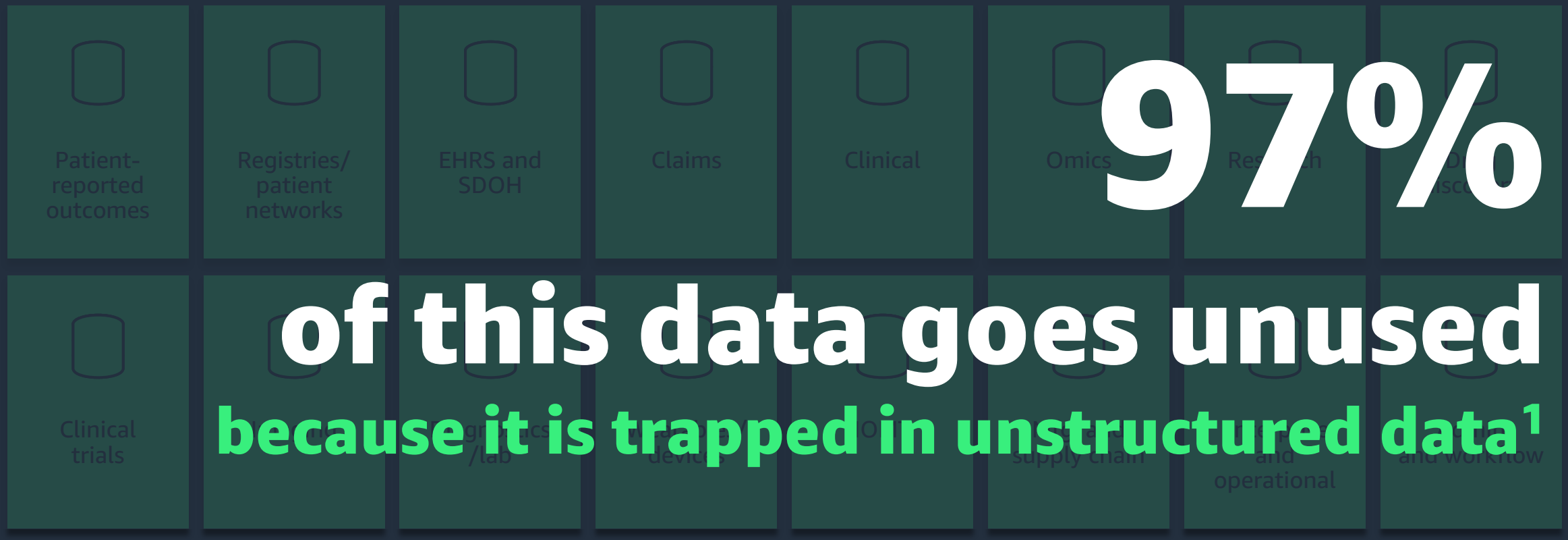
AWS Modern Health Datalake Strategy with Analytics in the Cloud

Bryan Marsh

Principal Healthcare Solutions Architect,
AWS Public Sector Healthcare

50 petabytes of healthcare data generated annually by hospitals



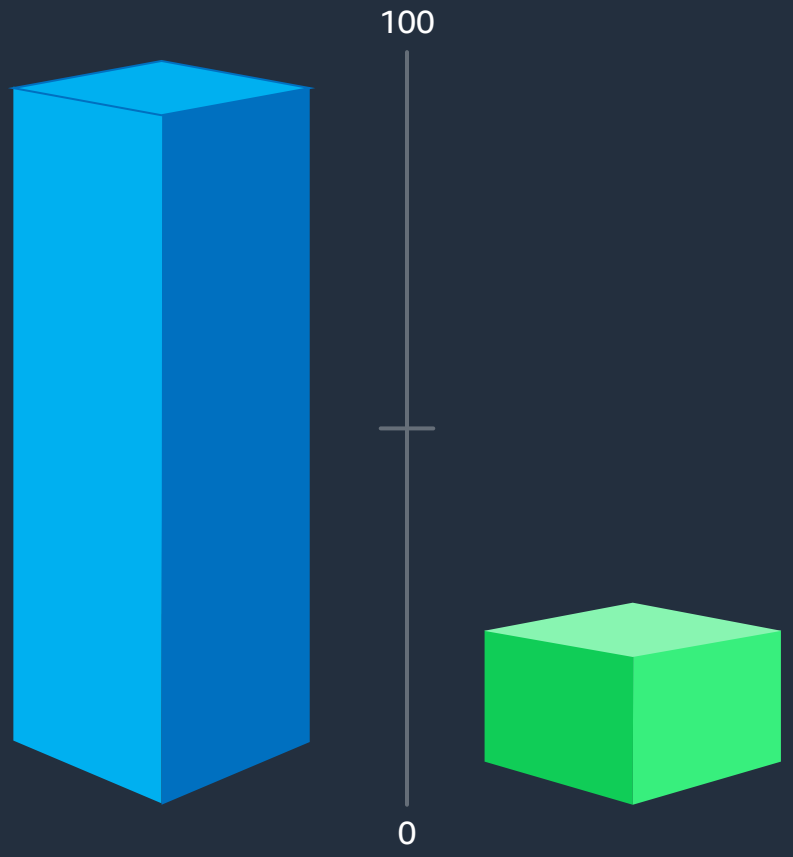
97% of this data goes unused

because it is trapped in unstructured data[1]

