



Wider einige gängige Ansichten zur juristischen Informationserschließung

Bettina Mielke

Landgericht Regensburg
93066 Regensburg
bettina.mielke@lg-r.bayern.de

Zusammenfassung

Ausgehend von einer Betrachtung in der Rechtsinformatik geläufiger Annahmen zur juristischen Informationserschließung beschreibt der Beitrag wesentliche Ergebnisse einer empirischen Studie der Retrievaleffektivität von Recherchen in juristischen Datenbanken. Dabei steht die Frage nach der Notwendigkeit einer intellektuellen Erschließung einerseits, der Effektivität der sogenannten Stichwortsuche andererseits im Mittelpunkt. Die Ergebnisse der Studie, bei der auch ein Vergleich zwischen einem Informationssystem auf der Basis eines Booleschen Retrievalmodells mit einem System auf der Basis statistischer Verfahren vorgenommen wurde, legen den Schluss nahe, dass in der rechtsinformatischen Fachliteratur analytisch begründete Annahmen wie die Gefahr zu großer Antwortmengen bei der Stichwortsuche empirisch nicht zu belegen sind. Auch zeigt sich keine Überlegenheit intellektueller Erschließungsverfahren (Beschlagwortung) gegenüber der automatischen Indexierung, im Gegenteil führt der Einsatz eines statistischen Verfahrens bei identischer Dokumentkollektion zu einer höheren Wiedergewinnungsrate (recall).

Abstract

Starting from a short analysis of wide-spread assumptions about law information systems this paper describes the main results of an evaluation study concerning the effectiveness of searches in juridical databases. The need for intellectual indexing as well as the effectiveness of keyword-based search strategies are at the heart of this study. Its results which also comprise a comparison of a retrieval system based on Boolean logic with a statistical search engine show that analytical assumptions about too large answer sets for keyword searches cannot be verified empirically. Furthermore, no indication for advantages of intellectual indexing can be seen. To the contrary, using a sta-



tistical search engine results in higher recall values search the same document collection.

1 Einleitung

Internet- bzw. webbasierte Informationssysteme sind heute auch im Bereich der juristischen Fachinformation in vielfältiger Weise verfügbar. Zwar haben sie die juristische Praxis noch nicht in einer mit anderen Fachdisziplinen vergleichbaren Weise durchdrungen, werden aber mehr und mehr zum alltäglichen Arbeitswerkzeug des praktizierenden Juristen.

Dabei darf nicht darüber hinweggesehen werden, dass der Aspekt der *generellen Verfügbarkeit* durch Zugang zum WWW und damit auch zu juristischen Informationssystemen von jedem Arbeitsplatz aus sowie der Aspekt der Verbesserung der hierfür entwickelten Benutzerschnittstellen im WWW bisher wenig an der grundsätzlichen Problematik der Erschließung relevanter Fachinformation geändert haben dürften.

Der folgende Beitrag befasst sich mit einigen gängigen Thesen zur juristischen Informationserschließung und stellt diesen die Ergebnisse einer empirischen Evaluierungsstudie gegenüber. Dabei zeigt sich, dass sich die allein aufgrund von analytischen Betrachtungen gewonnenen Einschätzungen empirisch nicht bestätigen lassen.

2 Literatursuche als notwendiger Bestandteil der juristischen Praxis

Die Literatursuche ist ein notwendiger Bestandteil der juristischen Praxis. Sie beginnt mit der Suche nach den einschlägigen und zu einem bestimmten Zeitpunkt gültigen Gesetzesbestimmungen, was bei selten vorkommenden Nebengesetzen durchaus aufwändig sein kann. Auch die Suche nach Gerichtsentscheidungen und Literaturbeiträgen ist in nahezu allen juristischen Tätigkeitsfeldern unumgänglich - man denke allein an die Haftung des Rechtsanwalts, der seinen Mandanten falsch berät, da er die neueste obergerichtliche Rechtsprechung nicht kennt.

Die traditionell wichtigste juristische Sekundärliteratur in Deutschland stellen Kommentare und Handbücher dar, die Gerichtsentscheidungen und wissenschaftliche Rechtsliteratur auswerten und einen ersten Überblick zu den vor-

handenen Rechtstexten geben.¹ Daneben besteht ein überaus vielfältiges Spektrum von Fachzeitschriften und sonstigen Sammlungen, in denen Gerichtsentscheidungen und Aufsätze veröffentlicht sind und die über verschiedene Register inhaltlich erschlossen sind, z.B. die Fundhefte für Zivilrecht, Systematischer Nachweis der Deutschen Rechtsprechung und Zeitschriftenaufsätze oder die Karlsruher Juristische Bibliographie.²

Juristische Informationssysteme stellen nicht nur eine Alternative zur traditionellen Informationsbeschaffung dar, sondern eröffnen neue, direkt vom Arbeitsplatz aus erreichbare Informationsquellen, die hinsichtlich ihrer Leistungsfähigkeit über die traditionellen Sammlungen und Nachschlagewerke weit hinausgehen.

