

# Von der Literaturdokumentation zum Indexing

## ***Ein Interview mit Dr. Robert Fugmann zu seinen Erfahrungen und Einsichten einer langen Beschäftigung mit der Inhaltserschließung von Dokumenten***

Am 24. Juni 2013 besuchten Angelika Fallert-Müller und Walter Greulich im Auftrag des DNI Herrn Dr. Robert Fugmann an dessen Wohnort in Idstein im Taunus, um von ihm, der sich seit Mitte der 1950er-Jahre mit der Informationserschließung von Dokumenten beschäftigt, zu erfahren, wie er mit dem Thema Indexing in Berührung kam und was ihn daran besonders fasziniert. Auch um seine Einschätzung der Indexing-Situation in Deutschland wurde er gebeten.

Robert Fugmann war in seinem aktiven Berufsleben Leiter der zentralen wissenschaftlichen Dokumentationsabteilung der Hoechst AG sowie Lehrbeauftragter und Gastprofessor an mehreren Fachhochschulen und Universitäten im In- und Ausland. Im informationswissenschaftlichen Kontext wurden ihm zahlreiche Ehrungen zuteil, die deutlich machen, mit welchem Engagement er sich den Themen Klassifizierung und Indexing gewidmet hat. Unter anderem hat er erhalten: die Gmelin-Beilstein-Gedenkmünze der Gesellschaft Deutscher Chemiker, den Herman-Skolnik-Preis der American Chemical Society und den Ranganathan Award for Classification Research 1994. Ferner ist er Autor des „Best J. ASIS Paper 1985“ der American Society for Information Science.

Über Indexing und Klassifikationstheorie hat er mehrere Bücher sowie zahlreiche Artikel in Fachzeitschriften veröffentlicht.

### ***1 Wie sind Sie als Chemiker zum Thema Indexing gekommen?***

Das Thema ist mir sozusagen in die Wiege gelegt worden. Mein Vater war im Ersten Weltkrieg Meldereiter, also wenn man so will Informationsüberträger. Und ich war im Zweiten Weltkrieg Soldat bei der Nachrichtentruppe der Fallschirmtruppe. Später habe ich dann in Jena, Würzburg und Hamburg Chemie studiert. Nach dem Studium bin ich zur Hoechst AG

(damals noch Farbwerke Hoechst AG) in Frankfurt gegangen und war in der Pharmaforschung tätig.

Hier musste ich eines Tages ein Arzneimittel herstellen, für das ich als Vorprodukt Glyoxylsäure benötigte, welches ich ebenfalls selbst erzeugen musste. Das war, auch wenn es sich bei Glyoxylsäure um eine einfache Verbindung handelt, recht aufwendig. Nachdem ich also mühsam eine kleine Menge gewonnen hatte, sah ich zufällig bei meinem Kollegen im Nachbarlabor eine 5-Liter-Chemikalienflasche, fast voll mit Glyoxylsäure. Auf meine Frage, wo er die her hat, meinte er, dass sie im Technikum hergestellt werde. Ich ging also zum Technikum, der Techniker zog einen Auftragblock heraus und fragte, wie viele Fässer ich denn gern hätte. Als ich das meinem Vorgesetzten erzählte, meinte er, das komme vor und sagte: „Machen Sie etwas dagegen!“ Das wichtigste war in meinen Augen, dass man die Chemikalien, die es in einem Werk gibt, auch finden kann. So habe ich mit Lochkarten angefangen, Chemikalien zu dokumentieren. Im Jahr 1956 wurde bei Hoechst der erste große elektronische Rechner, der für damalige Zeiten unglaublich leistungsfähige IBM 705, installiert, den wir nutzten. Es entstand das erste Schlagwortvokabular, das wir GREMAS nannten (**G**enealogisches **R**echerchieren durch **M**agnetbandspeicherung). Das System wurde so erfolgreich, dass nicht nur die werksinternen Stoffe in unserer Dokumentationsabteilung damit erfasst werden sollten, sondern auch die Literatur. Bis zu 14 Chemiefirmen arbeiteten zusammen bei der Literaturerfassung und es wurde ein Institut, das IDC (Internationale Dokumentation für Chemie), gegründet, das in Sulzbach/Taunus etabliert war und in dem zeitweilig bis zu 30 Chemiker tätig waren.

