

B 6 Informationsaufbereitung III: Referieren (Abstracts – Abstracting – Grundlagen)

Rainer Kuhlen

B 6.1 Grundbegriffe

B 6.1.1 Abstracts

Was ein **Abstract** (im folgenden synonym mit Referat oder Kurzreferat gebraucht) ist, legt das *American National Standards Institute* in einer Weise fest, die sicherlich von den meisten Fachleuten akzeptiert werden kann: „An abstract is defined as an abbreviated, accurate representation of the contents of a document“ (Lit. 03, S. 19); fast genauso die deutsche Norm DIN 1426 (Lit. 21.): „Das Kurzreferat gibt kurz und klar den Inhalt des Dokuments wieder.“ Abstracts gehören zum wissenschaftlichen Alltag. Weitgehend allen Publikationen, zumindest in den naturwissenschaftlichen, technischen, informationsbezogenen oder medizinischen Bereichen, gehen Abstracts voran, „preferably prepared by its author(s) for publication with it“ (Lit. 03, S. 19). Es gibt wohl keinen Wissenschaftler, der nicht irgendwann einmal ein Abstract geschrieben hätte. Gehört das Erstellen von Abstracts dann überhaupt zur dokumentarischen bzw. informationswissenschaftlichen Methodenlehre, wenn es jeder kann? Was macht den informationellen Mehrwert aus, der durch Expertenreferate gegenüber Laienreferaten erzeugt wird? Dies ist nicht so leicht zu beantworten, zumal geeignete Bewertungsverfahren fehlen, die Qualität von Abstracts vergleichend „objektiv“ zu messen. Abstracts werden in erheblichem Umfang von Informationsspezialisten erstellt, oft unter der Annahme, dass Autoren selber dafür weniger geeignet sind. Vergewöhnen wir uns, was wir über Abstracts und Abstracting wissen.

Ein besonders gelungenes Abstract ist zuweilen klarer als der Ursprungstext selber, darf aber nicht mehr Information als dieser enthalten: „Good abstracts are highly structured, concise, and coherent, and are the result of a thorough analysis of the content of the abstracted materials. Abstracts may be more readable than the basis documents, but because of size constraints they rarely equal and never surpass the information content of the basic document“ (Lit. 15, S. 3). Dies ist verständlich, denn ein „Abstract“ ist zunächst nichts anderes als ein Ergebnis des Vorgangs einer Abstraktion. Ohne uns zu sehr in die philosophischen Hintergründe

der Abstraktion zu verlieren, besteht diese doch „in der Vernachlässigung von bestimmten Vorstellungsbzw. Begriffsinhalten, von welchen zugunsten anderer Teilinhalte abgesehen, ‚abstrahiert‘ wird. Sie ist stets verbunden mit einer Fixierung von (interessierenden) Merkmalen durch die aktive Aufmerksamkeit, die unter einem bestimmten pragmatischen Gesichtspunkt als ‚wesentlich‘ für einen vorgestellten bzw. für einen unter einen Begriff fallenden Gegenstand (oder eine Mehrheit von Gegenständen) betrachtet werden“ (Lit. 02). Abstracts reduzieren weniger Begriffsinhalte, sondern Texte bezüglich ihres proportionalen Gehaltes. Borko/Bernier haben dies sogar quantifiziert; sie schätzen den Reduktionsfaktor auf 1:10 bis 1:12 (Lit. 10, S. 5). Wir wollen für das Folgende aus dieser Eingangsdiskussion zweierlei festhalten:

- a) Die Reduktion (die Abstraktion) kann nicht objektiv sein; wäre sie es, dann müsste es im Durchschnitt einen etwa zehnpromzentigen Bedeutungskern in Texten (nehmen wir diese einmal als primäre Referenz von Abstracts) geben, der in 90% Redundanz eingebettet wäre. Das aber ist nicht sehr plausibel. Wenn man Abstracts schreibt, selektiert man. Der Philosoph oben erwähnte den „pragmatischen Gesichtspunkt“, der die Abstraktion bzw. die Selektion steuert. Dokumentare würden vom Benutzerinteresse sprechen. Diese Forderung nach einem pragmatischen Primat (vgl. Lit. 62) bei der Erstellung von Abstracts ist allerdings bislang kaum einlösbar. Abstracts sind in der Regel vorfabrizierte, d.h. auf ein Fachgebiet und auf einen weitgehend anonymen Benutzer hin erstellte Abstraktionen. Ein Gegenbeispiel findet sich bei Borko/Bernier (Lit. 10, S. 18 ff.), in dem einem Text drei verschiedene „special-purpose“ Abstracts aus medizinischer, biologischer und chemischer Sicht zugeteilt wurden; aber dies ist eher eine Ausnahme und löst auch nicht das Anonymitätsproblem (vgl. allerdings den Typ des „slanted abstracts“, siehe Abschnitt B 6.2.2).
- b) Durch die unter a) angedeuteten „Abstraktionsleistungen“ wird ersichtlich, dass das einmalige, in der Regel vertextete Abstract keine quasi naturgegebene Form der informationellen Ab-

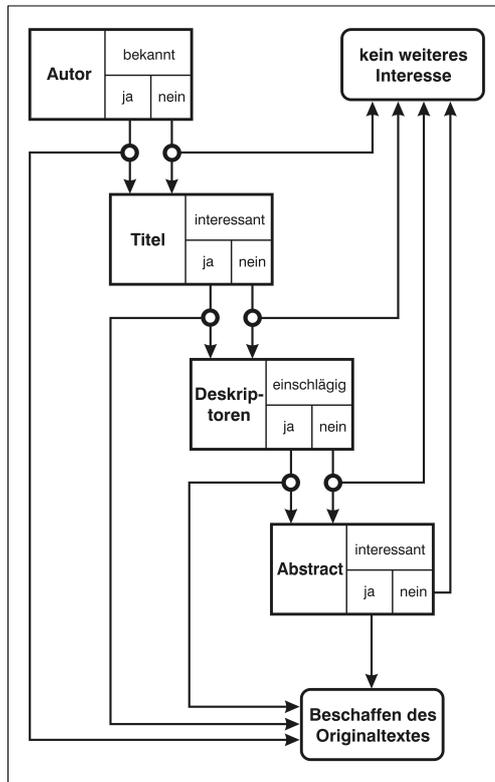


Abb. 1: Referenzielle Textakte

straktion sein muss. DIN 1426 bleibt bei den Texten und unterscheidet zwischen Inhaltsverzeichnis, Auszug, Zusammenfassung, Annotation, Kurzreferat (Abstract), Sammelreferat, Rezension bzw. Sammelrezension und Literaturbericht, während Borko/Bernier ihrer Phantasie freieren Spielraum lassen. Sie schlagen als „document surrogates“ vor: „abridgement, annotation, aphorism, axiom, brief, code, command, compendium, conclusion, databook, epitome, excerpt, extract, maxim, precept, precis, resume, review, selection, summary, summation, synopsis, and terse conclusion“ (Lit. 10, S. 5). Für diese Variabilität ist möglicherweise der neutralere, auch für multimediale und automatische Formen offene Begriff des Kondensierens bzw. des Kondensats geeignet (vgl. Lit. 60; Lit. 89). Automatisch erzeugte Abstracts bzw. Kondensate können variabler generiert werden und könnten sich daher flexibler an variable Benutzerbedürfnisse anpassen (Lit. 41; Lit. 44; Lit. 45; Lit. 83).

Bleiben wir aber bei den klassischen Abstracts. Was sind ihre wesentlichen Vorteile? Ihre Beliebtheit erklärt sich wohl dadurch, dass sie in ihrer Indikations- bzw. Referenzleistung, nämlich auf einschlägige Originaltexte hinzuführen und dabei die nicht-einschlägigen zu diskriminieren, anderen Referenzformen, z.B. Titeln von Dokumenten, eindeutig überlegen sind (vgl. Lit. 84). Abb. 1 beschreibt diese Leistungen als referenzielle Textakte. Abstracts haben per se in gewissem Umfang pragmatische Eigenschaften, sie fordern zum Handeln auf, entweder die Originaltexte zu ignorieren oder sie genauer zur Kenntnis zu nehmen („Abstracts assist readers in deciding whether they should consult the full text of the material that is abstracted“; Lit. 15, S. 4). Die Funktion von Abstracts erlischt, wenn Leser anfangen, sich mit dem Originaltext auseinander zu setzen.

Wir halten die Referenzfunktion für die wesentliche Eigenschaft von Abstracts. Zu einer Zeit, in der die Primärproduktion von Texten immer noch stark ansteigt und zunehmend mehr Volltexte in maschinellen Speichern aufbewahrt werden, in denen man nicht so leicht „browsen“ kann, sind dokumentarische Verfahren, die den Zugang zur Primärinformation offen halten, sehr wichtig. Borko/Bernier (Lit. 10, S. 6 ff.) führen insgesamt die folgenden Vorteile an:

- Abstracts ersparen Lesezeit. Borko/Bernier erwähnen erneut den durchschnittlichen Faktor von 1:10 Einsparung.
- Abstracts erlauben ein besseres Assimilieren und Integrieren des Gelesenen in die eigenen Wissensbestände, weil die Originaltexte zu umfänglich sind: „for putting words to work, condensations may be better vehicles than originals“ (Lit. 10, S. 7).
- Abstracts gestatten besser als andere Formen, wie z. B. Titel, die Selektion von relevanten Originaltexten (dies ist identisch mit der oben angeführten Referenzfunktion).
- Abstracts sind ein nützliches Hilfsmittel, Sprachbarrieren zu senken. Da viele der internationalen Referateorgane bzw. Datenbanken in englischer Sprache produziert werden, reichen einfache Englischkenntnisse aus, um zumindest (im Sinne der Referenzfunktion) die relevanten von den irrelevanten Originaltexten zu trennen und gegebenenfalls die ersteren übersetzen

zu lassen. Abstracts (darin vergleichbar Titeln) mit ihren in der Regel reduzierten syntaktischen Formen und beschränkten terminologischen Variationen sind auch schon bei dem heutigen Stand der Technik auf akzeptablem Niveau automatisch übersetzbar.