3 Einige gängige Thesen zur Informationserschließung

3.1 Notwendigkeit einer intellektuellen Erschließung

Das elektronische Publizieren von rechtlichen Texten, also Normen, Judikatur und Literatur, ist nach wie vor eines der wichtigsten Themen der Rechtsinformatik.³ Dabei geht die deutschsprachige Rechtsinformatik überwiegend davon aus, dass die intellektuelle Erschließung einen Mehrwert bringt und die automatische Erschließung nicht „die sehr hohe Recall und Präzision der intellektuellen Erschließung“⁴ erziele. So zeige z.B. die „hohe Akzeptanz einer juristischen Erschließung in Österreich“ deutlich, „dass außerhalb des öffentlichen Sektors (und oft auch dort) ohne einen Index juristische Arbeit nur erschwert möglich ist. Die nächsten Jahre werden eine zunehmende Unterstützung und Verbesserung dieser Indexierung durch neue Methoden der Rechtsinformatik bringen.“⁵ Diese Behauptungen werden dabei ohne empirische Grundlage getroffen, da Evaluierungsstudien zu diesem Thema fehlen. Im deutschsprachigen Raum gibt es bislang nur Endnutzerbefragungen, die nach

¹ Vgl. etwa Möllers, JuS 2000, S. 1203 (1208).

² Vgl. Mielke 2000, S. 15 f.; Möllers, JuS 2000, S. 1203 (1205 ff.).

³ Elektronisches Publizieren und Rechtsdatenbanken, Vorbemerkungen der Herausgeber, in: Schweighofer/Menzel/Kreuzbauer 2001, S. 97 ff.

⁴ Elektronisches Publizieren und Rechtsdatenbanken, Vorbemerkungen der Herausgeber, in: Schweighofer/Menzel/Kreuzbauer 2001, S. 98.

⁵ Elektronische Publizieren und Rechtsdatenbanken, Vorbemerkungen der Herausgeber, in: Schweighofer/Menzel/Kreuzbauer 2001, S. 98.

den persönlichen Einschätzungen der Systembenutzer fragen, aber keine Evaluierung der tatsächlichen Retrievalleistung vornehmen.⁶

Auch für die in Deutschland größte und wichtigste juristische Online-Datenbank *juris* werden Verbesserungsmöglichkeiten vor allem darin gesehen, weitere intellektuelle Indexierungsmaßnahmen vorzunehmen, z.B. durch die verbesserte Vergabe von Schlagworten und Normen bzw. durch den Aufbau einer juristischen Klassifikation.⁷

Dies ist nicht zuletzt deshalb erstaunlich, weil das System *juris* im Vergleich zu anderen Fachinformationssystemen sowie juristischen Informationssystemen in anderen Ländern⁸ bereits eine starke intellektuelle Indexierungskomponente aufweist. So werden nicht nur die bibliographischen Angaben erfasst, wie z.B. für die Rechtsprechungsdatenbank das Datum, das Gericht, die Gerichtsbarkeit, das Aktenzeichen, die Fundstelle der Entscheidung sowie die Entscheidungen, die die Ausgangsentscheidung zitieren. Daneben erstellen die Dokumentare Kurztexte (die sog. Orientierungssätze), ordnen den Dokumenten die Notation der *juris*-Sachgebietsgliederung zu und vergeben für alle Dokumente Normen und Schlagworte.⁹ Nach der Konzeption von *juris* war die Stichwortsuche, also die Suche mit den aus den Texten mittels des Programms *PASSAT* extrahierten und auf die Grundform zurückgeführten Wörter der Dokumente selbst¹⁰, lediglich dazu gedacht, die Schwächen der intellektuellen Indexierung auszugleichen.¹¹

3.2 Unzuverlässigkeit der Stichwortsuche

Wenngleich die Stichwortsuche bei *juris* letztlich nur als Kompensation der Schwächen der intellektuellen Aufbereitung gedacht war, steht sie heute eindeutig im Vordergrund, die Nutzung der übrigen Register - vielleicht mit

⁶ Vgl. etwa Moritz, BRAK-Mitt 3/1998, S. 115 ff.; ders., Nachrichten für Dokumentation 49 (1998), S. 490 ff.; Riehm et al. 1989, S. 490 ff.

⁷ Etwa Moritz, jur-pc 1995, S. 3000 ff.; Möller 1993, S. 147; Wolf, jur-pc 1992, S. 1744 (1748).

⁸ Hier ist vor allem das weltweit erste als Volltext-Datenbank aufgebaute juristische System *LEXIS* mit Rechtstexten aus 40 Ländern zu nennen, das vor allem in den USA und Großbritannien im Einsatz ist, Möllers, JuS 2000, S. 1203 (1205); Riehm et al. 1989, S.149.

⁹ Vgl. dazu genauer Mielke 2000, S. 25 ff. und S. 53 ff.

¹⁰ Vgl. dazu genauer Mielke 2000, S. 23 f. und S. 52 f.

¹¹ Vgl. dazu genauer Mielke 2000, S. 64.

Ausnahme des Normregisters - wird allgemein als gering eingeschätzt.¹² Dies ist auch deswegen nicht überraschend, da auch bei der Suche in herkömmlichen Medien überwiegend Schlagwortregister herangezogen werden.¹³

Bei der Stichwortsuche bei *juris* handelt es sich um eine Mischform zwischen Stichwort- und Schlagwortsuche, da die intellektuell vergebenen Schlagworte neben den automatisch extrahierten Stichwörtern in die Recherche miteinbezogen werden.

