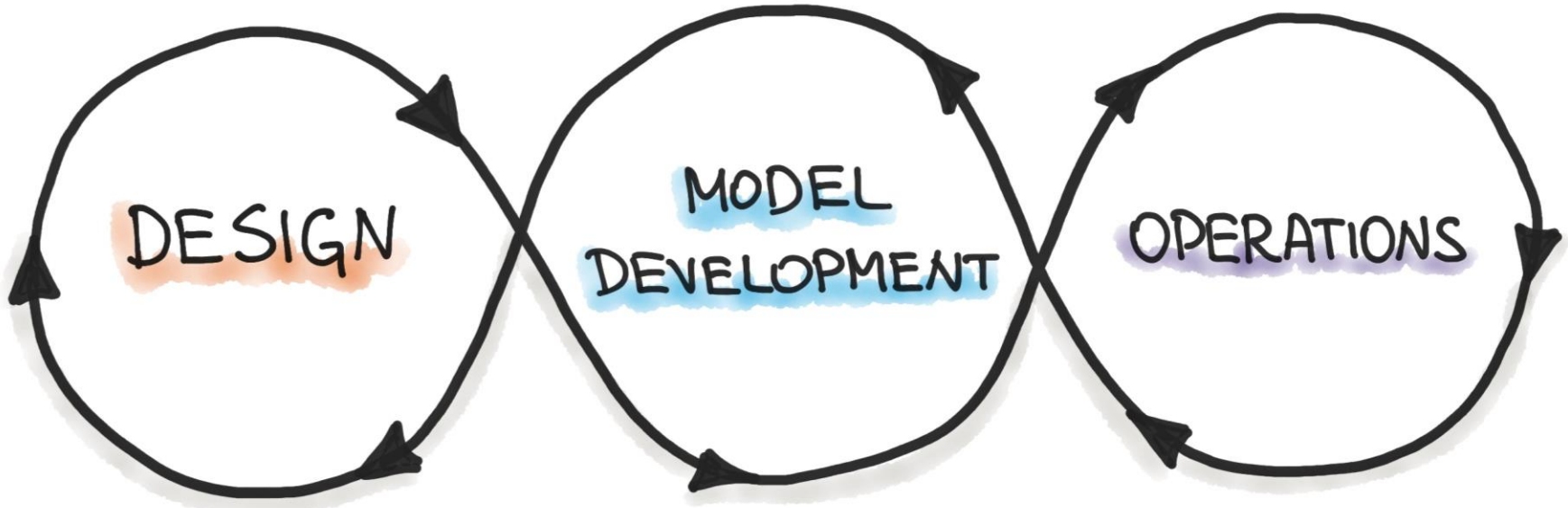
MLOps

Marcel Thomas
Kristian Paulic
Simon Jakschitz
Alexander Wallner
Vincent Mattes
Glenn Verhaag

Projekt 2 – Prof. Decker

MLOps



- Requirements Engineering
- ML Use-Cases Priorization
- Data Availability Check

- Data Engineering
- ML Model Engineering
- Model Testing & Validation

- ML Model Deployment
- CI/CD Pipelines
- Monitoring & Triggering

Use Case

Stock price prediciton

- Ziel: Vorhersage von Aktienkursen
- Annahme: Startbudget = 15.000€
- Annahme: Gebühren und Steuern werden ignoriert
- Einschränkung: Nur bestimmte Aktien werden gehandelt

