



Der Computer als Lernhilfe  
Foto: dpa

## Aufsatzanalyse per Computer

### Adieu Abschreiber

**Forscher an der Uni Würzburg haben ein Programm entwickelt, mit dem sich Aufsätze und Tests von Schülern und Studenten in Millisekunden bewerten lassen. Auch hinsichtlich ihrer Einzigartigkeit.**

*Von Jörg Donner*

Auf die Ergebnisse ihrer Vordiplomprüfung mussten die Studenten der Uni Würzburg in der Vergangenheit mindestens einige Tage warten. So lange dauerte es, bis beispielsweise die Arbeiten von 400 Psychologie-Studenten von mehreren Hilfswissenschaftler korrigiert und ausgewertet waren. In ein paar Jahren könnten die Studenten schon wenige Sekunden nach Abgabeschluss erfahren, welche Note sie erreicht haben. Ganz ohne Zutun von Hiwis oder Professoren. Das zumindest ist eines der Ziele der neuen Software „coNtext“ an dem die Forschungsgruppe um Dr. Wolfgang Lenhard vom Lehrstuhl für Psychologie arbeitet.

„Unser Hauptanliegen besteht darin, eine intelligente Lernsoftware zu entwickeln“, sagt Lenhard. „Damit erhalten Schüler und Studenten sofort eine qualifizierte Rückmeldung, wie gut ihre Arbeit ist und wo sie eventuell noch verbessert werden kann.“ So kann das Programm beispielsweise direkt nach Beantwortung einer Frage darauf hinweisen, ob noch wichtige Details ungeklärt geblieben sind oder gar eine Themaverfehlung vorliegt.

#### Geballtes Wissen aus Fachbüchern

**» Unsere Software korrigiert 400 Vordiplomklausuren innerhalb von vier Sekunden. «**

**Dr. Wolfgang Lenhard**

Grundlage des Programms ist der sogenannte Textkorpus. Er besteht aus einer Datenbank, die mit Informationen passend zum Thema gefüttert wird. Etwa 70.000 verschiedene Wörter aus 30.000 Absätzen in 14 Lehrbüchern bilden beispielsweise die Grundlage für die

testweise Auswertung einer Vordiplomklausur im Fach allgemeine Psychologie. Dabei wurden 50 Studenten jeweils acht Fragen gestellt und das Ergebnis sowohl vom Computer als auch von Professoren beurteilt. Das verbüffende Ergebnis: Eine hohe Übereinstimmung in der Beurteilung der fachlichen Leistungen der Studenten. Möglich ist die Computeranalyse durch eine rein statistische Berechnung. Aufgrund des Vorkommens von Worten in Fachtexten und -büchern bewertet coNtext deren Relevanz.

„Die Häufigkeit spielt nur eine untergeordnete Rolle“, sagt Entwickler Dr. Lenhard, „Artikel wie ‚der‘, ‚die‘ und ‚das‘ kommen zwar oft vor, sind aber mehr oder weniger wahllos über den Text verstreut. Wichtiger ist daher, in welchem Zusammenhang Worte stehen.“ Beispielsweise stünden „Bundeskanzlerin“ und „Merkel“ in einem sehr engen Kontext, die Worte ‚Bush‘ und ‚Massenvernichtungswaffen‘ kämen ebenfalls relativ häufig zusammen vor. „Computer“ und „Nietzsche“ seien dagegen sehr weit voneinander entfernt.

Durch diese Art der Auswertung entsteht ein sogenannter semantischer Raum mit 300 bis 500 Dimensionen. Darin werden die Worte grafisch angeordnet und mit Vektoren zu einem Ursprung verbunden. Je kleiner der Winkel zwischen den Vektoren ist, desto größer die Wahrscheinlichkeit, dass diese Worte im gleichen Zusammenhang auftreten. So sollte in einem Text, in dem „Merkel“ auftaucht auch das Wort „Bundeskanzlerin“ zu finden sein. Darüber hinaus merkt coNtext auch, wenn innerhalb eines Absatzes die Aussage vom eigentlichen Ziel abdriftet, also eine Themaverfehlung vorliegt.

Im besten Fall kann das Programm einen Schüler schon nach wenigen Sätzen darauf hinweisen, dass er offenbar einen falschen Lösungsansatz gewählt hat. „Besonders gut funktioniert die Auswertung bei wissensbasierten Texten, also etwa Biologie, Erdkunde oder Geschichte“, erklärt Lenhard. „Aus der Datenbasis, die für jedes Fach neue eingespeist werden muss, errechnet das Programm in etwa 20 Minuten das Auswertungsschema. Die Korrektur einer Arbeit dauert dann nur noch Millisekunden.“

### **Plagiate sind so gut wie ausgeschlossen**

Neben der schnellen Korrektur bietet coNtext einen weiteren Vorteil: Durch den Vergleich aller Arbeiten untereinander sind Plagiate und Gemeinschaftsarbeiten so gut wie ausgeschlossen. Ähnlichkeiten und Übereinstimmungen werden sofort erkannt, egal wie viele Prüflinge zum Test antreten. Bei 400 Kandidaten müssten alle Klausuren miteinander verglichen werden, um Gemeinsamkeiten zu erkennen. Rein rechnerisch also 400 mal 400 Möglichkeiten. Da der Vergleich in einer Richtung genügt, also Kandidat A mit Kandidat B und nicht noch zusätzlich Kandidat B mit A, bleiben etwa 80.000 Vergleiche. Für den menschlichen Korrektor eine unlösbare Aufgabe. „Es gibt bereits Programme, die kopierte Textstellen in Aufsätzen erkennen“, sagt Lenhard, „verändert man die Textstellen marginal, rutschen sie aber bereits durchs Raster.“

Anders bei coNtext: Hier würden auch Antworten als Plagiate enttarnt, die in der Satzstellung oder Wortwahl verändert wurden. Abschreiben wäre damit fast unmöglich, verdächtige Dokumente würden sofort gekennzeichnet und zur letzten Beurteilung an einem menschlichen Korrektor weitergeleitet.

Momentan befindet sich die Forschungsgruppe um coNtext noch in der Entwicklungsphase. Der Prototyp des Programm ist das Ergebnis des ersten Jahres des Projekts, das insgesamt auf fünf Jahre angelegt ist. Um das Programm in der Praxis zu testen sucht die Uni Würzburg noch Schulen in Unterfranken, die mit ihren Klassen von der fünften bis zur achten Jahrgangsstufe an der Entwicklung teilnehmen wollen.

sueddeutsche.de

im Netz

✉ [Universität Würzburg](#) - Kontakt zum Forscherteam

**Artikel drucken** ✉  
**Fenster schließen** ✉