

## Verbesserung der „Precision“ bei der juristischen Recherche

### I. Einleitung

Bei Beginn des Aufbaus der elektronischen Rechtsinformation zu Beginn der Achtzigerjahre, war eines der wichtigsten Ziele, die Systeme möglichst schnell mit ausreichend Material zu füllen (Gesetzestexte, Gerichtsentscheidungen, juristische Literatur aus Zeitschriften).

Heute, etwa 20 Jahre später, stellt sich der Dokumentenstatus in den wichtigsten österreichischen Rechtsinformationssystemen<sup>1</sup> wie folgt dar:

- RIS<sup>2</sup> Justiz: ca. 75.000 Originaltexte, ca. 120.000 Rechtssätze
- RDB<sup>3</sup>: ca. 1,8 Millionen Dokumente<sup>4</sup>
- RIDA<sup>5</sup>: ca. 125.000 Indextexte, ca. 170.000 Volltexte

Der jährliche Zuwachs beträgt ca. 10.000 Entscheidungen, ca. 3.000 Aufsätze und Buchbeiträge und ca. 500 Monografien.

Diese Zahlen sprechen für sich und machen deutlich, warum das Problem der juristischen Recherche inzwischen ganz wo anders liegt: Bei dieser Datenmenge wird es immer schwieriger, mit den klassischen Methoden des „information retrievals“ ein präzises bzw. auch nur ein überschaubares und damit in vernünftiger Zeit auswertbares Suchergebnis zu erzielen.

In diesem Beitrag werden die wichtigsten Suchmöglichkeiten nach Gerichtsentscheidungen, nämlich die Suche nach der Geschäftszahl, der Fundstelle, nach Normzitierten und nach Stichworten auf die Präzision des damit erzielbaren Ergebnisses untersucht. Außerdem werden mögliche Lösungen zur Verbesserung der Präzision der juristischen Recherche aufgezeigt.

### II. Definition von „Precision“<sup>6</sup>

Unter „Precision“ wird der Anteil der bei einer Datenbankrecherche nachgewiesenen Dokumente verstanden, die relevant sind.

Oder anders ausgedrückt:

$$\frac{\text{Zahl der nachgewiesenen relevanten Dokumente}}{\text{Zahl aller nachgewiesenen Dokumente}}$$

Dabei ist allerdings zu beachten, dass in der Informationspraxis die Bedürfnisse von Nutzer zu Nutzer bzw. je nach Situation völlig unterschiedlich sind. Bei einer wissenschaftlichen Recherche etwa sollte möglichst alles, was irgendwie von Interesse sein könnte, nachgewiesen werden. Bei der Suche eines Rechtsanwalts unter Zeitdruck hingegen sollte möglichst alles, was sich als irrelevant herausstellen könnte, nicht im Suchergebnis aufscheinen.

---

<sup>1</sup> Siehe dazu ausführlich *Jahnel/Mader*, Rechtsinformatik I<sup>4</sup> (2003).

<sup>2</sup> Rechtsinformationssystem des Bundes, [www.ris.bka.gv.at](http://www.ris.bka.gv.at).

<sup>3</sup> Die Rechtsdatenbank, [www.rdb.at](http://www.rdb.at).

<sup>4</sup> Diese – im Vergleich zum RIS – ungewöhnlich hohe Anzahl an Dokumenten kommt ua durch die Aufnahme des Hohenecker-Index als separate Dokumente (über 500.000 Leitsätze und Aufsatztitel!) und durch die Mehrfachpublikationen von Entscheidungen zustande.

<sup>5</sup> Rechtsindex & Datenbank, [www.rida.at](http://www.rida.at).

<sup>6</sup> *Salton/McGill*, Information Retrieval – Grundlegendes für Informationswissenschaftler (1987) 174 f.

