Building a Modern Data Strategy for the Public Sector

Getting Started with Intentionality

Carlos Rivero (he/him)

Executive Government Advisor
Worldwide Public Sector
rivercap@amazon.com

Today's Agenda

The need for modernization
What's in it for my constituents?

Challenges
Why is it so hard?

Architecture
How does data flow throughout my organization?

Management
How can modern data management principles help?

Intelligence
How can I empower my stakeholders to make better informed, data-driven, actionable decisions?

AWS Approach
How do we get there?

Use Case Deep Dive
Why develop and implement a modern data strategy?

Poverty Reduction

"In Washington state, more than half a million children live in families that struggle to make ends meet…this is unacceptable anywhere, but especially in a state with so much prosperity. We must do whatever we can to reduce poverty in Washington."

– Governor Jay Inslee

Improving Education and Workforce Outcomes

- Pennsylvania Governor Tom Wolf
- Longitudinal Data System (LDS)
  - Measure long term education and workforce outcomes
  - Produce insights to refine programs for continuous improvement

Improving Mental Health Care

Georgia mental health bill passed unanimously through state Legislature, signed by Governor Kemp orders private insurance companies and publicly funded programs to cover mental health disorders like other medical conditions.

Combatting Substance Use and Abuse

“The CDO shall focus initial efforts on developing a project for the sharing, analysis, and dissemination among and between state, regional, and local agencies of data related to substance abuse, with a focus on opioid addiction, abuse, and overdose.”

- Virginia SB 580 signed 2018

Combatting Homelessness

“The overlay between behavioral health and housing stability is as clear as day. Oregonians need a stronger, more accessible behavioral health system that meets them where they are and matches them with the appropriate level of care that they need” and the “homelessness, to jail, to hospital pipeline needs to be addressed.”

- Governor Tina Kotek.

Image credit to Statesmanjournal.com

What public sector organizations are looking for

Enhanced agency staff productivity

- Modernize legacy work flow processes to minimize manual work
- Simplify case review/ adjudication processes
- Introduce productivity tools – e.g. intelligent search, chat bots

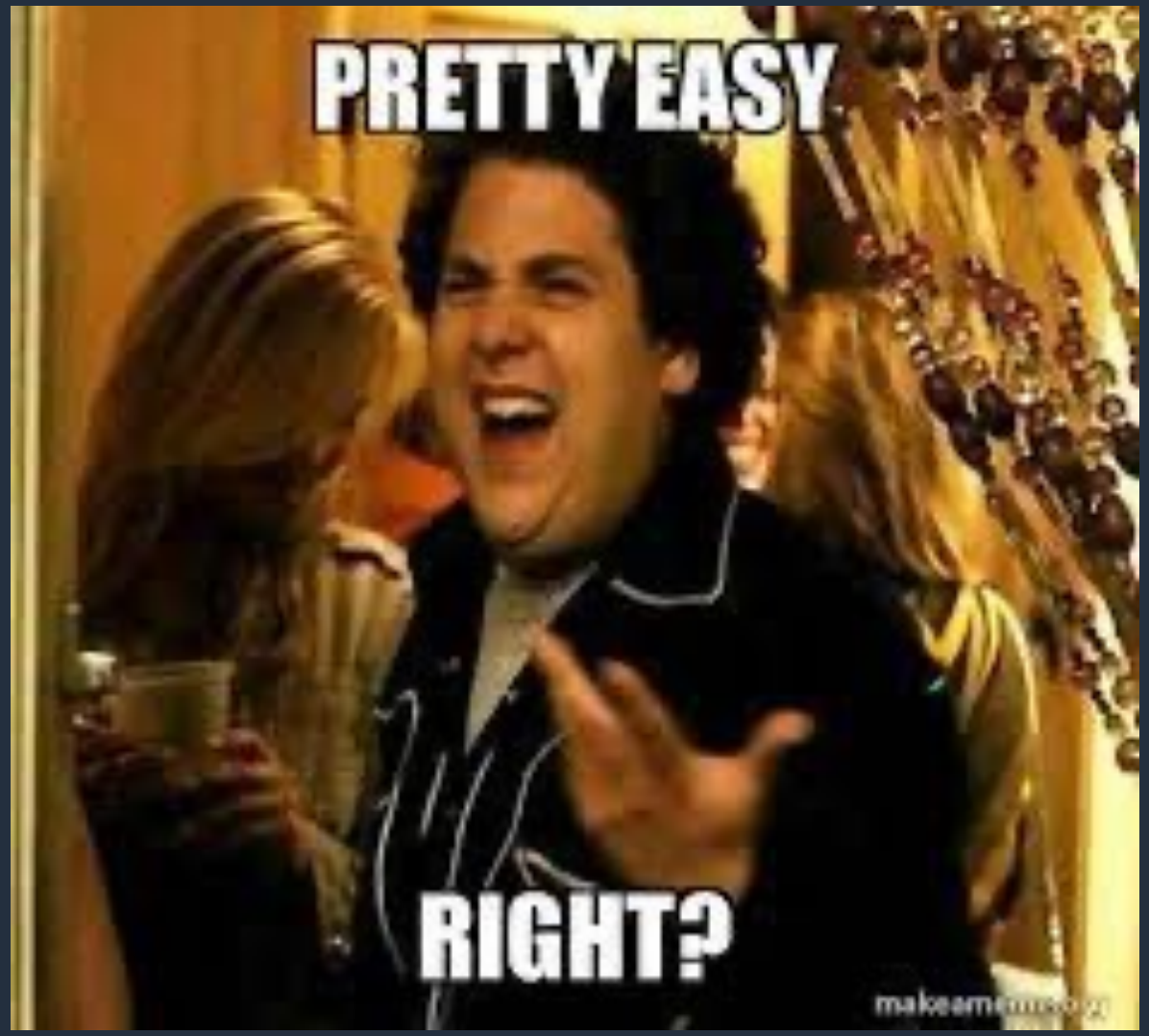
Enabled Program Leadership / Benefits administrators

- Enable data-driven decision-making and deep program insights
- Reduce fraud, waste and abuse
- Develop forecasting for enrollment, staffing, budgets, etc.
- Provide program compliance and reporting
- Adapt to new policy changes

Improved constituent experience

- Minimize time for benefits disbursement
- Provide self-service capabilities and real time status notifications
- Improve enrollment and eligibility verification & processing

aws



Challenges
Why is this so hard?

