

Stellenausschreibung

Studentische Hilfskraft

Thema: Semi-automatische Übersetzung und Mapping von deutsch- und englischsprachigen Fragebögen im Rahmen der LIFE-Studie

Das Leipziger Forschungszentrum für Zivilisationserkrankungen (LIFE) erforscht Erkrankungen, wie Depression, Diabetes, Allergien oder Herz-Kreislauf-Erkrankungen. Im Rahmen der Studie wurden zahlreiche Daten anhand von Fragebögen zur Gesundheit und zu den Lebensumständen der Leipziger Bevölkerung erhoben. Die verwendeten Fragebögen sollen nun mit biomedizinischen Ontologien annotiert und semantisch angereichert werden. Allerdings sind biomedizinische Ontologien zumeist nur auf Englisch definiert, während die in LIFE benutzten Fragebögen auf Deutsch verfasst wurden, so dass eine qualitativ hochwertige Zuordnung erschwert wird. In diesem Kontext suchen wir eine studentische Hilfskraft zur Mitarbeit an den folgenden Aufgaben:

- Konzeptionelle Planung des Datenmodells und Suche nach existierenden automatischen Übersetzungs- und Alignment-Tools
- Extraktion englischer Fragen aus Dokumenten verschiedener Formate (u.a. pdf, html) und Überführung in das definierte Datenmodell
- Erstellung eines korrekten Mappings zwischen deutschen und englischen Fragen
- Implementierung eines Tools zur automatischen Übersetzung von deutschen Fragebögen ins Englische in der Domäne
- Evaluierung eigener und existierender Alignment- und Übersetzungs-Tools bzgl. der Datenqualität

Interessenten melden sich bitte bei Dr. Anika Groß (gross@informatik.uni-leipzig.de) oder Dr. Ying-Chi Lin (lin@informatik.uni-leipzig.de).