

GLDV vergibt Preis für die beste studentische Arbeit

Die Förderung des wissenschaftlichen Nachwuchses gehört zweifellos zu den wichtigen Aufgaben der Gesellschaft für Linguistischen Datenverarbeitung. Wenn nun die GLDV im Rahmen der Vorbereitung ihrer Leipziger Jahrestagung '97 dazu aufgefordert hatte, studentische Arbeiten zu einem Wettbewerb einzureichen, dessen Ziel die Vergabe eines mit 1.000 DM dotierten "Preises für die beste Arbeit" war, dann war es im Prinzip bereits der zweite Anlauf. Betrachtet man allerdings Vorbereitung, Durchführung und Ergebnisse dieses Wettbewerbs, wo wird man berechtigterweise von der ersten offiziellen Vergabe des GLDV-Preises sprechen, den *Anke Kölzer* am Abend des 18.3. 1997 für ihre Diplomarbeit entgegennahm.

Die Jury hatte sich zwischen 3 Arbeiten zu entscheiden gehabt, die ohne Ausnahme auf außerordentlich hohem Niveau präsentiert wurden.

Wenngleich man sich mehr Einreichungen gewünscht hätte, kann man im Hinblick auf Qualität und thematischer Abdeckung nur von einer geglückten Werbung für die Idee des Wettbewerbs und die vertretenen Institutionen sprechen. Im einzelnen handelte es sich um folgende Beiträge:

- *Anke Kölzer (Universität Koblenz): "Lexana - Ein System zur Lexikon- und Grammatikanalyse für Kategoriale Unifikationsgrammatiken"*
- *Sergej Melnik (Universität Leipzig): „EPOS: Electronic Publishing mit OODB und SGML"*
- *Karsten Böhm; Timo Böhme (Universität Leipzig): "SQUIRREL: DBS-based Information Retrieval System for the WWW"*

Dem Bereich Information Retrieval ist die Arbeit von Karsten Böhm und Timo Böhme zuzuordnen. Sie beschreibt Konzeption und Implementierung eines Prototyps, der über das WWW den öffentlichen Zugriff auf Kollektionen von SGMLDokumenten ermöglicht (Zugriff unter <http://leipzig.informatik.uni-leipzig.de:8080/Squirrel/search.html>). Technisch gesehen handelt es sich um die Integration eines Information Retrieval Systems mit einem relationalen Datenbanksystem auf der Basis von Java- Werkzeugen und Programmierung, die das System portabel halten. Das Besondere der (graphischen) Oberfläche ist der "context tree", der die gerankten Rechercheergebnisse in einer inhaltlichen Cluster-Struktur anzeigt. Zur Erprobung des Systems wurden Manuals und News-Artikel in einer Kollektionsgröße von etwa 10 MB mit ermutigenden Ergebnissen benutzt.

Direkt das Tagungsthema "*Linguistik und neue Medien*" adressierte der Beitrag von Sergej Melnik. Mit EPOS wurde ein überzeugendes Konzept für eine gegenwärtig an vielen Stellen erkannte und bearbeitete Problemstellung mit einer ebenso überzeugenden Realisierung vorgestellt: Es geht um die datenbankgestützte Verwaltung strukturierter Texte und Daten, die u. a. über das WWW verfügbar gemacht werden sollen. EPOS geht von einer Modellierung und Auszeichnung (mark up) der Daten mittels SGML aus, die es dann in einer objektorientierten Datenbank verwaltet. Den konzeptionellen Kern des Systems bildet eine einfache Substitutionsprache, mittels der sich SGML-konform kontextabhängige Layout-Vorgaben formulieren lassen. Diese Layout-Varianten ermöglichen die Ausgabe als Print-Dokument, als Übersichts- und Navigationsdokument, als "gewöhnliches" WWW-Dokument oder auch als Eingabe- und Änderungsformular, direkt über das Web. Durch Implementierung in Java ist EPOS vollständig portabel. Neben der ursprünglichen Anwendung des Informatik-Vorlesungsverzeichnisses (<http://www.informatik.uni-leipzig.de>) wurden bereits weitere Anwendungen (Studienführer LD V /CL:<http://leipzig.informatik.uni-leipzig.de:8080/links/sf>) komplikationslos mit EPOS realisiert.

