SHERLOQ

KI-Lösung für Intelligente Dokumentenverarbeitung & Kundenkommunikation

Hallo Java User Group!

a solution by codecentric

KI-gestützte Automatisierung im Dokumentenprozess

a solution by codecentric

SHERLOQ

Wir machen…

- **Projekte** und nicht nur Technik
  Beratung, Integration, …
- **On-Prem** wenn gewünscht
  fast immer
- **KI** eingestellt auf **Kundendaten**
  train + evaluate + explain
- **faires Lizenzmodell**
  pay for success

statt…

- komplette **Produkt-Suite** für Input-Management
- reine **Cloud-Lösung** oder **Software-as-a-Service**
- **AI-as-a-Service** und **general-purpose-Modelle**

Wie sehen typische

SHERLOQ-Projekte

aus?

Beispiel #1

Klassifikation im Posteingang eines Energieversorgers

Hilferuf

In unserer Poststelle bleiben E-Mails
und Scans xxx Tage unbearbeitet liegen.

PoC

X Dokumente
Y Klassen

Beispiel #1

Klassifikation im Posteingang eines Energieversorgers

Hilferuf

In unserer Poststelle bleiben E-Mails
und Scans xxx Tage unbearbeitet liegen.

PoC

X Dokumente
Y Klassen

Risiko

Einbindung in bestehende
Input-Management-Suite?

Bedenken

Kundendaten aus dem Archiv
für das Anlernen einer KI?

Beispiel #1

Klassifikation im Posteingang eines Energieversorgers

Hilferuf

In unserer Poststelle bleiben E-Mails
und Scans xxx Tage unbearbeitet liegen.

PoC

X Dokumente
Y Klassen

Risiko

Einbindung in bestehende
Input-Management-Suite?

Bedenken

Kundendaten aus dem Archiv
für das Anlernen einer KI?

Beauftragung MVP

Beispiel #1

Klassifikation im Posteingang eines Energieversorgers

Ziel des MVP

Erkennung der **10 häufigsten Klassen**

Daten

- aus dem **Archiv**
- Scans & E-Mails als **PDFs** von der Input-Management-Suite
- **500-10.000** Dokumente/Klasse und **200.000** vom ganzen Rest

Integration

- KI als **REST-Service**
- Anbindung durch Kunden-IT (MuleSoft, RabbitMQ)

Infrastruktur

- On-Prem Linux-VMs für Training, Test, Prod

Beispiel #2

Extraktion aus Lieferscheinen

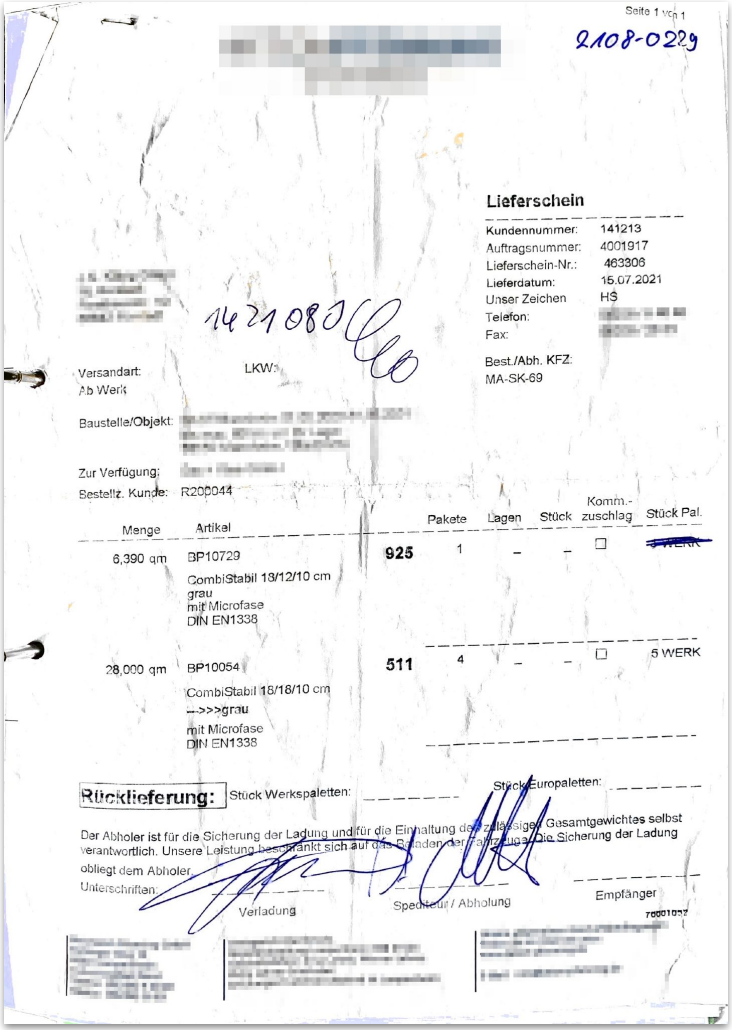
Ziel PoC

Verprobung der Extraktion von

- Lieferdatum
- Lieferscheinnummer
- Kreditor
- …

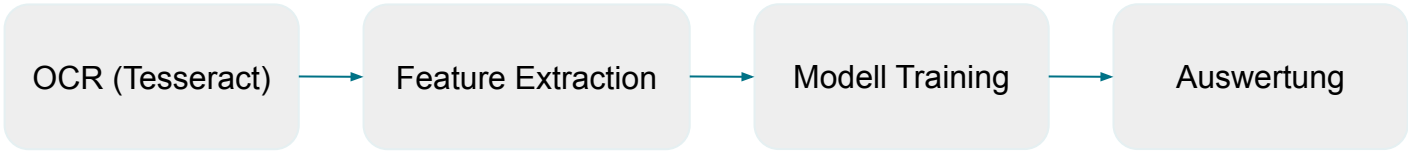
Daten

- **821** Seiten Scans / Fotos
- davon nur **70%** lesbar
- keine Werte / Annotationen
- “Macht damit, was ihr wollt.”

