

15.07.2021

**Bachelor/Masterarbeit zu vergeben:
“Data Science at the Frontline:
Wie sich Schlagzeilen vor und nach der Bundestags-
wahl 2021 bei Google ändern”
(Deutsch oder Englisch)**

Was ist das Thema?

Ziel des Projekts ist es, die Vielfältigkeit und Regionalität von Googles Schlagzeilen-Funktionalität zu analysieren. Hierzu werden die Schlagzeilen, die bei der Google-Suche mit vorgegebenen Schlagwörtern (z.B. „Kanzler“, „Kanzlerin“, Parteinamen, „Stuttgart 21“) angezeigt werden, über einen vierwöchigen Zeitraum um die Bundestagswahl 2021 automatisiert erfasst und für spätere Analysen gespeichert. Um regionale Unterschiede in den Schlagzeilen zu erfassen, werden die Daten von einem Server in Karlsruhe sowie einem Server außerhalb von Baden-Württemberg gesammelt. Die erfassten Daten werden anschließend hinsichtlich ihrer Vielfältigkeit und Regionalität ausgewertet und die Ergebnisse in der Abschlussarbeit aufgearbeitet.

Unter andere sollen folgende Forschungsfragen beantwortet werden: Wie vielfältig sind die Ergebnisse der Google-Schlagzeilen insgesamt und für jeden Suchterm? Wie verändert sich die Vielfalt über die Wochen? Welchen Anteil an den Schlagzeilen haben Domains aus Baden-Württemberg (z.B. swr.de)?

Die Abschlussarbeit umfasst sowohl die Implementierung der automatisierten Datenerfassung als auch die Analyse der Daten.

Als eine weltweit führende Forschungsgruppe auf dem Gebiet der Wissensrepräsentation verfügt der Lehrstuhl „Web Science“ über die Expertise und Ausstattung, um bei der Datenerfassung, Analyse und Vermittlung der gewonnenen Erkenntnisse zuverlässig Feedback an den Studierenden bzw. an die Studierende geben zu können. Die Arbeit wird von Dr. Michael Färber betreut. Bei Bedarf können weitere Expertinnen und Experten für die Betreuung herangezogen werden (z.B. Soziologen, ZKM Karlsruhe). Geplant ist, die Ergebnisse der Arbeit am Ende der Bearbeitungszeit auch wissenschaftlich zu veröffentlichen.

Was sollten Sie mitbringen?

- Grundlegende Kenntnisse und Erfahrungen im Bereich Data Science (z.B. etwas Erfahrung mit Python).
- Bereitschaft, mit der Abschlussarbeit ca. zum 15.08.2021 zu starten (die Datengewinnung sollte 2 Wochen vor der Bundestagswahl 2021 starten).

Keywords: search engines, information retrieval, data science, web science, data mining, election.

Kontakt:
Dr.-Ing. Michael Färber | Tarek Saier
{michael.farber,tarek.saier}@kit.edu