Die Arbeit von Anke Kölzer ist im engeren Sinn der Computerlinguistik zuzurechnen. Der Kontext dieser Diplomarbeit ist das System SYSLID, die Prolog-Implementation einer Kategorialen Unifikationsgrammatik (UCG), die zur Analyse sprecherunabhängiger deutschsprachiger Auskunftsdialekte eingesetzt wird (Kooperationsprojekt zwischen FORWISS, Erlangen und Daimler-Benz-AG, Forschung und Technik, Ulm).

Da in einer UCD das grammatische Wissen fast vollständig in den einzelnen Lexikoneinträgen kodiert ist, ist nur schwer nachzuvollziehen und zu kontrollieren, was die realisierte Grammatik tatsächlich leistet. Diesen Nachteil mittels eines geeigneten Werkzeuges zu kompensieren war die Aufgabe, und Lexana die Implementierung des Lösungsansatzes. Lexana generiert zu einem gegebenen UCG-Lexikon die bildbaren Phrasen und führt auf dieser Basis für die im Lexikon angelegten Strukturen eine Fehleranalyse durch. Die konkrete Anwendung für SYSLID bedeutete naturgemäß die Bewältigung einer daten- und rechenintensiven Aufgabe, deren (überraschende) Ergebnisse nachhaltigen Einfluß auf die weitere Arbeit mit und an SYSLID haben.

Bei der Preisverleihung, die den Höhepunkt einer gelungenen Abendveranstaltung in der Leipziger Moritzbastei darstellte, würdigte Gerhard Knorz (FH Darmstadt) stellvertretend für die Jury die erfreuliche Qualität der eingereichten Beiträge. Die Studenten Karsten Böhm, Timo Böhme und Sergej Melnik erhielten als Anerkennung neben einer Urkunde einen großzügigen Büchergutschein, und unter großem Beifall wurde Frau Anke Kölzer der GLDV-Preis überreicht. Die Arbeit wurde im übrigen vom Herausgeber in die Proceedings der Jahresta-

gung aufgenommen. Wenn die Stimmung dieser Veranstaltung sich auch nur in Ansätzen konkret materialisieren wird, wird der GLDV-Preis beim nächsten Mal neben der Qualität auch die angemessene Quantität vorzuweisen haben. Der GLDV-Preis 1997 für die beste studentische Arbeit als der Beginn einer wunderbaren Entwicklung... (sozusagen).

Gerhard Knorz, FR Darmstadt

Linguistik im deutschen WWW

Der Linguistik-Server-Essen LINSE, der im übrigen seine WWW-Adresse geändert hat (Neue URL: <http://www.linse.uni-essen.del>) bietet eine große Bandbreite von Informationen (siehe entsprechende Meldung im letzten Heft). Unter anderem enthält er eine umfangreiche Sammlung von Links auf Linguistik-Sites im WWW, und das weltweit. Wie Deutschland durch diese Brille aussieht, sehen Sie im folgenden. Neugierig geworden?

- Bamberg: Universität Bamberg, Lehrstuhl für Deutsche Sprachwissenschaft u. Ältere dt. Lit.
- Bayreuth: Universität Bayreuth, Literaturwissenschaft berufsbezogen
- Berlin: Babel (am Lehrstuhl Computerlinguistik der Humboldt-Universität zu Berlin)
Humboldt Universität Berlin, Computerlinguistik
Freie Universität Berlin (FU), Fachbereich Germanistik
- Bielefeld: Universität Bielefeld, Fakultät für Linguistik und Literaturwissenschaft, DaF
- Bochum: Universität Bochum, Sprachwissenschaftliches Institut
- Bonn: Gesellschaft für Linguistische Datenverarbeitung
Universität Bonn, Institut für Kommunikationswissenschaft und Phonetik
- Chemnitz: Universität Chemnitz, Germanistik
- Düsseldorf: Universität Düsseldorf, Allgemeine Sprachwissenschaft
Universität Düsseldorf, Germanistisches Seminar
- Duisburg: Universität Duisburg, Computerlinguistik
DISS - Duisburger Institut für Sprach- und Sozialforschung
- Erfurt: EESE (Erfurt Electronic Studies in English.c.)
- Erlangen-Nürnberg: Universität Erlangen-Nürnberg, CLUE - Computerlinguistik Uni Erlangen

