Self-Service für Datenbanken und Scheduling bei der Fraport AG

Martin Herold, Fraport AG

- Wer legt Datenbankobjekte an?
- Wie lange dauert es, bis ein Anwender eine neue Tabelle/View in der Datenbank bekommt?
- Wer konfiguriert die automatische Ausführung der ETL-Jobs?
- Wie oft verändert sich der Ablaufplan der eingeplanten ETL-Jobs?

Wie lief es früher?

Kunde fragt an, ein Entwickler schreibt irgendwann ein SQL und führt es in Prod aus.

Kunde fragt an, ein Entwickler konfiguriert.

- Wurde dokumentiert?
  Warum wurde die Änderung gemacht?
- Wer hat die Änderung gemacht?
  Jeder Entwickler braucht „Admin“-Rechte.
- Wurde die Änderung in allen
  Staging Systemen durchgeführt?
- Wie sah die Welt vor der Änderung aus?
  Was genau wird geändert? Wie kommt man zurück?
- Wurde ausreichend getestet?
- Hat jetzt jemand Zeit die Änderung zu machen?

Die Vision

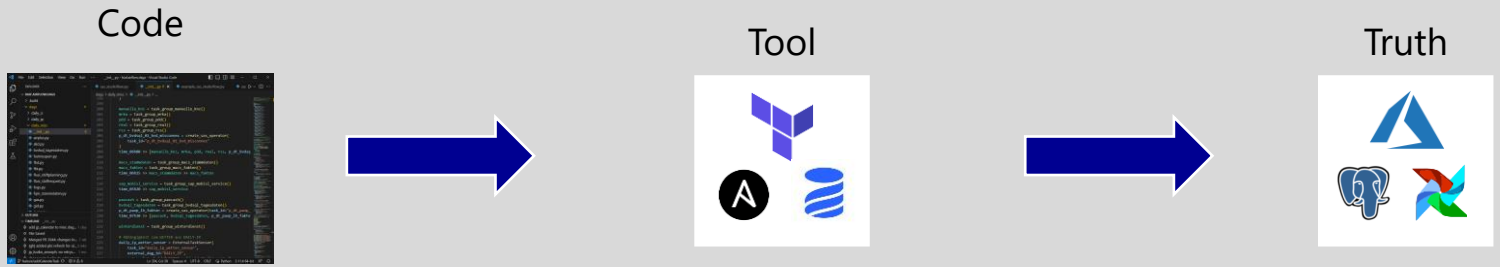
- Kunde möchte schnell sein, IT ist immer zu langsam und skaliert nicht genug
- IT möchte, dass keine Fehler nach Prod gehen
- Alles in Prod muss kontrolliert, getestet und freigegeben sein
- Transparenz schaffen
- Fail fast
- Automatisierung
- Dokumentation

Fraport Cloud Journey

Die Cloud Journey verändert die Arbeitsweise

Neue Prinzipien

- Infrastructure as Code -> Terraform
- Configuration as Code -> Ansible
- Database as Code -> Liquibase



Workload Scheduling

Was ist eigentlich Airflow?

