
AWS
re:Invent
NOV. 28 – DEC. 2, 2022 | LAS VEGAS, NV

COM306

Bringing software engineering rigor to data

Zainab Maleki

Technical Lead Consultant – Perth, Western Australia
Mechanical Rock

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Privileged

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Zainab Maleki



- 8 years of software engineering experience
- Variety of work includes frontend, backend, infrastructure, DevOps, and SRE
- Started data journey 3 years ago



https://www.linkedin.com/in/zainab-maleki

https://twitter.com/MalekiZainab

aws
© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.
MECHANICAL ROCK

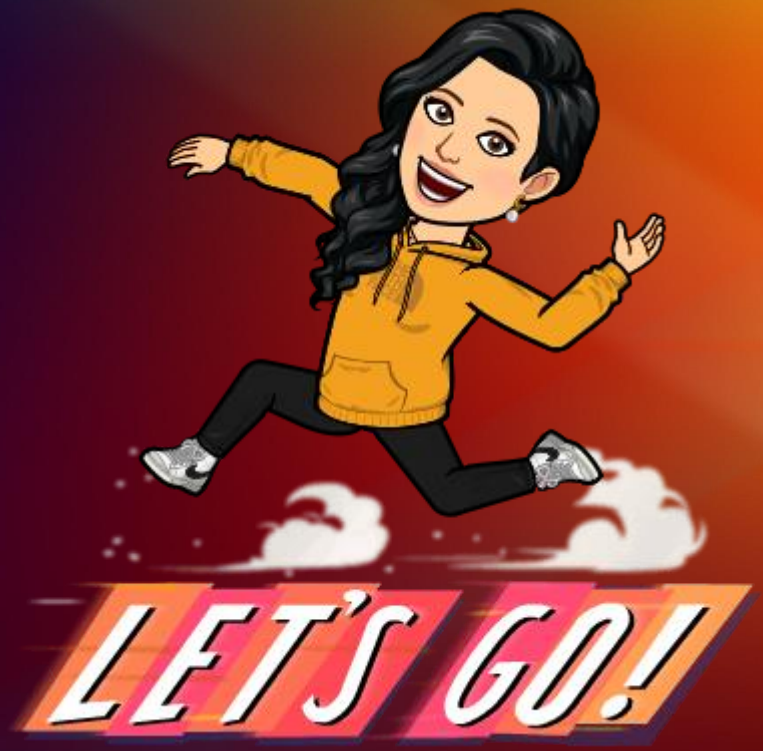
AGENDA



- Move fast with data
- DORA metrics in data
- Path to success
- Small wins

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Develop fast, move fast, and move with confidence



© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Think of it like building software applications