Tech stack

Trading

Alpaca

Language

Versioning

Experimentation

mlflow™

SDK/
Interface

Containerization

Model training

LUDWIG

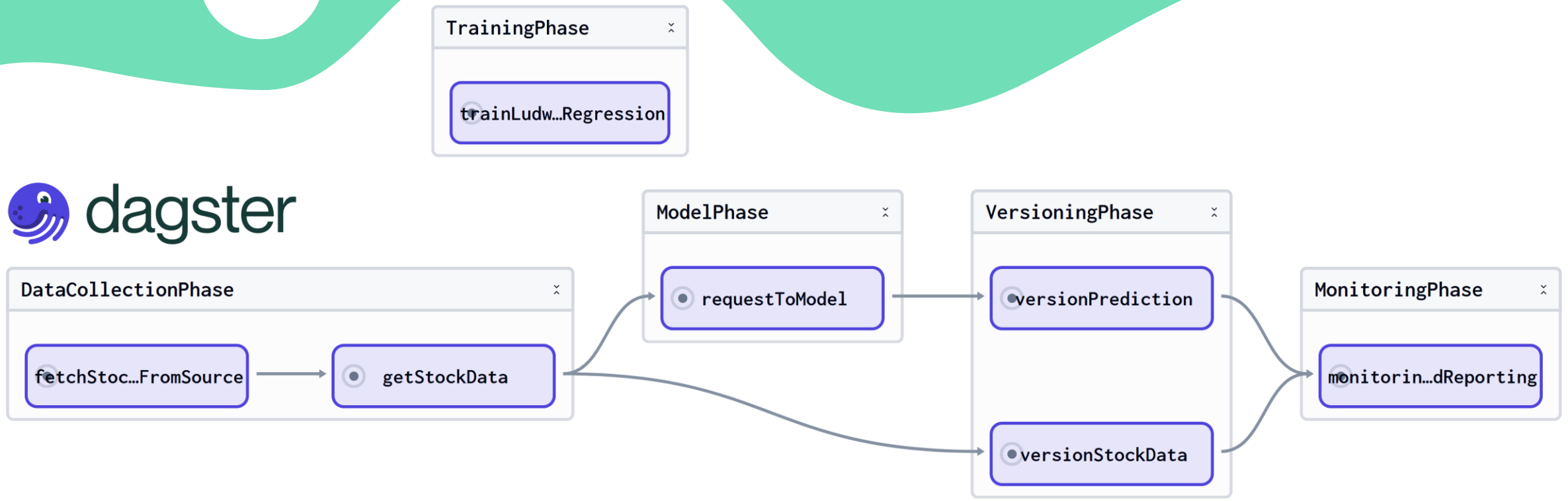
Data storage

MINIO