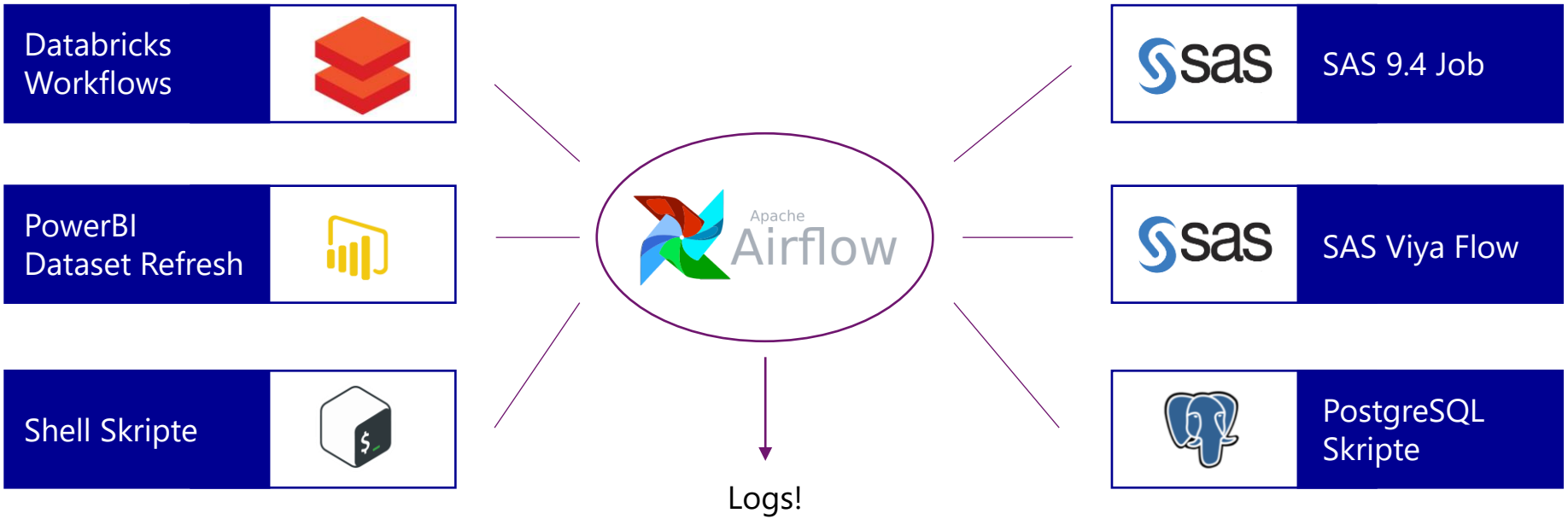
Fraport

Apache Airflow™

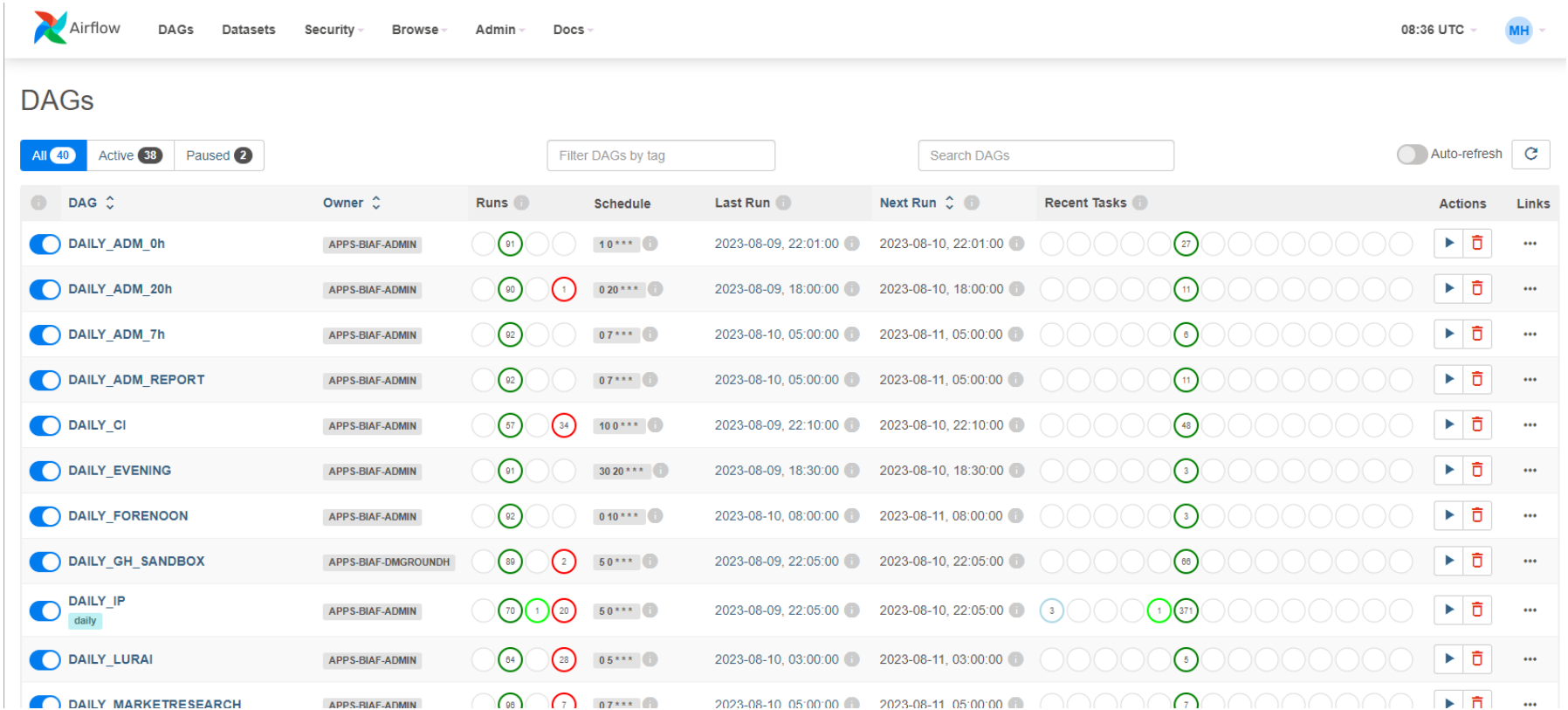
Airflow™ is a platform created by the community to programmatically author, schedule and monitor workflows.