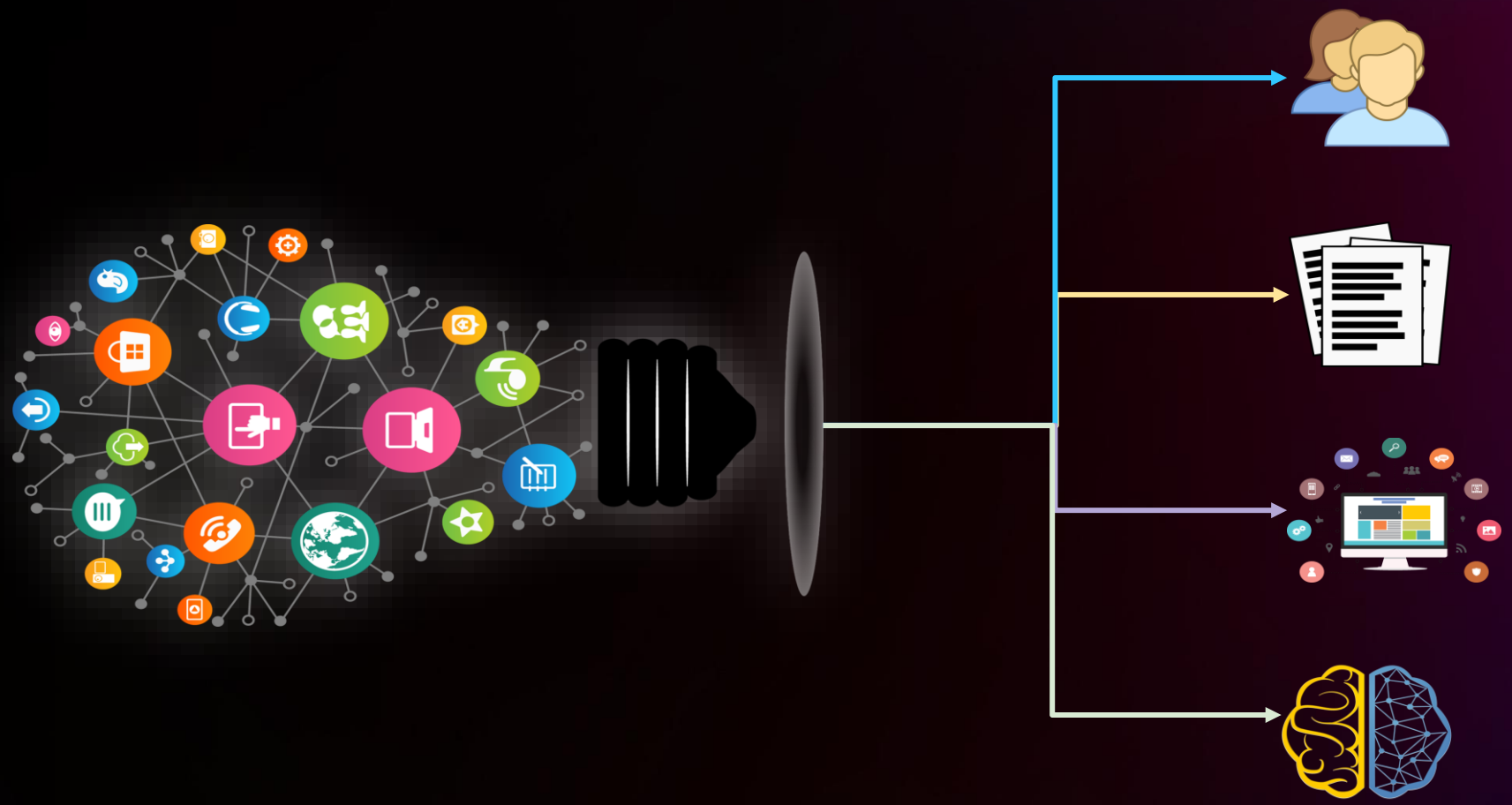
© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

You need to understand the differences



© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Data cleaning is hard



© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Testing is way more important than you think



© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

DORA

Through six years of research, the DevOps Research and Assessment (DORA) team has identified four key metrics that indicate the performance of a software development team

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

DORA metrics

Deployment frequency

- How often an organization successfully releases to production

Lead time for changes

- The amount of time it takes a commit to get into production

Change failure rate

- The percentage of deployments causing a failure in production

Time to restore service

- How long it takes an organization to recover from a failure in production

aws
© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.
MECHANICAL ROCK

DORA metrics

Metric name	Elite	High performing	Medium performing	Low performing
Deployment frequency	Multiple deploys per day	Between once a day to once a week	Between once a week to once a month	More than once a month
Lead time for changes	Less than an hour	Between one day to one week	Between one week and one month	More than one month
Change failure rate	0% to 15%	More than 15%	More than 15%	More than 15%
Time to restore service	Less than one hour	Less than one day	Between one day and one week	More than one week

aws
© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.
MECHANICAL ROCK



Accelerate State of DevOps 2021

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

How do we become elite data teams?



© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

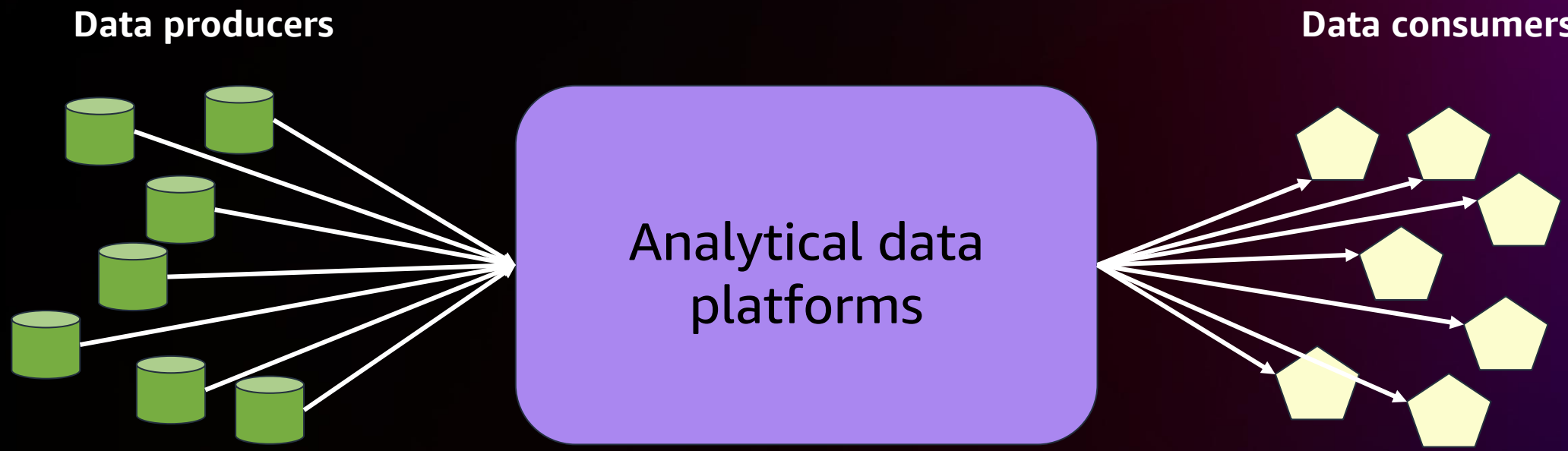
Slow-moving data platforms

CHALLENGE POINT #1



© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Slow-moving data platforms



aws
© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.
MECHANICAL ROCK

Slow-moving data platforms



© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Decentralize your data engineering teams

