



PUBLIC SECTOR SYMPOSIUM

BRUSSELS | MARCH 28, 2023

BBT102

Drive innovation with Data

Data driven organization – the journey to value

Samer Madfouni

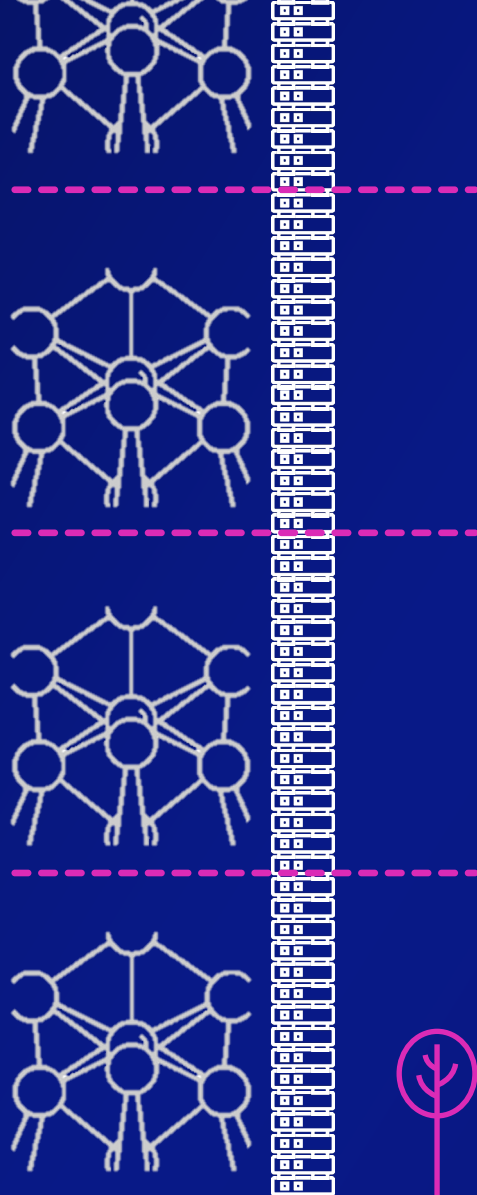
Head of Data – EMEA Public Sector
Amazon Web Services (AWS)

Marco Amabilino

Head of the Digitalisation Department
European Parliament



96 ZB
of data created in 2022



x **11.3 Million**



Why does it matter?



“Most of the CEOs I talk to regularly are focused on transforming their orgs, and part of that discussion is how to put data at the centre of decision making”

Adam Selipsky

CEO, AWS



“Customers are transitioning from the power of the cloud to the power of data”

Matt Garman

SVP Sales & Marketing, AWS



Why does it matter?



Make better decisions, faster



predicts production issues months in advance



turns insights into competitive advantage on the racetrack

Why does it matter?



Make better decisions, faster



Respond better to the unexpected

vyaire
MEDICAL

scales ventilator
production by 20x while
improving first pass
yield 3x

moderna

accelerates vaccine
development and scales
manufacturing

Why does it matter?



Make better decisions, faster



Respond better to the unexpected



Create better customer experience



brings fans closer to the action using data and ML



deepens customer engagement with personalized customer service

Why does it matter?



Make better decisions, faster



Respond better to the unexpected



Create better customer experience



Uncover new opportunities



expands to 59+ countries since launch



creates digital athlete to improve player safety

Why does it matter?