Data is a strategic asset for Healthcare organizations

99%

of businesses want to be data driven



26.5%

of businesses have been successful

Our point of view

Healthcare customers who are successful as data-driven organizations:

Treat data as an asset

Collaborate & share data across the enterprise and with third-parties to drive business and patient impact

Provide a federated data governance & compliance model between data producers and data consumers

Operationalize their data strategy to align not just processes and technology, but also their people

With AWS you can extract more value from your data

10+ years with dedicated healthcare and life sciences practice

15+ years as world's first, most comprehensive, and broadly adopted cloud platform

18+ years on average, AWS team leaders have been in the healthcare and life sciences industry

9 of the top 10 pharma companies use AWS for data analytics and ML

"Most mature, enterprise-ready provider, with the strongest track record of customer success and the most useful partner ecosystem"

Gartner® 2021

Build a modern health data foundation on AWS

Comprehensive

The most comprehensive set of services with optimal price performance for any use case

Integrated

Choices for integrating data including zero-ETL so you can easily connect to all your data

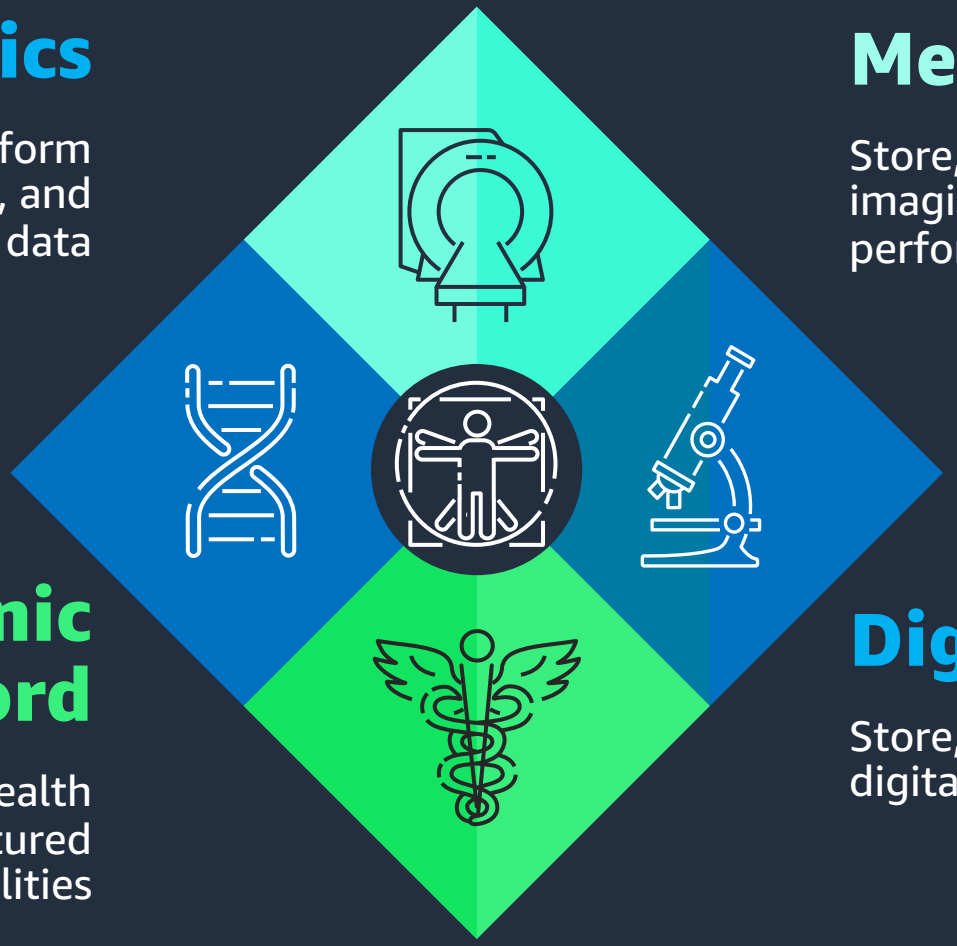
Governed

End-to-end data governance capabilities help you move faster with data

AWS healthcare multimodal point of view

Genomics

Store, analyze and transform genomic, transcriptomic, and other omics data



Medical Imaging

Store, analyze and stream medical imaging while preserving low-latency performance

Electronic Health Record

Store, analyze and align health data across unstructured modalities

Digital Pathology

Store, analyze and share securely digital pathology at petabyte scale

AWS offers a full generative AI and analytics stack

Health AI-powered applications and services

Machine Learning Tools & GenAI models as a service

Health Data Foundation (Databases & Analytics)

