

Wie gut suchen die österreichischen Rechtsdatenbanken?

Anton Geist

Juristische Fakultät, Universität Edinburgh (Postgraduate Student)

Arbeitsgruppe Rechtsinformatik, Universität Wien (Projektassistent)

Old College, Edinburgh, Großbritannien

home@antongeist.com

Schlagnorte: Rechtsdatenbank, Evaluierung, Retrieval Test, RDB, LexisNexis, RIDA

Abstract: Evaluierungen von österreichischen Rechtsdatenbanken haben es bisher aus nachvollziehbaren, aber praxisfremden Gründen vermieden, die Datenbanken anhand ihrer konkreten Suchergebnislisten zu vergleichen. Es wird eine kleine Studie des Autors beschrieben, die eine grobe Einschätzung der Rechercheleistung der Datenbanken ermöglichen soll. Das ideale Suchergebnis wird dabei mithilfe eines Gesetzeskommentars erstellt. Trotz der Ungenauigkeit dieser Retrieval Studie wird argumentiert, dass sich vor allem das allgemeine Verbesserungspotential der österreichischen Rechtsdatenbanken anhand ihrer konkreten Recall- sowie Precision-Werte zeigt.

1. Einleitung

In Österreich stehen eine Reihe kommerzieller Rechtsdatenbanken zur Verfügung, um Benutzer(innen) bei der juristischen Recherche zu unterstützen. Verschiedene Faktoren machen eine Bewertung der von den einzelnen Anbietern gelieferten Ergebnislisten schwierig und unpräzise. Von Jurist(inn)en verfasste, wissenschaftliche Bewertungen der Datenbanklandschaft klammern daher Vergleiche von Trefferlisten fast immer aus. Im Gegensatz dazu ist die Frage, wie gut oder schlecht eine Rechtsdatenbank in ihrem Datenbestand generell sucht, eine der ganz zentralen Überlegungen der (potentiellen) Benutzer(innen). Es ist somit zu überlegen, ob nicht – salopp formuliert – ein ungenauer Vergleichstest insgesamt immer noch besser ist als gar keiner.

Vor diesem Hintergrund beschreibt mein Artikel eine kleine Studie, die ich zur Einschätzung der Retrievalqualität von österreichischen Rechtsdatenbanken durchgeführt habe. Sie soll hauptsächlich als erste grobe Orientierung sowie als Diskussionsgrundlage für zukünftige „Retrieval Tests“ dienen.

2. Einschlägige Forschungsarbeiten

Was Österreich und Deutschland betrifft, so lassen sich die umfassenden Evaluierungen von Rechtsdatenbanken der letzten Jahre an einer Hand abzählen: In Deutschland haben sich Kraft, Noack/Kremer sowie Mielke mit dem Thema auseinandergesetzt, in Österreich bisher lediglich Liebwald. In Österreich sind zusätzlich zwei aktuelle wissenschaftliche Fachartikel zu nennen: Es handelt sich um Beiträge von Staudegger sowie Mader (siehe „5. Literatur“).

3. Studie zur Retrievalqualität österreichischer Rechtsdatenbanken

3.1 Die Ausgangssituation

Jede Evaluierung von Suchmaschinen ist von Beginn an mit einem, vermeintlich enormen, Problem konfrontiert: Jede Recherche hat eine große individuelle Komponente, es liegen nie völlig objektive Daten über ideale Trefferlisten vor. Jede(r) Benutzer(in) hat gegenüber der Suchmaschine einen anderen Anspruch an Vollständigkeit und Treffsicherheit der Ergebnisliste, zumal er oder sie ein unterschiedliches Vorwissen mitbringt. Folglich könnte

streng genommen nur jede(r) Suchende selbst in der konkreten Recheresituation nach Durchsicht aller (!) potentiell „relevanten“ Texte beurteilen, wie gut die ihm oder ihr zunächst angezeigte Trefferliste tatsächlich war.

