

18.08.2022

## Bachelor-/Masterarbeit zu vergeben: „Chronik 2050: Automatische Extraktion von erwarteten Ereignissen aus Webseiten“

### Herausforderung

In Wirtschaft, Politik und Gesellschaft ist es essenziell, Trends und erwartete Ereignisse frühzeitig (inklusive Eintrittswahrscheinlichkeit) erkennen zu können und ihre Auswirkungen richtig abschätzen zu können. Prominente Beispiele für die Einführung von Technologien sind etwa:

- „Das Telefon hat zu viele Schwächen, als dass man es ernsthaft für die Kommunikation in Erwägung ziehen kann“ (Western Union in einem internen Memo, 1876)
- "Es gibt keinen Grund, warum jeder einen Computer zu Hause haben sollte" (Ken Olsen, Gründer von Digital Equipment Corp, 1977)
- „Internet wird kein Massenmedium“ (Matthias Horx, Zukunftsforscher, 2001)

### Ziel der Arbeit

Ziel dieser Arbeit ist es, Prognosen über die Zukunft besser verstehbar, interpretierbar und diskutierbar zu machen. Hierfür sollen mit Hilfe von aktuellen Methoden des maschinellen Lernens und der Verarbeitung natürlicher Sprache Wissen über zukünftige Events und generelle Prognosen aus Webseiten identifiziert, thematisch und zeitlich kategorisiert, verknüpft und gespeichert werden. Auf diese Weise soll eine Datenbank mit zukünftigen Events aufgebaut werden, die als Grundlage für eine Chronik bzw. einen Kalender dienen soll, der die nächsten Jahrzehnte abdeckt. Die Chronik soll in Form einer Webseite verfügbar gemacht werden und als Grundlage für Online-Diskussionen dienen.

Im Zentrum der Arbeit steht die Extraktion von Events aus Textdokumenten. Während bisherige Ansätze vor allem aktuelle Events und Events aus vergangenen Jahrzehnten zum Ziel hatten, bringt die Extraktion von Events, die für die Zukunft im Text angedeutet sind, zusätzliche Herausforderungen, wie etwa die Einordnung der Plausibilität (Eintrittswahrscheinlichkeit) und die Gruppierung gleicher oder ähnlicher Events.

### Was sollten Sie mitbringen?

- Grundlegende Programmierkenntnisse.
- Interesse an Zukunftsthemen, Trendanalyse und Future Research.

### Wie wird die Arbeit umgesetzt und betreut?

Die Betreuung der Arbeit erfolgt über den Lehrstuhl Web Science am KIT-Institut AIFB in Kooperation mit dem Innovation Hub der Atruvia AG, dem IT-Dienstleister der genossenschaftlichen Finanzgruppe. Eine finanzielle Vergütung über die Atruvia AG ist vorgesehen.

Kontaktperson:  
**Dr. Michael Färber**  
michael.farber@kit.edu