

Michael Sonntag / Josef Schaitl

Das individuelle Gesetzbuch – Open Government Data aus dem RIS

Das Rechtsinformationssystem des Bundes (RIS) ist eine weltweit herausragende Dokumentation des österreichischen Rechts, doch ist seine Funktionalität auf die Darstellung einzelner Paragraphen, Urteile oder Gesetze eingeschränkt. Die Erweiterung des Webangebotes eines derartigen Drittanbieters, um z.B. mehrere Gesetze auf einmal als PDF darzustellen, kann auf mehrere Arten realisiert werden, doch besitzen diese unterschiedliche Vor- und Nachteile sowohl technischer wie auch juristischer Natur. Diese werden hier kurz diskutiert und es wird eine Beispiels-Applikation vorgestellt, mit welcher aus dem RIS individuelle Gesetzbücher generiert werden können.

Category: Articles
Region: Austria

Citation: Michael Sonntag / Josef Schaitl, Das individuelle Gesetzbuch – Open Government Data aus dem RIS, in: Jusletter IT next: 11. September 2014 – Lachmayer

Inhaltsübersicht

- 1 Einleitung
- 2 Das RIS: Technische Umsetzung
- 3 Open Government Data
- 4 Datenabruf von Dritt-Webseiten und ihre Weiterverwendung
 - 4.1 Screen Scraping
 - 4.1.1 Technische Umsetzung
 - 4.1.2 Rechtliche Beurteilung
 - 4.1.3 Vor- und Nachteile
 - 4.2 Proxy
 - 4.2.1 Technische Umsetzung
 - 4.2.2 Rechtliche Beurteilung
 - 4.2.3 Vor- und Nachteile
 - 4.3 Webservices
 - 4.3.1 Technische Umsetzung
 - 4.3.2 Rechtliche Beurteilung
 - 4.3.3 Vor- und Nachteile
 - 4.4 Gesetzbücher aus dem RIS
 - 4.4.1 Umsetzung der neuen Version
 - 4.4.2 Benutzerschnittstelle und Ergebnisdokument
 - 4.4.3 Erweiterungsmöglichkeiten
- 5 Zusammenfassung

1 Einleitung

[Rz 1] Das Rechtsinformationssystem des Bundes enthält eine Dokumentation des Österreichischen Rechtsbestandes. Darüber hinaus ist es seit 2004 auch die Plattform für die authentische elektronische Publikation des Bundesrechts. An beiden Teilen war maßgeblich Friedrich Lachmayer beteiligt. In seinem Sinne soll hier der Versuch gemacht werden, einen Teil seines Werkes weiterzuführen und auszuweiten.

[Rz 2] Das Rechtsinformationssystem des Bundes ist im Internet unter dem URL <http://www.ris.bka.gv.at/> kostenfrei für jeden (weltweit) über Webseiten zugänglich. Seit kurzem existieren auch «Apps», um einen einfacheren Zugriff von mobilen Geräten aus zu erlauben¹. Diese bilden jedoch «lediglich» die Möglichkeiten der Webseiten nach und bieten eine der typischerweise geringeren Bildschirmgröße angepassten Benutzerschnittstelle, besitzen jedoch keine Zusatzfunktionen. Auch die Webseiten erlauben lediglich eine Suche (wenn auch nach vielen Kriterien) nach einzelnen Paragraphen sowie den Zugriff auf ganze Gesetze. Dies allerdings sowohl zum derzeitigen wie auch früheren Zeitpunkten (in gewissem Maße fand eine Rückerfassung statt). Zusätzlich zu den Gesetzen bzw. Gesetzblättern sind weitere Daten verfügbar, z.B. Entscheidungen div. Gerichte und anderer Behörden (etwa der Datenschutzkommission), Landes-/Gemeinderecht (unvollständig!) sowie ausgewählte Erlässe. Ein ähnliches Verzeichnis, jedoch mit anderem Inhalt, stellt die Finanzdokumentation dar (Findok²). Bei ihr bestehen auch explizite Regelungen für eine kommerzielle Verwertung, was jedoch eine hohe Lizenzgebühr erfordert³.