Make better decisions, faster



Respond better to the unexpected



Create better customer experience



Uncover new opportunities



Improve efficiency

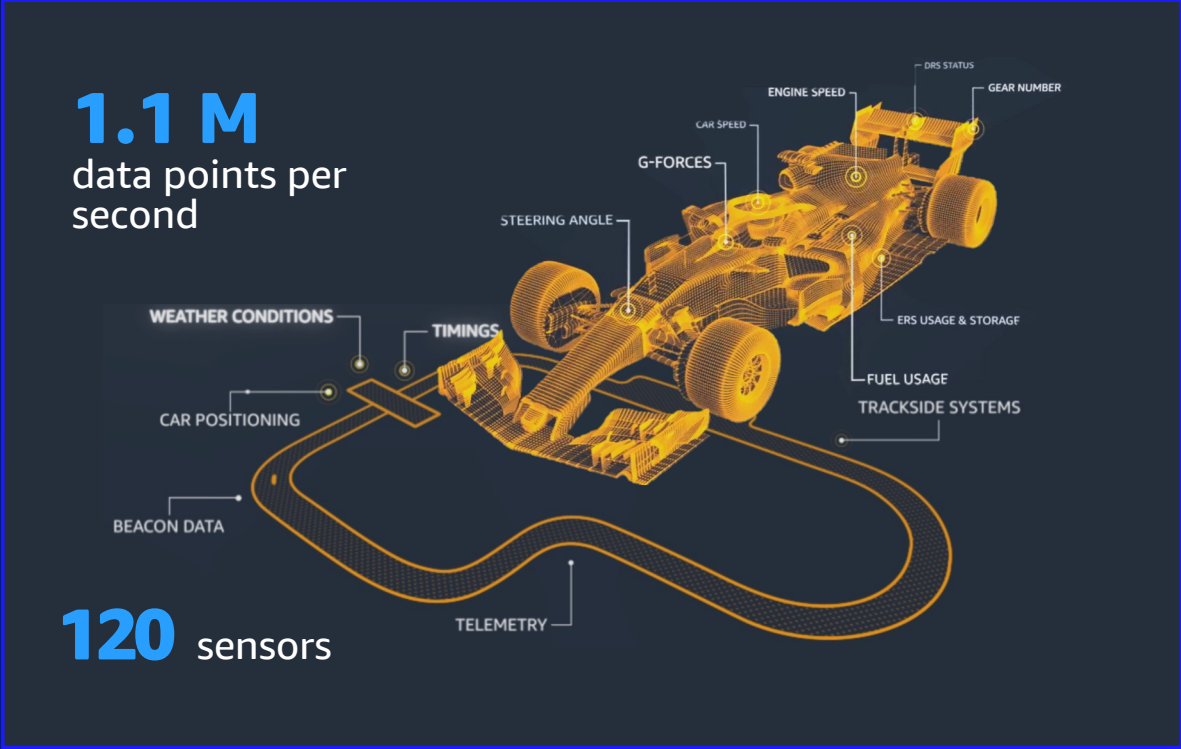
amazon.com

eliminates 915K tons
of packaging

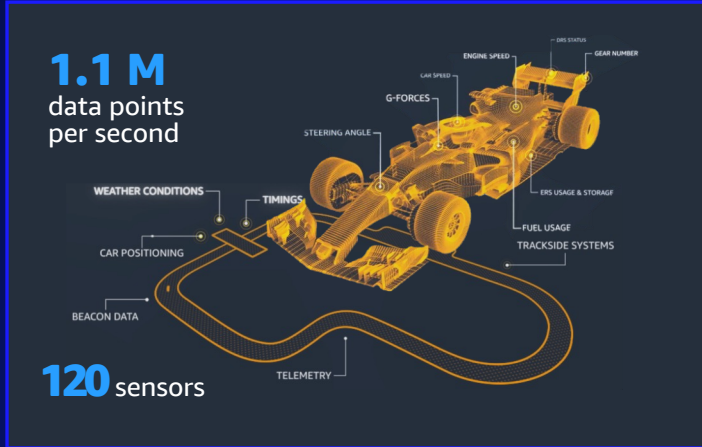
Expedia

reduces manual touch
rate for partner payment
from 17% to 0.06%

Data is the fuel for F1



Data is the fuel for F1



The business value of a modern data strategy

By making 10% more data accessible, a typical Fortune 1000 company will see a **\$65 million increase in net income**¹

415%
five-year ROI²

48%
reduced total cost of operations²

¹Read the article on [Forbes.com](https://www.forbes.com) | ²Read the [report](#)



And yet . . .

Even though **99%** of the 85 leading companies are investing in data, **only 24%** report being successful in creating a data-driven organization.

How to realize the vision of a **data-driven** organization?

Organization data evolution



How to realize the vision of a data-driven organization

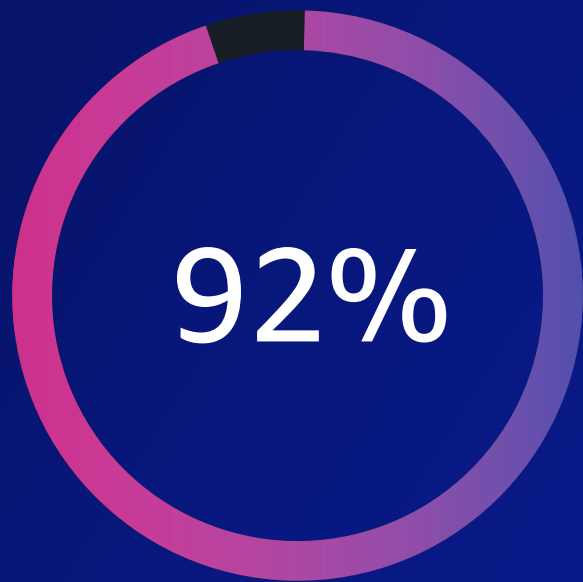
- Proclaim that “data is the new oil”
- Launch several big data initiatives, including an AI project
- Start by hoarding as much data as possible
- Hire a chief data officer (CDO) but change little else
- Create centralized function and funnel all data requests to them
- Repeatedly say that we are in the data business
- All of the above



Culture



Facts, not feelings!



of companies cite culture – people, process, organization, and change management – as the biggest impediment

Source: Big Data and AI Executive Survey 2021 by NewVantage Partners



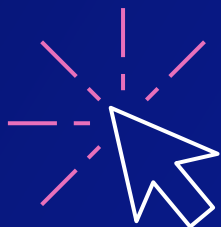
Engage

Executive sponsorship

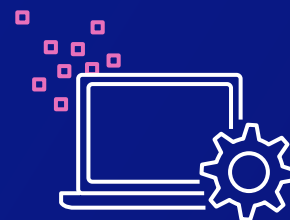
Executive engagement

Data to justify decisions

Data to guide decisions



Engage



Enable

Democratize access to data

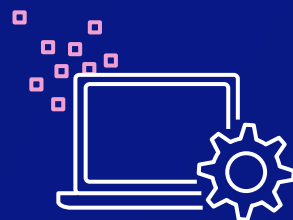
Democratize actions using data

Make data easy to use

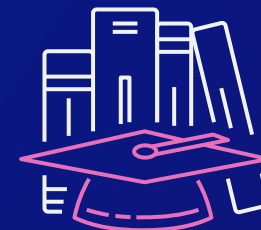
Make data enjoyable to use



Engage



Enable



Educate

Selective data literacy

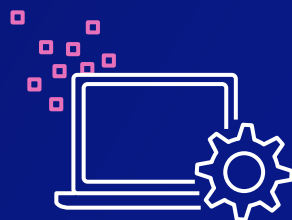
Data proficiency as a core skill

Lack of common definitions

Common understanding and vocabulary



Engage



Enable



Educate



Eliminate

Data as a departmental property

Data as an organizational asset

Reporting ghost towns

Data integrated into everyday workflows

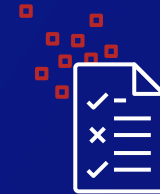


Mechanisms

Mechanisms work, not just good intentions



Governance



Data quality

Governance



Govern to enable, not restrict

How?

Good metadata = good data catalog

How?

Good metadata = good data catalog

Automate your bureaucracy

How?

Good metadata = good data catalog

Automate your bureaucracy

Encrypt everything

How?

