Adventures in CI/CD with Azure Synapse

Craig Porteous

Principal Data Engineer

https://craigporteous.com

@cporteous

https://github.com/cporteou

databricks PARTNER PROGRAM SOLUTIONS ARCHITECT CHAMPION

MVP Microsoft® Most Valuable Professional

Adventures in CI/CD with Azure Synapse

Craig Porteous

Principal Data Engineer

https://craigporteous.com

@cporteous

https://github.com/cporteou

databricks PARTNER PROGRAM SOLUTIONS ARCHITECT CHAMPION

MVP Microsoft® Most Valuable Professional

Let's thank our..

GOLD SPONSORS

quorum

SIOS

solarwinds

TIMEXTENDER

SILVER SPONSORS

ADVANCING ANALYTICS

redgate

POWER BI SENTINEL
Governance, Disaster Recovery and Auditing for Power BI

BRONZE SPONSOR

GE
GETHYNELLIS.COM

Community Partner Support

DATA SATURDAYS

ORGANISERS

Satya Jayanty

Enterprise Architect specialised in Data Platforms and Cloud Computing, Speaker, Mentor, and Community Organiser.

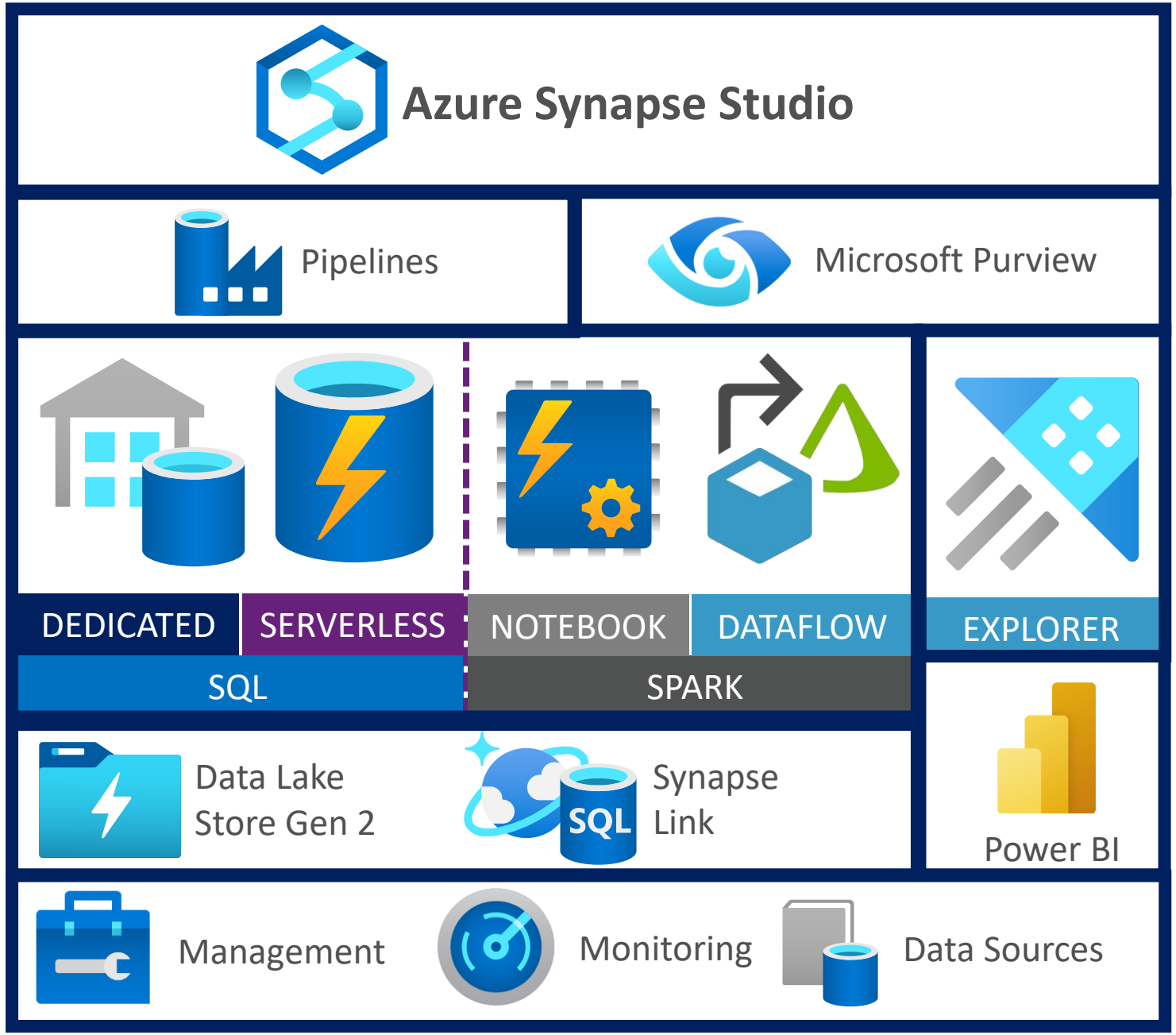
Tom Sykes

Senior Cloud Data Architect with many years of experience in the Cloud and Database industry, and Certified Trainer.

