



Opportunity for a Bachelor Project

Problem

Today's plagiarism detection systems are only able to identify copy&paste plagiarism, but not intelligently disguised forms of plagiarism such as translated, paraphrased, and idea plagiarism.

This means true plagiarisms often remain undiscovered for years – remember Karl-Theodor zu Guttenberg's Thesis.

SciPlore's Research

To alleviate this shortcoming, the research team of SciPlore introduced a new approach: Citation-based Plagiarism Detection (CbPD).

In contrast to the existing approaches, CbPD does not only examine textual similarity, that is the words occurring in a document, but makes use of citation patterns in scientific documents as a unique language independent fingerprint for similarity computation.

Evaluations of this new approach showed that citation patterns often show similarity even if texts have been translated or strongly paraphrased. We were able to prove this in analyzing the plagiarized doctoral thesis of Karl-Theodor zu Guttenberg. Our approach thus allows for the detection of strongly disguised academic plagiarism that otherwise would not be machine detectable using traditional approaches.

You

Collaborate with other student researchers and the SciPlore Team to complete your Bachelor project. You would learn about our CbPD algorithms and develop a web-based tool for the visualization and analysis of citations information to help detect plagiarism.

We are looking for the following qualifications:

- intermediate to advanced programming skills using JAVA and php
- additional programming experience is preferred but not required (C/++, Python and/or MySQL, Hibernate, Jersey, REST Web Services) are a plus,
- a passion for usability and the design of user interfaces is desired

If you have any questions, or want more information, just contact the project supervisor:

Dipl. Wirt. Inf. Bela Gipp, MBA (gipp@berkeley.edu)

Or visit: <http://sciplore.org/projects/citation-based-plagiarism-detection/>