Good metadata = good data catalog

Automate your bureaucracy

Encrypt everything

Centralize implementation, decentralize decisions

The modern data community

Producers

Teams that want to share data



Domain expertise

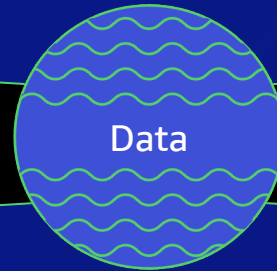
Data ownership and governance

Data quality

Metadata management

Marketplace

Team that runs the marketplace



Build security controls

Build and run the platform

Simplify onboarding

Enterprise datasets

Training and community

Consumers

Teams that want to use data



Execute business priorities

Business analytics development

Data discovery

Data pipeline development

Creation of new insights

Level of decentralization depends on maturity of skills, complexity of business, domain knowledge required, and pace of tech change

Amazon DataZone – data products at scale

Amazon DataZone enables **governed sharing of data products** across one or more organizations, departments, and data lakes



DATA PRODUCERS

Teams who want to share data

AMAZON DATA ZONE

Team who runs the data marketplace

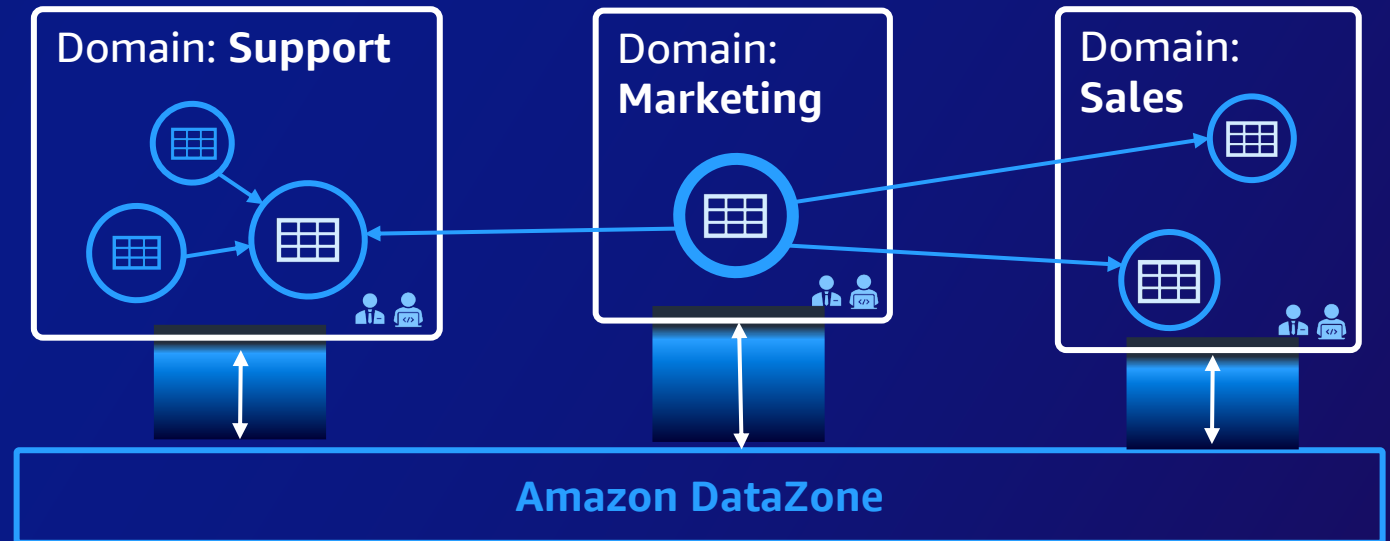
DATA CONSUMERS

Teams who want to use data

Decentralized and federated ownership

Data mesh

- Data-driven **agility** through increased **ownership and autonomy**
- Domain-local **governance** through **lightweight** centralized policy
- Isolate data resources with clear **accountability**
- Data a **product** that can exist in **any system**