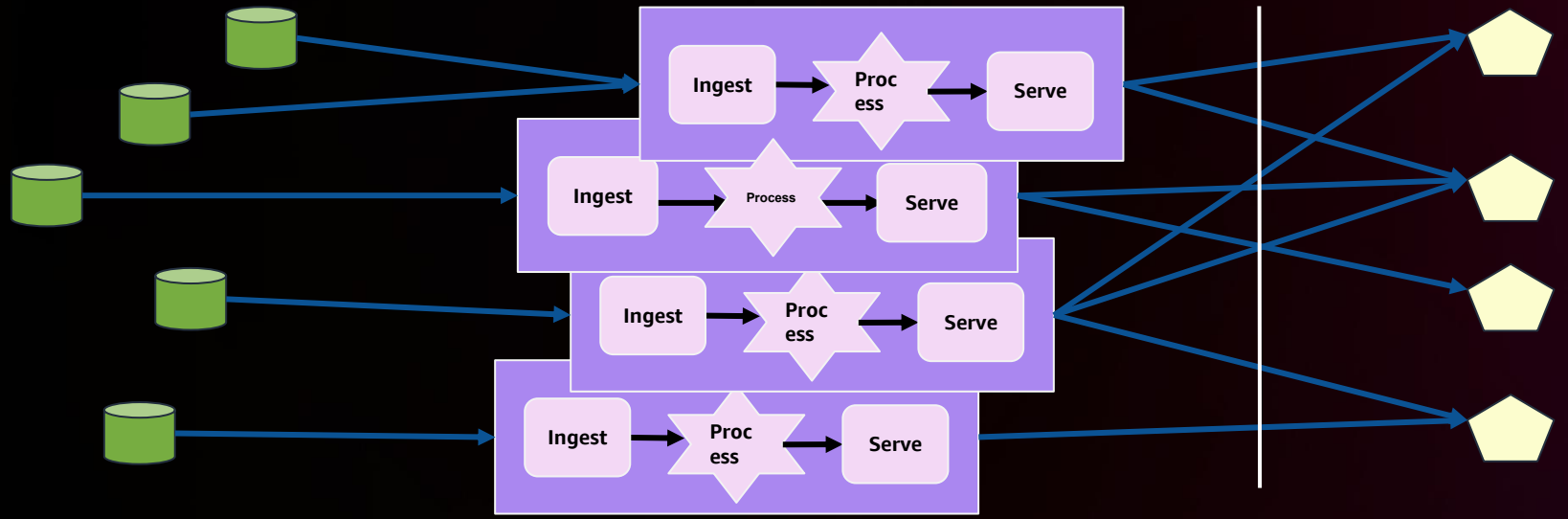
STRATEGY #1



© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

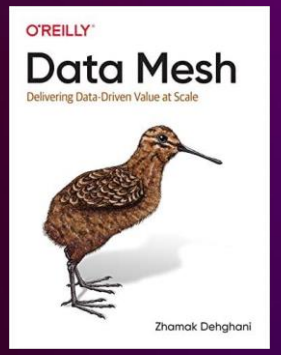
Data mesh

Data producers

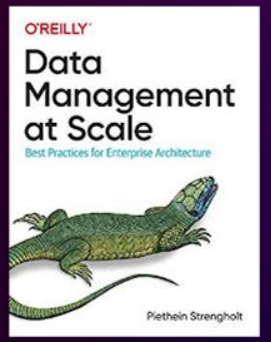


Data consumers

Consistent access layer



Data mesh



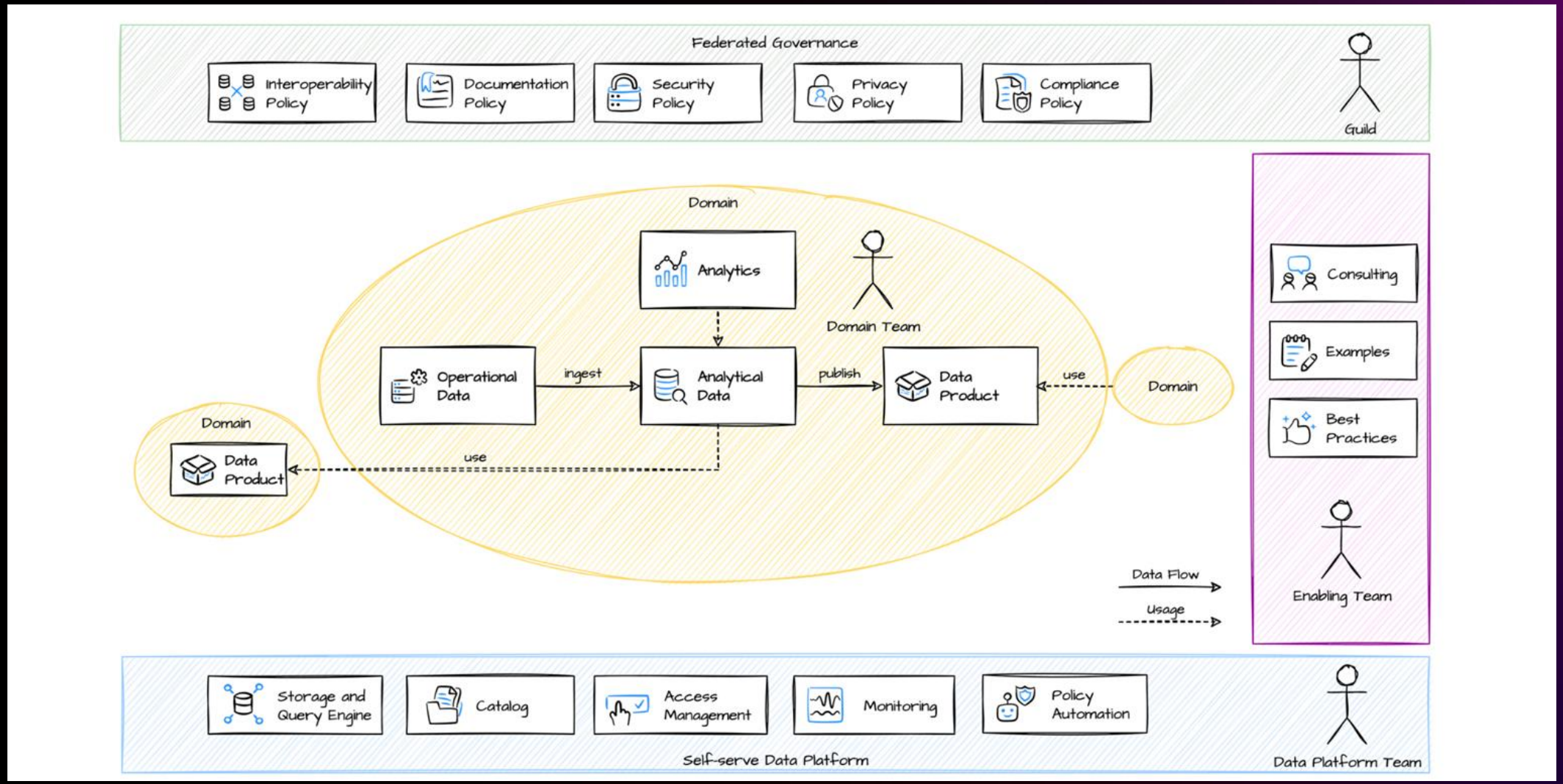
Domain-driven data design

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Separating roles and responsibilities between the platform and domains



© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.



Reference: https://www.datamesh-architecture.com/

