

AI-assisted data product engineering on AWS

The Thoughtworks Data Product Workbench

 thoughtworks



Thoughtworks at a glance

Core capabilities

- Legacy app & mainframe modernization
- New product development
- Data modernization: data-as-a-product and migration
- Enterprise AI: AI product build and AI strategy; In-house AI research Lab

Industries served

- Retail, CPG, travel and transport
- Automotive and manufacturing
- Software & technology (incl. gaming & media)
- Life sciences & medtech
- Banking and financial services

Our differentiators

AI/works™ platform

Thoughtworks' proprietary agentic development platform, built to accelerate legacy modernization and net-new product delivery. It draws on global context libraries and covers the full SDLC, from development through to operational runtime.

3/3/3 delivery model

3 days Product design

3 weeks Prototype

3 months MVP

All with fixed-price outcomes



Featured clients



Why Thoughtworks

1,400+ active AWS
certifications worldwide



**AWS Global Partner
of the Year 2025**
Data & Analytics



- Agentic AI Consulting Services
- Generative AI Consulting Services

23+
AWS competencies
and specializations

80% of our staff are
engineers building
cloud-native, design-led
systems

10+ years of AWS x
Thoughtworks
co-innovation, deploying
since the early days of
cloud

**Gartner Magic
Quadrant™**
Visionary winner in 2025,
2024 and 2023

10,000+

Thoughtworkers globally

18

countries

30+

years of engineering excellence

We didn't adopt data products. We **pioneered** them.

Data mesh and the data-product pattern the industry now follows originated at Thoughtworks. We have delivered them in production for years — and clients still come to us for advice on how to do it well.

We defined the category

- Coined data mesh and the data-product architectural pattern.
- Authored the practices teams now treat as the standard.

Proven in production

- Repeatable data-product delivery across industries and at enterprise scale.
- Deep engineering, platform and governance capability in one team.

The advisors others call

- Clients return to us to shape their data strategy and operating model.
- Recognized thought leadership in data mesh and data products.



Our point of view on

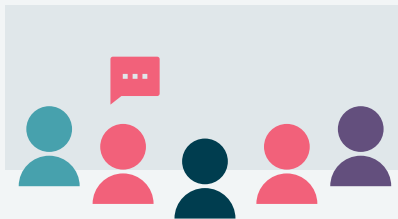
AI - Ready Data



AI at scale runs on data you can trust

Most AI programs stall before production, and the cause is rarely the model.

Most AI dies in the POC



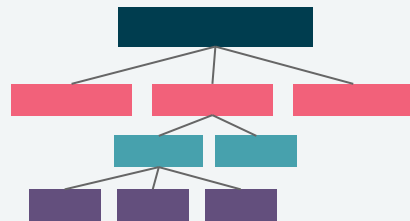
Without governed, trustworthy data, promising pilots never reach production.

Agents need context, not just access



They need clear meaning, shape and quality guarantees to act safely.

