

Tools der Korpuslinguistik: Möglichkeiten ihrer Anwendbarkeit und internetbasierten Nutzbarmachung

WORKSHOP

am 12. April 2012 in Düsseldorf

Im Rahmen des Workshops „Tools der Korpuslinguistik“ soll die auf der letzten Jahrestagung des Forschungsnetzwerkes *Sprache und Wissen* in Heidelberg entstandene Idee eines Workshops zur Anwendung und Entwicklung korpuslinguistischer Tools in der Diskurslinguistik umgesetzt werden.

Das erste Arbeitstreffen wird vom DFG-Projekt „Sprachliche Konstruktion sozial- und wirtschaftspolitischer Krisen in der BRD von 1973 bis heute“ durchgeführt und dient dazu, die unterschiedlichen Zugänge zum Datenmaterial kennenzulernen und Möglichkeiten ihrer Verknüpfung zu diskutieren. Dabei soll auf die Verknüpfung quantifizierender und qualitativer Methoden der Textanalyse eingegangen werden. Ein besonderes Interesse dieses Arbeitstreffens liegt auf den Möglichkeiten und Grenzen der Vernetzung der verschiedenen korpusanalytischen Werkzeuge, die im Hinblick auf die Idee der Programmierung einer eigenen Korpusanalyse-Cloud diskutiert werden soll.

Die Teilnehmer/innen sind gefragt, in einem kurzen Vortrag (*maximal 20 Minuten*) ein oder zwei ihnen gut vertraute textstatistische Verfahren (z.B. Kookkurrenzanalyse, Signifikanz/Spezifität des Vokabulars, n-Gramme, Konkordanzen, Multifaktorenanalyse u.a.) im Hinblick auf die Verknüpfung qualitativer und quantitativer Analyseverfahren, aber auch im Hinblick auf die Vernetzung unterschiedlicher Analysetools vorzustellen. Die Teilnehmer/innen werden gebeten, sich bei ihren Ausführungen kurz zu fassen, um so eine ausgiebige Diskussion der Werkzeuge und Methoden in ihrer möglichst innovativen Anwendung im Hinblick auf die Verknüpfung quantifizierender und qualitativer Methoden zu ermöglichen. In den Kurzreferaten soll zunächst das grundsätzliche Ziel eines Verfahrens vorgestellt und anschließend eine oder mehrere Ideen zur Anwendung auf die erwähnte Fragestellung vorgetragen werden.

Ein wichtiges Anliegen des Arbeitstreffens besteht darin, die Idee der Entwicklung einer gemeinsamen Analyseplattform zu diskutieren. Dabei geht es zunächst darum, bereits existierende Forschungsinfrastrukturen vorzustellen und über mögliche Anknüpfungspunkte zu diskutieren. Insbesondere sollen Ideen zu technisch-funktionalen, organisatorischen, finanziellen und lizenzrechtlichen Fragen besprochen werden. Die Intention dabei ist, die Möglichkeiten zur Entwicklung einer gemeinsamen Analyseplattform, zum Beispiel als Cloud, abzuwägen. Diese Plattform könnte etwa über eine Datenbankfunktion, textstatistische Messverfahren, POS-Tagging und eine manuelle Annotationsfunktion verfügen. Sie sollte die verschiedenen mehr oder weniger in Eigenregie entstandenen Module wie INGWER, LDA-Toolkit oder TextObserver integrieren und zu einem leistungsstarken Werkzeug der computergestützten Diskursanalyse kombinieren. Dabei sollte die Urheberschaft und die Wertschätzung der einzelnen integrierten Module auf keinen Fall verloren gehen. Im Idealfall wäre die Plattform mit einer Open Source-Lizenz ausgestattet.

Der gesamte Workshop soll eher informellen Charakter haben und vor allem der Diskussion von Ideen zur Weiterentwicklung und zu Kombinationsmöglichkeiten bestehender Tools und Module sowie dem persönlichen Kennenlernen dienen.