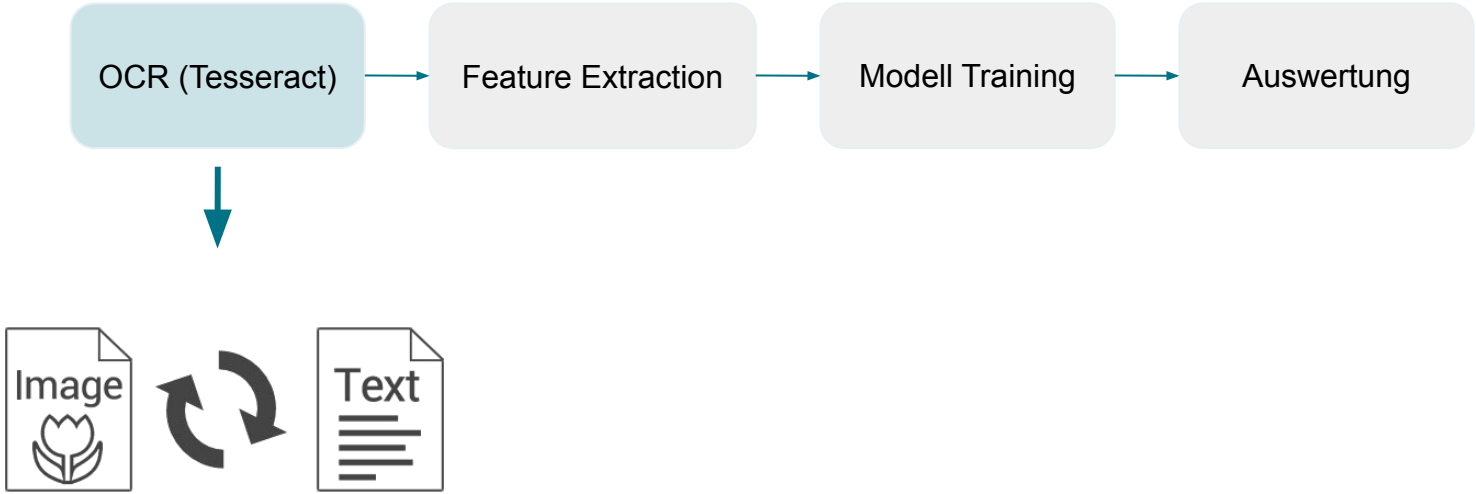


SHERLOQ
AI Überblick

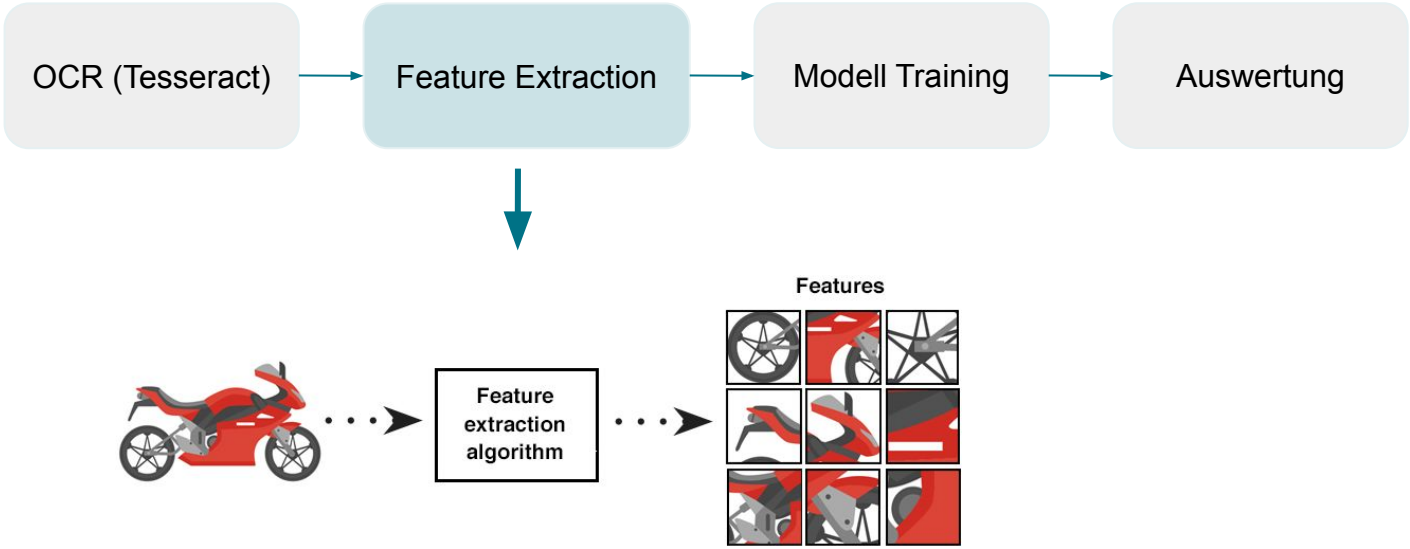
Von Kundendaten zum trainierten Modell



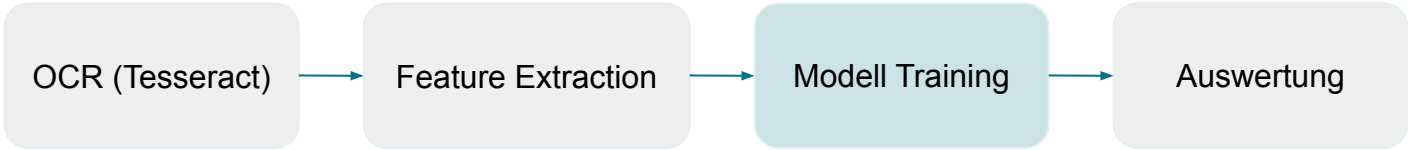
Von Kundendaten zum trainierten Modell

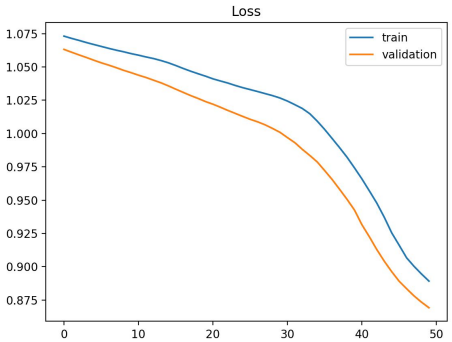


Von Kundendaten zum trainierten Modell

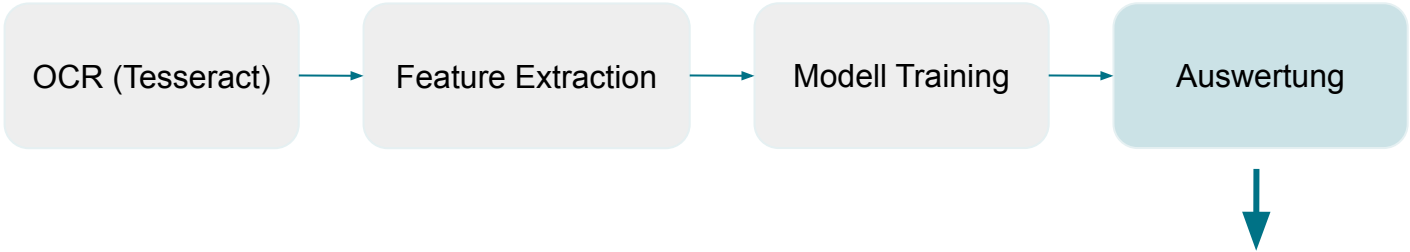


Von Kundendaten zum trainierten Modell





