

G.E.O.S (Project Lithium)

Meeting 15.07.2022

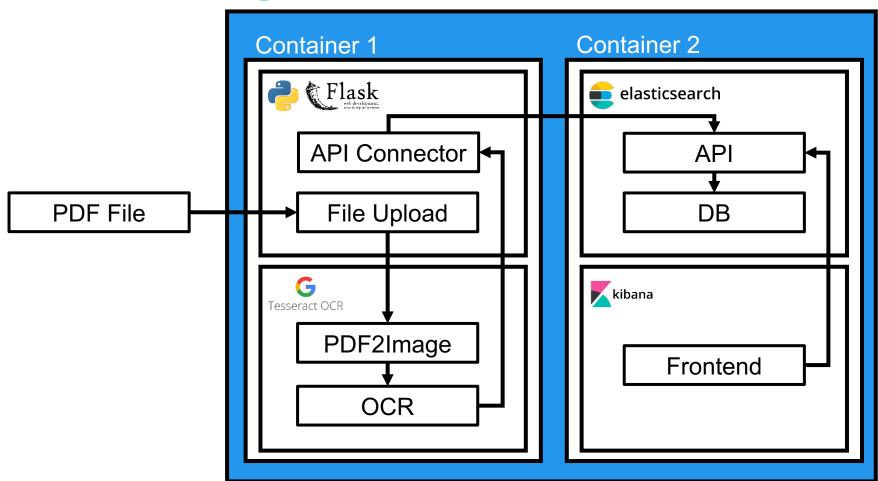
Software Architektur

- // Docker Container Umgebung (2 Container)
- Web App (Python Flask) für Zugriff auf OCR, Datenbank Anfragen
- Elasticsearch als Datenbank für Speicherung, Indexierung

