Language changes!
Want to use computer science to find out how?

Languages are evolving over time triggered by various factors like new technologies or high-impact events. Knowing about the evolution of a term can significantly decrease time needed for searching for information. Imagine making a search for St. Petersburg and not knowing that from 1914-1924 it was named Petrograd and then Leningrad. First in 1991 it changed back to Saint Petersburg. The typical user will not be aware of this complex development and thus doesn’t know why there are no documents returned for an 80 year period mentioning the city name.

M.Sc/Diploma thesis: Clustering for word senses detection

L3S Research Center, under supervision of Prof. Dr. Nejdl offers a M.Sc. or Diploma thesis in the scope of the LiWA project (http://www.liwa-project.eu/). The aim of the project is to develop methods for tracking evolution in terminology and using it for search. More specifically the thesis will be to implement word sense discrimination algorithms to model words and their senses. This task is closely related to clustering techniques and linguistic processing. Hierarchical clustering can help improve the quality of the word senses.

Are you interested in working with data mining and search technologies? Want to be part of a research team working with a new exciting research topic and get a good start for a future career?

The tasks:

I. Find and implement a word sense discrimination technique.

II. Apply hierarchical clustering to improve quality of current word sense discrimination techniques used.

III. Evaluate results using WordNet or user studies and compare to existing results.

You should have:
- Interest in semantic technologies and clustering techniques
- Some algorithmic and analytic skills
- Programming skills in Java

Are you interested or have questions? Contact us:

Nina Tahmasebi, Email: tahmasebi@L3S.de or Thomas Risse, Email: risse@L3S.de

Forschungszentrum L3S, Appelstr. 9a, 30167 Hannover