Orchestration

Monitoring

Recap Projekt 1



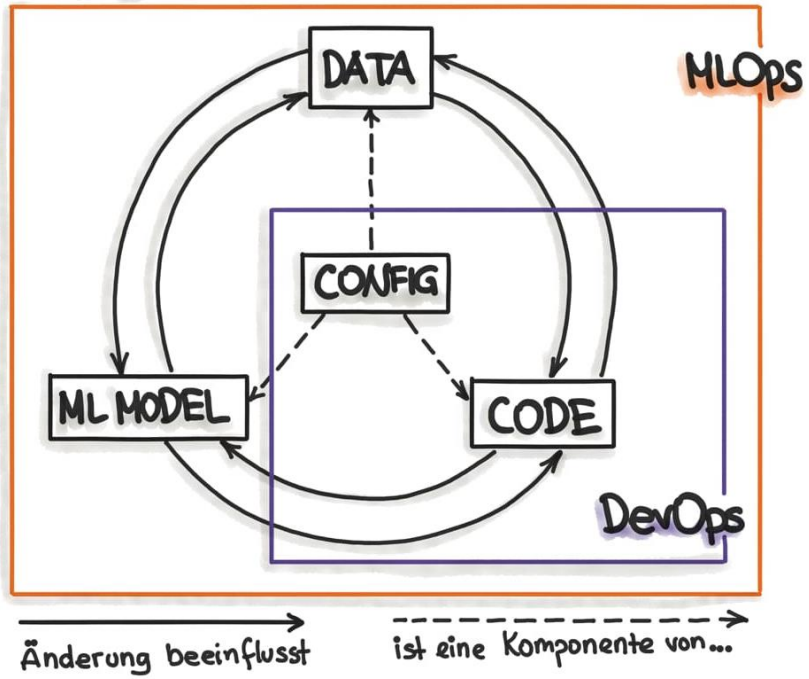
Reproduzier-barkeit

Anforderung:

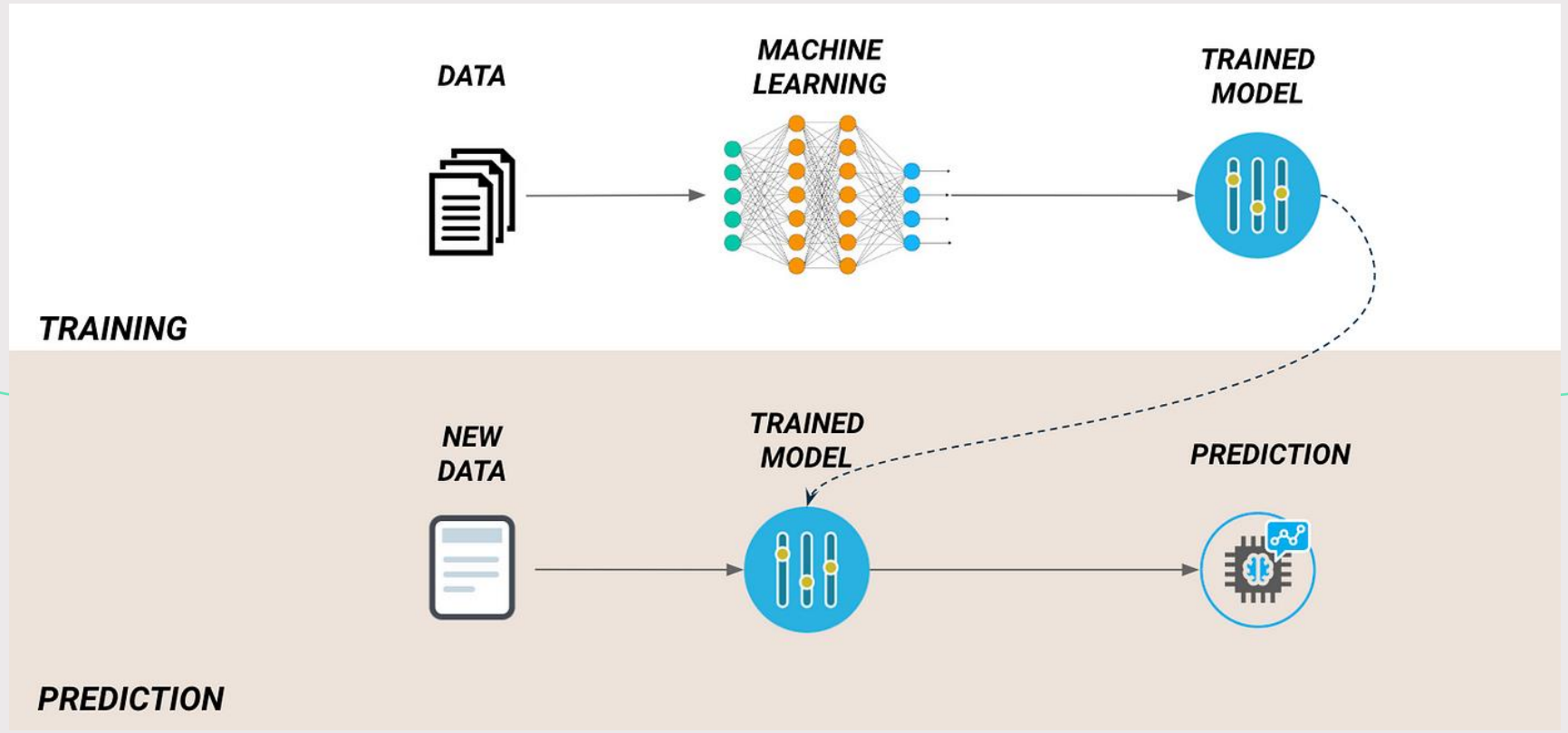
- Vollständiger Neustart der Infrastruktur

