

AWS re:Invent

NOV. 28 – DEC. 2, 2022 | LAS VEGAS, NV



LFS202

Gilead Sciences' journey from migration to innovation on AWS

Marc Berson

CIO
Gilead Sciences

Murali Vridhachalam

Head of Cloud, Data, and Analytics
Gilead Sciences

Ian Meyers

Director of Product Management, Amazon Analytics
AWS



Agenda

- Introduction – AWS for Health
- The migration to innovation journey
- Exploring a modern data strategy
- Gilead's journey from migration to innovation with AWS
- Deep dive into Gilead's data mesh approach



To make healthcare more accessible
by improving efficiency of care **is my
why**

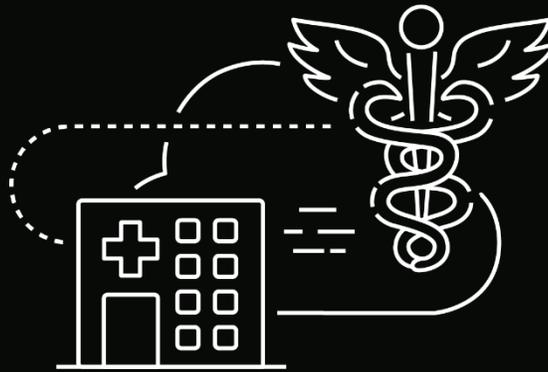
Ian Meyers



AWS for Health



Omics



Healthcare



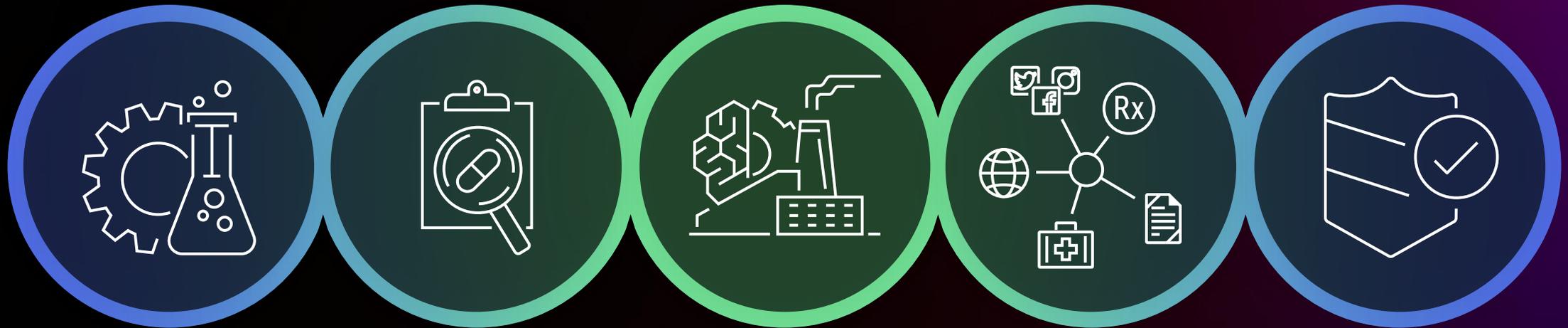
Life sciences

Accelerating innovations from benchtop to bedside

**AWS for Health
provide healthcare
and life sciences
organizations with
the right tools, data,
expertise, and
security to innovate
on behalf of the
patient**