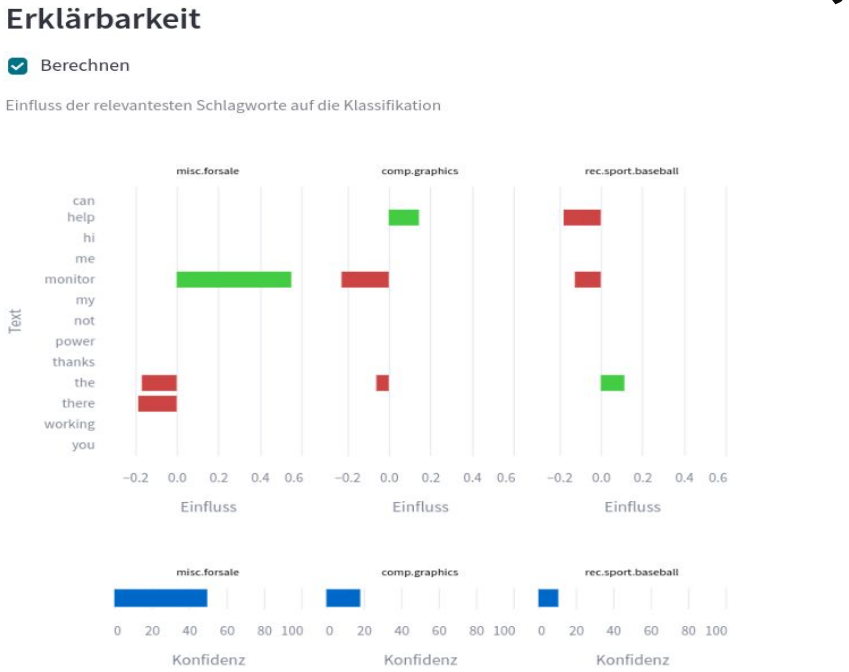
Von Kundendaten zum trainierten Modell



	Result
Accuracy	77.33%
Precision	78.76%
Recall	88.46%
F1	83.33%

Anwendungsfall Klassifikation

- **Training auf Kundendaten (AI-Lab)**
  - PDFs, E-Mails
  - xgboost, spacy, keras, …
- **Auswertungs-Dashboard**
  - explainable AI mit lime
  - Live-Test
- **Bereitstellung als REST-Service**
  - optionales Regelwerk



