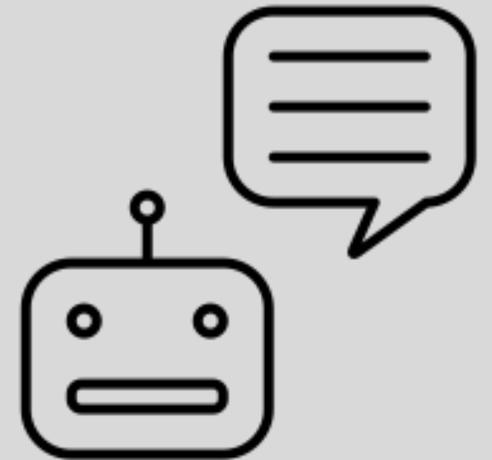


# Chatbot

Felicitas Castrian  
Merlin Sewina  
Paul Offner

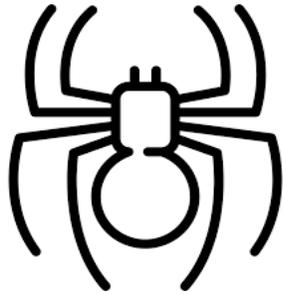


# Chatbot – Aufgabenstellung

Implementation eines Chatbots

- Antwort auf Supportanfragen (lowlevel)
- Verweise auf entsprechende Links

# Chatbot – Komponenten



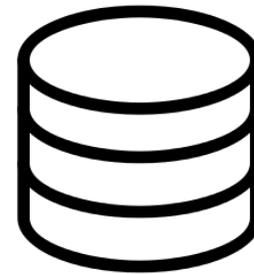
Crawler



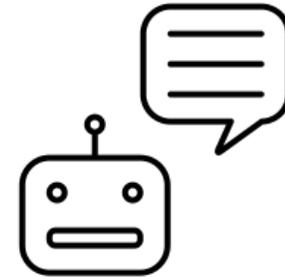
Parser



Analyzer



Datenbank

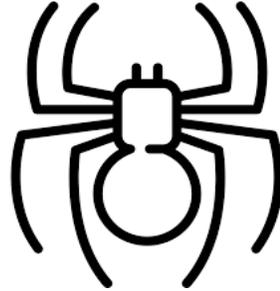


Chatbot

# Crawler



DKRZ-Seite



Crawler

- Speichert alle gefundenen HTML-Links mit Inhalt
- Problem mit PDFs, Fotos, Videos
- Langsam

# HTML-Parser



Parser

- Reduziert HTML-Seite auf reinen Text
- BeautifulSoup

# Analyzer



Analyzer

- Analysiert jeden geparsten HTML-Text auf Wörter (welche und wie viele)
- Probleme mit Rechtschreibfehlern, fehlendem einheitlichen Format

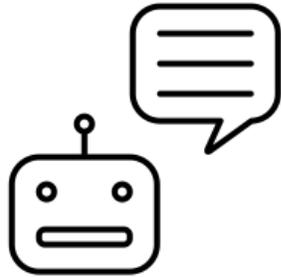
# Datenbank



Datenbank

- Alle Wörter aus Analyser
- Tabelle aller Wörter
- Tabelle aller Links
- Jedes Wort kennt seine Links

# Chatbot

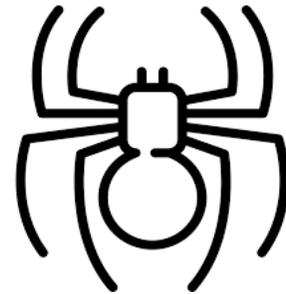


Chatbot

- Eingabe des Nutzers (aktuell über Console)
- Stopwords filtern
- Schlagwortsuche in der Dankenbank
- Ausgabe der Links



DKRZ-Seite



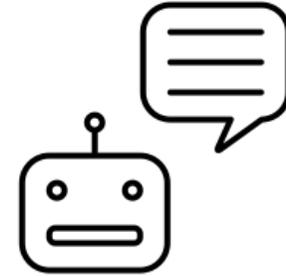
Crawler



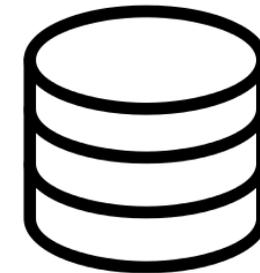
Parser



Analyzer



Chatbot



Datenbank

# Ausblick

- Mehr als nur Stopwords filtern
- Analyse von PDFs
- Nutzerinteraktion über Webschnittstelle