Bei Hoechst gab es große Unterstützung für GREMAS durch den Vorstandsvorsitzenden Winnacker sowie den Forschungsleiter Schultheis. Die angestellten Chemiker mussten nur dazu gebracht werden, das System zu nutzen. So wurden z. B. auch die Jahresberichte der Chemiker dokumentiert, und die Chemiker haben gemerkt, dass es jemanden gibt, bei dem man nachfragen konnte. Es kam vor, dass ein Patent angemeldet wurde und niemand wusste,

dass das Patent bereits vergeben war. Der Forschungsleiter drang daher darauf, dass niemand mehr ein Patent anmeldete, ohne vorher bei der Dokumentationsabteilung angefragt zu haben. Wir standen aber immer unter Druck, einen Effizienznachweis zu bringen. Dabei liegt die Wichtigkeit der Dokumentation auf der Hand: Eine Produktionsstätte ohne Dokumentation ist wie eine Fabrik, die Waren produziert und wenn sie sie produziert hat, nicht mehr wiederfinden kann und wieder von vorne anfangen muss. Der wirtschaftliche Druck muss groß genug sein, damit der Sinn der Dokumentation erkannt wird.

Anfangs habe ich mich um GREMAS nur neben der Labortätigkeit kümmern können und es stand auf Messers Schneide, ob ich Dokumentation betreiben würde oder Forschung und Entwicklung. Irgendwann wurde es zu viel, sodass ich mich entscheiden musste, und ich habe mich für die Dokumentation entschieden (1960).

Das Hineinfinden in das Thema Indexing hat sich allmählich entwickelt. Im Grunde ist der Index zum Buch „Theoretische Grundlagen der Indexierungspraxis“ (1993) der erste eigentliche Buchindex von mir gewesen. Bis dahin hatte sich allerdings so viel an Überlegungen zum Thema Indexing angesammelt, dass dies zu Papier gebracht werden musste. Und das Buch dazu bot den Einstieg in das praktische Indexieren.

1989 wurde die International Society for Knowledge Organization ISKO gegründet, in der das Buchindexieren immer wieder eine Rolle spielte. Die Interessengruppe Inhaltserschließung habe ich eine Zeitlang geleitet, und in diesem Zusammenhang ist das Buch entstanden.

Zum Thema Inhaltserschließung habe ich seit dem Aufbau des Dokumentationssystems GREMAS die Literatur aufmerksam verfolgt. So hat mich das Buch von B.C. Vickery „Classification and indexing in sciences“ (1958) sozusagen auf die Spur gebracht. Vickery und Ranganathan, die ich auch persönlich kennenlernen konnte, haben die theoretische Grundlage für das GREMAS-System geliefert.

In den 1950er- und 1960er-Jahren sind (von verschiedenen Autoren) grundlegende Arbeiten zur Inhaltserschließung erschienen, darunter auch viel Philosophisches. So ist z. B. der Ordnungsbegriff definiert worden:

- Driesch: „Ordnung ist das sinnvolle Beieinandersein der Teile eines Ganzen“ (was ich ergänzt habe durch „an vorhersehbarer Stelle“).
- „Information ist jede Nachricht, die sich für Empfänger als interessant erweist“ (frei nach dem russischen Informationswissenschaftler Vladuz 1960).

Ordnung ist das oberste Gebot in der Dokumentation. Ordnung ist subjektiv. Daraus folgt die Unmöglichkeit der Automatisierung der Indexierung, wenn man mehr als nur stümperhafte Ergebnisse erreichen möchte.

Um das mit einem Beispiel zu illustrieren: Nehmen wir an, ich halte mich auf dem Bahnsteig auf. „Der Zug nach München fährt heute ausnahmsweise nicht auf Gleis 3, sondern auf Gleis 5 ein.“ Wenn ich das zum ersten Mal höre, ist es für mich Information; wenn ich es zum zweiten Mal höre, ist es keine Information mehr und wenn ich es im Bett im Bahnhofshotel höre, ist es überhaupt keine Information, und zwar bereits von Anfang an keine.

Wichtig ist auch die Erkenntnis: Was man in der Literatur und im Internet findet, sind Nachrichten. Ob aus der Nachricht Information wird, hängt von der Bedarfslage und dem Wissensstand des Empfängers ab. Es gibt keine objektive Information. Auch Information ist subjektiv.

## ***2 Welche Typen von Indexen haben Sie erstellt und weshalb gerade diese Typen?***

Zu meinen eigenen Büchern habe ich nicht nur alphabetische, sondern auch systematische Register erstellt.