SHERLOQ

ML/DL für Text-Klassifikation

Entwicklung der Ansätze

Eingabe

Text als

- **Bag of Words / Wort-Statistiken**

Modelle

klassisches ML

dmlc XGBoost

- **Entscheidungsbäume** *on steroids*

LightGBM

ML/DL für Text-Klassifikation

Entwicklung der Ansätze

Eingabe

Text als

- Bag of Words / Wort-Statistiken
- **Zeichenfolge**
- **Wortfolge**

Modelle

klassisches ML

- Entscheidungsbäume *on steroids*

dmlc XGBoost

LightGBM

oder neuronale Netze / DL

- **Faltungsnetze**
- **rekurrente Netze**

PyTorch

TensorFlow

Transfer Learning

- vortrainierte **Wort-Einbettungen**

ML/DL für Text-Klassifikation

Entwicklung der Ansätze

Eingabe

Text als

- Bag of Words / Wort-Statistiken
- Zeichenfolge
- Wortfolge
- **Subtokenfolge**

Hugging Face

Modelle

klassisches ML

dmlc XGBoost

- Entscheidungsbäume *on steroids*

LightGBM

oder neuronale Netze / DL

PyTorch

- Faltungsnetze
- rekurrente Netze

TensorFlow

- **Transformer**

Hugging Face

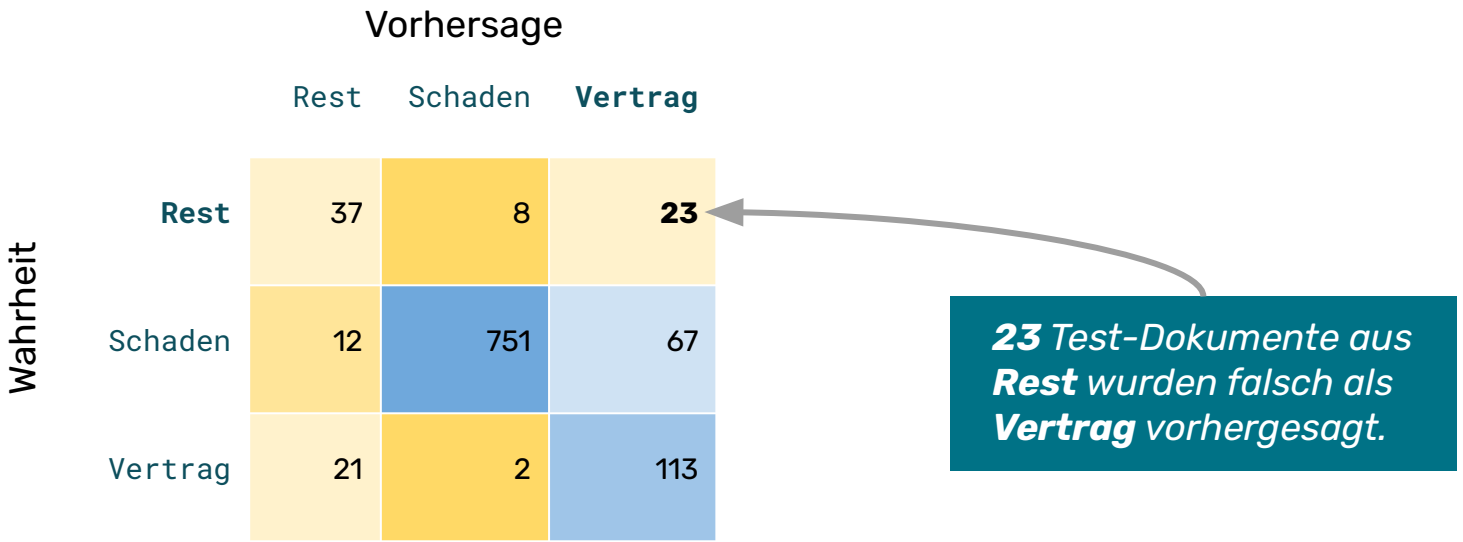
Transfer Learning

- vortrainierte Wort-Einbettungen
- vortrainierte **Sprachmodelle**

Hugging Face

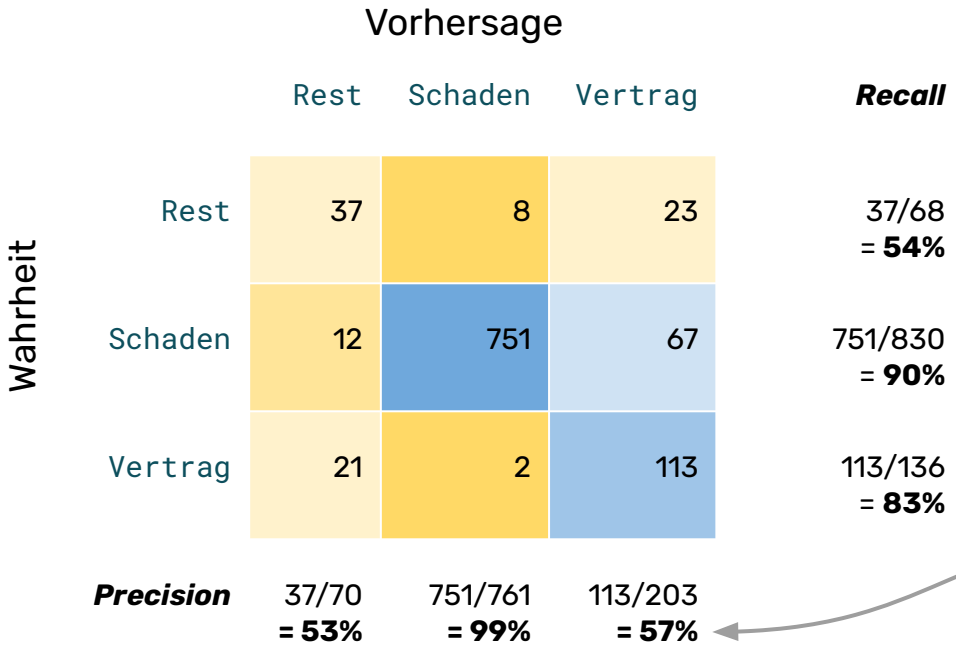
Auswertung Klassifikation

Konfusionsmatrix



Auswertung Klassifikation

Klassen-Metriken

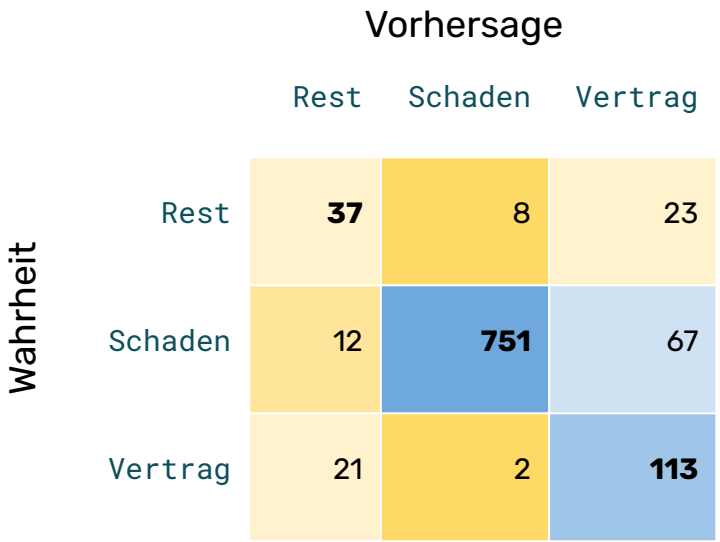


54% der **Schaden**-Dokumente wurden richtig erkannt

57% der **Vertrag**-Vorhersagen stimmten

Auswertung Klassifikation

Key Performance Indicators



Accuracy
"Diagonale / Alles"
87%

Macro F1
harmonisches Mittel von Precision und Recall, Durchschnitt über alle Klassen
72%