Challenges we are hearing from public sector customers

Demand for services is rising while resources and capacity to deliver them **aren't keeping pace**

Constituents increasingly expect government to **provide modern digital experiences** for conducting online transactions

Aging infrastructure for data capture, storage, and management **creates friction** for leveraging data for analytics and machine learning

Complex security, privacy, and compliance requirements create barriers to change and block adoption of many SaaS solutions

Risk averse culture and institutional inertia, slow innovation

aws

...and more challenges

BUSINESS CHALLENGES: Struggle to answer important Qs timely, limited transparency and auditability erodes trust, limited analytics capacity and capabilities to data share

DATA CHALLENGES: No defined governance models, poor data quality, no metadata for audit/compliance/provenance, no data cataloguing

OPERATIONAL CHALLENGES: Manual/ad hoc processes consume resources, inconsistent data manipulation, no 2-way pipelines, variation in data extraction and identity resolution, PII and PHI protection untraceable and vulnerable

FUNCTIONAL CHALLENGES: Roles and responsibilities not codified, not scalable

The challenges reflect legacy approaches to using data

BUSINESS CHALLENGES: Struggle to answer important Qs timely, limited transparency and auditability erodes trust, limited analytics capacity and capabilities to data share

DATA CHALLENGES: No defined governance models, poor data quality, no metadata for audit/compliance/provenance, no data cataloguing

OPERATIONAL CHALLENGES: Manual/ad hoc processes consume resources, inconsistent data manipulation, no 2-way pipelines, variation in data extraction and identity resolution, PII and PHI protection untraceable and vulnerable

FUNCTIONAL CHALLENGES: Roles and responsibilities not codified, not scalable

aws

Trends driving a change in data strategies

Growing exponentially

New Sources Velocity & Variety

AI/ML across the Data Value Chain

Diverse Data "Customers" and Personas

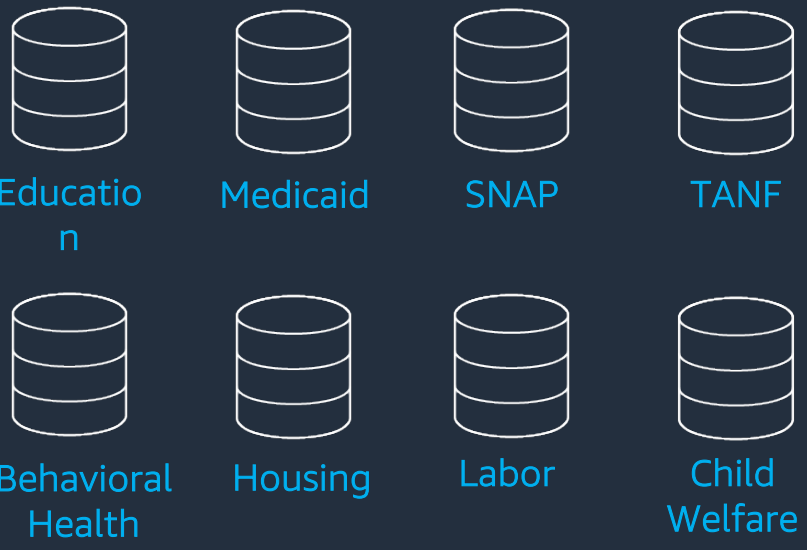
Insights Embedded in Workflows

Source: IDC Global DataSphere Forecast, 2022-2026: Enterprise Organizations Driving Most of the Data Growth

Autonomous business processes and data systems

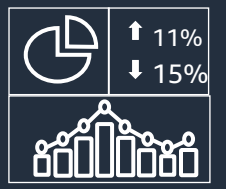
CURRENT STATE

Critical data inputs owned and managed by different departments and entities who haven't shared their data in this way before.



FUTURE STATE

- Data-driven decision making for resource allocation



- People are connected with the right services at the right time

- Data driven feedback loop supports continuous improvement

- Transparency, prudent use of taxpayer dollars

Government Agency

Law Enforcement

Medical Professional

Policy Maker

Innovative Research

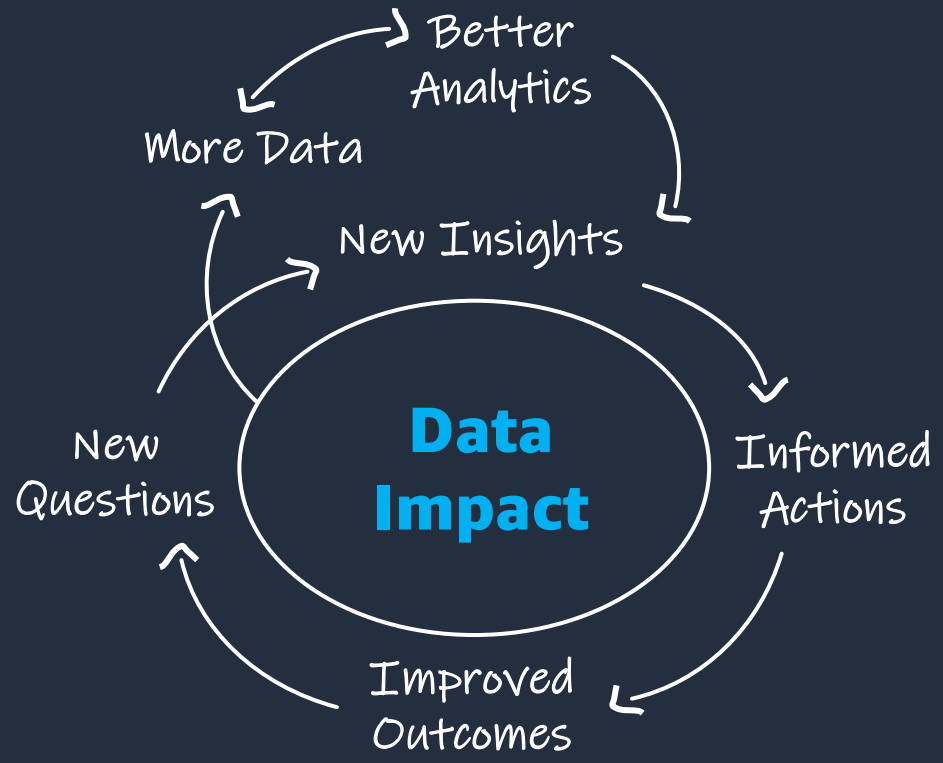
Social Worker

Academia Higher Ed

Community Leader

Public Stakeholder

The Data Driven Organization



"An organization that harnesses data as an asset, to drive sustained innovation and create actionable insights to improve policy making decisions that reflect outcomes constituents care about and trust government more."

