

Themenvorschlag für eine Masterarbeit in der Web Science Group
am Lehrstuhl Informationssysteme (Prof Naumann)

Zeitung der Zukunft - Aggregating News on Company Portfolios

Dieser Themenvorschlag entstand in einem attraktiven Forschungsumfeld am HPI (<http://hpi.de/naumann/>)
im Rahmen eines Drittmittelprojekts zur Exploration großer Datensammlungen.

Problem

- Credit Risk Manager und Kundenbetreuer in einer Bank müssen stets auf dem Laufenden sein
- Täglich unzählige Nachrichtenquellen nach hunderten Kunden durchsuchen ist kaum möglich
- Zu einer Neuigkeit gibt es etliche Berichte
- Abweichungen in den Berichten findet man nur durch aufmerksames Lesen

Lösungsansatz

- System zum kontinuierlichen Monitoring mehrerer Informationsquellen
- Berichte über Organisationen im Portfolio erkennen und richtig zuordnen
- Berichte zu Events/Fakten zusammenfassen
- Änderungen (falls vorhanden) erkennen
- Darstellung der Events im zeitlichen Verlauf, ggf. „Dauerthemen“ erkennen oder anderweitig bewerten

Über das Projekt:

Nach Enthüllungen wie den Panama-Papers stehen Journalisten vor der Aufgabe große firmeninterne Datensammlungen zu durchforsten und Zusammenhänge und Fakten herauszuarbeiten. Auch Spezialisten bei Audits stehen vor dem ähnlichen Problem, dass ein Gesamtüberblick ohne langwierige Einarbeitung unmöglich ist. In diesem Forschungsprojekt werden Ansätze entwickelt, um die Kerninformationen solcher Sammlungen automatisiert zu strukturieren und visualisieren.

Das Projekt schneidet die Forschungsfelder von **Textmining**, **Dokumentenklassifikation**, **Named Entity Extraction** und **Linking, Relationship Extraction**, sowie die **Analyse von (sozialen) Netzwerk Graphen** und **Data Visualisation**.

Kontakt:

Ich freue mich auch über Anregungen und komplett eigene Ideen,
die mit unstrukturierten Daten aus Text und Web zu tun haben.
Buzzwords: Ingestion, Extraction, Clustering, Visualisation.



Klingt interessant? Dann schreib' doch einfach eine E-Mail: tim.repke@hpi.de
Bei Fragen stehen wir vorab gern zur Verfügung.

Das Hasso-Plattner-Institut (www.hpi.de) ist eine in Deutschland einzigartige Forschungs- und Lehrereinrichtung, die weltweit beachtete Forschung betreibt. Als Digital Engineering Fakultät der Universität Potsdam bietet das HPI innovative Studiengänge im Bereich des Digital Engineering an und betreut Promotionsprojekte. Wir verstehen uns als eine „Education Company“ und sind ob unserer Besonderheit viel beachtetes Pilotprojekt, das im Fokus von Politik und Öffentlichkeit steht.