- e) Abstracts verbessern die Indexing-Qualität, sicherlich aber die Geschwindigkeit des Indexierens (Lit. 10; Faktor 2:4), falls – was in der Dokumentationspraxis häufig der Fall ist – nicht mehr auf den Originaltext zurückgegriffen wird (vgl. Lit. 67). In der Informationswissenschaft wird daher verschiedentlich das Erstellen von Abstracts als fundamentaler angesehen; in den meisten Gesamtdarstellungen wird zuerst das Abstracting beschrieben (z.B. Lit. 86). Auch automatische Verfahren des Indexing stützen sich häufig eher auf die Abstracts als auf die Volltexte ab. Dies ist zwar zunächst historisch bedingt (geringe Verfügbarkeit von maschinenverarbeitbaren Volltexten bis in die siebziger Jahre), ist aber auch systematisch, z.B. aus Konsistenz-/Kohärenz-Gründen erklärbar, bzw. negativ aus dem Fehlen geeigneter Volltextverfahren (vgl. Lit. 54; Lit. 60).
- f) Abstracts erleichtern die Erstellung von Übersichtsartikeln (*reviews, state-of-the-art reports*) und Bibliographien.

B 6.1.2 Abstracting

Abstracting (im folgenden synonym mit Referieren benutzt) ist ohne Zweifel im Bewusstsein derjenigen, die sich professionell und praktisch damit beschäftigen, eher eine Kunst als ein regelgeleiteter Routineprozess. Nicht umsonst nennt Edward T. Cressmins, mit der Erfahrung einiger tausend produzierter Abstracts im Hintergrund, sein ISI-Buch *The art of abstracting*: „The art of abstracting demands the application of extensive reading, thinking, writing, and editing skills“ (Lit. 15, S. 3). Vermutlich ist die Zubilligung von „Kunst“ etwas übertrieben (im englischen ist „art“ ja auch eher die Kunst des Handwerklichen), wenn man sich die Millionen in der Regel unter Zeitdruck produzierten Abstracts in den großen Referateorganen bzw. entsprechenden Online-Versionen anschaut: *Chemical Abstracts* mit einem Zuwachs von mehr als eine halbe Million DE pro Jahr; *Biological Abstracts* mit mehr als 250.000 DE pro Jahr Zuwachs. Auf jeden Fall – so die einhellige Meinung – setzt der Vorgang des Abstracting in hohem Maße Kompetenz

im Fachgebiet der zu referierenden Texte voraus: „Abstracts can only be made by persons fully aware of ... research problems, and, if possible, should be made by scientists and engineers“ (Lit. 52, S. 18).

Der Ausdruck „Kunst“ wird wohl auch deshalb gerne gebraucht, um die Skepsis bezüglich einer Automatisierbarkeit des Vorgangs des Referierens auszudrücken. Diese Skepsis spiegelt sich in der Stellungnahme eines Dokumentars wider, der anlässlich einer Erhebung von Abstract-Regeln seine Meinung zum automatischen Abstracting darlegte (vgl. Lit. 63, S. 134): Zur Analyse von „technisch/wissenschaftlichen Texten gehört Sachverstand, Berufserfahrung und Kenntnis vom Benutzerbedarf... Der mechanische ... Prozess kann die Texte immer nur in einer Ebene auswerten, die dem Wert wissenschaftlicher Arbeit oder Industrieerfahrung nicht gerecht wird. Hier ist der Einsatz des Rechners nach unserer Auffassung einmal tatsächlich so gefährlich, wie er in polemischen Darstellungen so oft beschrieben wird.“ Der mechanische Prozess wird der Kunst gegenüber gestellt. Nicht zuletzt die Forschungsansätze der Künstlichen Intelligenz (KI) und der Computerlinguistik (auch auf der Grundlage neuronaler Netzwerke) scheinen den Schluss zu erlauben, dass der mechanische Prozess vielleicht nicht mehr so mechanisch sein muss, sondern auf Verfahren beruhen kann, die der oben angeführten intuitiven Beschreibung des Dokumentars gar nicht so fern stehen: Dem Sachverstand wird durch leistungsstarke Wissensbasen mit Fakten und (auch heuristischen) Regeln Rechnung getragen; Kenntnisse vom Benutzerbedarf können in entsprechende Benutzermodelle oder durch Auswertung von Anfrageinformationen eingebracht werden. Berufserfahrung kann ein Rechner natürlich nicht erwerben, aber man kann ihm Lernstrategien, z.B. Verfahren der Analogiebildung und des induktiven und deduktiven Schließens, beibringen, mit denen er so etwas wie „Erfahrung“ gewinnen kann.

Wie dem auch sei, ob Kunst oder Mechanik, Abstracting, sei es intellektuell oder automatisch, wird nur dann Gegenstand dokumentarischer oder informationswissenschaftlicher Ausbildung sein können, wenn Aussagen darüber gemacht werden können, wie Abstracts von Dokumentaren oder Rechnern angefertigt werden können (Lit. 31). Dass über dieses z.B. in Regeln bzw. Handlungsanweisungen formulierte Wissen hinaus noch etwas als menschliche Leistung hinzukommen kann, was man dann

eben als Kunst oder Begabung bezeichnet, sei unbestritten. Besonders Cremmins stellt darauf ab, dass Abstracting nicht bloß eine Kopier- und Extracting-Funktion ist, sondern in hohem Maße eine produktive intellektuelle Funktion ist: „Abstractors must have a good grasp of and be able to apply the principles of analytical reading, logical thinking, informative writing, and concise editing (Lit. 15, S. 13).“ Die Abhängigkeit von der individuellen intellektuellen Leistung bringt natürlich auch Probleme mit sich. Schon in frühen Studien (z.B. Lit. 82; Lit. 77) hat man eine geringe Konsistenz zwischen verschiedenen Abstractern bei der Selektion von Sätzen aus Texten für Abstracts festgestellt, allerdings eine gewisse Präferenz für sogenannte *topic-Sätze*. Ähnliche Ergebnisse liegen aus dem *Indexing-Bereich* vor, so dass die etwas makabre These nahe liegt, dass automatische Verfahren (des Abstracting oder Indexing) möglicherweise schlechtere Leistungen als entsprechende intellektuelle erbringen, dafür aber verlässlicher, da konsistenter, sind.

Zu der Frage der Kunst des Abstracting gehört sicherlich auch die Qualifikation des Abstracters. Die Qualität des Abstracts hängt von der Vertrautheit des Referierers mit der Sprache des Originaldokuments, aber entscheidender wohl von seinem Wissen über den Gegenstandsbereich ab. Untersuchungen haben nach Mathis/Rush (Lit. 77) ergeben, dass Experten dazu neigen, zu wenig Information aus Texten herauszuholen, Laien dagegen nur unzureichend zwischen wichtigen und bekannten Daten unterscheiden können. Am besten scheint eine mittlere Qualifikation zwischen den Extremen „of expertise and passing knowledge“ (Lit. 77, S. 111) zu sein. Vielleicht ist es leichter, Experten das Referieren als Referierern Expertenwissen beizubringen. Entsprechend ist wohl auch die Einstellungspraxis von Dokumentationszentren und Abstract-Diensten – ob berechtigt oder nicht –, nämlich zuerst auf die Fachqualifikation und dann erst auf die informationsmethodische und literarische Qualifikation zu achten.

B 6.2 Entwicklung und Stand des Abstracting

B 6.2.1 Historische Entwicklung

Historisch gesehen ist das Entstehen von Kondensierungsverfahren, zu denen prominent das Referieren zählt, Ausdruck der Notwendigkeit, die seit

der Ausdifferenzierung der Bürgerlichen Gesellschaft in allen Fachgebieten ständig ansteigende Produktion von Texten überschaubar zu halten (vgl. Lit. 64; auch für das Folgende). Den Überblick müssen sowohl diejenigen behalten, die an der Produktion neuen Wissens arbeiten (Wissenschaftler) und dabei nicht mehr die Zeit haben, alles zu lesen, sondern nur noch das Relevante, als auch diejenigen, die Wissen anwenden und dabei nicht riskieren wollen, dass ihnen Wesentliches entgeht (z.B. Techniker, Journalisten, Politiker, Verwaltungsfachleute, Manager).

Nach allgemein anerkannter Meinung (vgl. Lit. 10; Lit. 77; Lit. 88) leiten sich verstärkte Anstrengungen, wissenschaftliche Untersuchungen in einen organisierten Zustand zu bringen, aus dem 17. Jahrhundert ab. Als wesentlicher Einschnitt wird Baccos Vorschlag zur Gründung eines „college of research“ angesehen. Der Gründung der „Royal Society of London“ (1665) folgten viele weitere europäische Gesellschaften. Die ersten „Proceedings“ (The Philosophical Transactions) als Überblicksleistungen erschienen schon im Gründungsjahr der „Royal Society“. Zu Beginn des 19. Jahrhunderts (1821) ist mit der Publikation der „Jahresberichte über die Fortschritte der physischen Wissenschaften“ durch Berzelius eine neue Systematisierungsform gefunden. Auch schon im 18. Jahrhundert gab es ähnliche Anstrengungen, z.B. das „Berlinische Jahrbuch für die Pharmacie und die damit verbundenen Wissenschaften“, das bis 1840 erschien. Dies waren allerdings Jahrbücher bzw. State-of-the-Art-Berichte.

Der erhöhte Bedarf nach kontinuierlicher Information in einem Fachgebiet findet seinen Niederschlag in den Abstracts des „Pharmaceutischen Centralblattes“ (1830). Danach steigt die Zahl der Referateorgane kontinuierlich an. Bis in die Mitte des 20. Jahrhunderts sind Referateorgane die beliebteste und vorherrschende Form der Referenzinformation (Lit. 04), und sie werden bis in die Gegenwart stark genutzt, natürlich immer mehr über den Zugriff über die Online-Datenbanken. Die Schätzungen über die Anzahl der regelmäßig produzierten Referateorgane schwanken so beträchtlich, dass genaue Angaben schwierig auszumachen sind. Gaus legt sich auf eine Zahl von 3.500 Referatezeitschriften zum Zeitpunkt 1983 fest (Lit. 39); dazu kommen sicherlich noch eine Vielzahl an organisationsinternen Referateblättern. Vickery/Vickery (Lit. 98, S. 305) führen allein für das Jahr 1979 für Großbritannien 339 Indexing- und

Abstracting-Dienste mit insgesamt 157 Verlegern an. Dass Abstracts weiter populär sind und benötigt werden, den Überblick zu behalten, sieht man auch daran, dass Suchmaschinen (nach Umfang) flexible, wenn auch wenig intelligente Summarizing-Verfahren anbieten.

B 6.2.2 Typen von Abstracts

DIN 1426 (Oktober 1988) (Lit. 21, S. 2) stellt Abstracts in den allgemeinen Kontext der Inhaltsangaben als „verkürzte Darstellung[en] des Inhalts [von] Dokument[en]“ und unterscheidet, wie schon erwähnt, zwischen:

- a) Inhaltsverzeichnis
- b) Auszug
- c) Zusammenfassung
- d) Annotation
- e) Kurzreferat (Abstract)
- f) Sammelreferat
- g) Rezension und Sammelrezension
- h) Literaturbericht.