Key Characteristics

Set 'Think Big' goals

Focus on delivering policy priorities with quality

Shared leadership conviction and Business-IT alignment on data ownership

Strong collaboration and agility concerning data products across data producers and consumers

Upskilled and empowered producers and consumers who self-serve

Privacy, security, compliance and federated governance without impeding innovation

aws

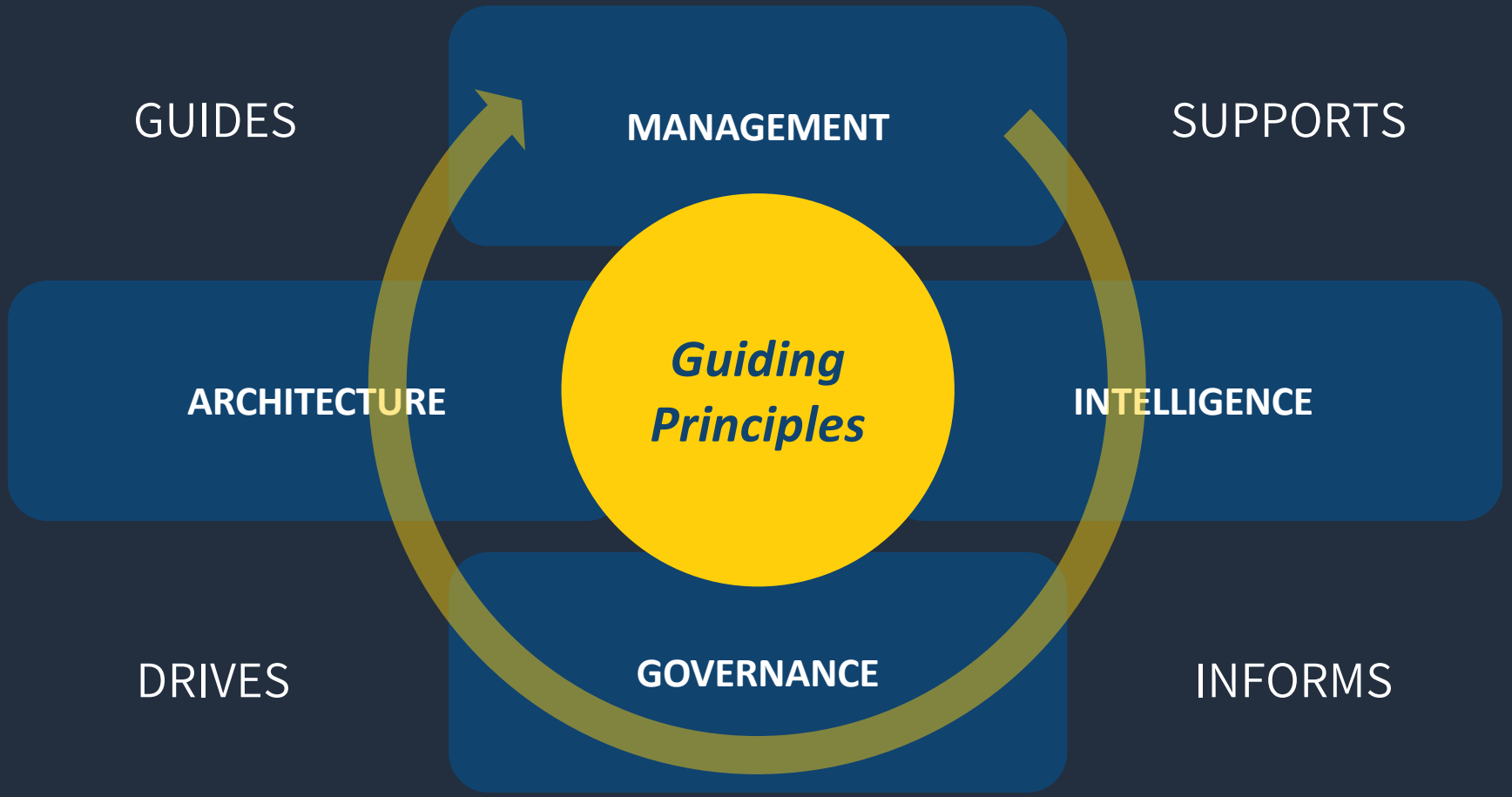
A suggested definition for modern data strategy

An **iterative** plan of **aligned actions** spanning mindset, people, process, and technology that **accelerates creating value using data** in direct support of strategic objectives

(Mindset **+** People **+** Process)

x Technology

Components of a modern data strategy



Establish guiding principles to communicate from the top

These represent the key tenets for your data strategy

Examples:

Empower the service agencies and divisions to solve their specific problems

Only move data when it is absolutely essential for supporting a specific use case.

Abstract technical complexities so teams can move fast.

Decentralized governance to improve agility while still protecting the data.

How does governance build trust?

