

Benutzerhandbuch zu COSMAS

Doris al-Wadi: COSMAS – Ein Computersystem für den Zugriff auf Textkorpora. Version R.1.3-1. Benutzerhandbuch. Mit einem Geleitwort von Gerhard Stickel. Institut für deutsche Sprache. Linguistische Datenverarbeitung. Mannheim, 1994. – XII, 278 S. Preis: 48,- DM.

Wörter verlangen Beispiele, die Beispiele gewährt, ohne welche ihre beste Kraft verloren gieng. wie könnten Stellen (loci) heissen, deren Stelle ungenannt bliebe? der Name ihres Urhebers reicht nicht aus, sie müssen aufgeschlagen werden können; aus der Leichtigkeit dieses Nachschlagens entspringt ein grosser Reiz, denn wie genau auch die Belege ausgehoben seien, der Leser hat nicht selten das Bedürfnis sie in ihrem vollständigeren Zusammenhang einzusehen: indem er weiter vordringt, findet er dicht neben den beigebrachten ausdrücken noch etwas anderes, unmittgetheilt gebliebenes, wodurch ihm das Verständnis vollends erschlossen wird. auch in der klassischen Philologie ist es hergebracht die Quelle anzuführen, aus der entnommen wurde. unbelegte Citate sind unordentlich zusammengepackt, unbeglaubigte, unbeeidete zeugen. (Deutsches Wörterbuch von Jacob Grimm und Wilhelm Grimm. Vorrede. 1. Band: A – Biermolke. – Leipzig: S. Hirzel, 1854. – Spalte XXXVI).

In der Vorrede zum Deutschen Wörterbuch von Jacob Grimm und Wilhelm Grimm wird im Jahre 1854 ein Wunsch geäußert, der in der Zeit der elektronischen Datenverarbeitung mit einem Rechensystem wie COSMAS schnell erfüllt werden kann: den wissenschaftlichen Wünschen heutiger Germanisten und Sprachwissenschaftler, Philologen und Psychologen eröffnen sich neue, ungeahnte Möglichkeiten des »Aufgeschlagenwerdens« und des »Nachschlagens«.

Mit COSMAS können ein einzelnes Wort, eine Wortgruppe, einzelne Zeichen und Zeichenketten, Wörter innerhalb eines definierten Rahmens in verschiedenen Abständen zueinander und vieles andere mehr gesucht werden, und meist in wenigen Sekunden schon zeigt COSMAS als Ergebnis der automatischen Belegrecherche die Suchobjekte im gewünschten Kontext und mit bibliographischen Beschreibungen, die den computerlesbar gespeicherten Textkorpora zur geschriebenen und gesprochenen deutschen Sprache entstammen. Eine erhebliche Erleichterung

für alle Studierenden, Lehrenden und Forschenden, die nicht mehr nach langem Suchen in Büchern, Zeitschriften und Zeitungen zu exzerpieren, sondern sich nur des Computersystems zu bedienen brauchen!

Jacob Grimm und Wilhelm Grimm hatten – ohne ein Retrieval-System wie COSMAS – noch beträchtliche Schwierigkeiten, die Beispiele in ihrem Wörterbuch nicht bloß zu erfinden, sondern korrekt zu zitieren. Für vage Zitatangaben wie die folgenden erntete Jacob öffentlich einige Kritik, die er mit den in der »Vorrede« vorgetragenen »Grundsätzen« zu beschwichtigen suchte:

- doch LUTHER sagt nur *adeler*, GÖTHER nur *adler*, SCHILLER könnte *aar* gebrauchen.
- Oft heisst es *zum abzuge blasen: der obergott war froh, befahl nun abzublase*. OPITZ;
- bei LOGAU heisst es: ...
- zuweilen geht *ander* in den begriff des *linken* über. *das ufer des flusses, wo wir stehn, heisst uns das rechte, das gegenüberliegende das andere oder linke*.
- ANDERN, n. *oculus*, ein seltnes, nur einmal gelesenes wort: ...
- wer soll nun des vertrauens anfang machen; SCHILLER;
- ANFÜHREN ... ein wort, eine stelle, ein buch, einen verfasser *anführen*, zum zeugen beibringen, citieren [ohne Quellenangabe!].
- ANFUNKELN, ein schönes wort ... mit wild anfunkelnden augen. VOSS;

