Microsoft Fabric and GitHub – The story so far

Kevin Chant

PATRONAT PREZYDENTA
MIASTA BYDGOSZCZY
Rafała Bruskiego

SQL Day Lite

SILVER SPONSOR

DATUMO

BRONZE SPONSORS

software one

C&F

PARTNERS

BYDGOSKI KLASTER INFORMATYCZNY

Microsoft

PyData Bydgoszcz

Agenda

- Bio
- Intro to Microsoft Fabric
- Intro to GitHub
- Configuring GitHub for Git integration
- Using GitHub with suggested CI/CD workflow options
- CI/CD for Data Warehouses

Kevin Chant

- Data Engineering Manager in the Netherlands
- Worked in IT since days of Windows 95
- Experience in various sectors
- Various certifications, MCT, MVP

- Twitter/Blue Sky: @kevchant
- LI: https://www.linkedin.com/in/kevin-chant/
- Blog: https://www.KevinRChant.com
- GitHub: https://github.com/kevchant

Intro to Microsoft Fabric

Microsoft Fabric

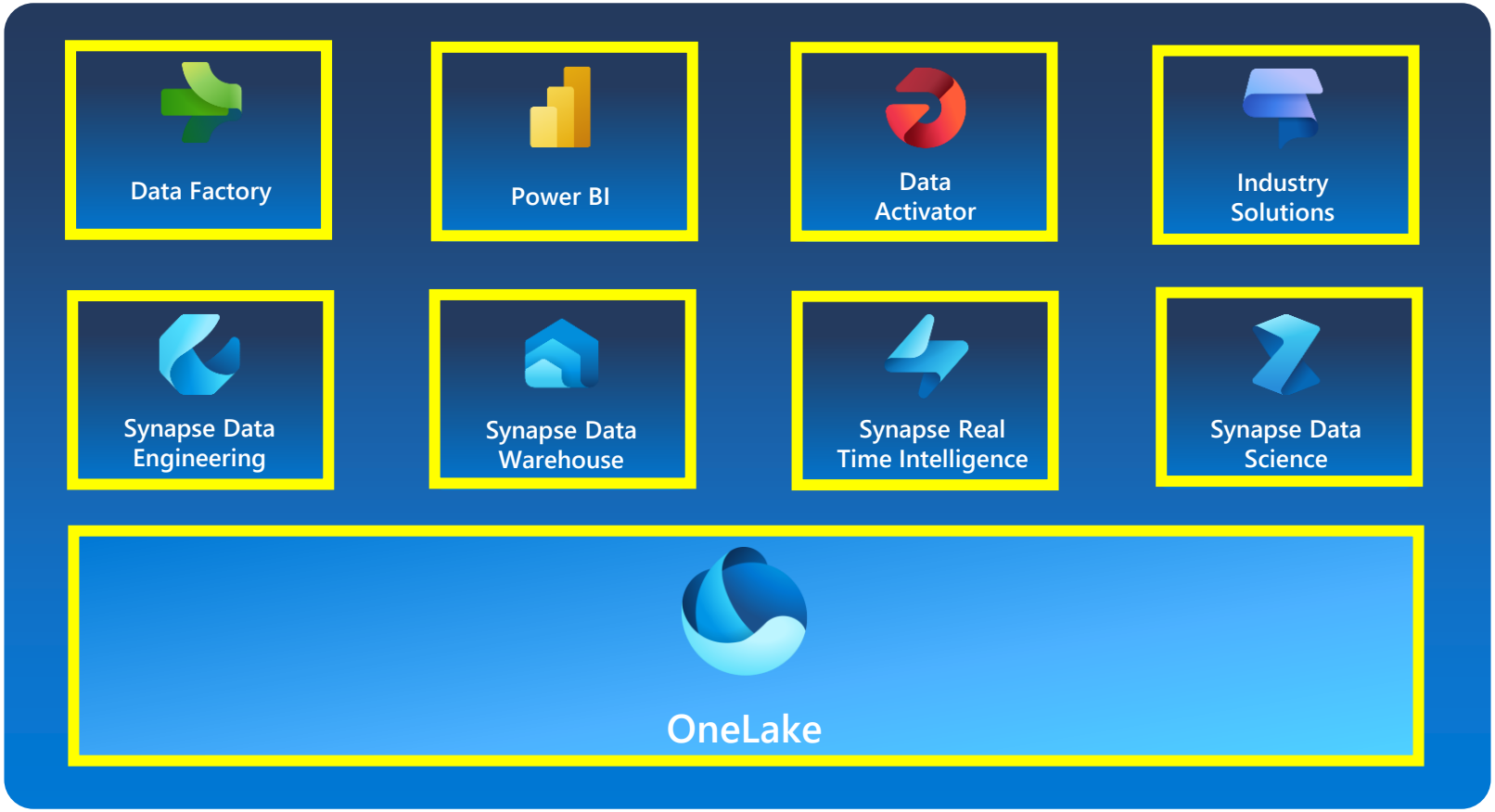
SQL Day Lite

Comp analytics platform

Lake centric and open

Empower every Business user

AI Powered



Microsoft Fabric

SQL DayLite

Intro to GitHub

GitHub

- Can be used for Application Lifecycle Management
- Popular within open-source community
- Various plans
- Large range of features