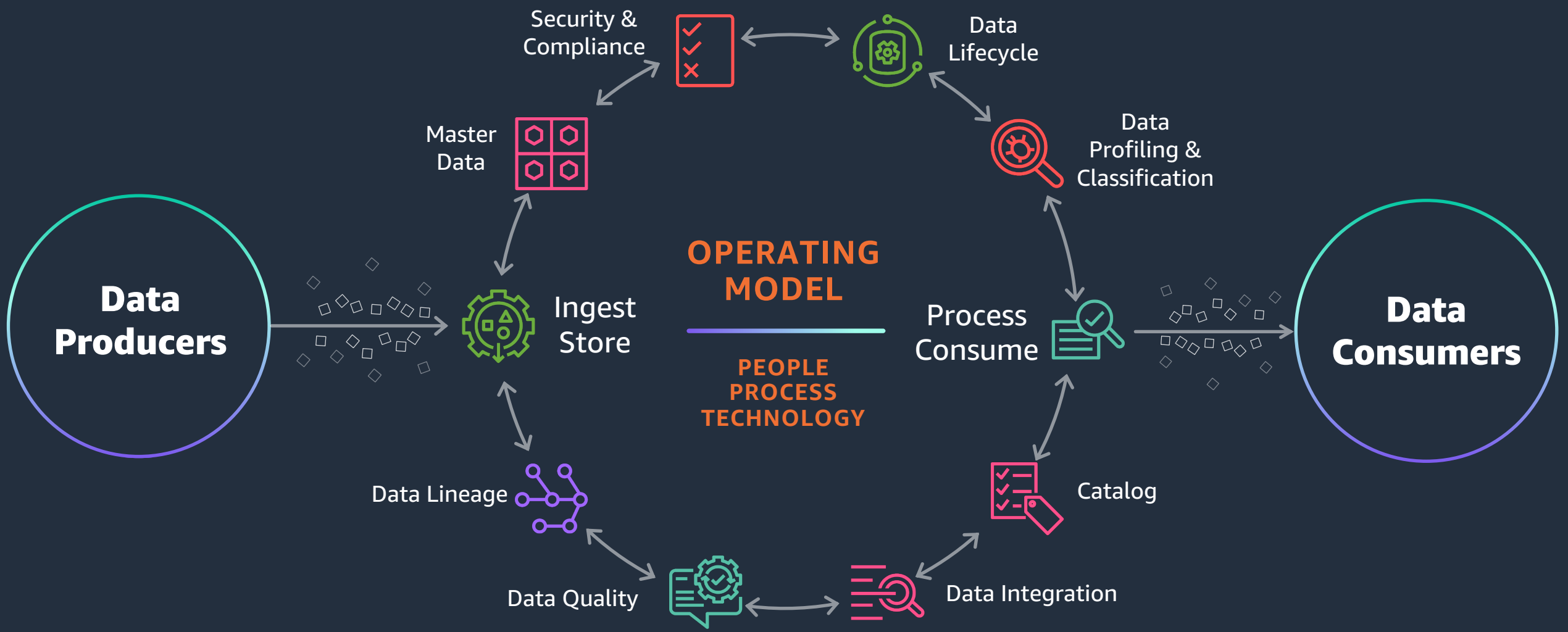
- Establish **accountability**
- Operate **transparently**
- Promote **visibility**

aws

Architecture

How does my data flow throughout the organization?

Build Your Operating Model



Modern Data Architecture Principles

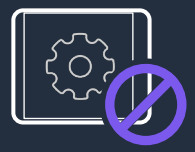
Scalable Data Environment

Purpose-built Storage and Analytics Services

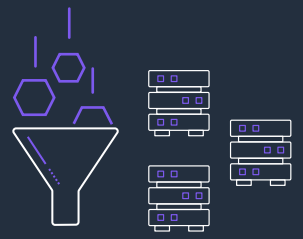
Unified Governance

Performant and Skill Aligned

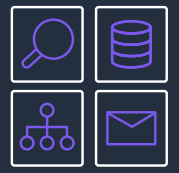
Seamless Data Movement- DB freedom



Infrastructure Decoupled from Business Domains and Use Cases



Distributed and Event-Driven



Loosely Coupled Services and Component Based Architecture

Persona-Oriented Consumption

Specify data management strategies

- Quality data: complete, clean, contextualized, and normalized
- Wisely and intentionally stored, secured, and organized data assets
- Component architecture with loose coupling and autoscaling
- Core initiatives and policy outcomes
- ML and applied data transport for security automation, logging, and audits
- Flexibility to enable functions across diverse organizations

Reflect those strategies in your modern data architecture

Durability and Availability

Replicate data across regions and availability zones to ensure your data is available globally with 99.999999999% durability and 99.99%+ availability

Security

Protect data with advanced encryption, fine grain access control (IAM), encryption key management (KMS), logging (CloudWatch / CloudTrail), and sensitive data discovery (Macie)

Object Level Controls

Attribution for fine-grain, object level control allows tagging of valuable data for replication and tiered storage, saving money, and increasing performance

Flexibility

Storing all data in one flexible and layered data environment avoids data silos and the cost of moving data or replicating data management processes

Operational Data Store

Creating an Operational Data Store (ODS) to access structured frequently used data for real time insights with off the shelf API

ML/AI

Once your data is in an AWS cloud services platform-automate data transport and security functions, and pull business insights faster and more efficiently with ML/AI

aws

Management
How can modern data management principles help?

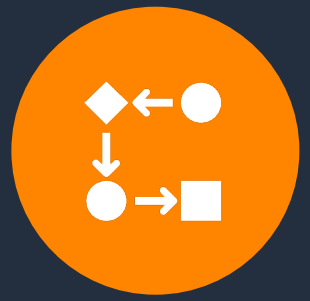
Master data management

Define

Documentation throughout the data lifecycle is critical. This includes the source of truth; the data entry mechanism, parameters, and data owners; and the data quality/definitions. Machine learning can help with a lot of this heavy lifting today.

Map

Know how the upstream and downstream impacts of data collection and data system design leads to more purposeful investments and improved data-driven action- data transformations and movements should be reusable, documented, and auditable.



Normalize

Data standards allow for quality entry and quality normalization. Transformations should occur from the cleaned original data whenever possible, whereas merging and novel dataset curation can occur from normalized data.

Protect

Defining data elements categorically as well as by attribute facilitates improved ABAC (attribute based access controls), easier auditing, and enhanced data security automation.

Maintain

Frequent policy changes and technology advancements will require regular review cycles of both data and infrastructure practices. These reviews should focus on continuous quality improvement and shared learning across program, IT and Data teams.

aws

Goal: Updated experience enables….