https://airflow.apache.org/

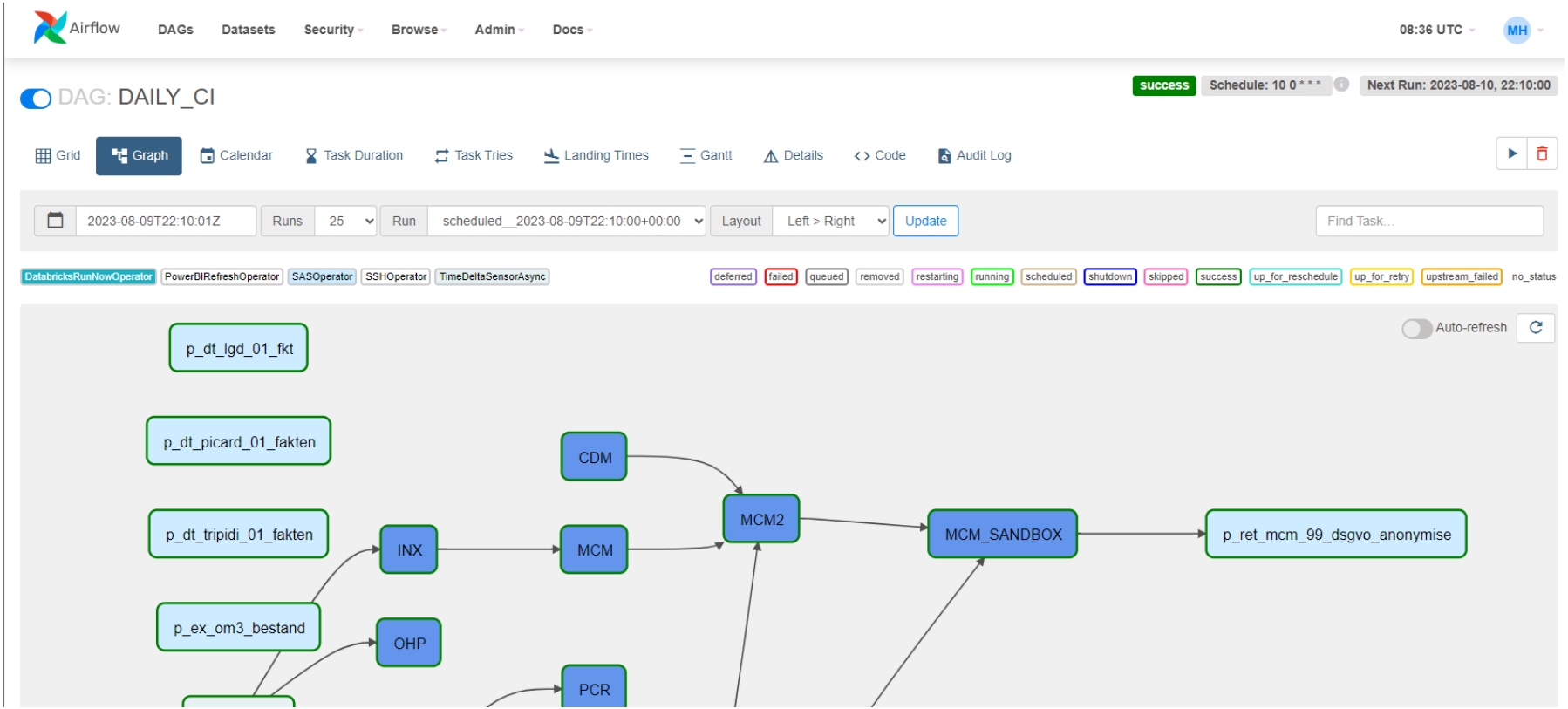
Welche Workload wird zeitgesteuert ausgeführt?