Gegen die Suche mit Stichwörtern bestehen erhebliche Vorbehalte. So wird grundsätzlich bezweifelt, dass diese Art der Suche für das deutsche Recht geeignet sei, da hier im Gegensatz zum amerikanischen Recht, in dem in weiten Bereichen *case law* vorherrscht, nicht die Suche nach älteren Gerichtsentscheidungen im Vordergrund stehe. Ohne hier in die Einzelheiten zu gehen, sei an dieser Stelle nur darauf hingewiesen, dass auch im deutschen Recht die große praktische Bedeutung von sog. „Präjudizien“ kaum zu bestreiten sein dürfte, wenn auch die herrschende Meinung eine normative Bindungswirkung von Präjudizien ablehnt.¹⁴

Neben diesem rechtsmethodischen Einwand wird die Stichwortsuche vor allem als zu unzuverlässig angesehen. Insbesondere bestünde die Gefahr, dass die Benutzer in großen Datenbanken mit vielen hunderttausend Dokumenten eine nicht mehr handhabbare Menge an Ausgabertexten auf ihre Suchanfrage erhalten.¹⁵ Sucht man z.B. Literatur zu der Rechtsfrage *Besetzungsrüge bzgl. Der Teilnahme eines blinden Richters bei der Augenscheinseinnahme*, erhalte man mit den Suchbegriffen *Richter* und *blind* nicht nur sämtliche prozessualen, gerichtsverfassungsrechtlichen, personalvertretungsrechtlichen, dienst-

¹² Vgl. dazu Kuhlmann, *Demokratie und Recht* 1988, S. 50 (59 f.); Moritz, *BRAK-Mitt* 3/1998, S. 115 (117); Wolf, *jur-pc* 1992, S. 1608 (1612); vgl. auch ausführlich Mielke 2000, S. 65.

¹³ Zu diesem Ergebnis kam eine Untersuchung von Infratest Sozialforschung, Jungjohann et al. 1974, S. 24 f. Vgl. auch Hoffmann 1993, S. 72.

¹⁴ Vgl. dazu ausführlich Mielke 2000, S. 68 f.

¹⁵ Vgl. etwa Moritz, *jur-pc* 1995, S. 3000 (3002): „Freitextsuche führt meist zu großer Vollständigkeit aber auch zu großem Ballast“; ders., *BRAK-Mitt* 3/1998, S. 115 (117). Ebenso Bernecker, *Verwaltungsrundschau* 1994, S. 332 (335); Buhlmann, *NJW-CoR* 2/1996, S. 98 (99); Schallbruch, *Jura* 1995, S. 670; Sellnick, *ThürVBl* 1997, S. 151 (152); Wagner-Döbler, *Informatik und Recht* 1988, S. 458 (460); Wolf, *jur-pc* 1992, S. 1608 (1615). Vgl. auch Tenopir/Ro 1990, S. 74 ff. Eine Ausnahme stellt die Untersuchung von Blair/Maron dar, bei deren Studie zu einer englischsprachigen Rechtsdatenbank die precision weitaus höher war als der recall, Blair/Maron, *Communications of the ACM* 28 (1985), S. 289 (293 ff.).

rechtlichen oder versorgungsrechtlichen Dokumente über blinde Richter, sondern auch solche, in denen sich ein blinder Richter scheiden ließ oder ein sehender Richter von einem Angeklagten entsprechend genannt wurde.¹⁶ Basierend auf dieser Überlegung wird vielfach nahegelegt, dass man mit der Stichwortsuche bevorzugt in Kurztexen und nicht in den Volltexten recherchieren sollte.¹⁷ So bietet die *juris* GmbH in ihrer Benutzeroberfläche *juris formular* als erste Suchmöglichkeit die Suche im Kurztext an, wodurch sich die Suche auf die von Dokumentaren erstellten Kurztexen beschränkt und die Volltexte für die Recherche ausgespart bleiben. Durch diese Gestaltung legt die *juris* GmbH nahe, dass es sich um die erfolgversprechendste Suchmöglichkeit mit Begriffen handelt. Durch eine Einschränkung auf rein formaler Ebene soll damit der Gefahr zu großer Treffermengen begegnet werden.

Den beiden häufig genannten Thesen

- a) die Stichwortsuche führt zu einer zu großen Antwortmenge
- b) im juristischen Bereich ist eine starke intellektuelle Indexierungskomponente erforderlich

soll anhand der Ergebnisse einer Evaluierungsstudie weiter nachgegangen werden. Dabei wurde die Effektivität von Recherchen mittels der Stichwortsuche in *juris* untersucht und mit der Effektivität von Recherchen in einer Datenbank, die statistische Retrievalverfahren nutzt, verglichen.¹⁸ Diese Studie ist dabei dem Umstand zu verdanken, dass ein Teilbereich des Datenbestands, nämlich etwa 400.000 Gerichtsentscheidungen, in beiden Datenbanken in gleichem Umfang gespeichert, aber in unterschiedlicher Weise inhaltlich aufbereitet und damit auf unterschiedliche Art zu recherchieren ist.

4 Evaluierung juristischer Informationssysteme

4.1 Juris vs. QSearch

Bei *juris* handelt es sich um ein „klassisches“ Boolesches Retrievalsystem, bei dem - wie bereits ausgeführt - zur Suche alle sinntragenden Begriffe der gespeicherten Dokumente (Gerichtsentscheidungen, Aufsätze, Gesetze und

¹⁶ Vgl. Wolf, jur-pc 1992, 1608 (1615) sowie ausführlich zu diesem Beispiel Mielke 2000, S. 66 f.