- ✓Automated data ingestion automation in near-real time
- ✓Durable and inexpensive storage
- ✓User friendly cataloguing, data attribution, and access management
- ✓Consistent data transformations with automation and logging
- ✓Meta data management for streamlined data consumer access
- ✓Reliable answers and consistent access to curated datasets and dashboards that update/refresh regularly
- ✓Secure data practices, encryption, ML for sensitive data protection, and auditing capabilities for compliance

AWS supports your modern data strategy

Comprehensive

Comprehensive set of services for storing and querying structured unstructured and vector data

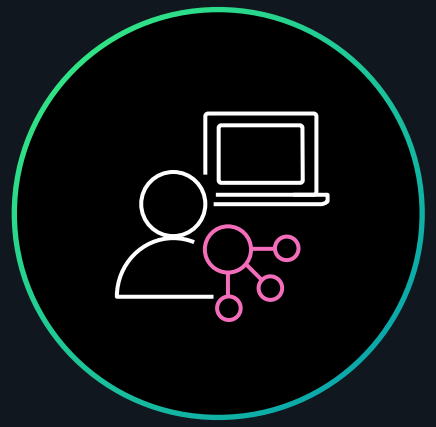
Integrated

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

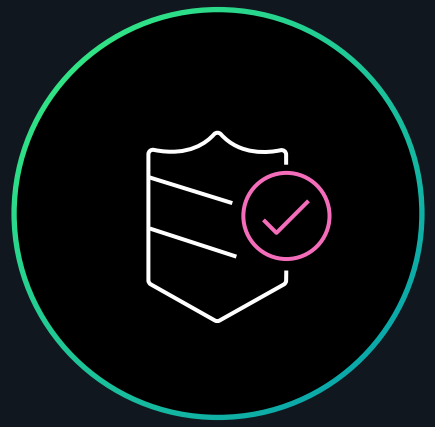
Governed

End-to-end data governance capabilities responsible AI and regulating user interactions with LLMs

