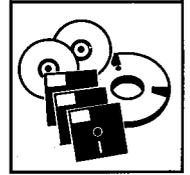
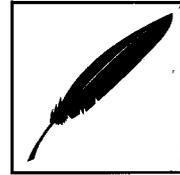


Zur praxisgerechten Erfassung und Verwaltung von Rechtsprechungsdokumenten bei der elektronischen Datenverarbeitung

Günter Krohn



I. Bedeutung und Stellenwert der Leitsätze.

Leitsätze werden in der gerichtlichen Praxis (womit hier die des Bundesgerichtshofs gemeint ist) für solche Entscheidungsteile vergeben, die eine über den Einzelfall hinausgehende – „grundsätzliche“ – Bedeutung haben und eine bestimmte Problemlösung entweder erstmals beschreiben oder zu früheren Lösungen in positivem, negativem oder änderndem Sinne Stellung nehmen. Die Einschätzung, ob dies im konkreten Fall anzunehmen oder zu verneinen ist, hängt sehr vom Problembewußtsein des entscheidenden Spruchkörpers ab. Der eine Spezialmaterie betreuende Spruchkörper neigt eher dazu, ein feines Raster seiner Rechtsprechung in Gestalt von Leitsätzen auszubreiten. Die Rechtsprechung in Wettbewerbssachen gibt hierfür ein Beispiel.

Wenn für die maschinenlesbare Dokumentation der Rechtsprechung vornehmlich auf den Inhalt der Leitsätze abgestellt wird, scheinen auf den ersten Blick daher alle wesentlichen Aussagen der Rechtsprechung gesichert zu sein. Indes, der erste Blick trügt. Soweit der Leitsatz die Rolle eines „Finders“ spielen soll, muß man sich vor Augen halten, daß die Entscheidungsinhalte durch die vergebenen Leitsätze in der Regel nur teilweise erschlossen werden. Die rechtsgrundsätzliche Frage legitimiert den Leitsatz. Sonstige fallbezogene Rechtsfragen, nicht zuletzt solche des für den Praktiker besonders wichtigen Prozeßrechts, ergeben sich nur aus den Gründen. Dabei ist zu beobachten, daß die Senate des BGH bereits einmal entschiedene Fragen bisweilen auch dann nicht für leitsatzwürdig halten, wenn diese in einer bisher „nicht ganz so“ abgehandelten Konstellation auftreten. Nicht selten sind gerade die hierzu gemachten Ausführungen von hohem didaktischem Wert, weil sie – anders als die abstrakten Leitsätze – Aussagen zu konkreten Falllösungen machen und hierbei die Entwicklungslinien der Rechtsprechung in einer mehr der Rechts-Anwendung – nicht der Rechts-Fortschreibung – verbundenen Weise offenlegen. Als Beispiel mögen die Versuche der Rechtsprechung dienen, die in der Praxis oft als hart empfundenen kurzen Verjährungsfristen des Gewährleistungsrechts durch das Ausweichen auf die richterrechtlichen Figuren des Verschuldens bei Vertragsschluß und der positiven Vertragsverletzung zu verdrängen. Um diese sonstigen Inhalte der Suche zugänglich zu machen, müßte man diese Teile der Entscheidungsgründe ebenfalls abspeichern, mindestens jedoch ihren wesentlichen Inhalt so wiedergeben, daß sie bei der Suche nicht ausfallen. Geschieht das nicht, so geht wesentlicher Argumentations-Stoff verloren.

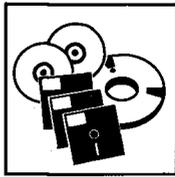
Im Blick auf das Revisionsrecht ist auf eine weitere Besonderheit hinzuweisen. Die Annahme einer Revision, die keine Frage von grundsätzlicher Bedeutung aufwirft und die im Ergebnis auch keine Aussicht auf Erfolg verspricht, kann der BGH ablehnen (§ 554 b ZPO iVm BVerfGE 54, 277). Von dieser Möglichkeit machen die Senate des BGH in unterschiedlichem Umfang Gebrauch. Da solche Nichtannahmebeschlüsse nicht begründet werden, wird ein nicht unwesentlicher Teil der revisionsrechtlichen Praxis seitdem nicht mehr dokumentiert. Die Wiedergabe des bestätigten Berufungsurteils mit beigelegter Beschlußformel des BGH trägt zur Aufhellung nicht immer viel bei. Denn der BGH braucht nicht die für das Berufungsgericht tragenden Gründe bestätigt zu haben, wenn er die Revision nur „im Ergebnis“ für nicht erfolgversprechend hält. Der u.a. für die Amtshaftung und die öffentlich-rechtliche Entschädigung zuständige III. Zivilsenat pflegt dagegen einem Teil seiner Nichtannahmebeschlüsse eine Begründung beizugeben. Über die dafür oder dagegen sprechenden Gründe ist hier nichts auszuführen. Festgehalten werden soll nur, daß in diesen Rechtsbereichen die Entwicklung der revisionsgerichtlichen Rechtsprechung nur derjenige noch einigermaßen zuverlässig überblicken kann, der diese Begründungen kennt. Da diese Beschlüsse – die übrigens durch die von den Richtern des BGH herausgegebene Rechtsprechungssammlung BGHR erfaßt sind – grundsätzlich keine Leitsätze erhalten, bleiben sie bei einer bloßen Leitsatz-Dokumentation gänzlich unberücksichtigt.

Die rechtsgrundsätzliche Frage legitimiert den Leitsatz.

