



# Data Mesh on AWS for Healthcare & Life Sciences

## Customer Challenges:

Healthcare and Life Science organizations want to make their data and analytic tools more easily accessible across teams and departments to encourage collaboration while ensuring data is secured for only authorized users.

These organizations also need to apply analytics, artificial intelligence, and machine learning (AI/ML) across disparate datasets to gain new scientific and business insights to accelerate novel discoveries and ultimately, improve patient outcomes. While effective collaboration is key, identifying and securely sharing relevant scientific and health data can still be a challenge.

## AWS Health for Data:

AWS helps to simplify collaboration by making it easier for customers to search, share, and discover data at scale across organizational boundaries. We provide use-case driven guidance and data reference architectures to help organizations choose technologies and partners that position them to scale while getting the most value out of their data. For example, customers can use Amazon HealthLake and Amazon Omics to help store and catalog FHIR and genomics data. Once the data is prepared, customers can soon use Amazon DataZone to unlock data across their organization securely and open access to analytic tools such as Amazon Athena and Amazon Redshift with a simple to use portal.

## Top Benefits:



**Create a scalable data foundation**



**Facilitate secure collaboration**



**Generate insights using analytics & ML**

## Featured AWS Services:



### Amazon HealthLake

Securely store, transform, query, and analyze health data in minutes



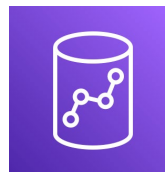
### Amazon Omics

Transform genomic, transcriptomic, and other omics data into insights



### Amazon DataZone

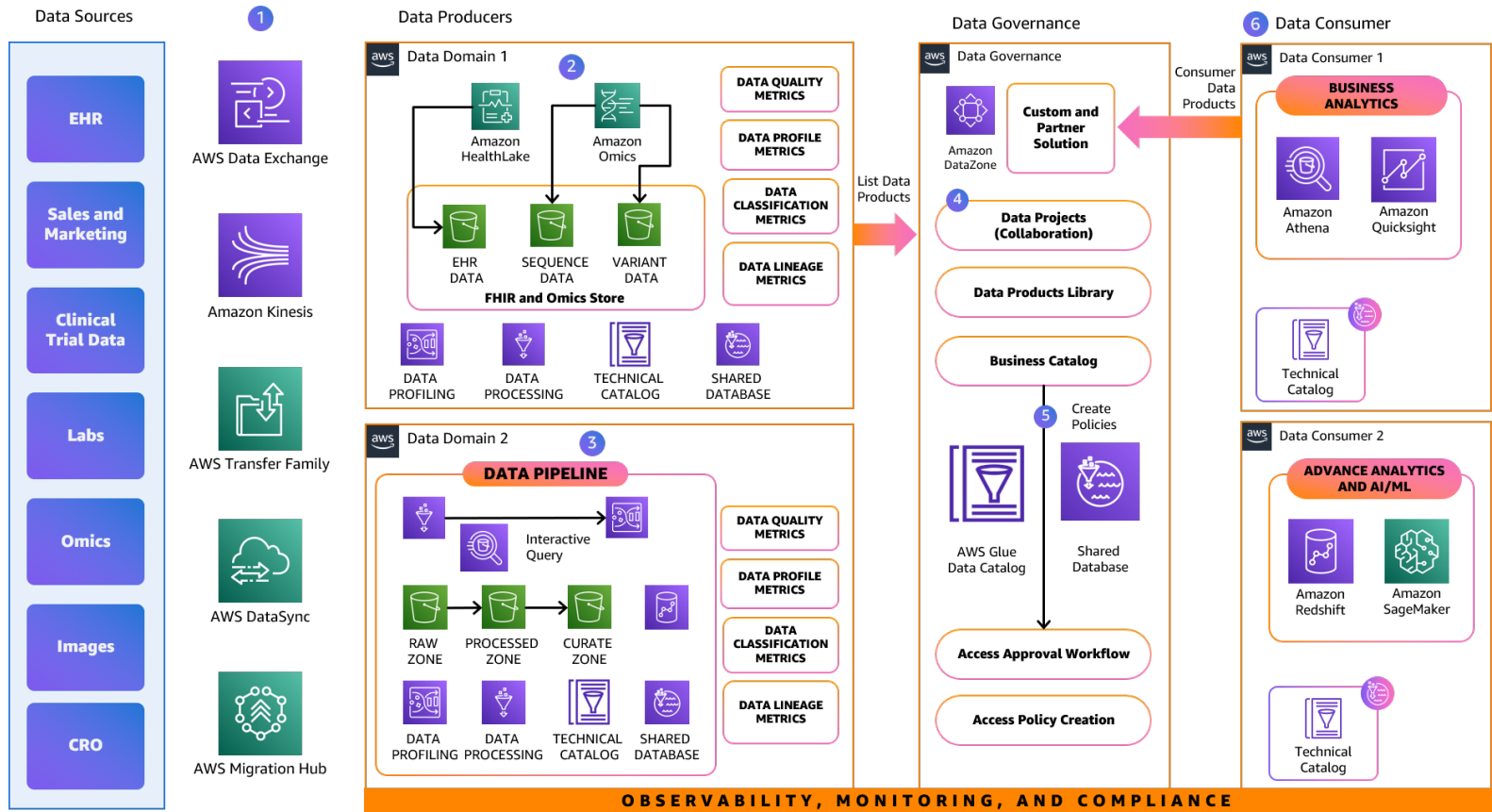
Unlock data across organizational boundaries with built-in governance



### Amazon Redshift

Optimize insights with the best price-performance for cloud data warehousing

# Data Mesh Architecture for Healthcare & Life Sciences:



**1/ Diverse set of data sources.** Ingesting data from external and internal data sources is the first step in creating data products. AWS has a suite of services that can help migrate and subscribe to external and internal datasets. This includes AWS Data Exchange, AWS migration service, and AWS Data Sync.

**2/ Data Producer 1.** EHR clinical data is ingested and stored in an Amazon HealthLake and HealthLake Analytics store. Omics data leverages Amazon Omics sequence and variant store and workflow. Data stores are cataloged and managed by AWS LakeFormation and key metrics are extracted to be listed in the business catalog.

**3/ Data Producer 2.** Other data sources are ingested using AWS Transfer family into raw Amazon S3 buckets. Data is cleaned, curated, transformed, and stored in a purpose built storage that can be published as a data product. Amazon Athena provides interactive query engine for data producers to transform data using SQL.

**4/ Metadata and Publish Data Products.** All into the central business catalog managed by Amazon DataZone, a custom solution or a partner solution.

**5/ Amazon DataZone or partner solutions self-serve workflow.** Accepts consumer request for data products and AWS LakeFormation shared database and tags are utilized for fine grained access controls.

**6/ Data consumers** access data using Amazon Athena and build consumer data products that can be listed in the business catalog as well for discovery and access.