Health Landing Zone & Optimized Infrastructure

AWS healthcare purpose-built services

AWS HealthOmics

Transform genomic, transcriptomic, and other omics data into insights

AWS HealthLake

Store, transform, transact, and analyze health data in minutes

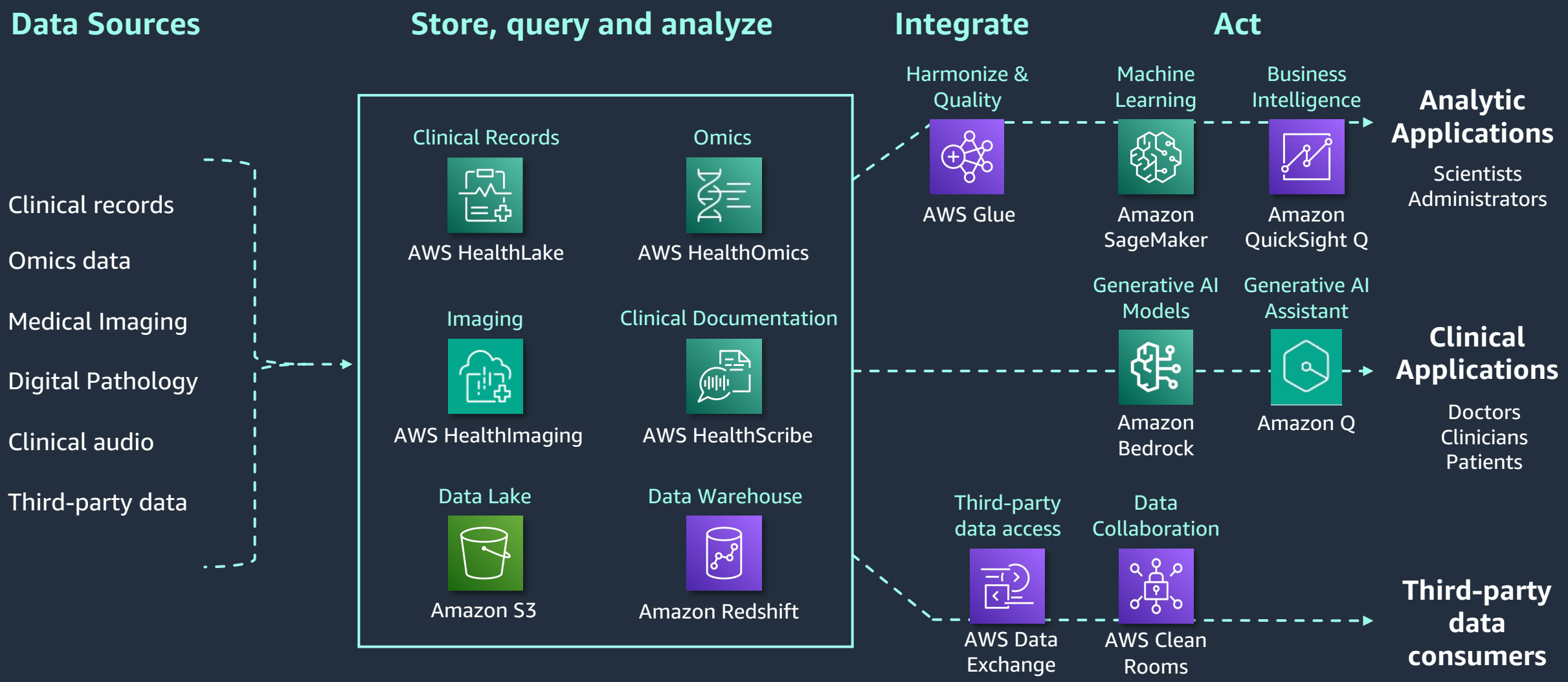
AWS HealthImaging

Store, analyze, and share medical images in the cloud at petabyte scale

AWS HealthScribe

Auto-generate clinical notes from patient-clinician conversations

AWS end-to-end health data in action



Catalog and govern | Amazon Data Zone

Amazon Bedrock supports leading foundation models

AI21 labs	amazon	ANTHROPIC	cohere	Meta	MISTRAL AI_	stability.ai
Contextual answers, summarization, paraphrasing	Text summarization, generation, Q&A, search, image generation	Summarization, complex reasoning, writing, coding	Text generation, search, classification	Q&A and reading comprehension	Text summarization, Q&A, text classification, text completion, code generation	High-quality images and art
Jurassic-2 Ultra	Amazon Titan Text Lite	Claude 3 Opus	Command	Llama 3 8B	Mistral Large	Stable Diffusion XL1.0
Jurassic-2 Mid	Amazon Titan Text Express	Claude 3 Sonnet	Command Light	Llama 3 70B	Mistral 7B	Stable Diffusion XL 0.8
	Amazon Titan Text Embeddings	Claude 3 Haiku	Embed English	Llama 2 13B	Mistral 8x7B	
	Amazon Titan Text Embeddings V2	Claude 2.1	Embed Multilingual	Llama 2 70B		
	Amazon Titan Multimodal Embeddings	Claude 2	Command R+ (Coming Soon)			
	Amazon Titan Image Generator	Claude Instant	Command R (Coming Soon)			