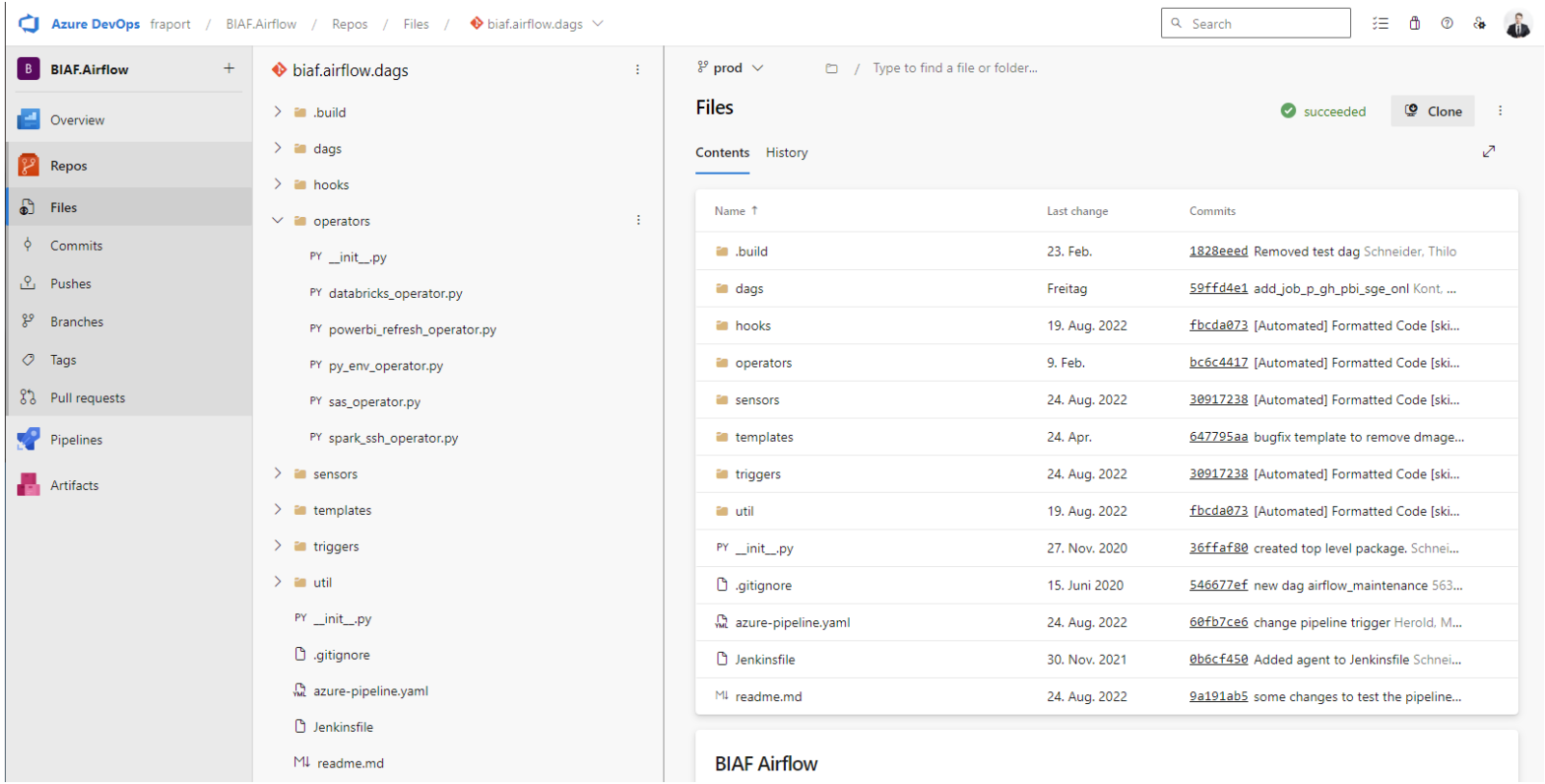
Airflow – Übersicht der DAGs



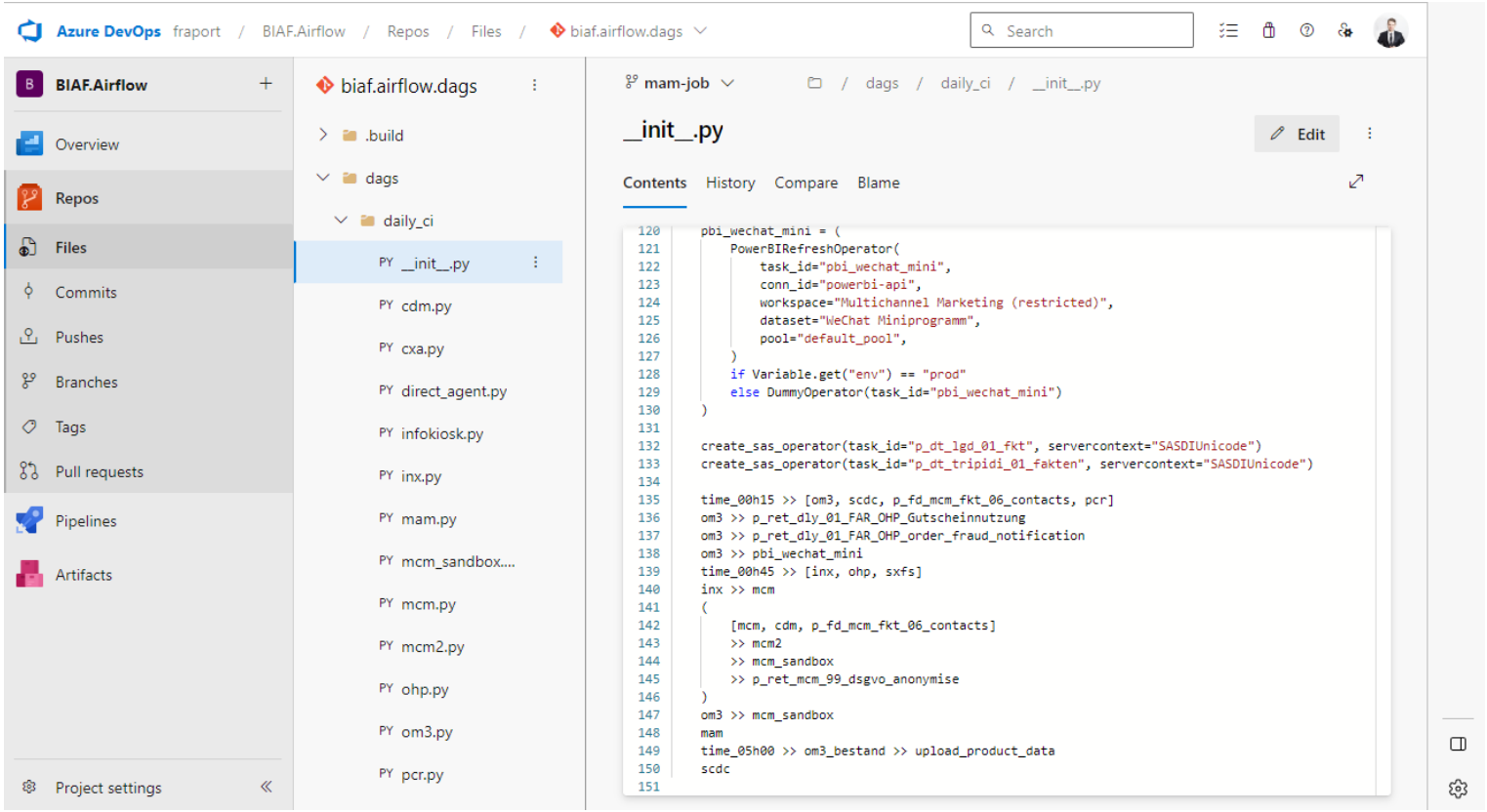
Airflow – Ein Beispiel für einen DAG



Airflow DAG Git Repository



DAG Beispiel



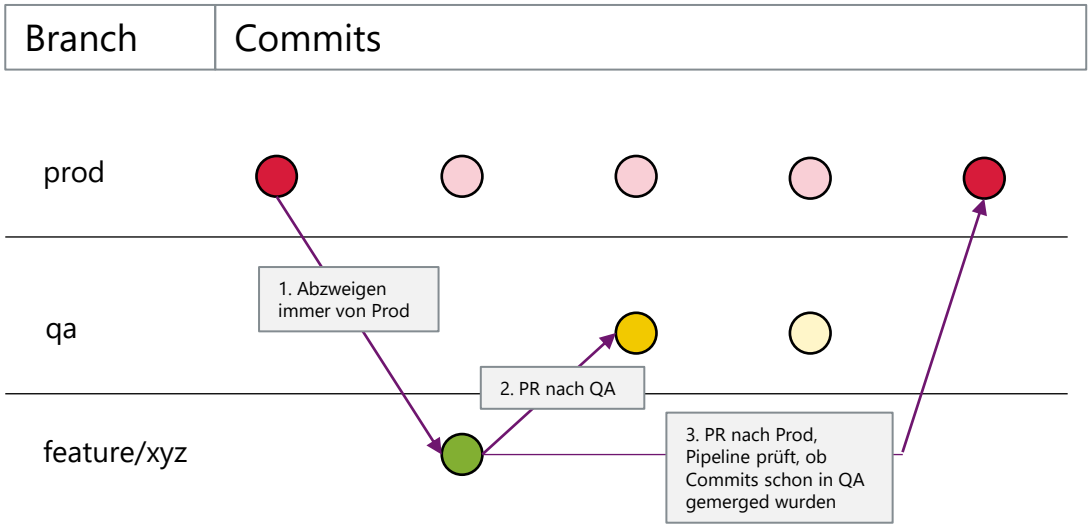
Konzept des Airflow Deployments

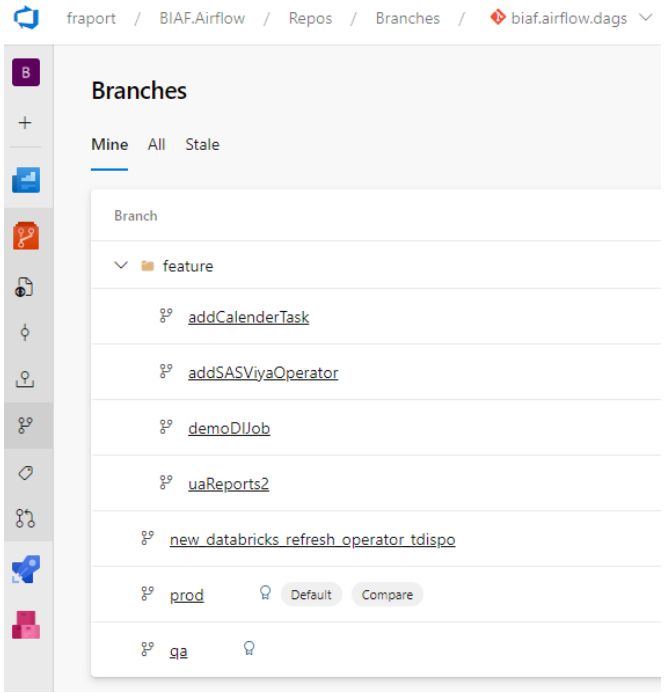
- Eine Änderung am Scheduling führt zu einem neuen Container Image für Airflow
- Jede neue Version durchläuft automatisch
