AWS for Genomics

Empowering genomic innovations at the intersection of technology and biology

Today's genomics landscape is rapidly evolving with the accelerated adoption of genomics by biopharma organizations, infectious disease tracing programs, and healthcare systems. Organizations that are addressing new market opportunities and embracing innovative genomics applications and technologies are looking to the cloud to stay agile, innovative, and economical.

For almost a decade, Amazon Web Services (AWS) has helped genomics organizations such as Ancestry, DNAnexus, GRAIL, Genomics England, and Illumina optimize their businesses and build scalable infrastructure to accelerate scientific discoveries.

Genomics organizations of all sizes and disciplines choose AWS to meet their unique business needs, which span from using AWS machine learning services to bridge limitations in NGS infectious disease applications, to leveraging AWS global infrastructure to seamlessly deploy globally in a secure and compliant manner.

Featured AWS customers

ancestry

Genomics england

GRAIL

illumina

mission bio

AWS for Genomics

AWS for Genomics provides secure purpose-built solutions, tools, and partner offerings across the genomics workflow, aligned to key solution area and industry use cases.

Reduce time to discovery

AstraZeneca leverages AWS to run over 51 billion statistical genomics tests in <24 hours.

Optimize cost

Illumina reduced costs by ~$400k/month on AWS.

Scale seamlessly

With AWS, Igenomix realized a 600% increase in sample capacity.

Achieve compliance globally

- HITRUST
- GDPR
- FedRAMP
- HIPAA
- ISO 27001
- ISO 3425

Purpose-built for the genomics industry

Hailing from research labs, post-doc programs, academia, genomic ISVs, government health organizations, academic hospitals, and innovative health startups, the AWS for Genomics team overlays industry tools, best practices, and partner solutions with AWS services to help organizations implement innovative technologies.

Reduce time to discovery

Genomics is a data-heavy discipline requiring extensive compute resources to analyze samples and extract meaningful insights.

To accelerate discoveries, AWS offers a robust suite of powerful compute and machine learning options that enable scientists to process more samples, run more complex analyses, and query at-scale. The only cloud provider to deliver 100 Gbps of networking throughput, AWS delivers high-throughput data ingestion, analysis, and interpretation services and tools designed to help genomics organizations get more from their data.

Genomics organizations such as Fauna Bio leverage the robust computation power of AWS to analyze multi-omic datasets, accelerate research, and uncover new discoveries.

Accelerate data analysis

With native integration to workflow tools, including Nextflow and Cromwell, organizations can orchestra AWS Batch processes to accelerate processing time for computational analysis. For example, with AWS Fred Hutch was able to reduce compute time from seven years down to seven days, accelerating the organizations research on developing therapeutics to fight cancer.

Simplify data interpretation

The full value of genomic data is recognized once its put into context. Population genomics programs across the globe leverage the security, flexibility, and scalability of AWS to host population-scale biobanks and provide democratized access to the industry. Using querying and machine learning services from AWS, scientists can rapidly query datasets, including the Cancer Genome Atlas (TCGA) and the Broad Institute's Genome Aggregation Database (gnomAD) hosted on AWS registry of open data, to rapidly extract insights and answers.

"In just 11 days, we were able to identify 130,000 novel RNA viruses on AWS – that's 10 times more than were identified over the past century of virology research"

UBC THE UNIVERSITY OF BRITISH COLUMBIA

"By hosting on AWS, we can democratize access to our data. All researchers need is a small budget to fund compute costs and access to a computer."

Genomics england

Optimize cost

While sequencing costs have fallen, costs associated with compute and storage have grown as organizations continues to increase throughput and explore more data-heavy applications, such as single-cell genomics. In contrast to on-premises setups that require large upfront investments and continuous CapEx, AWS enables genomics organization to use only what they need, when they need it.

AWS offers storage and compute offerings at a variety of price points. Services such as Amazon EC2 Spot Instances offer up to 90 precent discounts in comparison to on-demand compute prices. For long-term storage of infrequently used data, Amazon S3 Glacier provides secure data archiving starting at $1 per terabyte per month.

Genomics organizations also leverage AWS to reduce development and operational costs. By building its AI-based genomics intelligence platform on AWS, Emedgene reduced its costs of applying artificial intelligence to big genomics by 70 percent while accelerating model development and optimization.

Dedicate 20% more IT resources to software development

CodeX 科德施 GENETICS

70% cost reduction for AI implementation using AWS

emedgene

Scale seamlessly

Many factors impact spikes and lulls in genomics workloads. From academic grant cycles to flares in infectious disease, genomic organizations must be able to respond quickly to succeed. In today's competitive market, companies that plan their on-premises capacity based on maximum anticipated volume carry an enormous financial burden; whereas those that cannot accommodate demand spikes quickly face costly delays.

The elastic nature of AWS allows organizations to launch as many instances as needed and be ready to work in near real-time. Likewise, they can decrease capacity during down times to reduce costs.

Melbourne Genomics Health Alliance built its GenoVic software, a shared clinical system used by Victoria's decentralized health system, on AWS to enable seamless scalability as more laboratories join the alliance and expand their genomic testing services.

