

der gedruckten Volltexte können ab Februar 1985 durchgeführt werden. Der Zugriff über Bildschirmtext wird geprüft. Im Verlauf des Jahres 1985 werden Aufbau und Betrieb einer Textdatenbank einschließlich der Online-Nutzung digitalisierter und umgesetzter graphischer Informationen sowie teilweise von Volltexten der Patentschriften erfolgen. Die Pilotdatenbank wird in deutscher Sprache in STN-INTERNATIONAL – dem zur Zeit vom FIZ-Karlsruhe zusammen mit Chemical Abstracts Service aufgebautem Informations- und Rechnerverbund für Wissenschaft und Technik – einbezogen werden.

Dieser erste internationale Rechnerverbund im Bereich des Online-Fachinformationstransfers, sprich die Kooperation zwischen dem FIZ-Karlsruhe und der American Chemical Society im STN-Verbund (The Scientific Technical Information Network) bietet für die Nutzer hier, als auch jenseits des großen Teiches enorme Vorteile. So ermöglicht dieser Rechnerverbund, daß neben dem bibliographischen und dem Indexteil der Chemical Abstracts jetzt auch über den Karlsruher Knoten des Netzes der bisher ausschließlich von Columbus (Ohio) angebotene, zur Struktursuche in der Organischen Chemie verfügbare CAS-Online Informationsdienst des Chemical Abstracts Service anwählbar ist.

Weiterhin ist die Recherche in allen über den STN-Verbund angebotenen Datenbanken zum jeweils kostengünstigen nationalen Posttarif und in einer einheitlichen Abfragesprache möglich (die notwendigen Umstellungen werden sukzessive vorgenommen). Für die Suche in den unterschiedlichen Datenbanken ist ein „file-switching“, ohne daß das System verlassen werden muß, durchführbar. Erfreulich ist auch die Tatsache, daß die gewünschten Offline-Prints jeweils vom Eingangshost, für deutsche Nutzer also von Karlsruhe aus, versandt werden und so – falls die Post keinen Strich durch die Rechnung macht – bereits am nächsten Tag verfügbar sind. Da STN-INTERNATIONAL allen interessierten Partnern offensteht, ist bald mit weiteren Anbietern (aus Japan) im System zu rechnen.

Volltextdatenbanken, in der deutschen Datenbanklandschaft noch nicht häufig anzutreffen, bekommen in den USA ein immer stärkeres Gewicht. Interessant ist es überhaupt, die Entwicklung in dem „Online-Geburtsland“ USA zu betrachten. Das Online-Retrieval, eine Technik die noch keine zwanzig Jahre zählt, hat die Informationslandschaft in erheblichem Umfang verändert, jedoch nicht in dem Ausmaße wie es stets prognostiziert wurde; die Zuwachsraten verzeichnen längst nicht mehr die Anstiege der vorangegangenen Jahre. In den USA ist die Benutzerdichte rund 50mal größer als in der Bundesrepublik Deutschland und damit stellt sich die Frage: Liegt es an der Innovationsfreude der Amerikaner, den anderen Kommunikationsformen (auch bedingt durch große geographische Entfernungen, nicht so ausgeprägten Dokumentationseinrichtungen, weiter entwickelteren Informationstechnologien...), dem weitaus größeren und damit attraktiveren englischsprachigen Markt (keine Sprachbarrieren), dem professionellen Marketing oder anderen Gründen? Beachtlich ist auch die Tatsache, daß die Zahl der Nutzer im Bereich der Wirtschaftsinformationen bei weitem die Anwenderzahl im naturwissenschaftlichen und Technikbereich überwiegt.

Aber nicht nur in den USA, auch im benachbarten Ausland ist die Informationsbeschaffung via Datenbanken nicht so problematisch wie in einheimischen Gefilden. Kann das Datenbankangebot in Naturwissenschaft und Technik als recht gut und umfassend angesehen werden, wobei Verbesserungen immer noch möglich sind, so kristallisiert sich aber klar heraus, daß im Bereich der Wirtschaftsinformationen (incl. Recht) (siehe hier „Schumacher, Bericht zur IuD-Lage der Nation im geteilten FIZ-Deutschland, 1984“) enorme Lücken

vorhanden sind und diese nur annähernd vom ausländischen Angebot her geschlossen werden können. Aber auch hier ist ein – zwar sehr schmaler – Silberstreif am Horizont erkennbar, wenn man die Datenbankaktivitäten von Hoppenstedt, dem FIZ-Technik, I. P. Sharp und der Gesellschaft für betriebswirtschaftliche Information (GBI) betrachtet.