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Platform responsibilities

- Enabling the domains
- Platform security
- Centralized access layer
- Platform consistency
- Platform audit
- Building templates

Domain responsibilities

- Ingesting data
- Transforming data
- Serving their data
- Observability on their data
- Automated testing on their data
- Data masking, PII removal

aws
© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.
MECHANICAL ROCK

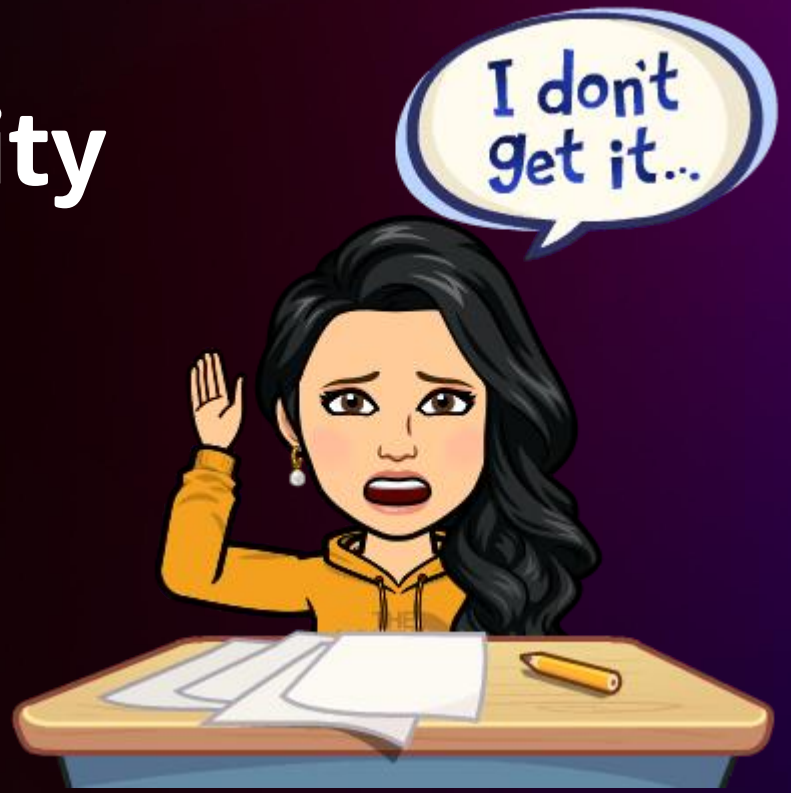
Build self-served, fully automated data platforms



© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Lack of appropriate capability

CHALLENGE POINT #2



aws

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

MECHANICAL ROCK

To build a strong platform, you need to combine your capabilities

STRATEGY #2



© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Slow platform adoption

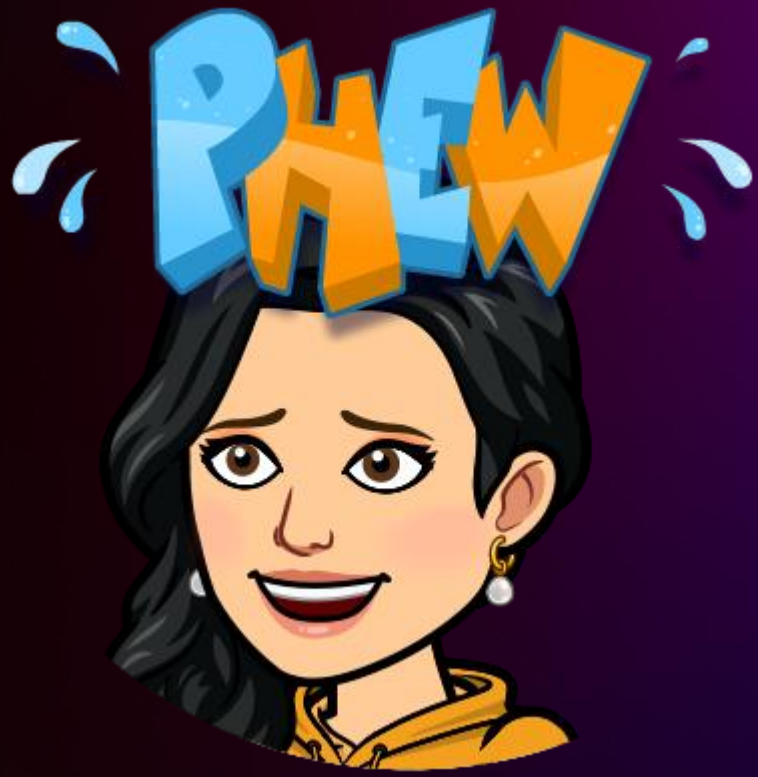
CHALLENGE POINT #3



© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Automate domain capabilities to allow easy adoption

STRATEGY #3



© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Built-in CI/CD

Built-in automated testing patterns

Built-in observability capability



aws

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

MECHANICAL ROCK

Over-protective data teams

CHALLENGE POINT #4



© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Build security and data discoverability into the platform

STRATEGY #4

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Detective vs. protective approach