- Essen: Universität Essen, LINSE (Linguistik-Server-Essen)
Universität Essen, Mediävistik
- Frankfurt: TITUS an der Universität Frankfurt a. M.
Universität Frankfurt, Arbeitsbereich Kognitive Linguistik
- Freiburg: Universität Freiburg, Computerlinguistik
Universität Freiburg, Pragmatisches Kolloquium
- Gießen: Universität Gießen, FB 09
- Göttingen: Universität Göttingen, Sprachwissenschaftliches Seminar
- Halle: Universität Halle, Germanistik
- Hamburg: Prowitec
Universität Hamburg, Fachbereich Sprachwissenschaften
- Karlsruhe: Pädagogische Hochschule Karlsruhe, Sprachlaborzentrum
- Kiel: Universität Kiel, Institut für Phonetik und digitale
Sprachverarbeitung
- Koblenz: Universität Koblenz, Computerlinguistik
Universität Koblenz, Germanistik
- Köln: SLIP (Studentisches Linguistik-Forum)
Universität Köln, Inst. für deutsche Sprache und Literatur mit
volkskundlicher Abteilung
Universität Köln, Sprachliche Informationsverarbeitung
- Konstanz: Universität Konstanz, Sprachwissenschaften
- Leipzig: Universität Leipzig, Germanistik
- Mannheim: COSMAS (am Institut für deutsche Sprache)
Institut für deutsche Sprache (IdS)
- Marbach: Schiller - Nationalmuseum/Deutsches Literaturarchiv,
Marbach am Neckar
- München: CIS (Centrum für Informations- und Sprachverarbeitung)
Goethe- Institut
,jetzt online' (am Goethe-Institut)
LEO Dictionary German <-> English
Universität München, Sprachwissenschaften (Fak. 14)
- Münster: Universität Münster, Institut für allgemeine
Sprachwissenschaft Universität Oldenburg, Fachbereich 11
- Oldenburg: Universität Oldenburg, Fachbereich 11
- Osnabrück: Universität Osnabrück, Computerlinguistik und Künstliche In-
telligenz
- Passau: Universität Passau, Linguistik
Deutsche Gesellschaft für Sprachwissenschaft (DGfS)
- Potsdam: Universität Potsdam, Institut für Germanistik
- Regensburg: Universität Regensburg, Institut für Germanistik

- Saarbrücken: SUSY MT system (am Rechenzentrum der Universität Saarbrücken, bescheidene automatische online- Übersetzung Deutsch-Englisch und Russisch-Deutsch. Als login gibt man "susytest3" ein, als Password "susyguest3".)
Universität des Saarlandes, Computerlinguistik und Phonetik
Verbmobil
- Siegen: Universität GH Siegen, PB 3: Sprach- und
Literaturwissenschaften
- Stuttgart: Institut für maschinelle Sprachverarbeitung (an der Universität
Stuttgart)
Universität, Institut für Germanistik/Linguistik
Universität Stuttgart, Institut für Linguistik/Romanistik
Universität Stuttgart, Institut für maschinelle Sprachverarbeitung
- Trier: Universität Trier, Linguistische Datenverarbeitung Universität
- Tübingen: Tübingen, Seminar für Sprachwissenschaft Universität
- Würzburg: Würzburg, Institut für deutsche Philologie Gesellschaft für
- Wuppertal: Angewandte Linguistik (GAL e.V.) Universität Wuppertal,
Sprach- und Literaturwissenschaften