### III. Suchmöglichkeiten nach Gerichtsentscheidungen

#### 1. Geschäftszahl

Auf den ersten Blick erscheint die Geschäftszahl (GZ) einer Gerichtsentscheidung als ein absolut eindeutiger Suchbegriff. In der Regel ist auch einer GZ ein einziger Text zugewiesen. Allerdings gibt es davon Ausnahmen. In einigen Fällen kommt es vor, dass zu ein und derselben GZ mehrere Texte, allerdings fast immer mit unterschiedlichem Datum, existieren:

Rechtssache	Datum	Parteien	Sachgebiet:	
<a href="#">C-186/01</a>	Urteil	2003-03-11	Dory	Sozialpolitik
<a href="#">C-186/01</a>	Schlußanträge	2002-11-28	Dory	Sozialpolitik
<a href="#">C-186/01 R. 1</a>	Beschluß	2001-10-24	Dory	Sozialpolitik

Ergebnis: 3

Hinter der GZ „C-186/01“ etwa verbergen sich drei verschiedene Texte, nämlich die Entscheidung über einen Antrag auf einstweilige Anordnung, der Schlussantrag des Generalanwalts und schließlich das Urteil des EuGH in der Sache selbst.

Im RIS führt die Eingabe einer GZ (hier am Beispiel des Verwaltungsgerichtshofs) dann nicht zu einem eindeutigen Ergebnis, wenn sowohl „Texte“ als auch „Rechtssätze“<sup>7</sup> ausgewählt werden.

1	<input type="checkbox"/>	<a href="#">TE VwGH Erkenntnis 23. April 2003 2002/08/0270</a>
2	<input type="checkbox"/>	<a href="#">RS* VwGH Erkenntnis 23. April 2003 2002/08/0270</a>
3	<input type="checkbox"/>	<a href="#">RS* VwGH Erkenntnis 23. April 2003 2002/08/0270</a>
4	<input type="checkbox"/>	<a href="#">RS* VwGH Erkenntnis 23. April 2003 2002/08/0270</a>
5	<input type="checkbox"/>	<a href="#">RS* VwGH Erkenntnis 23. April 2003 2002/08/0270</a>
6	<input type="checkbox"/>	<a href="#">RS* VwGH Erkenntnis 23. April 2003 2002/08/0270</a>
7	<input type="checkbox"/>	<a href="#">RS* VwGH Erkenntnis 23. April 2003 2002/08/0270</a>
8	<input type="checkbox"/>	<a href="#">RS* VwGH Erkenntnis 23. April 2003 2002/08/0270</a>
9	<input type="checkbox"/>	<a href="#">RS* VwGH Erkenntnis 23. April 2003 2002/08/0270</a>
10	<input type="checkbox"/>	<a href="#">RS* VwGH Erkenntnis 23. April 2003 2002/08/0270</a>
11	<input type="checkbox"/>	<a href="#">RS* VwGH Erkenntnis 23. April 2003 2002/08/0270</a>
12	<input type="checkbox"/>	<a href="#">RS* VwGH Erkenntnis 23. April 2003 2002/08/0270</a>

Bei der RDB schließlich führt die Aufnahme von Mehrfachpublikationen ein und derselben Entscheidung in mehreren Zeitschriften zu einer Redundanz (hier: 4 Ob 248/02b):

<sup>7</sup> Dabei handelt es sich um Kurzfassungen der Entscheidungen, die von den Evidenzbüros der Höchstgerichte erstellt werden.