  Funktionstests, Code Quality Checks, Code Formatierungen
- Ein volles Redeployment der Airflow Anwendung dauert ca. 10 Minuten
- Airflow kann ohne Downtime deployed werden
- Helm Deployment ermöglicht jederzeit einen Rollback
- Pipeline stellt sicher, dass keine Änderung nach Prod kann, die nicht in QA war

DAG Code ist einfach zu verstehen.

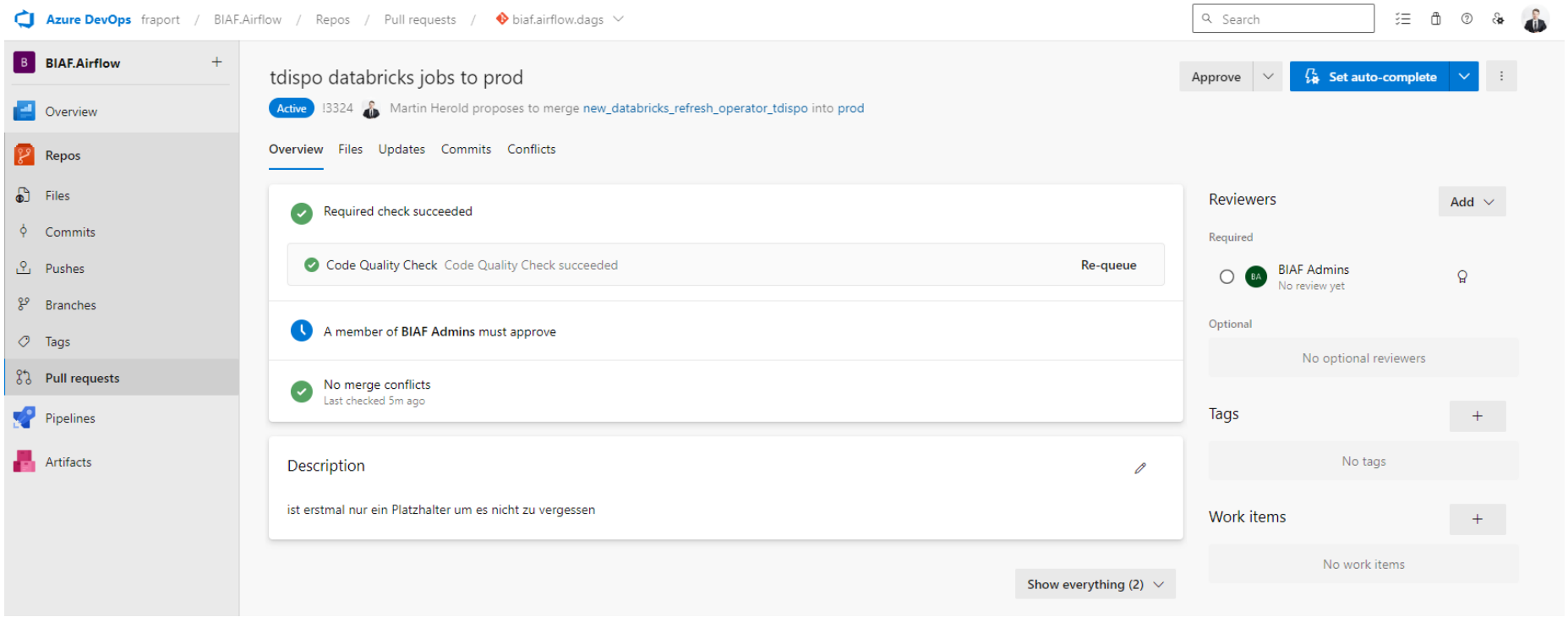
Git Workflow ermöglicht jedem eine Änderung durchzuführen