Supporting innovation and operational excellence across the value chain



Research
and discovery

Clinical
development

Manufacturing
and supply chain

Commercial and
medical affairs

Compliance

Migration to innovation journey



Migrate



Secure & comply



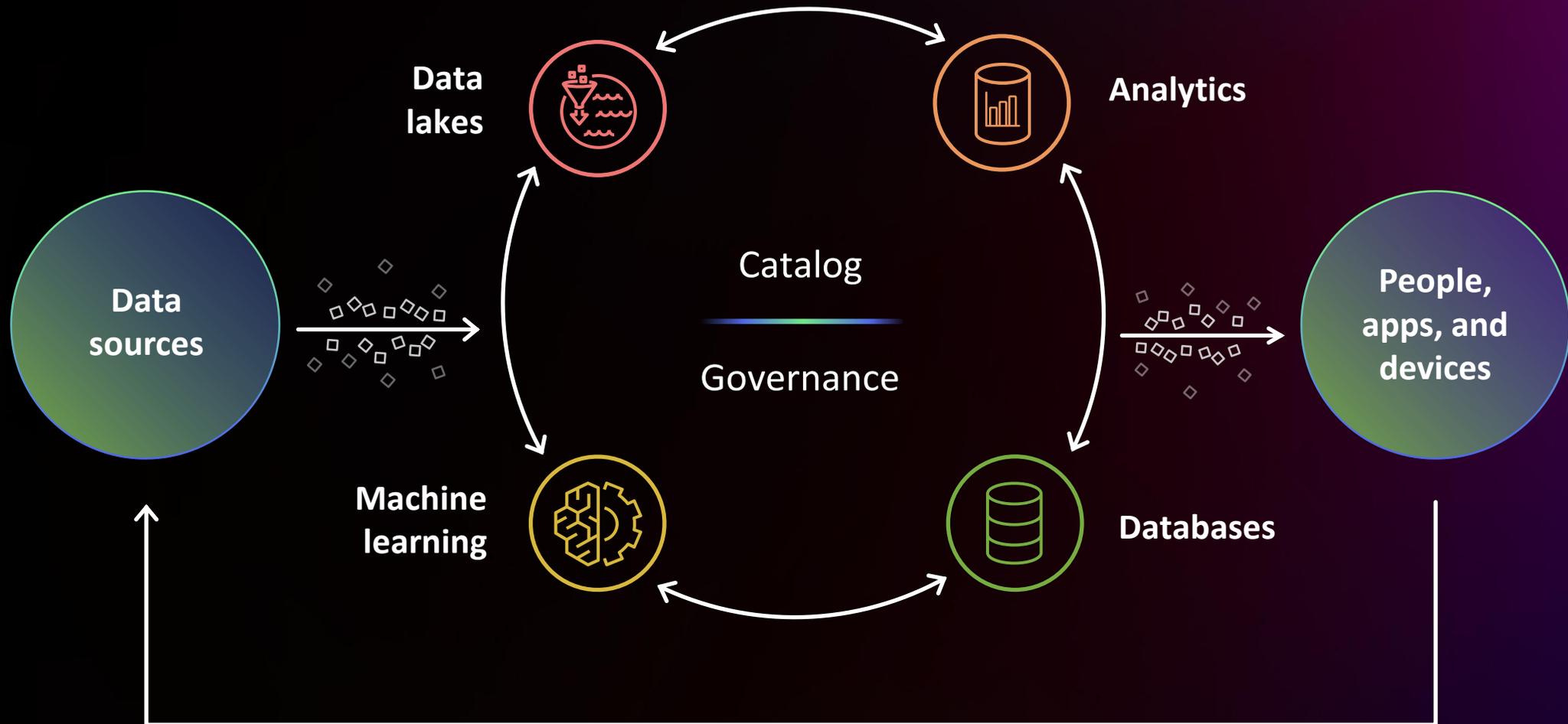
Unify



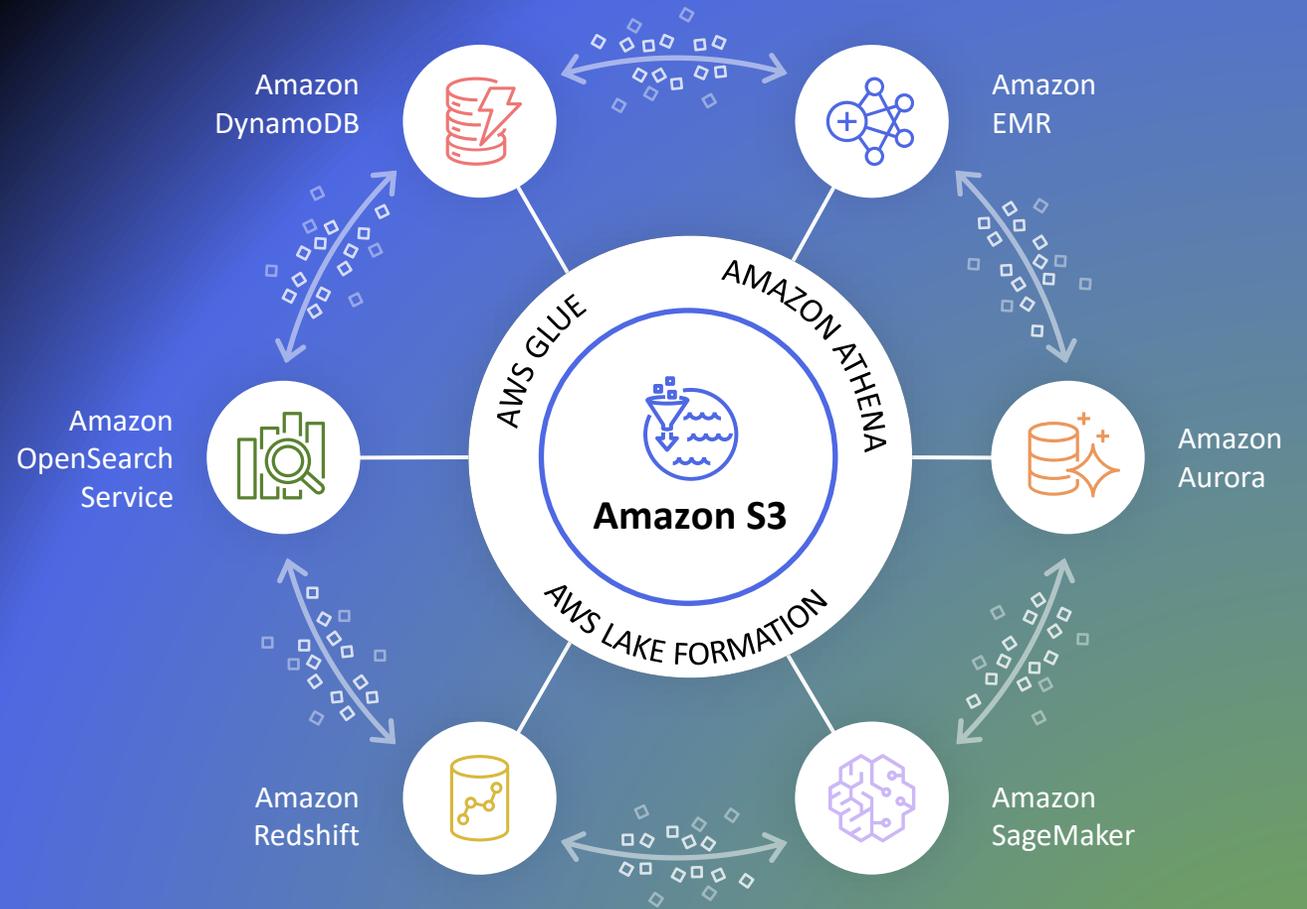
Innovate

Meet customers where they are
Help them get to where they want to be

Modern data strategy



Purpose-built services for a modern data strategy



Key focus areas across all services

Security, durability, availability

Simplicity and ease of use and operations

Price/performance

Data connectivity and integration

Data governance



Advancing cures to life-threatening diseases using cloud innovation and analytics is my why

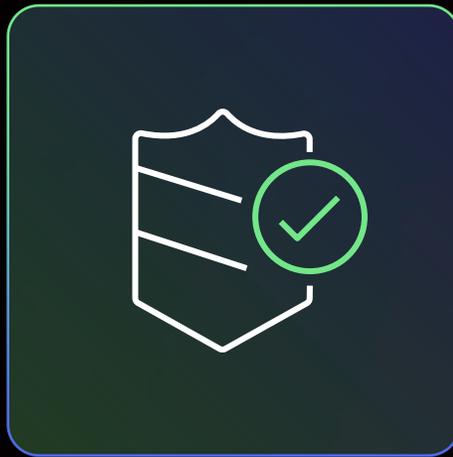
Marc Berson
Chief Information Officer
Gilead Sciences



Migration to innovation journey



Migrate



Secure & comply



Unify



Innovate

Gilead is leapfrogging others by being one of the fastest life sciences companies to move to the cloud

Why did Gilead choose AWS?

AWS technology was important...

... but passion for **transforming the industry**
and **inventing together** was even more
important

Value of cloud innovation

BENEFITING FROM AWS INVESTMENT AND CO-INNOVATION



**Partnership and investments in
co-innovation**

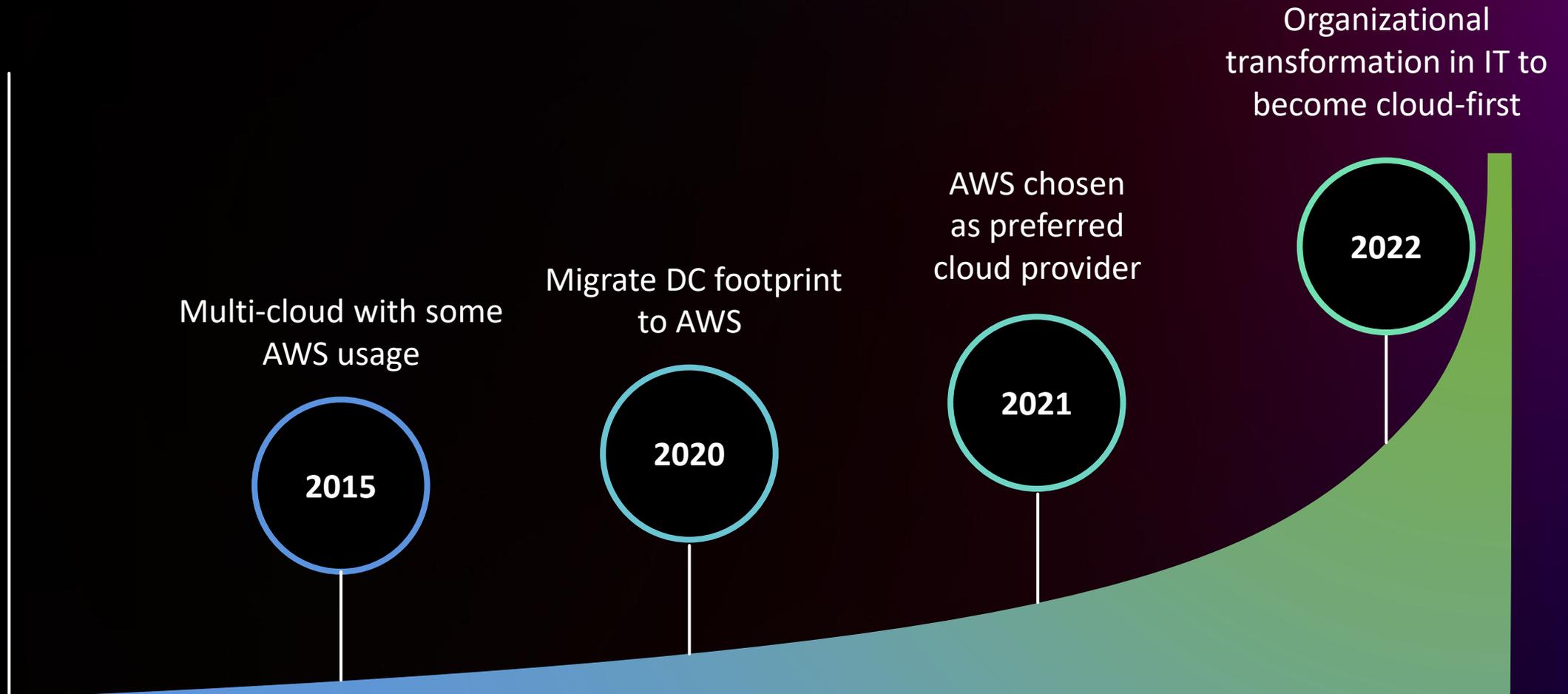


**Industry experts working hand-
in-hand with Gilead**



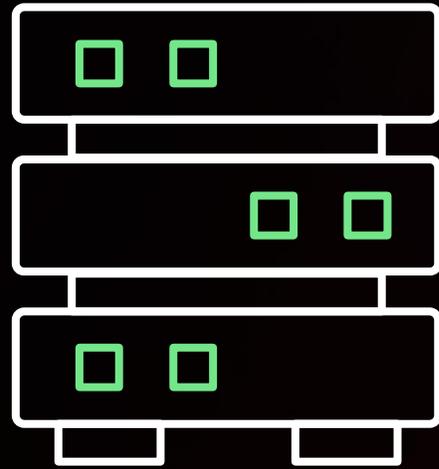
**Developing a data-driven culture
to benefit patients**

Gilead's journey with AWS



Perceived barriers for GxP compliance

GXP ON-PREMISES TRANSLATED TO GXP COMPLIANCE WHEN MOVING TO AWS



Myth debunked

Value realized



Enabled operating model transformation, including increased staff productivity



Increased operational resiliency



Increased agility to deliver new features

Future state vision



Efficiency

Integrate with ITSM and improve staff productivity

Adopt technologies like serverless compute to optimize costs



Democratization

Enabled role-based account provisioning catalog

Integrated data catalog and access management for access to data



Agility

Accelerate and adopt continuous delivery practices

Enable rapid self-service account provisioning



Sustainability

Continue to lower carbon footprint

Reduce data center footprint

Migration to innovation journey



Gilead is leapfrogging others by being one of the fastest life sciences companies to move to the cloud

