

Wege zum elektronischen CELEX-Kommentar

*Erich Schweighofer**

*Arbeitsgruppe Rechtsinformatik
Institut für Völkerrecht und Internationale Beziehungen
A-1090 Wien, Universitätsstraße 2
Erich.Schweighofer@univie.ac.at*

Schlagworte: EU-Datenbanken, CELEX, elektronischer Kommentar

Abstract: Die letzte Dekade hat Rechtsinformationssysteme sehr nahe dem Ziel einer vollständigen Rechtsdokumentation gebracht. Nunmehr müssen aber die vielfältigen Forschungen einer besseren Wissensrepräsentation in die Praxis umgesetzt werden. Der Beitrag gibt einen Überblick über die Möglichkeiten einer Verwirklichung am Beispiel der EU-Rechtsdatenbank CELEX.

1. Einleitung

Die Anforderungen an Rechtsinformationssysteme haben sich gewaltig verändert. Hat zu Beginn noch die Dokumentation als Leitwissenschaft gedient, so ist dies heute Wissensmanagement. Das Ziel des Wissensmanagements liegt in einer lebendigen und aktiven Behandlung des Wissens einer Organisation.¹ Die riesigen Datenwarenhäuser sollen von einer passiven Sammlung von Daten zu aktiven Komponenten in der Wissensorganisation werden.

Dies bedeutet einen grundlegenden Wandel in der Rolle von Rechtsinformationssystemen. Die bereits erreichte Rechtsdokumentation wird durch eine zunehmend automatisierte Strukturanalyse des Rechts² zu ergänzen

* Hauptverwaltungsrat in der Europäischen Kommission, ao. Univ.-Professor an der Universität Wien (karenziert). Dieser Beitrag enthält die persönliche Auffassung des Autors, die nicht notwendigerweise jener der Europäischen Kommission entspricht.

¹ So sehr einprägsam *Wimmer, M. and Traummüller, R.*, Trends in Electronic Government: Managing Distributed Knowledge, in: *Tjoa, A. M., Wagner, R. R. and Al-Zobaidie, A.* (Eds), DEXA'2000 Eleventh International Workshop on Database and Expert Systems Applications, September 2000, Greenwich, IEEE, Los Alamitos, CA, 340 ff. „Treating an organisation's information and knowledge as „living“ and active, rather than simple as a passive collection of records in huge data warehouses and repositories, is a main concern of managing distributed knowledge.“

² Vgl. *Schweighofer E.*, Rechtsinformatik und Wissensrepräsentation, Wien 1999.

sein. Diese Kombination, hier dynamischer elektronischer Kommentar genannt, dürfte den nächsten Qualitätssprung bei Rechtsinformationssystemen darstellen.

2. Der Wandel bei CELEX

Vor 30 Jahren wurde CELEX als Instrument für Bibliothekare und Dokumentalisten entwickelt. Ohne Informationsvermittler (oder Rechercheur) gab es keinen Zugang zum Rechtsinformationssystem. Die Betonung der Dokumentation in Europa zeigte sich auch in der Struktur der Datenbank, die bei CELEX als sehr komplex bezeichnet werden kann, und der Menge der zusätzlichen Aufbereitungen zum Volltext. Dieses Merkmal ist bei allen Entwicklungen in Europa zu beobachten und steht im Gegensatz zur starken Betonung tagfertiger Textsammlungen in den USA. Die Einbeziehung der Dynamik der täglichen Änderung der Rechtsordnung war (in Europa) noch weniger ausgeprägt.

Die Juristen waren bereits hoch zufrieden, wenn unscharfe Ergebnisse geliefert wurden, waren diese doch das einzige brauchbare Mittel, um der umständlichen und ungenauen Leserecherche zu entgehen, wenn eigenes Gedächtnis und Kommentare „versagt“ hatten. Ebenfalls existierte eine hohe Bereitschaft der Juristen, Suchergebnisse zu interpretieren und juristisch zu strukturieren.

Heute gibt es natürlich weiterhin Informationsvermittler und Dokumentalisten, aber in wesentlich spezialisierterer Funktion. Reine Recherchen wie Auffinden von Texten nach bibliographischen Angaben und Fundstellen kann das Rechtsinformationssystem wesentlich besser und billiger übernehmen.