Ein systematisches Register ist immer subjektiv, d.h., was verwandte Begriffe sind, ob und wie sie miteinander verwandt sind, schätzt jeder Indexer auf seine Weise ein. Das systemati-

sche Register dient mir bei der Arbeit, weil ich dadurch einen guten Überblick über das Vokabular habe. Dabei steht immer die Suche nach dem am besten passenden Schlagwort im Vordergrund („Cutter’s rule“, 1876). Denn alles was wichtig ist, sollte auffindbar sein. Das wird leider oft außer Betracht gelassen. Das bestpassende Schlagwort zu benutzen, gelingt einem nur dann, wenn man einen guten Überblick über sein Register hat. Dazu dient die Systematik. Der Leser wird vermutlich meistens das alphabetische Register benutzen, evtl. noch ein Schlagwortverzeichnis, in dem die Begriffszusammenhänge in den Grundzügen dargelegt sind. Zu einem guten Register gehören Untereinträge, deren Formulierung aber nicht programmierbar ist.

Vorteil der Untereinträge für den Leser: Er kann leicht feststellen, ob er das gesuchte Schlagwort in dem gewünschten Zusammenhang in diesem Buch finden kann. Das beschleunigt den Zugriff.

Die Untereinträge sind dabei normalerweise weder alphabetisch noch chronologisch geordnet. Ordnung ist erst notwendig, wenn sich so viele Untereinträge angesammelt haben, dass sie unübersichtlich werden. Dann werden mit Hilfe eines Pseudoschlagworts Untergruppen nach bestimmten Unterteilungsgesichtspunkten gebildet und die entsprechenden Unterschlagworte nach dem Rang einsortiert. Beispiele für Pseudoschlagworte: „unterteilt nach Herkunft“, „unterteilt nach Material“, „unterteilt nach Wirkung“ usw. Anhand des Pseudoschlagwort erkennt der Nutzer, ob sein gesuchter Zusammenhang dargestellt wird (Beispiel: Pestizide / natürlich vorkommende Pestizide).

Ein systematisches Register entsteht in mehreren Durchgängen. Im ersten Durchgang der Registererstellung kommt das Schlagwortvokabular zusammen. Tauchen bestimmte Schlagworte erst am Ende des Durchgangs auf, verlangen diese beim zweiten Durchgang erhöhte Aufmerksamkeit. Zu jedem Schlagwort werden dann die zugehörigen Ober- und Unterbegriffe aufgenommen, ebenso Nicht-Deskriptoren. Das sind Stichwörter, unter denen der Nutzer auch suchen könnte (und von denen auf passende Schlagwörter verwiesen werden muss).

Wichtig bei einem systematischen Register sind sein hierarchischer Aufbau und das dazu passende Layout.

Als großer Vorteil erweist es sich, nicht mit Seitenzahlen zu arbeiten, sondern die Absätze zu nummerieren und auf die Absatznummern zu verweisen. Die Nummern werden auf dem Rand neben den Absätzen abgedruckt. Die Absatznummerierung sollte dezimal sein, weil der Autor dann nachträglich beliebig Absätze einschieben und der Indexer für die neuen Absätze Nummern vergeben kann, ohne dass das bereits bestehende Nummernsystem davon beeinflusst wird. Da die Nummerierung fortlaufend ist, können Lücken entstehen, wenn Absätze verschoben oder gelöscht werden. Absatznummern haben für den Leser den Vorteil, dass er gezielter hingeführt wird. Allerdings muss ihm ein solches System erläutert werden.

Hilfreich war dabei eine Software, die ich seit langem zur Erzeugung systematischer Register verwende: LIDOS von der Firma LAND. Dieses Programm hat bereits Mitte der 1980er Jahre ein Thesaurus-Modul gehabt, mit dem sich z. B. auch das Problem der Polyhierarchie lösen ließ. Polyhierarchie bedeutet, dass ein Begriff mehrere Oberbegriffe hat. Nehmen wir z.B. Papier: es kann ein Papier im DIN-A-4-Format sein, Papier von 60 Gramm, Papier, das schwarz auf weiß bedruckt ist. Polyhierarchie ist nicht einfach zu handhaben, ihre Verwaltung hat Ähnlichkeiten mit relationalen Datenbanken, man betrachtet aber Beziehungen zwischen Begriffen, nicht zwischen Datensätzen. Dem Leser kann eine Polyhierarchie in einem systematischen Register dadurch deutlich gemacht werden, dass unter jedem Unterbegriff mehrere zugehörige Oberbegriffe erscheinen – und umgekehrt.