¹ <http://www.data.gv.at/anwendungen/risapp/> (9. Januar 2013)

² <https://findok.bmf.gv.at/findok/> (9. Januar 2013) Diese erlaubt es z.B. auch, Links auf mehrere Dokumente gleichzeitig zu erstellen. Allerdings führt dies lediglich zur Darstellung in Form einer Liste, analog einem Suchergebnis.

³ https://findok.bmf.gv.at/findok/https://findok.bmf.gv.at/hilfe/vertriebsinfo.htm#10_3_kommerzielle%20Nutzung (9. Januar 2013): €14.000 pro Jahr (Wert für 2010, also inzwischen ev. höher).

[Rz 3] Nicht möglich ist es jedoch im RIS, sich ein individuelles Gesetzbuch zusammenzustellen und als Gesamtdokument herunterzuladen. Es können lediglich von einzelnen Gesetzen, Paragraphen, Entscheidungen usw. PDFs generiert werden, die anschließend lokal manuell zusammengefügt werden müssten. Damit sind naturgemäß auch Zusatzfunktionen wie Inhalts- oder Stichwortverzeichnisse nicht automatisiert verfügbar. Ein individuelles Gesetzbuch, z.B. nach thematischem Zusammenhang, ähnlich wie von mehreren Verlagen in Papierform vertrieben, ist damit nicht mit vernünftigen Aufwand (Aktualisierung bei Novellen!) möglich.

[Rz 4] Genau dies wird aber u.U. gewünscht, um sich nicht nur selbst eine beliebige Sammlung nach eigenen Bedürfnissen zusammenstellen zu können, sondern um diese auch elektronisch zur Verfügung zu haben (z.B. für Volltextsuche) sowie einfach aktualisieren zu können. Damit kann ein Warten auf eine neue Auflage eines Druckwerkes vermieden werden. Zusätzlich könnte ein derartiges elektronisches Verzeichnis noch weitere Funktionen unterstützen:

- Auch Entscheidungen von Gerichten bzw. Behörden können aufgenommen werden (insofern im RIS enthalten), sowohl zusätzlich zu einem Gesetz/Paragraphen (z.B. vollautomatisch alle entsprechend annotierten Entscheidungen, aber ev. auch per Volltextsuche identifizierte) wie auch als reine Entscheidungssammlung.
- Eine elektronische Sammlung erlaubt ein einfaches Teilen mit anderen, sodass auch eher selten benötigte Sammlungen für Spezialgebiete analog «Crowdsourcing» vorbereitet sein könnten.
- Basierend auf einer konkreten Instanz einer Sammlung wäre es möglich Differenzen festzustellen, d.h. Benachrichtigungen zu verschicken wenn sich Änderungen durch Novellen ergeben, oder eine Gegenüberstellung zu erzeugen: Alter Text/Inhalt <-> Neuer Text/Inhalt.

2 Das RIS: Technische Umsetzung

[Rz 5] Bei dem Rechtsinformationssystem des Bundes handelt es sich um eine Web-Applikation, die auf eine Datenbank mit dem Bundesrecht im Hintergrund zugreift. Darüber hinaus wird über dieselbe Webseite (und ähnliche Suchmasken) auch Zugang zu weiteren Vorschriften, insb. dem Landesrecht angeboten. Hierbei wird jedoch auf andere Datenbanken zugegriffen, sodass das Format des Inhalts sich unterscheiden kann. Darüber hinaus sind noch Entscheidungen (insb. der obersten Gerichte aber auch sonstiger Behörden) teilweise vollständig, teilweise nur als Auswahl zugänglich. Auch Erlässe und Verordnungen sowie Vormaterialien wie Begutachtungsentwürfe und Regierungsvorlagen sind enthalten, wiederum in einem eigenen Datenformat. Die Zulieferung der Daten erfolgt von den jeweiligen erzeugenden Stellen, insb. um sich selbst die Umsetzung und den Betrieb einer entsprechenden Plattform zu ersparen.