As well as repositories...

- GitHub Actions
- GitHub issues
- Wikis

GitHub Plans (formerly products)

- GitHub Free for personal accounts
- GitHub Pro
- GitHub Free for organizations
- GitHub Team
- GitHub Enterprise
  - Cloud
  - Server

General Security considerations for GitHub

- Use secrets and Azure Key Vault with GitHub Actions.
- Enable code and secret scanning.
- GitHub has document on what to do if sensitive values appear in repository.
- Recommend GitHub Enterprise for production use.
- Consider GitHub certification training.

GitHub

Configuring GitHub for Microsoft Fabric Git integration

About Microsoft Fabric Git integration

- Allows you to synchronize supported items within a Fabric workspace with a Git repository
- Support of various Fabric items vary
- Supports Azure DevOps & GitHub repositories
- Currently only supports GitHub.com offerings
- Requires Fabric or Power BI Premium capacity

Configuring GitHub for Microsoft Fabric Git integration

- Requires Personal Access Token for each user.
- Supports “Branch out with workspace functionality”

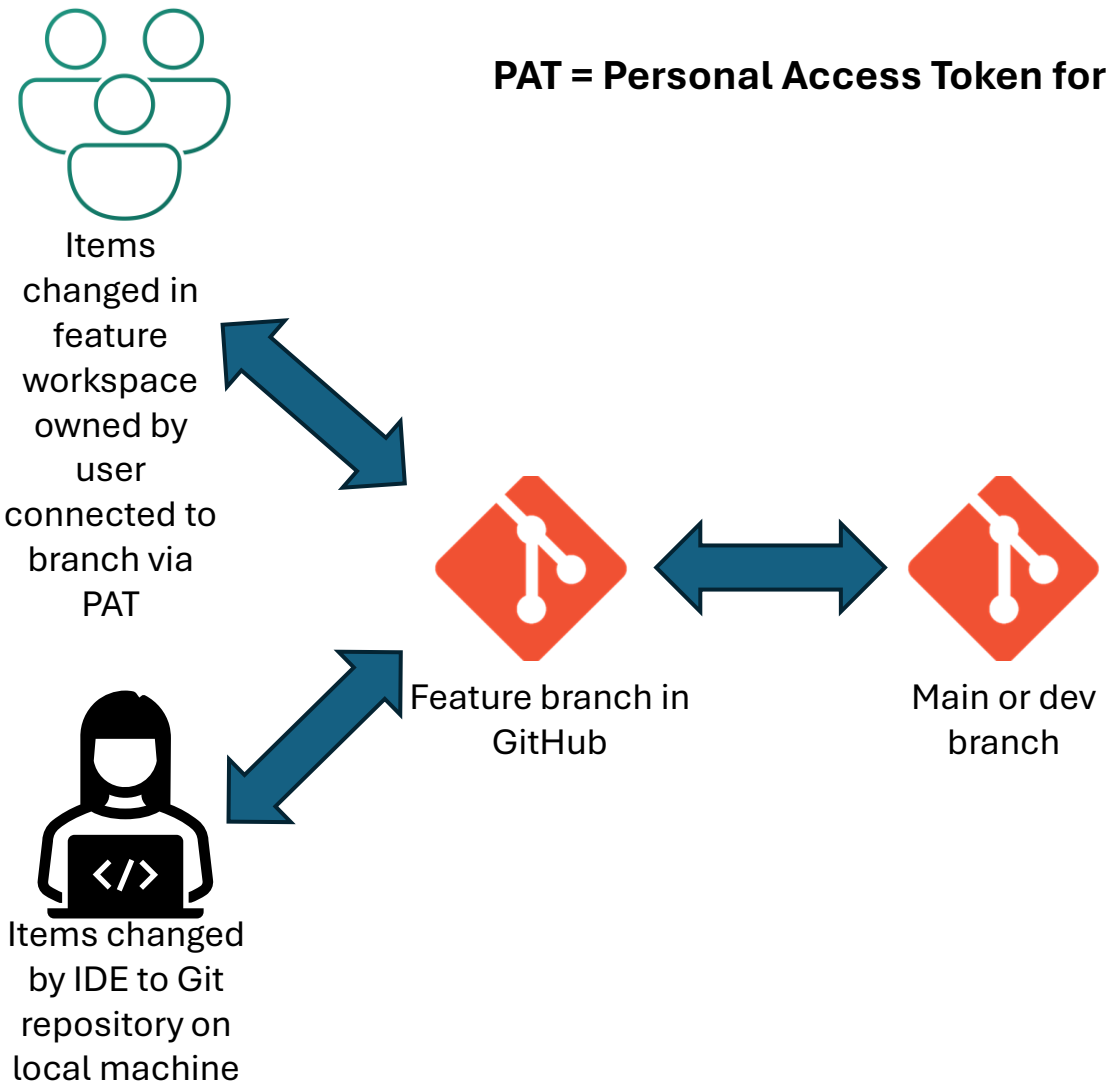
Security considerations for Git integration