Das Problem der „nicht ganz so“ abgehandelten Konstellationen

Nichtannahmebeschlüsse und Begründung: Auch ein Dokumentationsproblem

Dr. Günter Krohn ist Vorsitzender Richter am Bundesgerichtshof.



*Leitsätze mit unterschiedlichen
Bezeichnungen für dieselben Begriffe*

*Eine Gefahr: Bloßes
„Leitsatz-Wissen“*

*Volltextspeicherung: Eine Archivie-
rungsnotwendigkeit – keine Gewähr
für das Wiederfinden*

*Ein ernstes Defizit: Es fehlt ein an-
erkannter juristischer Thesaurus*

Die Benutzerphantasie ist gefordert

*Die Lösung bei BGH-DAT:
Einheitliche inhaltsbeschreibende
Kürzel*

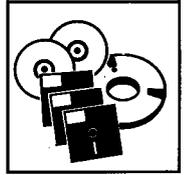
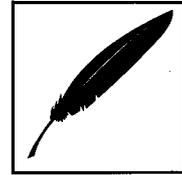
Es ist im übrigen aber auch nicht immer damit getan, die vergebenen Leitsätze in ihrem speziellen Wortlaut zu erfassen. Dieser ist oft sehr fallbezogen, nicht zuletzt auch durch die von den Parteien verwendeten Begriffe mitgeprägt. Das Gericht erliegt dann leicht der Versuchung, grundsätzliche Aussagen in den „Wortschatz“ des konkreten Prozesses zu kleiden.

So ergeben sich dann Leitsätze mit unterschiedlichen Bezeichnungen für dieselben Begriffe, etwa über die „Bürgschaft auf erstes Anfordern“ bzw. „auf erste Anforderung“ oder den „anderweitigen“ bzw. „anderweiten“ Ersatz. Für die elektronische Datenverwaltung schafft eine solche Praxis selbst bei gebotener Möglichkeit der Teilwort-Suche erhebliche Fehlerquellen. Nach alledem muß eine Abspeicherung nur der Leitsätze erhebliche Bedenken aufwerfen. Zu ihnen tritt noch die Besorgnis, daß das Finden auch der „richtigen“ Leitsätze den Benutzer nur mit dem Ergebnis der richterlichen Würdigung, nicht auch mit dieser Würdigung selbst bekannt macht. Hier ist die Gefahr der Vermittlung bloßen „Leitsatz-Wissens“ deutlich zu sehen. Ohne Kenntnis der den Leitsatz tragenden Gedanken wird dieses Leitsatz-Wissen ein durch die Leichtigkeit der computergestützten Suche in einem gewissen Sinne „eingepacktes“ Ergebnis fremden Nachdenkens. Das vom BGH aus edukatorischen Gründen bisweilen gewählte Abfassen von „Zur Frage ..“-Leitsätzen zwingt zwar zum Nachlesen der Gründe, macht aber gerade dadurch die elektronische Ablage solcher Inhalte unergiebig.

II. Volltext – Wiedergabe

Die anfängliche Euphorie, alles finden zu können, wenn man alles im Volltext erfaßt hat, hat sich inzwischen etwas gelegt. An Boden gewinnt die Einsicht, daß die Speicherung der Dokumente im Volltext zwar für deren Archivierung unabweisbar ist, jedoch allein noch keine Gewähr dafür gibt, die für die Beantwortung einer bestimmten Rechtsfrage auszuwertenden Dokumente – zumal in einer zumutbaren Recherchezeit – auch wirklich vollständig zu finden. Dies hängt einmal damit zusammen, daß mehrere in einem Dokument verwendete Begriffe – sofern sie dort überhaupt in einer juristischen Bedeutung gebraucht werden –, die bei einer verknüpften Suche ermittelt werden, keine innere Beziehung zueinander haben müssen. Die formale Strategie des „Nahe bei“-Suchbefehls kann solche Zufallsverknüpfungen bis zu einem gewissen Grad vermeiden, verhindern kann sie diese nicht. Diese Scheinbezüge nehmen mit der Länge der Dokumente noch zu.

Hinzu kommt, daß es für die von den Gerichten und vom Schrifttum verwendeten Fachausdrücke und Bezeichnungen noch immer keinen anerkannten juristischen Thesaurus gibt, wodurch diesen Bezeichnungen die gleichen – oder gar noch mehr – Fehlerquellen innewohnen wie den Leitsätzen. Die darauf zurückzuführenden Mißerfolge bei der Suche selbst in großen Archiven könnten zu einem erheblichen Umfang vermieden werden, wenn bei der Eingabe der Dokumente darauf gesehen würde, neben den im Dokument verwendeten Bezeichnungen auch die synonymen und „quasi-synonymen“ Begriffe mit aufzunehmen. M.E. kann diese Aufgabe nur von juristischen Praktikern bewältigt werden, die neben ihrer materiellen Sachkenntnis das nötige Problembewußtsein für die Wichtigkeit dieser Aufbereitungsaufgabe haben. Da in dieser Richtung jedoch noch keine verwertbaren Ergebnisse – im Sinne einer Verständigung über einen Wortschatz gleichbedeutender und verwandter Rechtsbegriffe – vorliegen, hängt es von der Phantasie des einzelnen Benutzers ab, wie er unter Zuhilfenahme von Wildcards und Trunkierung seine Suchbefehle strukturiert, um nicht am abweichenden Wortlaut zu scheitern. Die Bequemlichkeit dieser Suchmethode hat aber auch ihre Schattenseiten. In dem Maße, in dem die Suche durch Verwendung von Wildcards u.ä. unschärfer wird, werden zusätzlich Wörter des allgemeinen Sprachschatzes gefunden, die wegen ihrer rein textlichen Bedeutung die methodische Suche erschweren, weil sie aus dieser Sicht eher als „Füllwörter“ erscheinen, die man eigentlich in eine „Stoppliste“ aufnehmen und dadurch von der Indizierung ausnehmen sollte. Das ist indessen, wie keiner näheren Darlegung bedarf, praktisch nicht machbar. Der Weg zu „intelligenter“ Software, die diesen Schwierigkeiten vorbeugen könnte, ist wohl noch weit. Die – bereits praktizierte – Methode, über ein dem Datenverwaltungs-System beigegebenes Wörterbuch auch die sinntragenden Wort-Stämme zu erfassen, markiert einen Anfang dieses Weges, macht indes den Aufbau des beschriebenen Thesaurus nicht entbehrlich. In der Rechtsdatenbank BGH-DAT wird demgegenüber versucht, die Gefahr unter-