Ziel:

- Sicherung und Wiederherstellbarkeit von Code, Data, Models und Funktionalität der Pipelines
- Verbessern der Prozesse im Kontext “Desaster Recovery “



Model Training - Struktur

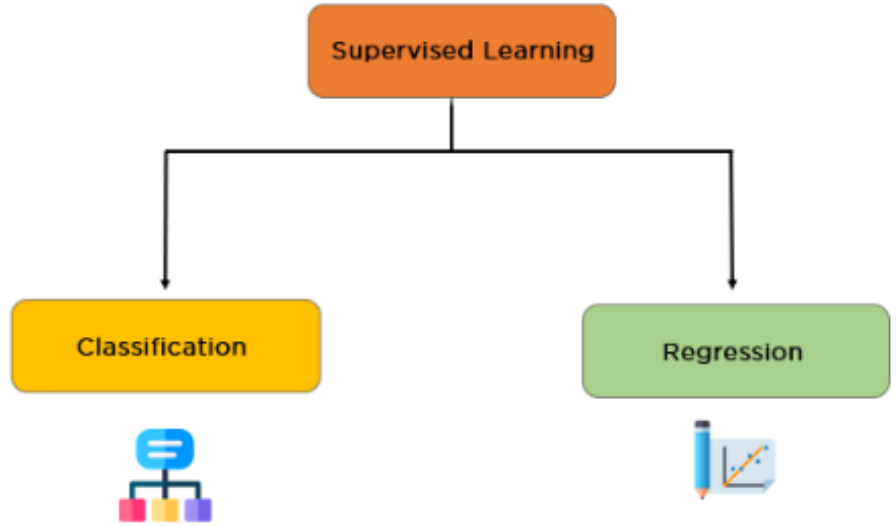


Model Training - Tuning

Mögliche Modeltypen

Ziel:

- Welcher Typ passt zu unserem Problem?
- Auswertung der Performance anhand von Machine Learning Metriken (R^2, Precision, MAE …)
- Tuning Ansätze (Epochenerweiterung, Preprocessing …)



Default

Provide Feedback | Add Description

metrics.rmse < 1 and params.model = "tree" | Time created | State: Active | Datasets | Sort: Created | Columns | Group by

Table | Chart | Evaluation Experimental

Run Name	Created	Dataset	Duration	Source	Models
Apple_regressor_rot_2024-09-25_17-32-05	3 days ago	-	1.9min	__main__...	-
IBM_regressor_rot_2024-09-24_20-23-38	4 days ago	-	3.9min	__main__...	-
Microsoft_regressor_rot_2024-09-24_20-18-59	4 days ago	-	4.1min	__main__...	-
SAP_regressor_rot_2024-09-24_20-12-15	4 days ago	-	3.0min	__main__...	-
Amazon_regressor_rot_2024-09-24_20-11-15	4 days ago	-	6.4min	__main__...	-
Nike_regressor_rot_2024-09-24_20-07-39	4 days ago	-	7.5min	__main__...	-
SAP_regressor_rot_2024-09-24_19-57-30	4 days ago	-	1.8min	__main__...	-
Tesla_regressor_rot_2024-09-24_19-46-48	4 days ago	-	2.3min	__main__...	-
Google_regressor_rot_2024-09-24_19-26-32	4 days ago	-	4.7min	__main__...	-
IBM_regressor_2024-09-24_18-44-51	4 days ago	-	26.1min	__main__...	-
IBM_regressor_blue_2024-09-23_13-59-39	6 days ago	-	1.1min	__main__...	-
Google_regressor_test_2024-09-23_13-19-04	6 days ago	-	22.8s	__main__...	-
Google_regressor_2024-09-21_09-22-33	8 days ago	-	8.5min	__main__...	-
Amazon_regressor_blue_2024-09-18_13-40-52	11 days ago	-	1.2min	__main__...	-
Google_regressor_blue_2024-09-09_21-41-52	19 days ago	-	23.4s	__main__...	-
Google_regressor_blue_2024-09-09_20-16-26	19 days ago	-	23.2s	__main__...	-