Wenden wir uns aber jetzt dem Online-Alltag zu. Jedes Rechercheergebnis ist eine Funktion des Datenbankinhaltes und der Qualifikation des Rechercheurs. Unter dem Gesichtspunkt der ständig zunehmenden Bedeutung des Produktionsfaktors „Information“ muß daher dem Stellenwert der Datenbankerstellung und der Informationsvermittlung eine weitaus größere Beachtung – was seinen Niederschlag auch im neuen Fachinformationsprogramm finden sollte – beigemessen werden. Das gilt ebenso für das Berufsbild des Informationsvermittlers bzw. Informationsbrokers. Eine durchaus reizvolle Aufgabe, die sich hier allen, die in der Aus- und Weiterbildung von Onlinern oder Informationsvermittlern tätig sind, stellt. Unterschiedliche Inhalte und Formen der Datenbanken, die verschiedenen Retrievalsysteme und das rasche Vordringen neuer Technologien (z. B. PCs) werden auch zukünftig hohe Anforderungen an diesen Berufsstand stellen.

Aber nicht nur Theorie, besonders die Praxis wurde in Neuisenburg ganz groß geschrieben. Neueste Hard- und Software wurde auf der Ausstellung präsentiert. Somit konnten die Besucher einen genauen Überblick über das heutige und in näherer Zukunft verfügbare Angebot an Datenbanken sowie Abruf-, Speicher- und Weiterverarbeitungsmöglichkeiten für Informationen erhalten. Die GID präsentierte die für das neue IuD-Terminal, das die alte IuD-Dialogstation der Firma Nixdorf ablöst, ausgewählte Hard- und Software. Die eventuell noch offenen Fragen der Onliner konnten in der Diskussion mit den Ausstellern direkt vor Ort am Terminal geklärt werden. Weitere detaillierte Informationen erfuhren die Interessenten während der Product Reviews, so etwa, daß die Quo Vadis-Datenbank Seminare der GID, die u. a. die Termine sämtlicher Online-Trainingskurse in der Bundesrepublik Deutschland enthält, nun auch online und über Btx allgemein zugänglich ist. Trotz des kompakten Tagungsplanes bot sich natürlich auch die Möglichkeit des Gedankenaustausches im Gespräch mit dem Fachkollegen.

Anzumerken ist noch, daß die erstmals vom Lenkungsanschluß veranstaltete Vortragsreihe „Online für Neulinge“ regen Zuspruch fand. Abschließend kann man allen Onlinern und Online-Interessierten nur raten, die Zeit vom 5. bis 7. 3. 1985 im Terminkalender dick anzustreichen – dann findet nämlich die 7. Frühjahrstagung der Onliner statt.

Dipl.-Ing. Brigitte Hennemann-Böckels, AZ Direct Marketing Bertelsmann GmbH, Carl-Bertelsmann-Str. 161, 4830 Gütersloh 1.

*Maschinelle Übersetzung; Computerlinguistik; Künstliche Intelligenz; Übersichtsbericht; Tagung*

## **Verfahren und Systeme der Automatischen Übersetzung**

**„Tutorial on Machine Translation“ vom 2. bis 6. April 1984 in Lugano\***

### **Allgemeiner Überblick**

Zweck der Veranstaltung war es, einen Überblick über die Entwicklung der Automatischen Übersetzung (im folgenden

\* Das Tutorial wurde organisiert vom Istituto per gli Studi Semantici e Cognitivi (ISSCO). (Institut für semantische und kognitive Studienforschung.)

AÜ) und den gegenwärtigen Stand, vielleicht ergänzt um einige Zukunftsperspektiven, zu geben. Verhältnismäßig viel Raum wurde dabei der historischen Perspektive eingeräumt, möglicherweise auch aus der Einschätzung heraus, daß die meisten wesentlichen Prinzipien der AÜ schon in der Frühzeit der Computeranwendung entwickelt worden seien. Dieser historische Überblick sollte damit vielleicht ein Ersatz für die fehlende systematische Einführung sein, die, unabhängig von früheren oder gegenwärtigen Ansätzen, auf einer Theorie der AÜ basieren müßte. Diese Theorie fehlt aber ganz offensichtlich. Ja nicht einmal besteht Einigkeit darüber, ob eine solche einheitliche Theorie überhaupt möglich oder sinnvoll sei und ob sie sich an der humanen Übersetzungsleistung orientieren müsse oder nicht. So war die Veranstaltung letztlich eher eine Reihe von Einzelvorträgen zu Einzelprojekten, die alle für sich natürlich sehr interessant waren, deren Darstellung dadurch aber recht redundanzlastig wurde. (Mit zwei Ausnahmen wurden zu allen Vorträgen schriftliche Fassungen vorgelegt.)