Zusammen mit den zum Teil sehr unterschiedlichen Datenbeständen der österreichischen Anbieter digitaler Rechtstexte hat diese Subjektivität der Relevanz bisher dazu geführt, dass meines Wissens kein öffentlich zugänglicher Retrieval-Test österreichischer Rechtsdatenbanken verfügbar ist.

Die subjektive Komponente jeder Evaluierung von Suchmaschinen sollte jedoch nicht dazu führen, auf Retrievaltests von vornherein zu verzichten. Das Hauptkriterium für das Verwenden einer Suchmaschine ist ihre Retrievalleistung. Dementsprechend müsste ein – wenn auch zwingend subjektiver – Vergleich anhand von Trefferlisten immer noch wünschenswerter sein als der bloße Hinweis darauf, dass ein solcher aufgrund von mangelnder Objektivierbarkeit nicht möglich sei.

3.2 Die Vorgaben für den Retrieval Test

Nachdem noch keinerlei entsprechende Daten vorliegen, bestand mein Ziel darin, die kommerziellen Rechtsdatenbanken in Hinblick auf ihre Rechercheleistung zu testen. Konkret sollten die Datenbanken RDB, LexisNexis sowie RIDA untersucht werden.

Ich musste folgende Ausgangssituation schaffen: Erstens war eine Suchanfrage nötig, die auch in der Praxis so denkbar wäre und die von allen Datenbanken gleich gut gelöst werden könnte. Das bedeutet, dass die Anbieter, was potentiell relevante Dokumente betrifft, einen identen Grunddatenbestand aufweisen mussten. Da die kommerziellen Volltextanbieter (RDB, LexisNexis) den Großteil ihrer Zeitschrifteninhalte exklusiv anbieten, fiel die juristische Fachliteratur als Suchinhalt somit von vornherein weg.

Zweitens musste ich einen Weg finden, eine ideale Ergebnismenge zu definieren, die jener des/der Durchschnittssuchenden möglichst nahe kommen würde. Juristische Recherchehandbücher nennen einheitlich den juristischen Kommentar als beste Anlaufstelle für eine detaillierte und umfassende Recherche. Auch in der juristischen Praxis wird dieser Ansatz verfolgt: Hat man die relevanten Normen für einen Recherchefall identifiziert und den einschlägigen Gesetzeskommentar durchgearbeitet, so fühlt man sich meist ausreichend vorbereitet. Es lag somit nahe, die ideale Trefferliste in irgendeiner Form mit Hilfe eines Gesetzeskommentars zu erstellen.

3.3 Fiktive Suchanfrage und ideale Treffermenge

Ich definierte meine fiktive Recheresituation als wohnrechtliche Judikurrecherche. In diesem Bereich verfügen alle Anbieter über denselben (RIS-)Grunddatenbestand. „Finden Sie alle wichtigen OGH Urteile betreffend § 1096 ABGB zwischen 2005 und 2007!“ lautete die fiktive Rechercheaufgabe.

Die zu erwartenden Treffermengen sollten für mich als Einzelperson handhabbar bleiben, was zu den Einschränkungen auf eine Gericht (OGH), eine Norm (§ 1096 ABGB) sowie einen Zeitraum von drei Jahren (2005-2007) geführt hat.

Die Auswahl des § 1096 ABGB, der die Erhaltungspflichten zwischen Mieter und Vermieter dispositiv regelt, ist auch deswegen erfolgt, weil in diesem Bereich ein aktueller und renommiertes Gesetzeskommentar zur Verfügung steht, nämlich der Manz Online-Kommentar zum Wohnrecht von Prader (siehe „5. Literatur“). Der Erstellung der idealen Trefferliste aus dem Kommentar habe ich die These zugrunde gelegt, dass Entscheidungen dann wichtig sind, wenn sie im besagten Kommentar genannt werden. Durch manuelles Durchsuchen der Kommentierung zu § 1096 ABGB wurde eine Gesamtmenge von 50 Judikaturzitate extrahiert. Diese 50 OGH-Entscheidungen nennt der Kommentar in den Erläuterungen zu § 1096 ABGB, wobei es sich „netto“ um 27 verschiedene Entscheidungen