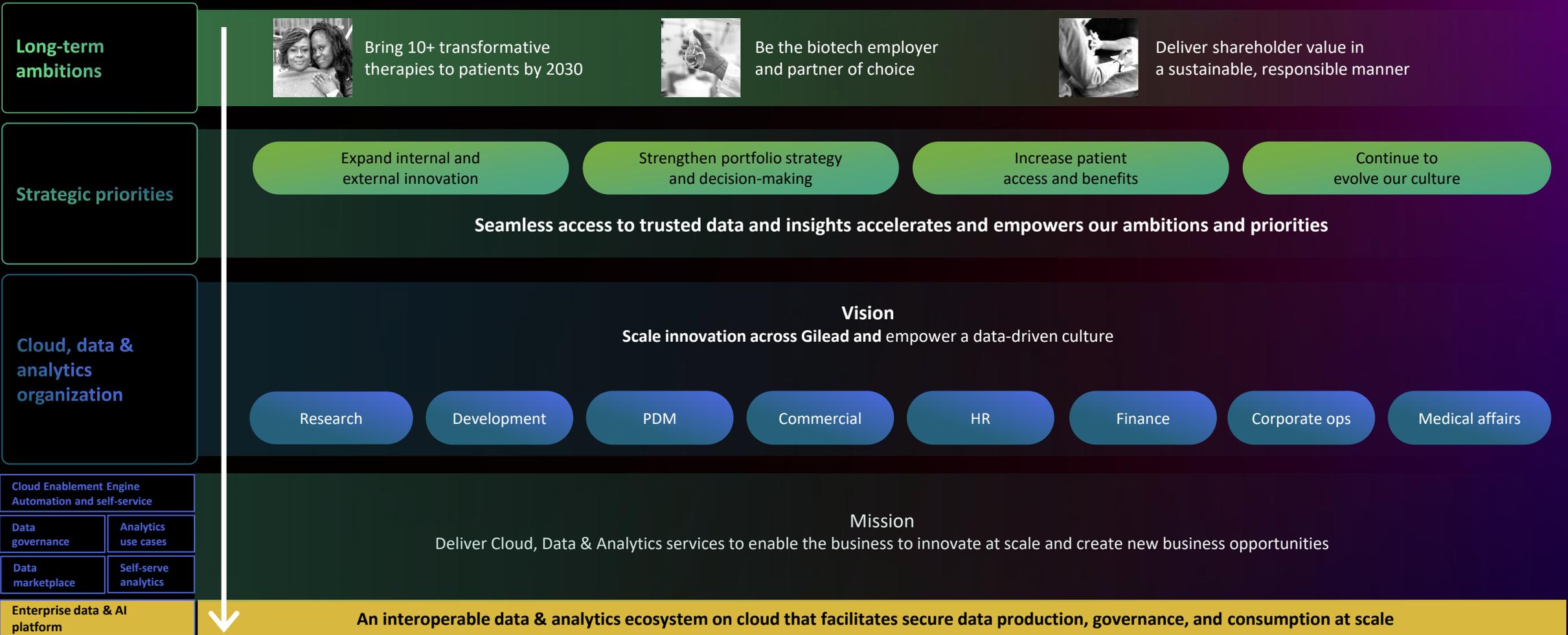


The confluence of cloud, data, and analytics plays a crucial role in accelerating drug discovery and commercialization

Is my why

Murali Vridhachalam
Head of Cloud, Data & Analytics
Gilead Sciences

Gilead's long-term ambitions & strategic priorities



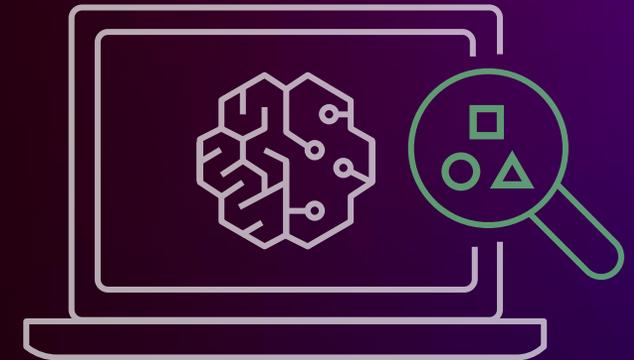
Examples of innovation on AWS



**Enterprise semantic search
(Morpheus)**



Gilead Data Marketplace



**Enterprise Data & AI Platform
(Gilead-DnA)**

Morpheus

Enterprise search application that uses deep learning for semantic search and domain-specific question answering



Increased staff productivity and streamlined data management



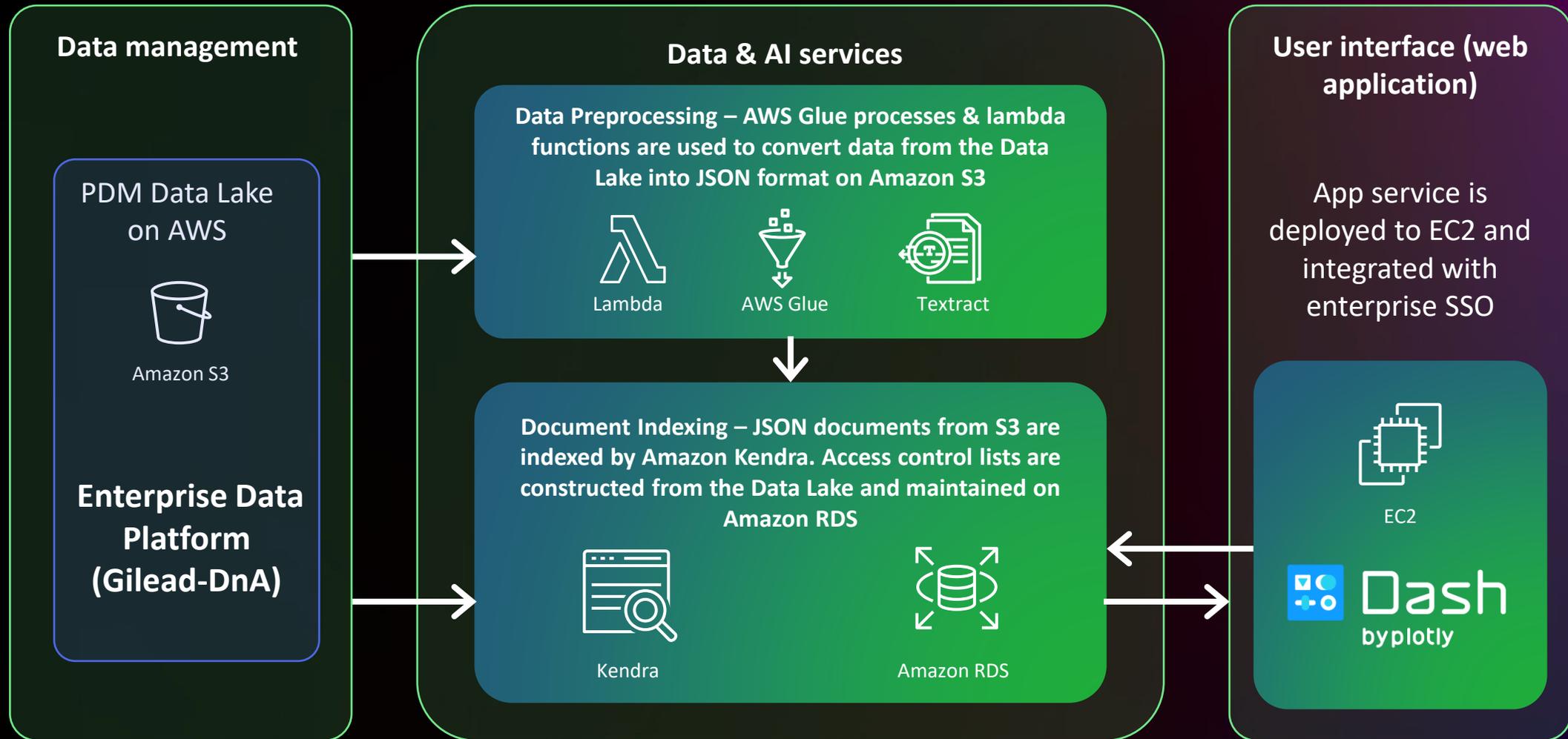
Scalable enterprise search tools that centralize and make it easier to find the right data



Ability to organize and provide results from structured and unstructured data across enterprise systems and knowledge repositories

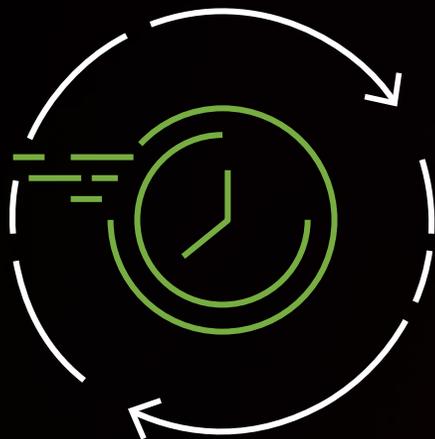


How was it built?