Governance provides guardrails to innovate faster



Find, access, and share
the right data

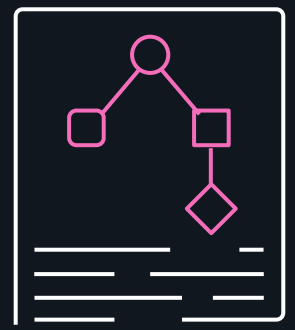


Keep data
safe and secure

Enable appropriate
audits and controls

Audit dataset usage by user and business use case

Business user



Audit dataset usage, business use case; monitor costs across projects

Intelligence
Empower stakeholders to make data-driven decisions

AWS provides an end-to-end data foundation

Data sources

- IOT/ DEVICES
- APP/ LOGS
- THIRD-PARTY DATA

Integrate

Store, query, analyze

DATABASES

Amazon Aurora

Amazon DynamoDB

Amazon Kinesis & Amazon MSK

DATA WAREHOUSE

Amazon Redshift

DATA LAKE

Amazon S3

BIG DATA

Amazon EMR

Act

BUSINESS INTELLIGENCE

Amazon QuickSight

MACHINE LEARNING

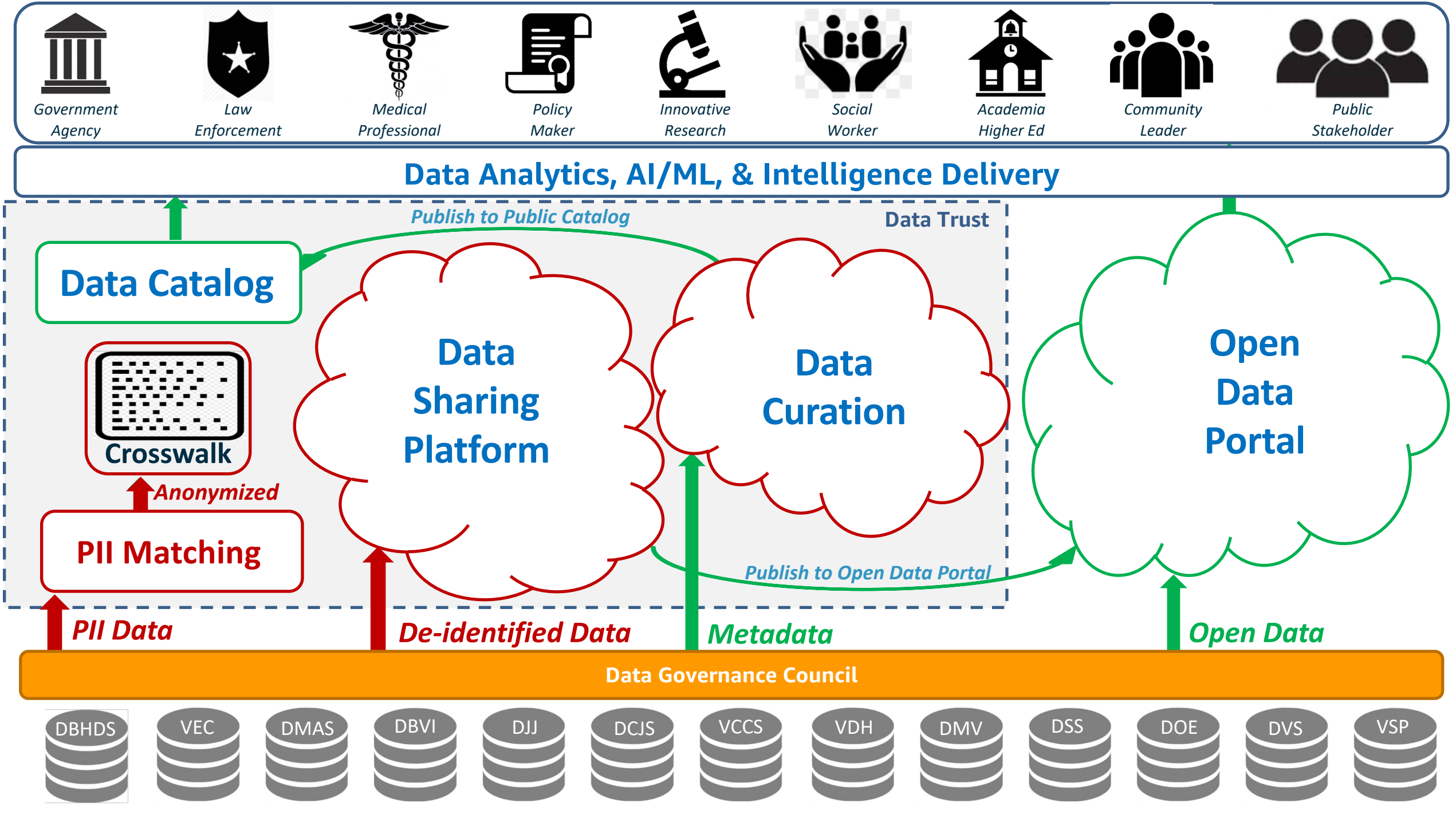
Amazon SageMaker

GENERATIVE AI

Amazon Bedrock

- PEOPLE
- APPS
- DEVICES

Catalog and govern | AWS Lake Formation, Amazon DataZone


Government Agency
Law Enforcement
Medical Professional
Policy Maker
Innovative Research
Social Worker
Academia Higher Ed
Community Leader
Public Stakeholder
Data Analytics, AI/ML, & Intelligence Delivery
Publish to Public Catalog
Data Trust
Data Catalog
Crosswalk
Anonymized
PII Matching
Data Sharing Platform
Data Curation
Open Data Portal
Publish to Open Data Portal
PII Data
De-identified Data
Metadata
Open Data
Data Governance Council
DBHDS
VEC
DMAS
DBVI
DJJ
DCJS
VCCS
VDH
DMV
DSS
DOE
DVS
VSP

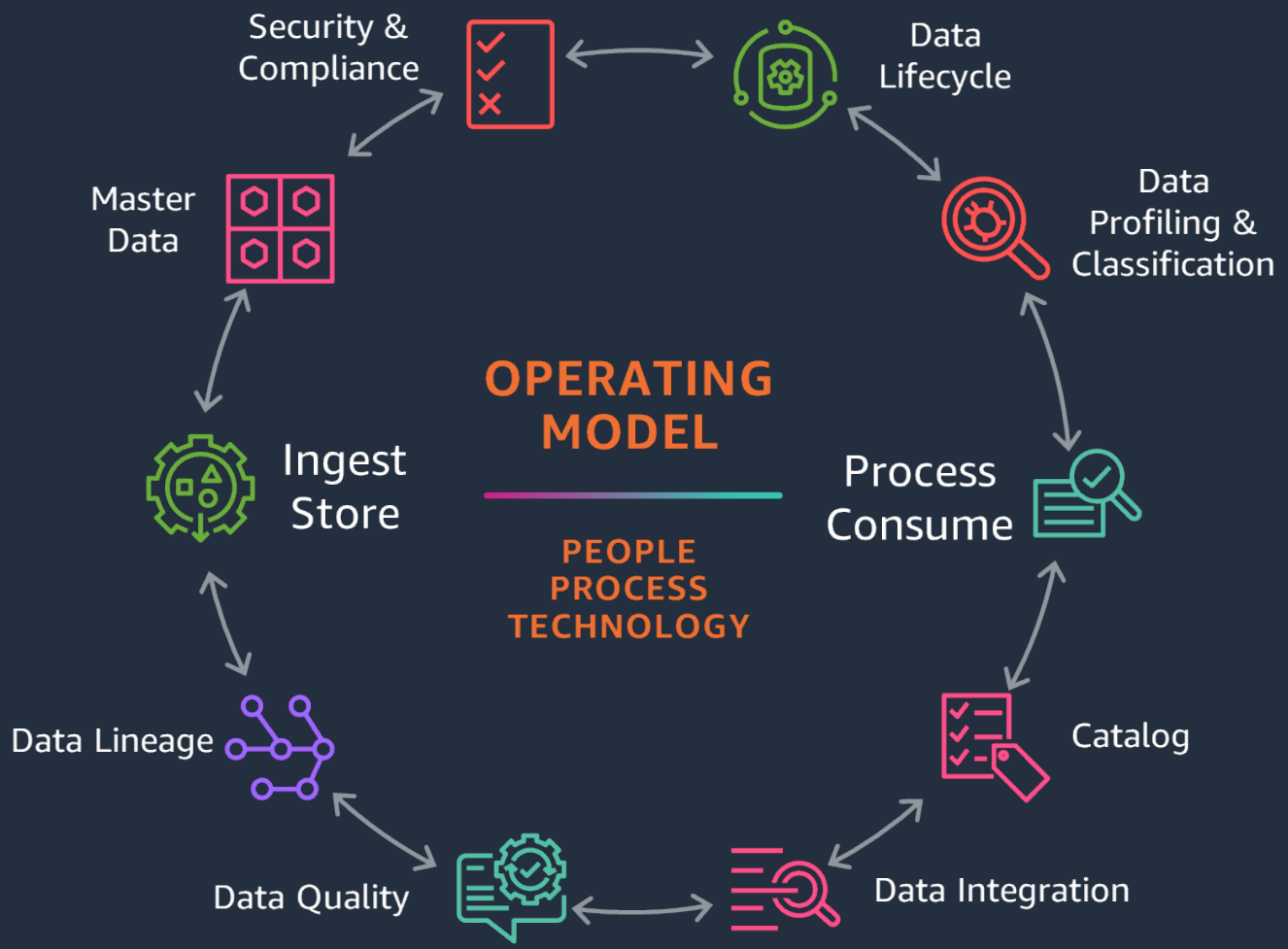
How do we get there?

Roadmap for Getting Started

1. Think Big, Start Small, Scale Fast.
2. Engage the right people.
3. Leverage best practices.

1. Think Big, Start Small, Scale Fast

Microservice based operating model that is flexible, adaptive, and scalable



THINK BIG, START SMALL, SCALE FAST

1. Architect your environment to support the wider data strategy
2. Implement incrementally based on business initiatives and use cases that drive the data strategy
3. Further evolve data and intelligence capabilities over time

5 Start Small Examples

Agency Data Exchange

360° View of Constituents

Continuum of Care Data Hub

Targeted Digital Comms with Data Collection

Migrate document-based data sources to data lake for analytics

WHAT GOVERNMENT DECISION MAKERS CARE ABOUT				
When we bring defined data together from different departments what longitudinal patterns of success, or gaps in service can we observe and make available to the public and policy makers?	Which families are we serving in multiple programs, and are all their needs met with the current service array? Are our service offerings aligned with the needs?	Are providers able to follow through on referrals and provide families with the right array of services proactively? What are the gaps in care that lead to reactive service needs that could have been prevented?	Which of our safety-net recipients have moved and need to update their address leading up to a policy change implementation or recertification?	What factors of service or program access have improved permanence across kinship placements for foster care youth removed from parental custodial care? Does this vary by geography?

