

# Juristische Suchmaschine Lawsearch Enterprise – drei Herausforderungen

Franz Kummer<sup>1</sup> und Blaise Dévaud<sup>2</sup>

**Abstract:** Die juristische Suchmaschine Lawsearch Enterprise bewältigt den Informationsfluss einer juristischen Organisation (Anwaltskanzlei, Gericht, Rechtsdienst, Notariat etc.) mit einem Single Point of Search. Drei technische und organisatorische Herausforderungen stehen im Vordergrund: Auswahl und Integration der Informationsquellen; Beschaffung und Verwendung der Metainformationen; Erkennung und Verlinkung der juristischen Referenzen. Der Beitrag zeigt konkrete Lösungsansätze dieser drei Probleme und gibt dabei einen Überblick über die Funktionsweise der juristischen Suchmaschine Lawsearch Enterprise.

**Keywords:** Lawsearch; Suchmaschine; Recht, Wissensmanagement; Indexierung; Metadaten; Verlinkung; Parsing.

## 1 Einleitung

Im Jahre 2015 hat die Schweizerische Bundeskanzlei über 6'000 Seiten Rechtsvorschriften publiziert. Dies alleine entspricht 16 DIN A4 Seiten pro Tag. Dazu werden jeden Monat mehrere hundert relevante Gerichtsentscheide, juristische Beiträge in Fachzeitschriften und Monographien veröffentlicht. Ganz zu schweigen von Tausenden von Word, PDF und E-Mail Dateien, die jede juristische Organisation produziert und mehr oder weniger sorgfältig intern ablegt.

Die Bewältigung dieses Informationsflusses ist das Hauptproblem eines effizienten juristischen Wissensmanagements.

Systemisch gesehen besteht der Informationsfluss einer juristischen Organisation aus externen und internen Informationsquellen. Die externen Quellen sind insbesondere die Gesetzgebung, die Rechtsprechung und die juristische Literatur; die internen Quellen bilden die Büro-Dokumente, E-Mails zu den behandelten Rechtsfällen sowie eine allfällige eigene Bibliothek. Alle diesen Quellen sind stark heterogen: Gesetzgebung entsteht auf verschiedenen Staatsebenen; Rechtsprechung wird durch verschiedene Gerichte publiziert; und Literatur – wenn überhaupt digital verfügbar – ist meistens in bezahlten Systemen einzelner Verlage abgelegt (Paywall). Auch die internen Informationen lassen

---

<sup>1</sup> Jurist; Lehrbeauftragter am Kompetenzzentrum Informatik und Recht, Universität Bern; Gründer und Inhaber der Weblaw AG; Weblaw AG, Laupenstrasse 1, 3008 Bern, Schweiz; [www.weblaw.ch](http://www.weblaw.ch); [franz.kummer@weblaw.ch](mailto:franz.kummer@weblaw.ch).

<sup>2</sup> Jurist und Wirtschaftsinformatiker; Projektleiter bei der Weblaw AG; Weblaw AG, Laupenstrasse 1, 3008 Bern, Schweiz; [www.weblaw.ch](http://www.weblaw.ch); [blaise.devaud@weblaw.ch](mailto:blaise.devaud@weblaw.ch).

sich selten einheitlich verwalten: aus historischen oder operativen Gründen sind die Dokumente in separaten E-Mail-Archiven, Geschäftsverwaltungssystemen, Datei-Ablagen und Bibliothekslösungen verteilt.

In einer dermassen komplizierten Informationslandschaft spielt eine Suchmaschine eine zentrale Rolle.

Die juristische Suchmaschine Lawsearch Enterprise<sup>3</sup> der Firma Weblaw AG ist entstanden, um alle wichtigen Informationsquellen einer juristischen Organisation in einer einzigen Suchmaske verfügbar zu machen. Die Benutzer können sich mit diesem Single Point of Search über die (neusten) externen Veröffentlichungen informieren und gleichzeitig in den kompletten internen Archiven relevante Informationen zu den eingehenden Rechtsfällen finden.