schiedlicher Schreib- oder Ausdrucksweisen durch inhaltsbeschreibende Abkürzungen, die jedem Datensatz beigegeben werden, in Grenzen zu halten. So findet z.B. das Kürzel „PrdH“ alle von der Produkthaftung handelnden Datensätze, mögen sie Begriffe wie „Produkthaftung“, „Produzentenhaftung“ oder gar „Produktenhaftpflicht“, verwenden.

Ein besonderes Problem: „Neue“ Terminologie Die Methode der Volltext-Indizierung bringt auch dann nicht die erwarteten Ergebnisse, wenn in den abgespeicherten Dokumenten besondere, z.T. ausgefallene juristische Bezeichnungen verwendet werden, die dem allgemeinen Wortschatz (noch) nicht oder nicht mehr geläufig sind und daher dem einzelnen Benutzer bei der Suche nicht präsent sein müssen. Das gilt namentlich für den Fall, daß neue Fachausdrücke auftauchen, die von einem Gericht verwendet werden, die sich aber in der Fachwelt noch nicht durchgesetzt haben. Schließlich – und dies dürfte der in der Praxis häufigste Fall sein – gibt es Rechtsfiguren, die unter unterschiedlichen Bezeichnungen auf dem Markt sind, ohne daß vorauszusagen wäre, für welche Bezeichnung sich das Gericht im Einzelfall entschieden hat. Ich denke hier u.a. an die juristische Nomenklatur für die persönliche Haftung des Vertreters, die in der gerichtlichen Praxis mit Begriffen wie „Vertrauenshaftung“, „Sachwalterhaftung“, „Durchgriffshaftung“, „Vertreterhaftung“, „Verschulden bei Vertragsschluß“ oder auch „culpa in contrahendo“ jegliche Einheitlichkeit vermissen läßt. In solchen Fällen bietet sich auch kein „Einheits“-Begriff an, der die in den Einzelbegriffen steckenden Nuancierungen überspielen könnte. Eigentlich müßten daher die Bestände von Datenbanken, die auch zeitlich weiter zurückliegende Dokumente enthalten, fordaufend an die Entwicklung der juristischen Fachsprache angepaßt werden. Das damit verbundene ständige Durchforsten des Datenbestandes bei gleichzeitiger sensibler Kontrolle des Fortschritts der Rechtswissenschaft ist schon redaktionell kaum vorstellbar. Eine gewisse Abhilfe können hier die sog. *Makros* bieten, die es ermöglichen, gegebenenfalls mit einer ganzen Liste von Begriffen gleicher oder ähnlicher Bedeutung zu suchen, um hierdurch eine abweichende Wörtwahl im zu suchenden Dokument zu überspielen. So könnte etwa im obenstehenden Beispiel unter dem Sammelbegriff „Vertrauenshaftung“ mit der ganzen Wortliste gesucht werden. Allerdings müßten dann diese *Makros* der Entwicklung der Rechtswissenschaft entsprechend gepflegt werden; das erscheint mir aber in Ansehung des damit verbürgten Erfolges noch ein vertretbarer Aufwand.

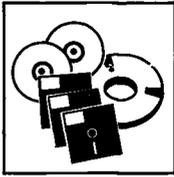
Eine weitere Unsicherheit bei der Suche in Volltext-Datenbanken ergibt sich, wenn sie darauf angelegt sind, die gesamten Inhalte ohne Rücksicht auf ihren Sinngehalt unterschiedslos zu erfassen. Jede gerichtliche Entscheidung enthält „Zeilenfüller“, die für eine juristische Recherche ohne jede Bedeutung sind. Der in so gut wie jeder zivilrechtlichen Entscheidung enthaltene Satz „Die Kostenentscheidung beruht auf § 91 ZPO“, ist solch ein „Füller“, der zum „Datenmüll“ zu rechnen ist und daher grundsätzlich ausgesondert werden müßte. Er steht – wenn tausendfach mit abgespeichert – einer übersichtlichen Suche nur im Wege und sollte daher nur erscheinen, wenn wirklich ein nicht schon vom Gesetzeswortlaut gelöstes Problem des Kostenrechts auftaucht. Dieses Beispiel steht für viele ähnliche Formeln, die nur Selbstverständlichkeiten ausbreiten und daher in einer Rechtsdatenbank – außer im Gesetzestext – nichts zu suchen haben.

