

Studentische / Wissenschaftliche Hilfskraft (8h / Woche)

Stellenausschreibung SHK / WHF

Zusammen mit Wissenschaftler:Innen des Institute for Artificial Intelligence an der Universität Bremen arbeiten wir an der Extraktion und Repräsentation von Wissen, welches Roboter für Manipulationsaufgaben wie z.B. das Schneiden von Früchten [1] benötigen. Unser Ziel ist es, das extrahierte Wissen in einem Wissensgraph zu repräsentieren und *actionable* zu machen. Im ersten Schritt unserer Arbeit liegt der Fokus auf der Extraktion von relevantem Wissen aus Internet-Quellen wie z.B. WikiHow.

Für die Arbeit mit WikiHow wurde bereits ein Tool in Java entwickelt [2]. Für zukünftige Erweiterungen und komplexere Analyseschritte soll dieses Tool im Rahmen dieser Stelle in Python übersetzt werden. Zusätzlich sollen Methodiken des statistischen Natural Language Processings (z.B. Coreference Resolution [3]) auf die Daten angewendet werden, um die Extraktion zu verbessern. Auch die Verknüpfung zu Ressourcen des Semantic Web (z.B. FrameNet [4]) ist angedacht.

Die Stelle richtet sich an Studierende der Informatik oder einem vergleichbaren Studiengang und ist auf **8 Stunden pro Woche** ausgelegt. Die Stelle startet zum **nächstmöglichen Zeitpunkt**. Es besteht die Möglichkeit eine Abschlussarbeit aufbauend auf dieser Tätigkeit zu schreiben.

Ihr Profil:

- Programmiererfahrung in Python
- *(optional)* Programmiererfahrung in Java
- *(optional)* Erfahrung mit (statistischem) NLP und/oder Ressourcen des Semantic Web
- Interesse an den Themen des Projekts (Robotik, Wissensextraktion, NLP) & wissenschaftlichem Arbeiten

Interessiert? Dann schicken Sie Ihre Bewerbungsunterlagen (kurzes Anschreiben und Lebenslauf) per Mail an: jtöberg@techfak.uni-bielefeld.de

Related Work

[1] <https://food-ninja.github.io/FoodCutting/>

[2] <https://github.com/Food-Ninja/WikiHow-Instruction-Extraction>

[3] R. Sukthankar, S. Poria, E. Cambria, and R. Thirunavukarasu, 'Anaphora and coreference resolution: A review', *Information Fusion*, vol. 59, pp. 139–162, Jul. 2020, doi: 10.1016/j.inffus.2020.01.010.

[4] J. Ruppenhofer, M. Ellsworth, M. Schwarzer-Petruck, C. R. Johnson, and J. Scheffczyk, 'FrameNet II: Extended Theory and Practice', International Computer Science Institute, 2010, Online: <https://framenet.icsi.berkeley.edu/>.

Die Arbeitsgruppe Semantische Datenbanken untersucht und entwickelt Methoden, die es Maschinen ermöglichen, relevante Kenntnisse sowie sprachliche Fähigkeiten zu erwerben. Unter Verwendung von Methoden des *Natural Language Understanding* und des *maschinellen Lernens* streben wir nach Maschinen, die in der Lage sind, Wissen durch das Lesen unstrukturierter Texte zu erlangen. Dafür konzentriert sich die Gruppe insbesondere auf Methoden der Informationsextraktion, semantischen Analyse, Ontologie-Lernens, Sentiment-Analyse, Entity Linkings sowie Fragebeantwortung.

Mehr Informationen finden Sie auf unserer Website: <http://www.sc.cit-ec.uni-bielefeld.de/>