Reinvent how you work with Amazon Q

BUSINESS

Amazon Q Business

KNOWLEDGE SEARCH

SUMMARIZATION

CONTENT CREATION

EXTRACT INSIGHTS

RESEARCH & ANALYSIS

Amazon Q in QuickSight

UNDERSTAND DATA

BUILD & REFINE VISUALS

BUILD CALCULATIONS

EXECUTIVE SUMMARIES

CREATE DATA STORIES

DEVELOPERS

Amazon Q Developer

PLAN APPLICATION

CODE GENERATION

UNIT TESTING

SECURITY SCANNING

CODE REMEDIATION

CODE MIGRATION

TROUBLESHOOTING

DEVELOPER KNOWLEDGE

SPECIALIZED USERS

Amazon Q in Connect

AGENT ASSIST

Amazon Q in AWS Supply Chain

SUPPLY CHAIN

Healthcare innovations using generative AI

Medical Research

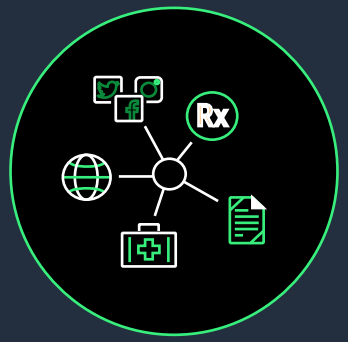
Patient to trial matching

Multi-modal data analysis

Clinical Efficiency

Longitudinal patient records for full patient picture

Automate medical image interpretation



Operational Efficiency

Auto-generate referral letters, clinical coding, and prior authorization

Intelligent document processing

Patient Experience

Patient outcome prediction

Personalize patient discharge instructions and treatment plans

Digital Health

Patient care concierge

Remote care management

Improving patient care with machine learning at Beth Israel Deaconess Medical Center

Challenge:

Beth Israel Deaconess Medical Center (BIDMC) needed a way, within digital health records, to collect, recognize, correctly place documentation in file sections and automate flags of incomplete patient consent, history, and physical forms to decrease errors and processing time for staff.

Solution:

Using AWS machine learning and AI models electronic health records can be shared between facilities, scanned and filed, with flags being automatically sent to staff for follow-up on incomplete forms reducing errors, while alleviating potential surgery delays and staff resourcing bottle necks.

Benefits:

- Created cost efficiencies by reducing staff processing time
- Less re-scheduling and/or delays due to incomplete paperwork

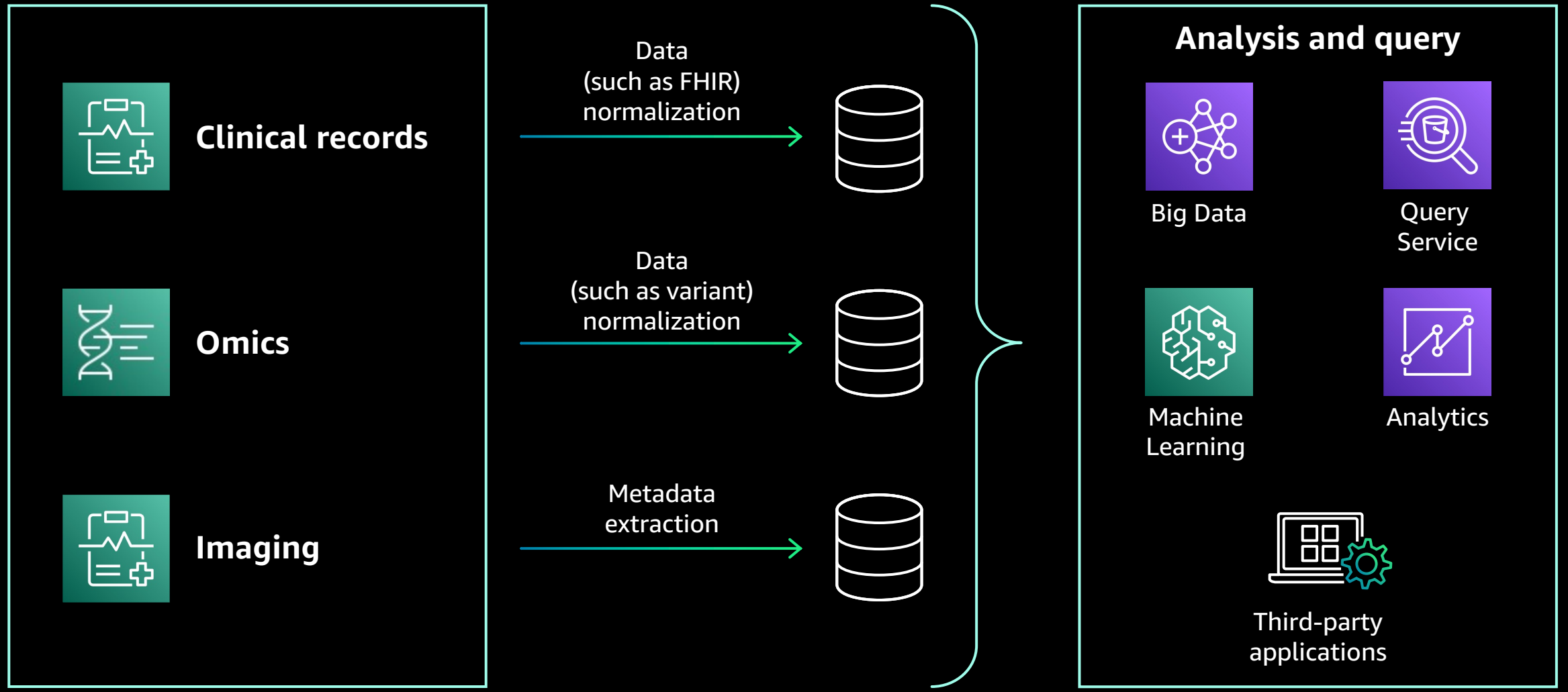
Beth Israel Lahey Health
Beth Israel Deaconess Medical Center