Show

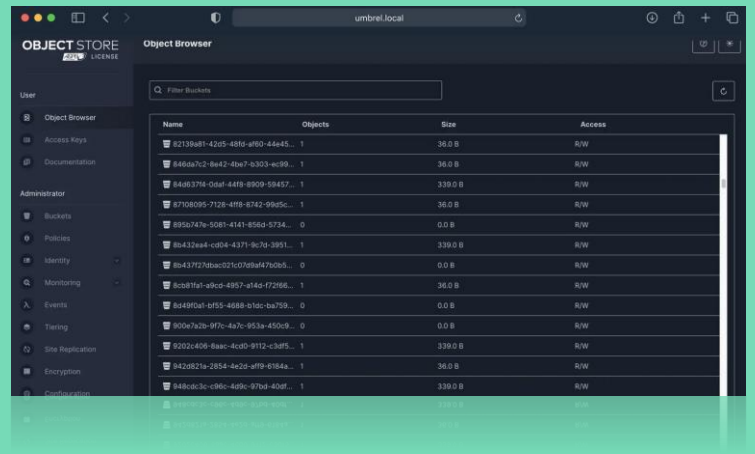
Trading

Alpaca

Abfragen der **Account-balance**

Trade anhand der **Predictions** ausführen

MINIO

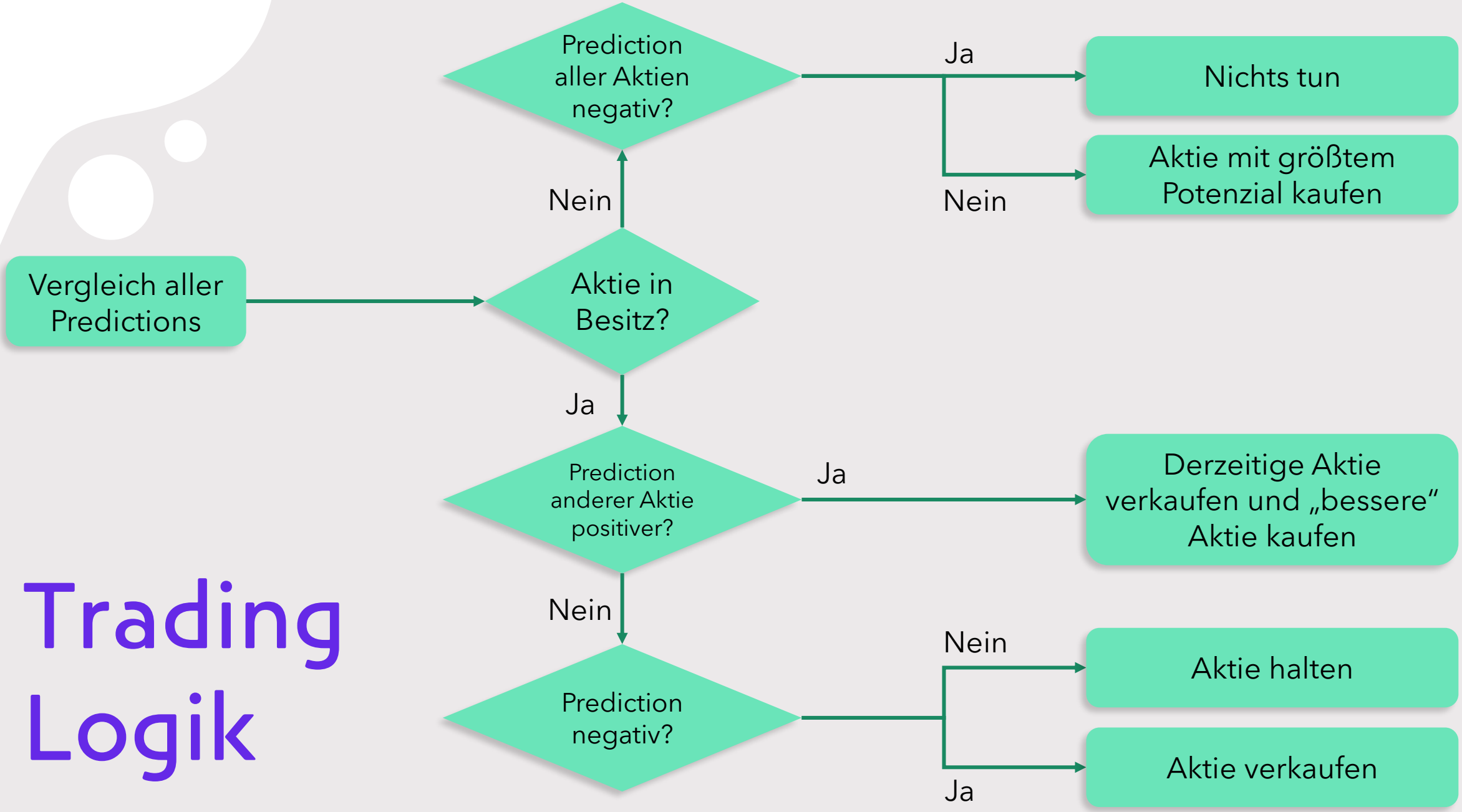


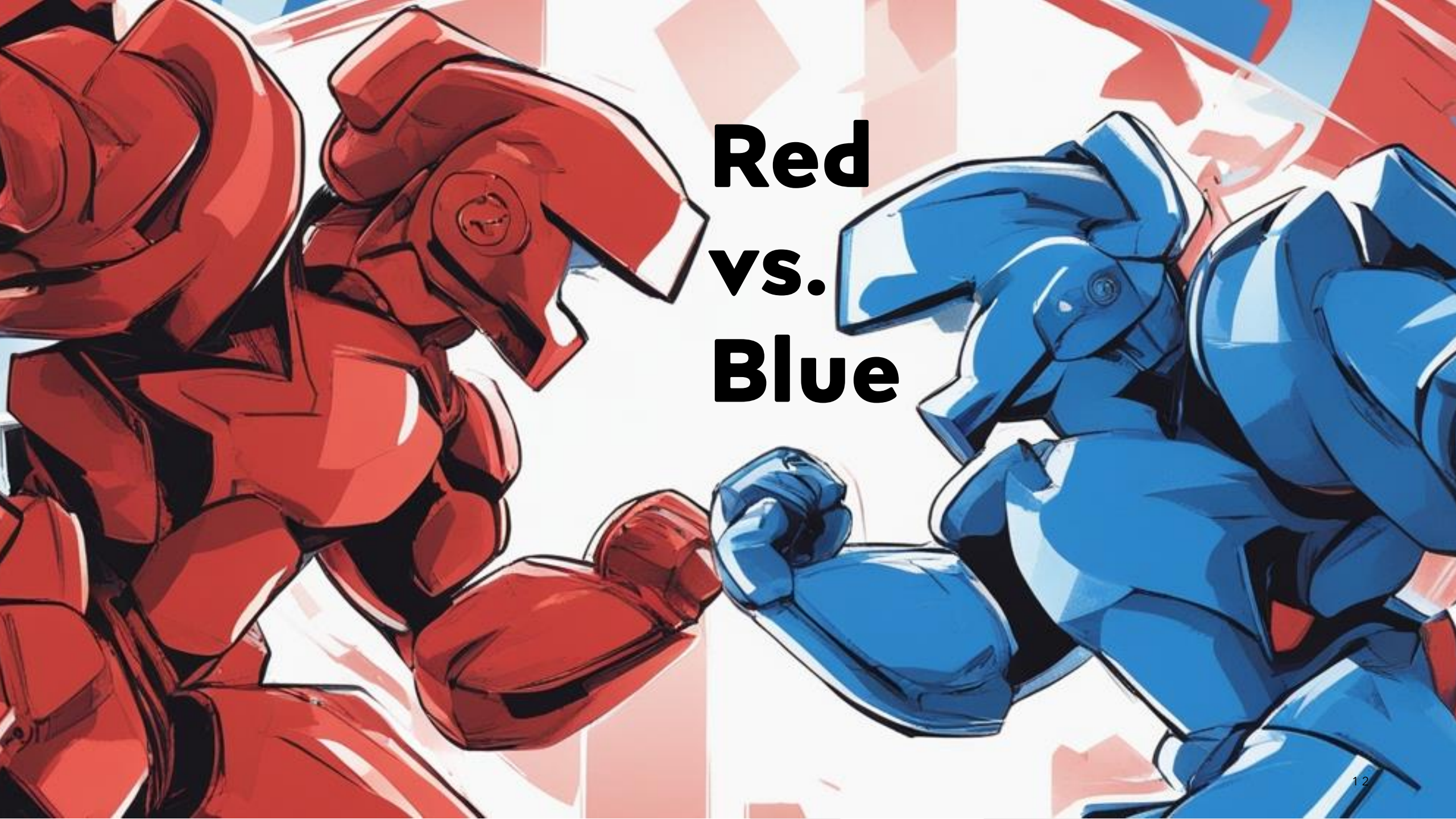
Predictions beziehen

TradingPhase

tradeScript

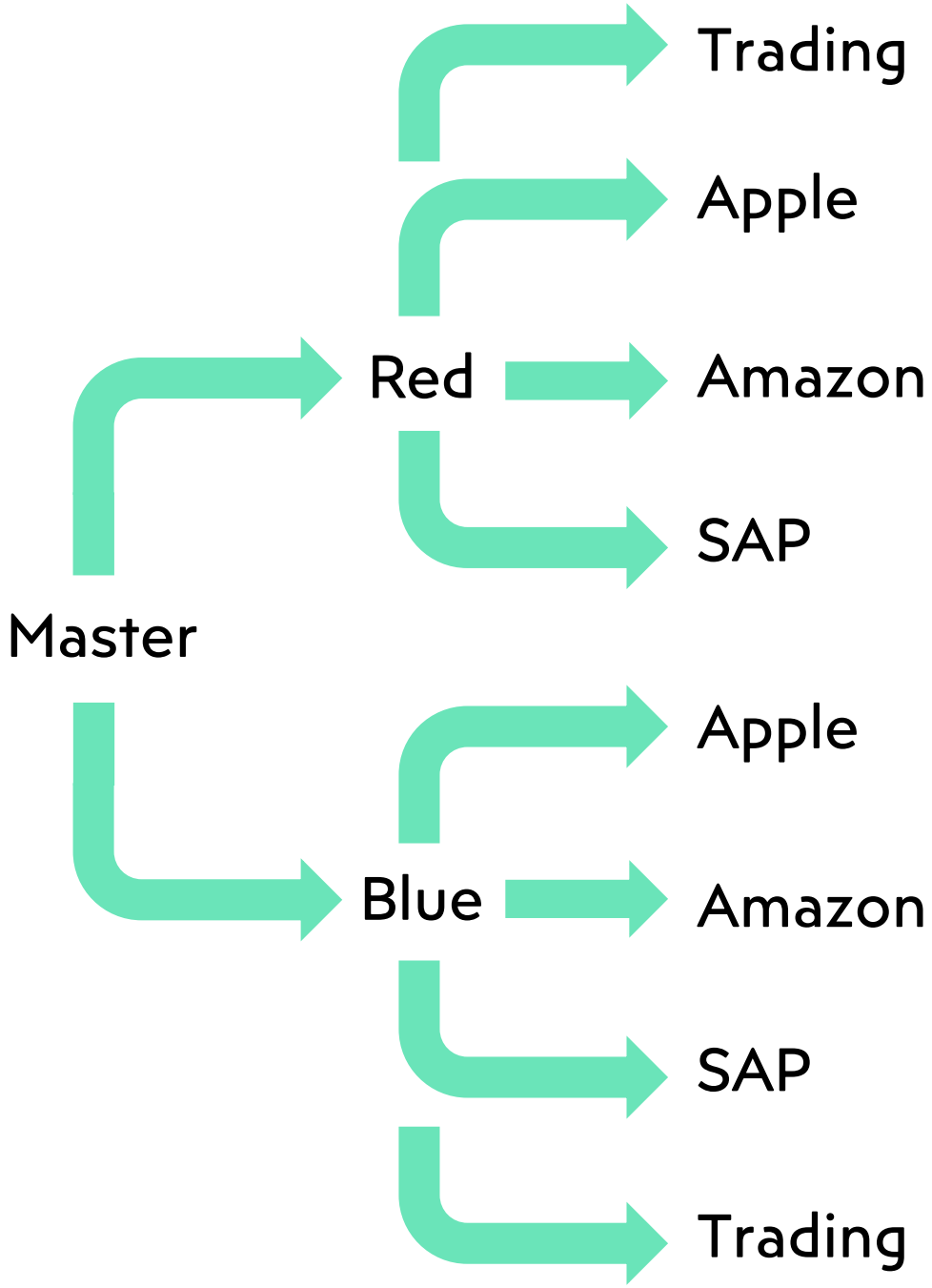
Trading Logik




Red
vs.
Blue

Branching

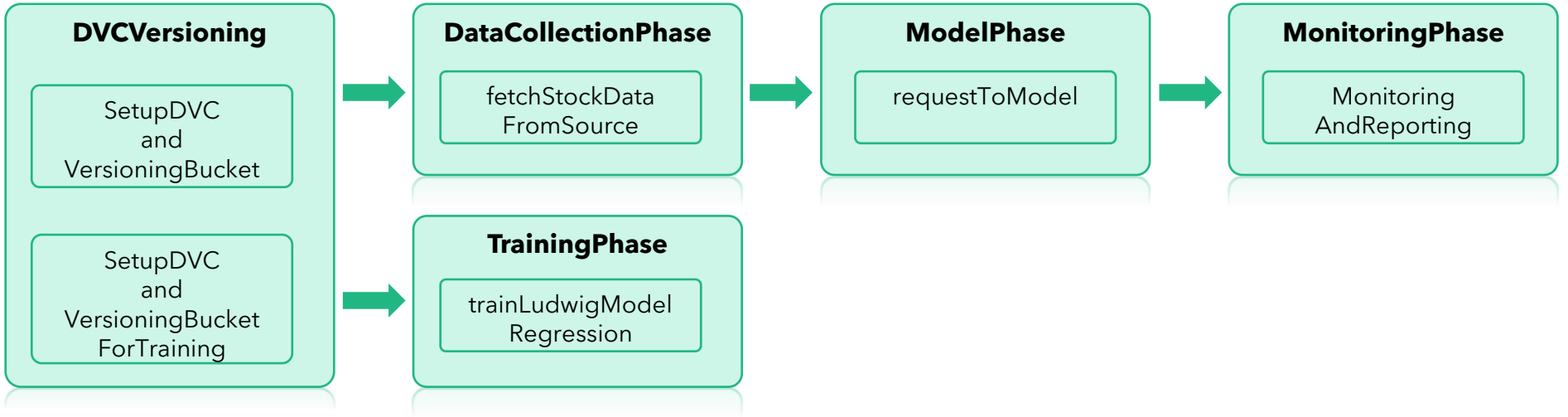
- Eine Pipeline pro Branch