**Tools der Korpuslinguistik: Möglichkeiten ihrer Anwendbarkeit und internetbasierten
Nutzbarmachung**

PROGRAMM

12.04.2012

Vortragsraum der Universitätsbibliothek Düsseldorf

Universitätsstr. 1/**Gebäude 24.41**/40225 Düsseldorf

Geokoordinaten (Grad): Lat N 51° 11' 18.394799" / Long. 6° 47' 41.557199"

(Notfallnummer: 01577-1525-467)

11:00h

- 1. Einführung** (David Römer/Ronny Scholz/Alexander Ziem) **und Vorstellungsrunde**
- 2. MARC KUPIETZ**
„KorAP vor dem Hintergrund großer Forschungsinfrastrukturen: Wie Daten, Methoden und Forscher zusammenfinden“
- 3. GISELA MINN**
„Internet-gestützte Zusammenarbeit in Forschungsverbänden – Das Forschungsnetzwerk und Datenbanksystem „Fremdheit und Armut“

13:00 – 14:00h Mittagspause

- 4. ANDREAS BLÄTTE**
„Plenardebattenkorpus und CQPweb: Variable Partitionierung und doch begrenzte Möglichkeiten“
- 5. DAVID RÖMER**
„INGWER – eine Datenbank und Annotationswerkzeug zur Kombination quantifizierender und qualitativer Methoden“

15:00h – 15:15h Kaffeepause

- 6. RONNY SCHOLZ**
„Textobserver – Multivariate Datenverarbeitung als quantifizierende heuristische Methode zur Stilanalyse thematischer Textkorpora“
- 7. FRIEDEMANN VOGEL**
„LDA-Toolkit - Korpuslinguistische Arbeitsumgebung für linguistische Diskurs- und Imageanalysen“

16:30h – kurze Kaffeepause (5min)

- 8. KRISTIN KUCK**
„Analyse von Metaphern korpuslinguistisch - Möglichkeiten und Grenzen“
- 9. VASCO SCHMIDT**
„Vision einer "Linguistik-Cloud“

Diskussion zu Möglichkeiten und Grenzen einer gemeinsamen Korpusanalyse-Plattform im Rahmen eines DFG-Projektes oder/und von CLARIN.

18:30h ENDE

Tools der Korpuslinguistik: Möglichkeiten ihrer Anwendbarkeit und internetbasierten Nutzbarmachung

Liste der Teilnehmer/innen:

1. **JOHANNES ANGERMÜLLER** (Soziologie, Mainz)
2. **ANDREAS BLÄTTE** (Politikwissenschaft, Duisburg/Essen)
3. **NOAH BUBENHOFER** (Korpuslinguistik, IDS Mannheim/semtracks)
4. **MUSTAFA DOGAN** (Digitale Bibliothek SUB Göttingen)
5. **JUAN GARCES** (Digital Humanities Göttingen, GCDH)
6. **KRISTIN KUCK** (Germanistik, Trier)
7. **MARC KUPIETZ** (Korpuslinguistik, IDS Mannheim)
8. **CHRISTIAN KREUZ** (Germanistik, Trier)
9. **GISELA MINN** (Geschichte, Trier)
10. **MARCUS MÜLLER** (Germanistik, Heidelberg/Paderborn)
11. **DAVID RÖMER** (Germanistik, Trier/Düsseldorf)
12. **VASCO SCHMIDT** (Mathematik, Sprachwissenschaft, SAP)(*Zusage unter terminlichem Vorbehalt*)
13. **RONNY SCHOLZ** (Soziologie; Germanistik, Trier)
14. **ALEXANDER SOHN** (Statistik, Göttingen)
15. **CHRISTINA STEIN** (Germanistik/Geschichte, Trier)
16. **FRIEDEMANN VOGEL** (Germanistik, Heidelberg)
17. **MARTIN WENGELER** (Germanistik, Trier)
18. **ALEXANDER ZIEM** (Germanistik, Düsseldorf)

Zur Anfahrt siehe auch:

BAHN: <http://www.ub.uni-duesseldorf.de/home/info/anfahrt/bahn>

und <http://www.ub.uni-duesseldorf.de/home/info/anfahrt/plan.html>

AUTO: <http://www.ub.uni-duesseldorf.de/home/info/anfahrt/auto/>