Ungoverned data misleads

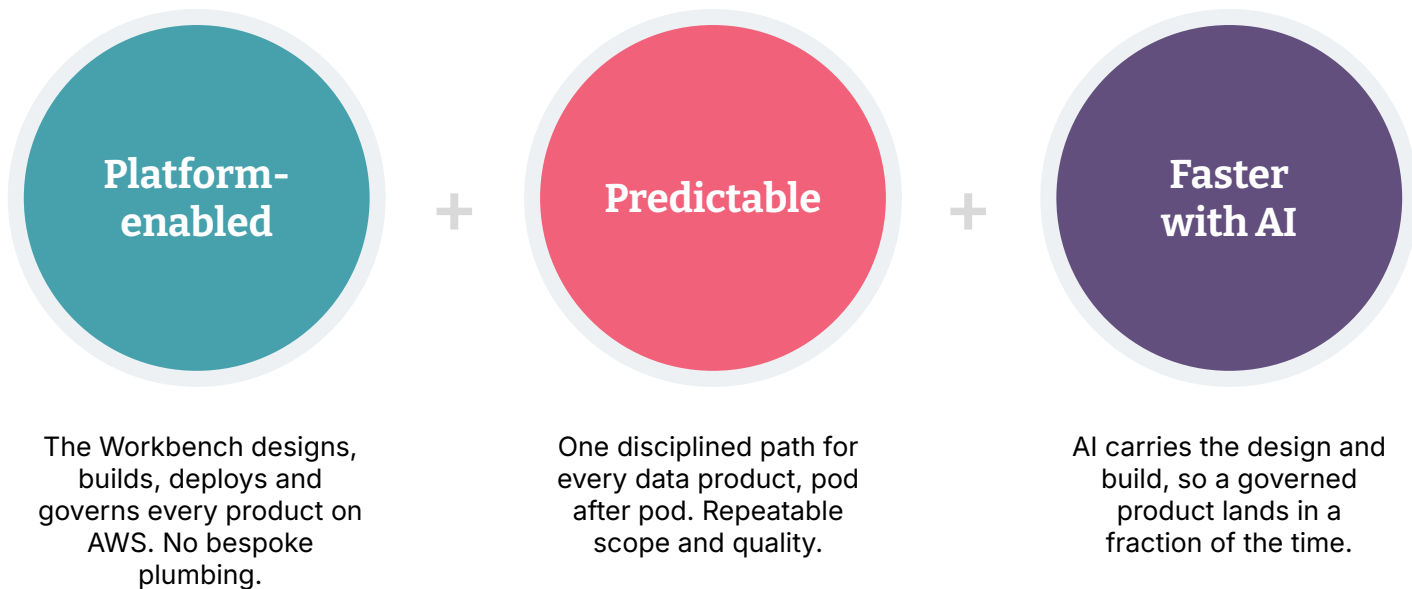


It produces confident, wrong answers that erode trust in every AI initiative.

Trust is the bottleneck. Governed, reusable data products are how you clear it.

Data products are the foundation for AI at scale

Three shifts make governed data products deliverable today.



Proven outcomes with the Workbench

Delivered in production, across industries.



60 to 80%

Faster delivery



Hours

Idea to deployed product



8 to 10

Data products per use case



Native

Databricks on AWS



What this means for you

Faster time to value from your existing AWS investment, enterprise-grade governance by default, and data that both your teams and your AI agents can trust.

Data Product Workbench

Solution overview



Where AI meets data engineering

An AI-assisted engineering environment for governed data products.



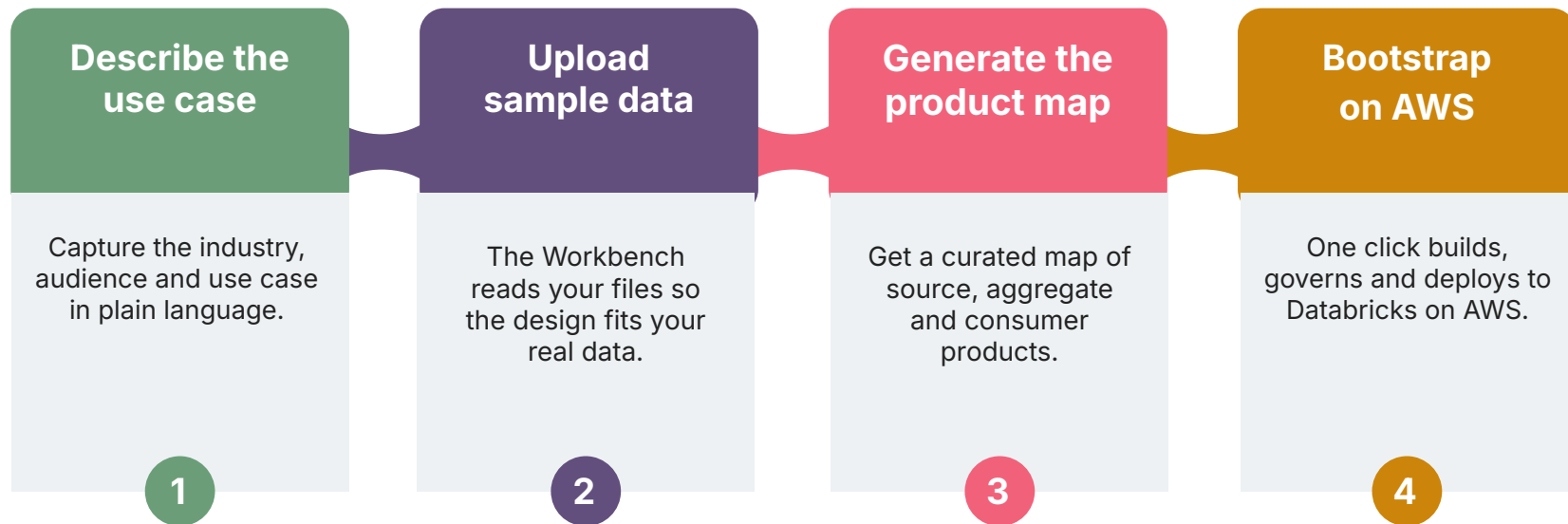
You describe a business use case and add sample data.