Deutscher Terminologie-Tag wertet LDV-Forum aus

Der *Deutsche Terminologie- Tag (DTT e. V)* gibt für seine Mitglieder einen Rundbrief heraus, in denen über Entwicklungen in den Bereichen Terminologie, Übersetzen und Dokumentation berichtet wird. Darüber hinaus wurde ein Literaturvermittlungsdienst eingerichtet, der terminologierelevante Literatur dokumentiert und beschafft und auf Anfrage Literaturrecherchen durchführt. Für beide Serviceleistungen werden in Zukunft das LDV-Forum im Hinblick auf einschlägige Beiträge ausgewertet.

Die Arbeit des Deutschen Terminologie-Tages hat zum Hintergrund, daß infolge der rasanten wissenschaftlich-technischen Entwicklungen der jüngsten Zeit die fachlichen Kommunikationsprobleme immer größer werden. Aufgrund der wachsenden Spezialisierung wird es nicht nur immer schwieriger, Expertenwissen für Laien zugänglich zu machen, auch die Verständigungsschwierigkeiten von Experten untereinander nehmen zu. So braucht z.B. die technische Fachsprache Millionen von Fachausdrücken, um die Flut der technischen Neuerungen in all ihren Komponenten sprachlich fassen zu können. Die Neuprägung solcher Fach-

LDV-Forum Bd.14, Nr. 1, Jg. 1997

ausdrücke läuft dabei meist völlig unkoordiniert ab, was in einzelnen Bereichen bereits zu einem regelrechten Begriffswirrwarr geführt hat.

Die Probleme nehmen noch einmal ganz erheblich zu, wenn an der technischen Kommunikation mehrere Sprachen beteiligt sind. Übersetzte Fachtexte sind häufig noch unverständlicher als die Originaltexte; Fachwörterbücher hinken um Jahre hinter der technischen Entwicklung her; unterschiedliche Normen in den verschiedenen Ländern sorgen für weitere Schwierigkeiten. Die Mehrsprachigkeit hat durch die letzte Ausweitung insbesondere im Norden weiter zugenommen (elf Amtssprachen bedeuten 110 mögliche Sprachenpaare); der Beitritt weiterer Länder mit Schwerpunkt in Mittel- und Osteuropa ist nur eine Frage der Zeit. Politik, Wirtschaft und Verwaltung arbeiten heute weltweit, und alle sind auf eindeutige Kommunikation angewiesen.

Eine Verstärkung der Zusammenarbeit - innerhalb der Europäischen Union, aber auch darüber hinaus zwischen allen Bereichen, in denen Terminologiearbeit geleistet wird, ist erkennbar erforderlich. Noch immer wird sehr viel kostspielige Doppelarbeit geleistet mit der Folge, daß wichtige Aufgaben nicht in Angriff genommen werden können, weil es an Zeit und Geld fehlt.

Der Deutsche Terminologie-Tag e. V. sieht es als seine Aufgabe an, diese Probleme bewußt zu machen und Wege zu verbesserter Kommunikation durch Terminologiearbeit aufzuzeigen. Er tritt in öffentlichen Veranstaltungen und - von Fall zu Fall - in Form von Expertengruppen in Erscheinung und versteht sich als ständiges Forum zur Behandlung terminologischer Belange, das Einzelpersonen, Firmen und staatliche Stellen im In- und Ausland im Bereich der Terminologie durch Beratung und Koordination unterstützt.

Der DTI e.V. arbeitet eng mit anderen auf dem Gebiet der Terminologie tätigen Institutionen zusammen, so z.B. im Rat für Deutschsprachige Terminologie (RaDT), einer Initiative der Deutschen, Österreichischen und Schweizerischen UNESCO-Kommissionen. Weiter ist nunmehr der DTI e.V. für den deutschsprachigen Raum im Vorstand von Infoterm vertreten, in dem weltweit die Interessen der Terminologie wahrgenommen und gefördert werden. So hat der DTI e. V. seit 1989 insgesamt 5 Symposien veranstaltet. Das bislang letzte Symposium "Terminologie für ein vielsprachiges Europa" fand im April 1996 in Köln statt und hatte folgende zentrale Themen: *die europäische Sprachenpolitik, mehrsprachige Terminologiearbeit, der Einfluß des Englischen auf die Terminologie anderer europäischer Sprachen sowie Möglichkeiten und Grenzen praktischer Terminologiearbeit.* Das nächste DTI-Symposion ist für 1998 geplant.