160,000 employees, **40** business units in **70** countries

Built a **Common Data Hub** in AWS and over **1,000** projects

+200 TB of data, **+12K** datasets, **+2K** data consumers

Automated energy predictions

Maximized **wind farm** **energy** production



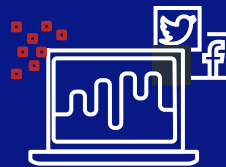
Data quality



Quality by design

Data quality

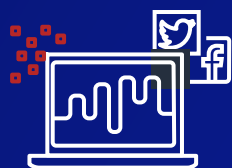
Garbage in, garbage out



Incentivize better quality
at the source

Data quality

Garbage in, garbage out



Incentivize better quality
at the source

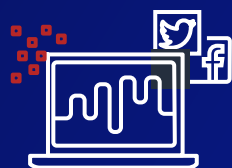


Ask for what you need,
not everything

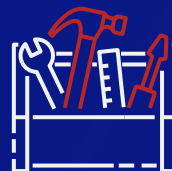


Data quality

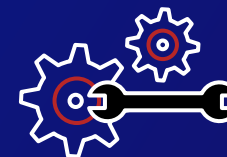
Garbage in, garbage out



Incentivize better quality
at the source



Ask for what you need,
not everything



Automate inferential data
and quality checks

Data quality

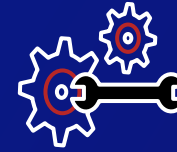
Garbage in, garbage out



Incentivize better quality at the source



Ask for what you need, not everything