[Rz 6] Die Inhalte des RIS (z.B. Bundesrecht) sind so organisiert, dass jeder Paragraph / Entscheidung / Erlass / ... ein eigenes Dokument darstellt und eine eindeutige und einmalige Dokumentennummer erhält. Diese ändert sich nicht mehr, sondern bei einer Novellierung dieses Paragraphen wird ein neues Dokument mit verändertem Inhalt angelegt, das ähnliche Metadaten enthält (also u.A. zum gleichen Gesetz gehört, dieselbe Paragraphennummer besitzt etc., aber z.B. ein anderes Datum des Inkrafttretens). Dokumente im RIS werden niemals gelöscht (eingeschränkter Lebenszyklus), da sie aus Dokumentationsgründen erhalten bleiben müssen. Sie sind intern in XML repräsentiert, wobei über die Webseite nur ein Abruf von daraus erzeugten HTML, PDF und RTF-Dokumenten möglich ist. Einzelne Großbereiche können (kostenpflichtig!) als Gesamt-

auszug auch im XML-Format bezogen werden (siehe oben analog der FinDok — ursprünglich auf CD-ROM, heute per USB-Festplatte oder Online-Zugang⁴).

3 Open Government Data

[Rz 7] Im Sinne der Offenheit von öffentlichen Verwaltungen (auch wenn es hier inhaltlich um Gesetzgebung bzw. Judikatur geht!) wurde ein «besserer» Zugang zum RIS eröffnet. Basierend auf den «Open Data-Prinzipien»⁵ sind hier insb. folgende Punkte besonders relevant:

- Primärquelle: Es besteht ein «direkter» Zugang zur Datenbank, sodass man nicht auf potentiell veraltete Kopien (siehe oben: käuflich erworbene Extrakte auf Datenträgern) angewiesen ist.
- Maschinenlesbarkeit: Wie oben erwähnt, sind die Daten im Ursprung in XML gespeichert, doch über die Webseiten ist ein Zugriff auf diesen «Quelltext» nicht möglich. HTML, PDF und RTF sind jedoch für eine maschinelle Weiterverarbeitung weniger geeignet. Über diese Schnittstelle werden die Dokumente hingegen im XML-Format ausgeliefert.
- Verwendung offener Standards/Dokumentation: Durch den Einsatz von XML und genauer Definition des Datenformats wird eine sinnvolle Weiterverarbeitung erst möglich, ohne vorher eine umfangreiche Inhaltsanalyse vorzunehmen. Weiters ist bei offizieller Dokumentation auch von Stabilität des Formats (seltene Änderungen) auszugehen.

[Rz 8] Die Maschinenlesbarkeit führt jedoch zu einem zusätzlichen potentiellen Problem, da in machen Gesetzen (z.B. der Straßenverkehrsordnung) auch Grafiken enthalten sind. Während diese in einer Webseiten-Darstellung keine Schwierigkeiten bereiten (Auslieferung durch den Server als Grafikdatei), können sie im XML problematisch sein. Konkret wurde dies so implementiert, dass die Grafiken als Binärdaten direkt in das XML eingebettet sind — und zwar als Bilder (mögliche Datentypen: GIF, JPG, TIFF, PNG), d.h. es handelt sich nicht um Vektordaten, wodurch automatisierter Verarbeitung Grenzen gesetzt sind (z.B. sind Vergrößerungen aufgrund fester Bildauflösung stark eingeschränkt). Doch spielt dies für die vorliegende Anwendung keine Rolle, da Bilder ohnehin nur dargestellt, aber nicht inhaltlich ausgewertet werden. «Exotische» Formate würden bei der Interpretation bzw. Umsetzung auch viel leichter zu Schwierigkeiten (passende Programmbibliothek verfügbar?) führen. Spezifiziert⁶ ist auch, dass eingebettete Elemente direkt PDF-Dateien sein können (siehe auch später zu sich daraus ergebenden Implementations-Schwierigkeiten).