Über die Symposien hinaus veranstaltet der DTI e.V. - vorwiegend für seine Mitglieder Workshops und Seminare und arrangiert Besichtigungen terminologierelevanter Einrichtungen.

Ganz aktuell arbeitet der Deutsche Terminologie-Tag mit Infoterm (Internationales Informationszentrum für Terminologie, Wien) zusammen, um im Rahmen des EU-Programms MLIS (Multi-Lingual Information Society) den gegenwärtigen Stand der Terminologearbeit in Wirtschaft und Industrie, Forschung und Entwicklung sowie in der öffentlichen Verwaltung in der Bundesrepublik Deutschland zu erheben.

Der Deutsche Terminologie-Tag e. V. ist auch im World Wide Web zu finden. Die Adresse lautet: <http://info.uibk.ac.at/c/c6/c613/dtt.html>.

Christa Womser-Hacker - Habilitation im Fachgebiet Informationswissenschaft

Seit dem 19. Februar hat Frau Dr. Christa Womser-Hacker mit so mancher Kollegin und manchem Kollegen so manches Glas Sekt geleert: An diesem Tag nämlich stand ein besonderer Vortrag auf der Tagesordnung: *"Interaktions- und Visualisierungsformen vager Anfragen für das Faktenretrieval"*. Eingeladen hatte die Habilkommission der Philosophischen Fakultät IV der Universität Regensburg, und um es kurz zu machen: Der Vortrag war der gute Abschluß einer guten Arbeit!

Die Habilitationsschrift selbst trägt den Titel: *"Das Mimor-Modell. Mehrfachindexierung zur dynamischen Methoden-Objekt-Relationierung im Information Retrieval"*. Kern der Arbeit ist, dem Titel gemäß, das sog. Mimor-Modell, das auf empirischen Erkenntnissen aus der Information-Retrieval-Forschung basiert. Die Mimor-Komponente ist als Zwischenschicht in ein Information Retrieval-System eingliederbar und verfolgt das Ziel eines systeminternen, während des Einsatzes stattfindenden Selbsttests, der mit einer dynamischen Zuordnung von informationellen Objekten und Repräsentationsverfahren einhergeht. Mimor bezieht neben Texten auch verschiedenartige Faktendaten ein und setzt mehrfach repräsentierte Informationseinheiten (z.B. verschiedenartig gewichtete oder manuell erzeugte Terme) und Relevance Feedback voraus. Die Gesamteffektivität wird intern unter Rückbezug auf Benutzerurteile gemessen und unter Berücksichtigung des Kontexts (z.B. der Art der Objekte) auf die parallel angewandten Verfahren verteilt. Das System lernt die Zusammenhänge zwischen Repräsentationsverfahren und Informationseinheiten, was möglicherweise zur Aussteuerung bestimmter Verfahren und zur dynamischen Annäherung an eine optimale Methoden-Objekt-Konfiguration für ein bestimmtes Anwendungsfeld führt.

Das LDV-Forum gratuliert der frisch habilitierten Informationswissenschaftlerin ganz herzlich!

Rainer Hammwöhner mit Venia Informationswissenschaft

Am 20.11.1996 hat Dr. Rainer Hammwöhner von der Konstanzer Informationswissenschaft sein Habilitationsverfahren erfolgreich abgeschlossen. Die Fakultät hat ihm die Venia im Fach Informationswissenschaft verliehen. Herzliche Gratulation!

Dr. Hammwoehner, PD vertritt im WS 96/97 eine Professur an der Universität Regensburg.