All topics are discussed in such a way that it is possible to receive benefit from the reading of the essays without too much of theoretical background.

Rolf G. Henzler

SEELBACH, Dieter: **Computerlinguistik und Dokumentation.** Key-Phrases in Dokumentationsprozessen. = UTB Uni-Taschenbücher 468. München: Verlag Dokumentation 1975. 151 p. ISBN 3-7940-2643-8

Gibt es hier eine Antwort auf die Frage, ob Ergebnisse linguistischer Forschungen gegenwärtig eine Chance haben, in Anwendungsbereiche der elektronischen Datenverarbeitung einzudringen, also über ein experimentelles Stadium hinauszukommen? Die Dokumentation, hier vor allem der Problemkreis der Dokumenterschließung (Indexing), ist bestimmt durch einen hohen Zeit- und Kostenaufwand für die intellektuelle/manuelle Dokumentdeskription. Linguistische Verfahren sollen daher eine Automatisierung unterstützen oder gar erst ermöglichen. (Ob allerdings die Kosten sich reduzieren lassen, ist noch offen).

Der Verf. stellt ein Verfahren vor, durch das die noch wenig zureichenden bisherigen Methoden – meist auf der Vergabe von Einzel-Schlüsselwörtern (= Key-Words) beruhend – durch die automatische Zuweisung mehrgliedriger Nominalsyntagmen (= Key-Phrases) präzisiert und damit die Retrieval-Ergebnisse verbessert werden sollen. Maschinell erkennbare Text- (Satz-) Segmente werden automatisch in standardiserte Paraphrasen transformiert, d. h. in die Form komplexer, i. a. mehrere Präpositionen und Nomina enthaltender Nominalgruppen (z. B. CONSUMPTION OF BEER IN YUGOSLAVIA) überführt. Die Verwendung der expliziteren Präpositionalgruppenform soll dabei Mehrdeutigkeiten vermeiden helfen:

WATER TREATMENT → TREATMENT OF/WITH WATER

Trotz der vielfältigen Analyseprozeduren – durch eine Reihe von Tabellen und Listen ist die große Praxisnähe der Untersuchung dokumentiert – wird eine günstige Kosten-Nutzen-Relation angestrebt. Gemessen an der geringen Tiefe linguistischer Operationen (es wird mit Wortbildungsregeln und syntaktischen Oberflächen-Transformationen/Permutationen Harris'scher Art gearbeitet) erscheint die Effizienz beachtlich.

Angestrebt ist neben einer Verbesserung der (automatischen) Deskriptorenvergabe (Schlüsselwortvergabe) zu Dokumenten mithilfe der syntagmatischen Verknüpfung von Einzeldeskriptoren die Herstellung benutzerfreundlicher (gedruckter) Register (Verbesserung von KWIC- und KWOC-Verfahren). Dabei werden im Text vorkommende Begriffe nach einer morphologischen Suffix-Analyse auch semantisch grob nach „Vorgang" und „Eigenschaft" zu klassifizieren versucht. Auf eine Auseinandersetzung mit ähnlichen, wenn auch eher intellektuell orientierten Verfahren in der Dokumentation (und entsprechenden Literatur) wird aber weitgehend verzichtet. In einem Ausblick wird auf eine mögliche Anwendung des vorgestellten Regelsystems auf die maschinelle Übersetzung von Key-Phrases hingewiesen, die die noch halbintellektuellen Abstract-Übersetzungen nach Art des TITUS-Systems im Bereich der Key-Phrases ablösen helfen sollen.

Wenn auch das spezifische Thema nur als Exemplum für die Problematik der Automatisierung von Dokumentationsprozessen gelten kann, wenn auch der mittlerweile weitgespannte Bereich der Computerlinguistik (von der Wortformenreduktion bis zu sprachlichen Aspekten der künstlichen Intelligenz) hier nur ausschnitthaft (vorwiegend syntaxbezogen) angesprochen ist: die Erläuterungen zu den verschiedensten Bereichen in Linguistik, Datenverarbeitung und Dokumentation in einer gewissen Zusammenschau rechtfertigen vielleicht den allgemeinen Obertitel, wenn es auch nicht beabsichtigt erscheint, durch diese Hinweise eine grundlegende Einführung in den Bereich der Computerlinguistik und Dokumentation zu ersetzen.

Fazit: Ausführliches (Tabellen, Listen, Regelsysteme) Material, praxisnahe Resultate, weiterführende Vorschläge, die weitgehend realistisch, d. h. anwendbar scheinen. Doch noch keine schlüssige Antwort auf die Ausgangsfrage: es fehlen die „harten" Zahlen einer Effizienz-Untersuchung. Welcher Anwender nutzt dies Möglichkeiten aus?

Harald Zimmermann

INTERNATIONAL FEDERATION FOR DOCUMENTATION: **Cumulated UDC Supplement 1965–1975. Vols. I–V.** The Hague, Netherlands: FID 1976. 548 p., hfl. 250,– (each vol. hfl. 50) = FID 543; ISBN 92-66-00543-6 (for the 5-volume set).

In addition to the P-notes, containing the proposals for new or revised UDC-notations, and issued in about 30–40 notes each year since 1931, the FID Classification Department issues "Extensions and Corrections to the UDC" ever since 1951 in a semi-annual mode and cumulates these every three years so that the final issue of a three-year-period replaces the foregoing 5 ones.

The cumulation under review contains all the alterations of the last 10 years, that is the cumulations from the Series 6, 7, 8 and No. 1 of Series 9 of the Extensions and Corrections to the UDC. In 1967 a similar 'Cumulated UDC Supplement" had been issued in 6 volumes, 795 pp., covering the changes from 1954 until 1964. Thus, the present volumes are the natural continuation of this practice.

The coverage is as follows: Vol. I: Auxiliaries + Classes 0/3; Vol. II: Classes 5; Vol. III: Classes 6/62; Vol. IV: Classes 63–676 and Vol. V: Classes 677/9.

Each volume can be obtained separately. The language is either English or German or French, depending on the language of the proposals accepted. Thus, very much effort has still to be put into the translation of the texts if new full editions in different languages are to be developed from this basis. From 1976 on the Extensions and Corrections to the UDC will only be published once a year.

It seems to be impossible to review any of the contents of the new tables, although the new numbers for "Classification and Indexing" (p. 70 of vol. I) (025.4.01/.49) which include also "Information retrieval techniques" (025.4.03) could be a very adequate "victim".

https://doi.org/10.5771/0943-7444-1976-1-45 https://www.inlibra.com/de/page/agb - Open Access - (cc) BY