[Rz 9] Ein wichtiger Aspekt bei Open Government Data ist immer die Lizenz, denn die Offenheit bedeutet nicht zwangsläufig, dass die Daten auch kostenlos verfügbar sind. Hier sticht das RIS leider nicht besonders hervor, da im Katalog der offenen Daten Österreichs dieses Feld leer ist⁷. Da nicht davon auszugehen ist, dass keinerlei Lizenz existiert oder jegliche beliebige Nutzung erlaubt ist, muss auf das RIS selbst zurückgegriffen werden. Doch auch dort sind keine

⁴ <http://www.ris.bka.gv.at/UI/Info.aspx>(9. Januar 2013)

⁵ <http://www.data.gv.at/hintergrund-infos/open-data-prinzipien/>(9. Januar 2013)

⁶ RIS OGDService Handbuch (Stand 30. Oktober 2012) http://data.bka.gv.at/RIS/Documents/RIS_OGD_Dokumentation.pdf(9. Januar 2013)

⁷ <http://www.data.gv.at/datensatz/?id=31430a9f-c8ba-4654-ab68-c9c3dff0361b>(9. Januar 2013)

Informationen auffindbar. Auf individuelle Anfrage hin⁸ wurde mitgeteilt, dass die Datenbanken Bundesrecht sowie Bundesgesetzblätter «im Internet frei verfügbar» sind, bei allen anderen Datenbanken jedoch für eine Veröffentlichung die Zustimmung der jeweiligen dateneinbringenden Stelle erforderlich wäre (insb. Gerichtsentscheidungen⁹). Dies stellt einen besonderen Hemmschuh dar und es sollte eine einheitliche Lizenzlinie gefunden — und explizit dokumentiert — werden. Dies sollte insb. deshalb unproblematisch sein, da Entscheidungen in Österreich ohnehin nur anonymisiert veröffentlicht werden und sie über das RIS selbst bereits weltweit frei ohne Anmeldung oder Gebühren zugreifbar sind. Besonders zu berücksichtigen wäre hierbei, dass sich im RIS auch Rechtssätze befinden, d.h. Zusammenfassungen wichtiger Elemente von Urteilen. Im Gegensatz zu den Urteilen selbst existiert für diese keine Ausnahme im Urheberrecht, sodass sich hierfür zusätzliche Rechtsprobleme stellen könnten, z.B. im Hinblick auf kommerzielle Verwertung.

[Rz 10] Die RIS-Daten sind über den Webservice anonym und ohne Anmeldung (kein «Zugriffstoken» o.Ä. erforderlich) abzufragen. Die Schnittstelle erscheint jedoch noch verbesserungsbedürftig. So wird das XML-Dokument etwa als HTML-codierter Inhalt eines Elements in einem XML-Dokument zurückgeliefert und bei Suchen werden die Ergebnisse immer «seitenweise» zurückgeliefert, d.h. unter Angabe der Anzahl an Suchergebnissen pro Seite (Auswahl aus vier Werten) und einer Seitennummer. Es wird also, offenbar der Einfachheit halber, direkt die existierende Webschnittstelle per Webservice «nachgebildet».

[Rz 11] Eine große Einschränkung ist, dass (zumindest derzeit, die Schnittstelle ist schon für mehr vorbereitet) über den Webservice lediglich das Bundesrecht abgefragt werden kann. Weder das Landesrecht (potentiell schwieriger, da anderes Format und unterschiedliche Datenquellen) noch Judikatur (sollte technisch deutlich einfacher zu lösen sein) sind im Augenblick darüber zugreifbar.