Ein mögliches Git Branching Konzept, welches sich bei Fraport etabliert hat:





Beispiel Pull Request



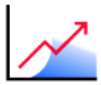
Datenbankanpassungen

Fraport

Liquibase

Flexible database change

Easily define changes in SQL, XML, JSON, or YAML.



Version control for your database

Order changes and standardize development.

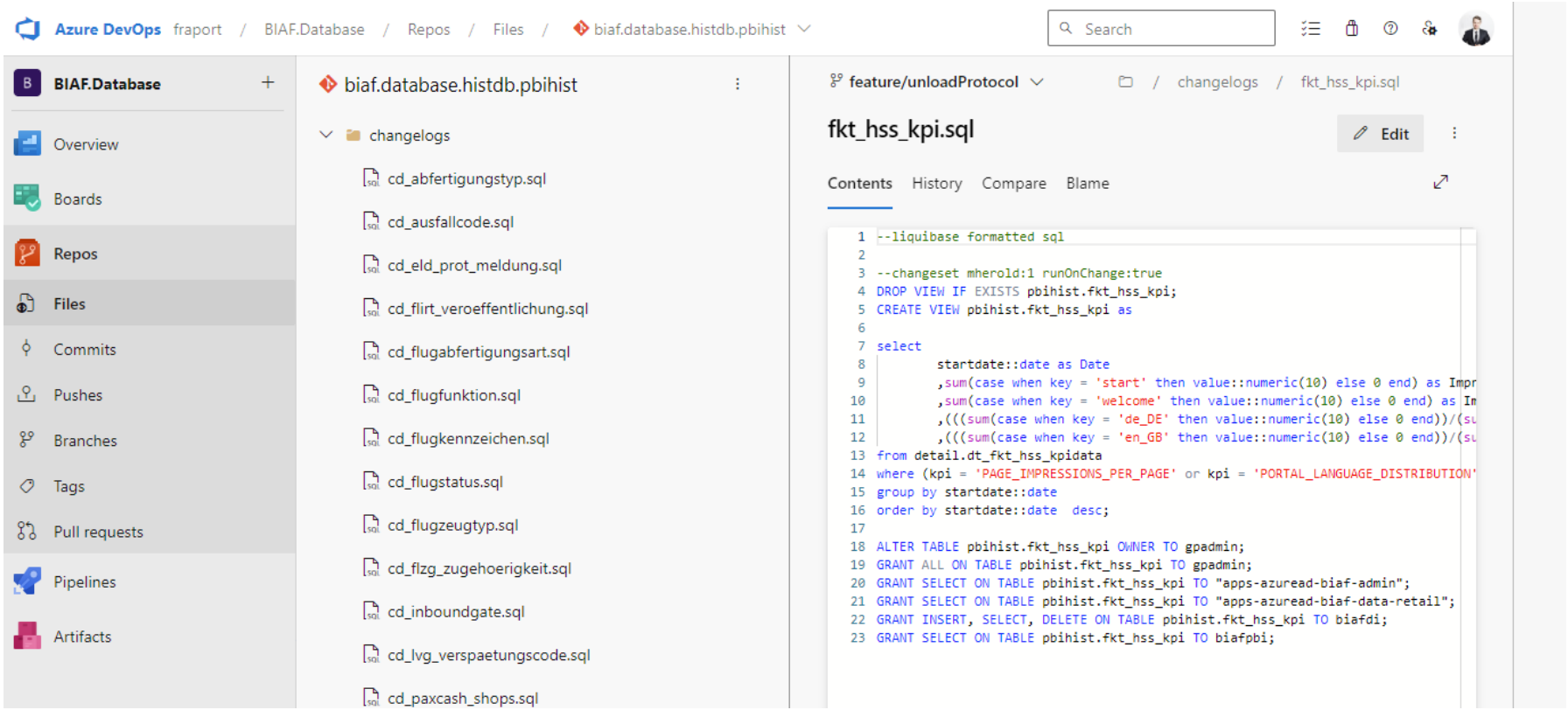
Built for developers

Control when, where, and how database changes are deployed.

Liquibase works with 50+ databases.