Die Ausrichtung der Volltext-Abspeicherung am ungekürzten Inhalt des Originaldokuments bringt jedoch noch weitere Probleme mit sich. Die Sprache der Juristen mag in gewisser Hinsicht eine „Kunst“-Sprache sein, sie kommt indessen nicht ohne den Gesamtwortschatz der Sprache aus, aus der sie gewachsen ist. Dies führt überall dort zu Mehrdeutigkeiten, wo die verwendeten Begriffe der Sprache auch sonst geläufig sind, ihr mehr oder minder zufälliges Auffinden als singulärer Begriff also nicht gewährleistet, daß damit auch das gesuchte „Fachwort“ gemeint ist. Ein Beispiel möge dies verdeutlichen: Der Satz „Die Klage über die unverständliche Ausdrucksweise der Juristen ist allgemein zu vernehmen“ klingt unverdächtig, jedoch können seine einzelnen Teile bei einer Textsuche verwirrende Folgen haben. Das Wort *Klage* ist juristisch „besetzt“; seine Verwendung in einem anderen Zusammenhang als dem des Prozeßrechts (oder darauf Bezug nehmender Teile des materiellen Rechts) bringt Mehrdeutigkeit in scheinbar eindeutige Begriffsinhalte. Bei einer nur darauf ausgerichteten Suche erschiene als „Treffer“, was in Wirklichkeit keiner ist. Das Adjektiv *allgemein* grenzt für die Juristen gemeinhin einen Bereich vom *Besonderen* ab. Kommt im Text des Dokuments etwa noch das Wort „Norm“ vor, so könnte bei einer Verknüpfung der beiden Begriffe leicht der Eindruck entstehen, als handele es sich um ein Dokument, das etwas über die allgemeine Wirkung oder Bedeutung einer Norm aussagt. Auch die Worte *Ausdrucksweise* und

Schwer einzuschätzen: Mehrdeutige Terminologie

Störend bei der Recherche: „Zeilenfüller“

Ein Beispiel: „Die Klage über die unverständliche Ausdrucksweise der Juristen ist allgemein zu vernehmen“



*Rechtsprechung: Gottlob noch keine
seelenlose Subsumptionsmaschine*

vernehmen scheinen, zumal durch ihre räumliche Nachbarschaft, in einem juristischen Text eher auf eine Frage der Beweisaufnahme oder -würdigung hinzudeuten als auf den oben dargestellten Zusammenhang. Wollte man diese Begriffsüberschneidungen radikal ausschalten, so müßten in juristischen Texten alle Worte der Alltagssprache, die an der gesetzten Stelle eine Fachbegriff „simulieren“ vor einer Datenübernahme ausgeschieden werden. Man kann sich vorstellen, daß EDV-Programme erstellt werden können, die eine solche Aufgabe der „Textbereinigung“ exakt erfüllen. Das wäre dann aber wohl das Ende einer bei allem Gesetzesvollzug doch schöpferischen Rechtsprechung, die – trotz aller Belastung und allen Leistungsdrucks – gottlob noch nicht zu einer seelenlosen Subsumptionsmaschine geworden ist. So geht es also nicht.

Andererseits würde es sich aber der juristische Praktiker, der sein „Handwerk“ (dies ist in einem positiven Sinne gemeint!) beherrscht, dem aber die EDV – bislang – noch nicht viel bedeutet, strikt verbitten, wenn wir von ihm forderten, er möge künftig seine Urteile und Beschlüsse in einer computergerechten Sprache absetzen. In der Tat ginge es nicht an, daß der Jurist sein Stil- und Sprachgefühl – soweit ihm die Hektik der Alltagsarbeit überhaupt noch dessen Einsatz ermöglicht – ausgerechnet auf dem Altar der EDV opfern müßte. Hilfe muß also von der Seite der EDV kommen, allerdings ist hierfür nach meiner Auffassung die juristische Vorbearbeitung der zu speichernden Dokumente eine unverzichtbare Voraussetzung.

III. Die Bedeutung der juristischen Bearbeitung aus dem Gesamthalt des Dokuments

Die Konsequenz dieser Überlegungen sollte sein, daß alle zu speichernden Inhalte – ob Leitsatz oder Volltext – mit einem juristischen Überbau versehen werden, der alle wesentlichen Aussagen über das aufzunehmende Dokument enthält. Dieser Überbau müßte vor allem die Suchrichtungen kanalisieren, indem die gemachten Inhaltsangaben das Dokument nach Bedeutung und Aussage sowohl konkretisieren als auch eingrenzen. Über das hier „Wesentliche“ wird sich freilich kein allgemeiner Konsens erzielen lassen. Hier entscheiden einfach fachliches Können und Phantasie des eingebenden Bearbeiters, der im übrigen mit den Leistungsmöglichkeiten der zur Verfügung stehenden elektronischen Werkzeuge hinreichend vertraut sein sollte. Auch hier ist also Erfahrungswissen im Bereich der EDV in der Justiz dringend gefordert.

Die Inhaltsangabe darf sich aus den oben angeführten Gründen nicht auf die in dem Dokument verwendeten Worte und Begriffe beschränken. Sie muß so beschaffen sein, daß der zu vergebende „Index“ die Problemlösung in ihrem juristischen Kontext kennzeichnet. Ich halte es in diesem Zusammenhang für falsch, das abzuspeichernde Dokument vor allem durch Daten zu kennzeichnen, die es zu einem „Unikat“ machen. Ich möchte dies an dem Rechtsprechungs-Beispiel demonstrieren, das auf dem Workshop „Informationstechnik am Arbeitsplatz für Juristen 1989“ in Saarbrücken für das JURISTISCHE LITERATURARCHIV JULIA gegeben wurde (vgl. die gleichlautende Schrift, herausgegeben von van Raden/Weihermüller S.99 [102]). Der dort wiedergegebene Leitsatz des OLG Frankfurt a.M. lautet:

„Bei einem Ratenkredit liegt ein auffälliges Mißverhältnis zwischen Leistung und Gegenleistung auch dann vor, wenn der vereinbarte Zins den marktüblichen Zins um mehr als 12%-Punkte übersteigt. Zur Berechnung der marktüblichen Zinsen bei Krediten über 60 Monaten Laufzeit.