100%	RDB-ENTS	17.12.2002	<a href="#">VRInfo 2003: OGH 17. 12. 2002, 4 Ob 248/02b</a>	€ 11kB	<input type="checkbox"/>			
OGH zur Zulässigkeit von Links auf Wetterkarten. Der Wetterinformationsdienstleister Meteo-data Wetteranalysen GmbH sorgte schon vor längerem für Aufsehen, indem er Website-Betreibern (egal ob Private oder Unternehmer) - welche ...								
100%	RDB-ENTS	17.12.2002	<a href="#">ecolex 2003: OGH 17. 12. 2002, 4 Ob 248/02b</a>	€ 16kB	<input type="checkbox"/>			
3. Die so geschaffene Möglichkeit begründet auch weder Rufausbeutung noch bringt die Bekl die Kl dadurch in zu missbilligender Weise um die Früchte ihrer Arbeit, kann doch auch die Kl aus der Ausgestaltung des Copyright-Vermerks als Link auf ihre eigene Homepage Vorteile dadurch ziehen, dass ihre...								
100%	RDB-ENTS	17.12.2002	<a href="#">ÖBI-LS 2003: OGH 17. 12. 2002, 4 Ob 248/02b</a>	€ 6kB	<input type="checkbox"/>			
100%	RDB-ENTS	17.12.2002	<a href="#">ÖBI 2003: OGH 17. 12. 2002, 4 Ob 248/02b</a>	€ 36kB	<input type="checkbox"/>			
Die Kl betreibt ein Dienstleistungsunternehmen zur Erstellung von Wetterkarten, Wetterprognosen, Wetteranalysen und Wettergutachten; sie bietet auch eine direkte Einstellung von Wetterprognosen jeweils gewünschter Länder oder Regionen auf die Website ihrer Kunden gegen Entgelt an. Seit 1997 ...								
100%	RDB-ENTS	17.12.2002	<a href="#">ÖBI-LS 2003: OGH 17. 12. 2002, 4 Ob 248/02b</a>	€ 3kB	<input type="checkbox"/>			
100%	RDB-ENTS	17.12.2002	<a href="#">MR 2003: OGH 17. 12. 2002, 4 Ob 248/02b</a>	€ 27kB	<input type="checkbox"/>			
6. Wird im Weg des Framing auf fremde Webseiten verwiesen, die durch einen Copyright-Vermerk gekennzeichnet sind, liegt darin keine sittenwidrige unmittelbare Leistungsübernahme. Es wird damit auch kein fremdes Arbeitsergebnis erschlichen, durch Vertrauensbruch erlangt oder zur Behinderung des ...								
100%	RDB-ENTS	17.12.2002	<a href="#">wbl 2003: OGH 17. 12. 2002, 4 Ob 248/02b</a>	€ 21kB	<input type="checkbox"/>			
Sind die auf einer Site abrufbaren und mittels Link auf einer anderen Site sichtbar gemachten Wetterkarten Werke iSd § 1 UrhG und hat der Urheber unbeschränkte Werknutzungsrechte daran übertragen, hat der "Linksetzer" urheberrechtlich dafür einzustehen, dass ...								
100%	RDB-ENTS	17.12.2002	<a href="#">KRES: OGH 17. 12. 2002, 4 Ob 248/02b</a>	€ 14kB	<input type="checkbox"/>			
4. Der Verlust möglicher Werbeeinnahmen ist nur ein unbeabsichtigter Nebeneffekt und stellt für sich allein keine Wettbewerbswidrigkeit dar. Die klagende Partei betrieb ein Dienstleistungsunternehmen zur Erstellung von Wetterkarten, Wetterprognosen, Wetteranalysen und ...								
100%	RIS-RESA	17.12.2002	<a href="#">OGH 17. 12. 2002, 4 Ob 248/02b</a>	€ 8kB	<input type="checkbox"/>			
100%	RIS-RESA	17.12.2002	<a href="#">OGH 17. 12. 2002, 4 Ob 248/02b</a>	€ 3kB	<input type="checkbox"/>			
100%	RIS-ENTS	17.12.2002	<a href="#">OGH 17. 12. 2002, 4 Ob 248/02b</a>	€ 25kB	<input type="checkbox"/>			

## 2. Fundstelle

Entscheidungen werden häufig nicht (nur) mit der GZ, sondern auch mit der Fundstelle in Fachpublikationen zitiert. Erfolgt die Zitierung ausschließlich mit einer Fundstelle, dann kann der Text der Entscheidung nur in den Datenbanken RDB und RIDA, nicht aber im RIS gefunden werden, weil im RIS keine Literaturfundstellen gespeichert werden.

Wie sieht es nun mit der Präzision dieses Suchbegriffes aus?