Outcomes

INCREASED SPEED AND ACCURACY

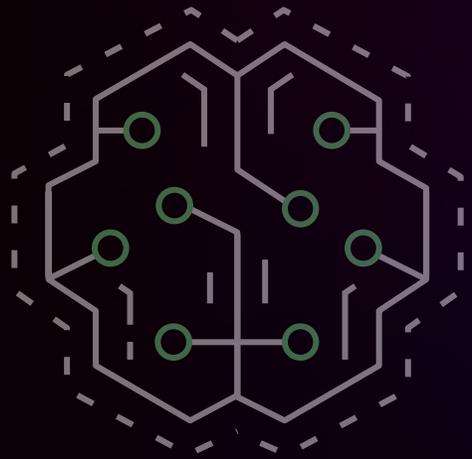


**Reduced search times
by ~50%**



**Increased accuracy of
search results**

A few examples of innovation on AWS



**Enterprise Semantic Search
(Morpheus)**



Gilead Data Marketplace

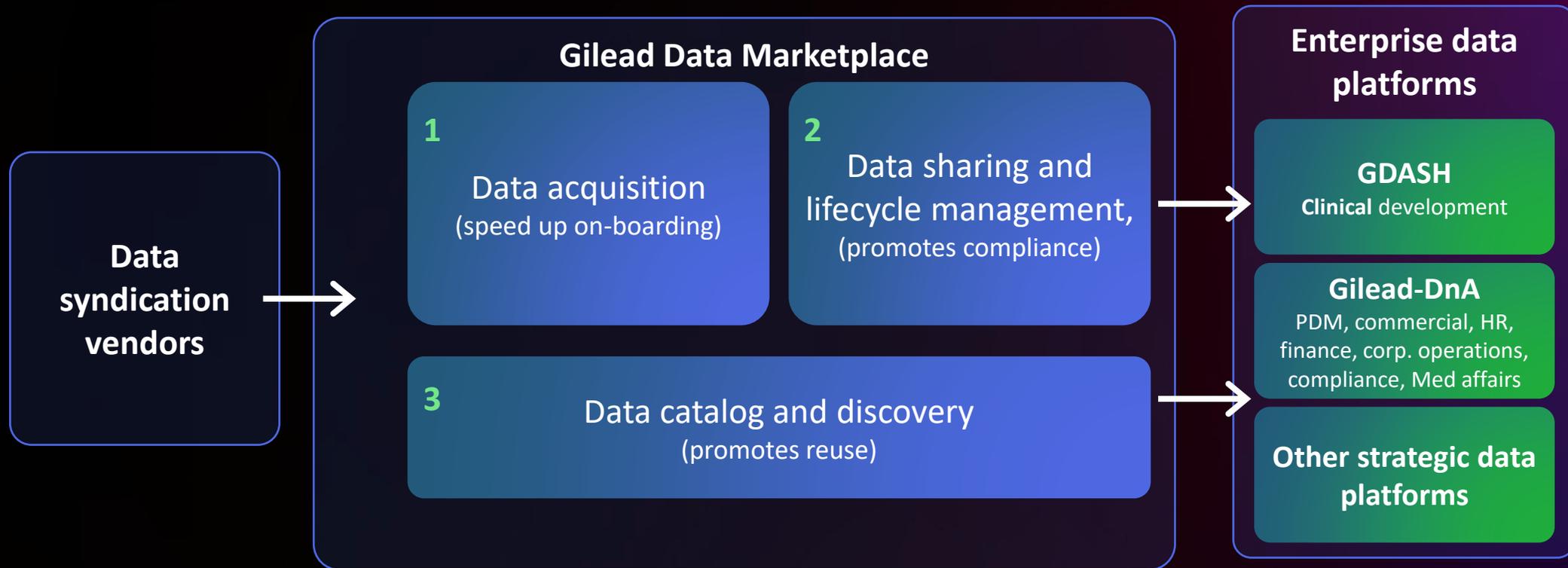


**Enterprise Data & AI Platform
(Gilead-DnA)**

Gilead Data Marketplace

WHAT IS GILEAD DATA MARKETPLACE

A streamlined mechanism to acquire and share data purchases from vendors, that (1) speeds up onboarding, (2) promotes reuse, and (3) improves compliance with licensing terms

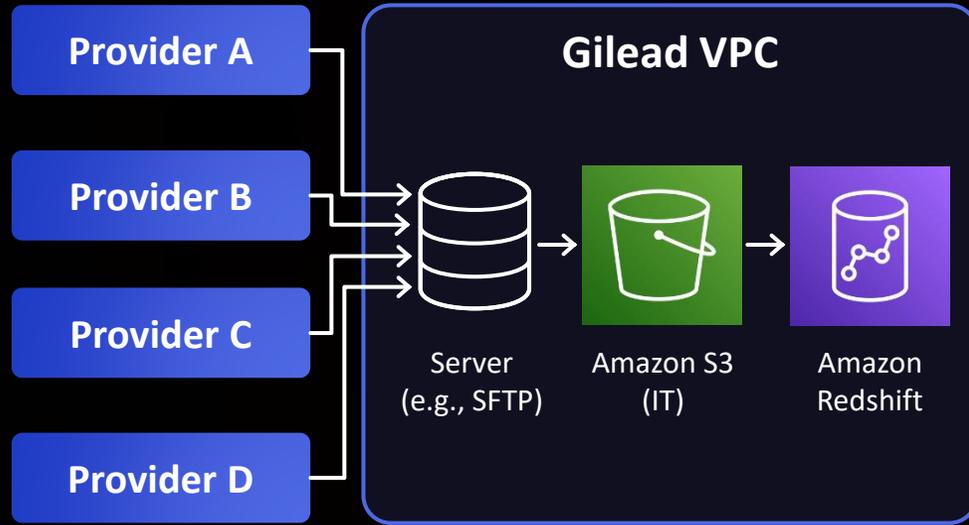


Gilead Data Marketplace

ADX COULD ACCELERATE THE THIRD-PARTY DATA DELIVERY BY MORE THAN 50%, REDUCE OPERATION COST, AND STREAMLINE DATA PROCUREMENT PROCESS

Without ADX

Subscriber manages infrastructure (i.e., servers/S3) while accessing integrating data from providers into their analytics platform(s)



Subscriber **manages the server operation** that includes storage, latency, security, access controls, and compliance

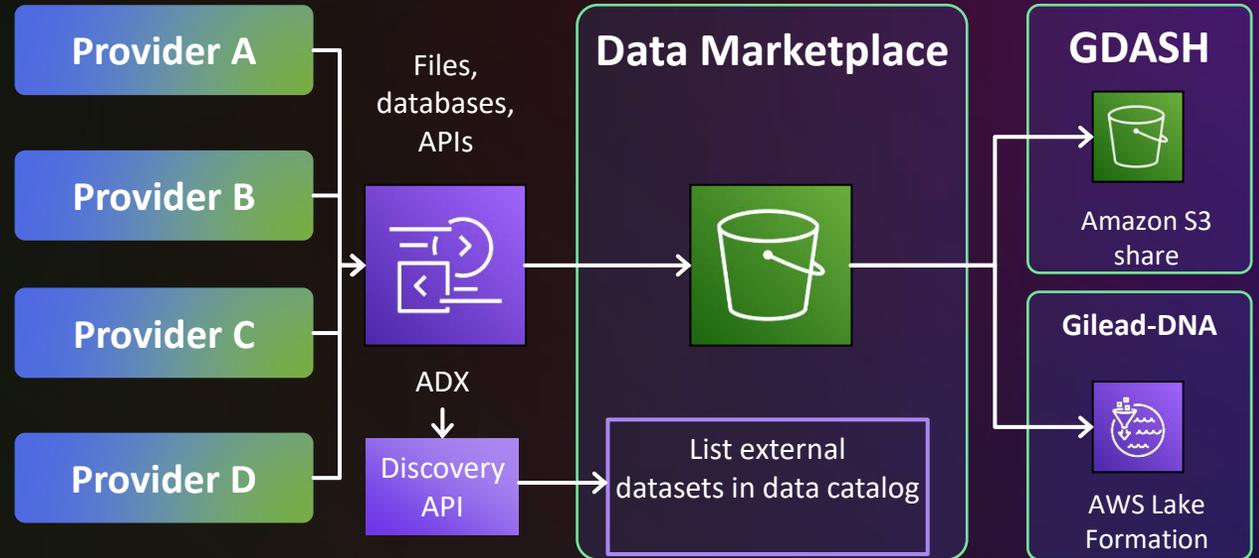
Subscriber has to **account for the variability** during the files from the server to providers (e.g., files updates, format)

Subscriber **maintains ETLs** to transfer the files from the server to S3 Subscriber and then to Amazon Redshift

Subscriber stakeholders **manually operate** data procurement life cycle including finding data providers, catalog, manage entitlements, etc.