- Can’t create repository in specific region.
- Enable code and secret scanning.
- Review PAT strategy.
- Review who can manage connections and gateways in Fabric.

Configuring GitHub

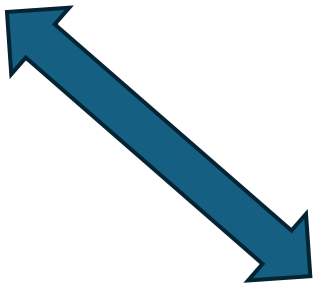
Using GitHub with suggested CI/CD workflow options

Recommended development process



Taking it one step further for Power BI reports

Power BI report changed in feature workspace owned by user connected to branch via PAT



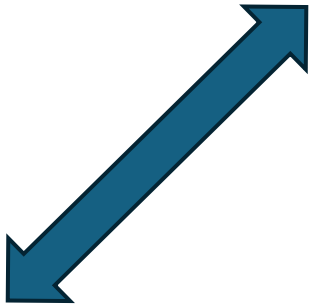
PAT = Personal Access Token for GitHub

Pull request raised

CI checks passed

GitHub

Power BI changed locally in Power BI Desktop to Git repository on local machine



Feature branch in GitHub

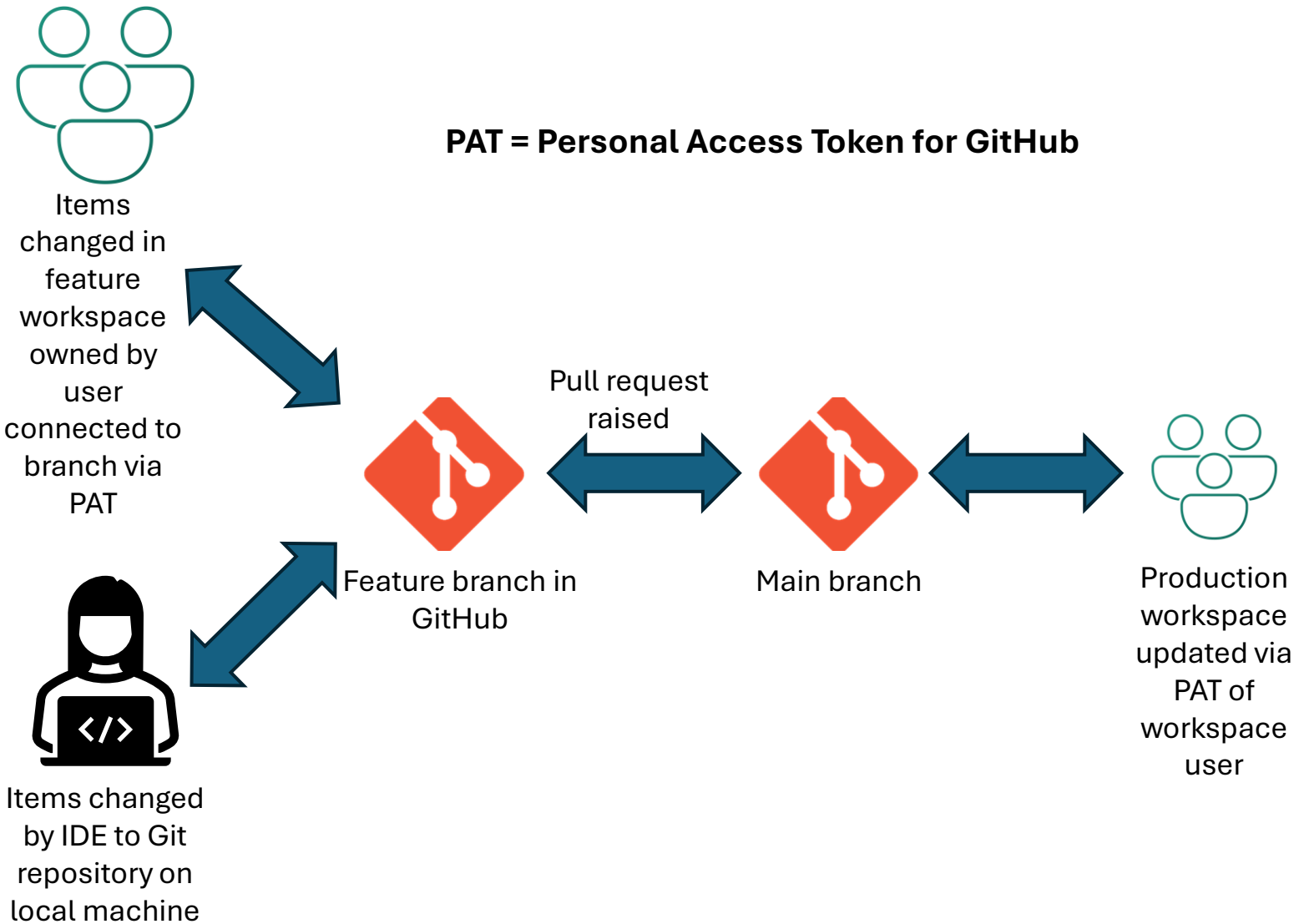
GitHub Actions workflow runs. Performing CI tests using Tabular Editor and PBI Inspector

Main or dev branch updated

Release Option 1 – Git-based deployments

- Deploy to multiple workspaces that are connected to same Git repository.
- Achieved by using multiple branches.

Release Option 1 – Diagram



Release Option 2 – Git-based deployments using build pipeline

- Deployments to different workspaces orchestrated by GitHub Actions

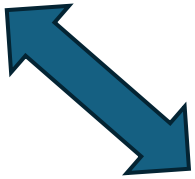
- Recommendation is that each workflow contains a build and release process.
  - Build for unit tests
  - Release to perform update

- Consider own GitHub Runners and environments.