Wird eine Entscheidung unter Angabe der Seite einer Fachzeitschrift zitiert, kann dies eindeutig sein, muss aber nicht, wie das folgende Beispiel zeigt. In der SWK (Sozial- und Wirtschaftskartei) sind im Jahrgang 2003 auf Seite 141 sieben Entscheidungen des Verwaltungsgerichtshofs in stark verkürzter Form abgedruckt:

Autor/Gericht/GZ	Fundstelle(n)
1. VwGH, 98/13/0180	ÖStZB 2002/632, 798 = SWK 2003, R 7 = SWK 2003, 141
2. VwGH, 2000/13/0029	ÖStZB 2002/852, 1068 = SWK 2003, 141 = SWK 2003, R 7
3. VwGH, 96/13/0168, 96/	SWK 2003, R 7 = SWK 2003, 141
4. VwGH, 98/13/0017	SWK 2003, R 7 = SWK 2003, 141
5. VwGH, 98/13/0105	SWK 2003, R 7 = SWK 2003, 141
6. VwGH, 98/13/0182	SWK 2003, R 7 = SWK 2003, 141
7. VwGH, 98/13/0195, 98/	SWK 2003, R 7 = SWK 2003, 141

Werden die Entscheidung allerdings mit einer laufenden Nummer versehen, wie dies inzwischen in zahlreichen Fachpublikationen der Fall ist, dann ist das Fundstellenzitat eindeutig. In diesem Fall sollte allerdings zusätzlich auch die Seite in der Zeitschrift zitiert werden, um dem Benutzer ein leichteres Auffinden der Textstelle zu ermöglichen: zB *ecolex* 2003/109, 247 = Entscheidung Nummer 109 auf Seite 247.

Ein weiteres Problem mit Fundstellenzitaten ergibt sich in den Fällen, in denen die jeweilige Quelle nicht in der Datenbank enthalten ist, wie dies zB bei den Amtlichen Sammlungen (SZ, SSt, VfSlg, VwSlg) in der RDB der Fall ist (Ausnahme: SZ [Sammlung Zivilrecht] vor 1986).

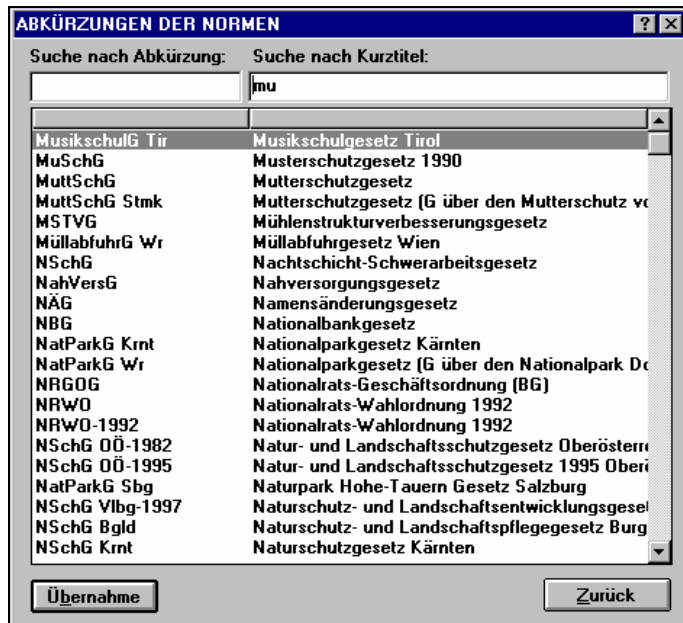
Fundstellenzitate mit Entscheidungsnummern in diesen Sammlungen sind präzise nur in der Datenbank RIDA recherchierbar.

### 3. Normzitate

Normzitate - womit die Gesetzesstellen gemeint sind, auf denen die jeweilige Entscheidung basiert – umschreiben in der Regel das juristische Problem präziser als Stichworte. Es entfallen dabei nämlich die Erschwernisse von Synonymen, unterschiedlichen Fachbegriffen für dasselbe Rechtsproblem, zusammengesetzte Worte, Verwendung von Einzahl oder Mehrzahl etc. Allerdings gibt es etliche Normzitate, die so häufig sind, dass sie allein deshalb kein verwertbares Suchergebnis liefern können: zB § 1295 ABGB für Schadenersatzrecht oder § 1 UWG für unlauteren Wettbewerb.

