

# **Hinweise zu Abschlussprojekt / Ausarbeitung**

Seminar: Computerlinguistische Analyse von Twitterdaten

Tatjana Scheffler

3.7.2013

## **Kursanforderungen:**

Vortrag + Projekt und Ausarbeitung

Ausarbeitung: ca. 10 Seiten

Abgabetermin der Ausarbeitung: **31.8.2013**

- Projektthema vorher mit mir abklären!
- Erreichbarkeit bei Problemen/Fragen: ab 12.8.

## **Aufbau/Inhalt der Ausarbeitung (Checkliste):**

- Fragestellung, Motivation
- Literatur/State of the Art (kann hier relativ kurz sein)
- Daten (welche, wie erhoben)
- Methoden/Vorgehensweise
- Ergebnisse/Evaluation
- Literaturverzeichnis

## **Wissenschaftliches Schreiben (Tipps)**

Theory → Specific Hypothesis → Experiment → Data →

... Go back to beginning and start again!

- Ein wissenschaftliches Papier ist kein Krimi! Vorgehensweise und die wichtigsten Ergebnisse (Zahlen!) sollten schon im Abstract bzw. in der Einleitung stehen.
- Klarheit und Nachvollziehbarkeit sind am wichtigsten. Einleitung und Zusammenfassung sollten auch für sich alleine sinnvoll sein.
- Behauptungen müssen normalerweise begründet werden. Dies geschieht entweder durch korrektes Zitieren des Papiers, aus dem die Einsicht stammt, oder durch Beleg mit Hilfe von eigenen Daten oder Schlussfolgerungen.
- Geschickte Verwendung von Tabellen und Bildern ist oft sehr hilfreich. Tabellen und Bilder müssen im Text erwähnt, erklärt und eingeordnet werden.
- Korrekte Zitation aller verwendeten Literatur ist zwingend! Es sollte nachgeprüft werden, dass das Literaturverzeichnis und die Referenzen im Text vollständig sind. In der (Computer)Linguistik sind direkte Zitate unüblich. Eher verwendete Erkenntnisse umformulieren bzw. ihre Bedeutung für die neue Studie ausdrücken.
- Verwenden Sie ein Textverarbeitungsprogramm, mit dessen Hilfe Querverweise, Zitationen und Nummerierungen korrekt dargestellt werden und mit dem Sie Formeln, etc. einbinden können. Im Hinblick auf weitere Seminare, Bachelor- und Masterarbeit lohnt sich die Einarbeitung in Latex (siehe unten). Bei Latex-Problemen helfe ich gerne, wenn Ihre anderen Hilfsquellen versiegen.

## **Mögliche Projekt-/Ausarbeitungsthemen:**

- Spamerkennung in Tweets
- Verbesserte Tweet-Suche (z.B. durch Synonyme)
- Tonalitätsanalyse (z.B. "2013" vor/nach Silvester)
- Übertragung eines der behandelten Themen auf deutsche Twitterdaten
- Dialoge auf Twitter
- Generierung von Hashtags
- Analyse eines linguistischen Phänomens
- Identifikation von ortsrelevanten Wörtern
- Topikerkennung/Trenderkennung
- NLP für Twitter: POS-Tagging/Parsing/NE Recognition ...
- Reguläre Hausarbeit – Vergleich von Ansätzen, Evaluierung (deutsche Daten)
- ... (eigene Ideen)

## **Links:**

Wissenschaftliches Schreiben, unter anderem:

<http://linguistik.de/linktipps/thema/?q=Wissenschaftliches%20Schreiben&t=1>

<http://www.wissenschaftliches-arbeiten.org/>

J. Wallenberg, Writing a Linguistics Essay (Folien):

<http://www.staff.ncl.ac.uk/joel.wallenberg/NatureOfLanguage/1007Essay-Oct-30-2012.pdf>

Latex:

<http://latex-project.org/ftp.html>

<http://www.essex.ac.uk/linguistics/external/clmt/latex4ling/> (Latex for Linguists)

<http://www.tex.ac.uk/tex-archive/info/symbols/comprehensive/symbols-a4.pdf> (Latex Symbols)