Automated compliance testing



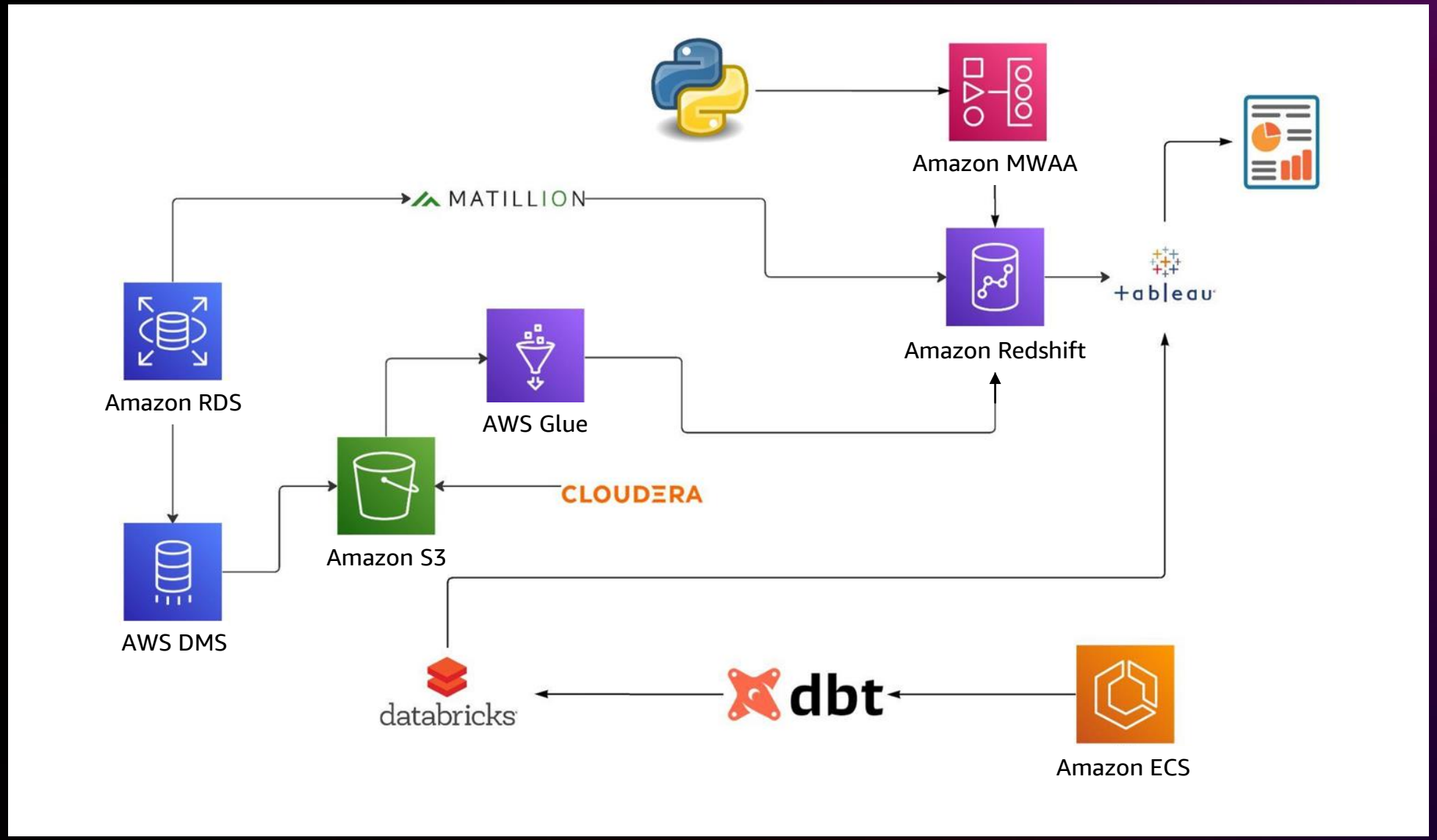
© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