"…using machine learning services like Amazon SageMaker, researchers at BIDMC will build deep learning models that are capable of making highly accurate predictions of where and when space will free up in the hospital for unexpected patients."

Learn more

Demo

Multiomics and multimodal analyses



How does Amazon HealthLake work?

Data sources

Claims
Electronic health records
Wearables + devices
Clinical notes
Real-time or batch

Convert data to FHIR

Amazon HealthLake

Data is structured, tagged, and indexed

Validation
US Core, CARIN BB, ABDM Profiles

SMART on FHIR for access

APIs
Read, search
Bulk export

APIs
Create, update, delete
Bulk import jobs

Integrated Comprehend Medical

Analytics
Athena + LF integration for SQL access on transformed data for analytics

Security, Identity, Compliance OAuth2.0/ OpenID

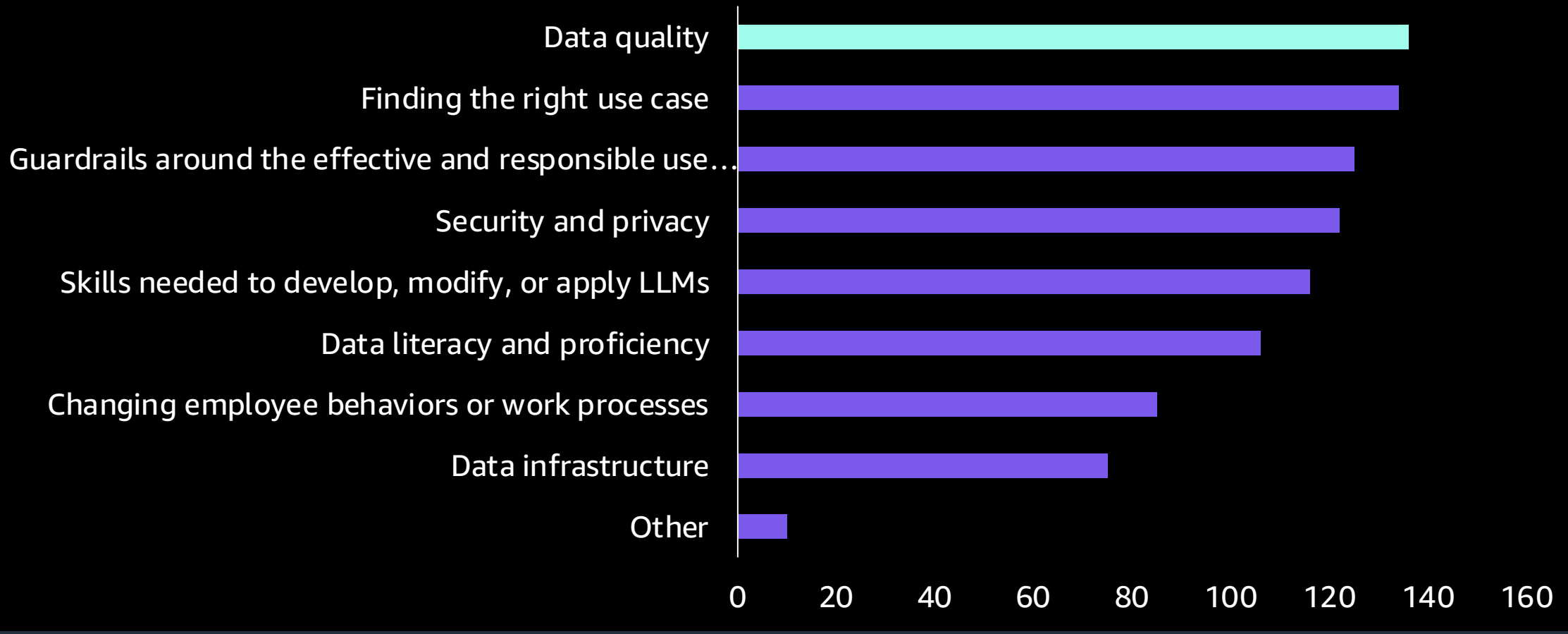
Healthcare apps/portal

Patient access API
Provider directory
Population health & analytics
FHIR Viewer for LMR
Third-party apps

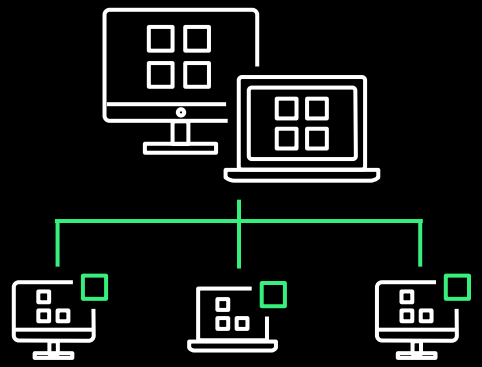
BI dashboard/ML models

Third-party apps

But generative AI and analytics needs quality data



AWS Data Harmonization and Data Quality

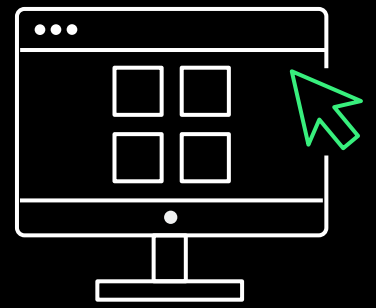


Serverless and Scalability

Serverless infrastructure and tested at scale to manage quality of 60 PB data lakes

Data Quality

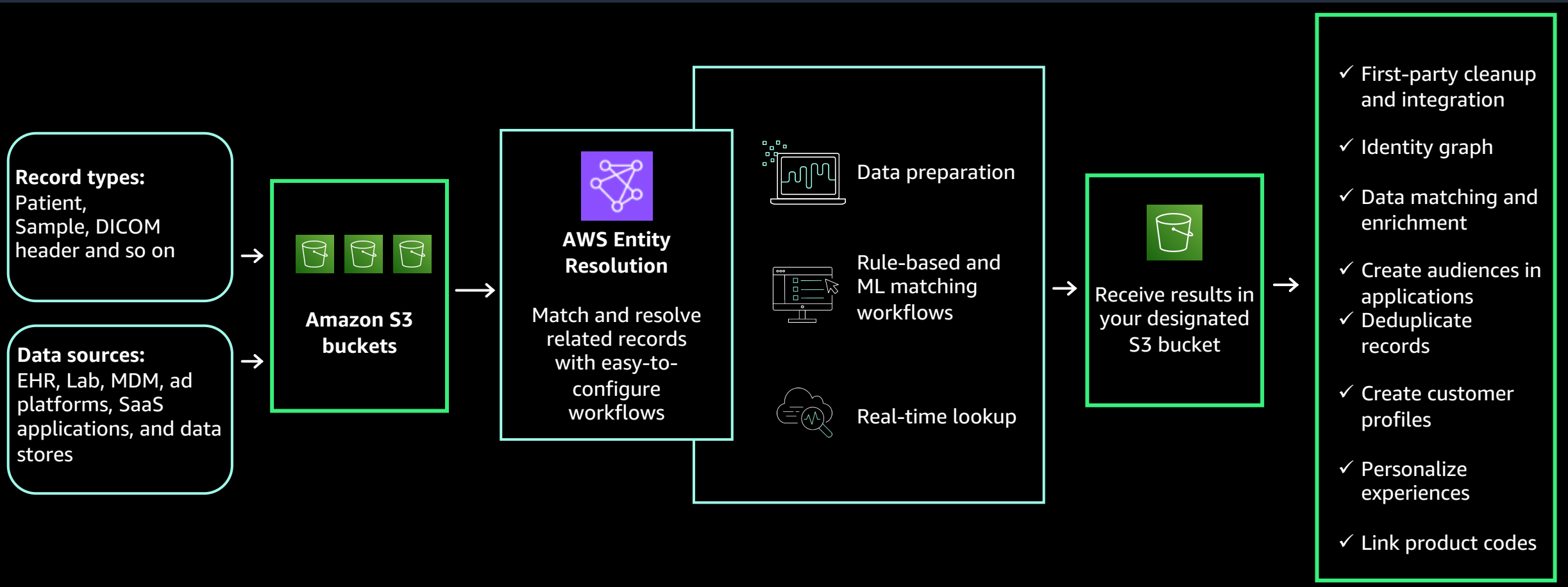
Out-of-the-box data quality rules and actions to take when quality deteriorates



Multiple Persona Support

Data Scientists, Data Engineer, Data Analyst

How does AWS Entity Resolution work?