¹⁷ Vgl. etwa Buhlmann, NJW-CoR 2/1996, 98 (99): „Wegen der hohen Trefferzahl eignet sich Volltextsuche meist nicht als alleiniges Suchkriterium“ oder Bettschneider, Versicherungswirtschaft 1998, 550 (551): „Die Suche im Kurztext führt zu exakteren Ergebnissen, daher sollte sie stets bevorzugt vor der Recherche im Gesamttext durchgeführt werden.“ Vgl. dazu auch Mielke 2000, S. 146.

¹⁸ Zum Testaufbau genauer Mielke 2000, S. 105 ff., sowie Mielke 2001, S. 104 ff.

Rechtsverordnungen, Gesetzesmaterialien etc.) in ihrer Grundform sowie die zusätzlich von Dokumentaren intellektuell bzw. manuell vergebenen Suchkriterien zur Verfügung stehen. Bei dem System *QSearch*¹⁹ handelt es sich um ein natürlich-sprachliches System, bei dem der Benutzer mit natürlichsprachlich formulierten Sätzen anfragen kann. Für die Inhaltserschließung kommen statistische Verfahren zum Einsatz, die mit Wortfrequenzen und Termgewichtungen arbeiten.²⁰ Eine manuelle oder intellektuelle Aufbereitung der gespeicherten Dokumente ist nicht gegeben. Die Ausgabe erfolgt nach Relevanz geordnet (*ranking*).

4.2 Testanordnung

An der nachfolgend beschriebenen Evaluierungsstudie nahmen insgesamt 22 Doktoranden zum Zivilprozessrecht teil. Die Testpersonen erhielten zwölf Sachverhalte von Entscheidungen zu zivilprozessualen Rechtsproblemen aus den Jahren 1987 bis 1992. Alle Urteile und Beschlüsse sind veröffentlicht und in beiden Datenbanken enthalten. Die in den Entscheidungen veröffentlichten Sachverhaltsangaben bildeten die Aufgabentexte. Dabei ist der Testaufbau nicht so zu verstehen, dass nach einer *einzelnen* vorgegebenen Entscheidung gesucht werden soll. Vielmehr dienten die gewählten Aufgabentexte dazu, ein allgemeines rechtliches Problem - ohne Vorgabe bestimmter Suchbegriffe - aufzuzeigen, zu denen die Testpersonen relevante Entscheidungen suchen sollten. Den Testpersonen war dabei nicht bewusst, dass sich in der Datenbank eine zu den jeweiligen Sachverhalten ergangene Entscheidung befindet.²¹

Als Sachverhalt 12 dienten beispielsweise folgende Angaben:²²

Die Parteien streiten über die Wirksamkeit der Zustellung eines Vollstreckungsbescheides. Die Klägerin macht aus abgetretenem Recht einen Anspruch aufgrund eines Schuldaner-

¹⁹ Diese Datenbank wurde von dem deutschen Recherchedienst der COMPLEX GmbH genutzt, deren Geschäftsanteile seit 1. September 2000 von der Foris AG übernommen wurden. Der Recherchedienst wird dabei weiterhin über Telefon und Fax betrieben, vgl. <http://www.Foris.de>, Stand: 28. Mai 2002. Zum Recherchedienst der COMPLEX GmbH vgl. Mielke 2000, S. 157 ff.

²⁰ Eine auf der US-amerikanischen Patentschrift für dieses System basierende Beschreibung der wichtigsten statistischen Verfahren findet sich in Mielke 2000, S. 154 ff.

²¹ Das Testdesign ist ausführlich beschrieben in Mielke 2000, S. 107 ff.

²² Die Angaben stammen aus dem in *BGHZ* 111, 1 ff. (=NJW 1990, 1666 f.) veröffentlichten Urteil des VIII. Zivilsenats vom 14. März 1990, Aktenzeichen: VIII ZR 204/89. Alle weiteren Sachverhalte sind in Mielke 2000, S. 188-195 abgedruckt. Eine weitere Übersicht zu den ausgewählten Sachverhalten findet sich in Mielke 2000, S. 197-203.

kenntnisses geltend, durch das der Beklagte nach ihrer Darstellung Kaufpreisforderungen für Warenlieferungen der Zedentin bestätigt hat. Der Beklagte bestreitet sowohl die Warenlieferungen als auch die Echtheit seines Namenszuges unter dem Schuldanerkenntnis. Die Klägerin erwirkte gegen den Beklagten zunächst einen Mahnbescheid und anschließend einen Vollstreckungsbescheid über 8.217,33 DM nebst Zinsen und Kosten. Der Vollstreckungsbescheid wurde am 29. November 1984 in der Wohnung des Beklagten an seine damalige Lebensgefährtin, Frau T., übergeben.

Der Beklagte hatte sich im Jahr 1982 von seiner ersten Ehefrau getrennt und lebt seitdem mit Frau T. und deren minderjährigen Kindern zusammen. Im Dezember 1983 bezog er mit Frau T. und ihren Kindern die Wohnung, in der die umstrittene Zustellung stattfand. Dort führte Frau T. – die der Beklagte zwischenzeitlich geheiratet hat – den gemeinsamen Haushalt, für dessen finanzielle Bedürfnisse der berufstätige Beklagte sorgte.

Der Beklagte hat gegen den Vollstreckungsbescheid, von dessen Existenz er erst aufgrund einer Lohnpfändung erfahren haben will, am 21. Juli 1986 Einspruch eingelegt und in dem anschließenden Verfahren beantragt, den Vollstreckungsbescheid aufzuheben und die Klage abzuweisen. Gegen die Wirksamkeit der Zustellung hat er eingewandt, zum Zustellungszeitpunkt habe zwischen Frau T. und ihm noch keine gefestigte Lebensgemeinschaft bestanden, denn er habe »jederzeit wieder gehen können.« Die Klägerin hat beantragt, den Einspruch als unzulässig zu verwerfen.