Automate inferential data and quality checks

Better data



Better insights



The data network effect



Better outcomes



Better actions



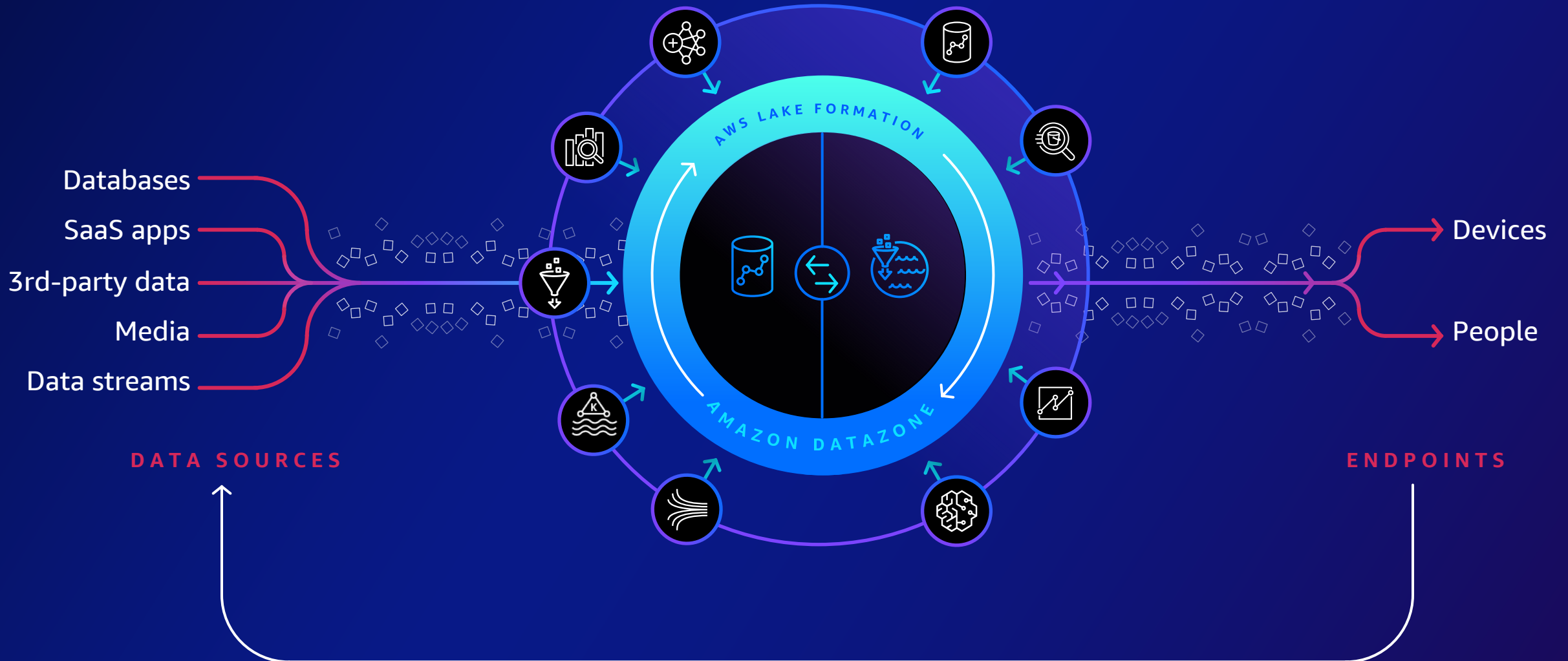
Execution

Think big but execute small

Traditional approaches don't scale



End-to-end data journey



Modern data strategy on AWS

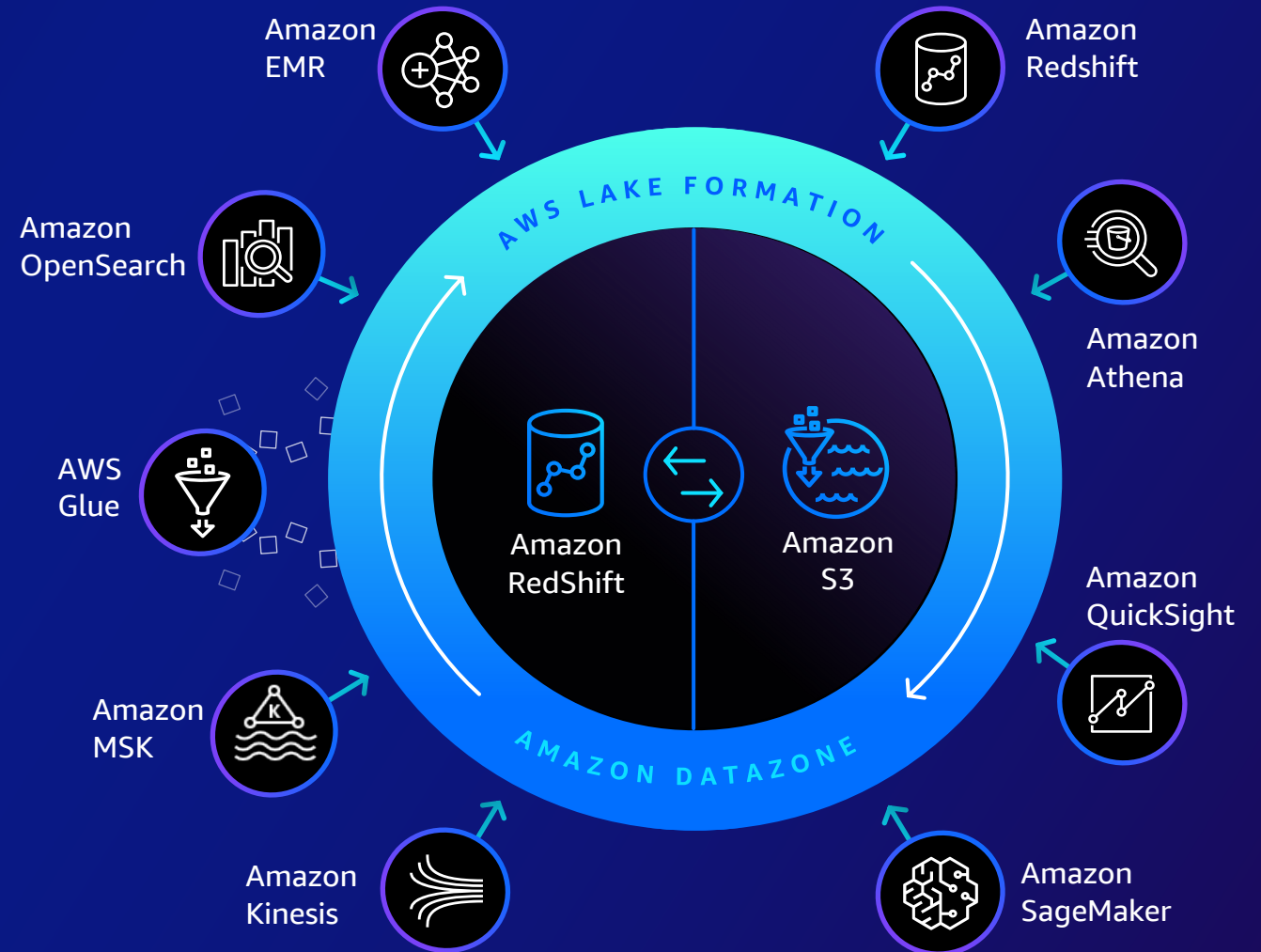
Data at any scale

The best price-performance

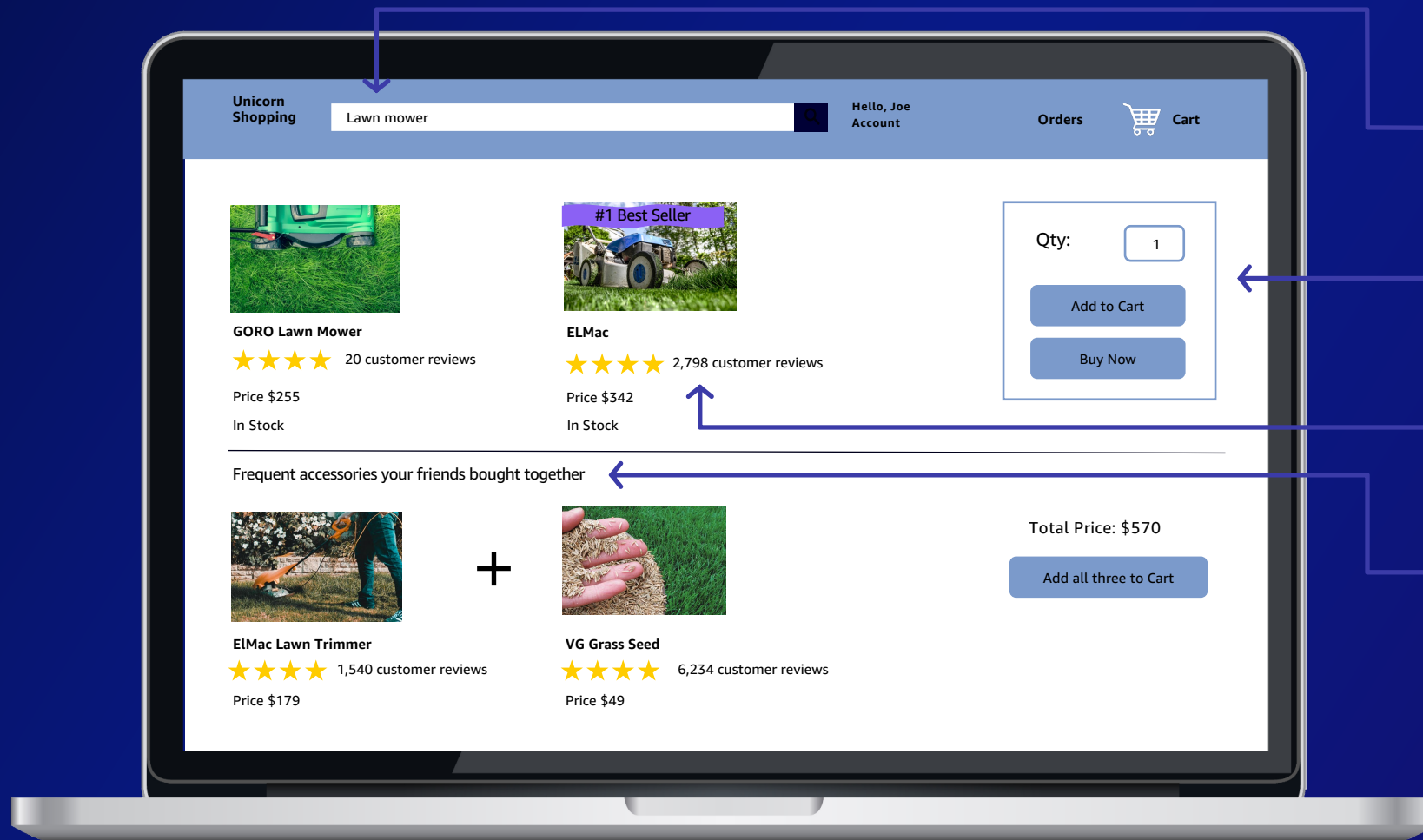
Seamless data access

Unified governance

AI and ML to solve business challenges



Modern cloud-based applications



Search
Index optimized store

Shopping cart
Relational database

Customer reviews
Key-value database

Recommendations
Graph database



European Parliament

EP Archive Dashboard - fast digital access to over 850,000
parliamentary documents

The European Parliament (EP)

THE EUROPEAN PARLIAMENT (EP) IS THE ONLY EU INSTITUTION DIRECTLY ELECTED BY EUROPEAN CITIZENS

- 705 MEPs
 - 26 Committees
 - 7 Political groups
 - 45 Delegations
- Parliament's Secretariat
 - 8132 total number of staff (May 22)