In diesem Zusammenhang sind vor allem die Formen b, c, d, e einschlägig, so dass deren Beschreibungen nach DIN 1426 hier wiedergegeben werden sollen:

Auszug: Ein Auszug ist die verkürzte Wiedergabe eines Dokuments durch ausgewählte, repräsentative Teile. Hierbei können auch Teile ausgewählt werden, die im Originaldokument nicht direkt aufeinander folgen (Sätze, Abschnitte, bei Filmen: Schnittfolgen).

Zusammenfassung: Eine Zusammenfassung ist die Darstellung der wesentlichen Ergebnisse und Schlussfolgerungen eines Dokuments oder von Teilen eines Dokuments und steht meist am Ende des Textes, den sie im allgemeinen zu ihrem Verständnis voraussetzt. Dadurch unterscheidet sie sich vom „Kurzreferat“; die Benennungen „Zusammenfassung“ und „Kurzreferat“ sollten daher nicht synonym benutzt werden.

Annotation: Die Annotation ist eine möglichst kurze allgemeine Charakterisierung eines Dokuments. Sie ist bestimmt durch folgende Merkmale:

- a) Sie ist möglichst redundanzfrei, d.h. sie enthält keine Angaben, die aus dem Titel eines Dokuments oder in Verbindung mit dem Titel erschlossen werden können.

- b) Sie dient dem weiteren Verständnis des Titels eines Dokuments unabhängig von bestimmten Benutzerbedürfnissen und soll den Hauptgegenstand des Dokuments verdeutlichen.
- c) Sie ist rein deskriptiv, braucht nicht aus vollständigen Sätzen zu bestehen und darf nur Angaben enthalten, die aus dem Dokument erschlossen werden können.

Kurzreferat (Abstract): Das Kurzreferat gibt kurz und klar den Inhalt des Dokuments wieder. Das Kurzreferat soll informativ ohne Interpretation und Wertung (Ausnahme siehe kritisches Referat) und auch ohne die Originalvorlage verständlich sein. Der Sachtitel soll nicht wiederholt, vielmehr, wenn nötig, ergänzt oder erläutert werden. Es müssen nicht alle Inhaltskomponenten des Dokuments dargestellt, sondern es können diejenigen ausgewählt werden, die von besonderer Bedeutung sind.

Da das Abstract hier im Vordergrund steht, wollen wir dessen Hauptformen und später seine Merkmale weiter diskutieren.

Bezüglich des Verfassers wird allgemein zwischen Autoren- und Fremdreferaten unterschieden. Cremmins (Lit. 15, S. 5) nennt die letzteren Access-Abstracts, die, in der Regel von Informationsspezialisten geschrieben, einen besseren Einstieg in Publikationsdienste bzw. Online-Datenbanken ermöglichen. Autorenreferate sind nach Rowley (Lit. 86) zu den „homotopic abstracts“ zu rechnen, welche dadurch charakterisiert sind, dass sie zur gleichen Zeit wie die Originaltexte publiziert werden. Daher sind „homotopic abstracts“ meistens Autorenreferate. DIN 1426 legt besonderen Wert darauf, dass ein Autorenreferat deutlich als solches zu kennzeichnen sei. Die Mehrheit der Fachleute ist skeptisch bezüglich des Informationswertes von Autoren-Abstracts (vgl. auch Lit. 10, S. 13), auch wegen der rechtlich schwierigen Situation bei der Übernahme von Autorenreferaten in Referateorgane bzw. Online-Banken. Allerdings spricht der Zeitvorteil („homotopic“) zugunsten der Autorenreferate.

Nach dem inhaltlichem Bezug wird unterschieden nach

- informativen Referaten
- indikativen Referaten
- informativ-indikativen Referaten als Zwischenform.

DIN 1426: „Das informative Kurzreferat gibt so viel Information wieder, wie Typ und Stil des Dokuments zulassen. Es gibt insbesondere Auskunft über das behandelte Gebiet, Zielsetzungen, Hypothesen, Methoden, Ergebnisse und Schlussfolgerungen der im Originaldokument enthaltenen Überlegungen und Darstellungen, einschließlich der Fakten und Daten.

Das indikative Referat gibt lediglich an, wovon ein Dokument handelt. Es weist den Leser auf die im Dokument behandelten Sachverhalte hin und deutet die Art der Behandlung an, aber gibt nicht konkrete Resultate der im Dokument enthaltenen Überlegungen oder dargestellten Untersuchungen wieder.

Das informativ-indikative Referat ist eine Mischform, die den Benutzer über ausgewählte Sachverhalte informiert und andere Sachverhalte nur erwähnt.

Im Normalfall soll informativ referiert werden. In begründeten Fällen, z.B. bei längeren Texten, wie Übersichtsdarstellungen, Literaturberichten und vollständigen Monographien, kann ein indikatives

Referat verwendet werden. Für das informativ-indikative Referat soll man sich insbesondere dann entscheiden, wenn Beschränkungen bezüglich der Länge des Kurzreferats oder Typ und Stil des Dokuments ein informatives Referat nicht möglich machen. Dabei können Sachverhalte exemplarisch oder nach spezifischen Benutzerbedürfnissen oder wegen ihres Neuigkeitswertes herausgehoben werden.“

Beispiele für diese drei Typen von Kurzreferaten sind in der Norm auf S. 6 angeführt. Wir wollen hier ein Borko/Bernier (Lit. 10) entnommenes Beispiel für ein modulares Abstract anführen (Abb. 2), da in ihm die wesentlichen Formen (auch der Typ des kritischen Referats) zusammengeführt worden sind.

Wir haben eingangs erwähnt, dass Abstracts in der Regel mit Blick auf einen anonymen Benutzerkreis hin produziert werden. Eine Ausnahme stellen die sogenannten „slanted abstracts“ dar, die meistens für betriebsinterne Zwecke auf einen genau definierten Benutzerkreis hin erstellt werden. Dabei ist das Abstraktionskriterium das Benutzerprofil der

Ablation of fiberglass-reinforced phenolic resin. R. E. Rosensweig and N. Beecher. American Institute of Aeronautics and Astronautics Journal 1, 1802 - 9 (1963).

Annotation:

A model is developed for charring and melting a composite material with glassy ablation combined with char-layer-molten-glass reactions.

Indicative:

Variables in ablation of a fiberglass-phenolic-resin composite include glass ablation and plastic pyrolysis, flow of melt, mass loss, reaction-heat absorption, mass injection, and coupling between pressure and chemical reaction. Mathematical development and approximation are discussed. Parametric examinations are made.

Informative:

Melting and pyrolysis and other chemical reactions are considered in this theory of ablation of phenolic-resin-fiberglass composite. In this theory, reaction occurs in a surface film in which carbon from pyrolysis of the resin reacts with the glass. For IRBM reentry, there is little temperature drop in the reaction zone, usually less than 1 % and 6 % maximum. Depth of the reaction zone was one-thousandth that of the thermal thickness. The unreacting runoff in the melt was 40 - 80 % and was a function of the possible reaction-enthalpy level. More than 99 % of the material reaching the reaction zone was affected. At 1400 - 2000°C the reaction assumed was: $\text{SiO}_2 + 3 \text{C} = \text{SiC} + 2 \text{Co}$. Up to 25 % increase in the ablation rate appeared only at lower reaction rates. Changing reaction enthalpy three times changed the reaction rate less than 10 %. The value calculated according to this theory for peak reentry ablation rate was 38 % below the experimental value.

Critical:

This theory of ablation of carbon-contaminated glass extends the work of Bethe and Adams (Cr. Avco-Everett Research Laboratory, Research Report No. 38, Nov. 1985) on glasses. Experimental ablation was 38 % greater than that calculated by this theory. Thorough error analysis was not included. Spalding (Aero Quarterly 237 - 74 (Aug. 1961)) and Scala (General Electric Co. MSVD, report R59SD401 (July 1959); ARS Journal 917 - 24) have treated similar problems.

Abb. 2: Modulares Abstract (Quelle: Figure 1.4 aus Lit. 12, S. 17)

Zielgruppe (vergleichbar damit der SDI-Dienstleistung). Solche Abstracts müssen also keine Zusammenfassungen von Texten sein, sondern filtern die Texte gezielt nach bestimmten Informationen. Ähnliche Ansätze sind auch in der Yale-Schule der KI beim automatischen Textverstehen und -zusammenfassen verwendet worden (vgl. Lit. 18), bei denen Texte „überflogen“ („skimming“) werden und eine detailliertere Analyse von Textpassagen erst dann unternommen wird, wenn es Hinweise dafür gibt, dass sie für vorabdefinierte Interessen relevant sein könnten. Rowley rückt diese „slanted abstracts“ in die Nähe von „missionoriented abstracts“ („an abstract which is prepared for a mission-oriented abstracting service, or an abstracting service that has been charged to cater for the application of a specific branch of knowledge“; Lit. 86, S. 16) oder von „findings-oriented abstracts“, bei denen nicht spezielle Gruppeninteressen als Filter dienen, sondern das Interesse an Forschungsergebnissen. Weitere Formen sind leicht vorstellbar. In eher populären Fachzeitschriften kommen auf Titelseiten oder bei Inhaltsverzeichnissen häufiger sogenannte „highlight abstracts“ (vgl. Lit. 86, S. 17) zum Einsatz, die in Ergänzung zum Titel von Beiträgen oder zu systematischen Stichwörtern mit einigen „highlights“ aus den Artikeln das Interesse der Leser erwecken sollen. Dabei wird auf eine Zusammenfassung des gesamten Originalbeitrags in der Regel aber kein Wert gelegt.

Zur Vervollständigung sei noch auf das Strukturreferat hingewiesen, das allerdings in der Praxis trotz seines hohen Informationswertes bislang sehr wenig zum Einsatz kommt, da die Erstellung im intellektuellen Prozess aufwendig ist. Das Strukturreferat, auch Positionsreferat, ist „einheitlich nach vorgegebenen Kategorien in einer meist vorgegebenen Reihenfolge gegliedert. Beispiele solcher Kategorien sind: Zielsetzung, Gegenstand, Verfahren und Methode, Ergebnis, Anwendung, Zeitraum, geographischer Raum“ (DIN 1426). Die einzelnen Kategorien können, falls das Originaldokument dies erlaubt, mit Texten, aber auch mit Deskriptoren oder anderen informationellen Einheiten gefüllt werden. Nützlich ist diese Form sowohl aus Sicht des Herstellers des Abstracts (vorgegebene Kontrollstruktur) als auch des Rezipienten (erleichtert die Orientierung). Abb. 3 zeigt ein weitgehend formal gehaltenes Strukturreferat, wie es z.B. für die Beschreibung von Proceedings-Beiträgen verwendet werden könnte. Wie man am Eintrag „Vortragender“ erkennt, können die Strukturen auf weiteren Ebenen ausdifferenziert werden. Die lineare Darstellung stößt allerdings dann bald auf Grenzen.