OLG Frankfurt, 9.1.1987, 22 U 80/86, NJW-RR 87, 304.–

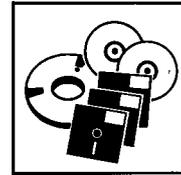
Vorgeschlagen wird aaO als „naheliegend“ die Abspeicherung der Bezeichnungen des Gerichtsorts, des Datums, des Aktenzeichens und der Fundstelle, weil diese Daten „nur auf dem Ordner dieses Leitsatzes vorkommen“. Daneben sollen die Schlagworte *Ratenkredit*, *BGB 138*, *Sittenwidrigkeit* und *Marktüblicher Zins* vergeben werden.

Dazu ist zu bemerken: Die Angaben über Gerichtsort und Aktenzeichen erscheinen mir recht fernliegend. Sie werden sich bei einer systematischen Suche eher als hinderlich erweisen. Dagegen erlauben Fundstelle und Datum schon eine erhebliche Eingrenzung der einschlägigen Dokumente nach Erscheinensjahr und Veröffentlichungsmedium. Hinsichtlich der vorgeschlagenen materiellen Deskriptoren bin ich ganz anderer Ansicht als die Verfasser. Der Begriff der *Sittenwidrigkeit* stellt ein juristisches Werturteil für einen komplexen Sachverhalt dar. Er baut von der Tatbestandsseite her auf rechtlichen Voraussetzungen auf, die ihrerseits durch Schlagworte gekennzeichnet müssen. Unverzichtbar erscheinen mir danach zunächst die weiteren Schlagworte

*Der Index muß die Problemlösung
im juristischen Kontext kennzeichnen*

*Ein Beispiel von der GI-Tagung
1989*

*Die „Verortung“ der Entscheidung
erfordert einen beachtlichen
Aufwand an „Verschlagwortung“*



Leistung, Gegenleistung und *Mißverhältnis*. Ohne diese unterstützenden Angaben hängt auch der vergebene Begriff: *Marktüblicher Zins* in der Luft. Da das Mißverhältnis zwischen Leistung und Gegenleistung entscheidend am *Marktzins* gemessen wird, müßte auch dieses Schlagwort erscheinen, ebenso der Begriff *Vertragszins*; dagegen erscheint der *Marktübliche Zins* als eine mehr zufällige Wortwahl, deren Verwendung bei einer Suche zwar zu der eingegebenen Entscheidung führt, jedoch weitere repräsentative Funde auf diesem Rechtsgebiet nicht verbürgt. Die Berechnung des Marktzinses bei längerfristigen Krediten läßt weiter an die Unterschiede zwischen der *Uniformmethode* und der *finanzwissenschaftliche(n) Methode* denken, jedenfalls aber an eine *Methode* der *Berechnung* des *Zins(es)*, sodaß zumindest diese Schlagworte noch vergeben werden sollten. Neben dem vorgeschlagenen Schlagwort *Ratenkredit* erscheint mir des weiteren der Hinweis auf: *Darlehen* und auf: *BGB § 607* ebenso notwendig wie auf: *Laufzeit* und *Kredit*. Selbst durch diese zusätzlichen Angaben wird der Entscheidungsinhalt aber noch nicht befriedigend ausgeschöpft sein. Man müßte bei einem Ratenkredit etwa noch prüfen, ob die Entscheidung über den weiteren Inhalt des Darlehensvertrages etwas Wichtiges aussagt (etwa über unwirksame AGB-Klauseln), über die Anrechnung und Verteilung einer Restschuldversicherungsprämie, über eine Vermittlungsgebühr, über ein mögliches (offenes oder verdecktes) Packing u.ä.m. Im Blick auf die Konditionen des Kreditmarktes müßte man wohl auch an die Schlagworte *Niedrigzinsperiode* und *Hochzinsperiode* denken.

Ich führe dieses Beispiel nur an, weil es deutlich macht, wie eine gerichtliche Entscheidung gegebenenfalls „verortet“ werden müßte, um für spätere Recherchen auch wirklich zur Verfügung zu stehen. Es zeigt m.E. eindringlich, welche Schwierigkeiten in der Praxis auftauchen, wenn man eine Entscheidung nur über solche Schlagworte und Begriffe erfaßt, die dem Leitsatz und den äußeren Entscheidungsdaten entnommen sind, ohne zumindest auch einen Teil der Entscheidungsgründe mit in den Index aufzunehmen.

Befriedigende Ergebnisse sind hier nur zu erzielen, wenn der in der jeweils angesprochenen Rechtsmaterie kundige Jurist bei der Vergabe des Schlagwort-Thesaurus im Einzelfall selbst Hand anlegt. Die dabei auftauchenden Probleme ähneln übrigens sehr den Schwierigkeiten, denen sich die Entwickler von juristischen Expertensystemen ausgesetzt sehen. Das wundert denjenigen nicht, der in der geglückten Analyse der Arbeitsweise des juristischen Praktikers eine strukturelle Vorbedingung für solche Expertensysteme sieht.