handelt, die im Kommentar zwischen 1mal und 4mal zitiert wurden. Diese 27 Entscheidungen bildeten somit die Messlatte, wenn es um die folgenden Ergebnislisten der Rechtsdatenbanken ging. Der Einfachheit halber wurde die Häufigkeit der Zitate nicht weiter beachtet.

3.4 Die Ergebnisse der Rechtsdatenbanken

Vor dem Hintergrund der 27 durch den Kommentar als „wichtig“ ausgewiesenen OGH-Entscheidungen wurden die österreichischen Rechtsdatenbanken LexisNexis, Uni-RDB, Business-RDB, sowie RIDA zur Rechercheunterstützung genutzt.

Den Datenbanken wurden über ihre jeweiligen Suchmasken jene Informationen mitgeteilt, die auch bei der Benützung des Kommentars verfügbar waren: Gesucht wurde nach dem Dokumenttyp „Entscheidungen“ vom Autor/Gericht „OGH“, betreffend die Gesetzesnorm „§ 1096 ABGB“. Der Zeitraum wurde mit „2005-2007“ festgelegt. Sofern Datenbanken standardisierte Listen (RIDA) oder spezielle (Judikatur-)Suchfelder (alle anderen) für bestimmte Informationen zur Verfügung stellten, wurde von diesen Gebrauch gemacht.

Zur Sicherheit habe ich zunächst durch Eingabe der – dem Suchenden freilich nicht bekannten – (eindeutigen) Geschäftszahlen festgestellt, ob alle 27 Entscheidungstexte in allen Datenbanken prinzipiell verfügbar wären. Nachdem dies der Fall war, wurde nun eine Datenbank nach der anderen abgefragt und die Trefferlisten verglichen. Auch die Ergebnislisten der Rechtsdatenbanken mussten um mehrfache Fundstellen bereinigt werden, nachdem auch die ideale Treffermenge keine Mehrfachnennungen zulässt.

Die Ergebnislisten der 4 Datenbanken (Stichtag: 1. Juni 2008 [Business-RDB] bzw. 15. April 2008 [alle anderen]) fielen sehr unterschiedlich lang aus: Die Uni-RDB lieferte 41 Treffer, wobei sich die Zahl auf 26 Treffer reduzierte, nachdem mehrfache Treffer ausgeschieden wurden. Die Business-RDB lieferte 127 Treffer, wobei es sich „netto“ um 40 Entscheidungen handelte. Die RIDA gab eine Treffermenge von 29 Entscheidungen aus. Bei RIDA gab es keinerlei Mehrfachtreffer, nachdem RIDA Mehrfachfundstellen in einem Eintrag zusammen fasst. LexisNexis bildete eine Trefferliste aus 6 Dokumenten, wobei es sich „netto“ um 4 Entscheidungen handelte. Diese Trefferliste von LexisNexis war identisch mit jener, die das öffentliche RIS Rechtsinformationssystem mit dieser Abfrage produzierte. Diese Zahlen sagen freilich noch wenig über die Qualität der Ergebnisse aus. Geht man von den 27 idealen (Kommentar-)Treffern aus und überprüft, ob und wenn ja in welchen Datenbank-Ergebnislisten diese aufscheinen, so zeigt sich folgendes Bild:

8 der 27 Entscheidungen konnten von keiner Datenbank gefunden werden. In der Ergebnisliste der RIDA fanden sich 16 „wichtige“ Entscheidungen, in jener der Uni-RDB 12. Die Datenbank LexisNexis lieferte mit 2 einschlägigen Treffern das schlechteste Ergebnis. Die Business-RDB Datenbank lieferte mit 19 Treffern das beste Ergebnis. Das Finden dieser 19 Treffer erforderte allerdings bei einer Brutto-Trefferliste von 127 Einträgen einige Zeit.