The EP Historical Archives

THE EUROPEAN PARLIAMENT (EP) IS THE ONLY EU INSTITUTION DIRECTLY ELECTED BY EUROPEAN CITIZENS

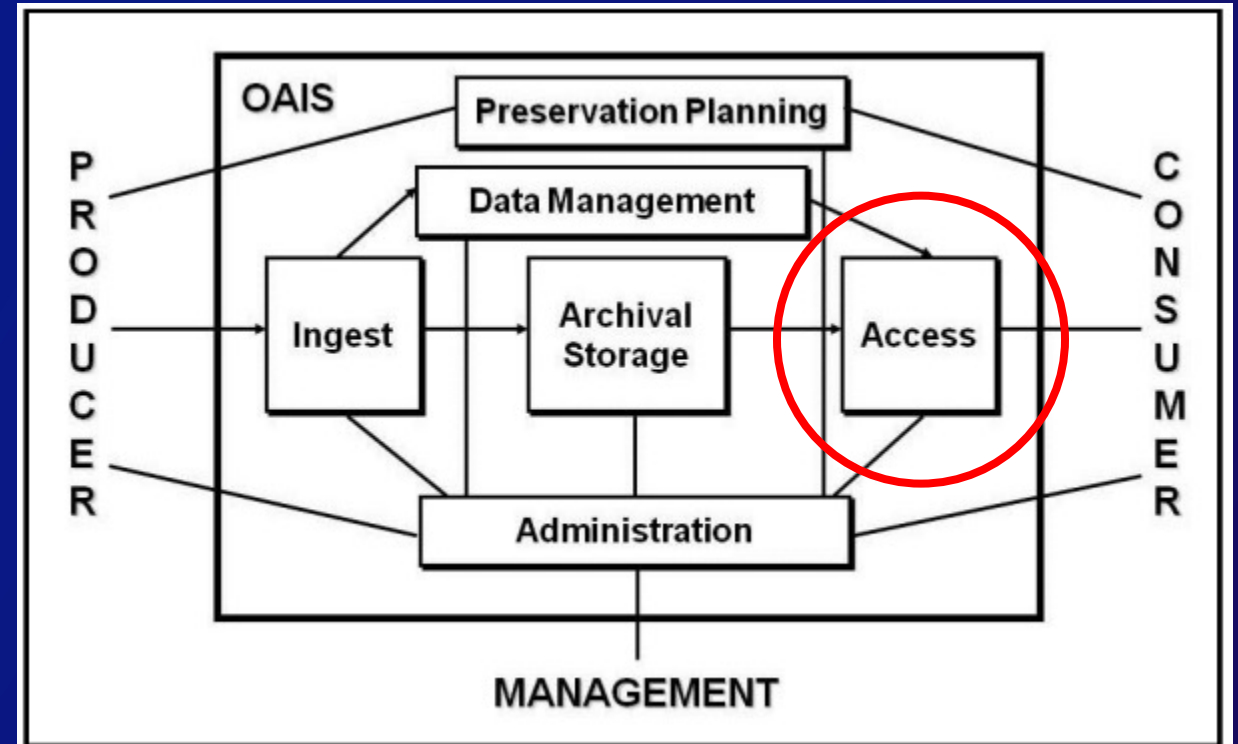
- The archives function was created in 1952 by the Common Assembly
- Located in Luxembourg
- Holdings: 3.282 km of physical documents
- > 1TB of digital documents
- Stock rooms: 7 main rooms
- Archive management system: ArchiDoc (INETUM) with > 5 million references



Archive management

AS PART OF THEIR SERVICE TO THE PUBLIC, THE EP ARCHIVES ENABLE RESEARCHERS AND CITIZENS TO VIEW ARCHIVED DOCUMENTS

- **Ingest:** Preparing documents for storage and management within the archive
- **Archival Storage:** storage, maintenance and retrieval of documents. Digitization of documents to PDF/A files
- **Access:** search and retrieval of documents
- **Data Management:** administration of the database containing documents metadata
- **Administration:** day-to-day operation of the archive



EP Archives Dashboard objective

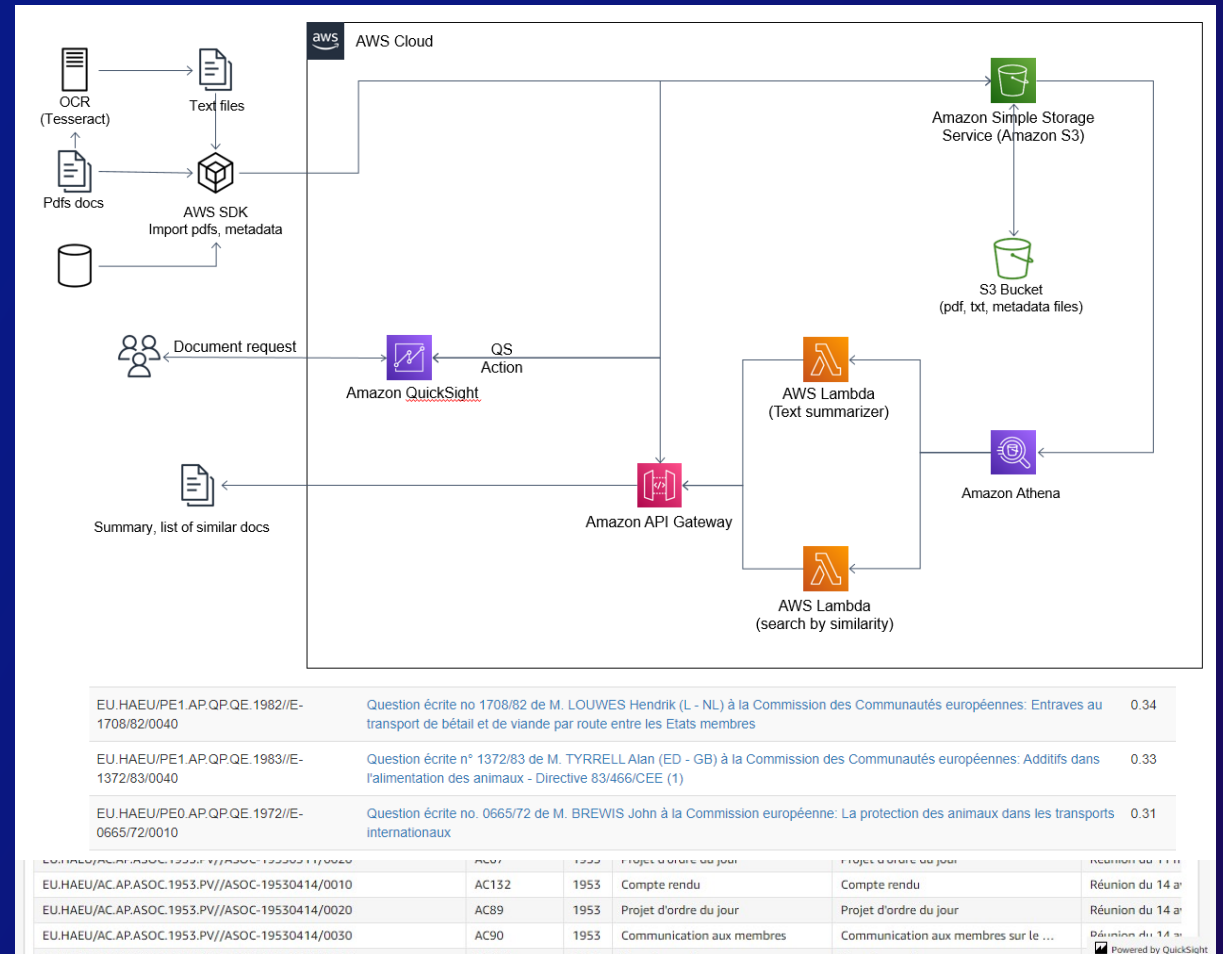
DEVELOP AN INTUITIVE WEB INTERFACE TO SEARCH AND RETRIEVE PUBLIC ARCHIVED DOCUMENTS