Release Option 2 – Diagram

Items changed in feature workspace owned by user connected to branch via PAT

PAT = Personal Access Token for GitHub



Feature branch in GitHub

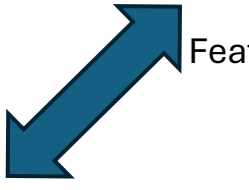
Pull request raised

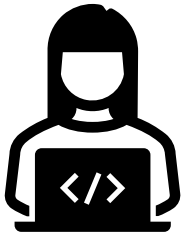
Main branch

GitHub

GitHub Actions workflow runs

Production workspace updated via PAT of workspace user



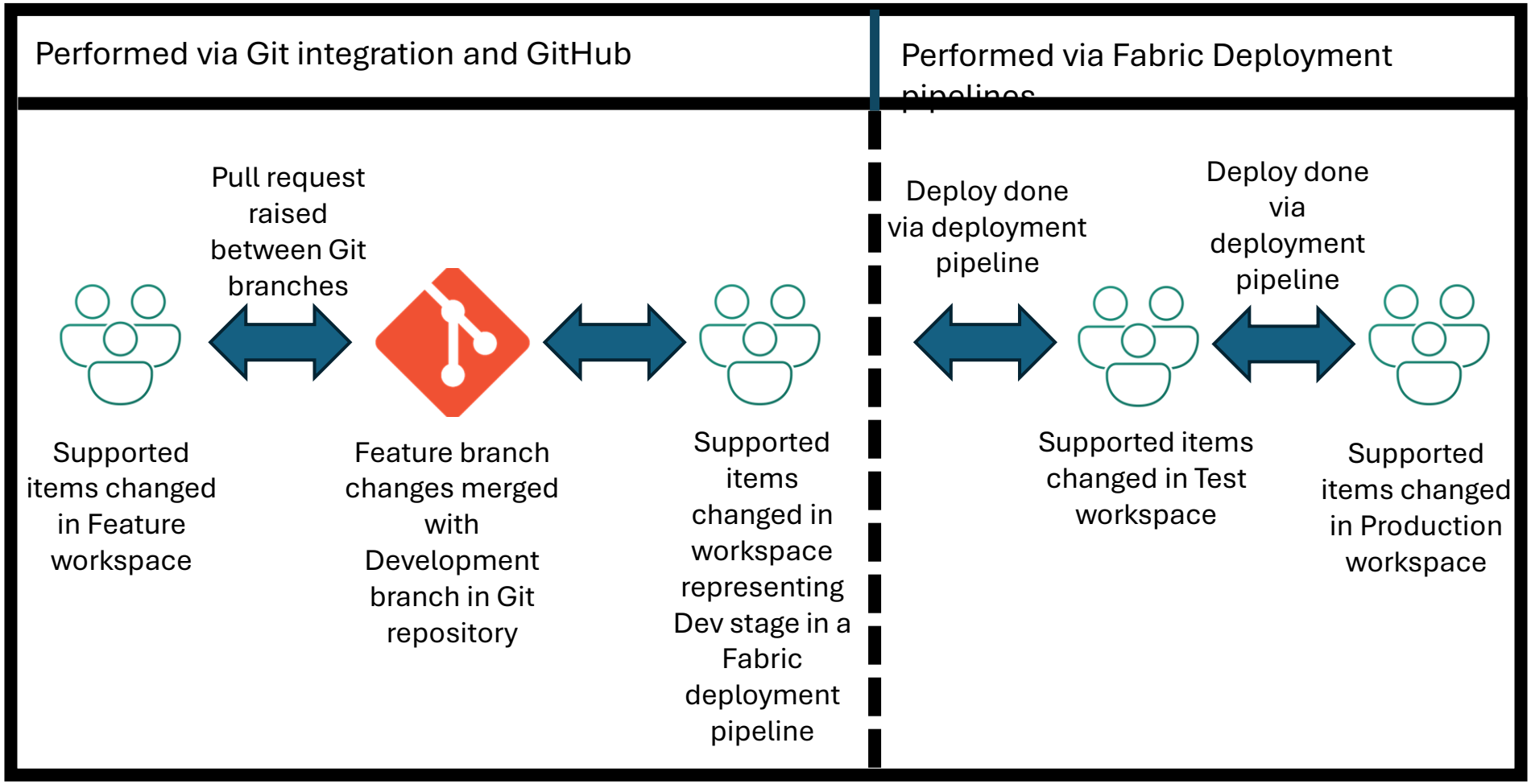