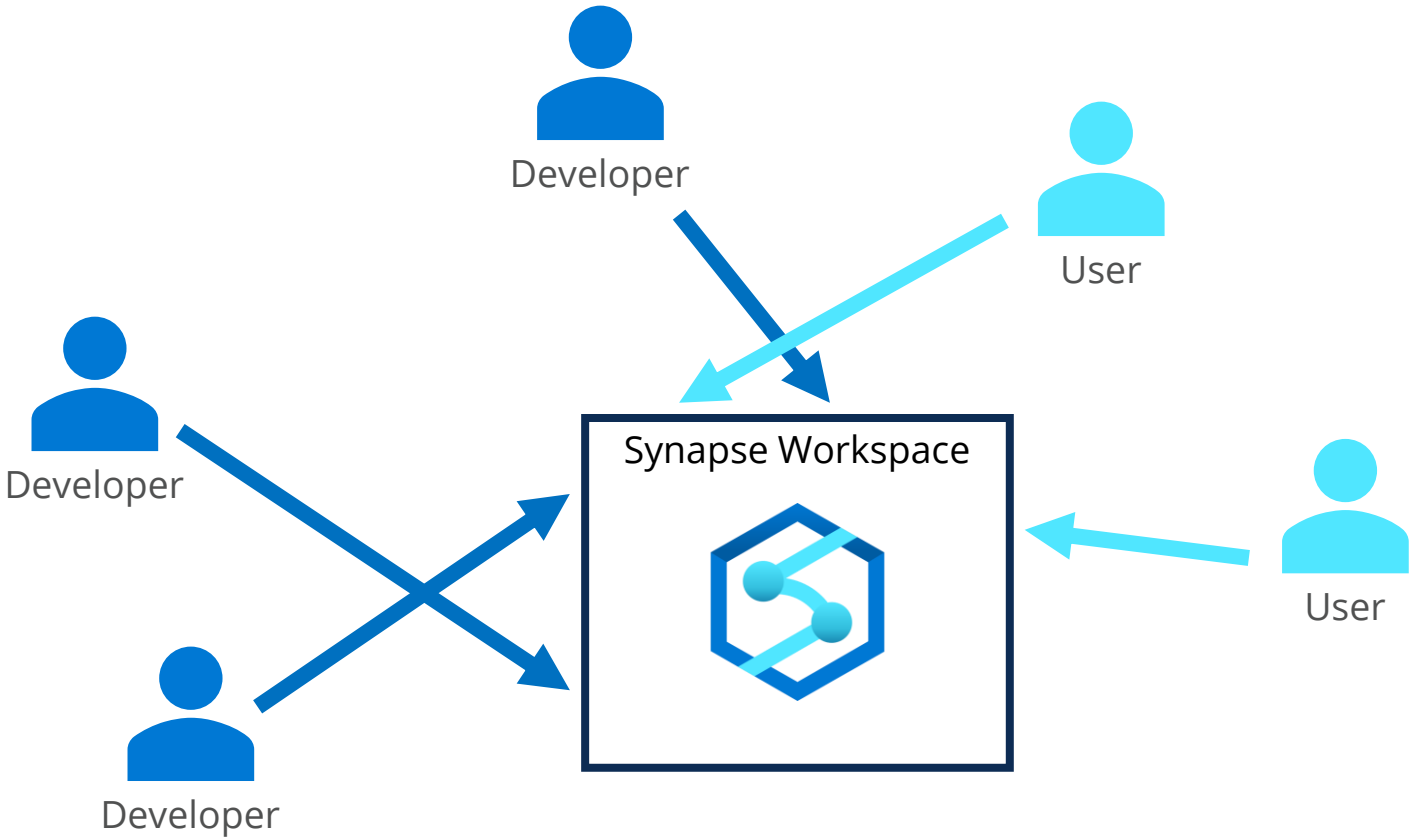
Volunteers:

- Syama Rudra
- Aditya Yembuluru
- Venkatesh Gaddam
- Samrat Ravuri
- Saikrishna Prathipati
- Saimohan Inala
- Mark Adamson