The Workbench designs a full set of source, aggregate and consumer data products, then builds and deploys them directly onto Databricks and AWS.

One click provisions the catalog, applies governance and publishes API and MCP output ports. It uses open standards throughout, so everything it produces is yours to own and portable.

How it works

From a plain-language brief to a live, governed product on AWS.



What makes the Workbench different

Built for AI-ready data, native to AWS.

AI-assisted design

An AI copilot proposes designs, standardizes metadata and defines SLOs.

Open standards

Open Data Product Standard specs and open data contracts keep you portable.

Native to Databricks on AWS

Products deploy to Unity Catalog with governance provisioned automatically.

Agent-ready by default

Every product exposes API and MCP ports for human and AI consumption.

Built on AWS

The Workbench generates and deploys products natively onto your AWS estate.



Inputs

Use case + sample data



Workbench

AI-assisted design + specs (ODPS/ODCS)



CI/CD

GitHub Actions pipeline



Databricks on AWS

Unity Catalog, RBAC, medallion



Output ports

API + MCP

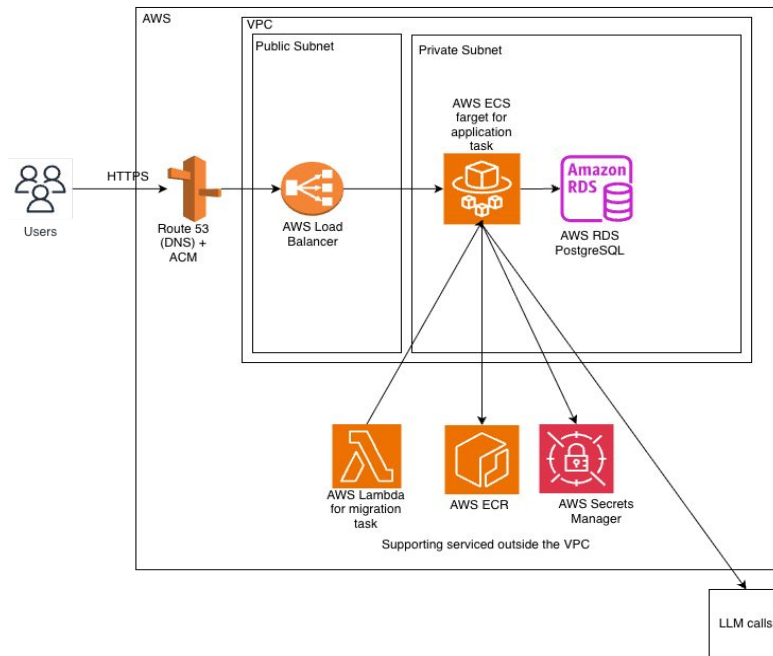
Consumers

Analysts and BI tools, applications, and AI agents all consume the same governed product. Governance, masking and lineage are applied as the product deploys. No migration, no lock-in.