Rush University System for Health Creates a Population Health Analytics Platform on AWS

Challenge:

Rush University System for Health (RUSH) is committed to addressing the underlying causes of the 16-year life expectancy gap among minority and lower-income residents of Chicago's West Side. RUSH sought to build a comprehensive analytics solution to identify and inform scalable interventions for equitable healthcare based on clinical, cardiometabolic, and social needs.

Solution:

With support from Amazon Web Services (AWS), RUSH developed the Health Equity Care & Analytics Platform (HECAP). This platform transforms, aggregates, and harmonizes data from different sources to reflect the complex interplay of clinical and social factors on patient health.

Benefits:

- Aggregates data from multiple sources using HIPAA-eligible services
- Builds a complete patient profile using clinical, social, and patient-generated data
- Produces risk scores and prediction modeling to guide clinical and community intervention
- Advances health equity for minority and underserved patient populations

RUSH

"We have a great opportunity to start bringing in more data from different sources and use the power of AWS to scale massively across our system, significantly benefiting the care of our patients in Chicago."

Learn more

aws

Demo

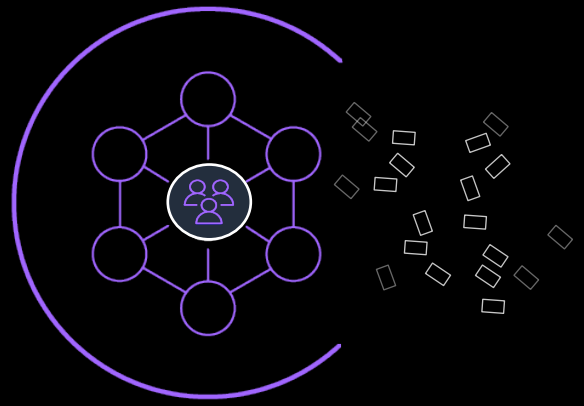
Governance provides guardrails to innovate faster

Find, access, & share
the right data

Keep data
safe & secure

Enable appropriate
audits & controls

How does Amazon DataZone work?



Data producers

Bring data from different sources, and across accounts/regions

Amazon DataZone

Domains

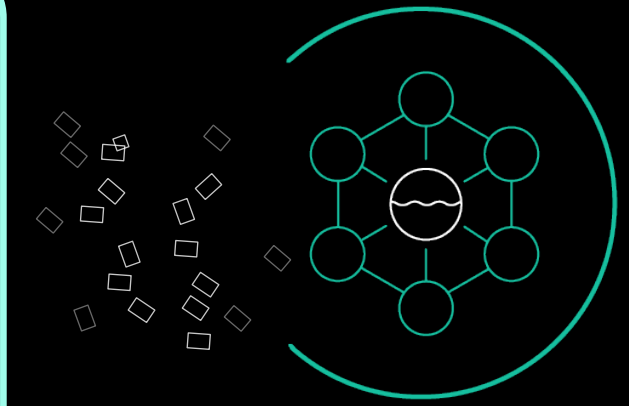
Data portal

APIs

Business data catalog

Projects & environments

Governance and access control

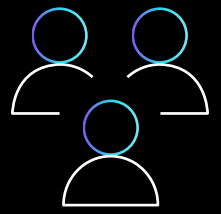


Data consumers

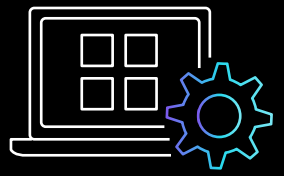
Simplify access, collaboration and consumption using different tools

How does AWS Clean Rooms work?

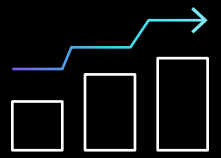
Create your own clean room, add participants, and start collaborating in **just a few steps**



Collaborate with hundreds of thousands of companies **without sharing or revealing underlying data**