Erklärbarkeit der Klassifikation

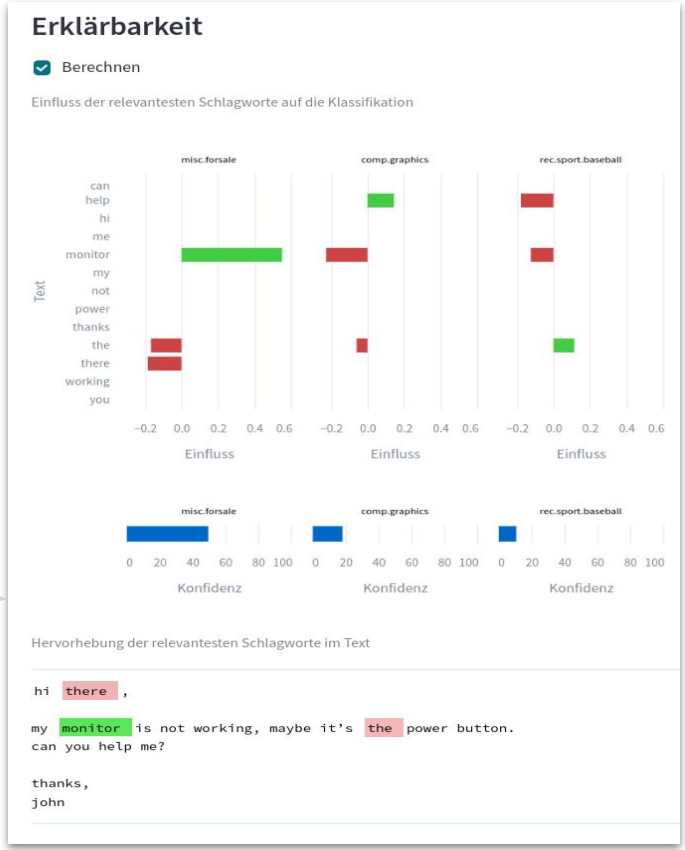
Beispiel-Dokument

trainiertes Modell

eXplainable-AI-Bibliothek

`Local Interpretable Model-agnostic Explanations`

Welcher Textteil spricht für/gegen welche Klasse?



Anwendungsfall Extraktion

SHERLOQ

spezifische Entitäten

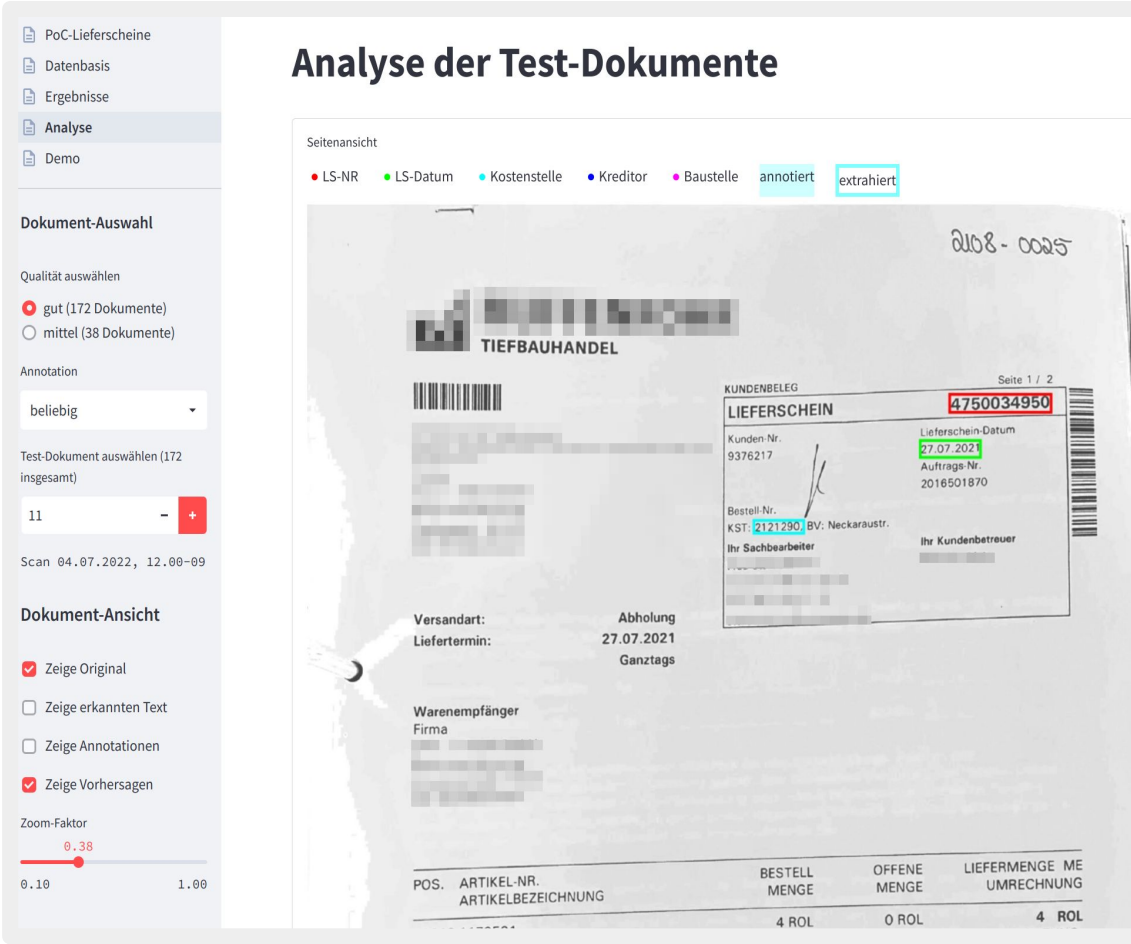
- Lieferdatum, Belegnummer, etc.