"Using Amazon EFS, we don't have to worry about scaling research workloads—the system can grow automatically to meet our researchers' needs, no matter what the compute and storage requirements are."

ancestry

Achieve compliance globally

DNA is the most personal source of data. As genomics becomes an increasingly global practice, organizations must consider both domestic and international regulatory standards.

AWS takes a security first approach to enabling genomics storage and compute, and continues to invest in new availability zones around the globe to enable organization to store and analyze their data locally.

The AWS Global Cloud Infrastructure is the most secure, extensive, and reliable cloud platform, offering over 200 fully featured services from data centers globally.

With 84 Availability Zones, 2X more than any other cloud provider, across 310 Points of Presence, genomics organizations can take full advantage of cloud and maintain data sovereignty.

To help genomics organization adhere to industry compliance and regulations, AWS maintains the following compliance certifications: HITRUST, GDPR compliance, FedRAMP, HIPAA, ISO 27001, ISO 3425.

HITRUST | FedRAMP | HIPAA

GDPR | ISO 27001 | ISO

To help its customers scale their science, DNAnexus uses AWS to reliably and securely analyze more than 1,000,000 samples per year.

DNAnexus

"Some of our customers have sensitive data that cannot leave their region due to local regulatory requirements. We can help them store that data in their country or region by using AWS."

FABRIC GENOMICS

"Data stays in our clients' environments, and all of the AWS safety features keep it secure."

Lifebit reengineered the traditional model for securing data — bringing its compute engine and analytics to the data itself – enabling customers like Genomics England to securely democratize access to its data on the AWS Cloud.

Access to industry tools and datasets

Technical and business experts from AWS collaborate with genomics customers to address common operational and business challenges, such as providing best practices for deploying analysis workflows at scale and integrating with open-source projects such as Cromwell or Nextflow. AWS works commercially with ISVs, such as Illumina's DRAGEN™ Bio-IT Platform and Sentieon, to make their applications available through native and platform services. Finally, the AWS Open Data team works with public data providers such as gnomAD and the NIH to make their datasets available to all AWS customers at the click of a button.

A shared approach to security

The AWS' shared responsibility model helps you manage data privacy, reliability, and security both within the cloud and of the cloud itself. Work with a team of industry specialists to understand and improve your cloud security posture at every stage of the genomics lifecycle—from data creation, collection and processing, to storage.

Customer

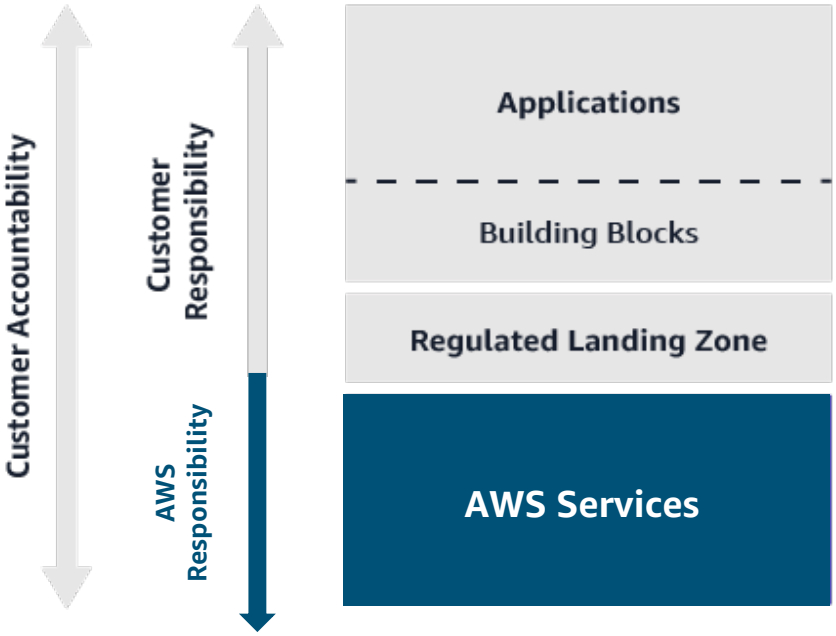
Responsibility for security **"IN" the cloud**

Genomics organizations control access to and management of their data, includes data access permissioning.

AWS

Responsible for security **"OF" the cloud**

AWS is responsible for protecting the infrastructure that runs all of the services offered in the AWS cloud.



illumina®

"At the very foundation, we can count on the AWS Shared Responsibility Model to ensure that our underlying cloud infrastructure maintains enterprise-level security and compliance."

Your cloud migration journey starts here.

AWS' team of genomics experts and AWS Professional Services specialize in helping life sciences organization migrate to the cloud and developing solutions that meet their unique business needs.

Visit AWS.com/Genomics to learn more.

About AWS for Genomics

AWS for Genomics matches the needs of the genomics industry with innovative technologies to provide scalable, secure, and cost-effective tools that accelerate genomic discoveries. From solutions to migrate and securely store genomic data in the AWS cloud, to tools to accelerate secondary and tertiary analysis, to services to integrate genomic data into multi-modal datasets, AWS for Genomics offers a curated portfolio of validated AWS and AWS Partner solutions across the genomics workflow.

To learn more, visit: aws.amazon.com/genomics