Die Form von Strukturreferaten erinnert an Strukturen, wie sie aus der Wissensrepräsentationstechnik bei *Frame-Sprachen* oder Skripts üblich sind. Eine prinzipielle Erweiterung des Strukturrefera-

VORTRAG	
Thema	Graphische und textuelle Erzeugung flexibler Kondensate aus Textwissensstrukturen. Die Systemleistung von TWRM-TOPOGRAPHIC als Prototyp eines neuartigen Informationssystems
Anlass	Deutscher Dokumentartag
Zeit	29.9.1988
Ort	Aachen
Vortragender	Name: Rainer Kuhlen Institution: Universität Konstanz, Fachgruppe Informationswissenschaft Adresse: D-7750 Konstanz, Postfach 5560 Qualifikation: Hochschullehrer
Zielgruppe	Dokumentare, Informationswissenschaftler, KI-ler, Informatiker, Linguisten, Datenbasisproduzenten, Datenbankanbieter
Veröffentlichung	vermutlich in den Proceedings des Dokumentartags
Honorar	nicht vorgesehen

Abb. 3: Strukturreferat für Beiträge zu einem Proceedings-Band

tes ist dann möglich, wenn bei einer Organisation von Kondensierungsstrukturen als Hypertexte die Einträge in Tiefe und Umfang variabel und multimedial gestaltet werden können. Aber damit löst sich prinzipiell die textuell-diskursive Struktur von Referaten/Kondensaten auf (wie auch schon in Abb. 3 angedeutet). Dieser Fall ist in der Norm 1426 noch nicht vorgesehen.

Zum Abschluss der Typendiskussion sei noch daran erinnert, dass Textzusammenfassungen nicht naturgegeben textueller Art sein müssen. In der Geschichte des Information Retrieval ist verschiedentlich mit graphischen Darstellungs- und Ausgabeformen experimentiert worden (Lit. 45; Lit. 64). Vor allem angesichts der referenziellen oder indikativen Funktion des Retrieval (also auf die einschlägigen Texte zu verweisen) spricht einiges dafür, neben Texten auch Graphiken einzusetzen, auch unter der Annahme, dass strukturierte graphische Darstellungen den Verständnisprozess fördern können.

Wir weisen hier nur auf zwei Stationen in der Entwicklung hin, deren Möglichkeiten wegen des fortgeschrittenen Technologiestandes erst heute voll ausgeschöpft werden können. Schon in den frühen Arbeiten von Doyle (Lit. 22; Lit. 23) wurde versucht, Wissen in Assoziationsnetzen/-karten graphisch darzustellen, wobei die dermaßen kondensierte Darstellung gleich auf ganze Bibliotheken abzielen sollte. Die „Wissens“netze sollten entsprechend dem damaligen Stand der statistischen Linguistik nach Assoziationsmaßen auf der Basis von Konkurrenzen berechnet werden, wobei die Intensität einer relationalen Beziehung durch Entfernung der Knoten bzw. durch gewichtete und unterschiedlich graphisch gestaltete Kanten dargestellt wurde. Dienten die Netze zunächst beim Retrieval der Orientierung über den Beständen, also der Unterstützung der Frageformulierung, wobei auch hierfür einfache Stilmittel, wie unterschiedliche Fonts für die Konzepte darstellenden Knoten, benutzt wurden, so sollten sie auch zu diagrammatischen Repräsentationen individueller Dokumente (Lit. 23, S. 385) oder in einer Endstufe zu kondensierten Darstellungen ganzer Bibliotheken führen. Mehr Linguistik wurde von Strong (Lit. 91) im Kontext des ADAM-Projektes der Ohio State University eingebracht, wobei sowohl eine vollständige Syntaxanalyse zum Einsatz kam als auch semantische Relationen auf der Basis der Fillmore'schen Kasusgrammatik, angereichert durch semantische Merkmale, verwendet und graphisch dargestellt

wurden. Hierbei wurden als Stilmittel vor allem unterschiedliche Kantentypen verwendet. Allerdings beruhen die bei Strong (Lit. 91) angeführten graphischen Surrogate der Texte zwar wohl auf algorithmischen Verfahren, sind aber graphisch nicht automatisch generiert (vgl. auch Lit. 93).

B 6.2.3 Merkmale von Abstracts und Regeln zu ihrer Erstellung

Die allgemeinen Merkmale von Referaten sind in der Fachliteratur weitgehend unbestritten, auch wenn Einigkeit darüber besteht, dass sie nur annäherungsweise zu realisieren sind (z. B. Konflikte zwischen Vollständigkeit und Kürze). Die Ausprägungen dieser allgemeinen Merkmale variieren natürlich je nach Abstract-Typ, so dass die Norm zur Vermeidung von Missverständnissen vorschlägt, die Bezeichnung „Referat“ nur in Komposita und Zusammensetzungen zu benutzen (z.B. indikatives Referat). Wir folgen auch hier der Formulierung der Norm:

- a) **Vollständigkeit.** Das Kurzreferat muss für den Fachmann des jeweiligen Bereichs ohne Rückgriff auf das Originaldokument verständlich sein. Alle wesentlichen Sachverhalte sollen auch im Hinblick auf die maschinelle Recherche im Kurzreferat explizit enthalten sein. Hierzu gehören: Hypothese, Zielsetzung, Gegenstand, Verfahren und Methode, Ergebnis, Schlussfolgerung, Anwendung, Zeitraum, geographischer Raum. Der Titel soll jedoch im Kurzreferat nicht wiederholt werden. Um Fehlinformationen der Benutzer und Ballast bei der maschinellen Recherche zu vermeiden, sollen Nebenthemen nur dann berücksichtigt werden, wenn das Originaldokument ausreichend informative Aussagen enthält und ein besonderer Grund vorliegt.
- b) **Genauigkeit.** Das Kurzreferat soll genau die Inhalte und die Meinung der Originalarbeit wiedergeben, d.h. es soll weder die Akzente des Originals verschieben noch im Original nicht enthaltene Angaben bringen.
- c) **Objektivität.** Das Kurzreferat soll sich jeder Wertung enthalten. Zwar bedeuten Auswahl und Darstellung der referierten Sachverhalte immer einen subjektiven Einfluss des Referenten, doch soll dieser soweit wie möglich reduziert werden. Deshalb empfiehlt es sich, den Aufbau des Kurzreferats am Aufbau des Originaldokuments auszurichten sowie die Terminologie des Autors und direkte Zitate zu übernehmen.

- d) **Kürze.** Das Kurzreferat soll so kurz wie möglich sein. Überflüssige Redewendungen sind daher zu vermeiden. Allgemein bekannte Abkürzungen sowie – in einigen Fachbereichen – der Einsatz von graphischen Hilfsmitteln (z.B. Strukturformeln) können ebenfalls zur Kürzung der Inhaltsangabe beitragen.
- e) **Verständlichkeit.** Das Kurzreferat soll verständlich sein. Dies ist zu erreichen durch: Verwendung möglichst weit verbreiteter Fachausdrücke, Verwendung national oder international eingeführter Nomenklaturen, Maßangaben, Formelzeichen, Symbole und Abkürzungen, Vermeidung ungebräuchlicher Fachausdrücke (vom Verfasser der Originalarbeit neu eingeführte Ausdrücke sollen erwähnt und erläutert werden).

DIN 1426 regt ausdrücklich an, im konkreten Fall eigene Regelwerke zu erstellen, die je nach Verwendungszweck die Merkmalsbeschreibungen eindeutig machen sollen. DIN 1426 ist also nicht als unmittelbare Handlungsanweisung zu verstehen, sondern bedarf der jeweiligen Anpassung. Entsprechend viele Regelwerke mit zum Teil detaillierten Anweisungen liegen vor (vgl. Lit. 63 mit einem Verzeichnis von Regelwerken, die in der Dokumentationspraxis zum Einsatz kommen). In der angelsächsischen Welt wird ISO 214 (Lit. 53) häufiger direkt als Regelwerk verwendet.

Dass Regelwerke manchmal ganz knapp gehalten werden können, wird bei Borko/Bernier (Lit.12, S. 47) belegt (Abb. 4).

In Form von Handlungsanweisungen fasst Cremmins (Lit. 15, S. 10) seine langen Erfahrungen (und die des ANSI) bei der Herstellung von Abstracts wie folgt zusammen:

- Prepare an abstract that access services can reproduce with little or no change, copyright permitting.
- State the purpose, methods, results, and conclusions presented in the original document, either in that order or with initial emphasis on results and conclusions.
- Make the abstract as informative as the nature of the document will permit, so that readers may decide, quickly and accurately, whether they need to read the entire document.
- Unless otherwise instructed, use fewer than 250 words for most papers and portions of mono-

DDC Rules for Preparing Abstracts

Outline

In brief:

1. Always an informative abstract if possible
2. 200 - 250 words
3. Same technical terminology as in report
4. Contents
 - a. Objectives or purpose of investigation
 - b. Methods of investigation
 - c. Results of investigation
 - d. Validity of results
 - e. Conclusions
 - f. Applications
5. Numerals for numbers when possible
6. Phrases for clauses, words for phrases when possible
7. No unconventional or rare symbols or characters (...)
8. No uncommon abbreviations
9. No equations, footnotes, preliminaries
10. No descriptive cataloguing data
11. Security Classification
12. Dissemination controls, if any
13. Review it

Abb. 4: Regelwerk für die Erstellung von Abstracts (Quelle: Lit. 10, S. 47)

graphs and fewer than 100 words for notes and short communications. For long reports and theses, do not exceed 500 words.

- Avoid including background information or citing the work of others in the abstract, unless the study is a replication of their work.
- Do not include information in the abstract that is not contained in the textual material being abstracted.
- Verify that all quantitative information or qualitative information used in the abstract agrees with the information contained in the full text of the document.
- Use standard English and precise technical terms, and follow conventional grammar and punctuation rules.
- Give expanded versions of lesser known abbreviations and acronyms, and verbalize symbols that may be unfamiliar to readers of the abstract.
- Omit needless words, phrases, and sentences.

