

Repetitorium Computerlinguistik: Wissenschaftliches Arbeiten

Für die Anfertigung einer Abschlussarbeit in der Computerlinguistik gelten andere Regeln als in anderen Studiengängen. Nicht zuletzt der anfallende praktische Programmieranteil und dessen Evaluation stellen hierbei eine Herausforderung dar. Um die Studenten beim Schreiben ihrer Bachelorarbeit zu unterstützen, präsentiert Frau Sedinkina Tipps für wissenschaftliches Arbeiten.

Gliederung und einzelne Kapitel

Als erstes geht Frau Sedinkina auf die Gliederung der Bachelorarbeit ein. Mit dieser kann sich der Leser einen Überblick über den Inhalt verschaffen. Hier ist vor allem ein logischer Aufbau wichtig: Allgemeines soll vor Konkretem, die Motivation vor der Ausführung beschrieben werden. Überschriften sollen im Nominalstil gehalten und Wiederholungen möglichst vermieden werden.

Frau Sedinkina zeigt eine Gliederungsvorlage, die von Forschungsaufsätzen inspiriert wurde. In der dort ersichtlichen Reihenfolge geht sie näher auf die einzelnen Unterkapitel der Arbeit ein.

Die Einleitung der Bachelorarbeit sollte nicht länger als eine Seite sein. Hier wird der Leser an das Thema herangeführt. Es soll davon ausgegangen werden, dass dieser selbst Computerlinguist ist, aber auf ein anderes Gebiet spezialisiert ist. Anstatt der Einleitung ein eigenes Kapitel zu widmen, lassen sich die Gliederungspunkte "Kurzübersicht", "Thematischer Hintergrund" und "Fragestellung" auch zu einem Kapitel zusammenfassen.

Unter dem Gliederungspunkt "Thematischer Hintergrund" wird auf die Relevanz des Themas eingegangen, z.B. indem auf Lücken in der Forschung hingewiesen wird, die der eigene Forschungsbeitrag zu schließen vermag. Dabei wird auch, in Form einer Fragestellung, die Forschungsfrage formuliert.

Nach einem Kapitel zur Einbettung und Zusammenfassung relevanter Literatur folgt der Hauptteil, der etwa zwei Drittel der Arbeit ausmachen soll: Das Kapitel über den eigenen Forschungsbeitrag.

Dieser soll in den Unterkapiteln unter Einbezug von Formalismen, Beiweisen und Konzepten im Detail beschrieben und danach mittels Experimenten quantitativ untersucht werden. Dafür können z.B. bekannte Evaluationsmaße wie Precision, Recall oder F1-Measure verwendet werden.

In der einseitigen Zusammenfassung werden schließlich die wichtigsten Ergebnisse betont.

Die Kenntnis um die Lektüre der Arbeit darf dabei vorausgesetzt werden. Eine Einordnung des Beitrags in den Forschungskontext sowie ein Ausblick runden die Zusammenfassung ab.

Literaturrecherche

Wie trifft man seine Auswahl an Fachliteratur?

Dafür empfiehlt Frau Sedinkina den Studenten vor allem, sich beim jeweiligen Betreuer der Arbeit Literaturhinweise einzuholen. Auch sollte nach Konferenzen und Journalen recherchiert werden, die sich mit ähnlichem Thema befassen. Hierfür stellt Frau Sedinkina den Studenten URLs zu bekannten Konferenzen und Journalen bereit. Abgesehen davon könne in der Bibliothek nach Lehrbüchern Ausschau gehalten oder die Suchmaschine "Google Scholar" verwendet werden.

Für die schnelle Einschätzung der Eignung gegebener Literatur für die Bachelorarbeit schlägt Frau Sedinkina eine altbewährte Technik vor: Das Querlesen. Dabei sollen pro Paper lediglich die wesentlichen Abschnitte wie z.B. Abstract und Summary gelesen und dabei nicht mehr als 5 Minuten investiert werden. Erste Einschätzungen zur Relevanz jedes Papers können dann z.B. tabellarisch zusammengefasst werden, um basierend darauf schließlich eine Untermenge von Papers näher in Betracht zu ziehen.

Schreibstil

Die Arbeit soll sachlich und größtenteils im Präsens geschrieben sein. Lediglich bei als vergangenes Ereignis dargestellten Verweisen ("Hearst (1992) entwickelte das erste Verfahren zur Extraktion taxonomischer Relationen aus Texten.") auf Fremdliteratur sowie vereinzelt bei Beschreibungen des Experiments ("Es wurden zwei Modelle mit Parametern 1, 2 und 3 trainiert") sollte das Präteritum verwendet werden.

Schreibprozess

Nach klassischer Vorgehensweise erstellt man zunächst die Gliederung und befüllt anschließend die einzelnen Punkte mit Inhalt. Aber nicht immer eigne sich dieses Vorgehen, da sich die Gliederung ohnehin meistens im Laufe des Schreibens der Arbeit nochmal ändert. Außerdem kann der Schreibfluss behindert werden, wenn versucht wird, sich zu streng an einen Gliederungspunkt festzuhalten.

Als alternatives Vorgehen, das diese Blockade zu verhindern versucht, schlägt Frau Sedinkina den Schreibprozess nach Bolker vor. Nach diesem wird zuerst das Material beschafft und erst danach gegliedert. Für die Materialbeschaffung ist es wichtig, in den Schreibfluss zu kommen. Dabei kann es helfen, einfach „drauf los“ zu schreiben, ungeachtet dessen, ob das Resultat etwas mit dem Thema zu tun hat. In mehreren Iterationen kann der Text danach dann überarbeitet und gegliedert werden.

Anschließend sollte man dritte Personen Korrekturlesen lassen.

Auch Kombinationen aus dem klassischen Ansatz und dem Bolkers seien denkbar. Probieren geht hier über studieren.

Bewertungskriterien

Abschließend zählt Frau Sedinkina Kriterien auf, anhand derer eine Bachelorarbeit in Computerlinguistik üblicherweise bewertet wird. Darunter fällt z.B. der Grad des logischen Aufbaus der Arbeit, der Abdeckung wichtiger Literatur sowie der Eigenständigkeit des Studenten bei der Anfertigung der Arbeit. Diese Aufzählung sei aber nicht vollständig, weswegen jedem Studenten geraten wird, sich diesbezüglich näher beim jeweiligen Betreuer zu informieren.