- Aufteilung nach Team und Aktie
- Custom Config files etc. pro branch
- Versionierte Artefakte werden pro Pipeline in die entsprechende Branch gepusht



Pipelines

Model Training & Prep

eine Pipeline je Aktie und Team → ***6 Stk.***



Trading

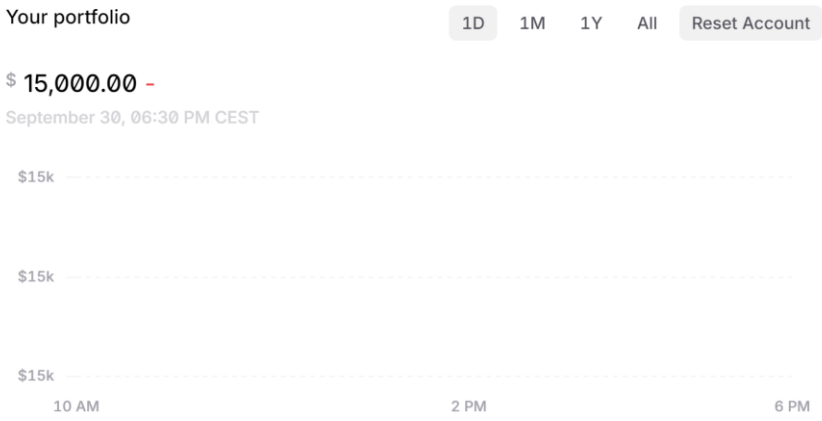
eine separate Pipeline je Team → ***2 Stk.***

8 Stk.

Model- und Tradingperformance

TEAM BLUE

```
------------------------------ Statistics ------------------------------

Simulation period: 2024-01-02 + 180 days

Apple, Model 1
Mean Absolute Error (MAE): 43.5727
Percentage of correct Gain/Loss predictions: 44.44%

SAP, Model 1
Mean Absolute Error (MAE): 13.3266
Percentage of correct Gain/Loss predictions: 45.0%

Amazon, Model 1
Mean Absolute Error (MAE): 15.92
Percentage of correct Gain/Loss predictions: 47.22%

------------------------------------------------------------------------
```