Beim Register zu dem Büchlein von Schlemper aus dem Bereich Quantenphysik habe ich ein Verfahren angewandt, das ich in meinem Buch „Das Buchregister: Methodische Grundlagen und praktische Anwendung“ beschrieben habe: Entstanden sind ein alphabetisches Schlagwortregister und ein „informatives“ Register. Von der Indexingtechnik her wurde ebenfalls nicht mit Seiten- sondern mit Absatznummern gearbeitet. Darüber hinaus wurden die Register nicht gedruckt, sondern ins Internet gestellt.

Eine Besonderheit des informativen Registers, die dem Nutzer häufig als Erstes auffällt, ist die Länge und die Vielzahl der Untereinträge. Sie bestehen im Gegensatz zu einem herkömmlichen Schlagwortregister nicht nur aus zwei bis vier Worten. Wir sind ja aus der Zeit heraus, in der man mit dem Schuhkarton und der Karteikarte Register machte, der Platz knapp war und der Verleger sagte, dass das Register nicht mehr als 3 % des Buchumfangs beanspruchen darf. Diese Einschränkungen bestehen alle nicht mehr, wenn man das Register elektronisch veröffentlicht, also ins Internet stellt. Ein weiterer Vorteil ist, dass weder Verlag noch Indexer unter Zeitdruck stehen, denn das Buch kann ja längst erschienen sein, bevor das Register fertig ist. Die Qualität des Registers wird besser, weil man viel gründlicher arbeiten kann. Herr Schlemper lobte das Register: „Jetzt gehen mir erst einmal die Augen auf, was in meinem Buch alles drinsteht.“

***3 Was sind Ihre persönlichen Lieblingstitel in der Indexing-Fachliteratur, die Sie anderen empfehlen würden? Und welche Erkenntnisse aus dem Studium von Indexing-Literatur sind für Sie besonders wichtig?***

*Zeitschriften:*

- Journal of the American Society of Information Science
- Journal of Documentation
- Nachrichten für Dokumentation
- Zeitschrift der Deutschen Gesellschaft für Informationswissenschaft und Praxis
- Knowledge Organization (Fachblatt der ISKO)
- indische Fachzeitschrift, die der indische Dokumentationspapst S.R. Ranganathan [Großvater des WDR-Redakteurs Ranga Yogeshwar] herausgegeben hat.

*Bücher:*

- Wellisch
- Bar-Hillel (israelischer Linguist)
- Vickery
- Ranganathan

Ranganathan betonte: Die Grammatik ist wichtig, nicht nur der Wortschatz!

GREMAS für die Chemie kann als Beispiel genommen werden: Den Wortschatz bilden die Atome, die Grammatik sind die Atomverknüpfungen zum Molekül. So können etwa Kohlenstoffatome vom Aldehydtyp mit „L..“ benannt werden, die Kohlenstoffatome vom Carboxyltyp mit „N..“. Das Aldehyd-Kohlenstoffatom selbst heißt dann „LH.“, das Carboxylkohlenstoffatom selbst „NN.“. An dritter Stelle steht dann immer, ob es Glied einer Kohlenstoffkette ist oder Substituent an einem Ring, „A“ an dieser Stelle bedeutet immer Kettenglied, „LHA“ oder „NNA“ sind also Kettenglieder. Dies ist der Wortschatz. In dem Grammatik-Term für die Kette kommt in abgekürzter Form zum Ausdruck, welche Kohlenstoffatome in ein und derselben Kette vorkommen und welche dort nicht vorkommen. „Kette LL“ bedeutet, dass dort zwei Kohlenstoffatome vom Aldehydtyp vorkommen. „Kette LLNNN“ würde bedeuten, dass dort zwei Kohlenstoffatome vom Aldehydtyp und drei Kohlenstoffatome vom Carboxyltyp vorkommen und keine weiteren Kohlenstoffatome anderen Typs. Erst durch diese Grammatik werden die Beschreibung genügend wiedergabegetreu und die Recherche genügend treffsicher.

Die Wissenschaft der Verknüpfung zwischen Gegenständen ist ein Fachgebiet der Mathematik: die Topologie. In der Dokumentationsabteilung von Hoechst waren zwei Mathematiker beschäftigt, die ein topologisches System entwickelt haben: TOSAR (topologische Darstellung von synthetischen und analytischen Relationen). Es stellt eine Art Grammatik für die Verknüpfung von Gegenständen allgemeiner Art zur Verfügung. Das war besonders hilfreich

zur Herstellung von Polymeren, denn bei ihnen lässt sich gut angeben, wie sie aus Monomeren (beispielsweise entweder aus Butadien oder aus Styrol) aufgebaut werden.