Freiform-Dokumente

- Kandidaten-Ranking / custom-NER

Unterstützung beim Labeln

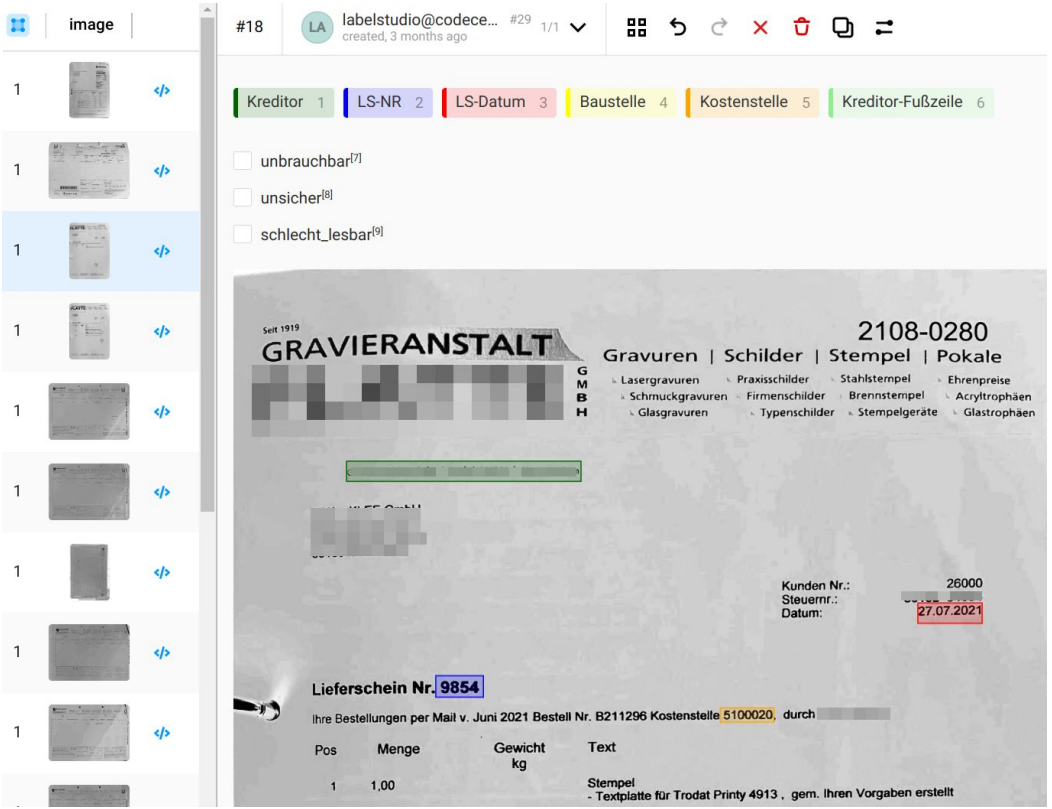
Label Studio



Extraktion Labeling

Labelstudio-Oberfläche mit selbstdefinierten Labels

- vorinstallierten Labeling-Templates
- kollaboratives Arbeiten



Anwendungsfall Extraktion

SHERLOQ

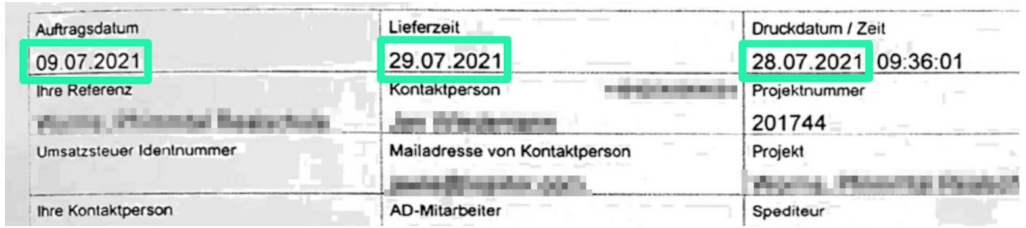
Machine Learning Methoden:

Kandidaten-Ranking

Kandidaten mit **RegEx** extrahieren

Merkmale jedes Kandidaten ermitteln (**Position** + **umliegender Text**)

ein ML Modell anhand der Kandidaten-Features **trainieren**



class	page	x	y	text	text_left	text_above
LS-Datum	1	261	884	09.07.2021		Auftragsdatum
LS-Datum	1	684	886	29.07.2021		Lieferzeit
LS-Datum	1	1137	885	28.07.2021		Druckdatum / Zeit

Anwendungsfall Extraktion

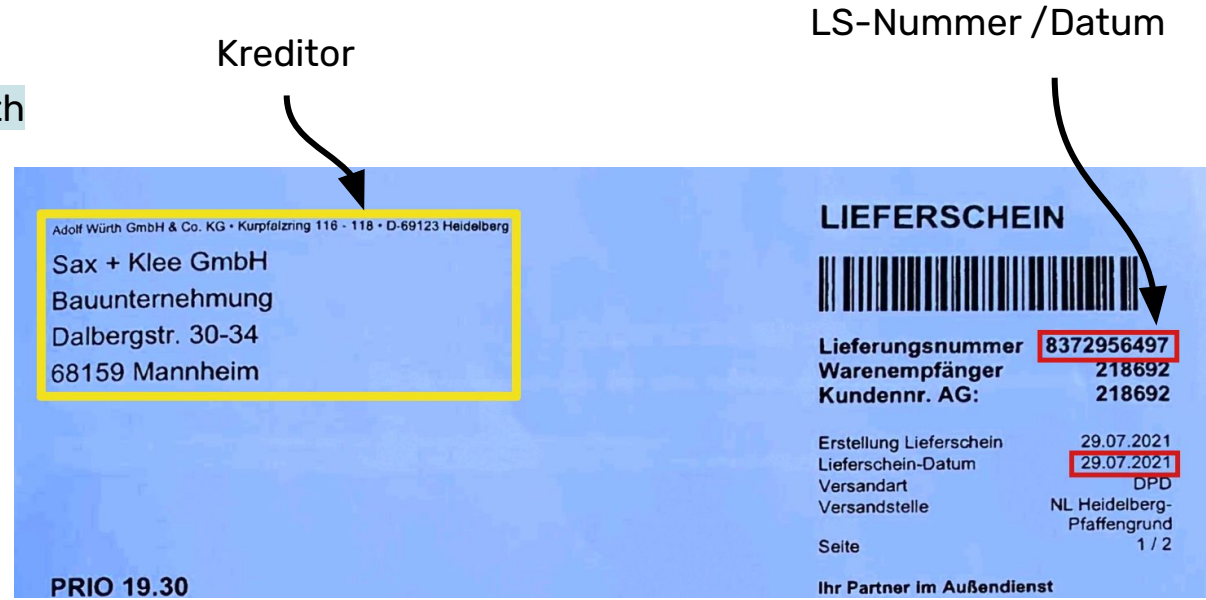
SHERLOQ

Machine Learning Methoden:

