

22.06.2017

Protokoll für das Kolloquium vom 19.6.2017

In dieser Sitzung wurde drei Bachelorarbeiten von den Studenten in Form von kurzen Vorträgen mithilfe einer kleinen Präsentation am Beamer vorgestellt.

1. Vortrag:

Ines Röhrer, BA-Betreuer: Dr. Maximilian Hadersbeck

Thema: Musik und Ludwig Wittgenstein: Semantische Suche in seinem Nachlass

Ines gab zu Beginn ihres Vortrages eine kleine Einleitung in ihr Thema. Dabei erklärte sie, dass der Fokus ihrer Arbeit sich auf die Suchmaschine WiTTFind und den Nachlass von Ludwig Wittgenstein richtet. Fortführend beschrieb sie WiTTFind, wobei es sich wie wir auch schon aus anderen Vorträgen gelernt haben um eine Suchmaschine für den Nachlass von Ludwig Wittgenstein handelt. In dieser Suchmaschine gibt es zwei unterschiedliche Suchoptionen, nämlich eine regelbasierte und eine für semantisches Suchen. Sie führte das Beispiel der Suchkategorie Farbe heran, da sie eine ähnliche Struktur wie die dort verwendete für ihre eigene Arbeit benutzte.

Dann beschrieb sie die unterschiedlichen Teile des Wittgenstein-Nachlasses. Er besteht aus öffentlich zugänglichen Teilen in Deutsch und Englisch sowie aus noch nicht veröffentlichten Teilen. Die aktuelle Version von WiTTFind arbeitet nur mit den öffentlich zugänglichen Teilen des Nachlasses. In ihrer Arbeit benutzte sie jedoch auch Teile, die noch nicht frei verfügbar sind. Dabei erwähnte sie, dass die nicht frei zugänglichen Teile noch um ein Vielfaches größer sind als die veröffentlichten. Der Nachlass besteht aus zwei Arten von Texten, Manuskripte und Typoskripte. Diese sind mit den Bezeichnungen Ms und Ts markiert und enthalten zusätzlich noch eine einzigartige Identifikationsnummer.

Im folgenden stellte Ines dann Motivation und Ziel ihrer Arbeit vor. Hier geht es um eine Erweiterung der semantischen Suchoptionen von WittFind. Da Musik eine wichtige Rolle in Wittgensteins Leben spielte, soll ihre Arbeit es möglich machen, musikalische Erwähnungen im Nachlass zu untersuchen, weshalb sie dafür ein neues Modul zur Suche von musikalischen Begriffen für die Webseite erstellen und die Musikbegriffe strukturieren sollte. Ein weiteres Ziel war die Untersuchung über Ontologien für diese Musikbegriffe und die Erkennung von Relationen zwischen diesen Ausdrücken.

Sie fuhr dann fort mit einer Erläuterung ihrer Vorgehensweise bei dieser Aufgabe. Wichtiger Teil davon war die Erweiterung des Webfrontends. Dafür diente ein Tutorial zur Aufsetzung eines lokalen Webserver, Erweiterungen der HTML-Datei nach dem Vorbild der Farbsuche und Erweiterungen für die semantische Suche in der Javascript-Datei. Auch die Struktur des Webfrontends war an die Farbsuche angelehnt, wobei die Kategorien hauptsächlich auf dem frei zugänglichen Teil des Nachlasses entstanden. Die Basis für die Auswahl der Musikbegriffe war dabei die Hausarbeit eines Studenten der Musikwissenschaften. Dabei ging es vor allem darum, die richtigen Begriffe auszuwählen. Es musste darauf geachtet werden, dass nicht alle Begriffe übernommen werden können, da sie durch Ambiguität die Frequenzberechnung stark verfälschen würden.

Daraufhin gab Ines einen Überblick über die Kategorien, in die die gefundenen Begriffe eingeteilt werden sollten. Diese sind Komponisten, Instrumente, Gattungen, Intervalle, Bezug zu Komposition und Sonstige Begriffe, eine Kategorie, in die Musikbegriffe einsortiert werden sollten, deren Kategorie noch nicht klar oder noch nicht implementiert ist. Hier erwähnte Ines, dass es sich

keinesfalls um eine vollständige Liste von Kategorien handelt und diese noch erweitert, während andere möglicherweise wieder in Überkategorien zusammengefasst werden sollten. Ines erklärte hier dass die Begriffe in Wordclouds angezeigt werden sollten in welchen die Frequenz der Begriffe durch die Größe der Wörter symbolisiert wird. Für diese Darstellung sind Frequenzberechnungen der Ausdrücke nötig. Dafür werden die Wörter in ein Dictionary mit der Herkunftsdatei als Value gespeichert. Diese sollten dann beim Textdurchlauf gezählt werden. Dabei war vor allem auch die Vollformenextraktion wichtig um die richtige Frequenz zu ermitteln. Da ihr zu Beginn ihrer Arbeit kein Vollformenlexikon zur Verfügung stand benutzte sie Endungslisten für die Erkennung von flektierten Formen. Später versuchte sie es noch mit einem Lexikon, dies führte jedoch zu schlechteren Ergebnissen. Im Anschluss beschrieb Ines noch einige Probleme bei der Frequenzberechnung.

Danach erklärte Ines kurz dass sich im Verlauf ihrer Arbeit der Fokus etwas auf Kontextextraktion verschob. Hierfür zog sie mehrere Methoden in Betracht und entschied sich dann es mit einem Ringbuffer und Listenoperationen zu versuchen. Die Extraktion des Kontextes sollte gleichzeitig mit der Frequenzberechnung stattfinden. Sie gab dann zu jeder Methode einen kurzen Überblick und Erklärung.

Im Vergleich erwies sich der Ringbuffer als weniger effizient weshalb sie die Listenoperationen in der finalen Fassung ihres Systems verwendete.

Der letzte wichtige Teil ihrer Arbeit waren Ontologien um die Relationen zwischen den Musikbegriffen darzustellen. Sie entdeckte während ihrer Arbeit eine schon bestehende Musikontologie mit welcher ihr eine erste Modellierung gelang. Zwar sprengte eine Weiterentwicklung den Rahmen ihrer Arbeit doch sah sie gute Aussichten für die zukünftige Entwicklung dieses Systems.

Zum Schluss sprach Ines über die Erkenntnisse die sie während ihrer Arbeit erlangt hatte. So erklärte sie dass sich die Ziele ihrer Arbeit in ihrem Verlauf stark verändert hatten, die Ergebnisse jedoch zufriedenstellend waren.