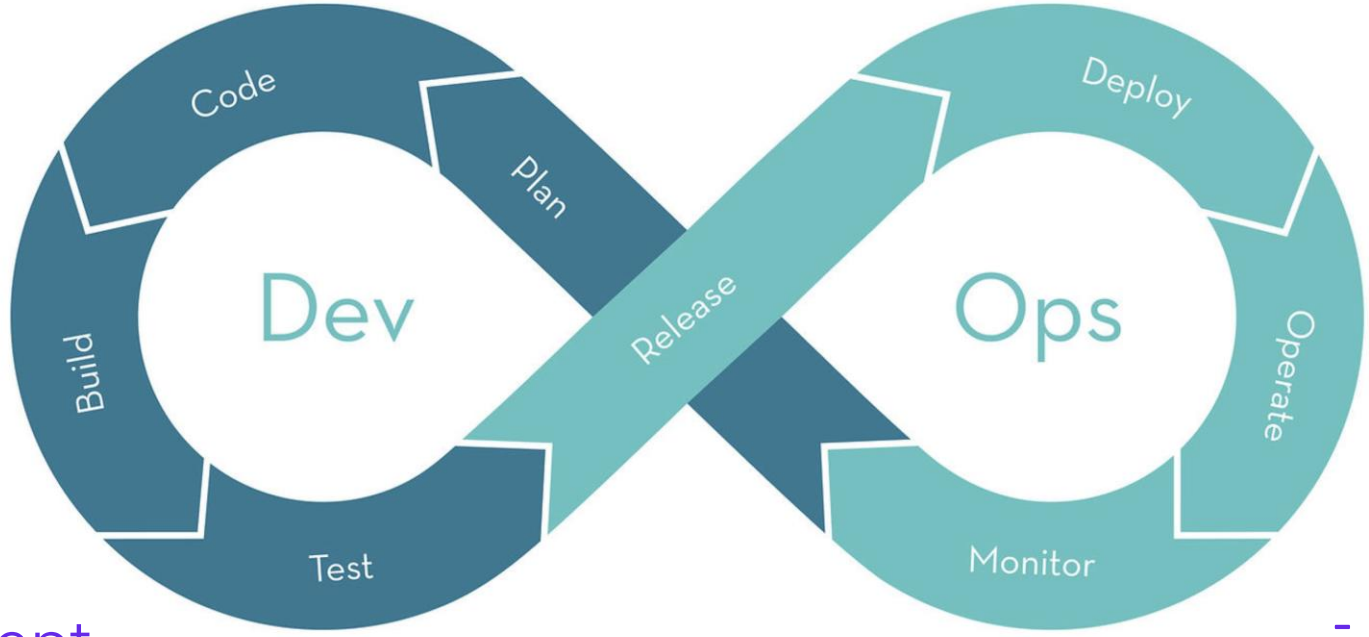


TEAM RED

```
------------------------------ Statistics ------------------------------

Simulation period: 2024-01-02 + 180 days

Apple, Model 1
Mean Absolute Error (MAE): 47.2786
Percentage of correct Gain/Loss predictions: 44.44%

SAP, Model 1
Mean Absolute Error (MAE): 120.9709
Percentage of correct Gain/Loss predictions: 45.0%

Amazon, Model 1
Mean Absolute Error (MAE): 378.8522
Percentage of correct Gain/Loss predictions: 52.22%

------------------------------------------------------------------------
```

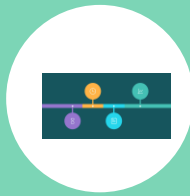
Recap Projekt 2

- Konzept
- Entwicklung
- Proof of Concept



- Skalierbarkeit
- Zuverlässigkeit
- Weiterentwicklung

Challenges



- Komplexität
- Islands of Experts
- Skalierbarkeit
  - z.B. Models auf verschiedene Aktien
- Verzögerungen im Zeitplan
- Fehlende Flexibilität
  - z.B. im Fall von Code-Änderungen/Bugfixing

Lessons Learned

- ‘Ops’-Aspekt größte Herausforderung
- Bugfixing VOR Rollout
- Klare Voraussetzungen definieren bevor das Projekt gestartet wird
- Klarere Dokumentation

Key Takeaways & Ausblick

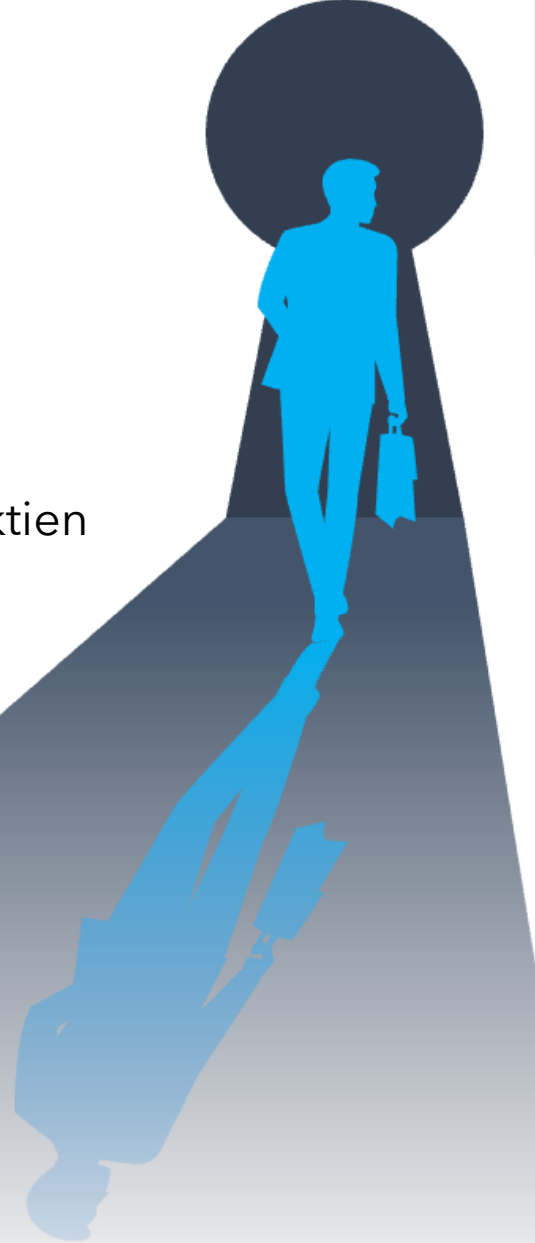
01 Weiterentwicklung der Models

02 Live-Trading als Folgeprojekt

03 Automatisierung diverser Pipeline-Steps

Mögliche neue Features:

Kombination von mehreren Aktien



ML-Ops Projekt benötigt stetige Weiterentwicklung