Data Product Workbench deployment on AWS

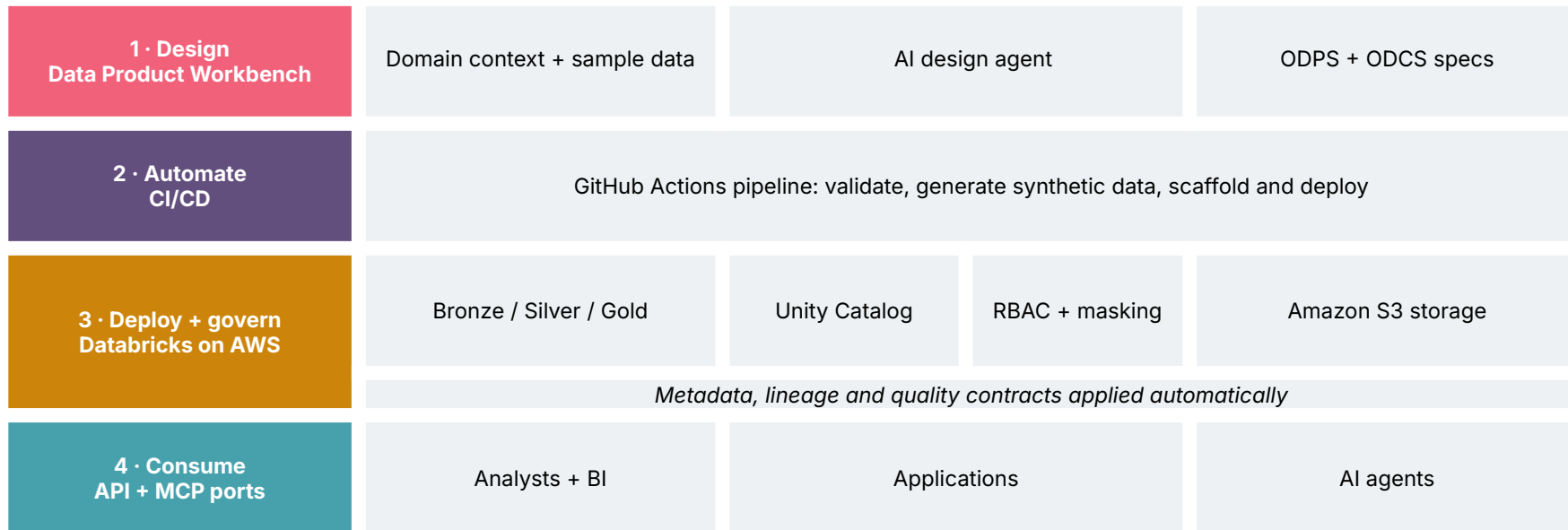
Technologies

- **AWS ECR** — Private container registry storing the Workbench Docker image, pulled by ECS on deployment.
- **AWS RDS (PostgreSQL)** — Managed Postgres 16 database for application data.
- **AWS Secrets Manager** — Stores LLM API keys, JWT signing secret, DB credentials, and admin credentials, injected securely into containers.
- **AWS ECS Fargate** — Serverless container compute running the app and one-shot DB migration tasks.
- **ALB** — Internet-facing Application Load Balancer routing HTTP/HTTPS traffic to ECS tasks.
- **VPC** — Isolated network with public/private subnets, NAT/Internet Gateway, and security groups controlling traffic between tiers.
- **Route 53** — DNS service wiring a custom subdomain to the ALB, upserted by the release workflow post-deploy.
- **ACM (Certificate Manager)** — Issues/manages the TLS certificate attached to the ALB's HTTPS listener.
- **AWS Lambda** — Deployment-time function that runs and polls the ECS DB-migration task as a CloudFormation custom resource.
- **LLM** — Pluggable AI provider (e.g Anthropic Claude, or OpenAI) powering the design and polish agents.



Reference architecture on AWS

How the Workbench designs, deploys and governs data products on Databricks and AWS.



See it in action

A live underwriting profitability use case, from brief to a deployed product.

1

Describe the use case

Banking and financial sector, actuaries and underwriting officers.

2

Upload sample data

Drag and drop 7 CSVs of raw source data.