Stellt man die im Beispielsfall für erforderlich gehaltenen Schlagworte/Begriffe zusammen, so kommt man auf die stattliche Anzahl von mindestens 17 beschreibenden Hinweisen. Von diesen kommen im konkreten Leitsatz immerhin 8 vor; sie würden daher bei einer Völlindizierung des Leitsatzinhaltes erfaßt. Die restlichen Hinweise müßten dagegen den Entscheidungsgründen entnommen werden bzw. eigenständig gebildet werden. Andernfalls würden sie ohne Erfassung des weiteren Inhalts der Entscheidungsgründe für spätere Suchen nicht verfügbar sein.

IV. Anforderungen an ein „juristengemäßes“ Dokumenten-Verwaltungs-Programm

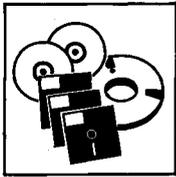
1. Die Datenzugriffe sollten von einem gröberen bis zu einem feineren Raster reichen. Dabei sollten die „natürlichen“ Suchwege des Praktikers unterstützt werden. Insofern steht m.E. nicht die „Suche auf den Punkt“ im Vordergrund. Es mag zwar hilfreich sein, etwa im Unterhaltsrecht direkt danach zu suchen, ob im Datenbestand ein Dokument über die „Sättigungsgrenze“ vorkommt. Ebenso kann es sich anbieten, mit der Suche in einer sehr kleinen Menge zu beginnen, um in ihr vielleicht schon den erwünschten „Treffer“ zu landen. Im allgemeinen geht die juristische Suche jedoch auf anderen Wegen vor sich: Rechtsinstitut (Beispiel: Verschulden bei Vertragsschluß), Gesetzesbegriff (Beispiel: BGB) und einzelne Norm (Beispiel: 607) werden wohl den Einstieg bilden. In der so gefundenen Teilmenge sollte dann nach Maßgabe eines feineren Rasters gesucht werden können. Ich denke dabei in erster Linie an Schlagworte/Stichworte, aber auch an das Datum und den Spruchkörper. Diese Kriterien müßten bei der Suche sowohl verengend („Und“-Verknüpfung) als auch erweiternd („Oder“-Verknüpfung) eingesetzt werden können; sinnvoll erscheint in bestimmten Fällen auch eine „Und Nicht“-Verengung.

Selbstverständlich müssen diese Suchkriterien dokumentbezogen einsetzbar sei. Dies

Beim Thesaurus muß der kundige Jurist selbst Hand anlegen

Nicht alle nötigen beschreibenden Hinweise finden sich im Leitsatz

Recherche-Einstieg: Rechtsinstitut, Gesetzesbegriff, Norm



Wünschenswert: Transparenz der Zwischenergebnisse

Ideal: Eine „gestaffelte“ Recherche

Unverzichtbar: Geschwindigkeit in großen Datenmengen

Der Zugriff auf Feldinhalte muß gesichert sein

Die Lösung in ADIMENS GT: Feldqualifikatoren

stellt indessen nur eine Minimalforderung dar. Denn genauere Ergebnisse lassen sich erst erzielen, wenn auch in bestimmten Bereichen innerhalb des Dokuments gesucht werden kann.

2. Die gleichzeitige Verknüpfung mehrerer Kriterien in einem einzigen Suchvorgang führt schnell in kleinere Datenmengen. Dies kann zwar erwünscht sein, hat jedoch den Nachteil, daß die Suchergebnisse jeder einzelnen „Stufe“ unerkannt bleiben. Möglicherweise hätte deren Kenntnis Anlaß geben können, ab einem bestimmten Punkt in eine andere Richtung und ggfls. mit anderen Kriterien zu suchen. Die Suche mit mehreren Vorgaben kann zwar mit jeweils verminderter Zahl der Sucheingaben wiederholt werden; dies erschien mir aber gekünstelt, weil vom Endergebnis her bestimmt.

3. Dagegen gestattet die gestaffelte, in Stufen vor sich gehende Suche eine eher spielerische Recherche, die unter größtmöglicher Transparenz verläuft. Man wählt z.B. einen möglichst breiten Einstieg (Beispiel: Darlehen; Ratenkredit; Kredit; BGB § 607, in einer „Oder“-Verknüpfung) und verengt dann diesen „Maximalbestand“ schrittweise durch die Eingabe spezifischer Suchkriterien (Beispiele: Sittenwidrigkeit; Marktzins; Vertragszins, etc). Wichtig ist hierbei, daß das System in der Lage ist, jeweils die mit der vorletzten Eingabe gefundene Teilmenge festzuhalten; auf diese Weise werden unter Wahrung der bisherigen Zwischenergebnisse Irrwege sofort erkannt; sie können dann leicht durch Änderung der späteren Suchbefehle vermieden werden. Eine solche Suchstrategie hat gegenüber der erstgenannten Methode neben höherer Transparenz wohl auch die größere Flexibilität für sich.

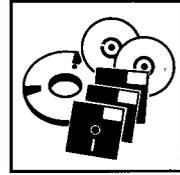
4. Beide Methoden werden indes bei größeren Datenmengen (ich denke an mehr als 4 MB) nur befriedigen können, wenn sie eine schnelle Suche gewährleisten. Man wird daher nicht daran vorbei kommen, die aufzufindenden Datenteile in Indexlisten aufzunehmen, die einen sofortigen, dokumentbezogenen Zugriff auf das indizierte Datenmaterial erlauben. Damit sind wir bei einem Dokumenten-Retrieval-Programm als Wunschobjekt angelangt.

