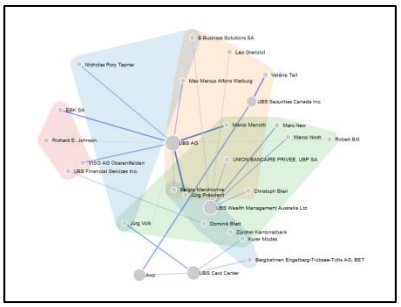
Relation Visualization

Entwicklung einer Visualisierung von extrahierten Relationen zwischen Enitäten aus Texten

Student	Norman Süsstrunk
Referent	Martin Studer
Korreferent	Albert Weichselbraun
Master Research Unit	Software and Systems, HTW Chur



Darstellung von Relationen

Diese Projektarbeit entstand im Auftrag des Forschungsprojektes WISDOM (Web Intelligence for Improved Decision Making) an der HTW Chur. Wisdom kombiniert Konzepte aus Business Intelligence und Web Intelligence Systemen. Es benutzt dazu neben firmeninternen Daten auch frei verfügbare Daten aus dem Internet und im speziellen aus Social Media Kanälen wie Facebook oder Twitter. Es werden moderne Verfahren aus der maschinellen Sprachverarbeitung (Natural Language Processing) eingesetzt, um die Textdaten zu strukturieren und bezüglich Trends zu analysieren. Dies ermöglicht wirtschaftliche Trendanalysen. Die Relation Extraction Komponente extrahiert Beziehungen zwischen Entitäten in Texten. Als Entitäten in Texten werden Personen, Organizationen und geographische Orte bezeichnet. Relationen bezeichnen die Beziehungen zwischen den Entitäten. Die Analyse einer grossen Sammlung von Wirtschaftsnachrichten führte zu einer Vielzahl von extrahierten Relationen.

Das Ziel dieser Projektarbeit war, diese Relationen mittels einer Visualisierung darzustellen. Die Visualisierung sollte in eine Webapplikation integrierbar sein. Für das Verständnis der zu visualisierenden Daten wird ein Überblick über das Verfahren der Relation Extraction gegeben. Anschliessend werden die entwickelten Konzepte der Visualisierung vorgestellt. Aus den Daten wurden weitere Metriken abgeleitet, um die Visualisierung bezüglich ihrer Aussagekraft zu optimieren. Mit Einsatz des D3 Framework wurde ein Prototyp entwickelt, um die Konzepte zu überprüfen.

Die Projektarbeit wurde im Rahmen des Master-Studienganges Master of Science in Engineering (MSE) an der Hochschule für Technik und Wirtschaft HTW Chur erstellt.