Regelwerke betonen meistens lediglich die formalen Eigenschaften von Abstracts (z.B. Umfang, Syntaxmuster). Man kann ihnen kaum entnehmen, woraus denn der eigentliche intelligente Prozess (die Kunst) besteht, der zu einem guten Abstract führt. Wir haben schon auf die vier Prinzipien Cremmins des analytischen Lesens, des logischen Denkens, des informierenden Schreibens und des konsistenten Editierens hingewiesen, die er detailliert ausführt (vgl. auch Lit. 16). In Abb. 5 sind die wesentlichen Einsichten zusammengefasst.

In verschiedenen Arbeiten ist darauf aufmerksam gemacht worden (vgl. Lit. 29; Lit. 30; Lit. 31), dass für den Prozess der Erstellung von Abstracts sowohl die Ergebnisse der eher praktisch bzw. deskriptiv ausgerichteten Schreib- und Leseforschung (Lit. 01; Lit. 35; Lit. 47; Lit. 05; Lit. 78) und der theoretischen Textlinguistik (vor allem Kintsch, van Dijk; Lit. 58; Anwendungen: Lit. 19; Lit. 20) für das Abstracting einschlägig sind. Dabei muss darauf geachtet werden, inwieweit Aussagen über die allgemeine Produktion von Texten auf den Spezialfall der Produktion von informierenden Texten, der beim Abstracting gegeben ist, angewendet werden können. Es scheint jedoch sinnvoll zu sein, die bislang überwiegende Beachtung der Fachkompetenz von Abstractern durch stärkere Berücksichtigung der kognitiven und sprachlichen Momente des Texterstellungsprozesses und seiner Rahmen-

bedingungen (z.B. Benutzung von technischen Hilfsmitteln, integrierte Autorenarbeitsplätze, informationelle Absicherung durch Zugriff auf externe Ressourcen, Kontakte mit Autoren und Distributoren) zu ergänzen. Das Schreiben von Abstracts ist eine hochwertige kognitive und sprachliche Leistung (Lit. 57) und sollte in seinem Umfeld entsprechend dem Stand des technisch und methodisch Möglichen unterstützt werden. In der Habilitationsschrift von Endres-Niggemeyer (Lit. 31) wird dieser kognitive Ansatz auf der Grundlage reicher empirischer Studien als Grundlage für eine erweiterte Theorie des Abstracting ausgebaut. Aus der Textlinguistik scheint vor allem das von Kintsch/van Dijk 1983 (Lit. 58) vorgelegte Modell zum Verstehen und Zusammenfassen von Texten für das Abstracting produktiv zu sein. Texte werden in diesem Modell aus einer allgemeinen semantischen Struktur (macrostructure) abgeleitet: „The result of the comprehension process is a well-organized text base in the memory of the reader. For every reader, the macrostructure may be different, depending on goals and strategies of comprehension. The reduction process is performed by macrostrategies. They cooperate with other supporting strategies of text comprehension. At each level of reduction, they produce a coherent text, consisting of macropropositions. The derived text is shorter and more abstract and may – at a certain point of reduction – correspond to an informative abstract.“ (Lit. 30)

Stages	Techniques	Results
1. Focusing on the basic features of the materials to be abstracted	Classifying the form and content of the materials	Determination of the type of abstract to be written, the relative length, and the degree of difficulty
2. Identifying relevant information (sometimes done simultaneously with Stage 1)	(a) Searching for cue or function words and phrases, structural headings and subheadings; (b) expanding the search based on the results of (a)	Identification of a representative amount of relevant information for extraction
3. Extracting, organizing, and reducing the relevant information	Organizing and writing the extracted relevant information into an abstract, using a standard format	Preparation of a concise, unified, but unedited abstract
4. Refining the relevant information	Editing or review of the abstract by the originator or editorial or technical reviewer	Completion of a good informative or indicative abstract

Abb. 5: Stufen im menschlichen Abstracting-Prozess (Quelle: Lit. 15, S. 17)

Eine explizit informationswissenschaftliche Theorie des Abstracting bzw. der „Aboutness“ (Lit. 06) hat unter dem Einfluss der Prager Textlinguistik (vor allem unter Verwendung der textuellen Progressionsmuster von Danes, Lit. 17) Hutchins (Lit. 50; Lit. 51) vorgelegt (vgl. auch die folgenden Ausführungen zur Textlinguistik mit Blick auf automatisches Generieren und Kondensieren).

Im weiteren Sinne zu den Regeln des Abstracting gehören auch Kriterien, nach denen Originaldokumente zum Zwecke des Abstracting ausgewählt werden (Abstract-Würdigkeit). Rowley (Lit. 86, S. 22f) hat die folgenden Punkte zusammengestellt:

- a) pertinent to the interests of clients
- b) novel contributions to a given field of endeavour
- c) final reports, or other reports well supported by sound methodology and convincing evidence
- d) those which convey information that is likely to be difficult to access, such as foreign documents or internal reports and memoranda and other documents with a limited circulation
- e) significant advances and reviews
- f) those which contain information located in sources known to be reliable, such as the information in professional and technical journals with an established reputation
- g) sources, in particular journals or reports issued by a specific organization, for which the abstracting agency has undertaken to give comprehensive coverage.

B 6.2.4 Abstracting und Online-Banken

Interessanterweise hat das Aufkommen der Online-Informationsbanken die informationsmethodische Beschäftigung mit Referaten nicht eingeschränkt. Zwar mag es sein, dass Online-Informationsbanken das sukzessive Ende von Referateorganen in ihrer gedruckten Form bedeuten können, der angesprochene referenzielle Informationswert von Abstracts beim Überfliegen von Retrievalergebnissen scheint den Nutzern von Informationssystemen jedoch nach wie vor hoch zu sein. Entsprechend wird auch heute noch viel methodische Energie beim Formulieren von Regeln zum Referateerstellen aufgebracht. Viele Regeln aus der Doku-

mentationspraxis sind erst neueren Datums (vgl. Lit. 61). Zunehmend wird bei diesen Regeln der Tatsache Rechnung getragen (auch in der neuen Version von DIN 1426 von 10/88), dass Abstracts beim *Online-Retrieval* zum Einsatz kommen.

In diesem Zusammenhang sind die Empfehlungen einschlägig, die Fidel (Lit. 36) aufgrund einer Untersuchung an 57 Datenbasen zusammengestellt hat (Abb. 6). Diese 57 Datenbasen waren eine Teilmenge aus insgesamt 123 Datenbasen, die alle im Online-Betrieb recherchierbare Abstracts enthielten. Lediglich bei diesen 57 Datenbasen enthielten die Abstracts-Richtlinien der dafür zuständigen 36 Datenbasenproduzenten ausdrückliche Freitext-Anweisungen, die darauf abzielen, den Informationswert von im Freitext-Retrieval invertierten Dateien möglichst hoch zu halten (vgl. auch Lit. 99), vor allem durch die Verwendung von Begriffen aus einem kontrollierten Vokabular für das Formulieren von Referatetexten.

Die fortschreitende Automatisierung von Informationssystemen ist laufend Anlass, den Wert von klassischen dokumentarischen Techniken neu zu bedenken. In mittlerer Perspektive scheinen dabei die Produkte des automatischen Textkondensierens noch keine realistische Konkurrenz zum intellektuell erstellten Referat zu sein. Eine solche Konkurrenz sind aber offenbar für viele Datenbankanbieter des internationalen Informationsmarktes die Verfahren des Freitextretrievals, die wegen der fallenden Kosten der intellektuellen Arbeit attraktiv zu sein scheinen. Beruht die dafür erforderliche Invertierung der Textwörter auf den Referatetexten, so ergibt sich eine beim Retrieval zuweilen nützliche Koalition zweier Verfahren. Beruht sie auf den Volltexten selber, so entsteht eine in der Regel zuungunsten der Abstracts ausgehende Konkurrenzsituation. Volltextdatenbanken sollen – so die Annahme der Anbieter – die Diskriminierungs- und Referenzleistung von Abstracts übernehmen und sie dadurch sogar übertreffen, dass die Referenz dann tatsächlich zu den vollen Texten führt. Volltextdatenbanken werden nicht zuletzt auch durch den zunehmenden Einsatz von CD-ROM (Volltextdaten mit entsprechender Retrievalsoftware auf optischen Speichern) attraktiv (vgl. Lit. 48). Ogleich man sich in der Forschung einig ist (vgl. Lit. 75; Lit. 55; Lit. 08; Lit. 62; Lit. 94; Lit. 80), dass durch eine bloße Volltextinvertierung keine zufriedenstellenden Retrievalergebnisse zu erzielen sind, mag die Einschätzung und das Verhalten der Datenbankanbieter verständlich sein: fehlen

THE CONTENT OF ABSTRACTS

General statements

Use 'important' concepts and terms (e.g. those which will enhance free-text retrieval, those for which a document gives enough information, or key words).

Index terms

Co-ordinate concepts used in abstracts with assigned descriptors.

- (a) Assign concepts in abstracts that are identical to descriptors.
- (b) Assign concepts in abstracts that complement descriptors (e.g. relevant terms that are missed in descriptor indexing and in titles, terms that are more specific than descriptors, or a particular type of term that is important to the subject area, such as geographic names).
- (c) Assign concepts in abstracts that both complement and are identical to descriptors.

Enhance indexing independent of any index language used.

Check lists

Follow a list of retrieval-related elements that should be included in abstracts.

Forms of check lists:

- (a) Categories that should be included in abstracts (e.g. materials, properties and processes) and the conditions under which they should be included (e.g. only when they are discussed elaborately, or whenever mentioned).
- (b) Specific and particular guidelines (e.g. ,whenever dealing with a new product, mention the company name).

THE LANGUAGE OF ABSTRACTS

Use of author language

Use author language.

Do not use author language.

Use standardised and concrete terms specific to a subject area.

Use both author language and synonyms.

Relationship to index language used

Co-ordinate terms in abstracts with descriptors.

Complement descriptors with terms in abstracts (e.g. use synonyms or more specific terms).

Use specific and well-accepted terms for particular categories (such as materials, processes and products).

Practices to avoid

Do not use negative (e.g. use sick instead of not healthy).

Do not list terms which have a common last word as series (such as 'upper, middle, and working class').

Word forms

Follow local language practices (e.g., change American spelling for British databases).

Always spell out terms in certain categories (such as processes, materials, products).

When a term and a descriptor are the same, record the term in the form used by the descriptor.