Die Recherche selbst ist schwieriger geworden. Norm- und Entscheidungssammlungen werden immer umfangreicher, aber trotzdem müssen relevante Normen und Entscheidungen genau berücksichtigt werden. Der Jurist muss mehr recherchieren, hat aber weniger Zeit für die Interpretation der Suchergebnisse. Dies ist einerseits durch den Zeitmangel, andererseits die Fülle der Ergebnisse von Recherchen in Rechtsinformationssystemen bedingt. Eine brauchbare Unterstützung der Arbeit durch Rechtsinformationssysteme wird erwartet, kann aber mit den derzeitigen Methoden nicht gewährleistet werden.

Der Dynamik der Rechtsordnung muss wesentlich besser Rechnung getragen werden. Die Rechtsinformationssysteme müssen alle Texte in Echtzeit, dh sofort und gut suchbar bereitstellen. Die boolesche Suchlogik hat ihre Vorteile, sollte aber durch bessere Methoden wie „Google“ zumindest ergänzt werden.

Die wesentlichen Änderungen wären die Strukturanalyse in Echtzeit, dh Textverlinkung, Textklassifikation, Textbeschreibung und semiautomatische Textanalyse tagfertig bzw nach maximal zwei Wochen.

Nur am Rande sei erwähnt, dass die Integration des Rechtsinformationssystems in bestehende Prozessmodelle des e-Government sehr wichtig ist. Mit E-Recht (Rechtsetzung und elektronisches Gesetzblatt in Österreich) wird die Produktion wesentlich verbessert. Bereits weitgehend verwirklicht ist die Unterstützung der Dokumenterstellung mit Anwaltssoftware (Übernahme, Recycling etc von Texten).

3. Vom Textarchiv zum Rechtskommentar

Ohne Zweifel existiert eine riesige Lücke zwischen der Qualität des Textarchivs und jener des Rechtskommentars. Rechtsinformationssysteme liefern Texte nach Auswahl des Benutzers, aber keine Strukturierung eines Wissensgebietes wie ein Kommentar. Unzweifelhaft bieten Rechtsinformationssysteme wie EUR-Lex oder CELEX neben dem Dokumentenarchiv auch Suchmöglichkeiten, die Ergebnisse vergleichbarer Qualität eines Kommentars bringen. Die Schwierigkeit liegt nur darin, diese Suchstrategien auch so zu beherrschen, dass sowohl Qualität als auch Verlässlichkeit ausreichend gegeben sind.

EUR-Lex und CELEX bieten einen bibliographischen Zugang zu allen EU-Dokumenten nach Dokumentennummer, CELEX-Nummer oder Fundstelle im Amtsblatt. Die boolesche Suchlogik unterstützt bei Volltextsuche (bei CELEX auch im Titel). Sehr hilfreich ist die Suche mit dem Fundstellennachweis des geltenden Gemeinschaftsrechts. EUR-Lex bietet nur diese eingeschränkten Suchmöglichkeiten. CELEX bietet weiters die Suche mit Verweisen³ und die Strukturierung nach Dokumenttypen. Die vielen weiteren CELEX Felder⁴ können nur in der Expertenversion und mit ausreichender Kenntnis dieses Potentials genutzt werden. Diese Form der Wissensrepräsentation durch umfassende Feldstrukturierung hat sich wegen der hohen Schulungskosten in der Praxis nicht durchgesetzt. Unter Einsatz von XML bieten sich jedoch wichtige Chancen, die einer Nutzung harrten.

³ Diese ist jedoch etwas umständlich angelegt. Zuerst muss das Dokument aufgerufen werden, sodann werden Hypertextlinks zu zitierten Dokumenten angeboten.

⁴ Vgl im Detail *Schweighofer, E.*, Wissensrepräsentation in Information Retrieval-Systemen am Beispiel des EU-Rechts, Wien 1995 (Drucklegung: 2000).

3.1. Vom suchbaren Dokumentenarchiv zum Werkzeug des Wissensmanagements

Der Wandel vom *Stand der Technik*, dh einem suchbaren Dokumentenarchiv zum Werkzeug des Wissensmanagements, sollte in zwei Schritten erfolgen.