5. Um Zufallszusammenhänge so weit wie möglich auszuschalten, sollte das Programm weiter in der Lage sein, strukturierte Suchbefehle abzuarbeiten.

a) Gemeint ist damit vor allem, daß auf bestimmte Teile in den Dokumenten zugegriffen werden kann, denen wegen der „Feldinhalte“ eine besondere, unverwechselbare Bedeutung zukommt. Wenn beispielsweise in einer bestimmten Zeile jedes Dokuments eine Inhaltsbeschreibung gegeben wird, darf man sicher sein, daß ein in dieser Zeile aufgeführtes Rechtsinstitut nicht nur beiläufig im Gesamttext vorkommt, sondern dokumentenspezifisch verwendet wird. Praktisch setzt dies allerdings voraus, daß der bereits beschriebene juristische „Überbau“ der diese Aussage machen soll, dem Daten-Verwaltungsprogramm vorgegeben wird.

Hier möchte ich den Blick auf Datenbank-Programme lenken, die diese Aufgabe bestens erfüllen können. Mit ihrer Hilfe können die Einzeldokumente entsprechend aufbereitet und verwaltet werden. Allerdings müssen sie in der Lage sein, auch umfangreichere Textteile aufzunehmen, wie dies z.B. bei der relationalen Datenbank ADIMENS GT der Fall ist. Nimmt man hier die in jedem Dokument behandeltem Gesetze und Rechtsbegriffe in bestimmte (Merkmals-)Felder auf, die dem Text vorangestellt sind, so hat man später die Möglichkeit, beim Export dieser Datensätze unter Verwendung entsprechender Mischformulare felderspezifische Daten auszulesen, die vom Dokumenten-Retrieval-Programm als solche erkannt werden, etwa, indem in dem Mischformular dem Feld *BGB* ein sonst nicht vorkommendes Sonderzeichen beigegeben wird, mit dessen Hilfe die Suche auf den Inhalt der einschlägigen Felder beschränkt wird. So könnte etwa die Zugabe des „\$“-Zeichens ein solches Feld definieren. Der Feldinhalt wäre dann zeilenbezogen abzurufen, was sich als eine Anwendung der „Nahe-bei“-Suche darstellen würde. Es verschlüge dabei nichts, wenn etwa in dem *BGB*-Feld verschiedene Vorschriften stünden; denn die feldbezogene Suche würde alle Eingaben zu überspringen haben, die zwischen der Feldbezeichnung und dem ersten zutreffenden Suchkriterium stehen. Es wäre auch denkbar, in ein solches Feld mehrere

Gesetze und Vorschriften aufzunehmen; ein falsches „Mischen“ nicht zusammengehörender Gesetze und Vorschriften würde beispielsweise vermieden, wenn der auf die Gesetzesbezeichnung folgende Befehl – die einschlägige Norm als Zahl – nur innerhalb des Zwischenraums wirken würde, der durch den ersten Buchstaben der ASCII-Tabelle begrenzt wird.



b) Die feldorientierte Suche im Gesamtdatenbestand ist eine spezielle – allerdings sehr wirksame – Spielart des „Nahe-bei“-Suchens. Diese Option dürfte in einem Dokumenten-Retrieval-Programm, das juristische Texte verwaltet, nicht fehlen. Die zugrundeliegende Idee ist, daß räumlich nahe beieinander stehende Begriffe eine erhöhte Wahrscheinlichkeit in sich tragen, daß sie auch in einem Sinnzusammenhang stehen. Die einfachste Form ist der „Neben“-Suchbefehl, der so beschaffen sein sollte, daß er zwei nebeneinander stehende Begriffe unbeschadet ihrer Reihenfolge auffindet. Dabei müßte das Programm dem Verwender die Möglichkeit geben, bestimmte Zeichen zu negieren. Wird etwa das „-“ Zeichen als ein Zeichen definiert, das bei der „Neben“-Suche vernachlässigt werden soll, so kann mit der Eingabe „positive Vertragsverletzung“ auch die Textstelle „Vertragsverletzung, positive“ gefunden werden. Darüber hinaus sollte die „Nahe-bei“-Suche auch für eine vom Benutzer definierbare Intervall-Suche eingesetzt werden können, die innerhalb einer bestimmten Anzahl von Zeichenketten wirksam ist.

Unverzichtbar: Die „Nahe-bei“-Suche

c) Damit enden die Wünsche des Juristen an die Ausstattung eines solchen Dokumenten-Retrieval-Programms noch immer nicht. Man muß sich vergegenwärtigen, daß das zu verwaltende Material – etwa wenn es mit einem Scanner eingelesen worden ist – nicht stets und durchgehend gewährleistet, daß alle zu indizierenden Datenteile von Leerzeichen eingerahmt sind, die dem Programm eine klare Separation ermöglichen. Zeichen wie „.,;:)?!“ um nur einige zu nennen, können einen Zugriff auf Zeichenketten verhindern, an deren Beginn oder Ende ein solches Zeichen steht. Für die Verwaltung der aus Buchstaben bestehenden Zeichenketten sollten daher diese Zeichen als „Trennzeichen“ wählbar sein und zwar so, daß der Benutzer diese Liste selbst zusammenstellen kann. Daneben sollten Trennzeichen auch für Ziffern einstellbar sein, jeweils mit der Folge, daß ein nicht zum Trennzeichen bestimmtes Zeichen dann auch als Teil der Zeichenkette verwaltet wird. So wird es z.B. sinnvoll sein, für Zahlenkombinationen den „Punkt“ nicht als Trennzeichen zu definieren, um beispielsweise die Datumschreibweise „1.12.89“ als einheitliche Zeichenkette verwalten zu können. Dagegen würde die Vorgabe der runden Klammer als Trennzeichen es ermöglichen, Zahlen – ohne Benutzung einer Wildcard – auch dann zu finden, wenn sie von solchen Klammern eingerahmt sind. Für die Verwaltung von Buchstaben-Zeichenketten gelten diese Ausführungen entsprechend.