Ein weiteres Hindernis für den Einsatz von Normziten als Suchbegriffe ist das Problem der richtigen Abkürzung. Das bekannteste Beispiel dafür ist „MSchG“: Diese Abkürzung kann für Markenschutzgesetz, Musterschutzgesetz oder Mutterschutzgesetz stehen. Besonders schwierig zu lösen wird die Frage der Abkürzungen bei Landesgesetzen, weil diese in den AZR<sup>8</sup> nicht enthalten sind.

Abhilfe schaffen kann hier die Abkürzungsliste von RIDA, die laufend aktualisiert wird und zurzeit etwa 2500 Eintragungen enthält:



Ein weiteres Problem der Normzitate besteht darin, dass Sie nur dann in einer Datenbank erfolgreich als Suchkriterium eingesetzt werden können, wenn sie nicht nur im Fließtext vorkommen, sondern in einem eigenen Feld als Suchbegriff zur Verfügung gestellt werden. Dies erfordert allerdings eine intellektuelle Bearbeitung der Entscheidungstexte.<sup>9</sup>

In den meisten Zeitschriften werden die Normzitate vor den Entscheidungstext gestellt, allerdings manchmal etwa in folgender Form:

§§ 12, 32 Abs 1 und 94 Abs 1 Z 1 GBG<sup>10</sup>.

<sup>8</sup> Friedl/Loebenstein, Abkürzungs- und Zitierregeln der österreichischen Rechtssprache und europarechtlicher Rechtsquellen<sup>5</sup> (2001).

<sup>9</sup> Was auch bei den meisten Rechtsdatenbanken erfolgt, wenn auch in unterschiedlicher Intensität. Vgl dazu Jahnel, Indexsuche und Volltextsuche im Vergleich, in: Schweighofer/Menzel/Kreuzbauer (Hrsg), IT in Recht und Staat (2002) 197.

<sup>10</sup> GBG = Grundbuchgesetz.

Als Abkürzung für ein und denselben Paragraphen des Notariatsaktgesetzes finden sich in der Zeitschriftenliteratur mehrere unterschiedliche Schreibweisen:

§ 1 Abs 1 lit d NZwG / § 1 Abs 1 lit d NotZwG/ § 1 lit d NZwG.

Beides erschwert eine gezielte Suche nach Normzitate in elektronischen Datenbanken. Der einzige Ausweg besteht darin, für eine Einheitlichkeit der Normzitate zu sorgen. Dies ist im Rahmen der Zeitschriftenpublikationen bei der Vielzahl der Medien und Autoren nicht möglich, sondern kann nur in der Datenbank selbst erfolgen. So wird etwa in der Normenliste von RIDA nur eine einzige Schreibweise für eine Gesetzesabkürzung erlaubt.

#### 4. Stichworte

Stichworte stellen in der Praxis – auch bei der juristischen Recherche – die am häufigsten verwendeten Suchbegriffe dar. Bei der Formulierung der Suchstrategie mit Stichworten sind allerdings verschiedene Erschwernisse zu berücksichtigen:

- unterschiedliche Endungen durch verschiedene Fälle;
- Einzahl, Mehrzahl, unregelmäßige Mehrzahlbildung;
- zusammengesetzte Wörter;
- mehrgliedrige Suchbegriffe (zB: „verbotene Ablöse“ oder „Einstweilige Verfügung“);
- Aufspüren von Synonymen.

Zu diesen Problemen kommt noch ein weiteres hinzu, weil die Begriffsverwendung in der Terminologie der Höchstgerichte keineswegs einheitlich erfolgt. So wird etwa in den Entscheidungen des VwGH neben dem am häufigsten verwendeten Begriff „Lenkeraskunft“ gleichbedeutend „Lenkererhebung“ bzw „Lenkeranfrage“ verwendet. Dies kann bei der Eingabe der Suchbegriffe – hier am Beispiel von RIDA – etwa so berücksichtigt werden:

The screenshot shows the 'SUCHSCHIRM: Version RIDA plus II' window. It contains several input fields and options:

- Autor/Gericht:** dropdown menu
- Fundstelle:** dropdown menu
- Geschäftszahl:** dropdown menu
- von Datum:** date field (empty)
- bis Datum:** date field (24.09.2003)
- RIDA-Nummer:** text field (empty)

The search criteria section is titled 'SUCHBEGRIFFE (Norm/Stichwort)' and 'Verknüpfung':

SUCHBEGRIFFE (Norm/Stichwort)	Verknüpfung:
1. Lenkeraskunft	ODER
2. Lenkererhebung	ODER
3. Lenkeranfrage	ODER
4. § 103 Abs 2 KFG	UND
5. Tatort	
6.	