Custom Named Entity Recognition with

spaCy

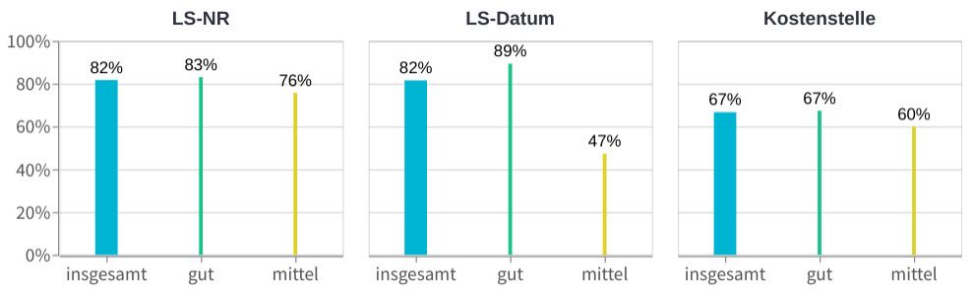
CNN-LSTM Model Architektur



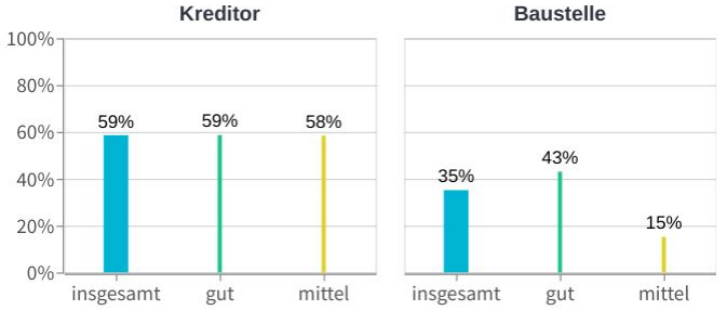
Ergebnisse der Extraktion

SHERLOQ

Kandidaten-Ranking



Custom Named Entity Recognition with ***Spacy***



a solution by codecentric

Von Daten zur KI-Lösung

Herausforderungen und Tools

Daten annotieren

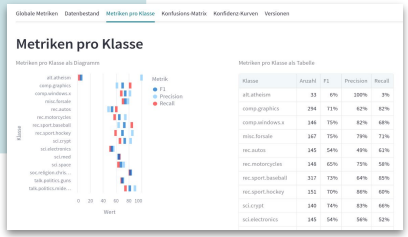
Label Studio

Experimente tracken

mlflow

Ergebnisse analysieren

Streamlit

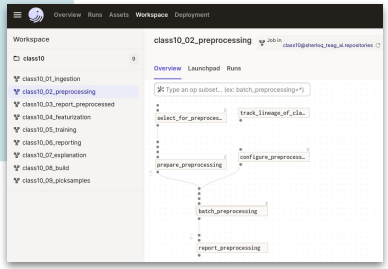


Daten verwalten

MINIO

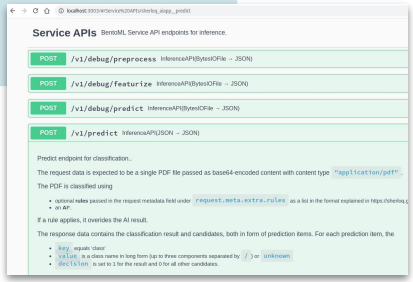
ML-Pipelines orchestrieren

dagster



BENTOML

AI-Service bauen



Daten versionieren

Pipelines konfigurieren

HYDRA

SHERLOQ Engineering

SHERLOQ Bausteine

Training (AI-Lab)
Training der AI Modelle

- bekannter Werkzeugkoffer
- schnelles Deployment
- skalierbar
- ressourcen-intensiv
- PoC Umgebung (On-Demand)
- KI-Checkup (Nach-Training)

Prediction
Betrieb der AI Modelle

- Individuelle AI Modelle/Preprocessing
- wiederverwendbare Integrationen (WebService, PubSub, Kofax)
- Daten Sammeln (für KI-Checkup)
- skalierbar