Diverting Mental Health Crisis From Jails

FULTON COUNTY, GA & TYLER TECHNOLOGIES

1. Connects and shares data across **criminal justice, courts, housing, and behavioral health** systems
2. Gives agencies more comprehensive view of an individual's **treatment journey**
3. Connects individuals in a crisis to better services and **alternatives to arrest**
4. Cloud allows quick and endless scale to expand and add new data sets over time

Improving Health Care Quality

MICHIGAN HEALTH INFORMATION NETWORK (MIHIN) & CLOUDTICITY

1. **MIHIN** passes **2M** patient health information messages weekly
2. Connects patient electronic health information with providers and payers
3. Speed and scale to meet growing demand for new data sources (e.g. SDOH)
4. Patient matching process improved and standardized quality data sharing

Helping People Achieve Wellness

VIRGINIA FRAMEWORK FOR ADDICTION ANALYSIS AND COMMUNITY TRANSFORMATION (FAACT)

1. Cross agency data sharing converges 65 data sets from **public health, public safety, and criminal legal system**
2. Delivers **insights** about contributing factors, actionable intelligence, and **enhances timely and effective responses.**
3. Data sharing culture and governance framework **built for resilience**

Moving the needle on retention

MARYVILLE UNIVERSITY

1. IT staff participated in data lake and modern data architecture **skills development**
2. Aggregated student touchpoint data from the **SIS, LMS, and CRM** into a data lake
3. Processing and machine learning to **identify at-risk students** from behaviors
4. Fed insights into communication platform for **early intervention and nudging**

2. Engage the right people

Stakeholder Engagement: “The 5 keys”

Organizational Leadership

Roles with maximum/high responsibility for mission and data within the organization in scope

Examples: Chief Data Officer, Cabinet Secretaries, Mayor’s staff, Commissioners, Legislators

Business Users

Business roles that use data solutions, and roles that facilitate consumption of the solutions by the business.

Examples: Business analysts, program data analysts, frontline operators

Program Management

Business and technical roles that are accountable for a specific data domain(s) within the organization in scope

Examples: Executive Directors or Agency/Department heads, Program Directors, Line of Business leaders

Data Solutions Engineering

Developers that build data products and solutions.

Examples: Director of data engineering, Principal data architect, Principal data engineer, data scientist

IT Infrastructure and Operations

Technical roles responsible for the infrastructure of the data platform and operations.

Examples: Director of IT, Director of engineering, Director of Operations, Principal DevOps Engineer

aws

3. Embrace Best Practices

Best practices to combat fears

Reasonable Fear	Best Practice	Benefit
Vendor lock-in	Microservices	Agility & extensible
Cool new technology	Component Architect.	Safe to change
Dependent monolith	Decoupling	Operational efficiency
Policy change	Loose Configuration	Flexibility & queuing
Staff size	Event Driven Automation	Cost management
Expanded adoption		Machine learning & automation
Replicability	Containers	Standard practices
Traceability	Infrastructure as Code	Logging & auditing
	Metadata	

Best practices for governance

Governance	Best Practice	Benefit
Master Data Management	Data Documentation	• Standardization → Normalization • Hierarchy for resolution & truth
	Attribute Assignment	• Security automation • Ease of update
Security	Data Protection	• Multiple layers of encryption (at rest, in movement) • ML logging – flag anomalies, PII/PHI, avoid error
	Resilience	• No single dependencies • Hardening, replicated data stores • Automated scaling built in elasticity

AWS Public Sector Enablement Programs

How can AWS help you overcome these challenges?

The journey to becoming a data-driven organization

Beginner

- Disconnected data and business strategy
- Reactive business operations
- siloed, fragmented data landscape
- Limited skills

First big challenge

Experimenter

- Data sponsorship growing
- Isolated data and AI initiatives
- Pockets of talent and value
- No connection between data and AI architectures

Harder than anticipated

Adopter

- Senior stakeholders engaged and advocate data
- Cross functional teams forming
- Value being realized, not consistently tracked
- Common data and AI standards, architectures, platforms forming

How to elevate?

Scaler

- Business actively investing in data
- Data as strong part of culture
- Cross functional teams and data-AI communities
- Active skill development
- Standardized governance and architectures integrated across data and AI

Data Driven

- Data and AI as significant part of the business value proposition
- Proactive actions, automated decision making
- Widespread autonomy in innovating with data and AI
- Product teams embedded in the business
- Autonomous use of standardized platforms, ethics, and governance

How can AWS help you on your data journey?

Get peer-level executive guidance

aws Executive Government Advisors

- ✓ Mental models and strategies based on the first-hand experience of former public sector CXOs
- ✓ Get peer-level sounding board and sparring partner

Inspire and accelerate your cultural transformation

Design and build a modern data strategy

aws Data and Digital Transformation

- ✓ Create an organizational vision for innovation with data to drive business outcomes
- ✓ Define the first pilot, learn, and build

Jump-start the modern data strategy flywheel

Develop an analytics Proof of Concept

aws Analytics Acceleration Lab

- ✓ Create tangible deliverables to accelerate strategic databases, analytics, and ML initiatives
- ✓ Leave with an architecture, working prototype, path to production, and deeper knowledge of AWS services

Quickly plan, execute, and see results for analytics POCs

aws

AWS WWPS Executive Government Advisors

Jayson Dunn

Executive Government Advisor

Danielle Hinz

Executive Government Advisor

Procurement

Learn more

Morgan Reed

Sr. Manager, Executive Advisory Team

Carlos Rivero

Executive Government Advisor

Data

Maria Thompson

Executive Education Advisor

Cybersecurity

Next steps to guide you on your data journey

Align key mission stakeholders

aws Public Sector Data Day

- ✓ Educate your high-level stakeholders through exposure to mental models and strategies based on the first-hand experience of former public sector CDOs
- ✓ Qualifying criteria:
  - ✓ Executive Sponsor