Wenn man davon absieht, daß aus Kostengründen kein japanischer Vertreter eingeladen wurde, so daß die dortige, zweifellos wichtige und intensive Entwicklung (am Ende des Vortrages von E. Ananiadou und S. Warwick wurde darauf verwiesen, daß es z. Z. in Japan 18 Projekte zur AÜ gibt) ausgeblendet blieb, so kann man doch annehmen, daß die Veranstaltung einen repräsentativen Einblick gegeben hat. Es war daher möglich, sich einen Überblick über den „Stand der Kunst“ zu verschaffen.

Im einzelnen wurden an Systemen vorgestellt: SYSTRAN (P. Wheeler), SUSY (D. Maas), TAUM/METEO und TAUM/AVIATION (P. Isabelle); eine deutsch-französische Version von ARIANE-78/GETA (J. P. Guilbaud), ROSSETA (Forschungsvorhaben von Philips, J. Landsbergen); METAL (J. Slocum); EUROTRA (M. King/S. Perschke) und GETA allgemein (Ch. Boitet).

Ergänzt wurden diese Systemdarstellungen, zusätzlich zu einem Versuch eines Einführungsvortrages von G. Sampson und den beiden historischen Beiträgen von B. Buchmann und E. Ananiadou/S. Warwick, durch zwei Vorträge zu Software-Aspekten der AÜ, davon der eine (J. L. Cocharde/D. Petitpierre) sehr einführend und der andere sehr hoch, also sehr wenig detailliert, angelegt war (R. Johnson, M. Rosner), zwei Vorträgen zu linguistischen Entwicklungen (E. Wehrli und H. A. De Roeck), beide sehr elementar und keineswegs auf dem neusten Stand, zwei Vorträgen zu Künstliche-Intelligenz-Ansätzen (der von P. Shann geplant, der von Nirenburg zusätzlich angeboten) und einem Vortrag von A. Melby über eine „translator workstation“, realisiert auf einem IBM-PC. Insgesamt war das Programmangebot also reichhaltig.

### Versuch einer generellen Einschätzung der AÜ-Situation

Sicherlich konzentriert sich die fachliche und fachpolitische Welt der AÜ zur Zeit auf die Erwartungen bezüglich EUROTRA. Nicht zuletzt deshalb ist das Tutorial vermutlich auch organisiert worden. Eine fachlich fundierte Bewertung müßte mit viel Sorgfalt und Aufwand betrieben werden. Dies ist in dem dafür erforderlichen Ausmaß hier nicht möglich. Die folgenden Bemerkungen zum „state of the art“ sind also eher impressionistisch und sicherlich subjektiv.

#### 1.

Es existieren bislang kaum Erfahrungen mit multilingualen Systemen. Der ursprüngliche Forschungsansatz, Entwick-

lung einer universalen Interlingua, wird nur in einer sehr langfristigen Perspektive verfolgt. Es bestehen offenbar einige Defizite in der Entwicklung einer Theorie der Transfer-Komponente für multilinguale Systeme (z. B. in welcher Form können Strukturbeschreibungen und ambigue lexikalische Einträge transferiert werden?).

#### 2.

Realistische Erfolge lassen sich offenbar durch Beschränkungen auf Fachgebiete technischer Art erzielen (Beispiele: TAUM-METEO, METAL). Dabei scheinen sich auch unter Akzeptanzgesichtspunkten AÜ-Systeme in Humanübersetzungs-Umgebungen einbetten zu können. Eine Fidelity-Quote von 80% scheint realistisch zu sein. Post-Editing ist, wie auch bei Humanübersetzungen, unverzichtbar. AÜ-Systeme können sinnvoll für – bei Humanübersetzern nicht beliebte – Routinetexte, aber auch für solche mit verhältnismäßig beschränkter Syntax, aber mit reichem Fachvokabular und -jargon eingesetzt werden. Der Bedarf zielt bei diesen Fachtexten auf qualitativ hohe und fachlich zuverlässige Übersetzungen ab. Das Relevanz-Kriterium („lohnt es sich, den Originaltext übersetzen zu lassen?“) reicht offenbar nicht aus. Das Konzept der vorstrukturierten Sprache (sub-language wie bei TITUS) wird nur gelegentlich verfolgt (XEROX arbeitet an einer solchen sub-language, um ausländischen Kunden die Lektüre englischer Fachtexte, z. B. Manuals, zu erleichtern). In der Regel werden Texte so akzeptiert, wie sie erstellt werden (vgl. allerdings Punkt 5).