4 Datenabgriff von Dritt-Webseiten und ihre Weiterverwendung

[Rz 12] In vielen Fällen ist es erwünscht, Informationen, die sich auf den Webseiten Dritter befinden, automatisiert weiterzuverarbeiten. Dies kann sich auf die reinen Inhaltsdaten beziehen, aber auch auf Subframes oder gar ganze Webseiten. Aus Sicht der Sicherheit ist dies besonders gefährlich, da z.B. der Einbau einer fremden Anmelde-Seite in die eigene Webseite dazu führen könnte, dass das Passwort der Dritt-Seite dem Betreiber der einbettenden Seite bekannt wird. Daher besitzen Browser starke Sicherheitsvorkehrungen, um derartige Konstellationen zwar prinzipiell zuzulassen, aber Sicherheitsprobleme dennoch zu verhindern¹⁰. Technisch gesehen unterscheidet sich aber der unerwünschte Zugriff auf das Passwort nicht davon, sonstige Daten (z.B. Bewertungen von Produkten/Dienstleistungen oder Preis-/Lieferinformationen) auszulesen. Doch vielfach ist ebenso die umgekehrte Aktivität erforderlich: Eine Webseite muss «bedient» werden, einerseits um überhaupt an die gewünschten Informationen zu gelangen (typ. Beispiel: Ausfüllen

⁸ Im Rahmen eines anderen Projektes: E-Mail Antwort von Dr. Helga Stöger (12. Juli 2006).

⁹ Dies wird dadurch erhärtet, dass deren schriftliche Einverständniserklärung für den Bezug ihrer Daten aus dem RIS als Gesamtextrakt erforderlich ist: <http://www.ris.bka.gv.at/UI/Info.aspx> (9. Januar 2013).

¹⁰ Beispielsweise durch SOP (Same Origin Policy): Zugriff ist nur auf Daten möglich, die aus derselben Quelle (dem gleichen Server: Protokoll, Host und Port müssen übereinstimmen) stammen. https://developer.mozilla.org/en-US/docs/JavaScript/Same_origin_policy_for_JavaScript (9. Januar 2013).

eines Suchformulars), andererseits um Aktionen auszulösen (wie etwa das Buchen eines Fluges über eine Dritt-Webseite durch «klicken» eines Links).

[Rz 13] Drei Varianten der Umsetzung sollen hier sowohl aus technischer als auch rechtlicher Sicht untersucht werden — Eine Kombination von Themen die auch Friedrich Lachmayer immer verfolgt hat.

4.1 Screen Scraping

[Rz 14] Beim so genannten «Screen Scraping» wird eine Webseite von einem Programm abgerufen und anschließend ihr Quelltext, der HTML-Code, analysiert.

4.1.1 Technische Umsetzung

[Rz 15] Da es sich hierbei um einen normalen Seitenabruf handelt, erhält das Programm auch die normale Webseite (und nur diese: Z.B. eingebettete Frames müssen separat abgerufen werden): Mit allen Formatierungen, Links auf Grafiken, Navigationshinweisen, Textcodierungen etc. Das Extrahieren von Informationen ist daher sehr komplex und erfolgt typischerweise über den Kontext, z.B. mit Regeln wie «In der Tabelle mit der ID <xyz> ist im zweiten Absatz das fettgedruckte Wort die gewünschte Information». Sollte nun aus irgendeinem Grunde die Tabelle umbenannt, ein weiterer Absatz eingefügt oder die Formatierung verändert werden, so muss man auch das Auswertungsprogramm anpassen. Sind die Informationen extrahiert, so müssen sie noch interpretiert werden, etwa die Umwandlung von einem Text in eine Zahl («1,990.90» → 1990,90) oder der Abgleich mit einer Liste fester Werte (welche Kategorie bedeutet «größere Mengen auf Lager» als Lieferbarkeitshinweis?).

[Rz 16] Die Bedienung einer Webseite ist ebenfalls komplex zu realisieren, da einerseits die korrekten Elemente (Felder, Buttons, Links etc.) für die gewünschte Aktion gefunden werden müssen (Probleme wie oben), andererseits deren «Aktivierung» nicht immer einfach ist, da u.U. umfangreiches Verhalten eines Browsers nachgebildet werden muss (bei Formularen müssen auch versteckte Felder mitgesendet werden, Links werden oft über JavaScript aktiviert etc.). Die Alternative, das komplette Fernsteuern eines Browsers, z.B. über ein Plugin, ist äußerst komplex, langsam und wird in Produktiv-Systemen nicht eingesetzt (siehe jedoch unten bei Proxy).