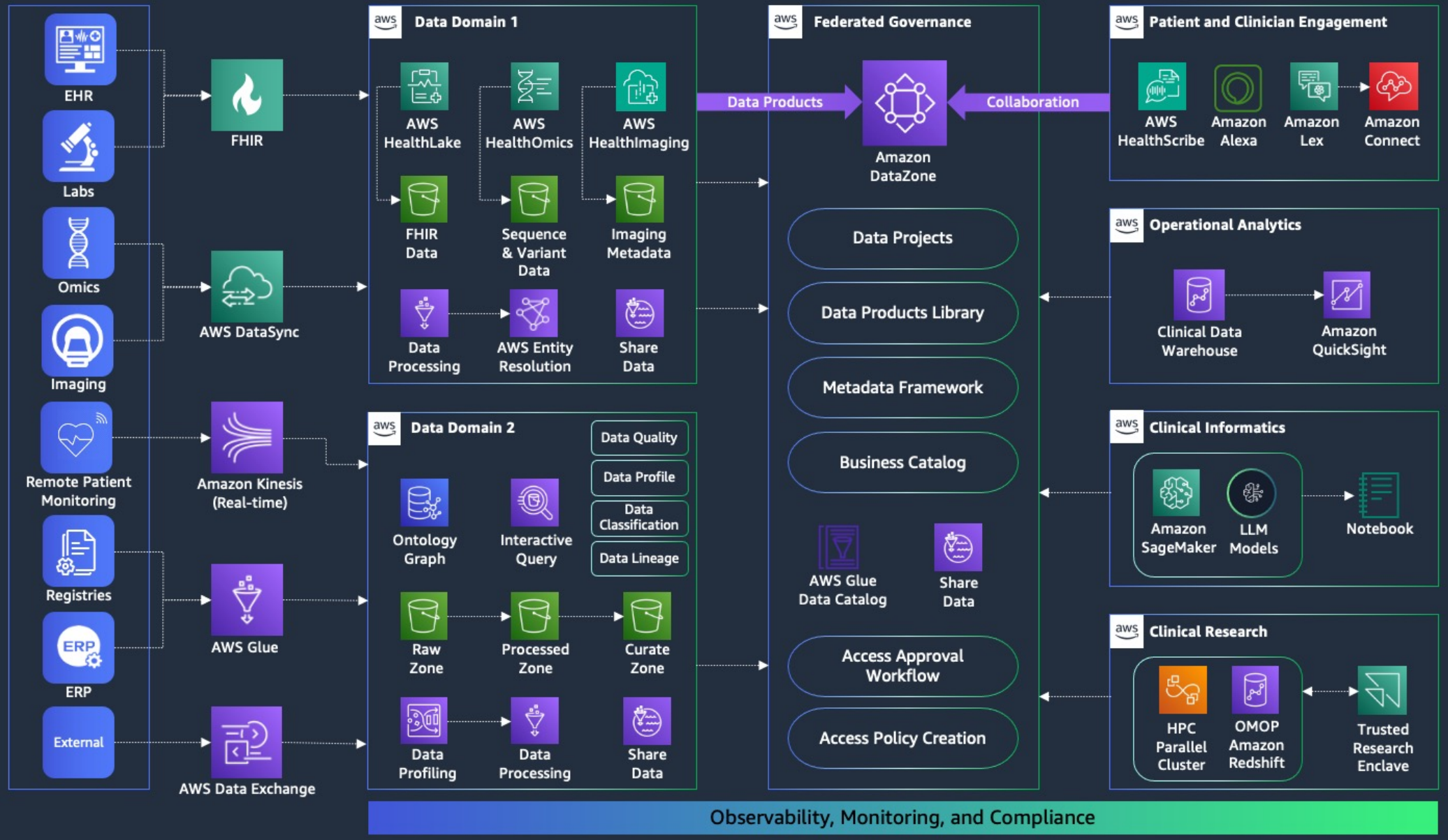


Protect underlying data with a broad set of **privacy-enhancing controls** for clean rooms



Use **flexible SQL** analysis rules and **privacy-enhancing machine learning** (ML) to meet your business needs

AWS Modern Health Data Architecture



AWS Health Data Accelerator – 11 weeks

Foundational Data Lake

Prebuilt data ingestion pipelines for Epic and Cerner to enable data lake house architecture

Data Governance

Data sets ingested are cataloged in Amazon DataZone with governance workflows and fine grained access controls

GenAI Quick Starts

Includes GenAI models and quick starts to jump start insights on your data

CHOP accelerates pediatric research using AWS-powered data resource

Challenge:

The Children's Hospital of Philadelphia (CHOP) wanted to address the challenge of storing and organizing increasing amounts of clinical data so that researchers could access, study, and cross-reference it to facilitate medical breakthroughs.

Solution:

CHOP used AWS to build the Gabriella Miller Kids First Data Resource Center (KFDRC), a transformative healthcare discovery solution that coalesces cross-disciplinary medical research, advancing genomic, clinical, and imaging data availability toward deriving insights for children all over the world and across a wide spectrum of diseases.

Benefits:

- Provides the research community with access to genomic and associated clinical data
- Indexed 1.5 PB of genomic, clinical, and imaging data within 1 year
- Increased KFDRC's collaborative potential

Children's Hospital of Philadelphia℠

"All of our system is currently built on AWS... we went from zero to managing a few petabytes of genomic data within a year using this setup."

Learn more

Demo

Q&A

Please Provide Your Feedback

Thank you!

Bryan Marsh
brymrsh@amazon.com

Step 1: Select Healthcare
Step 2: Select AWS Modern Health Data Strategy

Learning Day Content

https://sanfrancisco2024.awslearningday.com/

Appendix