Search options on the right:

- Alles
- Entscheidungen
- Literatur
- Erlässe

Advanced search options (all checked):

- Rechtstrunkierung
- Linkstrunkierung
- Aufsatztitel durchsuchen
- Leitsätze durchsuchen
- VOLLTEXTSUCHE

Buttons at the bottom: **Listen einblenden**, **Löschen**, **Ändern**, **SUCHEN**, **Hilfe**, **Ende**

Die Präzision des Suchergebnisses bei Eingabe von Stichworten ist recht unterschiedlich. Sie wird umso höher, je spezieller der gewählte Suchbegriff ist bzw je mehr er vom allgemeinen Sprachgebrauch abweicht. Insbesondere von der Lehre oder Rechtsprechung entwickelte Begriffe wie zB „Linkhaftung“ oder „Schockschaden“ ermöglichen eine besonders zielgerichtete Recherche.

Problematisch wird die Suche nach Fachbegriffen dann, wenn diese sehr häufig in den Texten vorkommen (zB „Haftung“) oder sprachlich mehrdeutig sind (das Wort „links“ im Englischen und im Deutschen).

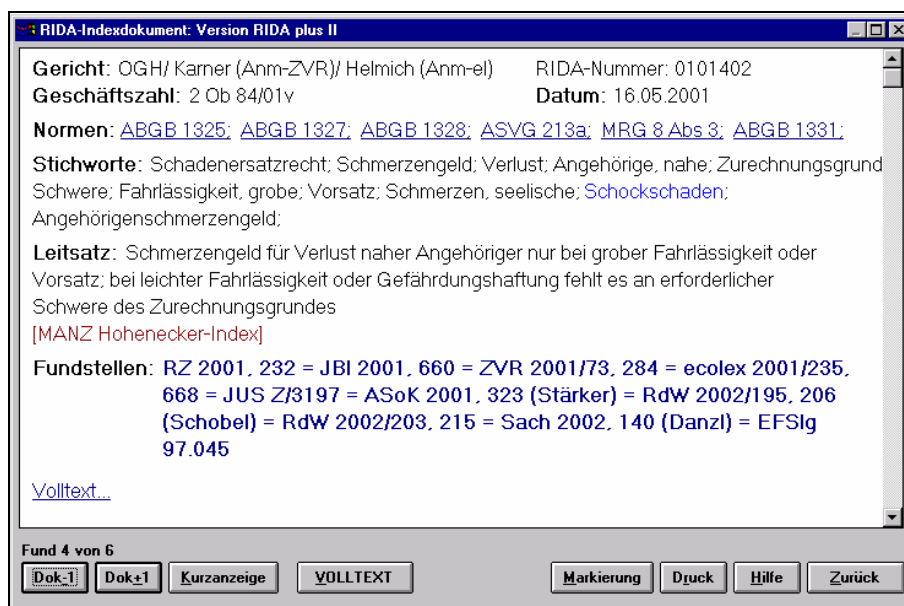
Folgende Lösungen sind denkbar:

- Vereinheitlichung durch Gerichte/Autoren
- Einschränkung der Suche auf Rechtsätze (RIS)
- Einschränkung der Suche auf Leitsätze/Kurztexte (RDB)
- Relevanzsortierung (RDB)
- Verschlagwortung durch Fachleute mit alternativer Volltextsuche (RIDA)

Eine Vereinheitlichung schon von Seiten der Gerichte bzw der Fachautoren wäre zwar wünschenswert, diese Forderung ist in der Praxis aber nicht durchsetzbar. Dies wegen der „Freiheit“ der Autoren einerseits und wegen der großen Datenmenge (siehe Punkt I.) andererseits.