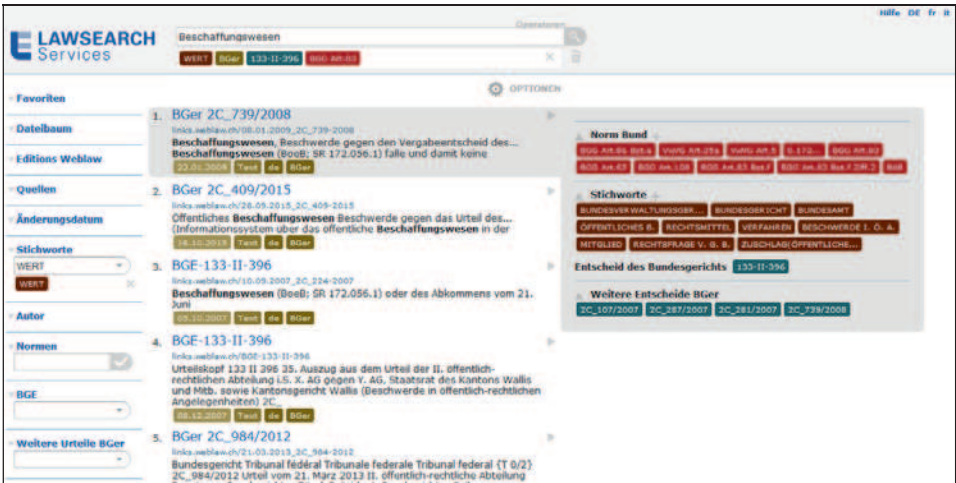


Abb. 1: Beispiel einer Suche mit der juristischen Suchmaschine Lawsearch. Textsuche („Beschaffungswesen“) kann mit einer mehrsprachigen Suche nach Stichwörtern und Referenzen (Bundesgerichtsentscheide BGE, Bundesnormen) kombiniert werden. In der Mitte werden Suchtreffer mit einem kurzen Textauszug angezeigt; rechts werden zu jedem Treffer relevante Metainformationen (Bundesnormen, Stichwörter, Rechtsprechungsreferenzen) gezeigt.

Die Gestaltung und Entwicklung der juristischen Suchmaschine Lawsearch Enterprise ist aber technisch und organisatorisch anspruchsvoll. Drei spezifische Herausforderungen stehen im Vordergrund: Auswahl und Integration der Informationsquellen; Beschaffung und Verwendung der Metainformationen; Erkennung und Verlinkung der juristischen Referenzen.

<sup>3</sup> Die Leser können die Webversion der juristischen Suchmaschine Lawsearch Enterprise bei den folgenden öffentlichen Instanzen testen: <http://www.weblaw.ch/competence/search.html> (Suchseite des Internet-Auftritts der Weblaw AG); <http://lawsearch.gr.ch/le/> (publizierte Gerichtsentscheide des Kantons Graubünden, Schweiz); <https://apps.vs.ch/le/> (publizierte Gerichtsentscheide des Kantons Wallis, Schweiz). Weitere Informationen: <http://www.weblaw.ch/competence/technology/lawsearch-enterprise.html>.