With ADX

ADX manages the entitlement and delivery of data across providers, as standardized "data products" directly into Gilead platform(s)



ADX, as a managed service, **facilitates secure, compliant data sharing and malware scanning** between third-party data providers and subscriber

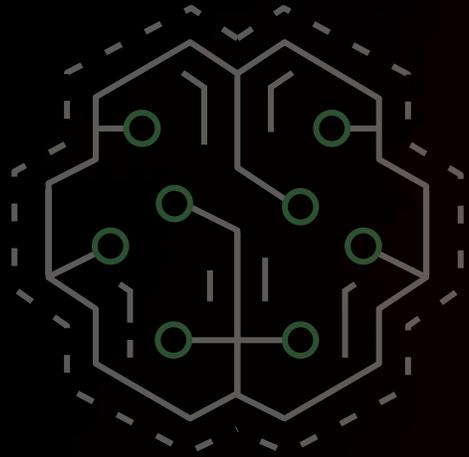
ADX **standardizes and automates the data delivery** between multiple third-party data providers and Gilead

ADX can facilitate 'no ETL' data sharing between vendors and Gilead

ADX streamlines **end-to-end data procurement** life cycles (i.e., request new data sets, manage/monitor entitlement, data sharing between vendor and subscriber)



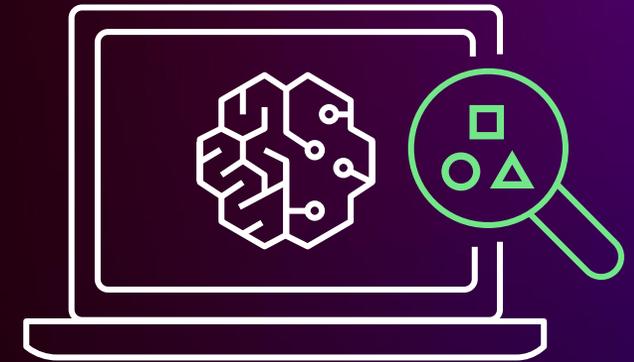
A few examples of innovation on AWS



Enterprise Semantic Search
(Morpheus)



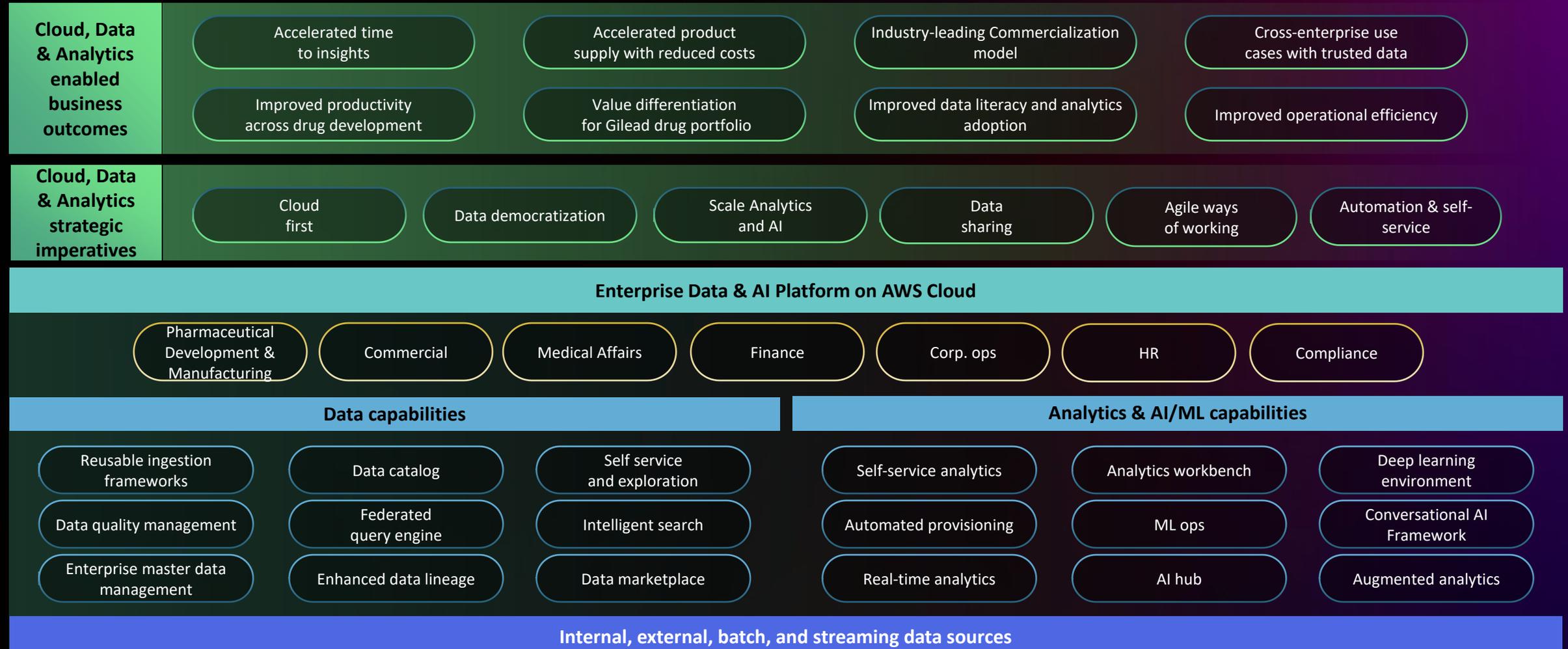
Data Marketplace



Enterprise Data & AI platform
(Gilead-DnA)

Gilead-DnA: Enterprise Data & AI Platform on AWS

ENABLING A DATA-DRIVEN CULTURE AT GILEAD



Why Gilead selected data mesh

SCALED APPROACH TO DATA MANAGEMENT

Lessons learned

Monolithic data platforms and approaches do not scale

Top-down data governance does not work

Reduce accidental complexity of data pipelines and copying data

Data movement and data copies increase costs and reduces trust in data

Adopt open data formats and avoid vendor lock-in

Adoption of tools that support open standards are imperative

Access to trusted data with speed along with data literacy leads to data-driven culture

Opportunities & Industry Trends

Adoption of AWS Cloud

New ERP rollout provides opportunities to simplify landscape

Fragmented Data landscape

Gilead-wide focus on deriving value from data and analytics

Gilead's aspirations require data management at scale

Emergence of Data mesh

Data platform guiding principles

Democratize access to trusted data

Minimize data movement and embrace storage and compute separation

Promote domain ownership and data sharing across the enterprise

Promote self-service and re-use of trusted data

Manage data as a product with enterprise data standards

Establish governance with cross-domain representation

Create a federated data consumption layer

Adopt cloud-first architecture with security compliance

Data mesh principles and outcomes



**Domain
ownership**



Data as a product



**Self-serve data
infrastructure**



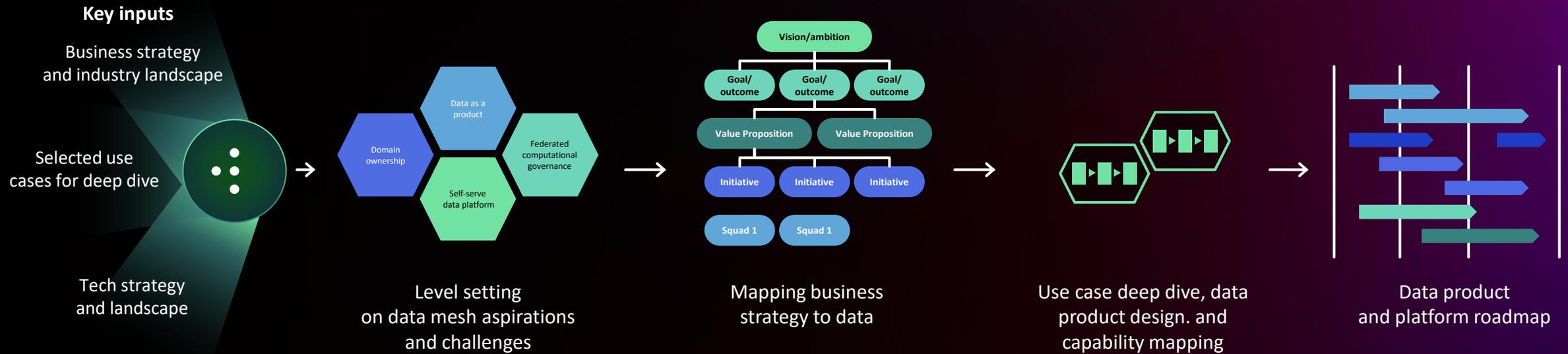
**Federated
computational
governance**

Respond gracefully to change in a
complex business

Increase the ratio of value from
data to investment

Sustain agility in
the face of growth

High-level approach for data mesh implementation



Prepare

Connect high-level business objectives to data strategy

Data strategy

OUTPUT #1
Lean value tree

OUTPUT #2
Aspiration and challenges

Deep dive on a single use case

OUTPUT #3
E2E flow to support a business use case

Co-create roadmap

OUTPUT #4
First version of Roadmaps for the data products + platform capabilities

Top-down design

START WITH EXPERIENCES AND WORK DOWN TO INFRASTRUCTURE

Mesh

Consumer-facing products are the face of the platform. They showcase data product standards and quality

Mesh experience products

Platform discovery

Main window to the platform for discovery
Example capabilities: search, lineage, request for access.
Build on top of data catalog

Data sharing

Managing access to data products
Capabilities such as: RBAC, tag-based, masking policies. Sharing derived and "Bring Your Own Data" products.

