

## Bachelor- und Masterprojekt: Entwurf und Entwicklung eines Dashboards zum Monitoring der Urkund Software

**Semester:** Sommersemester 2020

**Sprache:** Deutsch/English

### Motivation:

Bedingt durch die enorme Zunahme und Verfügbarkeit an wissenschaftlicher Literatur aber auch durch einen weniger sorgsamem Umgang mit Zitaten im Rahmen der Erstellung von wissenschaftlichen Arbeiten gehen Universitäten in jüngerer Zeit vermehrt dazu über sogenannte Plagiaterkennungssoftware (kurz: PES) zur Aufdeckung von Plagiaten unterstützend einzusetzen. Dabei werden die zur Begutachtung anstehenden Dokumente auf einer rein syntaktischen Ebene einem Vergleich mit bestehenden Veröffentlichungen unterzogen und ähnliche oder übereinstimmende Formulierungen dem Gutachter präsentiert, so dass dieser Anhaltspunkte für mögliche Plagiate erhält. Zur Einschätzung der Belastbarkeit der präsentierten Ergebnisse ist es für Gutachter unerlässlich einerseits über die berücksichtigten Quellen, andererseits aber auch über den funktionalen Status der Software im Bilde zu sein, um die präsentierten Ergebnisse der Software einzuordnen.

### Beschreibung:

Für die Erkennung von möglichen Plagiaten wird an der Universität Duisburg-Essen seit einiger Zeit die PES URKUND des schwedischen Herstellers Prio Infocenter AB eingesetzt. Die Software wird als Webanwendung im Sinne eines Software-as-a-Service angeboten und nimmt die zu prüfenden Dokumente per Upload, E-Mail oder auch über APIs entgegen um dann in Form eines Prüfberichts potentielle Plagiate mitsamt den Originalquellen und Textstellen auszuweisen. Der für die Überprüfung als Referenz genutzt Textkorpus besteht dabei nach Angaben des Herstellers u.a. aus verschiedensten Fachzeitschriften, Onlinequellen aber auch Code-Repositories was bspw. die Überprüfung von Plagiaten bei Abgaben zu Programmierübungen erlaubt. Der Einsatz und die Nutzung der Software haben in den vergangenen Monat aber immer wieder gezeigt, dass hinsichtlich der Nachvollziehbarkeit nicht immer für die Endbenutzer klar war, welche Quellen berücksichtigt wurden zumal einige Quellen zeitweise keine Berücksichtigung fanden, so dass Nutzer die Zuverlässigkeit der Software in Frage stellen. Um den Herausforderungen im Bereich Zuverlässigkeit und Nachvollziehbarkeit entgegenzuwirken, scheint es angemessen ein webbasiertes Werkzeug zu entwickeln, welches in Form eines Dashboards Auskunft über den Status einzelner Funktionalitäten der Software erkennen lässt (bspw. ob jüngste Veröffentlichungen in Journalen abgeglichen werden) und auch die Verwaltung und explorative Navigation von Quellen (bspw. ab welchem Jahrgang Fachzeitschriften berücksichtigt werden) ermöglicht. Mit dem vorliegenden Projekt soll ein solches Dashboard

**Institut für Informatik  
und Wirtschafts-infor-  
matik (ICB)**

**Lehrstuhl für Wirt-  
schaftsinformatik und  
Unternehmensmodel-  
lierung**

**Mario Nolte**  
Mario.Nolte@uni-due.de

R09 R04 H00  
Universitätsstraße 9  
45127 Essen

**[www.umo.wiwi.uni-due.de](http://www.umo.wiwi.uni-due.de)**

prototypisch entwickelt werden, wozu u.a. folgende Aufgaben zu berücksichtigen sind:

1. Klärung des Plagiatsbegriffes um differenziert die Unterstützungsmöglichkeiten beurteilen zu können.
2. Interviews mit Nutzern des PES um die Anforderungen an das Dashboard insb. hinsichtlich der oben benannten Themengebiete Nachvollziehbarkeit und Zuverlässigkeit zu erheben.
3. Konzeptualisierung des Dashboard Systems bspw. in Form von UML Use Cases, Klassendiagrammen und / oder Mock-ups und ggf. Rückabstimmung mit den Nutzern des PES.
4. Prototypische Entwicklung des Dashboard Systems, wobei je nach Anzahl der Projektteilnehmer und den durch URKUND zur Verfügung gestellten Schnittstellen die Funktionalität zu skalieren ist.

Neben dem dokumentierten Quellcode und einem Projektbericht, der organisatorische Details und Lessons Learned aus dem Projekt dokumentiert, wird von Ihnen ebenfalls die Ausarbeitung eines Konzeptpapiers erwartet, welches die vorgenommenen inhaltlichen Arbeiten des Projektes in strukturierter Weise zusammenfasst. Neben der Klärung grundlegender Begriffe, enthält es wichtige Zwischenergebnisse des Projektes (bspw. Synthese der Anforderungen aus den Interviews, konzeptuelle Modelle) sowie eine Anleitung zur Inbetriebnahme der Software

#### **Einstiegsliteratur:**

- Foltýnek, T., Dlabolová, D., Anohina-Naumeca, A., Razi, S., Kravjar, J., Kamzola, L., ... & Weber-Wulff, D. (2020). Testing of support tools for plagiarism detection. arXiv preprint arXiv:2002.04279
- Introna, L. D., & Hayes, N. (2011). On sociomaterial imbrications: What plagiarism detection systems reveal and why it matters. *Information and Organization*, 21(2), 107-122.
- Strathmann (2020) Mit Python Daten von beliebigen Websites auslesen; <https://www.heise.de/ratgeber/Scraping-Mit-Python-Daten-von-beliebigen-Websites-auslesen-4659822.html> (Zugriff auf Nachfrage beim Lehrstuhl)
- Weber, S. (2007). Das Google-Copy-Paste-Syndrom. Wie Netzplagiate Ausbildung und Wissen gefährden. Telepolis.

#### **Gruppengröße:** 1-2

**Bewerbung:** Bitte bewerben Sie sich mit einem kurzen Motivationsschreiben (etwa 1 DIN A4-Seite) und einer aktuellen Leistungsübersicht per Email an die Betreuer. Sie können sich individuell und in Form einer Gruppe bewerben. Wenn Sie sich als Gruppe bewerben, senden Sie bitte für jede/n Bewerber/in ein individuelles Motivationsschreiben und eine Leistungsübersicht und nennen Sie alle vorgeschlagenen Gruppenmitglieder. Bitte lassen Sie uns zur besseren Koordination wissen, falls Sie sich zeitgleich auf mehrere Projekte bewerben. Dieses Projekt kann auf Deutsch und Englisch bearbeitet werden.

Bewerbungsfrist: 14. Mai 2020