## 2 Auswahl und Integration der Informationsquellen

Das erste Problem ergibt sich aus der Verteilung in interne und externe Informationsquellen. Die internen Quellen einer juristischen Organisation sind extrem sensibel und dürfen das interne IT-System (bspw. einer Anwaltskanzlei) nicht verlassen. Dies hat zur Folge, dass die Suchmaschine beim Kunden innerhalb von seinem IT-System installiert und eingesetzt werden muss. Die externen Quellen – mehrheitlich frei im Internet verfügbar – sind sehr umfangreich (Hunderttausende von Gesetzen, Gerichtsentscheiden und Textdokumenten). Die Integration dieser Daten in den internen Suchindex jeder juristischen Organisation würde zu unnötigem und redundantem Ressourcenverbrauch führen (Speicherplatz, Bandbreite, Arbeitslast bei der Indexierung). Aus diesem Grund werden alle externen juristischen Quellen zentral bei der Weblaw AG indexiert und in einem gemeinsamen Suchindex gespeichert. Suchanfragen zu den externen Quellen werden aus jeder organisationsinternen Suchmaschine an den zentralen Suchindex weitergeleitet; die Suche wird im zentralen Index ausgeführt und die organisationsinterne Suchmaschine erhält die Suchergebnisse zurück. Dieses Vorgehen erlaubt eine effiziente Verwaltung der externen Quellen und respektiert die hohen Sicherheitsanforderungen für die internen Quellen.

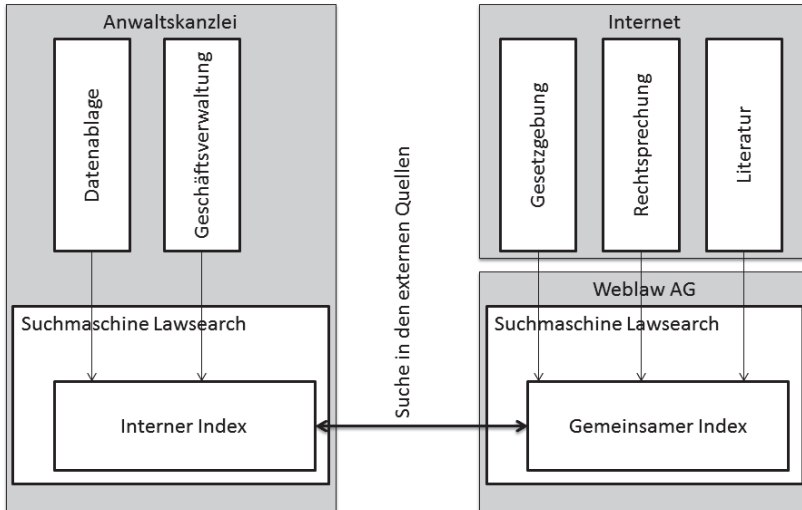


Abb. 2: Die internen Quellen sind intern beim Kunden indexiert; die externen sind zentral bei der Weblaw AG in einem gemeinsamen Index gespeichert. Anfragen zu den externen Quellen werden an den gemeinsamen Index weitergeleitet; die interne Suchmaschine erhält die Resultate zurück.

Sowohl die internen als auch die externen Quellen zeigen weitere spezifische Detailprobleme.

Die internen Informationen werden jeweils aus internen Drittsystemen des Kunden be-

schaft: Filesysteme; Intranet; Sharepoint; Geschäftsverwaltungs- und Zeiterfassungssysteme; E-Mail-Archive etc. Zu jedem dieser Drittsysteme muss eine Schnittstelle bestehen oder erstellt werden, öfters unter Berücksichtigung der Benutzerverwaltung und des Sicherheitskonzepts. Diese Tatsache erhöht die Wartungskomplexität der Suchmaschine. Andererseits können Skalenerträge erreicht werden, wenn mehrere Kunden die gleiche interne Technologie einsetzen.

Bei den externen Informationsquellen aus dem Internet ist im ersten Schritt eine intellektuelle Auswahl gefragt<sup>4</sup>: es sollen vorzugsweise Informationsquellen bestimmt werden, die für alle Juristen von Interesse sind. Quellen aus spezifischen Rechtsgebieten sind eingeschränkt, können aber auf Anfrage aufgenommen werden. Des Weiteren sind die externen Quellen dem üblichen Problem der Internet-Quellen ausgesetzt: die ausgewählten Webseiten/Datenbanken unterlaufen öfters formellen oder inhaltlichen Änderungen, was eine ständige Überwachung und Anpassung der spezifischen Crawler impliziert. Diese Situation ist in der föderalen Schweiz insbesondere akut: jeder Kanton entscheidet unabhängig von den anderen über die Publikation seiner Gesetzgebung und Rechtsprechung. Obwohl gemeinsame technische Standards bereits erarbeitet wurden<sup>5</sup>, finden sie in der Praxis noch kaum Anwendung. Der immense Kundennutzen einer aggregierten Suche in allen heterogenen Rechtsquellen rechtfertigt entsprechend diese hohen Integrationsaufwände.