Als erster Schritt sollen einfache Methoden umgesetzt werden. EUR-Lex wie CELEX sollen die Dokumente tagesaktuell bereitstellen; sie sollen Echtzeitinformationssysteme werden. Bei den vielen Amtssprachen ist dies aber immer ein Ressourcenproblem.

Der Wandel zu den benutzerfreundlichen Schnittstellen war mit einem Verlust an Suchmöglichkeiten verbunden. Die sinnvolle CELEX Struktur bietet jedoch bessere Suchmöglichkeiten. Als Beispiele seien die Suche in den Verweisfeldern, in der Verschlagwortung oder in der Zusammenfassung des Gerichtshofs angeführt. Bei Einsatz von XML könnten sich hier sehr sinnvolle Suchoptionen realisieren und die Suche in möglichst allen Textfeldern ermöglichen lassen. Volltextsuche sollte nicht bloß eine Wortsuche sein, sondern eine sinnvolle Kombination von Wörtern, die mit boolescher Logik und Abstandoperatoren Sprachstrukturen beschreiben. Gerade diese für juristische Anwendungen wichtige Suche wird derzeit nur in der Expertenversion angeboten. Ein hohes Potential ist in der Präsentation der Ergebnisse durch Orientierung an den Bedürfnissen der Benutzer gegeben. Die derzeit üblichen Titellisten der gefundenen Dokumente in chronologischer Reihenfolge können nicht der letzte Stand der Technik sein. Als gutes Beispiel einer Verbesserung kann *Prelex* genannt werden, bei welcher der Gesetzgebungsprozess in schön strukturierter Form präsentiert wird. CELEX bietet hingegen nur eine Ergebnisliste, die erst mühsam analysiert werden muss. Neben *Prelex* sollte überlegt werden, auch für die Anwendung und Umsetzung des Gemeinschaftsrechts ein *Prolex* zu schaffen, in welchem übersichtlich sowohl die – soweit nötig – innerstaatliche Umsetzung als auch nachfolgende Rechtsdokumente wie relevante EuGH-Judikatur oder nationale Rechtsprechung aufbereitet wird. Dazu gibt es bereits entsprechende Projekte.

Die Verknüpfung der Entscheidungen des EuGH durch Judikaturketten ist derzeit nur durch Verweise realisiert. Hier könnte durch intellektuelle Bearbeitung unter Unterstützung der Datenbank CELEX ein wesentlicher Mehrwert geschaffen werden. Die Ähnlichkeiten bei Judikaturdokumenten könnten mit neuronalen Netzen und intellektueller Nachbearbeitung in brauchbarer Form analysiert werden.

Der zweite Schritt ist wesentlich schwieriger, weil als Ziel der elektronische Kommentar für Juristen ins Auge gefasst wird.⁵ Eine Strukturana-

⁵ Vgl. *Schweighofer*, FN 2.

lyse des Rechts und der rechtlichen Materialien für juristische Aufgabenstellungen ist notwendig. Die Festlegung als auch die Abdeckung der juristischen Materialien Gesetz, Urteil und Literatur hat in wesentlich genauere Form zu erfolgen. Diese Materialien sind sachlich zu klassifizieren und zu verlinken. Die Zukunftskomponente, das Recht *de lege ferenda*, ist ebenfalls mit zu berücksichtigen.

Bis zu diesem Schritt ist allenfalls die Vollständigkeit sowie die Klassifikation und Verlinkung qualitativ höher als im Rechtsinformationssystem. Die wesentlichen Verbesserungen sind die semiautomatische Textanalyse sowie die Einbeziehung von Ontologien der juristischen Rechtsordnung.

Der Vollständigkeit halber sollte aber auch angemerkt werden, dass neben dem elektronischen Kommentar auch Amtshelfer für Bürger und Unternehmen eine ständig steigende Rolle spielen werden (zB <http://www.help.gv.at>). Amtshelfer zeichnen sich durch eine prozessartige Beschreibung des Rechts in bestimmten Lebenssituationen aus. Weiters werden eine Hilfe zur Erstellung relevanter Dokumente sowie Tipps und Hinweise zur Erledigung der Behördenwege gegeben. Rechtsinformationssysteme sowie elektronische Kommentare bieten hier nur eine Unterstützung für ein Anliegen des Bürgerservice.