#### 3.

AÜ-Systeme stützen sich auf elaborierte Wörterbücher ab, die von leistungsstarken Datenbanksystemen verwaltet werden müssen. Bezüglich der Qualität der Einträge zeichnet sich ein Methodenmix ab, d. h. es wird das genommen, was sich nicht gerade ausschließt. So vertragen sich offenbar ohne Probleme strukturalistische Informationen über Flexions- und Wortklassen mit Valenzrahmen und Subkategorisierungsregeln. Eine einheitliche Wörterbuch-Theorie ist nicht in Sicht, vielleicht auch nicht möglich. Ein ähnlicher Methodenmix ist bezüglich der Syntax festzustellen. Weitgehend akzeptiert ist die Vorstellung, daß „large-scale“-Systeme mit leistungsstarken Syntax-Grammatiken ausgestattet sein müssen, da semantik-gesteuerte Verfahren auf einer tiefen Ebene zum einen zu aufwendig seien und zum anderen keine einfachen natürlichsprachlichen Synthesen ermöglichen. Auffallend ist, daß schon fast „totgesagte“ einfache Phrasenstrukturgrammatiken offenbar eine AÜ-Renaissance erleben. Daneben werden aber auch andere Ansätze (Dependenz, aber auch Syntaxen auf Montague-Grammatik-Basis) verfolgt.

#### 4.

Parsing-Techniken sind so leistungsstark geworden, daß von dieser Seite her keine Restriktionen zu erwarten sind (Zielsetzung 1 s/Wort ist realistisch). Es zeichnet sich ab, daß die entsprechende Software so flexibel aufgebaut werden muß, daß unterschiedliche Strategien (single path, some/all paths, bottom-up, top down, seriell/parallel) bei unterschiedlichen Aufgaben verfolgt werden können. Es ist eindeutig, daß heutige AÜ-Systeme auf jeden Fall auf dem heute möglichen Software-Engineering-Niveau geplant werden müssen. Hierdurch wird gegenüber früheren Systemen, wie z. B. SYSTRAN, vor allem die ständig nötige Weiterentwicklung der Systeme begünstigt. Eine flexible Software ist vor allem auch deshalb wichtig, weil nach wie vor Übersetzungssysteme in erster

Linie als linguistische Systeme geplant werden, die von Linguisten entworfen werden und die deshalb besonders benutzungsfreundlich sein müssen.

5.

Unabhängig davon, ob AÜ-Systeme als interaktive oder Batch-Systeme zu konzipieren sind (für beides scheint es gute Argumente zu geben), sollten AÜ-Systeme auf den Stand der Textverarbeitung (hier im nicht-linguistischen Sinne gemeint) zurückgreifen können, ebenso auf fortgeschrittenen Terminal-Hardware-Komfort (Graphik, Mouse-Steuerung, Window-Technik usw.). Dabei scheint die Orientierung an Arbeitsplatzrechnern in der direkten Umgebung von Humanübersetzern zweckmäßig zu sein.

6.

Auffallend ist, daß die AÜ-Forschungsgruppen nach wie vor von den Forschungsgruppen der Computerlinguistik, Informationslinguistik oder Künstlichen-Intelligenz (KI) relativ getrennt arbeiten. Obwohl durchgängig die Bereitschaft gezeigt wird, auf neuere Entwicklungen der Linguistik oder KI oder der kognitiven Psychologie einzugehen, besteht doch die Gefahr einer gewissen Abkapselung, verbunden mit einer Tendenz zu einer pragmatizistischen Grundeinschätzung (Performanz wichtiger als Theorie). Da offenbar aber noch erhebliche theoretische Defizite in der Übersetzungstheorie bestehen, sollte zumindest zweigleisig (Grundlagenforschung, Entwicklungsforschung) verfahren werden. Die früh-

schnelle Festlegung auf Anwendungsziele ist bei zu ambitionierter Erwartung gefährlich (Beispiel TAUM-AVIATION).

7.