Eine besondere Art der externen Informationsquellen bilden Inhalte, die die Weblaw AG als juristischer Verlag selber produziert. Sie bestehen aus den Publikationen der Editions Weblaw<sup>6</sup> (juristische digitale Zeitschriften Jusletter (jusletter.ch), Jusletter IT (jusletter-it.eu), Richterzeitung (richterzeitung.ch) und Archiv für Schweizerisches Abgaberecht (asa.weblaw.ch); juristische Datenbank Banking and Finance (bf.weblaw.ch); digitaler Rechtsprechungskommentar (drsk.weblaw.ch)) sowie aus den bibliographischen Angaben zur juristischen Literatur (biblio.weblaw.ch). Diese letztgenannten Angaben entstammen der juristischen bibliographischen Datenbank<sup>7</sup> der Weblaw AG.

Alle Quellen sollen effizient und benutzerfreundlich durchgesucht und dargestellt werden. Diese Anforderung wird bei Lawsearch Enterprise durch den Einsatz von Suchreibern erfüllt. Unter dem Volltextsuchfeld werden verschiedene Reiter dargestellt: durch das Klicken auf die einzelnen Tabs werden die Resultate aus den entsprechenden Quellen angezeigt. Somit kann mit einer einzigen Eingabe der Suchbegriffe (und Filter) in allen verfügbaren Quellen das Resultat konsultiert werden. Weiter werden für eine Suche alle internen und externen Quellen detailliert in einem Suchbaum aufgelistet: so lässt sich die Suche auf einzelne Quellen einschränken.

---

<sup>4</sup> Eine Liste der aktuellen externen Quellen befindet sich unter [http://www.weblaw.ch/competence/technology/lawsearch-enterprise/juristische\\_quellen.html](http://www.weblaw.ch/competence/technology/lawsearch-enterprise/juristische_quellen.html).

<sup>5</sup> CHLexML heisst der Standard zur Darstellung von Gesetzen im XML Format; <http://www.svri.ch/de/CHLexML.html>. CHDecML für Rechtsprechung und CHDocML für Literatur befinden sich im Entwurfstadium).

<sup>6</sup> Siehe dazu die Verlagsinformationen unter [http://www.weblaw.ch/competence/editions\\_weblaw.html](http://www.weblaw.ch/competence/editions_weblaw.html).

<sup>7</sup> [http://www.weblaw.ch/competence/editions\\_weblaw/bibliotheksloesung/rechtsbibliografie.html](http://www.weblaw.ch/competence/editions_weblaw/bibliotheksloesung/rechtsbibliografie.html).



Abb. 3: Die Reiter unter den Suchfeldern erlauben eine Navigation in allen verfügbaren Quellen mit einer einzigen Suchanfrage. Der erste Reiter zeigt die Treffer aus den internen Quellen (Lokal), der zweite die Treffer aus den externen Quellen (lawsearch.ch), und die drei letzten Reiter zeigen die passenden bibliographischen Angaben (jeweils aus der Bibliothek vom eigenen Organisations-Standort; aus allen Bibliotheken der Organisation; aus der gesamten Rechtsbibliografie).