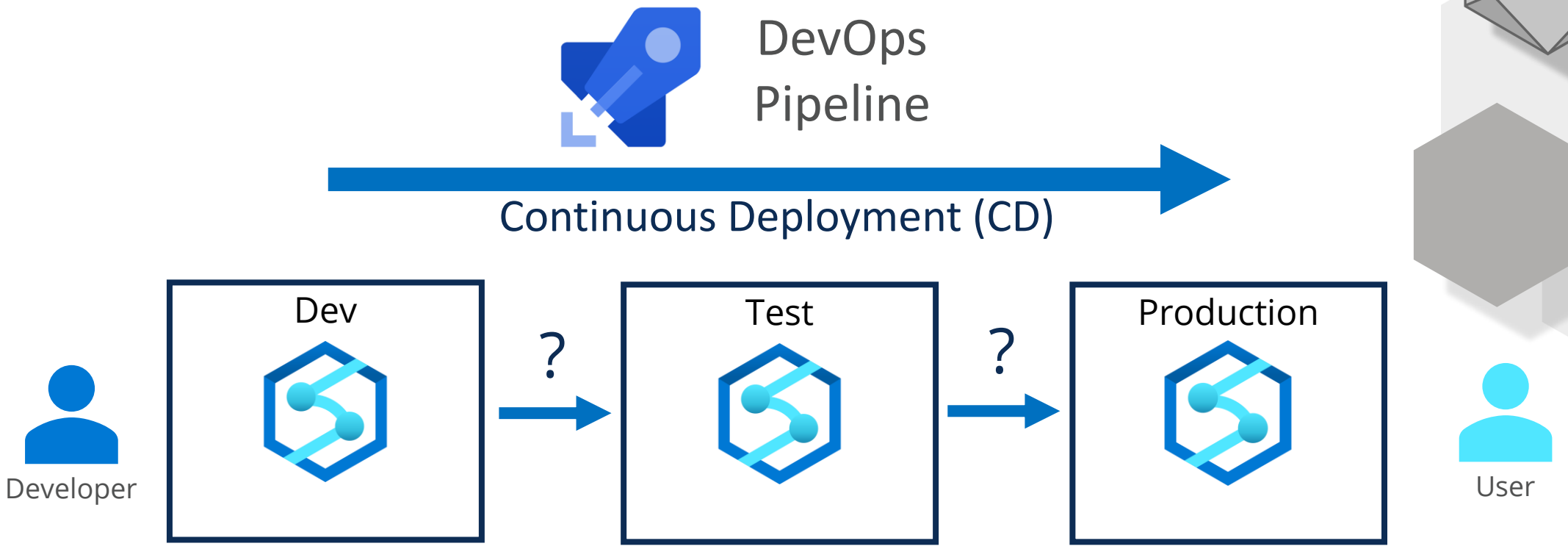
@SQLScot


Azure Synapse Studio
Pipelines
Microsoft Purview
DEDICATED
SERVERLESS
NOTEBOOK
DATAFLOW
EXPLORER
SQL
SPARK
Data Lake
Store Gen 2
SQL
Synapse
Link
Power BI
Management
Monitoring
Data Sources