Allgemeine Verknüpfungen bei Polymerverbindungen könnten z.B. folgende sein: Stoffe A und B werden mit C oder D zusammengebracht zu einem Polymer, das nun mit einem Weichmacher versetzt wird, und es entsteht ein neuer Stoff. Der wird bei 300° zu einem Faden gestreckt (in Spinnmaschine) und man erhält einen Polymerfaden. Der wiederum wird bei 1000° verkohlt, es ergibt sich ein Kohlenstofffaden. Dieser wird zu einem Gewebe gewebt, und das Gewebe wird in der Automobilindustrie oder Medizin oder woanders weiterverwendet. Das alles hat TOSAR festgehalten. Es ist auch gut geeignet, in Patentanmeldungen Zusammenhänge deutlich zu machen. Ich habe das Programm sogar gelegentlich verwendet, um komplizierte verwandtschaftliche Verhältnisse/Beziehungen zu verdeutlichen.

Übertragen auf das Indexieren bedeutet das: Begriffe und die Angabe der Eigenschaft „Oberbegriff“ oder „Unterbegriff“ gehören zum Wortschatz. Die Grammatik wäre, dass die Begriffe in ein und derselben Publikation vorkommen, oder in ein und demselben Absatz. Dabei sollte unterschieden werden, ob Begriffe mit Und- oder Oder-Verknüpfung vorkommen. Es wäre Aufgabe der Grammatik, solche Zusammenhänge zum Ausdruck zu bringen. Im GREMAS-System, in das bereits früh Grammatik-Verfahren eingeführt wurden, hat sich der Aufwand der Indexierung durch die Treffsicherheit bei der Recherche bezahlt gemacht.

Es besteht nämlich ein enger Zusammenhang zwischen dem Aufwand für die Indexierung einerseits und für die Recherche andererseits, auch in Abhängigkeit von der Zeit, die ein Informationssystem in Betrieb ist. Niedrig ist der Indexierungsaufwand dann, wenn man sich mit dem Wortschatz der Schlagwörter begnügt. Am Anfang läuft alles gut. Aber wenn die Recherchen zahlreicher werden und wenn der Speicher immer größer wird, wächst der Ballast an unzutreffenden Antworten, und es wächst auch der unbemerkte Verlust an Unaufgefundenem. Schließlich wird die Notbremse gezogen, und es muss ein Neuanfang gefunden werden. Wenn aber von Anfang an auch mit einer Grammatik der Schlagwörter indexiert wird, wenn

auch mit größerem Aufwand, dann bleiben Ballast und Verlust dauerhaft klein, und der größere Aufwand beim Indexieren lohnt sich auf längere Sicht. Allerdings wird der Wert dieses höheren Indexierungsaufwandes allgemein erst zu spät erkannt.

***4 Haben Sie irgendeine Idee, weshalb das Bewusstsein über die Bedeutung von Registern und die Fachkenntnisse über Registererstellung im deutschsprachigen Publikationswesen viel schwächer ausgeprägt sind als im anglo-amerikanischen Bereich?***

Deutschland ist eine Indexierungswüste, da auf Seiten der Verlage das Wissen um den Nutzen einer guten Indexierung sehr schwach ausgeprägt ist. Ansonsten habe ich keine Erklärung für das Phänomen, ich kann es nur mit Bedauern zur Kenntnis nehmen.

***5 Was müsste aus Ihrer Sicht im deutschen Publikationswesen grundsätzlich geschehen, damit sich die Situation verbessert, damit also mehr und bessere Register erstellt werden?***

Die Leute müssten stärker informiert werden, worin das Wesen der Indexierung besteht und welchen Nutzen man davon hat. Das ist das, was das DNI bereits macht. Dem Management muss gezeigt werden, dass Indexieren nicht einfach ist.

Bei Vorträgen habe ich manchmal folgende Anekdote zu Gehör gebracht:

Im Ministerium sollte eine Akte über die Einweihung einer Brücke gesucht werden, aber der zuständige Staatssekretär war in Urlaub. Unter E wie Einweihung war nichts zu finden, unter B wie Brücke ebenfalls nicht, auch nichts unter dem Namen des Flusses oder der Stadt. Schließlich wurde ein Telegramm an den Staatssekretär geschickt. Er kam, ging an seinen Aktenschrank – ein Griff, und er hatte die Akte. Frage: Wie konnten sie so schnell die Akte finden, worunter haben sie sie abgelegt gehabt? Ganz klar: unter H wie halb zwei, die Brücke

wurde halb zwei eingeweiht! Er hatte sich die Arbeit bei der Erfassung seiner Dokumente gar zu einfach gemacht.