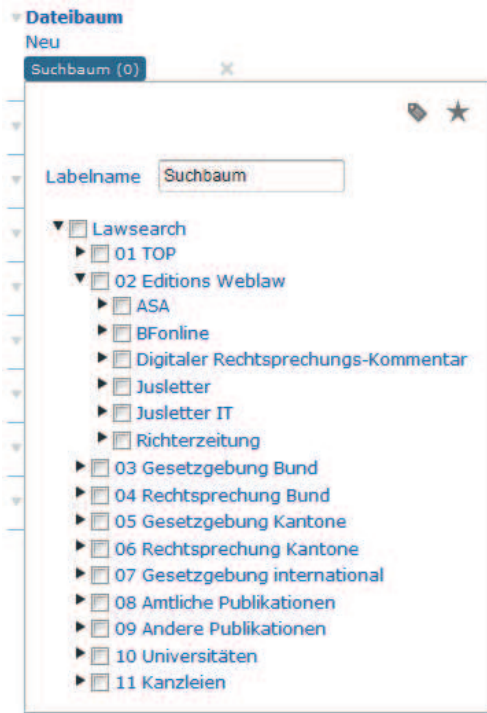


Abb. 4: Auswahl der externen Quellen im Suchbaum.

### 3 Beschaffung und Verwendung der Metainformationen

Reine Textsuchen in den verfügbaren Dokumenten sind für Juristen nicht ausreichend. Vielmehr benötigen sie bei der täglichen Arbeit die Möglichkeit, die Treffer nach gewissen Metainformationen zu filtern. Diese Metainformationen beziehen sich auf die juristi-

schen (Dokumentenart, Rechtsgebiet, Aktennummer, zuständiges Organ), technischen (Format, Erstellungsdatum, Änderungsdatum) und organisationsbezogenen (Verbindungen zu Projekten und Personen) Eigenschaften des Dokuments.

Bei den internen Dokumenten werden die Metadaten meist direkt über die Schnittstelle zu den Drittsystemen beschafft. Dabei besteht die technische Herausforderung darin, die Metadaten zu den Dokumenten aktuell zu halten. Dies ist insb. dann problematisch, wenn sich Metadaten ändern, ohne dass das Dokument an sich geändert wird. Dieses Problem wird durch einen Push-Mechanismus auf der Seite des Drittsystems gelöst: das Drittsystem benachrichtigt die Suchmaschine periodisch über die letzten Änderungen der Metadaten.

Bei den externen Dokumenten ist eine strukturierte Abfrage der Metadaten mittels eines Web-Services nur selten möglich. In den meisten Fällen müssen die publizierten Dokumente inhaltlich analysiert werden (Parsing), um die relevanten Informationen zu extrahieren. Aufgrund der verschiedenen Formate und Formatierungen der halbstrukturierten Text-Dokumente ist leider keine generische Lösung möglich: für jede Quelle muss ein dediziertes Parsing programmiert werden, was den Arbeitsaufwand bei den externen Quellen weiter erhöht.

Die hohe Anzahl der verfügbaren Quellen resultiert in einer entsprechend hohen Anzahl von verfügbaren Metadaten-Dimensionen. Es ist eine Design- und Usability-Herausforderung, alle Metadaten-Typen effizient und benutzerfreundlich in der Suchmaske verfügbar zu machen. Dieses Problem löst Lawsearch Enterprise mit dem Einsatz von den sog. Labels: es sind Suchfilter, die grafisch dargestellt und via drag-und-drop (oder Doppelklick) ins zweite Suchfeld gezogen werden (siehe dazu Abb. 4 „Ziehen Sie Labels hierhin um danach zu filtern“. Dieses zweite Suchfeld wird auch „Label-drop-space“ genannt). Mit diesem Lösungsansatz können beliebig viele Labels und Label-Typen miteinander kombiniert werden. Im Resultat liegt eine multidimensionale Facetten-Suche vor.



Abb. 5: Mehrere Labels können in der Suche gleichzeitig eingesetzt werden: dies entspricht einer multidimensionalen Facetten-Suche. Die Suche kombiniert den Begriff Beschaffungswesen mit dem mehrsprachigen Deskriptor Wert, der Quelle Bundesgericht, der Bundesgerichtsreferenz 133 II 396 sowie dem mehrsprachigen Erlasszitat Art. 83 (inkl. aller Substufen) BGG (Bundesgerichtsgesetz).