Für ein mittelfristig angelegtes AÜ-System sind ein effektives Management, koordinierte Organisation und evaluierende Kontrollbegleitung zentral. Das linguistische und softwaremäßige Know-how dürfte – bei realistischer Erwartung (siehe vor allem Punkte 2 und 5) – für die Entwicklung eines Systems der zweiten Generation prinzipiell vorhanden sein (man weiß, wie Wörterbücher auszusehen haben und wie sie verwaltet werden, wie Grammatiken aufgebaut werden und wie sie zu parsen sind u. a.). In mittlerer und erst recht in längerer Perspektive werden aber wohl auch Übersetzungssysteme wissensbasiert zu konzipieren und zu realisieren sein. In welcher Zeit diese wirklich als multilinguale Systeme in nichttrivialer (qualitativ und quantitativ gesehen) Umgebung realisiert werden – ob also z. B. EUROTRA zur Zeit seiner Fertigstellung Ende der achtziger Jahre nicht konzeptionell längst überholt ist –, ist kaum zu entscheiden. Die Forschung in Europa bewegt sich auf dem AÜ-Gebiet nur sehr zögerlich auf „intelligente“ Systeme zu.

Eine ausführliche Fassung des Berichts sowie die Manuskripte der Vorträge sind über das Informationszentrum der Gesellschaft für Information und Dokumentation (GID-IZ), Lyoner Str. 44-46, 6000 Frankfurt am Main 71, erhältlich.

Prof. Dr. Rainer Kuhlen, Universität Konstanz, FG Politik-/Verwaltungswissenschaft, Informationswissenschaft, Postfach 5560, 7750 Konstanz 1

### Datenbank für Betriebswirtschaft

Mit der deutschsprachigen Datenbank Belit (betriebswirtschaftliche Literaturdokumentation) soll eine weitere Lücke im Angebot an deutschsprachigen Wirtschaftsinformationen geschlossen werden. Diese Datenbank wird von der Gesellschaft für betriebswirtschaftliche Information (GBI) angeboten. Der Zugriff sei sowohl direkt im Computerverbund (online) als auch offline möglich. Zur Zeit enthält die Datenbank Belit über 70 000 Literaturhinweise. Geplant sei ein jährliches Wachstum um 10 000 Hinweise. Seit 1984 enthalten die gespeicherten Titel neben den bibliographischen Angaben und Schlagwörtern auch eine Zusammenfassung des ausgewerteten Artikels. Mehr als 450 europäische, davon allein mehr als 300 deutschsprachige Zeitschriften werden nach zehn Fachgebieten der Betriebswirtschaftslehre ausgewertet. Die Schwerpunkte reichen von Fragen zu Führung, Organisation und Planung, Finanzwirtschaft und Investition, über Materialwirtschaft und Logistik, Marketing und Handel hin zu Banken und Versicherungen beziehungsweise Dienstleistungen, Verwaltung und Büro. Belit soll vom Fachinformationszentrum Technik e.V. via Datex-P-Netz angeboten werden. (Blick durch die Wirtschaft, 18. April 1984)

### Betriebsstatistik für Fachinformationseinrichtungen

In der Reihe „Informationswissenschaft und -praxis“ der Gesellschaft für Information und Dokumentation mbH (GID) ist im August 1984 der 5. Band erschienen: Kuno Eustachi,

Erwin Gusel, Johann Lachhammer: Konzeption einer Betriebsstatistik für Fachinformationseinrichtungen. Frankfurt am Main: IDD Verlag, 1984. 320 Seiten. DM 48,-. ISBN 3-8151-0021-6.

Diese Studie ist das Ergebnis eines Forschungsprojekts, das Fachinformationseinrichtungen in die Lage versetzen soll, ihre Entscheidungen anhand statistischer Kennzahlen und/oder Auswertungen zielbezogen zu treffen und die Effizienz des Mitteleinsatzes zu steuern und zu kontrollieren. Der Band enthält einen theoretischen Teil, der der Erklärung und Fundierung des vorgeschlagenen Konzepts dient; ihm folgt der praktische Teil, der so aufgebaut ist, daß die Fachinformationseinrichtungen die entworfene Betriebsstatistik selbständig realisieren können.

### ISI-Award 1984 für Belkin, Seeger und Wersig

Dr. Nicklas J. Belkin (City University London), Priv.-Doz. Dr. Thomas Seeger und Prof. Dr. Gernot Wersig (beide Freie Universität Berlin) erhielten den ISI-Award 1984 für ihren Aufsatz „Distributed expert problem treatment as a model for information system analysis and design“ in: Journal of Information Science, Vol. 5, 1983, Nr. 5, S. 153-167. In diesem Artikel fassen die drei Autoren die Ergebnisse ihrer gemeinsamen Überlegungen zur „intellektuellen Architektur“ von Informationssystemen neuen Zuschnitts zusammen. Der ISI-Award ist ein vom Institute for Scientific Information (ISI) in Philadelphia, USA, gestifteter Preis für den besten Zeitschriftenaufsatz.