DEVELOPMENT CHALLENGES



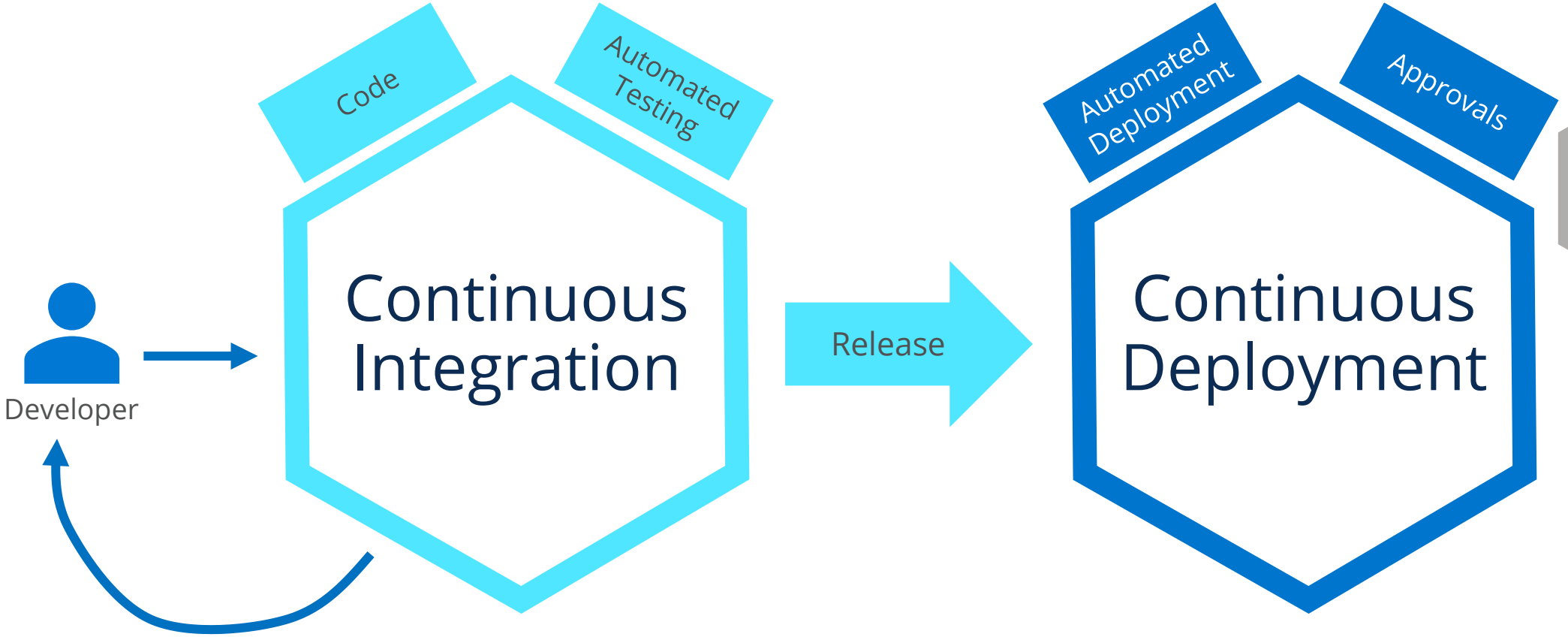
MULTIPLE ENVIRONMENTS



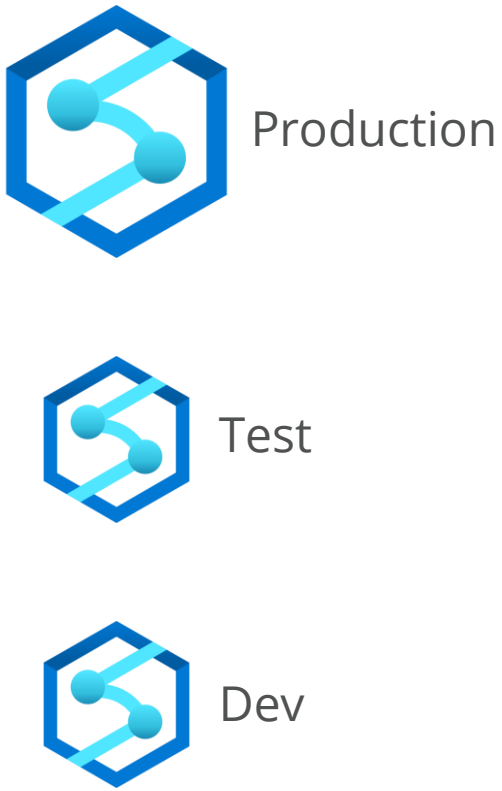
What is CI/CD?

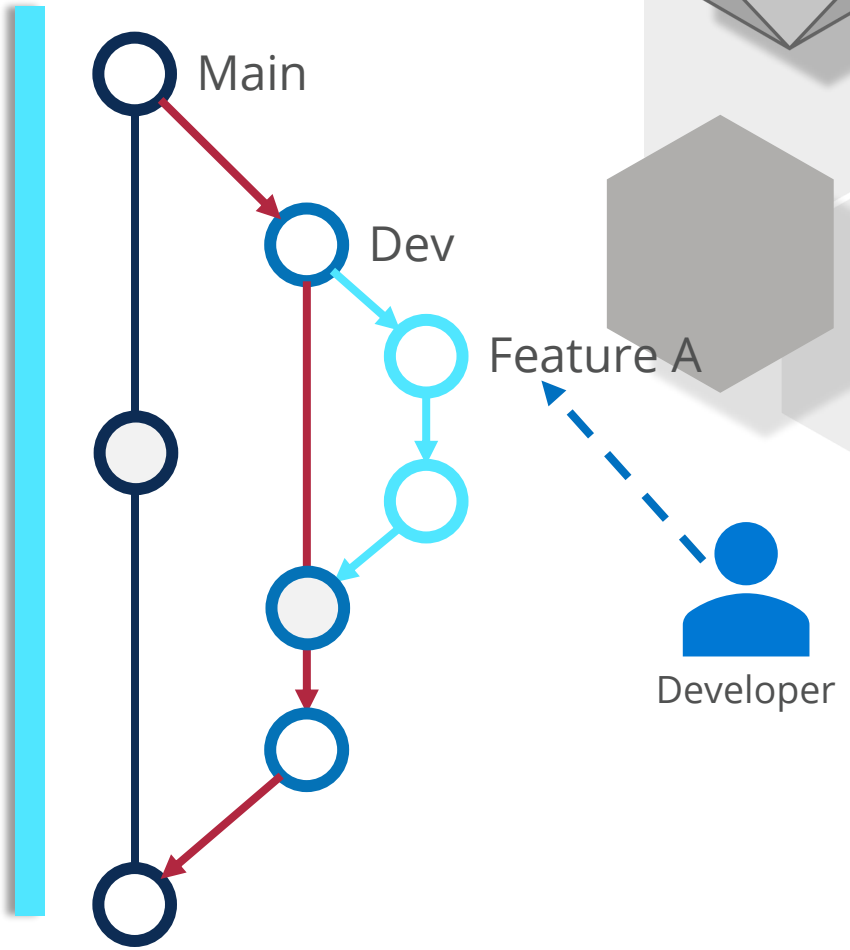
Continuous Integration

Continuous Deployment

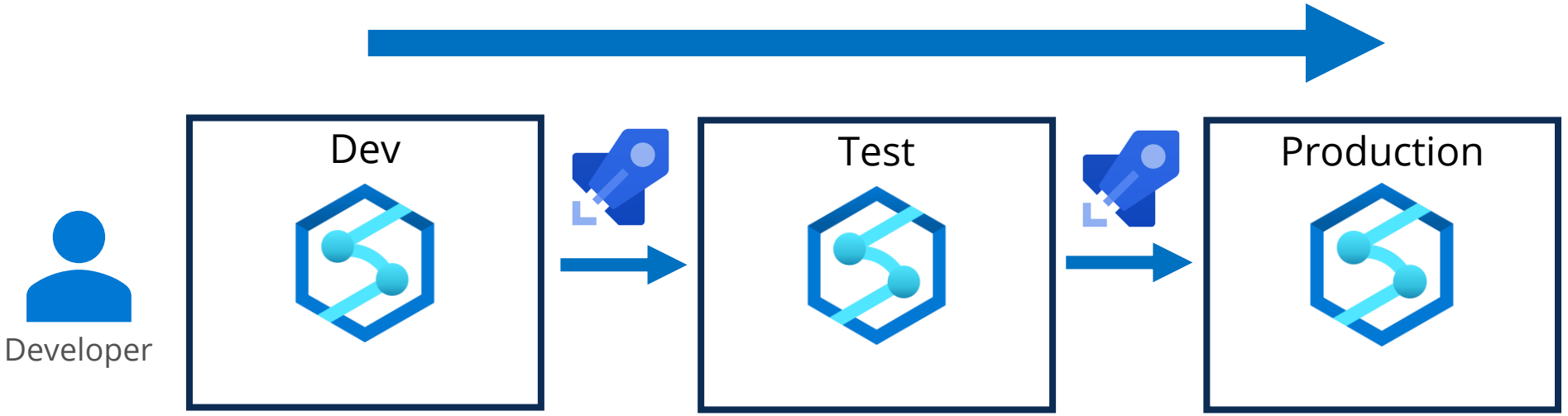

Code
Automated Testing
Continuous Integration
Developer
Release
Automated Deployment
Approvals
Continuous Deployment