Scorecard

Providing visibility into trustworthiness around understandability, data quality, etc.
Incentivize good behaviour.
Build into data catalog

Enhance user experience

Adoption dashboard, GUI data product creator for less technically savvy customers, and other long term goals
Build into data catalog where possible

Platform APIs

APIs enable self-service, enforce mesh standards, and allow all domains to bring their data to the mesh by separating how a data product is cataloged from how it is produced

Platform products

Data product registration API

One-step self-serve Data Product registration, regardless of infrastructure used. Single point of entry to the mesh enforces data product standards

Data product observability API

Standard interface for capturing runtime metadata such as pipeline health and data quality independent of infrastructure

Data product access API

Allow provisioning access in any workflow through automation or GUIs

Data product quality API

Update data quality metrics in the data catalog

Infrastructure

The platform provides self-serve spec-driven managed infrastructure for domains that want it

Infrastructure/utility products

Self-serve infrastructure

Provisioning of cloud infrastructure without any human coordination
No domain logic
Infra provisioning



End-to-End data product experience

PRODUCERS, STEWARDS, AND CONSUMERS



Govern

Enhance



Publish

Build

Data Product Scorecard

Data Access Policy Management

Master Data Management

Federated Governance

Data Product Authoring Experience

Data Product Shopping Experience (Discovery and Access Request)