Items changed by IDE to Git repository on local machine

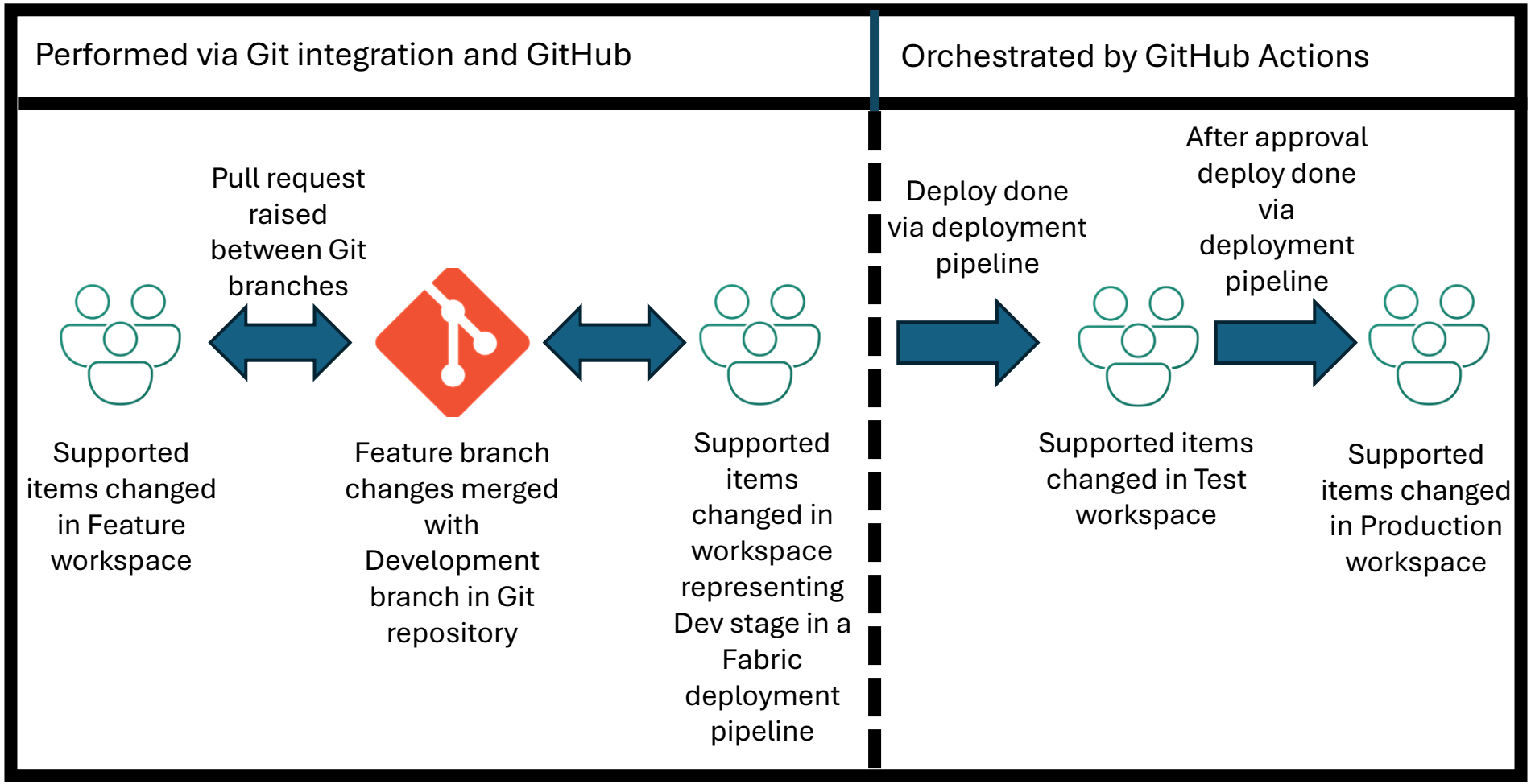
Release Option 3 – Deploy via Microsoft Fabric deployment pipelines

- Perform your pull request from feature branch to branch connected to a workspace representing Dev stage of a deployment pipeline.
- From there orchestrate using Microsoft Fabric deployment pipelines.
- Alternatively, orchestrate to different Microsoft Fabric deployment pipeline stages using GitHub Actions.