PREREQUISITES

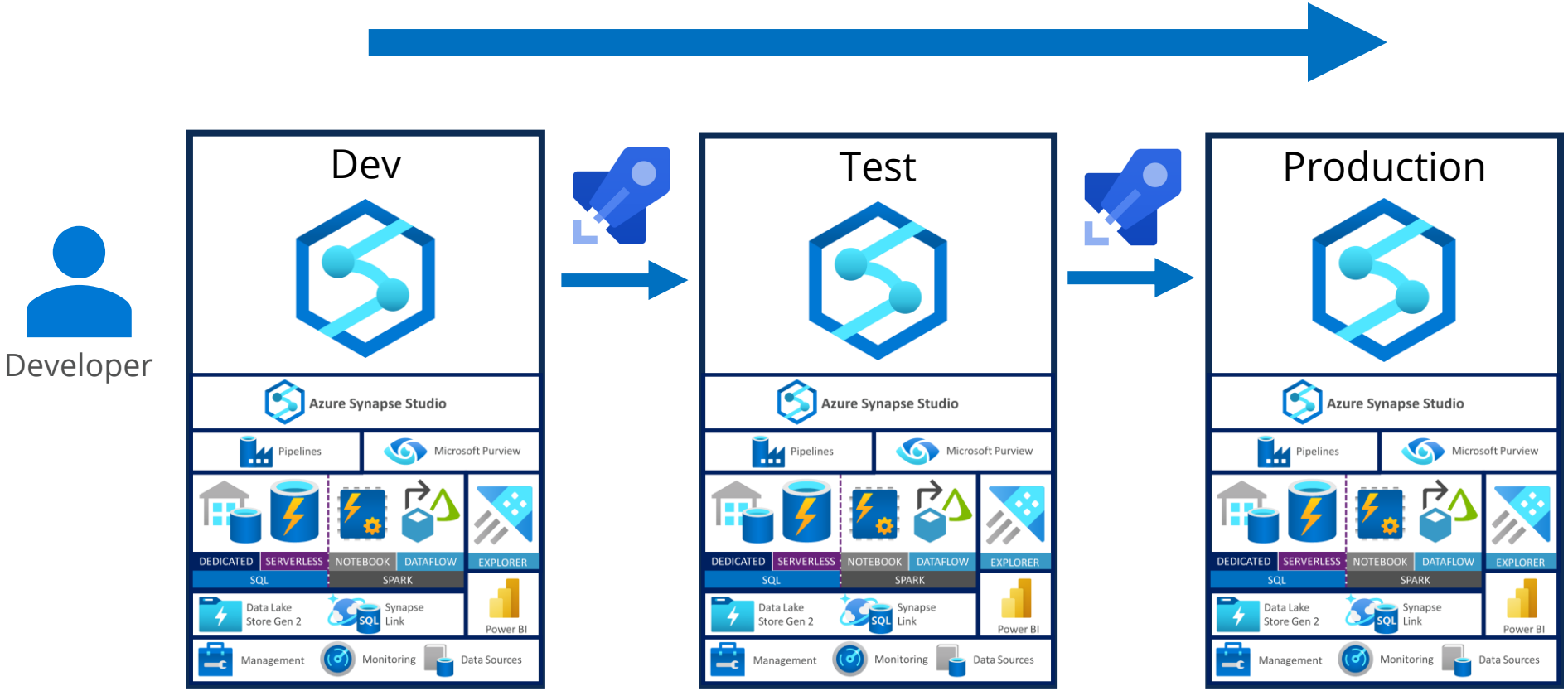




PROBLEM SOLVED



NOT SO SIMPLE



DEPLOYING

Self-hosted IRs

Purview connection

Artifacts

- SQL Scripts
- Notebooks
- Spark job definitions
- Pipelines
- Datasets
- Data Flows
- Linked Services
- Triggers
- Integration runtimes
- Credentials