In den 1960er-Jahren wurden die legendären ASLIB Cranfield-Experimente durchgeführt (Cyril Cleverdon), bei denen herauskam, dass man keine Indexierung braucht und alles per Volltextrecherche erledigen kann. Allerdings ist dabei ein grober Fehler unterlaufen: Die Experimentatoren haben bei abgespeicherten Dokumenten angeguckt, welche Wörter vom jeweiligen Autor benutzt wurden. Wenn man schon weiß, wie sich jemand ausgedrückt hat, dann ist das Rennen gelaufen! Das ist dasselbe wie beim Ostereiersuchen: Derjenige, der die Eier versteckt hat, kann sagen: Sehvermögen ist nicht wichtig. Denn er weiß ja, wo die Eier sind. Ich habe damals einen Aufsatz in den Nachrichten für Dokumentation als Antwort auf die Experimente publiziert, um die Volltextrecherche als „faules Ei des Kolumbus“ zu entlarven. Volltextrecherche ist schnell und billig. Aber nach einiger Zeit, wenn man nicht mehr weiß, wie die Autoren sich ausgedrückt haben, findet man nichts mehr. So lange die Erinnerung da ist, klappt alles wunderbar.

Der Amerikaner Charles L. Bernier hat schon in 1960er-Jahren klar gemacht: Die Entdeckungsrecherche muss beachtet werden! Erinnerungsrecherche klappt nur kurze Zeit.

***6 Was sagen Sie denjenigen, die für digitale Medien die Suchwort-Eingabe statt Register als völlig ausreichend ansehen? Und wie schätzen Sie die Bedeutung von Registern bei E-Books ein? Bei der Suche auf E-Readern (insbesondere Kindle-Geräten) werden dem Nutzer die Fundstellenlisten angeboten.***

Denen sage ich: Sie haben die Augen verschlossen vor der Zukunft, nämlich dass sie nach einiger Zeit nur noch sehr lückenhaft fündig werden können mit ihrer Suchwortmethode. Am Anfang ist alles latent gedächtnisgestützt. Das wird aber erst im Laufe der Zeit offenkundig, wenn man das wieder vergessen hat, was man gelesen hat und auch gar nicht wissen kann,

was man noch nicht gelesen hat. Man will ja auch Begriffe wiederfinden, deren Wortlaut man nicht mehr in Erinnerung hat.

Um noch einmal das Beispiel der Dokumentationsabteilung bei Hoechst zu bringen: Der Forschungsleiter bei Hoechst sagte: Es kommt mir keine Patentanmeldung über den Tisch, die nicht vorher von der Dokumentationsabteilung geprüft worden ist. Viele der angestellten Chemiker meinten: Das brauchen wir alles nicht, wir **machen** ja die Literatur! Sie hatten allerdings die Sicherheit, dass es mit uns eine Stelle gab, die ihre Arbeiten kritisch las und auch die internen Berichte analysierte. Wir wussten, an welchen Themen sie arbeiteten. Und dann haben wir sogar unaufgeforderte Recherchen gemacht und ihnen zu ihren Themen gesagt: Guckt euch das mal an, würde euch das denn nicht interessieren? Ihnen sind die Augen übergegangen, was sie alles nicht gewusst haben. Es gab daraufhin sogar Frage-Abonnements an die Literatur und an die internen Berichte. Auf diese Weise wurden die Leute auf dem Laufenden gehalten. All das war aber nur möglich, weil wir die Dokumente mit hohem Aufwand indexiert haben.

*Eigene Anmerkung: Auch E-Book-Reader könnten Leser von sich aus auf Zusammenhänge zwischen den auf dem Gerät befindlichen E-Books aufmerksam machen nach dem Motto: Wussten Sie schon, dass Sie zu dem Thema, das sie gerade lesen, auch in dem und dem Titel auf Ihrem Gerät etwas zu finden ist? Oder Überhaupt: Nach dem Einschalten könnten die Leser auf bestimmte Themen der Buchinhalte auf ihrem Gerät aufmerksam gemacht werden! Basis dazu könnten gute Indexe sein!*

Ein Register geht ja viel mehr in die Tiefe als das Inhaltsverzeichnis. Das wird uns auch immer entgegengehalten: Das Buch hat doch ein IHV, wozu brauchen wir noch einen Index? Das sind oberflächliche und undurchdachte Einwände. Die Leute stehen der Materie eben sehr fern. Das ist ja letzten Endes fast schon ein philosophisches Gebiet – mit all den Problemen, die damit verbunden sind. Und wenn ein Register elektronisch verfügbar ist, dann müsste man doch diese Möglichkeit viel mehr nutzen können.