4. Mögliche Schritte einer Verbesserung von CELEX

Trotz sehr hohem Vollständigkeitsgrad der Dokumentation der Rechtsakte in CELEX scheint in einigen Bereichen noch eine Verbesserung möglich. Die Verlinkung der Dokumente mit den nationalen Rechtsordnungen brächte eine wesentliche Erleichterung, wobei hier der Abschluss der einschlägigen Forschungsprojekte abzuwarten ist. Die Einbeziehung der Literatur wird – abgesehen von bibliographischen Hinweisen – aus urheberrechtlichen Gründen immer ein Wunschtraum bleiben.

Die Aufbereitung dieses umfangreichen multilingualen Textkorpus kann durch semiautomatische Textanalysen wesentlich verbessert werden. In vielen Fällen wäre eine Ablöse der üblichen intellektuellen Indexierung denkbar.

4.1. Verweise

Die Verweisstruktur in CELEX war immer hoch entwickelt, wenn auch wenig einsichtig für ungeübte Nutzer. Der Darstellung der Änderung von Rechtsakten, der Rechtsgrundlage von Rechtsakten sowie der Beziehung Rechtsakt und relevanter Judikatur wird großer Raum gegeben.

In der semiautomatischen Textanalyse wird mit Heuristiken gearbeitet. Verweise sind in einer sehr formalen Sprache gehalten und sind mit einem intelligenten Matching relativ einfach und mit ausreichender Genauigkeit zu finden. Damit ist die Frage des *ob* einer Automatisierung der Verweissanalyse⁶ gelöst, nicht jedoch das *wie*. Die Verweissystematiken von Rechtsinformationen wie RIS, CELEX oder *Shepard's* weisen auf mögliche Lösungen hin. Wünschenswert wäre nicht nur die Darstellung konsolidierter Rechtsakte, dh Stammakt eines Rechtsakts mit Änderungen, der Beziehung zwischen Rechtsnorm und Anwendung durch Verwaltung und Rechtsprechung, sondern auch der Innenbeziehungen der Rechtsprechung als solcher. Zur Judikatur sollte nicht nur festgelegt sein, zu welcher Norm diese ergeht, sondern auch, ob diese neu, klärend, bestätigend oder abändernd ist. Dadurch würde das Problem der vielfältigen Wiederholungen in Judikaturdatenbanken besser gelöst.

Alle bisherigen Verweissystematiken leiden an der hohen Quantität der Verweise. Ohne die qualitative Selektion ist die Information vielfach nicht sinnvoll nutzbar. Diese Lücke könnte durch die Kombination von semiautomatischer Textanalyse sowie nachfolgender intellektueller Bearbeitung gelöst werden. Zumindest die Verweissystematik würde die Qualität existierender Kommentare erreichen, wenn auch die Darstellung noch einiges zu wünschen übrig lassen wird.

4.2. Ähnlichkeitsklassen zwischen Dokumenten

Die Fülle von Dokumenten zwingt zur Nutzung automatischer Verfahren zur Klassifikation und Klassenbeschreibung. Dieser Dienst wird von vielen Rechtsinformationssystemen als intellektuelle Aufbereitung angeboten; der Fundstellennachweis des geltenden Gemeinschaftsrechts ist ein gutes Beispiel hierfür.

Die Rechtsinformatik bietet in der Kombination von Vektorrepräsentation von Dokumenten und Ähnlichkeitsanalyse mit neuronalen Netzen eine potentiell sehr effiziente Methode. Hierbei werden die Dokumente als Vektoren repräsentiert und mit Hilfe von Kohonennetzen Ähnlichkeiten von Dokumenten berechnet.⁷ Seit einiger Zeit ist sowohl die Beschreibung der automatisch gebildeten Klassen als auch deren hierarchische Strukturierung möglich. Dieses aufwendige Verfahren kennt leider noch einige technische Restriktionen. Die Anzahl von Klassen ist mit bis zu 1000 bereits von ausreichendem Umfang. Leider ist die Ähnlichkeitsberechnung aller Doku-

⁶ Hier wird insbesondere auf die Arbeiten des AustLII hingewiesen.