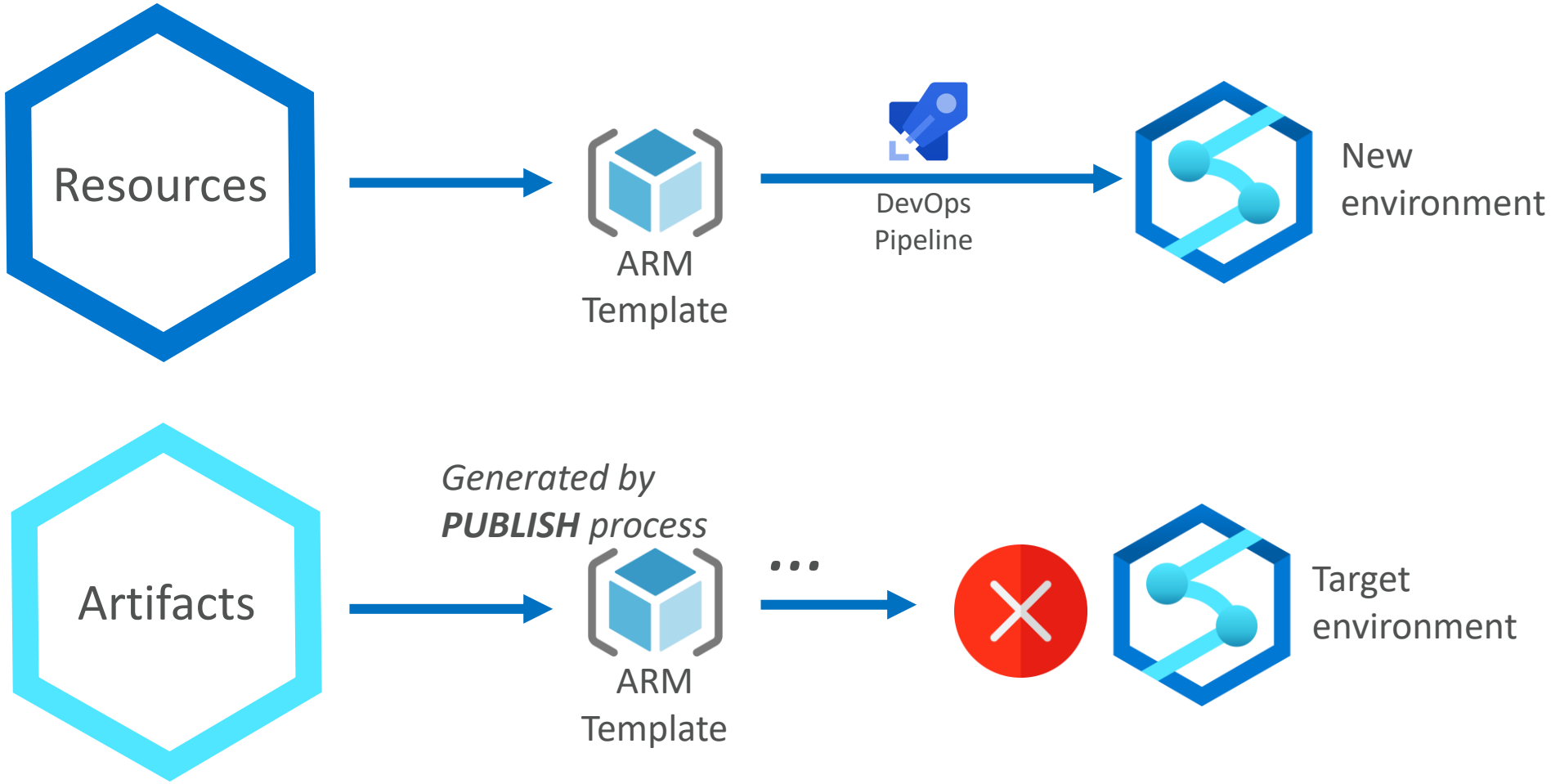
Resources

- Workspace
- SQL pools
- Spark pools
- Security
- Storage (ADLS)