Nötig auch: Freie Wählbarkeit von Trennzeichen

d) Eine Buchstaben-Zeichenkette muß schließlich auch dann noch gefunden werden, wenn sie am Zeilenende durch einen Bindestrich getrennt wird. Um zu vermeiden, daß dadurch sinnverwandte, ebenfalls durch einen Bindestrich getrennte Worte mit Großbuchstaben am Wortanfang miterfaßt werden, sollte der Trenn-Algorithmus auf die Fälle beschränkt sein, in denen nach dem Bindestrich am Zeilenende in der folgenden Zeile in kleiner Schrift fortgefahren wird.

Ein Sonderproblem: Der „Bindestrich“ am Zeilenende

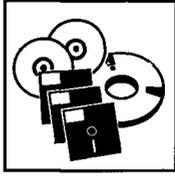
e) Der Einsatz von „wildcards“ erscheint unverzichtbar, soweit es sich um die Wortenden handelt („Rechts-Trunkierung“). Schwieriger dürfte sich die „Links-Trunkierung“ gestalten, allerdings wohl nur im Hinblick auf den benötigten Speicherplatz. Löst man dieses Problem dadurch, daß man alle Begriffe auch in Rechts-Links-Schreibweise indiziert, so gelangt man zu einer Verdoppelung der Index-Dateien. Wegen des ohnehin großen Speicherbedarfs der vorstehend beschriebenen Optionen dürfte dieser Wunsch daher eher in die „zweite Reihe“ gehören.

„Rechtstrunkierung“ unbedingt nötig „Links-Trunkierung“ als Wunsch „zweiter Reihe“

f) Das für die Indexierung des Datenbestandes benötigte Speichervolumen hängt grundsätzlich davon ab, ob (im Prinzip) jedes Wort indiziert wird oder nur bestimmte Teile des Bestandes.

Hier erscheint es mir nötig, aber auch ausreichend, alle Begriffe mit großem Anfangsbuchstaben und alle Zahlen in den Index aufzunehmen, von den kleingeschriebenen Wörtern jedoch nur diejenigen, die in eine Positivliste aufgenommen werden. Eine

Ein Teilindex genügt



*Wichtig: Automatische Separierung
von Einzeldokumenten*

solche Liste könnte etwa so erstellt werden, daß der Datenbestand als Text-Datei mit Hilfe eines geeigneten Textverarbeitungsprogramms bearbeitet und hierbei alle in eine Indexliste aufzunehmenden Worte gekennzeichnet werden. Daneben bedarf es einer Negativliste für alle mit großem Buchstaben beginnenden Worte, die – als nicht sinntragend – ausgeschieden werden sollen.

g) Das alles funktioniert freilich nur, wenn die aus der Datenbank zu exportierenden Datensätze vom Daten-Verwaltungs-Programm als Einzeldokumente behandelt werden, ohne daß man sich der Mühe unterziehen müßte, jeden Datensatz als eigene Datei aufzubereiten. Retrieval-Systeme, die auf große Dateien ausgelegt sind, erstellen idR eine Indexliste für eine Vielzahl selbständiger Dateien, die in Ordnern gesammelt werden können. Für eine Verwaltung juristischer Dokumente des hier vorausgesetzten Inhalts können sie kaum eingesetzt werden. Das hierfür in Frage kommende Dokumenten-Retrieval-Programm sollte in der Lage sein, ein an den Anfang jedes (aus der Datenbank exportierten) Datensatzes gestelltes Zeichen (etwa „???“) als Dokumentenanfang zu erkennen und die Datensätze entsprechend zu separieren. Auf diese Weise könnten die „Report“-Funktionen der Datenbanken für solche Dokumenten-Retrieval-Programme nutzbar gemacht werden.

*Von der Liste zum Dokument: Ein
entscheidender Arbeitsgang*

h) Von besonderer Wichtigkeit erscheint es mir auch, daß gefundene Dokumente nicht nur als Text sondern auch in Form einer Liste dargestellt werden können, in der jeder Datensatz eine Zeile einnimmt. Das Programm sollte weiter die Möglichkeit bieten, einzelne Dokumenten-Texte (mit Cursor oder Maus) aus der Liste heraus aufzurufen. Erfahrungsgemäß lassen sich aus einer solchen Liste schon vor einer Ansicht des „Volltextes“ einzelne Dokumente als unergiebig ausscheiden. Besonders hilfreich wäre es, diese Dokumente – wie unter Adimens GT und GEM via „Klemmbrett“ möglich – vom Bildschirm wegzubefördern und nur mit dem Restbestand des Suchergebnisses weiterzuarbeiten.

Nicht entscheidende Eingabemasken

i) Eine besondere Eingabemaske erscheint mir dagegen nicht unbedingt erforderlich. Die vorstehend geschilderten Funktionen lassen sich in einer einzigen Eingabezeile zusammenfassen, in die sowohl Zahlen als auch Begriffe eingegeben werden können. Es muß nur darauf geachtet werden, daß die Eingabezeile nicht zu kurz ausfällt. Eine Breite von 30 bis 35 Zeichen dürfte aber ausreichen, auch für den „Nahe-bei“-Suchbefehl, bei dem man nötigenfalls das Wortende durch Eingabe einer „wildcard“ ersetzen kann.