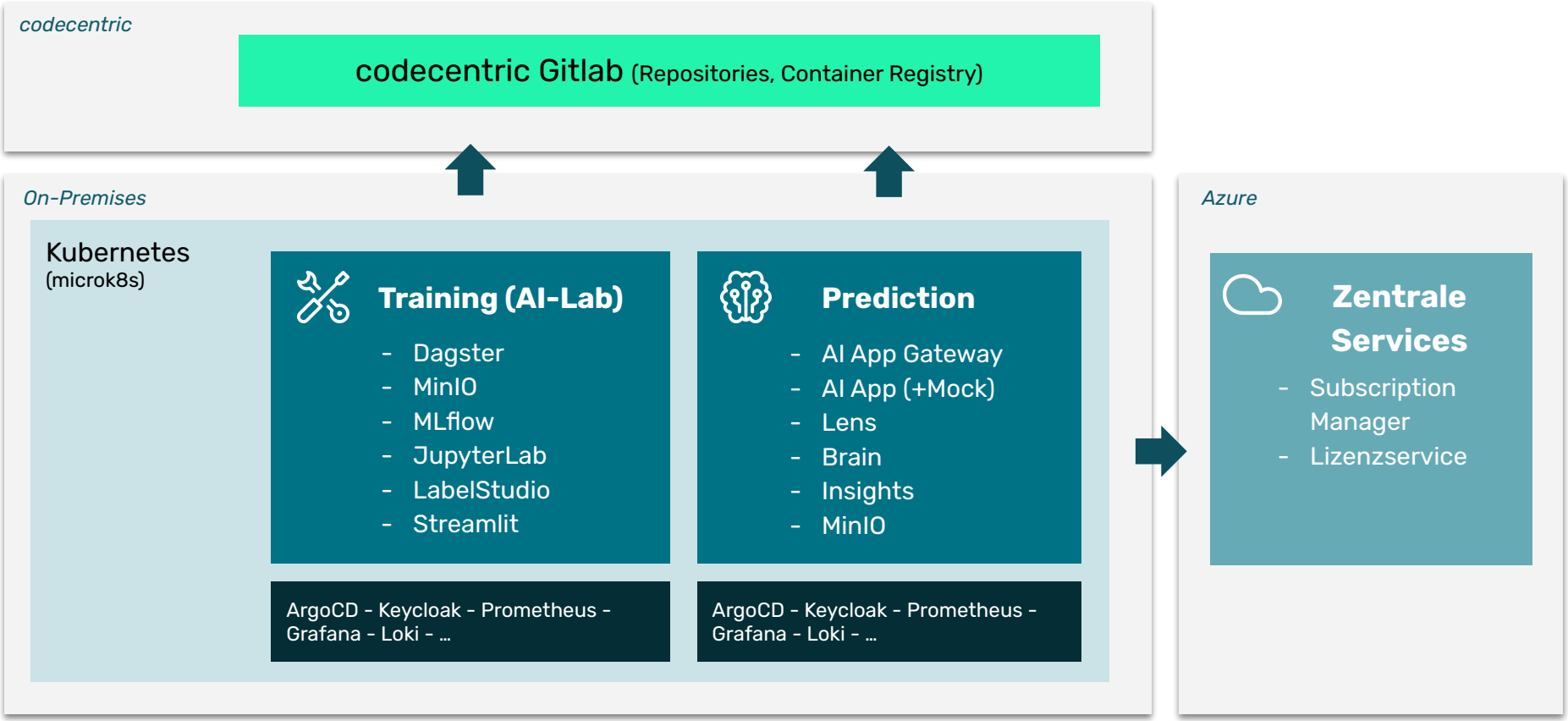
Zentrale Services
Kundenübergreifend

- Sammeln von Nutzungsdaten für die Abrechnung
- Erstellung von monatlichen Reports für die Abrechnung
- Lizenzierung
- …

Rahmenbedingungen

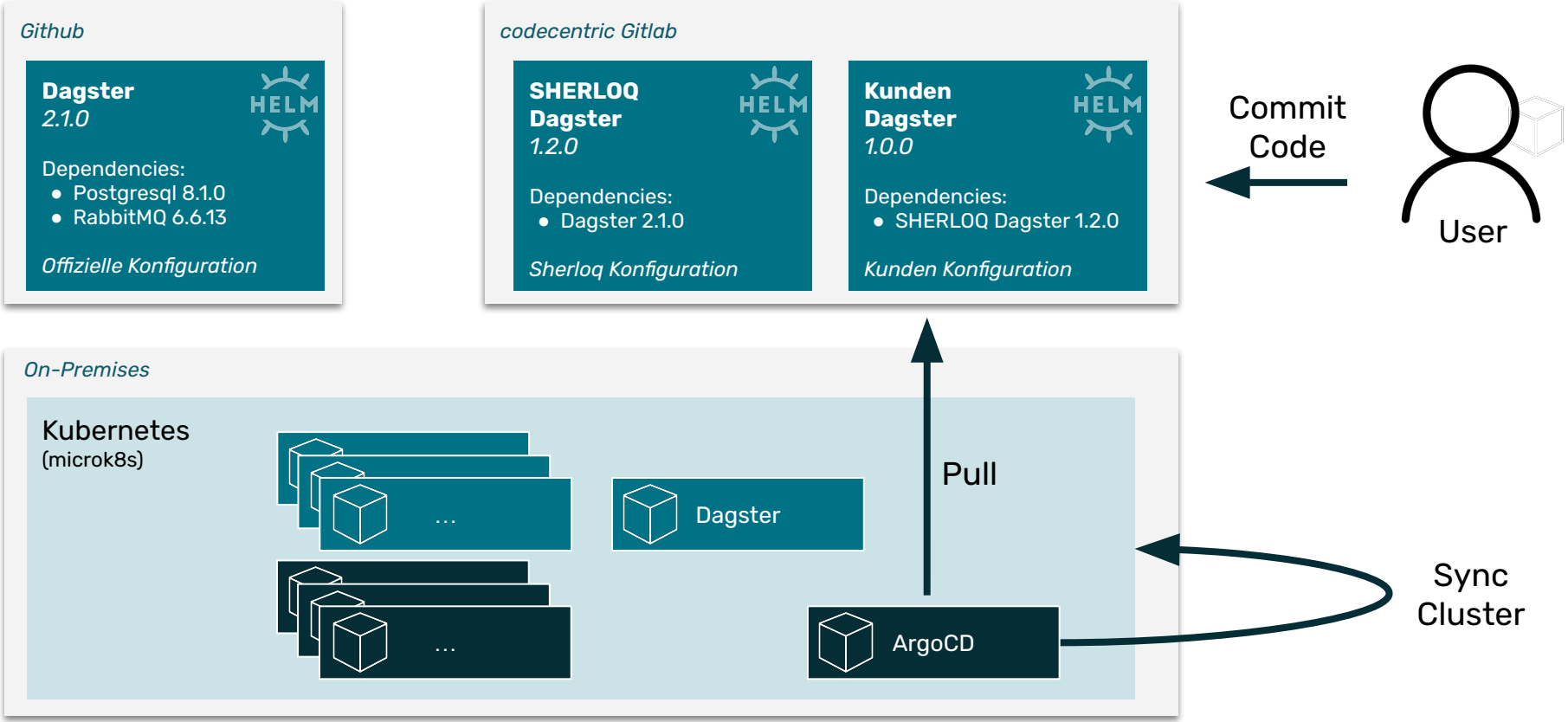
- On-Premises / Hybrid
- Konfigurierbar
- schneller Start für Data Scientists
- Authentifizierung / Autorisierung

SHERLOQ Engineering



SHERLOQ Deployment

Am Beispiel von Dagster



SHERLOCQ

a solution by codecentric

Elvira Siegel
Data Scientist
elvira.siegel@codecentric.de

Thomas Timmermann
Data Scientist
thomas.timmermann@codecentric.de

Johannes Voscort
Developer
johannes.voscort@codecentric.de

codecentric AG
Am Mittelhafen 14
48155 Münster

www.codecentric.de