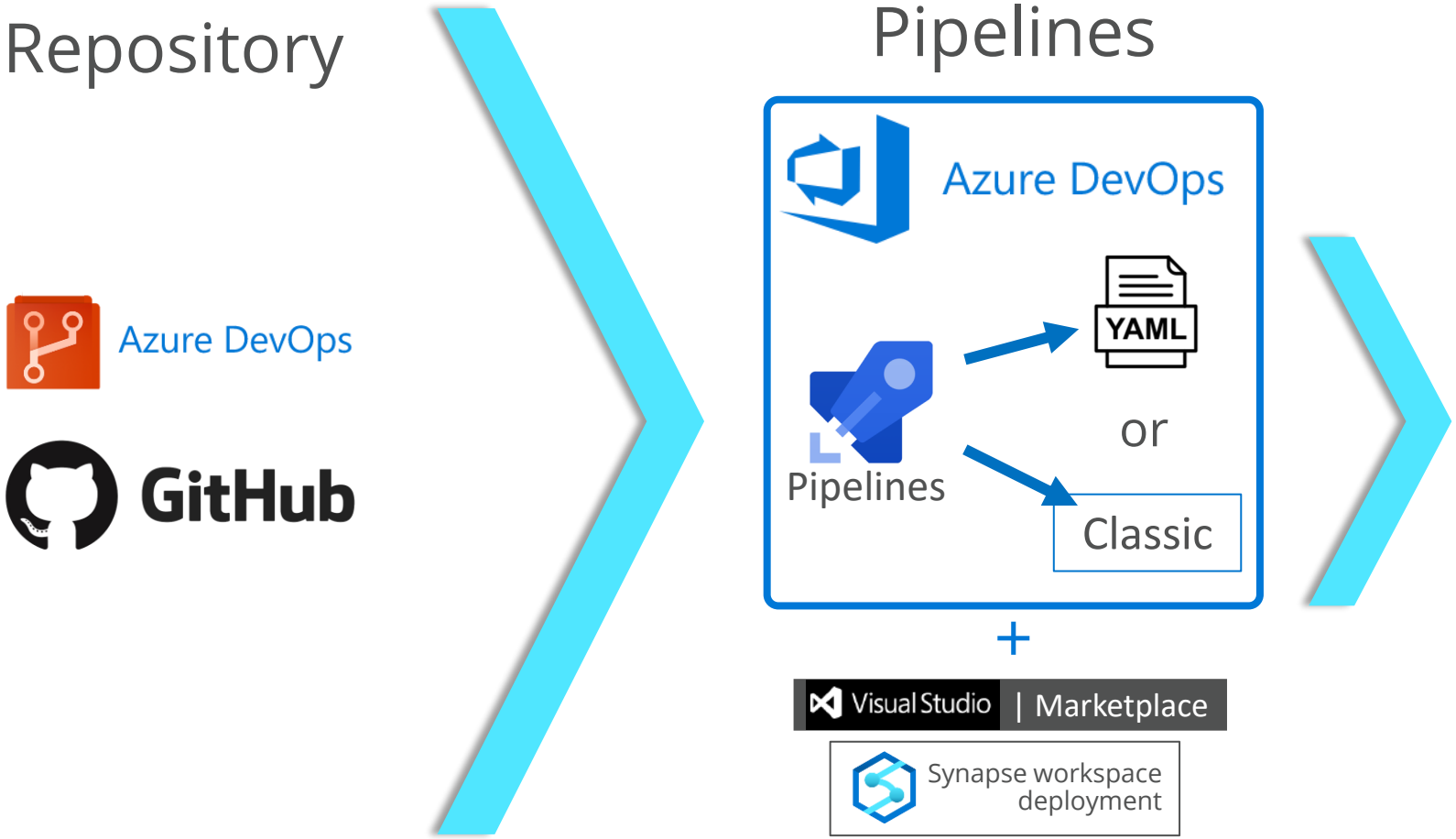
SQL dedicated pools

Database objects

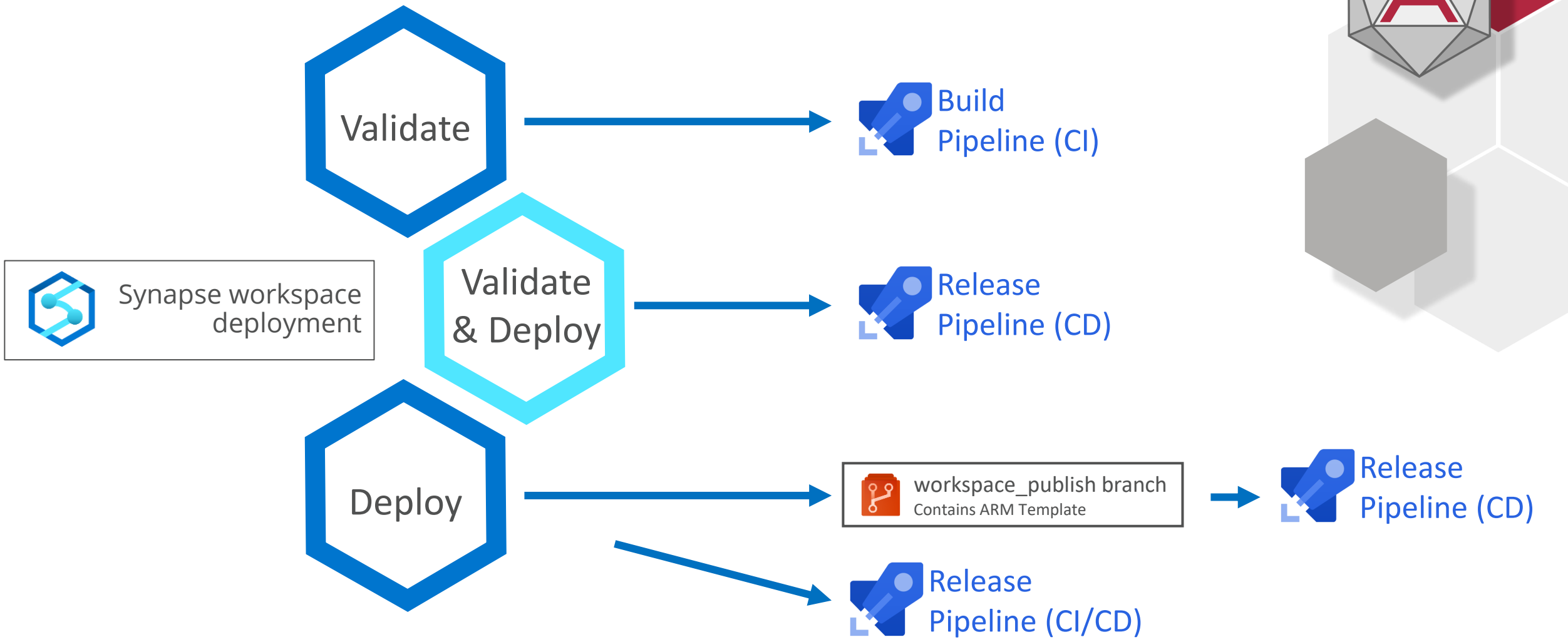
THIS IS NOT DATA FACTORY



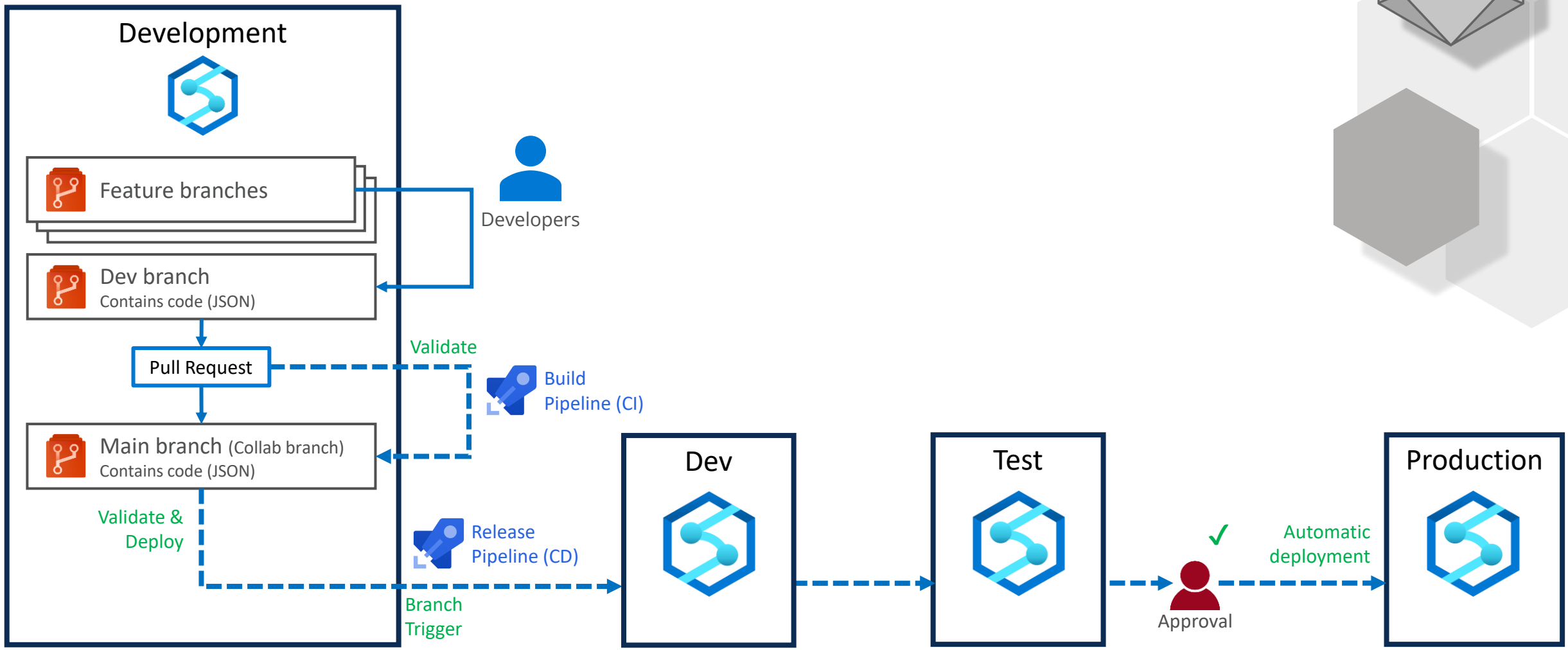
DEPLOYING ARTIFACTS



SYNAPSE EXTENSION MODES



AZURE SYNAPSE ANALYTICS CI/CD



DEMO

THINGS TO REMEMBER

- Integration Runtime's need to be the same type in all environments
- Deploying Artifacts as ARM templates will give inconsistent results
- Required resources (e.g. Pools) need to be in place before deploying artifacts
- The "Built-in" linked services will automatically "update" between environments
- Notebook pool connections need to be updated (manually or by Find/replace)
- Dedicated pools need to be online in target environment for CD pipeline
- Dedicated pool objects aren't handled – deploy with DACPAC step

THANK YOU

https://craigporteous.com

@cporteous

https://github.com/cporteou

Let's thank our..

GOLD SPONSORS

quorum

SIOS

solarwinds

TIMEXTENDER

SILVER SPONSORS

ADVANCING ANALYTICS

redgate

POWER BI SENTINEL
Governance, Disaster Recovery and Auditing for Power BI

BRONZE SPONSOR

GE
GETHYNELLIS.COM

Community Partner Support

DATA SATURDAYS

ORGANISERS

Satya Jayanty

Enterprise Architect specialised in Data Platforms and Cloud Computing, Speaker, Mentor, and Community Organiser.

Tom Sykes

Senior Cloud Data Architect with many years of experience in the Cloud and Database industry, and Certified Trainer.

Volunteers:

- Syama Rudra
- Aditya Yembuluru
- Venkatesh Gaddam
- Samrat Ravuri
- Saikrishna Prathipati
- Saimohan Inala
- Mark Adamson

@SQLScot

REFERENCES

https://marketplace.visualstudio.com/items?itemName=AzureSynapseWorkspace.synapsecicd-deploy

(5) Using CI/CD for Serverless SQL Pools in Azure Synapse Analytics – YouTube

CI/CD For Azure Data Factory With Azure DevOps - Craig Porteous