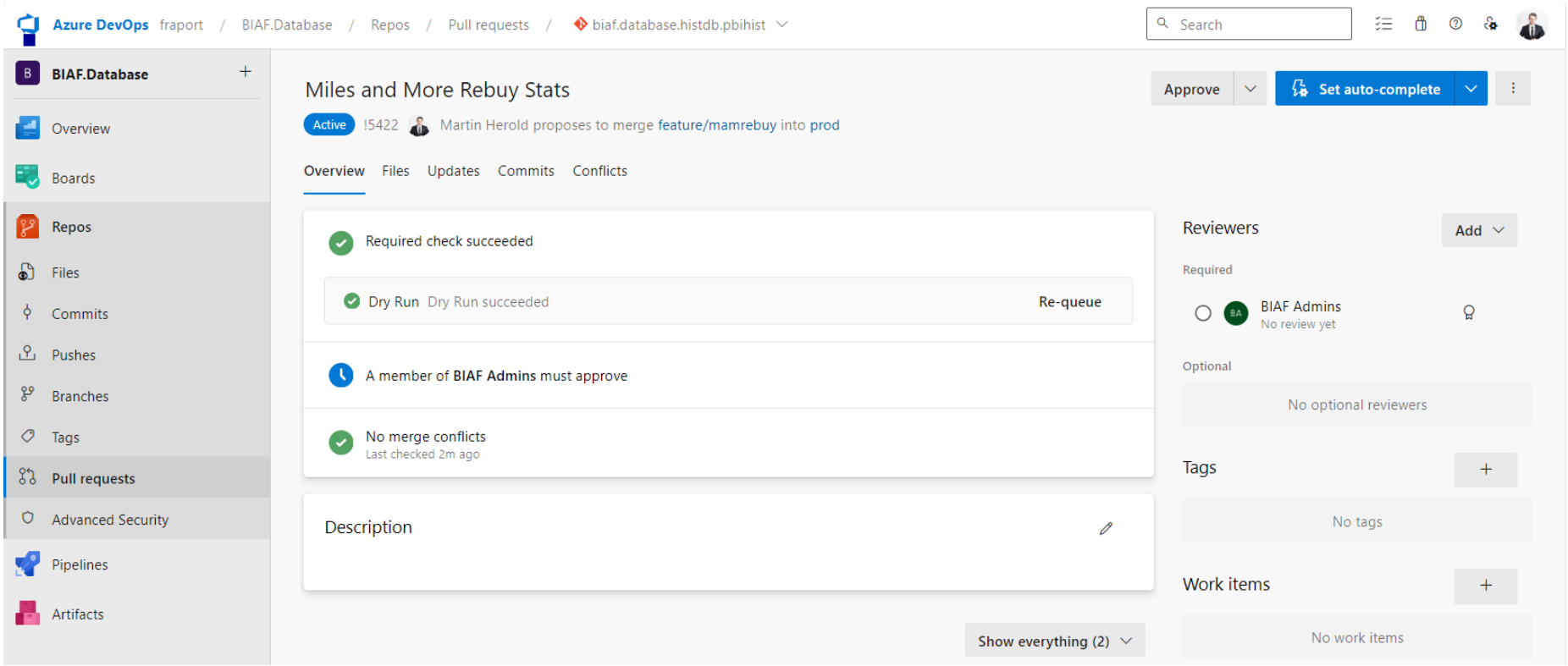
Beispiel – Liquibase Pipeline

Alle Datenbankobjekte sind als Changesets im Repo



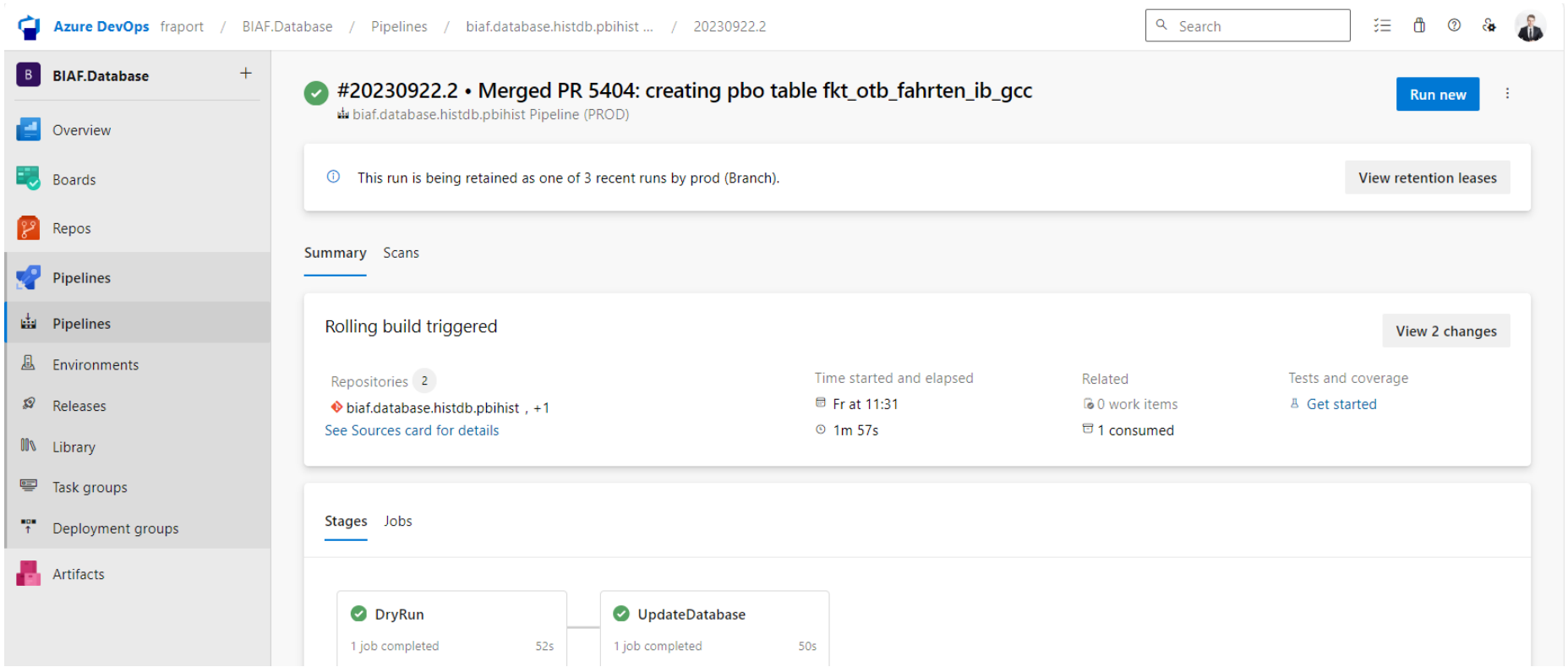
Beispiel – Liquibase Pipeline

Ein „Pull Request“ benötigt ein Review und eine Freigabe

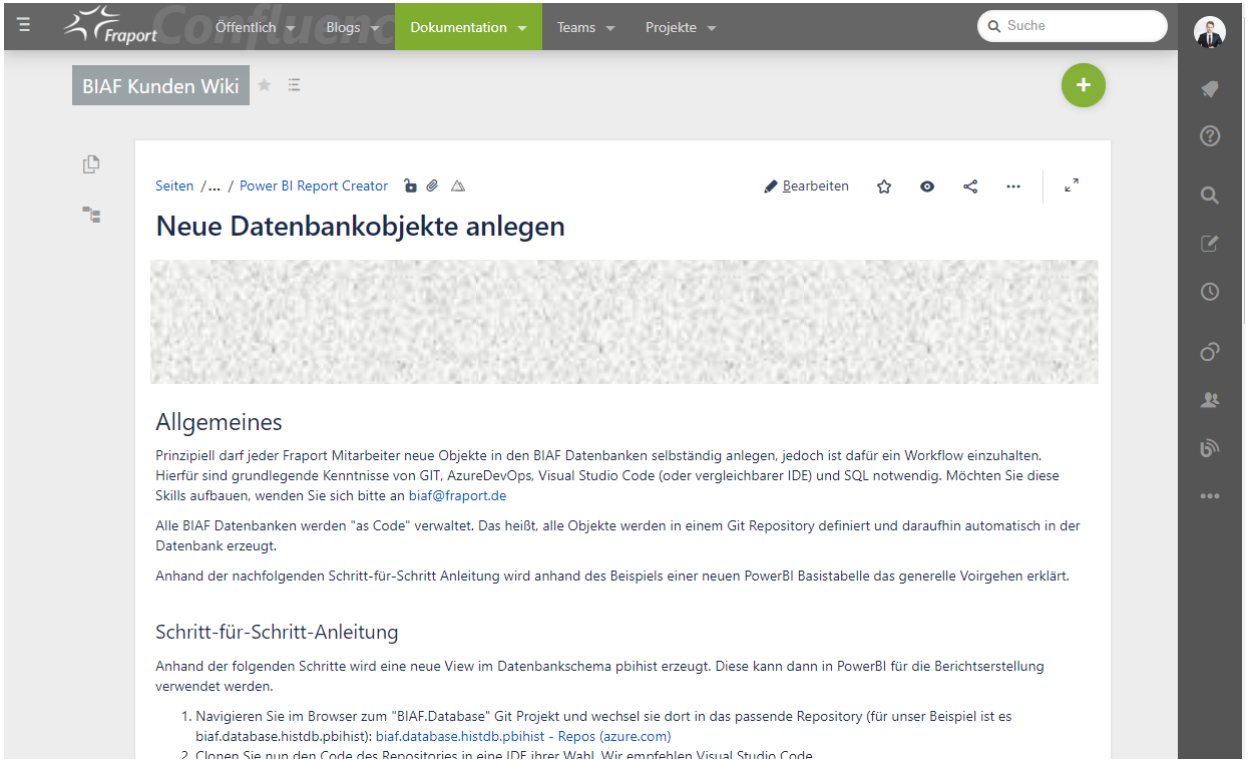


Beispiel – Liquibase Pipeline

Completed Pull Requests werden automatisch ausgerollt



Wissenstransfer? -> Wiki



Learnings

Wie machen wir es heute?

- Anfrage
  (Oft in Form von fertigem Code)
- Automatische Tests durch Pipelines
- Freigaben
  Automatisches Deployment stufenweise in alle Stages

- Dokumentiert
- Reproduzierbar
- Protokolliert
- Automatisch getestet
- QA und Prod sind gleich
- Entwickler brauchen keine Admin Rechte

Fraport DWH Umgebung

	@FRA
Wer legt Datenbankobjekte an?	IT, Analysten, Fachbereich, „Jeder“
Wie lange dauert es, bis ein Anwender eine neue Tabelle/View in der Datenbank bekommt?	< 30 Minuten
Wer konfiguriert die automatische Ausführung der ETL-Jobs?	IT, Analysten, Fachbereich, „Jeder“
Wie oft verändert sich der Ablaufplan der eingeplanten ETL-Jobs?	Same Day Delivery, mehrmals täglich

IT-Betriebsteams: 1 Person, rollierend jede Woche

Fraport

Gute Reise!
We make it happen