Allgemein üblich ist es, im Rahmen von Retrieval Tests Recall- und Precision-Werte zu errechnen¹: Der Recall-Wert fragt, wie viele der relevanten Dokumente in einer Trefferliste aufscheinen. Er misst somit die Vollständigkeit der Ergebnisse, sagt aber noch nichts über die Treffsicherheit einer Ergebnisliste aus. Diese Lücke füllt die Angabe der Precision, die die relevanten Treffer mit dem Gesamtumfang der Trefferliste vergleicht und somit die Treffsicherheit misst. Die entsprechenden Daten meines Retrieval Tests sind:

	Recall (Vollständigkeit)	Precision (Treffsicherheit)
RDB - Business	19/27 = 70%	19/40 = 48%
RDB - Universion	12/27 = 44%	12/26 = 46%
RIDA	16/27 = 59%	16/29 = 55%

¹ Eine genaue Beschreibung ihres bis heute oft zitierten Retrieval-Tests im juristischen Bereich geben Blair und Maron in ihrer Publikation aus dem Jahr 1985 (siehe „5. Literatur“).

LexisNexis	02/27 = 7%	02/04 = 50%
------------	------------	-------------

Abbildung 1: Recall und Precision der Ergebnislisten

4. Schlussfolgerungen und Diskussion²

Nochmals sei darauf hingewiesen, dass meine Studie keinen Anspruch erhebt, in irgendeiner Weise mit internationalen, groß angelegten, etablierten Retrieval Studien mithalten zu können. Vieles in der Methodik kann, soll und wird hoffentlich auch diskutiert werden.

Trotzdem glaube ich, folgende Aussagen aus den Testergebnissen herauslesen zu können:

- Österreichische Rechtsdatenbanken haben großes Verbesserungspotential, nachdem Recall und Precision beide noch weit von den angestrebten 100% entfernt sind.
- Die Businessversion der RDB Rechtsdatenbank liefert das vollständigste Suchergebnis. Die ausufernde Trefferliste macht dieses Suchergebnis allerdings unübersichtlich.
- Die Universion der RDB Datenbank liefert deutlich schlechtere Ergebnisse als die Business-Version.
- Die Datenbank LexisNexis stellt für diesen Recherchefall lediglich den RIS Datenbestand zur Verfügung.
- Die RIDA Datenbank erzeugt die übersichtlichsten Trefferlisten, indem sie Mehrfachfundstellen in einem Eintrag zusammen fasst.

Literatur

Blair, D. C., und Maron, M. E., An evaluation of retrieval effectiveness for a full-text document retrieval system. *Communications of the ACM*, 28(3):289–99, 1985.

Kraft, Matthias, Juristische Online-Datenbanken in der Praxis, 2. Auflage 2007.

Liebwald, Doris, Evaluierung juristischer Datenbanken, Wien 2003.

Mader, Peter, Bemerkungen zur österreichischen Datenbanklandschaft, *jusIT* 2008, 29.

Mielke, Bettina, Bewertung juristischer Informationssysteme, Köln 2000.

Noack, Ulrich; Kremer, Sascha, Die großen Vier: Kostenpflichtige Online-Dienste für Juristen im Test, 2. Auflage 2007.

Prader, Christian, MANZ Wohnrecht online - ABGB 2.05 § 1096 (www.rdb.at).

Staudegger, Elisabeth, Rechtsdatenbanken in Österreich, *MR* 2006, 183.

² Über jegliches Feedback unter der Adresse home@antongeist.com würde ich mich sehr freuen. Gerne stelle ich bei Interesse meine genauen Studiendaten zur Verfügung.