*Eigene Anmerkung: Das Problem ist, dass die Indexe zu E-Books im Augenblick meistens einfach nur digital umgesetzte Varianten der gedruckten Indexe sind. Auf einen solchen Gerät gibt es aber andere Bedürfnisse des Lesers, des Nutzers der Indexe. Zum Beispiel gibt es ja keine Seiten mehr. Ein Gerät zeigt ja die Seiten je nach gewählter Schriftgröße. Die Paginierung ist weg.*

Da würde die Paragraphennummerierung einspringen können. Das ist ja überhaupt ein Problem, dass die Paginierung nicht mehr stimmt, wenn man nachträglich einen Absatz einfügt.

Was die Fundstellen-Listen-Anzeige bei E-Books angeht, so hat bereits Eugene Garfield [einer der Begründer der Bibliometrie] in den 1960er-Jahren das Prinzip „keyword in context“ genutzt („KWIC“).

Bei schöngestiger Literatur braucht man natürlich keine Indexierung. Wenn trotzdem in Ausnahmefällen mal ein Index erzeugt wird, dann dient er ausschließlich dem Zweck der Erinnerungsforschung.

**7 Was halten Sie von den Recherchemöglichkeiten, wie sie von großen Fachdatenbanken wie SpringerLink (u. a. Landolt-Börnstein) oder Reaxys (Gmelin-Beilstein) angeboten werden? Beide setzen auf Volltextrecherche. Der einzige Zusatzpunkt, den beide im Bereich Chemie anbieten, ist die Recherche nach Strukturen. Man zeichnet die Struktur und dann wird per Mustersuche gesucht.**

Das haben wir auch bei Hoechst gehabt. Das macht auch Chemical Abstracts. Die Mustersuche ist programmierbar, denn das sind ja determinierte Prozesse. Man kann im Voraus genau sagen, wie die einzelnen Schritte vollzogen werden. Im Gegensatz dazu stehen die indeterminierten Prozesse, die nicht programmierbar sind – eine Aussage, die natürlich kein Informatiker auf sich sitzen lässt.

Aber bei der Volltextsuche mit den Wörtern aus dem Text des Gesuchten steht man vor einem indeterminierten Prozess, der nur behelfsmäßig programmierbar ist: Der Begriff einer Fragestellung kann in einem gesuchten Text nämlich auf unendlich viele verschiedene Arten ausgedrückt worden sein (Ausnahme: Autorennamen). Es ist unmöglich, sie alle in der Fragestellung zu berücksichtigen. Das geht nur für ausgesuchte Beispiele. Entsprechend lückenhaft wird das Rechercheergebnis sein. Aber dieses gibt seine Lückenhaftigkeit nicht zu erkennen. Die Informatiker führen zur Verteidigung der Volltextrecherche gerne einige ganz gut aussehende Beispiele an und verschweigen die Verlustanfälligkeit des Rechercheergebnisses. Wegen der Mehrdeutigkeit der Textwörter ist es auch ballastanfällig.

Ich kann zwar ein Suchwort nennen, z. B. Korrosion, und kann nach Korrosion suchen. Aber finde ich damit auch Aufsätze, in denen es heißt, dass die Schiffsböden so stark dem Angriff des salzhaltigen Meerwassers ausgesetzt sind, dass ganze Teile erneuert werden mussten? Oder dass die Öltanks vor dem Angriff der aggressiven Substanzen durch Magnesium kathodisch geschützt werden müssen? Das ist Korrosionsschutz! Oder: Die Rohre wurden mit PVC ummantelt, damit das Meerwasser keinen Schaden am Metall anrichten kann? Das ist immer wieder Korrosion. Die Tatsache dieser vielen Möglichkeiten wird vernachlässigt.