Ein möglicher Ausweg ist die Einschränkung der Suche auf bestimmte Teilbereiche der Texte, wie dies im RIS (Rechtssätze) oder RDB (Leitsätze/Kurztexte) möglich ist. Damit ist ein kleineres Suchergebnis gewährleistet, was allerdings auf Kosten der Vollständigkeit geht. Einen weiteren Ansatz zur Problemlösung bildet ein technischer, nämlich eine Sortierung des Suchergebnisses nach „Relevanz“ durch eine entsprechende Software, wie dies in der RDB seit Anfang 2004 geschieht. Dabei wird davon ausgegangen, dass ein Computerprogramm nach bestimmten Kriterien, wie zB Anzahl der Vorkommnisse der Suchworte, Gewichtung bestimmter Kategorien etc., eine „richtige“ Reihung vornimmt und auch bei einem großen Suchergebnis die relevanten Treffen am Anfang der Liste stehen. Schon allein wegen der Inhomogenität der gespeicherten Texte und der Relativität der „Relevanz“ (siehe Punkt II.) sind jedenfalls bei juristischen Texten gewisse Zweifel angebracht. Entsprechende Untersuchungen stehen allerdings noch aus.

Schließlich besteht noch die Möglichkeit einer Bearbeitung der Dokumente in Form einer Verschlagwortung und der Bildung eines kurzen Leitsatzes durch Fachleute, wie dies bei der Datenbank RIDA geschieht. Damit kann die Suche – je nach Situation – auf den Index eingeschränkt werden oder zusätzlich auf die Volltexte ausgedehnt werden:



Der Unterschied in der Präzision des Suchergebnisses wird an zwei Beispielen demonstriert (Stand: Oktober 2003):

Suche 1: „Links“ UND „Haftung“  
Indeksuche: 27 Treffer  
Indeksuche und Volltextsuche: 975 Treffer

Suche 2: Rechtsanwalt und Liebhaberei  
Indeksuche: 9 Treffer  
Indeksuche und Volltextsuche: 89 Treffer.

#### **IV. Zusammenfassung**

Die Anzahl der Dokumente in juristischen Datenbanken ist in den letzten Jahren sehr stark angestiegen. Dies führt dazu, dass es zunehmend schwieriger wird, bei Recherchen über den gesamten Volltext ein verwertbares Ergebnis zu erzielen. Auch mit dem Einsatz von Bool'schen Operatoren ist das Problem schwer in den Begriff zu bekommen. Immer wieder tritt der Fall auf, dass die Anzahl der Treffer zunächst für eine Auswertung zu hoch ist, das Hinzufügen eines weiteren Operators mit einer UND-Verknüpfung das Ergebnis aber auf Null reduziert.

Gerade bei juristischen Datenbanken liegt der Grund dafür ua darin, dass relativ häufig bestimmte Worte, die als Suchbegriff in Frage kommen, sowohl als juristischer Fachbegriff als auch in der normalen Umgangssprache verwendet werden. Daneben tritt noch das Problem, dass bei Gerichtsentscheidungen auch ähnlich gelagerte Fälle relevant sein können. In diesem Beitrag werden verschiedene Möglichkeiten vorgestellt, wie eine präzisere Abfrage in juristischen Datenbanken erreicht werden kann: zB Suche nach Normzitate, Einschränkung der Suche auf Rechtssätze (RIS) bzw auf Leitsätze/Kurztexte (RDB) oder die Erstellung eines "Meta-Dokuments" wie das Indextdokument der RIDA-Datenbank. Dabei zeigt sich, dass sich die Präzision zumindest bei juristischen Datenbanken durch eine qualitativ hochwertige intellektuelle Bearbeitung stark verbessern lässt.

#### **Kontaktadresse:**

ao. Univ.-Prof. Dr. Dietmar Jahnel  
Institut für Verfassungs- und Verwaltungsrecht  
Universität Salzburg, Kapitelgasse 5-7, A-5020 Salzburg  
E-mail: Dietmar.Jahnel@sbg.ac.at