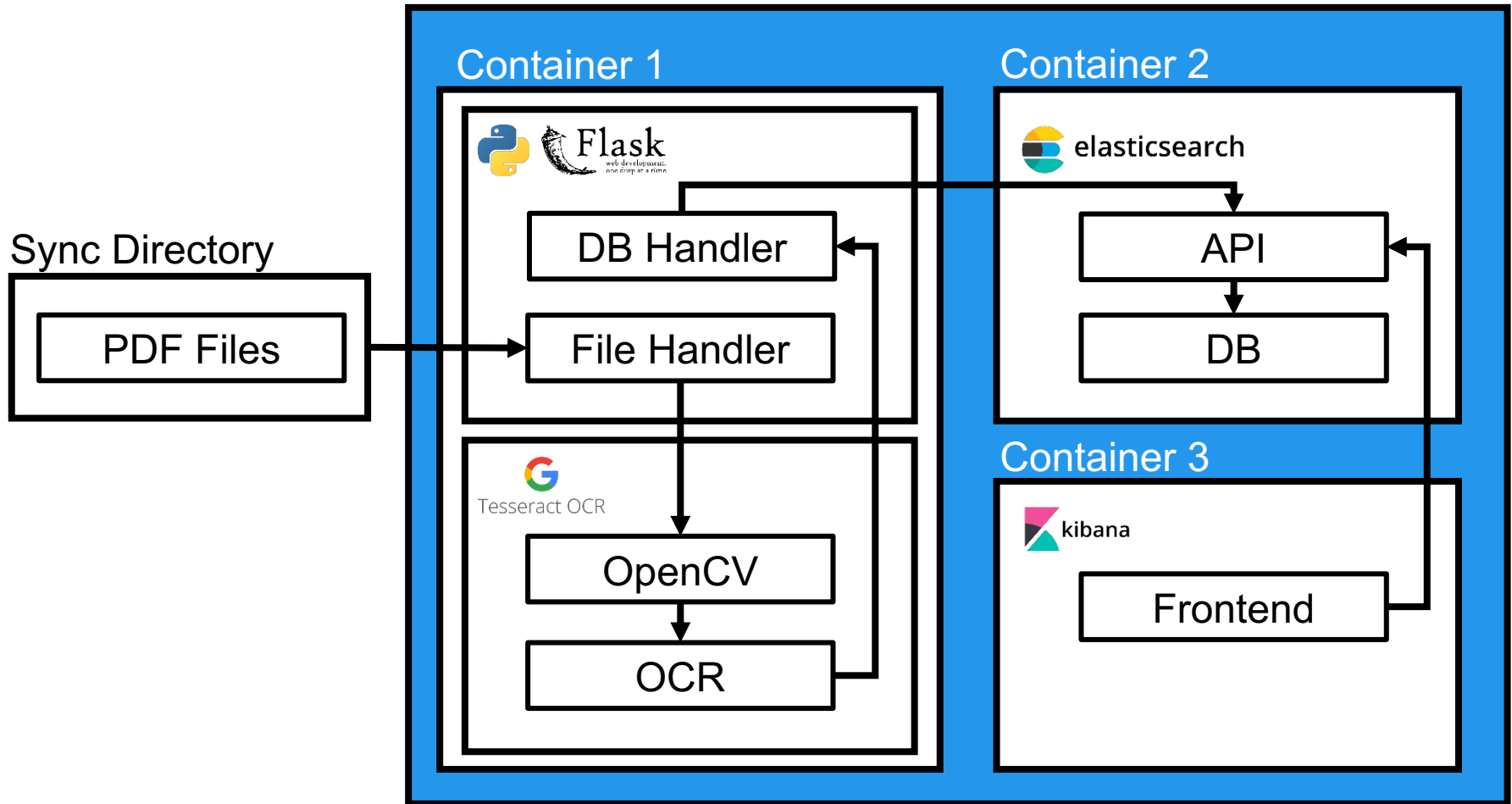


TerraSearch

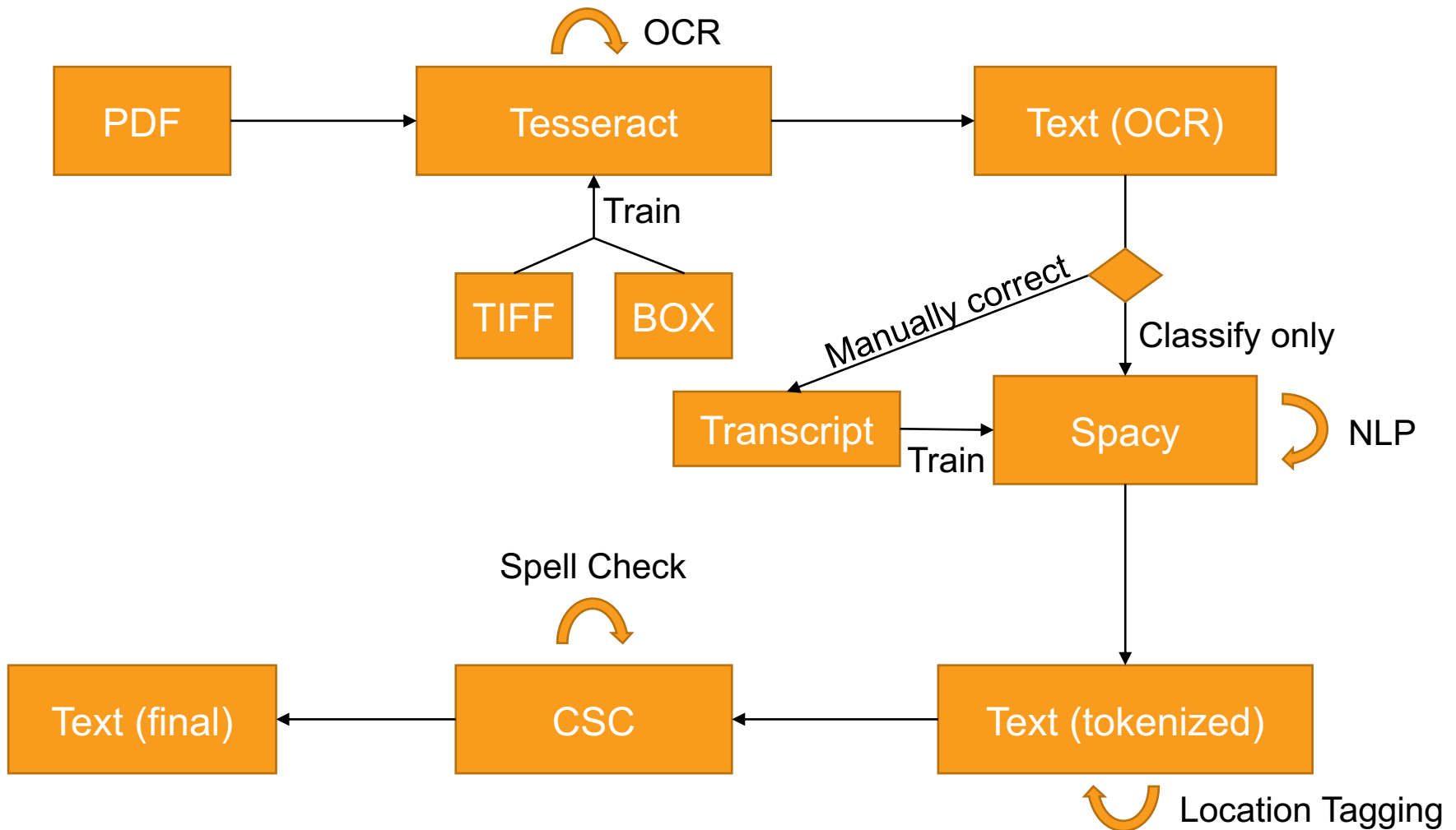
OCR und Suche in digitalisierten Bodenuntersuchungsdokumenten



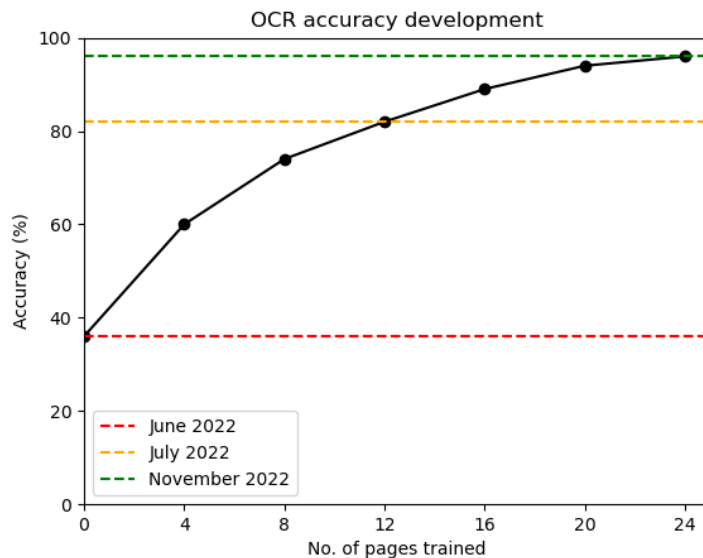
Software Architektur - Container



Software Architektur - Klassifikation



OCR Modell



- // Trainiert mit 8 Dokumenten (24 Seiten)
- // Erkennungsrate vor Training: 36%
- // Erkennungsrate nach Training: 96%
- // Herausforderungen:
 - // Schlecht lesbare Inhalte durch geringen Kontrast
 - // Niedrig-qualitative Scans (knicke, schief, stark gealtert)
 - // Unterschiedliche Schriftarten bzw. handschriftliche Formeln



Datenspeicherung & Analyse

- // Daten werden nach Klassifikation in Datenbank-Container abgelegt
- // Durchsuchbar nach Inhalt, Ort
- // Kartenansicht über Web App
- // Herausforderungen:
 - // Übertragung erkannter Locations in Karte (via Google Maps Geocoding → z.T. noch falsche Ergebnisse)
 - // Unterscheidung Ort des Gutachtens, andere Adressen (manuelle Verifikation notwendig) → „Human-in-the-Loop“



Web App

LIVE DEMO