Die Testpersonen sollten in einem *paper-and-pencil*-Test die Suchbegriffe angeben, die für eine Datenbankrecherche zu den in den Sachverhalten enthaltenen Rechtsproblemen geeignet erscheinen. Zum voranstehenden Sachverhalt, bei dem es um die Frage geht, ob eine Ersatzzustellung an den nichtehelichen Lebensgefährten des Zustellungsempfängers wirksam ist, konnte man etwa die Suchbegriffe *Zustellung*, *Ersatzzustellung*, *nichteheliche Lebensgemeinschaft* etc. verwenden. Im weiteren Testverlauf dienten die Anfrageformulierungen der Testpersonen als Suchanfragen an die *juris*-Rechtsprechungsdatenbank.²³

Der Test orientiert sich an einer vermittelten Recheresituation, bei der Juristen an professionelle Rechercheure ihre Suchanfragen richten, die dann die eigentliche Datenbankrecherche durchführen. Diese vermittelte Recherche dürfte immer noch in großen Teilbereichen ständige Praxis sein.²⁴ Selbst die *juris*-GmbH bietet für Personen oder Institutionen, die über keinen eigenen *juris*-Zugang verfügen, einen solchen Rechereservice an.²⁵ Auch bei dem Recheredienst der COMPLEX GmbH, der nun von der Foris AG übernommen worden ist, handelt es sich um eine vermittelte Recherche. Unabhängig davon, ob man diese Art der Recherche für angemessen hält oder nicht, ergibt sich für die Testanordnung somit eine gewisse Praxisnähe. Da offensichtlich

²³ Alle Anfrageformulierungen der Testpersonen finden sich in Mielke 2000, S. 205-216.

²⁴ Siehe dazu auch Mielke 2000, S. 108 f.

²⁵ Vgl. www.juris.de unter Bestellen - Rechercheauftrag, Stand 28. Mai 2002.

fehlerhafte Suchstrategien korrigiert wurden, konnten darüber hinaus technische Probleme oder Schwierigkeiten der Testpersonen mit der Abfragesprache aufgefangen werden.²⁶ Die Testergebnisse sind daher auch als unabhängig von Weiterentwicklungen der *juris*-Schnittstelle seit dem Testzeitpunkt in den Jahren 1995 und 1996 anzusehen.

Als Testmaterial standen 264 Aufgabenlösungen zur Verfügung. Da die Testpersonen zu einigen Aufgaben mehrere Alternativlösungen angeboten haben, ergaben sich insgesamt 312 Suchanfragen, die in einem ersten Schritt als Recherchen in der *juris*-Rechtsprechungsdatenbank dienten. Mit den Anfragen zur Hälfte der zwölf Sachverhalte, insgesamt 152 Suchanfragen, wurde in der *QSearch*-Datenbank recherchiert.

4.3 Testauswertung

Die mit den Suchanfragen der Testpersonen in der *juris*-Rechtsprechungsdatenbank recherchierten Dokumente wurden hinsichtlich der Größe der Antwortmengen und der Relevanz der ausgegebenen Dokumente für den jeweiligen Ausgangssachverhalt analysiert. Dabei wurden insgesamt 1.183 Gerichtsentscheidungen auf der Grundlage des Volltextes intellektuell auf ihre Relevanz für das zugrundeliegende Rechtsproblem untersucht.²⁷

Suchgenauigkeit vs. Wiedergewinnungsrate

Um der These näher nachzugehen, dass die Stichwortsuche insbesondere zu unzuverlässig ist und die Gefahr birgt, zu viele irrelevante Dokumente zu erzielen, wurden zunächst die in der *juris*-Rechtsprechungsdatenbank erzielten Antwortmengen in drei Klassen eingeteilt: Nullantworten, Ausgabemengen in einer akzeptablen Größenordnung sowie Antwortmengen mit zu vielen Dokumenten, da *juris* kein *ranking* der Antwortmenge nach inhaltlichen Kriterien kennt. Fraglich ist, ab welcher Grenze es sich noch um ein – was die Menge der nachgewiesenen Dokumente betrifft – akzeptables Ergebnis handelt. Nach der Literatur und auch nach Auffassung professioneller Rechercheure ist bereits eine Trefferzahl von über 20 Dokumenten zu groß und erfordert weitere Einschränkungen.²⁸ Für die hier vorliegende Studie wurde als

²⁶ Zu den vorgenommenen Korrekturen vgl. Mielke 2000, S. 116-121.

²⁷ Vgl. ausführlich auch zur Methode Mielke 2000, S. 130 ff.

²⁸ Möller 1993, S. 92; Viefhues, NJW-CoR 7/97, S. 343. Vgl. auch Turtle 1994, S. 215. Nach Jordan/Konradi-Martin, AnwBl 1994, S. 117 (119) ist ein Suchergebnis bei einer Treffermenge von fünf bis zehn Entscheidungen optimal. Riehm et. al. 1989, S. 235, gehen ebenfalls davon aus, dass man im Normalfall versucht sein wird, die Antwortmenge auf eine übersichtliche Zahl von 10 bis 20 Dokumenten einzuschränken. Relativ

Grenze für eine handhabbare Menge eine Trefferzahl von 30 Dokumenten angenommen.