Express terms both in their abbreviated and in their complete form.

Abb. 6: Empfehlungen zur Erstellung von Abstracts mit Blick auf ihren Einsatz in Online-Banken (Quelle: Lit. 36, S. 15)

doch nach wie vor entsprechende einsatzbereite Verfahren der Wissensrepräsentation von Volltexten oder der flexiblen automatischen Textkondensation.

Die intellektuelle Arbeit der Inhaltserschließung, hier bezogen auf das Referieren, erweist sich nach wie vor als Engpass des Information Retrieval. Offenbar besteht hier ein tatsächlicher Bedarf nach automatisierten Verfahren. Wir geben daher zum Schluss einige Hinweise auf Möglichkeiten des automatischen Abstracting bzw. des automatischen Textkondensierens, ohne im Zusammenhang dieses Beitrags auch nur annähernd beanspruchen zu können, einen repräsentativen Überblick über den Stand der Kunst des automatischen Abstracting/Summarizing zu geben (vgl. dazu Lit. 31; Lit. 33; Lit. 41; Lit. 44, mit Blick auf Anwendung für wissenschaftliche Literatur Lit. 95). Einen Überblick über „Automatic Abstracting & Summarizing Tools“ geben Lit. 97; vgl. Lit. 11; Lit. 32; Lit. 34.

Zwar sind auch heute noch nicht diese Verfahren derart robust, dass sie in der realen Dokumentationspraxis routinemäßig zum Einsatz kommen könnten. Jedoch sind die Konzepte so weit entwickelt und die experimentellen Arbeiten so weit vorwärts getrieben (vgl. Lit. 33; Lit. 40; Lit. 45; Lit. 65), dass sich die Fachleute in der Dokumentationspraxis eine Meinung über die Leistungsmöglichkeit dieser Verfahren bilden und entsprechende Maßnahmen für einen sinnvollen Transfer des Wissens in die Praxis einleiten können. Die Zeit scheint reif zu sein für größere Entwicklungsprojekte des automatischen Textkondensierens. Die (gewiss noch unzulänglichen) Summarizing-Angebote im Internet sind ein Hinweis darauf, dass Zusammenfassungsleistungen nützlich sind und im Spektrum elektronischer Dienste, wie auch Übersetzungsdienste, erwartet werden.

B 6.3 Automatische Verfahren des Abstracting

Nicht unumstritten in der Literatur ist die These von Mathis/Rush: „Before any attempt is made to automate the process of abstracting, it is important to understand how humans produce abstracts“ (Lit. 77, S. 113), da weder die Äquivalenz in der Leistung eine Identität der dafür verwendeten Methoden/Algorithmen notwendig bedingt (dies wird in der KI unter dem Gesichtspunkt der kognitiven Plausibilität von Verfahren diskutiert) noch es als

gegeben angenommen werden kann, dass die menschliche Leistung beim Referieren dergestalt befriedigend ist, dass sie für maschinelle Verfahren als Vorbild und Anregung dienen müsste (vgl. Lit. 61). Wir haben auf pragmatische Defizite, Starrheit in der Präsentationsform und auf Inter-Abstracter-Inkonsistenzen hingewiesen, so dass man argumentieren kann: „such emulation would lead simply to a faster rate of production of consistently poor abstracts“ (Lit. 77, S. 117). Es ist also zumindest fraglich, ob es viel Sinn macht, menschliche und maschinelle Abstracting-Leistungen zu vergleichen. Allerdings kann es aber durchaus sein, dass eine maschinelle Leistung nicht nur eher akzeptiert wird, wenn sie einer bekannten menschlichen sehr nahe kommt, sondern dass das Computerangebot auch objektiv als optimal angesehen werden kann, wenn es die kognitive Leistung von Menschen simulieren kann.

Der Wert des maschinellen Abstracting besteht aber vermutlich nicht in der Nachahmung menschlicher Leistung, sondern im Angebot neuer, flexibler Formen des Textkondensierens und des Reagierenkönnens auf variable Bedürfnisse. Dies war allerdings außerhalb der Reichweite der frühen, weitgehend auf statistischen Verfahren beruhenden Ansätze. (Für eine Darstellung dieser frühen Verfahren vgl. Lit. 103; Lit. 77; ein Gesamtüberblick bei Lit. 31; Lit. 40; Lit. 64)

Techniken des automatischen Abstracting waren bis Ende der siebziger Jahre weitgehend Extracting-Verfahren. Borko/Bernier stellten 1975 (Lit. 10, S. 14) in ihrer Untersuchung zum computer-gestützten Abstracting fest, „that the computer may select representative sentences, but it does not yet prepare abstracts – merely extracts“. Seit dem ersten Experiment mit dem Luhn-Algorithmus (Lit. 69) konzentrieren sich die Verfahren darauf, relevante Sätze im Gesamttext aufgrund vorgegebener Kriterien erkennen und in eine Rangfolge bringen zu können und in einer Sequenz auszugeben (vgl. Lit. 13; Lit. 25; Lit. 26; Lit. 27). Aus historischen Gründen sei das mit dem Luhn-Algorithmus erste automatisch produzierte Abstract angeführt (Abb. 7).

Als Kriterien für die Erkennung von relevanten Sätzen können z.B. dienen:

- Konzentration von Wörtern (laufende Textwörter, Grundformen, Stammformen) mit hoher absoluter oder relativer Häufigkeit (wobei die Wörter in der Regel mit einfachen morphologischen

Exhibit I

Source: The Scientific American. Vol. 196, No. 2, 86-94, February, 1957

Title: Messengers of the Nervous System

Author: Amodeo S. Marazzi

Editor's Sub-heading: The internal communication of the body is mediated by chemicals as well as by nerve impulses. Study of their interaction has developed important leads to the understanding and therapy of mental illness.

Auto-Abstract

It seems reasonable to credit the single-celled organisms also with a system of chemical communication by diffusion of stimulating substances through the cell, and these correspond to the chemical messengers (e.g. hormones) that carry stimuli from cell to cell in the more complex organisms. (7.0)

Finally, in the vertebrate animals there are special glands (e.g., the adrenals) for producing chemical messengers, and the nervous and chemical communication systems are intertwined: for instance, release of adrenalin by the adrenal gland is subject to control both by nerve impulses and by chemicals brought to the gland by blood. (6.4)

The experiments clearly demonstrated that acetylcholine (and related substances) and adrenalin (and its relatives) exert opposing actions which maintain a balanced regulation of the transmission of nerve impulses (6.3)

It is reasonable to suppose that the tranquilizing drugs counteract the inhibitory effect of excessive adrenalin or serotonin or some related inhibitor in the human nervous system. (7.3)

Abb. 7: Erstes automatisch produziertes Abstract (Extrakt) (Quelle: Lit. 69)

Verfahren der Reduktion auf Stamm- oder Grundformen normalisiert werden);

- Konzentration von „sprachlichen Einheiten“ mit bestimmten syntaktischen Merkmalen (Wortklassenzugehörigkeit, bevorzugt Substantive; Grundmuster von Nominalphrasen, Nominalgruppen mit angehängten Präpositionalphrasen (NP(art n) PP (präp NP))
- Konzentration von „sprachlichen Einheiten“, die mit Einträgen in einer Positiv-Wortliste oder auch mit Titelwörtern übereinstimmen.

Berücksichtigt werden können dabei auch die Position des Vorkommens relevanter Sätze. Als relevant werden häufig sogenannte *topic*-Sätze angesehen, also Sätze, die bevorzugt am Anfang oder am Ende einer größeren textuellen Einheit (Absatz, Abschnitt, Kapitel) vorkommen. Die zugrundeliegende Hypothese für die *topic*-Sätze-Präferenz ist, dass am Anfang einer größeren textuellen Einheit das Thema bzw. Unterthema angesprochen wird und am Ende eben dieser Einheit eine Zusammenfassung stehen sollte.

Der erste Schritt solcher Extracting-Verfahren besteht also aus dem Identifizieren von potentiell relevanten Sätzen. Der zweite Schritt besteht aus der Gewichtung der Abstract-Kandidaten-Sätze. Auch hier kann die volle Bandbreite statistischer Assoziations- und Clustering-Verfahren zum Einsatz kommen. In wie weit sie kognitiv plausibel sind, ist umstritten; sie sind jedoch hervorragend geeignet, durch Einstellen verschiedener cut-off -Werte unterschiedlich lange Abstracts zu produzieren. Die relevantesten Sätze werden dann in der Reihenfolge ihres Vorkommens im Originaltext ausgegeben.

Einen repräsentativen Überblick über diese statistischen und oberflächenlinguistischen Ansätze des automatischen Abstracting gewinnt man durch die verschiedenen Darstellungen des an der *Ohio State University* von verschiedenen Autoren (z.B. Mathis, Rush, Salvador, Zamora, später auch Strong) in den siebziger Jahren entwickelten ADAM-Systems. Hier sind auch zum ersten Mal (weitgehend auf strukturalistischen Grammatiken beruhende) text-syntaktische Experimente (Lit. 76) aufgrund der Einsicht durchgeführt worden, dass eine Auflistung extrahierter Sätze aus Lesbarkeitsgründen zumin-

dest einer kohäsiven Aufbesserung bedarf. Abstracts sind eben auch Texte, deren Sätze nicht unverbunden sind, sondern intersententiell verknüpft werden müssen, z.B. durch Koreferenzen oder Konjunktionen (Beispiele für automatisch produzierte Abstracts des ADAM-Systems bei Cremmins, Lit. 15) (vgl. Lit. 104).

Methodisch wurde in den 80er Jahren das Abstracting durch die KI-Forschung als exemplarisches Übungsfeld zur Simulation menschlicher intelligenter Leistungen angesehen. Entsprechend sind hier zahlreiche Summarizing-Arbeiten entstanden, z.B. DeJong (Lit. 18; Lit. 44; Lit. 66; Lit. 95; Lit. 34; Lit. 45); Hobbs (Lit. 49); Lehnert/Black/Reiser (Lit. 68); Cook/Lehnert/McDonald (Lit. 14). Einigkeit bestand darüber, dass hochwertige Leistungen der Analyse von natürlicher Sprache (und der Synthese in Form von Abstracts/Summaries) nicht nur durch sprachimmanente Verfahren der Grammatik und Semantik zu erreichen sind, sondern zusätzlich durch Wissen um den in dem jeweiligen Sprachfragment angesprochenen Kontext (in der KI spricht man von Weltwissen; in der Informationspraxis von Fachinformation). Dies angewendet auf Verfahren des automatischen Abstracting bedeutet, dass das Erstellen von Abstracts sich nicht nur auf die Texte selber beziehen kann, sondern Weltwissen heranziehen muss, weniger, um zusätzliches Wissen in den Abstracts unterzubringen (was ja nach der Abstract-Theorie untersagt ist, s. oben), sondern um das in den Texten niedergelegte Wissen überhaupt erst einmal durch Rechner identifizieren zu können. Dies macht den Paradigmenwechsel aus: Nicht länger sind Grundlage des Abstracting die als relevant extrahierten Sätze von Texten, sondern mit Hilfe von Weltwissen erkannte Textwissensstrukturen. Abstracts werden dann nicht mehr direkt aus Texten abgeleitet, sondern eher aus Wissensstrukturen generiert (vgl. Lit. 37; Lit. 38; Lit. 73; Lit. 92; Lit. 71; die Beiträge in Lit. 56, vor allem Lit. 85; Lit. 74; Lit. 89).