A cluster of mess

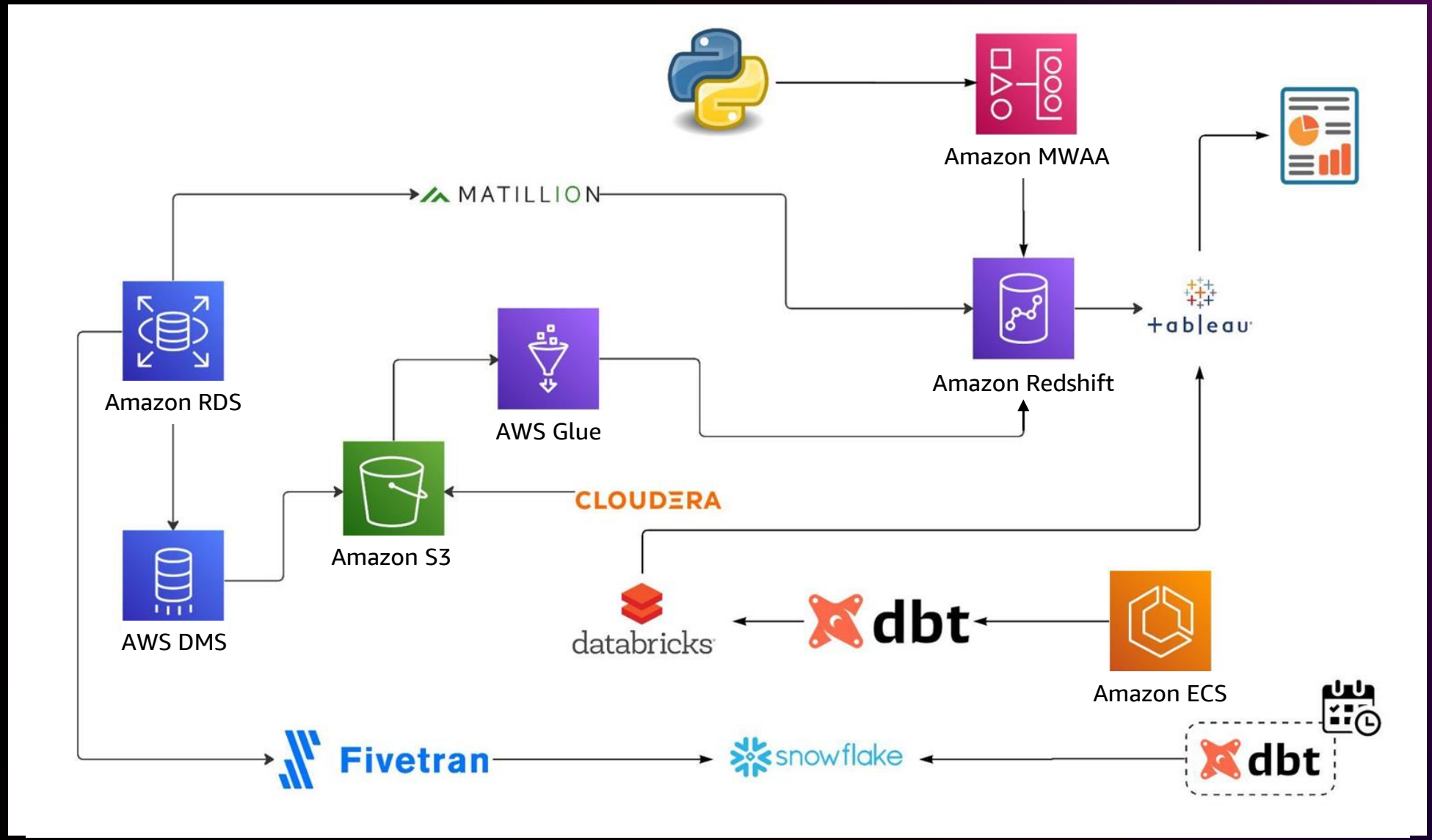
CHALLENGE POINT #5



© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.



© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.



© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Build it once, build it right

STRATEGY #5



© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

MECHANICAL ROCK



Good things take time

Small wins

- Start using Git/version control
- Setup automated deployments
- Add automated testing for all business logic
- Separate platform administration from data pipelines
- Start capturing data about the health of your data pipelines
- Remove your administration access



aws
© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Thank you!

Zainab Maleki

@zainab-maleki

@MalekiZainab



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

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Q&A



DataOps with AWS and Snowflake – https://bit.ly/3Ax2sGx
Snowflake CI/CD setup using AWS devtools



Achieving Data Mesh with DataOps for Snowflake – https://bit.ly/3OryOZ0
Demonstration on how dataOps can address concerns in data mesh



How to Use Metadata to Future-Proof Your Data Stack – https://bit.ly/3tKlCEL
Metadata driven data pipelines



Are you an Elite DevOps performer? Find out with the Four Keys Project – https://bit.ly/3TUZWAH
Four key metrics of DORA

aws
© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.
MECHANICAL ROCK