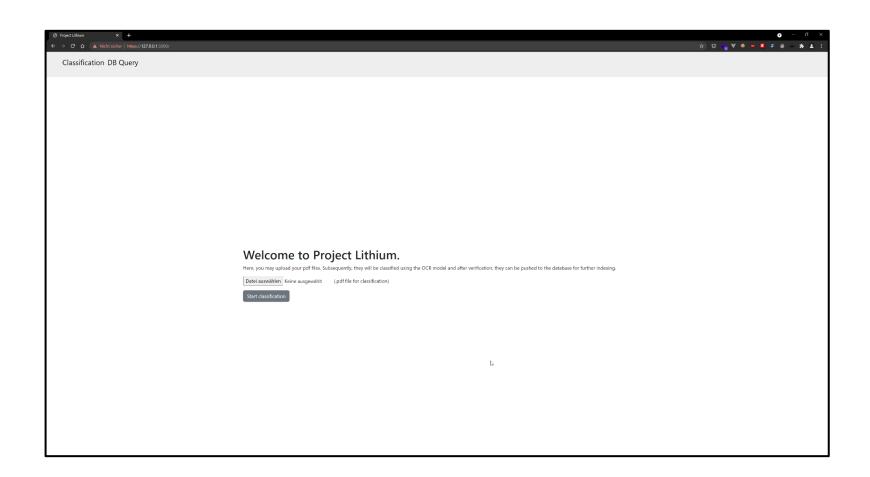


Software Architektur





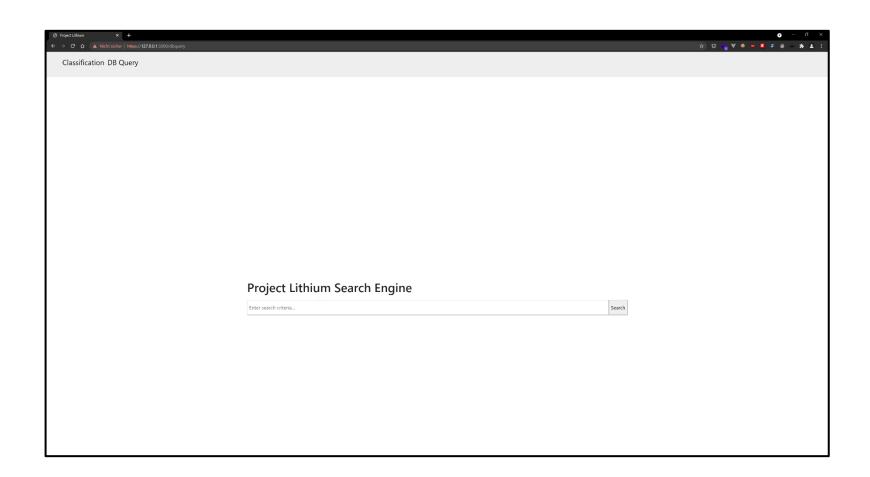




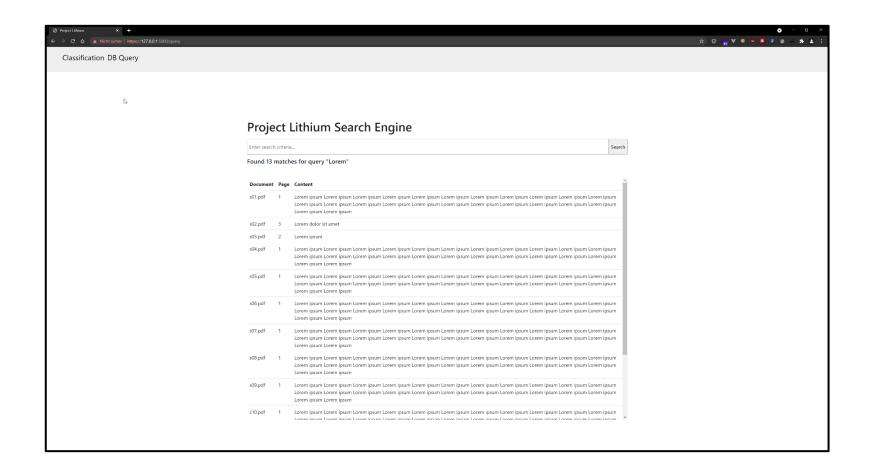






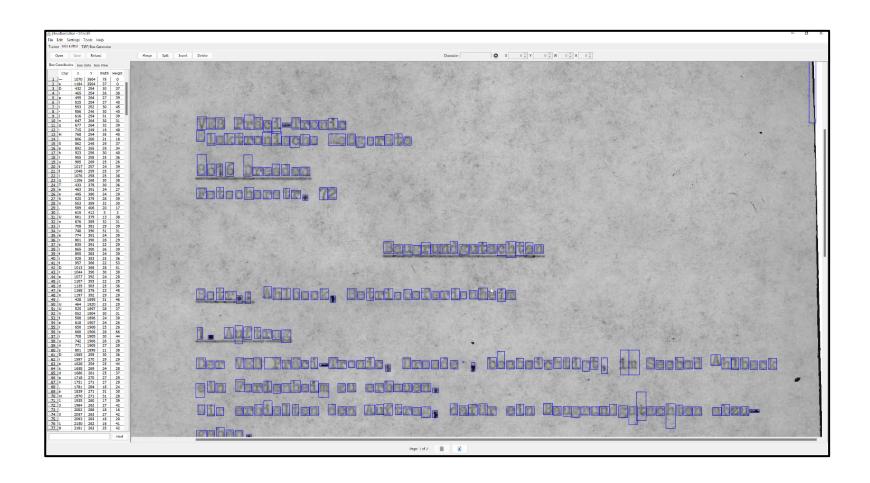








OCR Modell Training





OCR Modell

- Bisher 5 Dokumente (13 Seiten) "trainiert"
- # Erkennungsrate vor Training: 36%
- Erkennungsrate nach Training: 82%
- // Herausforderungen:
 - Schlecht lesbare Inhalte durch niedrigen Kontrast
 - // Handschriftliche Formeln



Datenspeicherung, Indexierung

- Daten werden nach Klassifikation in DB (Elasticsearch) abgelegt
- Filterbar nach Ort, Inhalt, etc.
- Später: Kartenansicht über Web App geplant
- // Herausforderungen:

 - Unterscheidung Ort des Gutachtens, andere Adressen