- Find archived documents without complex search rules
- Visual tools to search documents
- Access to stored documents
- Call to summarization and search by similarity functions (Parliamentary Questions)
- Availability of import tools capable of reading large csv files (metadata)
- Stable and quick response times

Solution architecture

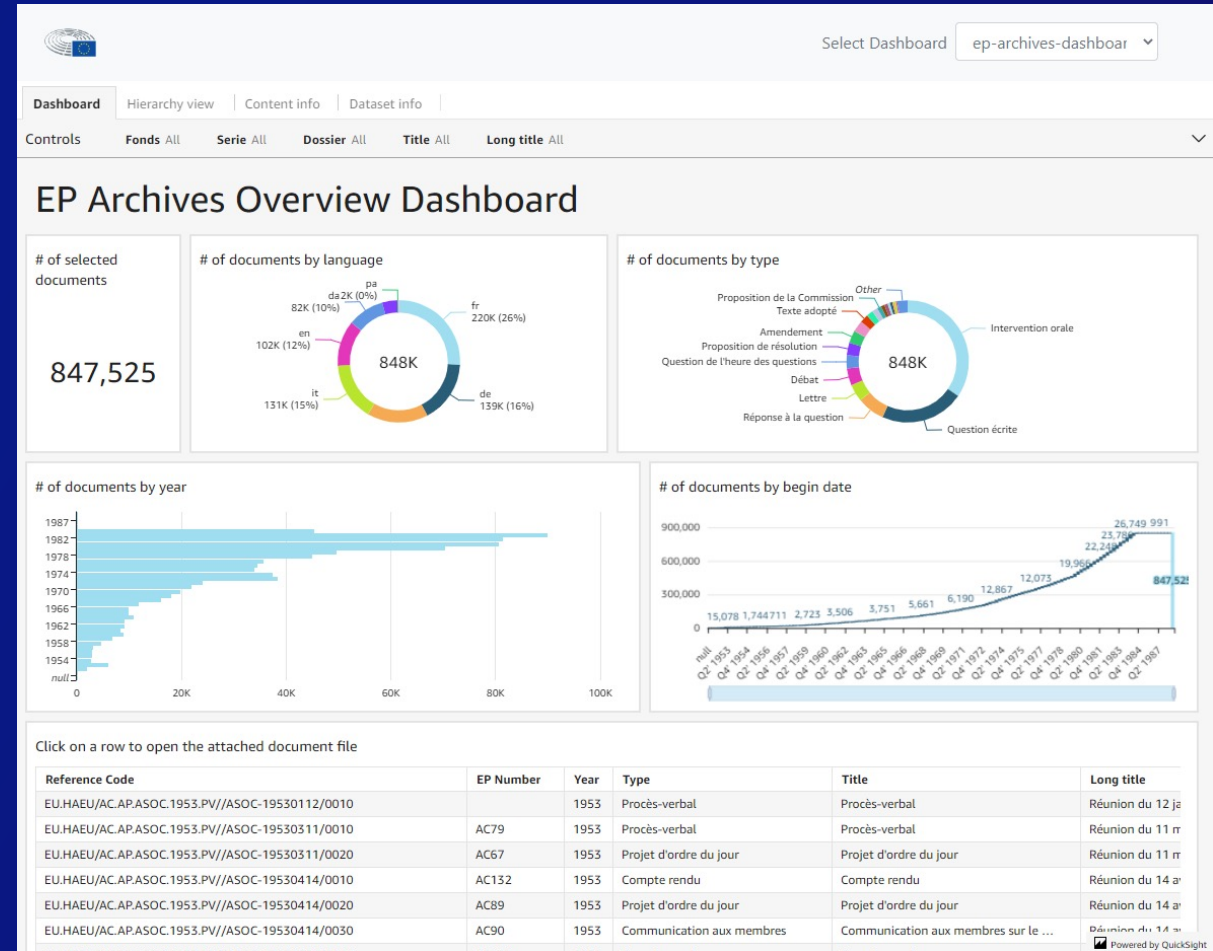
WE NEEDED A SCALABLE SOLUTION, QUICK TO IMPLEMENT, WELL DOCUMENTED WITH LIMITED MAINTENANCE COSTS

- Amazon QuickSight to visually search documents.
- Amazon S3 to store pdf and txt files.
- AWS Athena to query metadata.
- AWS Lambda to implement text summarization.
- AWS Lambda to implement similarity search.
- Amazon API Gateway to publish EP Archives API.



Results

- 850,000 public documents now available on the internet (1952-1984)
- Text extracted from 45,000 Parliamentary Questions
- 4113 document retrievals using the dashboard compared to the 227 requests received through the standard channels in 2022
- Increased users satisfaction
- Unexpected result: the dashboard is actively used internally to improve metadata quality

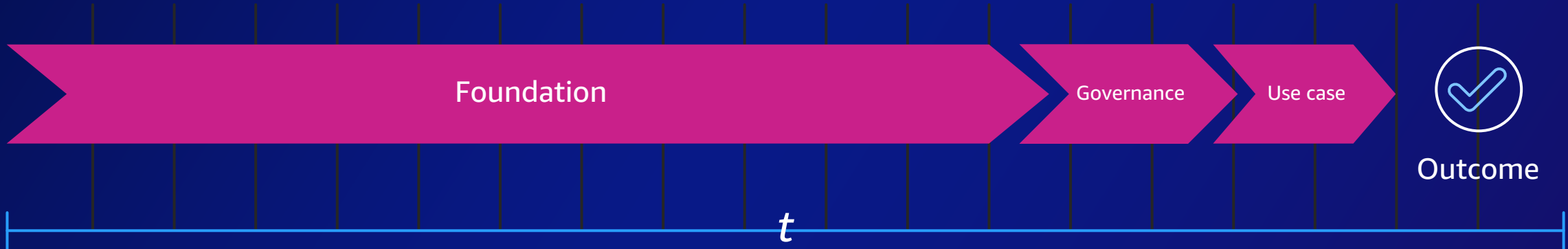


Try the Dashboard by scanning:

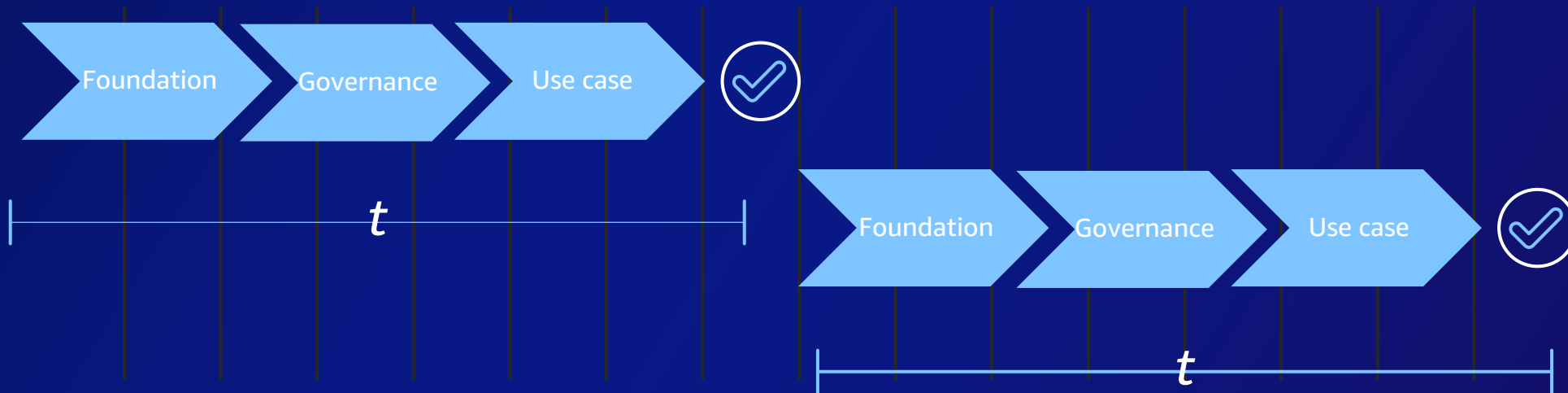


Time to value

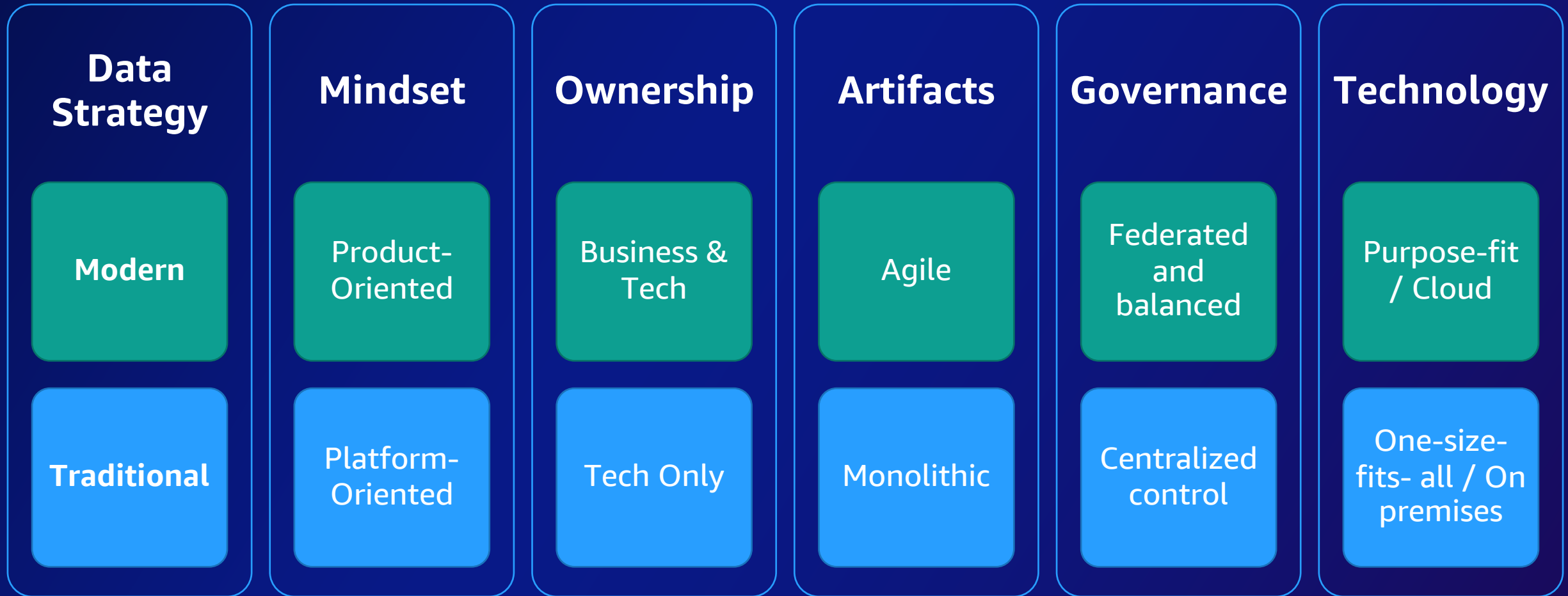
Traditional approach



Iterative approach



Modern vs. traditional data strategy



Want to build a data vision and strategy?



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

Jumpstart the data flywheel

Have a strategy and need help executing it?



- ✓ Joint engineering engagements between customers and AWS technical resources
- ✓ 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

Come with an idea, leave with a solution

Build skills to unlock the value of your data with AWS training and certification

Explore 180+ relevant trainings including:

Building Modern Data Analytics Solutions on AWS

(new collection of Classroom Trainings)



Data Analytics Fundamentals



Get AWS Certified



Talk to your AWS account team to learn more!

Thank you!

Samer Madfouni

 samerjm@amazon.ae

 <https://www.linkedin.com/in/samer-madfouni/>



Please complete the session survey in the mobile app