Dabei zeigt sich folgendes Ergebnis:

Nullantworten	15,15 %
Treffermengen zwischen 1 und 30	61,36 %
Treffermengen über 30	22,35 %
Entweder Nullantworten oder über 30 Treffer ²⁹	1,14 %

Es zeigt sich damit bereits auf dieser Ebene, dass sich die These, die Stichwortsuche führe zu sehr hohen Antwortmengen, nicht bestätigt. Vielmehr scheinen die Benutzer mit der Stichwortsuche – zumindest was die Anzahl der ausgegebenen Dokumente betrifft – gut zurecht zu kommen, wenn bereits in einem Test, der keine Kontrolle über die ausgegebenen Antwortmengen durch interaktives Recherchieren erlaubt und auf einer Datenbasis von ca. 400.000 Entscheidungen zu allen Rechtsgebieten erfolgte, überwiegend handhabbare Mengen gewonnen werden.

Da Benutzer im Fall akzeptabler Antwortmengen wesentlich seltener ihre Suchstrategie grundlegend ändern werden, als dann, wenn sie extrem hohe Ausgabezahlen oder Nullantworten erhalten, ist davon auszugehen, dass ein Großteil der Probleme der Testpersonen, die bereits mit der ersten Anfrage zu brauchbaren Treffermengen kommen, auch bei einer interaktiven Recherche mit der Möglichkeit einer Strategieänderung auftauchen.

Um eine Aussage zur Qualität der recherchierten Dokumente treffen zu können, wurden insgesamt 1.183 Judikate auf ihre Relevanz hin untersucht. Die Relevanzbewertung beschränkt sich dabei auf die Anfragen mit einer Antwortmenge zwischen 1 und 30 Dokumenten, da davon auszugehen ist, dass sich die Testpersonen eine solche Ausgabe ansehen würden, während sie bei den übrigen Trefferkategorien eine Änderung der Suchstrategie vornehmen würden.

Es zeigt sich dabei, dass die *precision* tendenziell höher ist als der *recall*. Von den ausgewerteten 174 Anfragen, die eine Antwortmenge zwischen 1 und 30 Dokumenten erzielen, wird in insgesamt 67 Fällen (entspricht knapp 39 %)

weit geht Janel, jur-pc 1995, S. 3183 (3188), wenn er als verwertbares Endergebnis maximal 50 Fundstellen ansieht. Unabhängig davon mag zu Forschungszwecken das Durchsuchen einer größeren Antwortmenge gerechtfertigt sein.

²⁹ Bezogen auf die verschiedenen Alternativanfragen zu einer Aufgabe.

eine *precision* zwischen 0,5 und 1 erreicht, darunter 43mal (entspricht knapp 25 %) die bestmögliche *precision* 1. Dagegen lässt sich nur für 45 Recherchen (knapp 26 %) ein *recall* von 0,5 und darüber berechnen. Einen *recall* von 1 erreichen lediglich 8 % der Anfragen. Während die *precision* einen Mittelwert von 0,4 erreicht, liegt der *recall* bei durchschnittlich 0,28. Dies bedeutet, dass durchschnittlich 40 % der ermittelten Dokumente relevant sind, gleichzeitig aber nur 28 % aller recherchierbaren relevanten³⁰ Dokumente gefunden werden können. Insgesamt zeigt sich, dass sich die These der mangelnden Genauigkeit als Hauptproblem der Stichwortsuche nicht bestätigen lässt.

Dass der *recall* weiter sinkt, wenn die Suche auf die Kurztexte beschränkt wird, was von einem Großteil der Literatur zur Vermeidung zu hoher Trefferzahlen vorgeschlagen wird³¹, zeigen die Ergebnisse einer Zusatzuntersuchung. Als Testmaterial dienten die hinsichtlich ihres *recall* besten Anfragen jeder Aufgabe, mit denen nun auf den Kurztext beschränkt recherchiert wurde. Als Ergebnis dieses Tests ist festzuhalten, dass sich zwar die Zahl der nicht relevanten Dokumente proportional stärker verringert als die Zahl der relevanten Entscheidungen, dies jedoch mit einem Verlust von ca. 60 % der relevanten Judikate einhergeht. Die Wahrscheinlichkeit, mit der Stichwortsuche relevante Dokumente zu finden, sinkt also erheblich, wenn die Suche auf den Kurztext beschränkt wird.³²

Ergebnisse des Vergleichstests juris – QSearch

Eine Auswertung der ermittelten Dokumentmengen erübrigt sich für die Auswertung des Vergleichstests auf der *QSearch*-Datenbank, da durch das *ranking* das Problem zu großer Antwortmengen entfällt. Ebenso sind aufgrund des *partial match*-Paradigma nur in extremen Ausnahmefällen Nullantworten zu erwarten. Um den gleichen Maßstab wie bei der Auswertung auf der *juris*-Datenbank zu erzielen, bilden die jeweils zuerst angezeigten 30 Dokumente die Grundlage für die Relevanzbewertung.³³ Bewertet man diese Dokumente nach ihrer Relevanz, zeigt sich, dass sich die Zahl der relevanten Dokumente bei der Recherche in der Datenbank mit statistischen Inhaltser-

³⁰ Als Bezugsgröße für die Zahl aller relevanten Entscheidungen diene die Gesamtzahl der von allen Anfragen zu einer Testaufgabe erzielten relevanten Dokumente, vgl. dazu genauer Mielke 2000, S. 95 f. und S. 132. Ein Vergleich mit den in den verschiedenen Kommentaren oder Lehrbüchern zitierten einschlägigen Entscheidungen zeigt, dass der *recall* weiter absänke, wenn man diese mit in die Gesamtzahl einbezöge, da überwiegend nicht alle zitierten einschlägigen Entscheidungen gefunden werden konnten, vgl. dazu genauer Mielke 2000, S. 134 f.