Bei meiner Vorlesung habe ich immer eine Preisfrage gestellt: Ich suche eine Äußerung eines prominenten Informationswissenschaftlers darüber, dass in Bälde die Notwendigkeit der Dokumentation immer stärker erkannt werden wird und dass es eine Aufwärtsentwicklung gibt. Wie suche ich das in der Literatur? Und in der Pause (halbe bis dreiviertel Stunde) hatten sie die Möglichkeit, die Köpfe zusammenzustecken und die verschiedenen Volltextmöglichkeiten zu skizzieren. Ich habe das dann eingesammelt und am nächsten Tag konnte ich verkünden: Keiner wäre mit seinen Suchwörtern fündig geworden. Keiner! Ich hatte ein Buch als Preis ausgesetzt, und das wurde trotzdem ausgegeben – das bekam derjenige, der sich die meiste Mühe gemacht hatte, möglichst viele Suchalternativen zu nennen. Eine Lösung dieses Such-

problems bei der Volltextrecherche gibt es nicht. Denn was man nicht sieht, gerät außer Betracht. Die Lücken werden nicht gesehen und werden als nicht existent hingenommen.

Es ist ein erkenntnistheoretisches, ja philosophisches Problem, wie der Mensch sich gegenüber dem vorhandenen Wissen verhält, wie er zu diesem Wissen beiträgt und wie er an dem Wissen interessiert ist. Er darf sich nicht einbilden, dass er bei Nichtinteresse kreative Arbeit leisten kann.

Schade ist, dass man Verlagen nicht verständlich machen kann, dass ein Buch, vor allem ein Sachbuch, erheblich an Wert gewinnt, wenn es ein gutes Register hat. „Ein Buch ohne Register ist wie ein Land ohne Landkarte“ hat mal jemand gesagt. Manch einer ist der Meinung, er brauche keine Landkarte, wenn er nie verreist. Doch das kann ja nicht der Standpunkt des nachdenklichen Menschen sein.

***8 Die Deutsche Gesellschaft für Informationswissenschaft und -praxis (DGI) hatte früher (unter dem Namen Deutsche Gesellschaft für Dokumentationswesen, DGD) eine breitere und beständigere Palette an Arbeitsgruppen. Gab es auch eine AG, die sich speziell mit Indexing beschäftigte?***

In dieser Gesellschaft meines Wissens nicht. Wohl aber in der ISKO (International Society for Knowledge Organization), die im Jahr 1989 gegründet wurde. Die dort eingerichtete Interessengruppe für Indexierungsfragen hat Subkonferenzen abgehalten. Auf der Jahrestagung gab es Vorträge zum Thema Indexieren. Ursprünglich gab es die Gesellschaft für Klassifikation und darin den begrifflich orientierten Teil und den technisch orientierten Teil („Numeriker“, die z.B. Häufigkeitsstatistiken gemacht haben und alles, was sich programmieren lässt). Die Techniker haben es geschafft, die Gesellschaft ganz auf die technische Schiene zu bringen. Der begrifflich orientierte Teil ist daraufhin ausgetreten und hat die ISKO gegründet.

**9 Was können Sie aus Ihrer über 50-jährigen Karriere in der Informationswissenschaft, in der das Indexieren stets einer Ihrer großen Schwerpunkte war, heutigen Indexern als Ratschläge mitgeben?**

Die Indexer, genauer die Buchindexer, müssen sich Bücher suchen, die danach verlangen, gut indexiert zu werden, weil sie Fundgruben für Wissen sind – ein Wissen, das erschlossen und zugänglich gemacht werden muss, und das kann nur intellektuell geschehen durch sachkundige, sorgfältige, allerdings auch aufwändige Indexierung. Und davon muss man den jeweiligen Verleger überzeugen. Und der Verleger wird sagen: „Es muss schnell gehen, und dem Register steht nur wenig Platz zur Verfügung.“ Dem muss man entgegenhalten: „Das ist kein Problem, drucken Sie das Buch, wir machen in aller Ruhe ein gutes Register und stellen es im Internet zur Verfügung. Sie müssen nur beim Druck des Buches darauf aufmerksam machen, dass in Kürze ein Register abrufbar sein wird, und zwar unter der und der Internetadresse.“ Auf diese Weise kann man den Einwänden begegnen. Aber der erste Schritt ist natürlich, ein Buch zu finden, bei dem sich das Indexieren anbietet. Volltextsuche im Text kann ja trotzdem möglich sein und schadet nicht (bei einer digitalen Ausgabe), aber man darf nicht dem Irrtum unterliegen, das sei eine gründliche und vollständige Erschließung des Inhalts.

*Vielen Dank, Herr Fugmann.*