Data Product Registration API

Data Product Observability API

Data Product Quality API

Data Product Grant Access API

Enterprise Data Catalog

Discover

Request Access

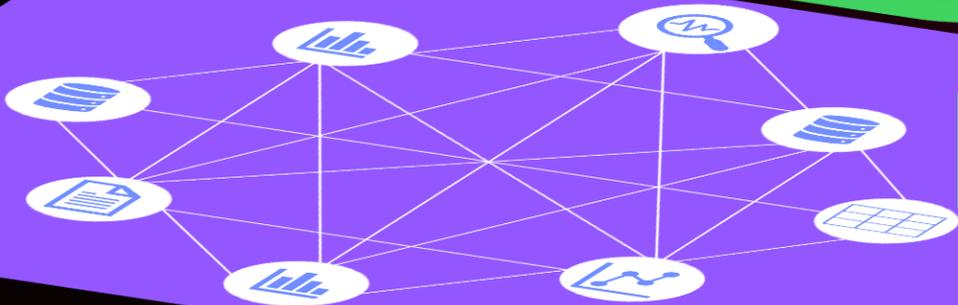


Data Consumer

Data Producer

Use

Data Products



Federated Query Engine

Data Mesh



Transparency and trust in data products

Characteristics

Before

After

Discoverable

Technical data in catalog only



Business, technical, and observability metadata

Addressable

Tied to infrastructure



Infrastructure independent

Trustworthy

Adoption apprehension



Lineage, SLOs, and quality in catalog

Secure

Lack of processes and feedback loops



Fine-grained access control
Standard approval workflows

Interoperable

Siloed LOB efforts



Cross-LOB coordination

Self-describable

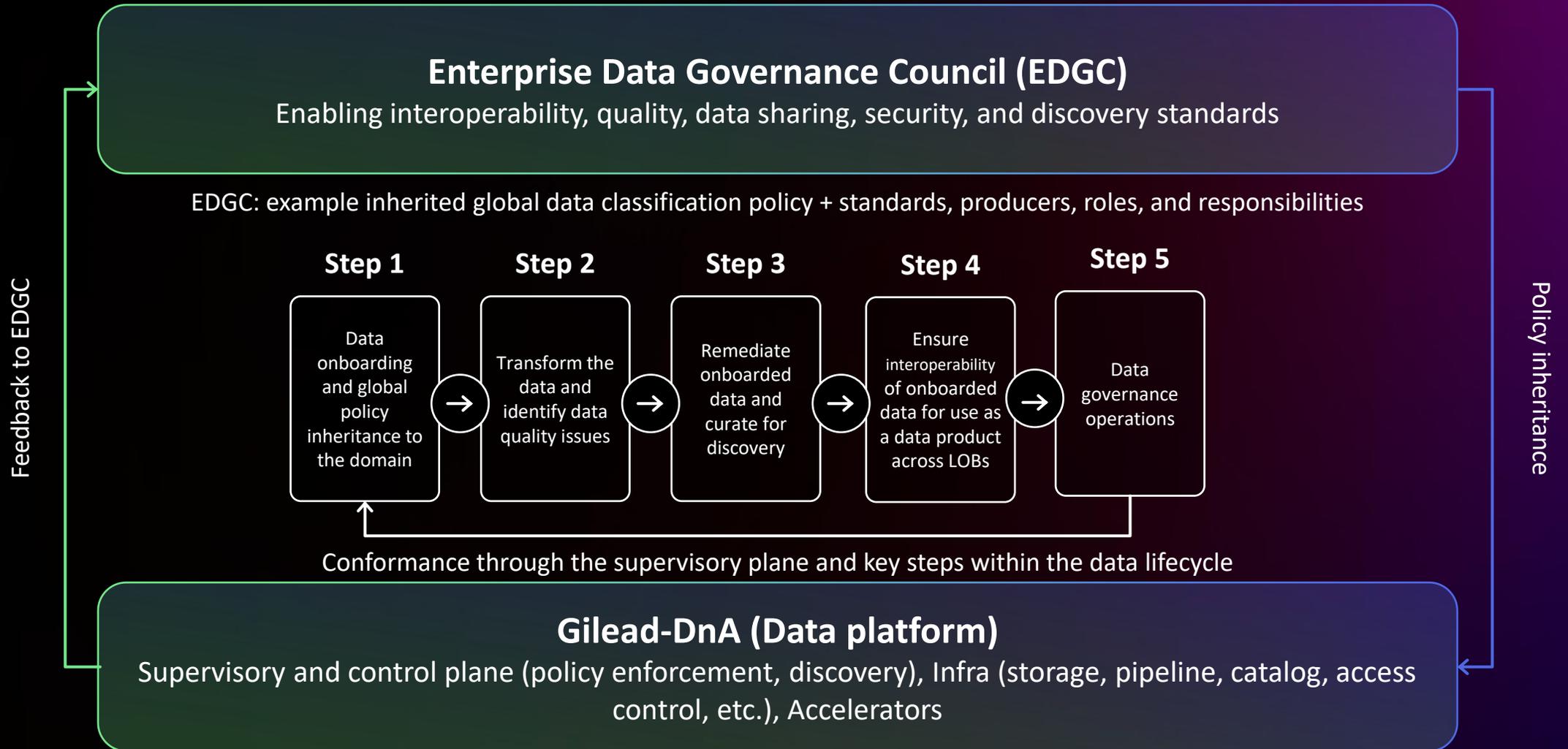
Domain SME help needed



Description in catalog

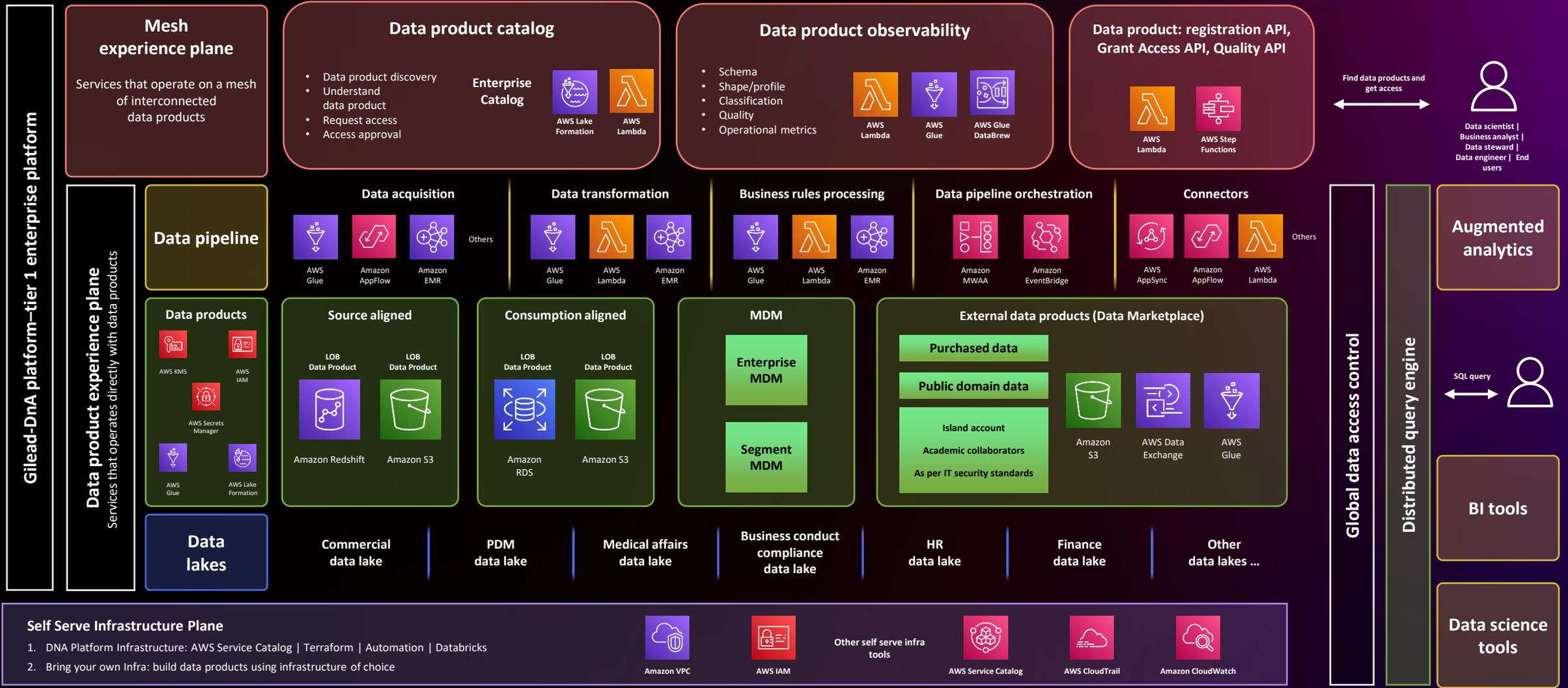
Data governance in a data mesh approach

COMPUTATIONAL GOVERNANCE AND MANAGEMENT



Gilead-DnA Platform: Data Management Services

DATA MESH SYSTEM CONTEXT



Data mesh adoption metrics

Data Mesh Principle	Objective	Measure	Metric
Data as a Product	Scale out data sharing across Gilead and generate value from data	Availability of Data Products	New Data Products published in Data catalog by Quarter
		Usability of Data Products	New Data Product Usages in Semantic Layer by Quarter
		User Satisfaction of Data Products	Avg. Rating of Data Products in Data catalog
Data Product Domain Ownership	Increase efficiency and effectiveness of data sharing	Domain Participation	New Data Producers in Data catalog by Quarter
		Interdependency between Domains	# Data Products shared between Domains in Data catalog
Self-Serve Platform	Increase domains' autonomy with lowered cognitive load and lower their cost of data ownership	Platform Adoption	New Producers by Quarter
			New Consumers by Quarter
		Domain Autonomy	# and usage of Self Serve Capabilities
		Platform Service Coverage	# of New Platform Features by Quarter
Federated Governance	Increase accountability for security, quality and integrity of Data	Domain Engagement in Data Governance	# of Data Product Owners who are active members of Data Stewardship Working Groups
		Confidence and Trust in Data Products	Data Quality Scorecards
			Meta Data Articulation Scorecards
		Compliance and Security of Data Products	# of Data Products with Classifications
			# of Data Products with Access Policies

Critical success factors in adoption of Data Mesh



Organizational model & Change management

Define the domains, organize teams around these domains

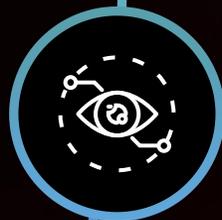
Define personas and make business teams assume data ownership responsibilities

Data culture: Right skills; Data governance committee with regular cadences



Sponsorship

Sponsorship at many levels—C-level sponsors supported by execution level champions are critical for success



Vision

Clarity of business use cases and goals around data;



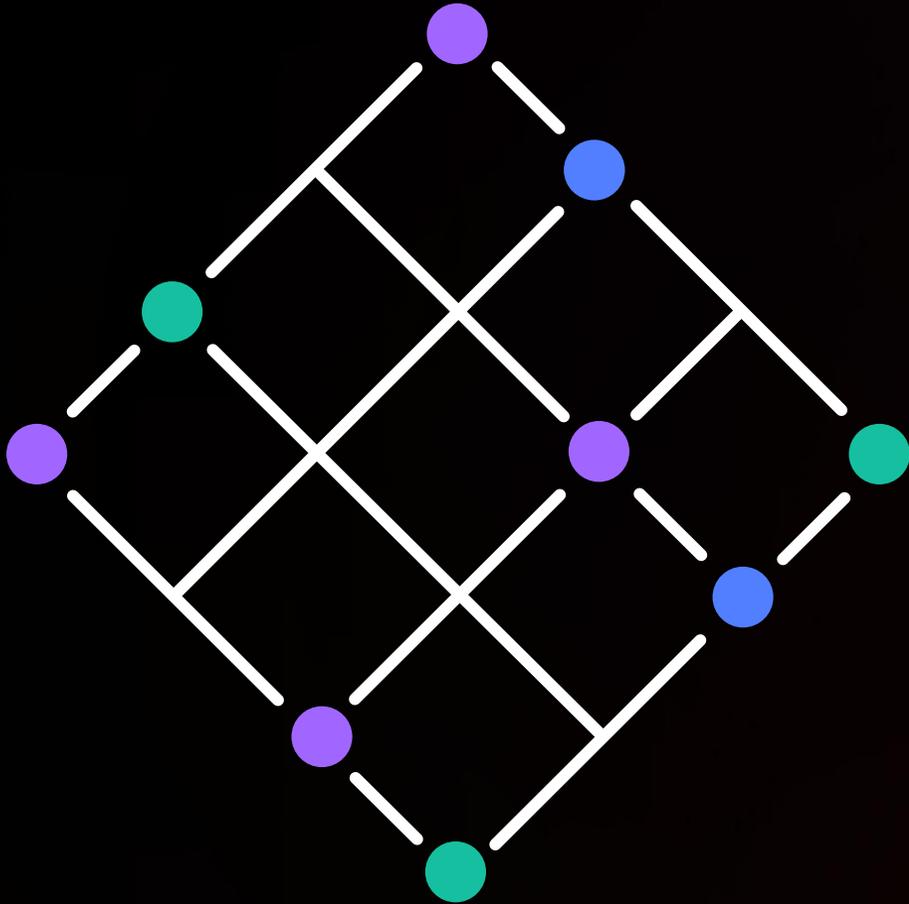
Well-informed stakeholders

Business and technical stakeholders are committed to adopt the vision

Key Takeaways

- Data mesh is real and not obsolete
- Mindset, organizational, and operational models are the biggest barriers to the adoption of data mesh
- Educate stakeholders and domain teams about Data Mesh
- Rapidly demonstrate value of data mesh for a use case – learn and adjust
- Gilead has successfully implemented a data mesh using the AWS ecosystem of services and a few non-AWS tools such as Starburst for Federated query engine and Collibra for data catalog
- Implement observability to maintain data quality
- Reduce the barrier for producers to create data products and for consumers to access data products
- Implement an adoption framework including metrics to measure
- Cloud-native skills are very important – choose the right implementation partner to adopt data mesh
- Manage change effectively

Realizing the benefits of data mesh



Enable organizations to **get value from data** at scale

Create a **business-oriented data product** that can support the top strategic goals

Allow business domains **federated governance** through lightweight decentralized policy

Encourage **data-driven agility**

Support the **sharing of data products** with the goal of **driving business results**

Accelerate data mesh with Amazon DataZone

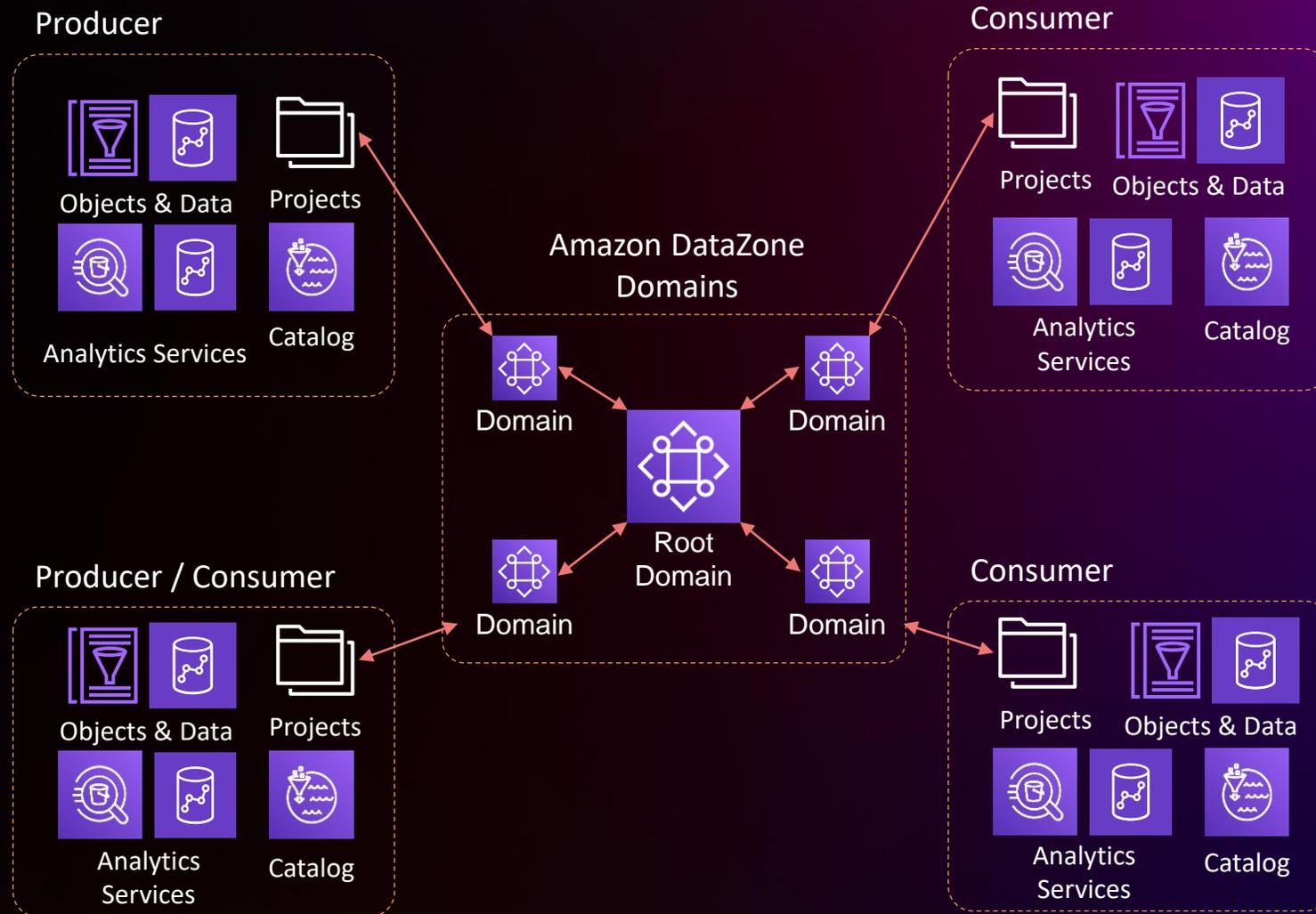
SCALE AND REMOVE BOTTLENECKS TO FULLY DEMOCRATIZE DATA