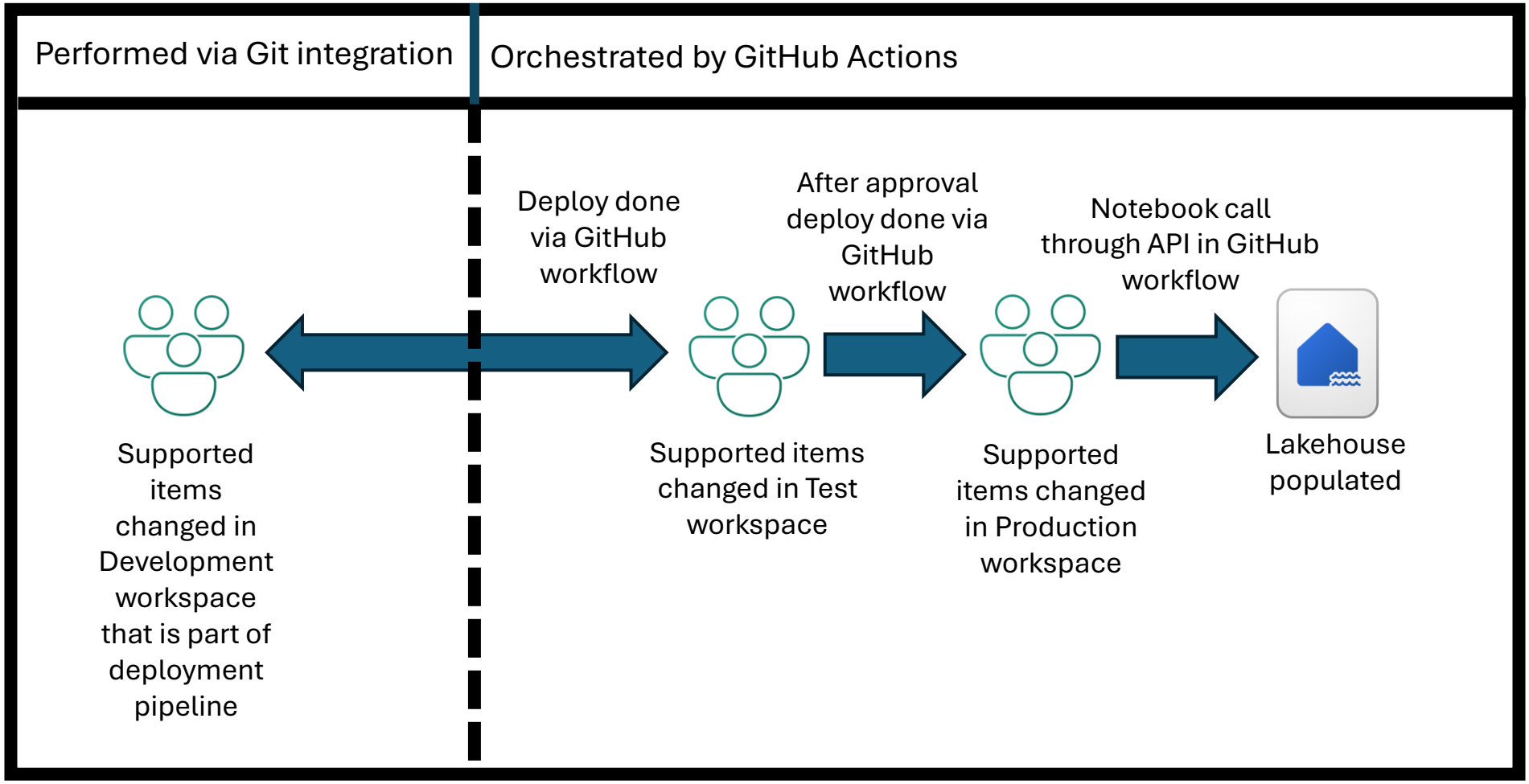
Release Option 3 – diagram



Release Option 3 – Orchestrated by GitHub Actions



