

David Kaumanns  
Hornstraße 24  
80797 München

Centrum für Informations- und Sprachverarbeitung  
Oettingenstraße 67  
80538 München

Abschlussbericht zum Gastvortrag

## **From XML structured texts to pattern-matching with PERL and the use of the search machine tool Windex, based on LUCENE**

in den Wittgenstein Archives der Universität in Bergen (WAB), Norwegen  
am 20. August 2013

gefördert als studentisches Forschungsprojekt im Rahmen des Projekts Lehre@LMU

"Wittgenstein Scholarship und Wittgenstein in Co-text" ist eine Kooperation zwischen dem Centrum für Informations- und Sprachverarbeitung (CIS) an der Ludwig Maximilians Universität München und den Wittgenstein Archives an der University of Bergen (WAB). Das Ziel des Projektes ist die Erstellung und Verwaltung einer digitalen Bibliothek, bestehend aus den Werken und dem Nachlass des Philosophen Ludwig Wittgensteins sowie der von ihm verwendeten Literatur. Auf diesen Daten werden computerlinguistische Werkzeuge entwickelt, um das Material für die Zwecke philosophischer und philologischer Recherche durchsuchbar zu machen. Die Koordinatoren des Projekts sind Dr. Maximilian Hadersbeck und Dr. Alois Pichler.

Webseite des Projekts:

<http://www.cis.uni-muenchen.de/forschung/ehumanities/research-group-co>

Webseite des WAB:

[http://wab.uib.no/wab\\_hadersbeck2013.page](http://wab.uib.no/wab_hadersbeck2013.page)

Teil meiner Magisterarbeit war die Konzeptualisierung und Implementierung einer Ideensuchmaschine auf dem digitalisierten Nachlass Ludwig Wittgensteins und seinem virtuellen Bücherregal. Das Programm ermöglicht eine linguistisch unterstützte Suche über dem Material und den bequemen Abgleich mit Wittgensteins mutmaßlichen Ideenquellen. Es wird vom Server des CIS zur Verfügung gestellt: <http://windex.cis.lmu.de>

Hierzu wurde ich von Herrn Hadersbeck zu einem Gastvortrag während seines mehrwöchigen Workshops am philosophischen Institut der Universität in Bergen eingeladen. Während meines Aufenthaltes vom 16.8. bis 21.8. besprach ich mit den Institutsangehörigen den aktuellen Stand des Projektes, insbesondere die aktuelle Quellenlage sowie ausstehende und gewünschte Funktionen in einer zukünftigen Version des Programms, und nutzte einen vom Institut zur Verfügung gestellten Arbeitsrechner zur Vorbereitung der Präsentation. Der rege Austausch mit den Philosophen vor Ort ermöglichte eine unmittelbare Rückmeldung und Kritik von den Nutzern, die zu vielen neuen Ideen führte.

In meinem Vortrag am Dienstag, 20.8.2013, erklärte ich die Konzepte und Ideen hinter meiner Arbeit. Hierzu gehörten auch die Probleme, die wir bei der Beschaffung externer Quellen hatten. Die technischen Hintergründe der Korpusbeschaffung und -aufbereitung gaben einen Einblick in die Probleme, die wir von computerlinguistischer Seite zu lösen hatten. Hierzu zählten auch die verschiedenen eingesetzten linguistischen Strategien. Insgesamt weckte der Vortrag ein besseres Verständnis für die Herausforderungen auf unserer Seite der Kooperation zwischen unseren Instituten und führte zu einer Reihe neuer gut fundierter Fragen, Kritik und Ideen von philosophischer Seite.

Die Vortragsfolien sind hier einsehbar (HTML5):

<http://www.cip.ifi.lmu.de/~kaumanns/windex-bergen>

Während der Einführungsveranstaltung am philosophischen Institut für die neuen Studenten am Montag, 19.8.2013, stand ich zusätzlich zusammen mit Patrick Seebauer für Fragen bezüglich des ERASMUS-Programms und den Studienmöglichkeiten in München zur Verfügung.

Zusammenfassend war der Aufenthalt in Bergen eine großartige Gelegenheit, unser Produkt einmal persönlich unter die Leute zu bringen und direkte Reaktionen und Meinungen zu erfahren sowie unsere Zielgruppe persönlich kennen zu lernen und Kontakte zu festigen. Ich hatte viele sehr interessante Gespräche mit den Philosophen, innerhalb und außerhalb des Universitätskontextes, über unser Gemeinschaftsprojekt und Kooperationen zwischen technischen und geisteswissenschaftlichen Einrichtungen im Allgemeinen, die mir in meinem kommenden akademischen Weg von großem Nutzen sein werden.