Literatur

- 01 Adler, M. J.; van Doren, C. (1972): *How to read a book: the classic guide to intelligent reading*. New York: Simon and Schuster, Inc.
- 02 Acham, K. (1971): Artikel „Abstraktion IV“. In: J. Ritter (Hg.): *Historisches Wörterbuch der Philosophie*, Bd. I A-C, Basel/Stuttgart: Schwabe & Co, Spalten 59-63
- 03 ANSI (1979): American National Standards Institute, Inc. *American National Standard for Writing Abstracts*. ANSI Z39.14 – 1979. New York: American National Standards Institute, Inc.
- 04 Baker, D. B.; Horiszny, J. W.; Metanomski, W. V. (1980): History of abstracting at chemical abstracts. *Journal of Chemical Information and Computer Sciences* 20 (4) 193-201
- 05 Beaugrande, R. de (1984): *Text production. Toward a science of composition*. Norwood, NJ: Ablex
- 06 Beghtol, C. (1986): Bibliographical classification theory and text linguistics: aboutness analysis. *Intertextuality and the cognitive act of classifying documents*. *Journal of Documentation* 42 (2) 84-113
- 07 Bernier, C. L.: Abstracts and abstracting. In: Dym (1985), 423-444
- 08 Blair, D. C.; Maron, M. E. (1985): An evaluation of retrieval effectiveness for a full-text document-retrieval system. *Communications of the ACM* 28 (3) 289-299
- 09 Borko, H.; Chatman, S. (1963): Criteria for acceptable abstracts: a survey of abstracters' instructions. *American Documentation* 14, 2, 149-160
- 10 Borko, H.; Bernier, C. L. (1975): *Abstracting concepts and methods*. New York/London: Academic Press, Inc.
- 11 Brandow, R.; Mitze, K.; Rau, L. F. (1995): Automatic condensation of electronic publications by sentence selection. *Information Processing and Management* Vol. 31 (5) 675-685
- 12 Cleveland, D. B.; Cleveland, A. D. (1983/2000): *Introduction to indexing and abstracting*. Littleton: Libraries Unlimited (3. Auflage 2000)
- 13 Climenson, W. D.; Hardwick, N. H.; Jacobson, S. N. (1961): Automatic syntax analysis in machine indexing and abstracting. *American Documentation* 12 (3) 178-183
- 14 Cook, M. E.; Lehnert, W. G.; McDonald, D. D. (1984): Conveying implicit content in narrative summaries. In: COLING 84: Proceedings of the 10th International Conference on Computational Linguistics & 22nd Annual Meeting of the Association for Computational Linguistics. Stanford Univ., CA, July 2-6: ACL, 5-7
- 15 Cremmins, E. T. (1982): *The art of abstracting*. Philadelphia, PA: ISI Press
- 16 Cremmins, E. T.; Trachtman, M. (1981): Writing good quality human and machine abstracts using the three r's of reading, rules, and relationships. In: L. F. Lunin et al. (eds): *The Information Community. Proceedings of the 44th ASIS Annual Meeting*, Washington, October 25-30, 1981. White Plains, NY: Knowledge Industry Publ., 157-159

- 17 Danes, F. (1974): Functional sentence perspective and the organization of the text. In: F. Danes (ed): Papers on Functional Sentence Perspective. Prague: Academia, 106-128
- 18 DeJong, G. F. (1979): Skimming stories in real time: an experiment in integrated understanding. Yale Univ. (Ph.D./Diss.)
- 19 Dijk, T. A. van (1979): Recalling and summarizing complex discourse. In: W. Burghardt; K. Hölker (eds): Text Processing: Papers in Text Analysis and Text Description. Berlin etc.: de Gruyter, 49-118 (= Research in Text Theory 3)
- 20 Dijk, T. A. van (1980): Macrostructures: an interdisciplinary study of global structures in discourse, interaction, and cognition. Hillsdale/NJ: Erlbaum
- 21 DIN 1426 (1988): Inhaltsangaben von Dokumenten. Kurzreferate, Literaturberichte. Deutsche Norm. Normenausschuss Bibliotheks- und Dokumentationswesen (NABD) im DIN Deutsches Institut für Normung e. V. Berlin: Beuth Verlag
- 22 Doyle, L. B. (1961): Semantic road maps for literature searchers. *Journal of the ACM* 4, 553-578
- 23 Doyle, L. B. (1962): Indexing and abstracting by association. *American Documentation* 4, 378-390
- 24 Dym, E. D. (ed.) (1985): Subject and information analysis. New York etc.: Dekker
- 25 Earl, L. L. (1970): Experiments in automatic extracting and indexing. *Information Storage and Retrieval* 6 (4) 313-334
- 26 Edmundson, H. P. (1969): New methods in automatic extracting. *Journal of the ACM* 16 (2) 264-285
- 27 Edmundson, H. P.; Wyllys, R. E. (1961): Automatic abstracting and indexing – survey and recommendations. *Communications of the ACM* 4, 226-234
- 28 Endres-Niggemeyer, B. (1985): Referierregeln und Referate – Abstracting als regelgesteuerter Textverarbeitungsprozess. *Nachrichten für Dokumentation* 36 (1) 38-50
- 29 Endres-Niggemeyer, B. (1988): Informationsorientiertes Schreiben oder die Produktion von textuell dargestelltem Wissen. *LDV-Forum* 5 (4) 21-37
- 30 Endres-Niggemeyer, B. (1988): Modelling content analysis as a task-oriented writing process. *Proceedings, Writing – an international workshop. Padova* 1.-2. Dez.
- 31 Endres-Niggemeyer, B. (1998): Summarizing Information. Berlin: Springer 1998 (Habilitationsschrift, Universität Konstanz: Abstrahieren, Indexieren und Klassieren. Ein empirisches Prozeßmodell der Dokumentrepräsentation)
- 32 Endres-Niggemeyer, B. (2000): SimSum: an empirically founded simulation of summarizing. *Information Processing & Management* 36 (4) 659-682
- 33 Endres-Niggemeyer, B.; Hobbs, J.; Sparck Jones, K. (eds.) (1995): Summarizing text for intelligent communication. Dagstuhl Seminar, 13.12.-17.12.1993. Dagstuhl-Seminar-Report, 79. Schloß Dagstuhl 1995
- 34 Endres-Niggemeyer, B.; Elisabeth Neugebauer (1998): Professional summarizing: No cognitive simulation without observation. *JASIS* 49 (6) S. 486-506
- 35 Faigley, L.; Chery, R. D.; Jolliffe, D. A.; Skinner, A. M. (1985): Assessing writer's knowledge and processes of composing. Norwood, NJ: Ablex
- 36 Fidel, R. (1986): Writing abstracts for free-text searching. *Journal of Documentation* 42 (1) 11-21
- 37 Fum, D.; Guida, G.; Tasso, C. (1986): Forward and backward reasoning in automatic abstracting. In: J. Horecky (ed.): COLING-82: Proceedings of the 9th International Conference on Computational Linguistics, Prague, 1982. Prague: Akademia, 83-88
- 38 Fum, D.; Guida, G.; Tasso, C. (1985): Evaluating importance: a step towards text summarization. In: IJCAI-85: Proceedings of the 9th International Joint Conference on Artificial Intelligence. Los Angeles, CA, USA, August 18-23, 1985. Los Altos, CA: Kaufmann, 840-844
- 39 Gaus, W. (1983): Referate. Lektion 6 von: Dokumentations- und Ordnungslehre. Berlin etc.: Springer-Verlag, 43-47
- 40 Hahn, U. (1986): Methoden der Volltextverarbeitung in Informationssystemen. Ein State-of-the-Art-Bericht. In *Lit.* 63, 195-216
- 41 Hahn, U.; Inderjeet, M. (2000): The challenges of automatic summarization. *Computer* 33 (11) 29-36
- 42 Hahn, U.; Reimer, U. (1986): TOPIC essentials. In: COLING-86: Proceedings of the International Conference on Computational Linguistics. Bonn, FRG, August 25-29, 1986, 497-503
- 43 Hahn, U.; Reimer, U. (1986): Semantic parsing and summarizing of technical texts in the TOPIC system. In *Lit.* 63, 153-193
- 44 Hahn, U.; Reimer, U. (1999): Knowledge-based text summarization: Saliency and generalization operators for knowledge base abstraction. In: Inderjeet Mani; Mark T. Maybury (Eds.), *Advances in Automatic Text Summarization*. Cambridge/MA, London/England: MIT Press, S. 215-232
- 45 Hammwöhner, R.; Kuhlen, R.; Sonnenberger, G.; Thiel, U. (1989): TWRM/TOPOGRAPHIC. Ein wissensbasiertes System zur situationsgerechten Aufbereitung und Präsentation von Textinformation in graphischen Retrievaldialogen. *Informatik. Forschung und Entwicklung* 4 (2) S. 89-10
- 46 Harbeck, R.; Lutterbeck, E. (1986): Inhaltsangaben in der Dokumentation. *Nachrichten für Dokumentation* 19, 2, 15-18