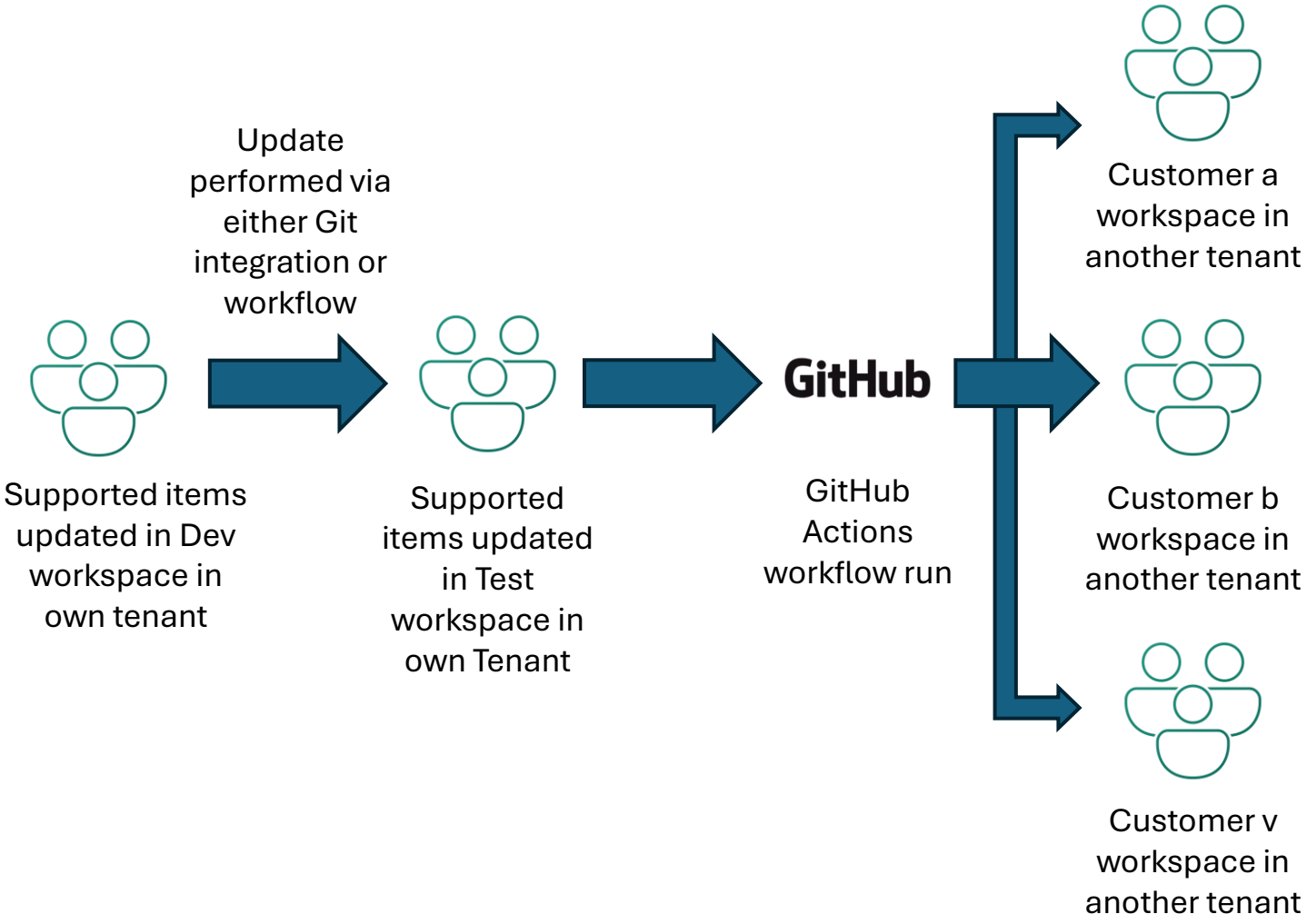
Release Option 3 – Another advantage of GitHub Actions