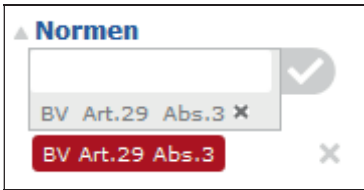


Abb. 6: Labels können beliebig von Hand erstellt werden. Hier wird ein Label erstellt, um alle Dokumente zu finden die den Artikel 29 Absatz 3 der Bundesverfassung zitieren (in drei Sprachversionen). Vorgeschlagen werden nur Referenzen, die sich effektiv im indextierten Dokumentenstamm befinden.



Abb. 7: Metainformationen zu jedem Treffer-Dokument werden als Labels dargestellt (rechts) und können direkt zur Verfeinerung der Suche benutzt werden.

## 4 Erkennung und Verlinkung der juristischen Referenzen

Lawsearch Enterprise bereichert jedes Dokument mit weiteren juristischen Metadaten (juristische Referenzerkennung, Veredelung). Die Suchmaschine erkennt automatisch die wichtigsten juristischen Referenzen (und Deskriptoren), die im Text des Dokuments erwähnt werden: zitierte Gesetzesartikel, „Fall-Nummern“ von Gerichtsentscheiden und zitierte Literatur aus juristischen Zeitschriften sowie weitere jur. Referenzen.

Die Erkennung erfolgt während der Indexierung der Dokumente und basiert auf dem Einsatz von regulären Ausdrücken. Voraussetzung der korrekten Erkennung ist eine einheitliche Zitierweise. Dies erweist sich in der Praxis öfters als problematisch: nicht alle Referenzen, insb. auf Kantonsebene, sind standardisiert. Des Weiteren kann der Sinn einer Referenz kontextabhängig sein (z.B. wird die gleiche Abkürzung zur Bezeichnung von verschiedenen kantonalen Gesetzen oder von verschiedenen historischen Versionen des gleichen Gesetzes verwendet). Die regulären Ausdrücke verarbeiten diese Abweichungen soweit möglich.

Die erkannten Referenzen werden zu drei Zwecken genutzt.

Sie können als Suchfilter verwendet werden. Dies erlaubt eine fortgeschrittene mehr-

sprachige Suche nach Dokumenten, die einen konkreten Gesetzesartikel oder einen bestimmten Gerichtsentscheid zitieren – diese Suchmöglichkeit wird in der Praxis besonders geschätzt.

Zweitens werden die erkannten Referenzen in einer Arbeitsversion des Dokumentes (sog. Cache-Version) direkt im Text hervorgehoben und verlinkt. Aufgrund der häufigen Änderungen der Internet-Adressen stellt die Verlinkung eine besondere Herausforderung dar. Um eine nachhaltige Verlinkung zu gewährleisten wurde ein Webservice<sup>8</sup> als Zwischenstufe eingebaut. Jeder automatisch erstellte Link zeigt auf die URL des Webservices mit der Referenz als Zusatzparameter<sup>9</sup>; der Webservice prüft die aktuelle URL für diese Referenz und leitet die Browser-Anfrage entsprechend weiter auf die korrekte Website. Der Webservice erlaubt eine zentrale Verwaltung der juristischen Links und eine zeitnahe, für die Benutzer kaum merkbare Anpassung bei Änderungen.

Drittens werden alle erkannten Referenzen in der Cache-Version gesammelt aufgelistet. Dies bietet eine automatische Zusammenfassung des Dokumentes und ist besonders bei längeren Texten hilfreich.