- 47 Hayes, J. R.; Flower, L. S. (1980): Writing as problem solving. *Visible Language* 14, 4, 388-399
- 48 Helgerson, L. W. (1986): CD-ROM search and retrieval software: the requirements and realities. *Library Hi Tech* 4 (4) 69-77
- 49 Hobbs, J. R. (1982): Towards an understanding of coherence in discourse. In: W. G. Lehnert; M. H. Ringle (eds): *Strategies for natural language processing*. Hillsdale, NJ etc.: Erlbaum, 223-243
- 50 Hutchins, W. J. (1977): On the problem of 'aboutness' in document analysis. *Journal of Informatics* 1 (1) 17-35
- 51 Hutchins, W. J. (1987): Summarization: some problems and methods. In: Jones, K. P. (ed.): *Informatics 9, Proc. of a Conference*. King's College Cambridge, March 26-27, 1987, London: Aslib, 151-173
- 52 IRRD (1982): Working rules. *International Road Research Documentation (IRRD)*, Chap. IV.2: Preparation of abstracts. January, 18-20
- 53 ISO 214 (1976): International Organisation for Standardisation: Documentation: Abstracts for publication and documentation. Genf: ISO
- 54 Jonak, Z. (1984): Automatic indexing of full texts. *Information and Management* 20 (5/6) 619-627
- 55 Karlgren, H.; Walker, D. E. (1983): The polytext system: a new design for a text retrieval system. In: F. Kiefer (ed.): *Questions and Answers*. Dordrecht etc., 273-294 (= *Linguistic Calculation* 1)
- 56 Kempen, G. (ed.) (1987): *Natural language generation. New results in artificial intelligence, psychology and linguistics*. Dordrecht: Nijhoff
- 57 Kennedy, A. (1983): On looking into space. In: Rayner, K. (ed.): *Eye movement in reading perceptual and language processing*. New York etc., 237-250
- 58 Kintsch, W.; van Dijk, T. A. (1983): *Strategies of discourse comprehension*. Orlando, FLA: Academic Press
- 59 Koblitz, J. (1972): *Referieren von Informationsquellen*. Lehrbrief. Berlin: Institut für Bibliothekswissenschaft und wissenschaftliche Information
- 60 Kuhlen, R. (1983): Volltextanalyse zum Zwecke des Abstracting. In: Deutsche Gesellschaft für Dokumentation (DGD) (ed.): *Deutscher Dokumentartag 1982: Fachinformation im Zeitalter der Informationsindustrie*. 27.-30.9.1982, Lübeck-Travemünde. München etc.: Saur KG, 300-312
- 61 Kuhlen, R. (1984/1986): Some similarities and differences between intellectual and machine text understanding for the purpose of abstracting. In: H. J. Dietschmann (ed.): *Representation and exchange of knowledge as a basis of information processes*. Proceedings of IRFIS 5 Heidelberg 5.-7. Sept. 1983. Amsterdam u.a., North-Holland, 87-109 (auch in *Lit.* 63, 133-151)
- 62 Kuhlen, R. (1985): Verarbeitung von Daten, Repräsentation von Wissen, Erarbeitung von Information. Primat der Pragmatik bei informationeller Sprachverarbeitung. In: B. Endres-Niggemeyer; J. Krause (eds): *Sprachverarbeitung in Information und Dokumentation*. Proceedings der Jahrestagung der Gesellschaft für Linguistische Datenverarbeitung (GLDV) in Kooperation mit der Fachgruppe 3 „Natürlichsprachliche Systeme“ im FA 1.2 der Gesellschaft für Informatik (01). Hannover, 5.-7. März, 1985. Berlin etc.: Springer, I-22 (= *Informatik Fachbericht* 114; Subreihe KI)
- 63 Kuhlen, R. (1986): *Informationslinguistik. Theoretische, experimentelle, curriculare und prognostische Aspekte einer informationswissenschaftlichen Teildisziplin*. Tübingen: Niemeyer
- 64 Kuhlen, R. (1989): Information Retrieval: Verfahren des Abstracting. In: Batori, I. S.; Lenders, W.; Putschke (eds.): *Computational linguistics. An international handbook on computer oriented language research and applications*. Berlin etc.: de Gruyter
- 65 Kuhlen, R. (1989): Graphische und textuelle Erzeugung informationeller Objekte als ein Beitrag zur Flexibilisierung und Individualisierung von Information. *Nachrichten für Dokumentation*
- 66 Kupiec, J.; Pedersen, J.; Chen, F.: A trainable document summarizer. Proceedings of the 18th annual international ACM SIGIR Conference on Research and Development in Information Retrieval. July 09-13, 1995, Seattle, Washington, United States, S. 68-73
- 67 Lancaster, F. W. (2003): *Indexing & Abstracting in Theory & Practice*. University of Illinois Graduate School
- 68 Lehnert, W. G.; Black, J. B.; Reiser, B. J. (1981): Summarizing narratives. In: *IJCAI-81: Proceedings of the 7th International Joint Conference on Artificial Intelligence*. Vancouver, Canada. August 24-28, 1981, Vol I. Menlo Park, CA: AAAI, 184-189
- 69 Luhn, H. P. (1958): The automatic creation of literature abstracts. *IBM Journal* April, 159-163
- 70 McGirr, C. J. (1978): Guidelines for abstracting. *Technical Communication* 2, 2-5
- 71 McKeown, K. R. (1986): Language generation. Applications, issues, and approaches. In: *Proceedings of the IEEE* 74 (7) 961-968
- 72 Maizell, R. E.; Smith, J. F.; Singer, T. E. R. (1971): *Abstracting scientific and technical literature*. New York: Wiley-Interscience (ISA 72-2147)
- 73 Mann, W. C. (1984): Discourse structures for text generation. In: *COLING-84: Proceedings of the 10th International Conference on Computational Linguistics*. Stanford, CA, 367-375
- 74 Mann, W. C.; Thompson, S. A. (1987): Rhetorical structure theory: description and construction of text structures. In *Lit.* 56, 85-96

- 75 Maron, M. E. (1982): Problems with full-text searching. In: Office Automation Conference, April 5-7. Moscone Center, San Francisco, 379-384
- 76 Mathis, B. (1972): Techniques for the evaluation and improvement of computer-produced abstracts. OSU-CISRC-TR-79-15. The Computer and Information Science Research Center, The Ohio State University, Columbus
- 77 Mathis, B. A.; Rush, J. E. (1985/1975): Abstracting. The nature and definition of abstracting. In *Lit.* 24, 445-484
- 78 Molitor, S. (1984): Kognitive Prozesse beim Schreiben. DIF (Deutsches Institut für Fernstudien). Tübingen: Universität Tübingen, Forschungsbericht 31
- 79 O'Connor, B. C. (1996): Explorations in Indexing and Abstracting: Pointing, Virtue, and Power. Libraries Unlimited Inc.
- 80 PADOK (1986): PADOK: Test und Vergleich von Texterschließungssystemen für das Deutsche Patent- und Fachinformationssystem. Endbericht. 1.1.1985-31.3.1986. Regensburg: Univ. Regensburg, Linguistische Informationswissenschaft
- 81 Pfeiffer-Jäger, G. (1980): Referat und Referieren: Linguistische Beiträge zu ihrer Applikation in der Information und Dokumentation. *Germanistische Linguistik* 1/2, 1-180
- 82 Rath, G. J.; Resnick, A.; Savage, T. R. (1961): The formation of abstracts by the selection of sentences. Part 1: Sentence selection by men and machines. *American Documentation* 12, 139-141
- 83 Reimer, U.; Hahn, U. (1988): Text condensation as knowledge base abstraction. In: Proceedings of the 4th IEEE Conference on Artificial Intelligence Applications, San Diego, CA. Washington, D. C. Comp. Soc. of the IEEE, 338-344
- 84 Resnick, A. (1961): Relative Effectiveness of document titles and abstracts for determining relevance of documents. *Science* 134, 1004-1005
- 85 Rösner, D. (1987): The automated news agency: SEMTEX a text generator for German. In *Lit.* 56, 133-148
- 86 Rowley, J. (1988; 1989): Abstracting and indexing. London: Clive Bingley (2. Auflage 1989)
- 87 Schwartz, C.; Eisenmann, L. M. (1986): Subject analysis. In: M. E. Williams (ed.): *Annual Review of Information Science and Technology (ARIST)* 21, 37-61
- 88 Skolnik, H. (1979): Historical development of abstracting. *Journal of Chemical Information and Computer Science* 19 (4) 215-218
- 89 Sonnenberger, G. (1988): Flexible Generierung von natürlichsprachigen Abstracts aus Textrepräsentationsstrukturen. In: Wiener Workshop „Wissensbasierte Sprachverarbeitung“. 4. Österreichische Artificial Intelligence Tagung, Wien 29.-31. August 1988. Berlin etc.: Springer, 72-82 (= Informatik Fachberichte 176, Subreihe KI)
- 90 Stibic, V. (1985): Printed versus displayed information. In: *Nachrichten für Dokumentation* 36 (4/5) 172-178
- 91 Strong, S. M. (1974): An algorithm for generating structural surrogates of English text. *Journal of the ASIS* 1, 10-24
- 92 Tait, J. (1985): Generating summaries using a script-based language analyzer. Cambridge: Acorn Computers
- 93 Taylor, S. L.; Krulee, K. (1977): Experiments with an automatic abstracting system. *Proceedings of the ASIS*. Vol. 14. White Plains, NY: Knowledge Industry Publ.
- 94 Tenopir, C. (1985): Full-Text databases. In: M. E. Williams (ed.): *Annual Review of Information Science and Technology (ARIST)* 20, 215-246
- 95 Teufel, S.; Moens, M. (2002): Summarizing scientific articles: experiments with relevance and rhetorical status. *Computational Linguistics* 28 (4) S. 409-445
- 96 Travis, I. L.; Fidel, R. (1982): Subject analysis. In: M. E. Williams (ed.): *Annual Review of Information Science and Technology (ARIST)* 17, 123-157
- 97 Varun, V. K.: Automatic Abstracting & Summarizing Tools – <http://itt.nissat.tripod.com/itt0202/ruoi0202.htm>
- 98 Vickery, B. C.; Vickery, A. (1987): Information science in theory and practice. London etc.: Butterworths
- 99 Wagers, R. (1983): Effective searching in database abstracts. *Online* 7 (5) 60-77
- 100 Weil, B. H.; Zarembek, I.; Owen, H. (1963): Technical-abstracting fundamentals 11. Writing principles and practices. *Journal of Documentation* 3 (2) 125-132
- 101 Wellish, H. H. (1980): Indexing and abstracting. An international bibliography. Santa Barbara, CA: ABC-Clio.
- 102 Wellish, H. H. (1984): Indexing and abstracting: an international bibliography. Teil 2 (für den Zeitraum 1977-1981). Santa Barbara, CA: ABC-Clio.
- 103 Wyllis, R. E. (1967): Extracting and abstracting by computer. In: Borko, H. (ed.): *Automated Language Processing*. New York etc., 127-179
- 104 Zechner, K. (1995): Automatic text abstracting by selecting relevant passages. Master's thesis, Centre for Cognitive Science, University of Edinburgh, Edinburgh