Inspire and accelerate cultural transformation

Assess data program maturity

aws Data Strategy Diagnostic

- ✓ Create a common understanding among mission and technology leaders of your organizations maturity level across 4 perspectives:
  - ✓ Mindset
  - ✓ People
  - ✓ Process
  - ✓ Technology
- ✓ Qualifying criteria:
  - ✓ Executive Sponsor
  - ✓ Mission Program Engagement

Know where you are on your data journey

Accelerate governance and engagement

aws Data Governance Accelerator

- ✓ Identify the right roles to participate in a governance structure appropriate for your organization
- ✓ Develop the charters and necessary decision-making processes
- ✓ Establish communication and escalation mechanisms
- ✓ Qualifying criteria:
  - ✓ Executive Sponsor
  - ✓ Mission Program Engagement
  - ✓ Defined Mission Use Case

Engage the right people in the right processes

aws

Key Concepts Learned

Modern data strategy is more than technology

Use cases that solve real world problems build the path to data program maturity

Challenges resulting from legacy thinking and approaches

Implement a technical architecture facilitates the flow of data and intelligence to stakeholders

Manage your operating environment to support access and security compliance and transparency

Think Big
Start Small
Scale Fast

Closing Thoughts

- **The Struggle is Real:** The amount of effort you are expending today to get this work done in your current environment is monumental.
- **Technology has evolved, so should your data strategy:** Understand how advancements give you fine grained control of the component parts of the data lifecycle.
- **People and Process have to evolve as well:** Don’t apply new technology to old ways.
- **Have the Right People in the room:** Engage the right leaders to participate.
- **Think Big:** In the public sector space, lives depend on it!
- **Start Small:** Pick a first use case. Rules of the road don’t change for the next one.
- **Scale Fast:** Tackle first use case and deliver, others will line up at your door.
- **Built for Change:** Flexible framework means you can pull in the next new cool tech.
- **Iterate:** Set a cadence to revisit your approach regularly.

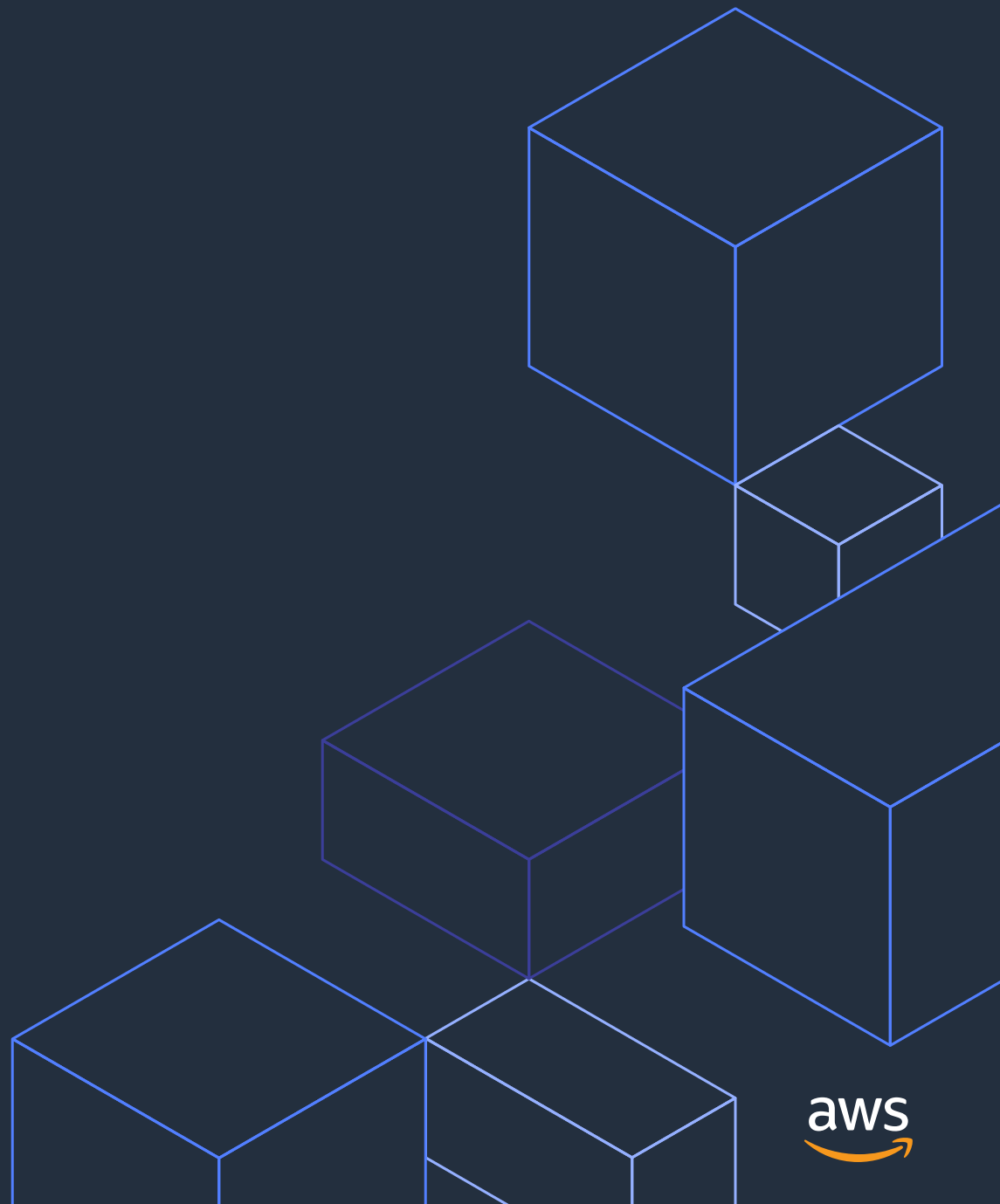
Next Steps

- What are the **goals** for your data program?
- What **use cases** can you identify that will help you accomplish your goals?
- Who are the right **stakeholders** to engage?
- How can you leverage the **AWS Public Sector Enablement Programs** to help your organization become data driven?

Please Provide Your Feedback

How did I do?
What can I do better?

Questions?



Interested in learning more?

Let’s Connect!

aws

Thank You!

Carlos Rivero
Executive Government Advisor
Worldwide Public Sector
rivercap@amazon.com