3

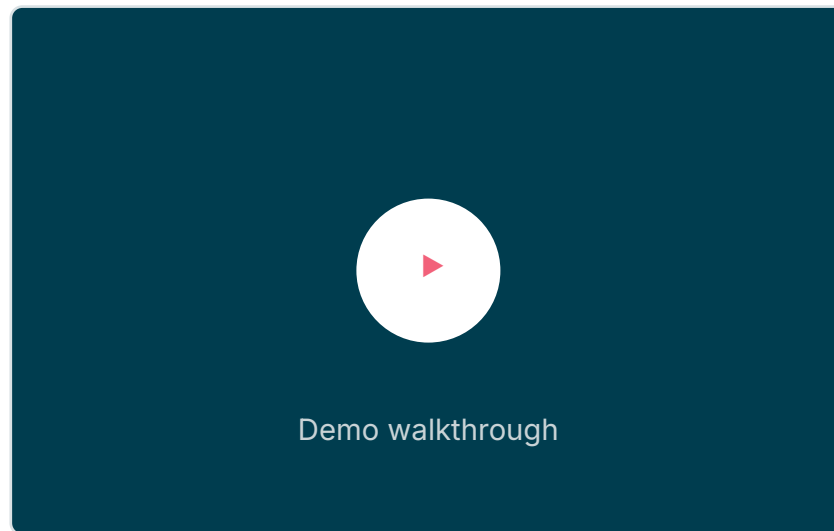
Generate the product map

Source, aggregate and consumer data products in seconds.

4

Bootstrap onto AWS

Deploys to Databricks on AWS with API and MCP ports.



Host the video on the site and add the link or a QR code here.

Client stories

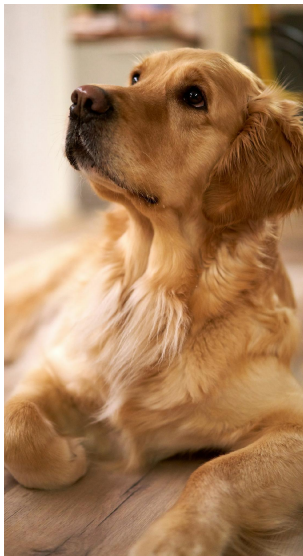
Proof in production





CLIENT STORY · ZOETIS VMRD

Data product workbench for reliable, **agent-ready** data delivered quickly



- **70–80%**

Faster delivery

First-cut implementation — 2–2.5 wks, was 8–10 wks

- **5 hrs**

Prompt to deploy

Spec + pipeline, was 1–2 wks

- **Same sprint**

Enhancement turnaround

Was a new cycle

— Background

At **Zoetis VMRD**, every data product was hand-built — bespoke specs, manual schema design and custom pipeline code, with no reuse or standards. Discovery and design alone consumed **60–70% of sprint capacity** before a line of transformation code was written, and quality varied widely across pods.

— Solution

- Built and deployed the **data product workbench** — a design + code-generation toolkit integrated with Open Code and GitHub Actions
- An **AI copilot** generates specs, attribute definitions and medallion-layer transformation code from domain prompts in minutes
- **Automatic CI/CD to Databricks** provisions Unity Catalog entries, permissions and output ports — no manual platform steps
- An **AIFSD guardrail model** governs AI-assisted changes to generated code, cutting per-feature cycle time

— Outcomes

- **70–80% faster first-cut delivery** — design to a working product in 2–2.5 weeks, down from 8–10
- A full **spec and pipeline generated from a prompt in ~5 hours**, replacing 1–2 weeks of manual work
- **Consistent, reusable contracts** across every pod — standardized schemas, lineage and quality SLOs at the platform layer
- Data products are **agent-ready**: MCP output ports let AI agents discover and query VMRD data autonomously

Appendix

Next steps for AWS Sellers



Why should an AWS seller care?

The Workbench drives durable AWS consumption and unlocks AI workloads.

Grows AWS and Databricks consumption

Every data product runs and scales on Databricks on AWS, Amazon S3 and core services.

Unlocks the AI workloads

Governed, AI-ready data is the prerequisite for Bedrock, SageMaker and agentic use cases.

Transactable in Marketplace

Procured through AWS Marketplace, drawing down committed spend.

Land and expand

One use case grows into a platform and enterprise-wide data transformation.

Aligned to AWS motions and funding

Easy to co-sell, and you do not have to lead delivery.

Grows AWS and Databricks consumption

Every data product runs and scales on Databricks on AWS, Amazon S3 and core services.

Grows AWS and Databricks consumption

Every data product runs and scales on Databricks on AWS, Amazon S3 and core services.

Grows AWS and Databricks consumption

Every data product runs and scales on Databricks on AWS, Amazon S3 and core services.≈

Expertise not required. Thoughtworks leads discovery, delivery and enablement. Your team opens the door and we prove the value together.

Available on AWS Marketplace

Procure the Workbench through your existing AWS agreement.