DECENTRALIZED
OWNERSHIP

FEDERATED GOVERNANCE

PEER-TO-PEER
DATA SHARING

SELF SERVICE
INFRASTRUCTURE



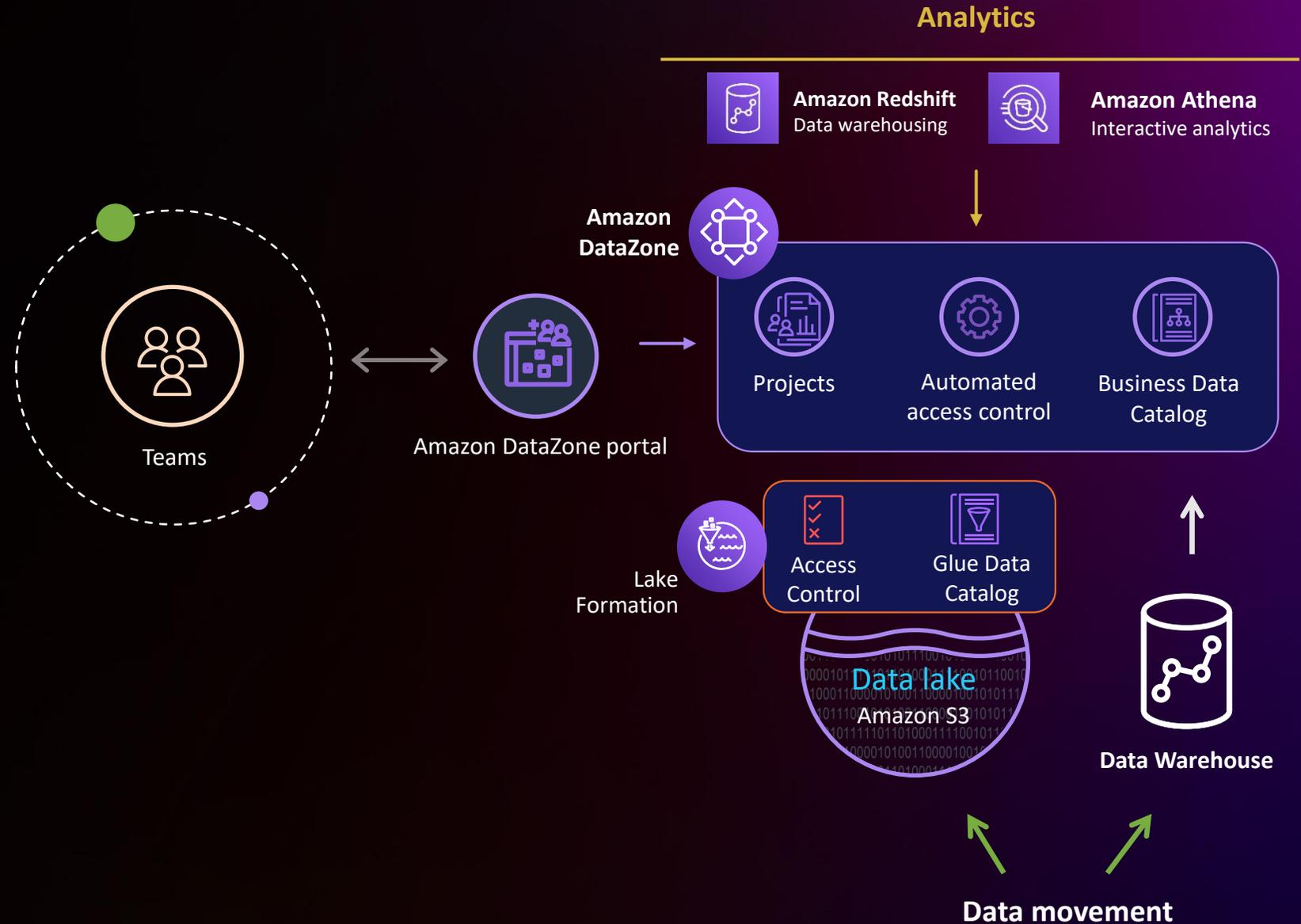
Amazon DataZone extends the AWS analytics stack

ORGANIZATION-WIDE
BUSINESS DATA CATALOG

GOVERNANCE AND
ACCESS CONTROL

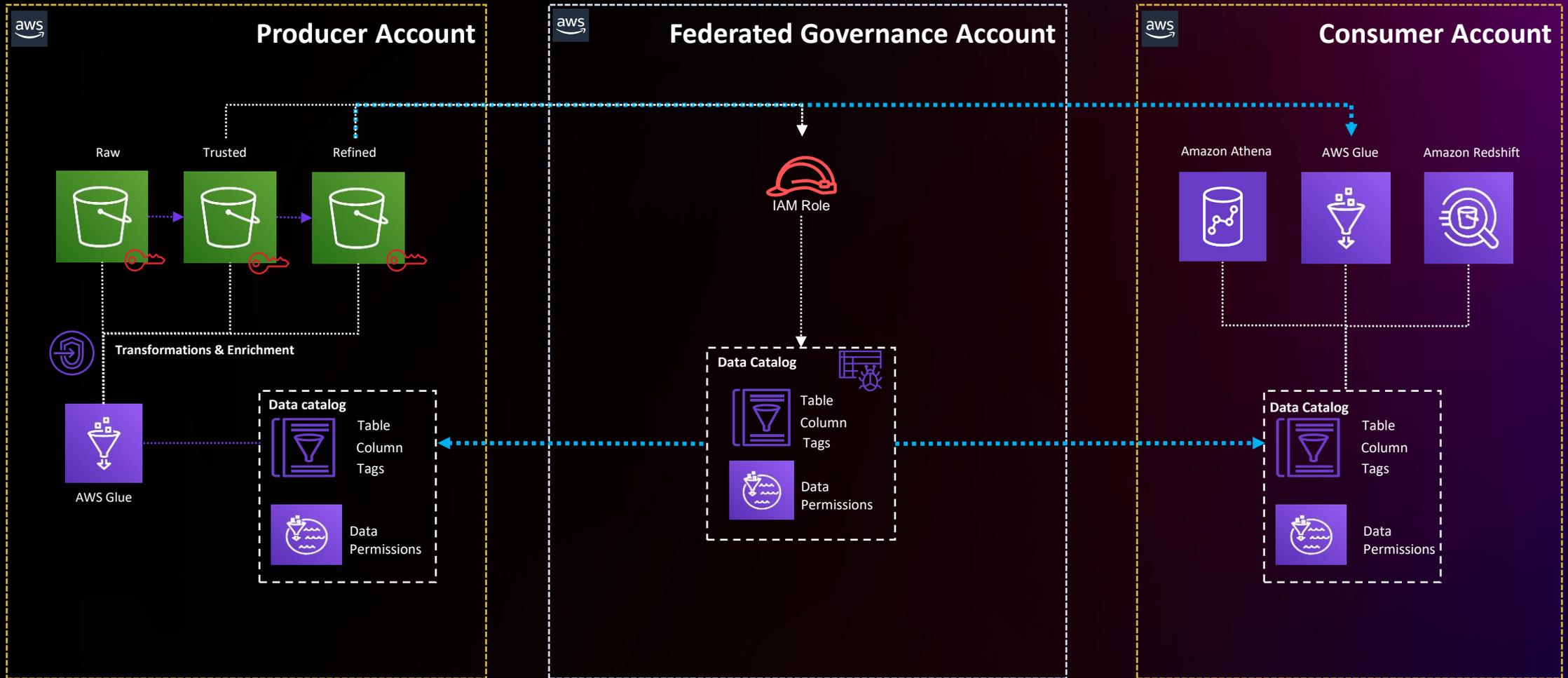
SIMPLIFIED ACCESS
TO ANALYTICS

DATA PORTAL



Build your data mesh with AWS Lake Formation

FINE GRAINED SHARING FOR ULTIMATE FLEXIBILITY



Thank you!



Please complete the session survey
in the **mobile app**