<p>Gegenstand  Art. 9<sup>GG</sup> BV ( <b>Beschaffungswesen</b> ),</p> <p>Staatsrechtliche Beschwerde gegen das Urteil des Kantonsgerichts Appenzell Innerrhodon, Abteilung Verwaltungsgericht, vom 6. September 2005.</p> <p>Sachverhalt:</p> <p>A.  Der Kanton Appenzell Innerrhodon holte im Einladungsverfahren bei zwölf Firmen Offerten für ein elektronisches Archivsystem ein. In die Schlussevaluation kamen vier Anbieter, darunter die Y. _____ AG, St. Gallen, und die X. _____ GmbH, Zürich.</p> <p>Im Auftrag von Landmann und Standeskommission erteilte der Ratschreiber den Zuschlag für das elektronische Archivsystem am 25. Februar 2005 zum Preis von Fr. 243'396.-- an die Y. _____ AG. Die X. _____ GmbH, deren Offerte unberücksichtigt blieb, erhob hiegegen am 6. März 2005 Beschwerde beim Kantonsgericht (Abteilung Verwaltungsgericht). Das Gericht befand, die Beschwerde genüge den gesetzlichen Anforderungen nicht und räumte der X. _____ GmbH Frist ein, um die Rechtschrift zu verbessern. Dies tat die Beschwerdeführerin mit Eingabe vom 22. März 2005, worauf der Kantonsgerichtspräsident der Beschwerde mit Verfügung vom 24. März 2005 zunächst aufschiebende Wirkung gewährte. Diese Verfügung hob er am 12. April 2005 wieder auf und entzog der Beschwerde die aufschiebende Wirkung, wogegen sich die X. _____ GmbH bis vor Bundesgericht zur Wehr setzte. Die am 25. April 2005 gegen diesen Zwischenentscheid des Kantonsgerichtspräsidenten erhobene staatsrechtliche Beschwerde blieb jedoch ohne Erfolg; sie wurde vom Geschäftsverzeichnis abgeschrieben, nachdem sich herausgestellt hatte, dass der Vertrag mit der berücksichtigten Konkurrentin, der Y. _____ AG, bereits am 15./22. April 2005 abgeschlossen worden war (vgl. Beschluss <sup>GG</sup><a href="#">2P.116/2005</a> vom 24. Mai 2005, insbesondere E. 1).</p>	<p><a href="#">BGG: Art.84</a> <a href="#">Art.86</a> <a href="#">Art.87</a>  <a href="#">Art.88</a> <a href="#">Art.156</a> <a href="#">Art.159</a></p> <p><a href="#">BGBM: Art.9</a></p> <p><b>Weitere Urteile BGER</b></p> <p><a href="#">2P.116/2005</a>  <a href="#">2P.218/2001</a>  <a href="#">2P.294/2005</a></p> <p><b>Leitentscheide BGE</b></p> <p><a href="#">125-II-86</a>  <a href="#">132-I-86</a>  <a href="#">125-II-518</a>  <a href="#">129-I-173</a></p>
----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

Abb. 9: In der Cache-Version (Arbeitsversion) eines Dokumentes werden die juristischen Referenzen erkannt, im Text hervorgehoben, verlinkt und auf der rechten Seite zusammenfassend aufgelistet. Die gesuchten Begriffe werden ebenfalls farblich gekennzeichnet.

<sup>8</sup> <http://links.weblaw.ch>.

<sup>9</sup> Z.B. „<http://links.weblaw.ch/de/SR-210 Art.3>“.



## **5 Fazit**

Die juristische Suchmaschine Lawsearch Enterprise wird täglich von hunderten Schweizer Juristen erfolgreich eingesetzt. Zu den Kunden zählen Anwaltskanzleien, Gerichte, Versicherungen und Rechtsdienste vieler Unternehmen. Sie ist ein relevantes Arbeitsinstrument bei der täglichen juristischen Arbeit.

Die gelungene Lösung der drei Herausforderungen hat zu diesem Erfolg wesentlich beigetragen. Die juristische Suchmaschine Lawsearch Enterprise ist eine Schweizer Success Story.