The banner features the Thoughtworks and AWS Partner logos at the top left. Below them, the text 'DATA PRODUCT WORKBENCH' is displayed in a small, white, sans-serif font. The main headline, 'AI-ready data products on AWS, in days', is prominently displayed in a large, white, sans-serif font, with 'AWS' highlighted in orange. A sub-headline in a smaller white font reads: 'Design, build and deploy governed data products on Databricks and AWS.' On the right side, a flow diagram illustrates the process: 'Use case' (grey box) connects to 'Sample data' (pink box) and 'Data product map' (grey box). Both 'Sample data' and 'Data product map' connect to 'Specs - ODPS' (grey box), which then connects to 'Databricks on AWS' (orange box). At the bottom left, an orange button contains the text 'Available on AWS Marketplace'. To its right, the text '60 to 80%' is shown in orange, with 'Faster delivery' in small white text below it. Further right, the word 'Hours' is in orange, with 'Idea to deployed' in small white text below it.

thoughtworks | **aws**
PARTNER

DATA PRODUCT WORKBENCH

AI-ready data products on AWS, in days

Design, build and deploy governed data products on Databricks and AWS.

Available on AWS Marketplace

60 to 80%
Faster delivery

Hours
Idea to deployed

How to buy

- Private offer against your committed AWS spend
- Draws down your AWS commitment
- Procurement in days, not months
- One use case, then expand

Search "Thoughtworks Data Product Workbench" in AWS Marketplace.

How to engage

Five steps — start with a call

1

**Call them
early**

2

**Joint
discovery**

3

**Opportunity
shaping**

4

**Funding
alignment**

5

**Joint
pursuit**

Learn more about our AWS Practice

<https://www.thoughtworks.com/en-gb/about-us/partnerships/cloud/aws>

Your Thoughtworks AWS alliance contacts | partnerships@thoughtworks.com for broader enquiries



Haowei Guo,
Global Head of AWS Alliances
haowei.guo@thoughtworks.com



Cody Carta-Bomersine,
Partner Acceleration (ATX)
cody.cartabomersine@thoughtworks.com



Adam Lang
Partner Sales for APAC
adam.lang@thoughtworks.com



Simone Thompson,
GVP, Strategic Partnerships
simone.thompson@thoughtworks.com



Alagu Perumall
Partner Sales (EU & IME)
alagup@thoughtworks.com



Andre Loureiro
Partner Sales for LATAM
andre.loureiro@thoughtworks.com



Specialized Funding for AI and Cloud Modernization

Discovery Phase

Assessment

AI Assessment | PoC | Catalyst Program

Acceleration Phase

GenAI Adoption

GenAIIC | Production Ready | SCA GenAI

Business Transformation

MAP

Migration Acceleration Program | VMware Modernization

Investment for Validation

- **Program Focus:** Identify business opportunities, build the business case, and define the cloud or AI adoption roadmap.
- **Program Structure:** Initial consulting phase to reduce the customer's upfront costs before starting an implementation or a large-scale migration.
- **Eligibility:** AWS-validated account, aligned business case, and funding approval according to the program's criteria.

Investment for Innovation

- **Program Focus:** Accelerate the development of Proofs of Concept and early production implementations of Generative Artificial Intelligence on AWS.
- **Program Structure:** Co-investment in professional services to validate and accelerate GenAI initiatives with high cloud consumption potential.
- **Eligibility:** Defined use case, architecture aligned with AWS services (primarily Amazon Bedrock), and approval from the GenAI team.

Investment for Transformation

- **Program Focus:** Reduce the cost and risk of migrating and modernizing applications, infrastructure, and data to AWS.
- **Program Structure:** Phased funding (Assess, Mobilize, and Migrate) based on expected incremental consumption.
- **Eligibility:** Completed technical assessment, opportunity registered with AWS, and business case validated by the AWS commercial team.

Top funding programs used by Thoughtworks. For further information, please reach out to: Beatriz Paiva (beatriz.paiva@thoughtworks.com) - Global Funding Lead

We look forward to working with you

Kiran Prakash

Business Data Product Owner
kprakash@thoughtworks.com

Brian Blanchard

Global VP, Cloud
brian.blanchard@thoughtworks.com