³¹ Vgl. oben Fußnote 17.

³² Zu den genauen Zahlen Mielke 2000, S. 143-145.

³³ Vgl. ausführlich zum Vorgehen Mielke 2000, S. 165 ff.

schließungskomponenten deutlich erhöht. Während die Anfragen bei *juris* nur 175 relevante Dokumente erzielen konnten, waren bei *QSearch* 425 relevante Dokumente zu ermitteln.³⁴

Auch die Zahl der **unterschiedlichen** relevanten Dokumente ist bei *QSearch* höher als bei *juris*. Bei *juris* wurden insgesamt 39 verschiedene relevante Entscheidungen recherchiert, bei *QSearch* 56. Bezieht man die durch die Recherche auf der *QSearch*-Datenbank gewonnenen zusätzlichen relevanten Dokumente in die Gesamtzahl der zu ermittelnden relevanten Dokumente mit ein, verringert sich die Wiedergewinnungsrate bei *juris* auf 20 %. Bei *QSearch* liegt die Rate dann bei 30 %, sie ist damit um die Hälfte höher.

Auch das *ranking* ist als überwiegend zufriedenstellend einzustufen: 80% der Anfragen führen innerhalb der ersten 30 angezeigten Entscheidungen zumindest zu einem relevanten Dokument, wobei sich 56,71 % der relevanten Dokumente bereits innerhalb der ersten zehn angezeigten Dokument befinden.³⁵ *QSearch* gibt damit innerhalb der jeweils ersten zehn angezeigten Dokumente mehr relevante Dokumente aus als *juris* insgesamt.

Insgesamt lassen sich die Ergebnisse wie folgt zusammenstellen:

	<i>Juris</i>	<i>QSearch</i>
Zahl der Anfragen, die die Ausgangsentscheidung finden	50 %	80 %
Gesamtzahl der nachgewiesenen relevanten Dokumente	175	425
Zahl der unterschiedlichen relevanten Dokumente	39	56
<i>Recall</i>	0,2	0,3

5 Fazit

Die Ergebnisse der empirischen Studie zeigen, dass entgegen der Einschätzung der Fachliteratur das Problem der zu großen Antwortmengen bei der Stichwortsuche nicht vorrangig ist. Dies ergibt auch die qualitative Auswertung des *juris*-Tests, wonach die *precision* insgesamt tendenziell höher ist als der *recall*. Damit müssen die Überlegungen zur Verbesserung juristischer Informationssysteme verstärkt dahin gehen, wie der *recall* zu erhöhen ist, was bisher in der Diskussion zu *juris* eine eher geringe Rolle spielt. Auch der Einschätzung vieler Autoren, dass die Suche im Kurztext erfolgversprechender ist, kann vor diesem Hintergrund nicht gefolgt werden.

³⁴ Die Zahlen beziehen sich dabei auf die Anfragen, mit denen in beiden Datenbanken recherchiert wurde.

³⁵ Zu den genauen Zahlen vgl. Mielke 2000, S. 167 und S. 174 ff.

Der Vergleich mit einem Informationssystem, das statistische Verfahren zur Inhaltserschließung nutzt, zeigt, dass die statistischen Verfahren zu einer Erhöhung der Wiedergewinnungsrate führen. Ein natürlichsprachliches System wie *QSearch* ist für den Nutzer damit nicht nur wesentlich komfortabler, sondern stellt auch das leistungsfähigere System dar. Da bei der Stichwortsuche in *juris* auch die intellektuell vergebenen Schlagwörter in die Recherche miteinbezogen werden, ist sogar davon auszugehen, dass der statistische Ansatz diese intellektuelle Indexierungskomponente erfolgreich ersetzt. Eine intellektuelle Inhaltserschließung bringt nur im Bereich der Erfassung formaler Angaben und der Zuordnung von Normen einen Zusatznutzen. Der Test liefert somit auch ein Argument gegen eine gängige Ansicht der deutschen und österreichischen Rechtsinformatik, wonach für Rechtsinformationssysteme intellektuelle Indexierungskomponenten unumgänglich seien. Diese Einschätzung hängt möglicherweise mit der im Vergleich zu anderen Fachgebieten sehr umfangreichen inhaltlichen Erschließung von Rechtstexten in den herkömmlichen Medien zusammen. Es lassen sich aber bislang keine empirisch abgesicherten Hinweise dazu finden, dass die Übertragung dieses Konzepts auf Informationssysteme erfolgversprechend ist.

Auch lässt der Interaktionsmodus von *QSearch* einen mobile access wesentlich realistischer erscheinen als der von *juris*, bei dem nur über ein relativ aufwändig gestaltetes Formular zu recherchieren ist. Damit bietet *QSearch* eine niedrigere Einstiegsschwelle in die Interaktion, die schon heute über einen sprachbasierten, wenn auch bisher nicht automatisierten, Recherchedienst zugänglich ist.

