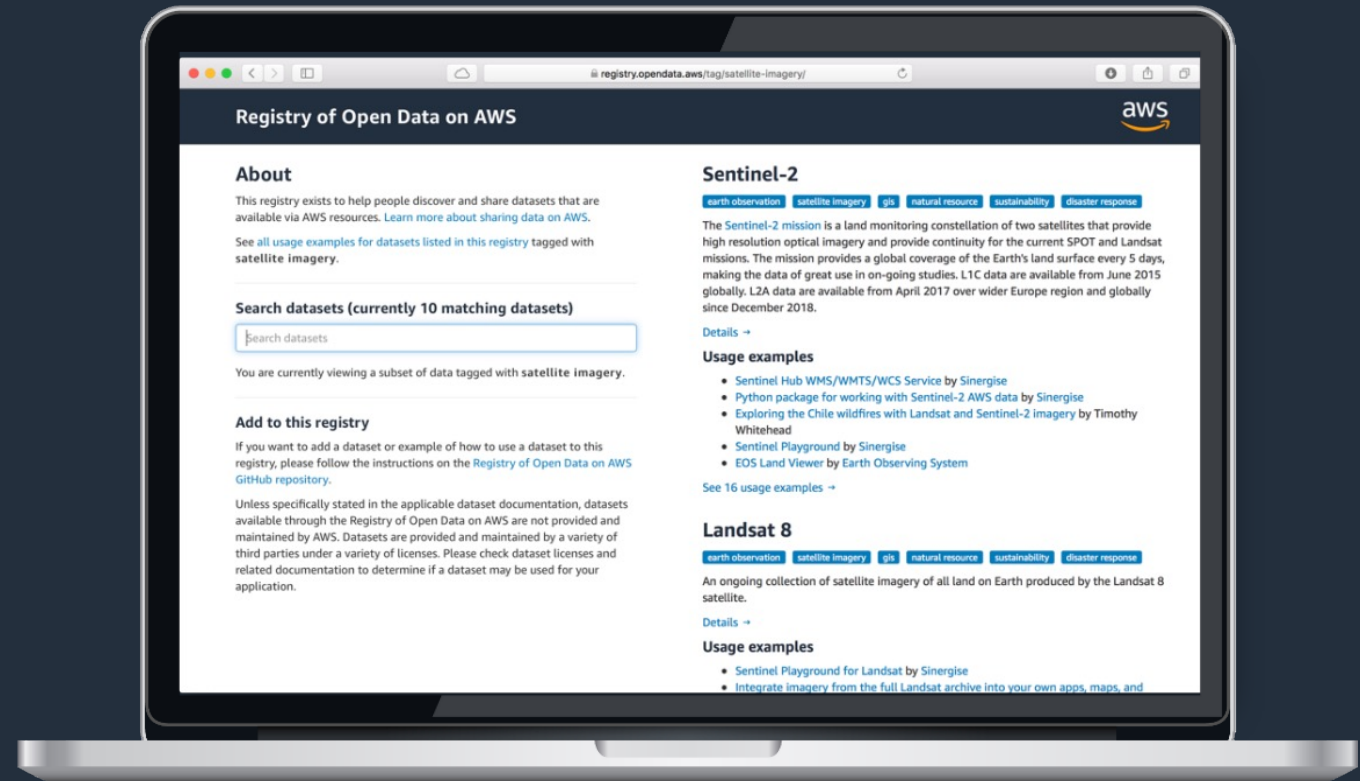
Industry Landing Zones

Open Data on AWS

AWS HOSTS PUBLIC DATASETS TO LOWER THE COST AND IMPROVE THE SPEED OF RESEARCH.

Examples include:

- 1000 Genomes Project
- The Cancer Genome Atlas
- International Cancer Genome Consortium
- Landsat 8
- Common Crawl
- SpaceNet
- OpenStreetMaps

Global Data Egress Waiver

Why | Researchers need predictable budgets

Who | Available to degree-granting and research institutions
Must use a research network (e.g. Internet2) or AWS Direct Connect

What | Waives charges for data downloads
Waives up to 15% of the customer total account bill
Data uploads are no-charge

How | Contact your AWS account manager
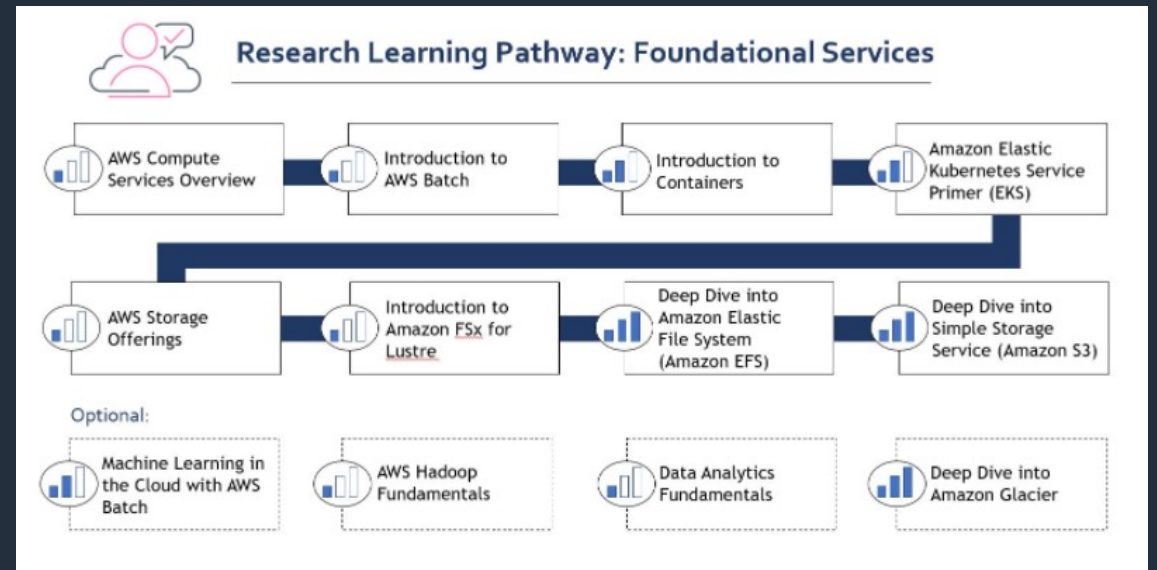Or via research portals (e.g. Arcus, GEANT)
Or via certain resellers (e.g. DLT/I2 NET+, VMWare on AWS)

No-cost Research Learning Pathway

Provides a structured, curated list of online courses tailored to researchers' needs

AWS Skill Builder provides 500+ no-cost digital courses

Q&A