⁷ Schweighofer E., Rauber, A. und Dittenbach, M., Automatic Text Representation, Classification and Labeling in European Law, in: Proc Int Conf on Artificial Intelligence & Law, ACM Press, New York, NY, 2001, 78-87.

mente eines Rechtsinformationssystem ohne Bildung von Subgruppen noch außer Reichweite, sodass sich ein Einsatz in kleineren Einheiten wie umfangreichen Suchergebnissen empfiehlt. Die bisherigen Forschungsergebnisse sind sehr vielversprechend, weisen jedoch auf das Problem der Optimierung des Vektors hin. Es bleibt sehr schwierig, mit statistischen Methoden (zB TF/IDF) gute Werte für die Qualität eines juristischen Begriffs zu bekommen.⁸ Derzeit werden modifizierte Vektoren getestet, wobei ein Schwerpunkt im Einsatz juristischer Ontologien besteht.

Für den elektronischen CELEX Kommentar könnte eine semiautomatische Erstellung des Fundstellennachweises oder die semiautomatische Erstellung von Judikaturklassen als erster Schritt in Angriff genommen werden.

Sehr hilfreich für eine Verbesserung der Klassifikation könnte die Einbeziehung juristischer Ontologien sein. Derzeit sind solche nur in sehr kleinen Teilbereichen verfügbar. Eine Nutzung für das Recht ist erst dann denkbar, wenn die Ontologien die notwendige Qualität erreicht haben, was noch etwas Zeit in Anspruch nehmen wird. Derzeitige Projekte zeigen das hohe Potential zum Aufbau eines elektronischen Kommentars.⁹ Langfristig könnte ins Auge gefasst werden, den Thesaurus EUROVOC zur Ontologie weiterzuentwickeln.

Die bisherigen Forschungen zur Textklassifikation¹⁰ haben gezeigt, dass brauchbare Unterstützung für Klassifikation und Contentanalyse geboten werden kann. In Verbindung mit intellektueller Nachbearbeitung kann auch die notwendige Qualität erreicht werden. Es bleibt aber weiterer Projektforschung vorbehalten, die Möglichkeiten des praktischen Einsatzes im Detail auszuloten.

5. Schlussfolgerungen

Beim derzeitigen Stand der Technik ist stets davon auszugehen, dass nur durch Zusammenarbeit zwischen Computer und Mensch ein sinnvolles Ergebnis erzielt werden kann. Juristische Texte und deren Interpretation sind einfach zu komplex, um diese vollständig in Algorithmen zu fassen. Aber Softwareprogramme können nach bestimmten Vorgaben Giga-

⁸ Schweighofer, E., Merkl, D. und Rauber, A., Some Remarks on Vector Representations of Legal Documents, in: A. M. Tjoa, R. R. Wagner und A. Al-Zobaidie, (Eds), DEXA 2000 Eleventh International Workshop on Database and Expert Systems Applications, September 2000, Greenwich, IEEE, Los Alamitos, CA, 2000, 1087-1091.

⁹ Von Engers, T. M. et al., POWER, Using UML/OCL for modeling legislation – an application report, in: Proc Int Conf on Artificial Intelligence & Law, ACM Press, New York, NY, 157-167.

¹⁰ Vgl Schweighofer et al., FN 7 und 8.

bytes von Texten durcharbeiten. Der Mensch muss aber die Vorgaben bestimmen und die Ergebnisse interpretieren.

EUR-Lex/CELEX haben als Textarchiv hohe Akzeptanz und Wertschätzung erzielt, trotzdem sind aber einige Verbesserungen nötig. Als einfache Schritte sind ein Echtzeitinformationssystem und die Nutzung des XML-Inputs zur besseren Suche und Präsentation zu nennen. Schon etwas schwieriger wird die bessere Visualisierung bei komplexen Prozessen sein. Schwierig wird die Strukturanalyse des Rechts für juristische Aufgaben, die noch durch Projektforschung auf ihre Tauglichkeit getestet werden muss. Semi-automatische Verfahren zur Klassifikation, Klassenbeschreibung und Verlinkung haben aber das Potential, dem Benutzer eine brauchbare Strukturierung der Dokumentflut zu geben und stellen einen wichtigen Schritt zum elektronischen Kommentar dar.