6 Literaturverzeichnis

- Bernecker, H. M.: Das juristische Informationssystem JURIS, Verwaltungsrundschau 1994, S. 332-337.
- Bettschneider, O.: Die Informationsgewinnung mit *juris*. Teil 1: Die *juris* WWW-Oberfläche - Speziallösung für Versicherer, Versicherungswirtschaft 1998, S. 550-553.
- Blair, D. C./Maron, M. E.: An Evaluation of Retrieval Effectiveness for a Full-Text Document-Retrieval System, Communications of the ACM 28 (1985), S. 289-299.
- Buhlmann, D.: Nicht auf Sand gebaut. *Juris* CD-ROM Baurecht im Test, NJW-CoR 2/1996, S. 98-107.
- Hoffmann, H.: Informationstechnik am Richterarbeitsplatz. Gerichts-EDV. Datenbanken. Berechnungsprogramme. Standardsoftware. Entwicklung juristischer Programme, München 1993.

- Jahnel, D.: Elektronische/konventionelle Suchmethoden - ein Gegensatz?, jur-pc 1995, S. 3183-3191.
- Jordan, H./Konradi-Martin, Ch.: Erstattung von juris-Datenbankrecherchekosten. Neue Wege der Rechtsprechung, AnwBl 1994, S. 117-119.
- Jungjohann, K., et al.: Informationsverhalten und Informationsbedarf von Juristen. Eine Erhebung von Infratest Sozialforschung, München, im Auftrag des Bundesministeriums der Justiz und der Gesellschaft für Mathematik und Datenverarbeitung. Teil 1: Analyse-Band, Berlin 1974.
- Kuhlmann, J.: Das juristische Informationssystem JURIS nach der Privatisierung. Auswirkungen auf anwaltliche Chancengleichheit, juristisches Arbeiten und juristisches Denken, Demokratie und Recht 1988, 50 (59 f.);
- Mielke, B.: Bewertung juristischer Informationssysteme. Evaluierung von *juris* im Vergleich zu einem statistischen Information Retrieval-System anhand zivilprozessualer Fragestellungen, Köln, Berlin, Bonn, München 2000.
- Mielke, B.: Wie effektiv sind Recherchen in juristischen Informationssystemen? Ein Vergleich von *juris* mit *QSearch*, in: Schweighofer/Menzel/Kreuzbauer (Hg.), Auf dem Weg zur ePerson, Aktuelle Fragestellungen der Rechtsinformatik 2001, Wien 2001, S. 101-112.
- Möller, T.: *juris* für Juristen. Optimierte *juris*-Nutzung mit den Mitteln der Rechtsinformatik unter Berücksichtigung überkommener juristischer Methodenlehre zugleich ein Beitrag zur Abbildung des deduktiven Hauptschemas der analytischen Begründungslehre als Computermodell. Diss. Jur. Saarbrücken 1993.
- Möllers, Th.: Die juristische Recherche in Bibliotheken und mit neuen Informationstechnologien, JuS 2000, S. 1203-1210.
- Moritz, K.: Optimierung der intellektuellen Ordnungskriterien bei *juris* am Beispiel des Arbeitsrechts: Sachgebietsgliederung, Normen, Schlagwörter, jur-pc 1995, 3000 – 3009.
- Moritz, K.: Nutzen eines *juris*-Anschlusses am juristischen Arbeitsplatz, BRAK-Mitt 3/1998, S. 115-119.
- Moritz, K.: Nutzen der juristischen Datenbank „*juris*“ am Richterarbeitsplatz, Nachrichten für Dokumentation 49 (1998), S. 490-492.
- Riehm, U., et al.: Endnutzer und Volltextdatenbanken. Empirische Untersuchungen zur Nutzung von Volltextdatenbanken in den Fachwelten Medizin, Recht und Wirtschaft. Kernforschungszentrum Karlsruhe, Juni 1989.
- Schallbruch, M.: *Juris*-Datenbanken vorgestellt, Teil 1: Die Rechtsprechungsdatenbank, Jura 1995, S. 670.
- Sellnick, H.-J.: Quo vadis juristisches Informationssystem? Ein Szenario zur Nutzung des Internet für Juristen, ThürVBl 1997, S. 151-154.
- Schweighofer, E./Menzel, Th./Kreuzbauer, G. (Hg.): Auf dem Weg zur ePerson, Aktuelle Fragestellungen der Rechtsinformatik 2001, Wien 2001.
- Turtle, H.: Natural Language vs. Boolean Query Evaluation: A Comparison of Retrieval Performance, in: Croft, W. B./van Rijsbergen, C. J. (Hg.): SIGIR '94. Proceedings of the Seventeenth Annual International ACM-SIGIR Conference on Research and Development in Information Retrieval, Dublin City University 1994, S. 213-220.
- Tenopir, C./Ro, J. S.: Full Text Databases, New York et al. 1990.

Wider einige gängige Ansichten zur juristischen Informationserschließung

Viefhues, W.: juris goes WINDOWS, NJW-CoR 7/97, S. 343-345.

Wagner-Döbler, R.: Karlsruher Juristische Bibliographie und juris: Ein Vergleich, Informatik und Recht 1988, S. 458-461.

Wolf, G.: juris – Ein denkbar einfacher Zugang zu allen Informationen, die Sie brauchen? jur-pc 1992, Teil 1: S. 1524-1536; Teil 2: S. 1568-1577; Teil 3: S. 1608-1619; Teil 4: S. 1676-1691; Teil 5: S. 1744-1752; Teil 6: S. 1801-1810.