

19.04.2015

Bachelor-/Masterarbeit

„*Scientification*: Automatisches Hinzufügen von Referenzen in Text“ zu vergeben!

Um was geht es?

Viele Autoren von Fachliteratur stehen vor der Herausforderung, die Angaben im Text mit Referenzen zu versehen, um auf Referenzliteratur (etwa auf die erstmalige Einführung von Konzepten) hinzuweisen. Durch die stark zunehmende Anzahl an Veröffentlichungen und das Aufkommen immer neuer Forschungsfelder ist es für Autoren immer schwieriger bzw. zeitaufwendiger, sich zunächst einen Überblick über die Referenzliteratur zu verschaffen, um anschließend die besten Referenzen für den eigenen verfassten Text auszuwählen und zu zitieren.

Im Zentrum der Abschlussarbeit steht daher die Konzeption und Implementierung eines Tools zum automatischen Hinzufügen (Annotieren) von Referenzen zu einem gegebenen Text (siehe Abbildung rechts).

Hierzu wird zunächst eine Menge von Veröffentlichungen analysiert und die Referenzen sowie die Kontextwörter (z.B. jeweils die zwei davorstehenden Wörter) extrahiert und in einer geeigneten Datenstruktur gespeichert. Soll ein Text mit Referenzen versehen werden, wird geprüft, welche (zuvor abgespeicherten) Kontextwörter im Text vorkommen und auf welche Referenzen geschlossen werden kann, so dass diese zum Text hinzugefügt werden können.

... by researchers in the field of information retrieval (IR)In the popular tf-idf scheme, a basic vocabulary of words is chosen.



... by researchers in the field of information retrieval (IR) (**Baeza-Yates and Ribeiro-Neto, 1999**)....
In the popular tf-idf scheme (**Salton and McGill, 1983**), a basic vocabulary of words is chosen.

Der Korpus an Veröffentlichungen zur Analyse wird bereitgestellt.

Was sollten Sie mitbringen?

- Interesse an der Anwendung und Konzeptionierung von *Text Mining*-Verfahren.
- Gute Programmierkenntnisse in Java.

Kontaktperson:
Michael Färber
michael.farber@kit.edu
Tel.: 0721/608 479 46