Das menüorientierte Programmsystem COSMAS (Corpus Storage, Maintenance, and Access System), für das Robert Neumann und Cyril Belica vom Institut für deutsche Sprache einen detaillierten Systementwurf und Entwicklungsplan konzipiert hatten, wurde in enger Zusammenarbeit mit der Softwarefirma MAKROLOG GmbH (Wies-

baden) erarbeitet und wird seit Mitte 1992 angewendet, um für unterschiedlichste sprachwissenschaftliche und andere sprachrelevante Untersuchungen die umfangreichen, computerlesbar aufbereiteten und als Volltextdatenbank organisierten Textmengen (Korpora) des Instituts für deutsche Sprache zu erschließen, zu pflegen und weiterzuentwickeln.

Die Programmentwicklung wurde von Andreas Herberger (MAKROLOG) und Cyril Belica (IDS) geleistet, und Robert Neumann ist bei der weiteren Vervollkommnung des Systems der verantwortliche Koordinator von Seiten des IDS.

Nach einer intensiven Testphase und nach fruchtbringenden Diskussionen mit Anwendern aus verschiedenen Disziplinen liegt jetzt das Benutzerhandbuch von Doris al-Wadi vor, das vor allem aus der Sicht des germanistischen Linguisten – sei er Muttersprachler oder, wie in vielen Fällen, Nichtmuttersprachler – den Zugang zum benutzerfreundlichen Retrievalsystem beschreibt, erklärt und an Beispielen aus der praktischen linguistischen Arbeit erläutert.

In einer sehr verständlichen Fachsprache werden die Benutzeroberfläche von COSMAS, die spezielle Tastaturverwendung, die konkreten Suchen, die mehrstufige Ergebnispräsentation (Textstatistik, KWIC-Übersichten, Einzelbelege mit beliebig groß gewähltem Kontext) und die einzelnen aufrufbaren Menüs beschrieben. Einen breiten Raum nehmen die Beispiele für das konkrete Umsetzen von Suchanfragen in die Suchanfragesyntax des Systems ein, denn COSMAS erlaubt auf Grund seines Lemmatisierungsprogramms (»Flexionsanalyse und Kompositazerlegung« von Cyril Belica) auch Recherchen nach Flexionsformen, nach Komposita und sonstigen Wortbildungsformen zu einem in der Suchanfrage enthaltenen unflektierten Wort oder Wortbildungs-morphem.

Als besonders benutzerfreundlich ist hervorzuheben, daß Doris al-Wadi ihre Beschreibungen durch abgebildete Bildschirmabzüge illustriert, so daß der Benutzer dieses Handbuchs sofort erkennen kann, wie der Bildschirm aussieht, wenn er sich z. B. durch Aufrufen des Menüpunktes »Fenster« frühere Suchen zeigen läßt. Ein Anhang mit einem Überblick über alle verwen-

deten Beispiele, eine Übersicht über alle Bildschirmabbildungen, eine Übersicht über die Textkorpora des IDS und ein Stichwörterregister sind zusätzliche Orientierungshilfen.

Auf den knapp 300 Seiten dieses Handbuchs für Anfänger, für Eingeweihte und für Experten kann man alles schnell und vor allem in einem von Laien wie Spezialisten akzeptierten Sprachstil finden, um mit dem System COSMAS, das bereits von ungefähr 50 Nutzern aus aller Welt on line genutzt wird, gut und erfolgreich arbeiten zu können.

Irmtraud Jüttner

Doris al-Wadi und Irmtraud Jüttner sind wissenschaftliche Mitarbeiterinnen in der zentralen Arbeitsstelle Linguistische Datenverarbeitung am Institut für deutsche Sprache.