Release Option 4 – For multiple customers/solutions

- Dev and test stages are managed in own Fabric tenant.
- Deployment to Prod workspaces in other tenants performed by GitHub Actions

Release Option 4 – Diagram



Demos

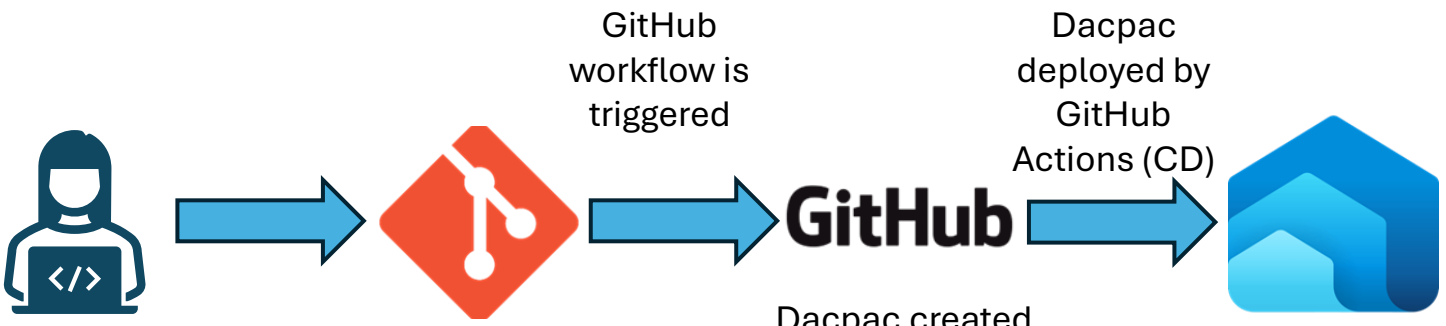
- Power BI Desktop projects
- Development process & deploying to multiple workspaces via Git integration
- Deploying via deployment pipelines

Alternative CI/CD method for Data Warehouses

CI/CD for Data Warehouses

- Can connect to Data Warehouse via connection string
- Allows deployment of traditional CI/CD methods (e.g. dacpac)
- Supports Database Projects created by MS Fabric Git integration

Diagram


GitHub workflow is triggered
Dacpac deployed by GitHub Actions (CD)
GitHub
Database project is updated locally
Update is synchronized to Git repository in GitHub
Dacpac created from stored Database Project by GitHub Actions (CI)
Schema in Data Warehouse updated

CI/CD for Data Warehouse

SQL DayLite

Questions?

Thank you

- Twitter/Blue Sky: @kevchant
- LI: https://www.linkedin.com/in/kevin-chant/
- Blog: https://www.KevinRChant.com
- GitHub: https://github.com/kevchant

Data Community