4.1.2 Rechtliche Beurteilung

[Rz 17] Bei der rechtlichen Beurteilung sind zwei Hauptaspekte relevant: Ist diese Art des Zugriffs (der Informationserwerb) erlaubt, sowie dürfen die gewonnenen Informationen verwendet und wieder öffentlich zur Verfügung gestellt werden? Der zweite Aspekt ist für alle Varianten des Zugriffs identisch und soll in dieser Darstellung ausgeklammert bleiben. Insbesondere sind für dessen Beurteilung Urheber- und Wettbewerbsrecht relevant: Urheberrecht da es sich um Vervielfältigung/öffentliche Zurverfügungstellung von Werken handeln könnte, und Wettbewerbsrecht im Hinblick auf Ausbeutung fremder Leistungen.

[Rz 18] Der Fokus dieser Untersuchung liegt daher auf der Informationsgewinnung, welche wiederum zwei wichtige Diskussionspunkte enthält:

1. Kann der Zugriff auf eine Webseite auf Menschen beschränkt werden, d.h. können Computer

als «Besucher» ausgeschlossen werden?

[Rz 19] Ein solcher Ausschluss ist vertraglich problemlos möglich — insofern ein Vertrag existiert. Ist daher eine Benutzeranmeldung erforderlich, bevor Zugriff auf die gewünschten Daten erlaubt wird, so kann, z.B. in AGBs, vereinbart werden, wie der Dienst genutzt werden darf. Hierbei kann auch ein automatisierter Zugriff ausgeschlossen werden. Allerdings ist die Definition eines solchen Zugriffs nicht trivial, denn auch im Normalfall greift der Mensch nur indirekt und «automatisiert» über einen Webbrowser auf die Webseite zu. Und bei einem Abruf durch ein anderes Programm steht am Ende die Entscheidung eines Menschen für diesen Vorgang und fast immer auch als Konsument des Ergebnisses dahinter. Es wird daher in der Praxis u.A. versucht, den Vorgang technisch zu beschreiben und diesen auf Browser-Interaktion einzuschränken oder sonstige «automatisierte Mittel» auszuschließen. In diesem Kontext sind Suchmaschinen relevant, die exakt diese Tätigkeit ausüben: Automatisierter Abruf von Webseiten und Extraktion des Inhalts zur Weiterverarbeitung. Hierfür existiert der (inoffizielle!) Standard der «robots.txt»-Datei (oder äquivalente META-Tags in der Webseite selbst), mit dem derartigen Besuchern mitgeteilt werden kann, ob/welche Seiten von einem automatisierten Besuch bzw. Indexierung (separat einstellbar!) ausgeschlossen sein sollen. Doch wird diese Beschränkung als nicht rechtlich zwingend angesehen, sondern lediglich als Empfehlung/Wunsch (insb. da mit Suchmaschinen gerade kein Vertrag besteht!).

[Rz 20] Dazu sind weiters die Entscheidungen zur Google-Bildersuche¹¹ relevant, weil dort ähnliche Fälle entschieden wurden: Wenn etwas ins Internet gestellt wird und keine Gegenmaßnahmen (siehe unten) ergriffen wurden, so ist von einer Einwilligung für eine Speicherung sowie Zugänglichmachung von Vorschaubildern auszugehen, was zwangsweise auch den bloßen Zugriff beinhaltet. Üblicherweise wird man allerdings — im Gegensatz zu Suchmaschinen — nicht davon ausgehen können, dass eine Webseite für einen automatisierten Zugriff in Form von Screen Scraping zugänglich gemacht wurde, ebenso werden von den interessanten Daten kaum «Vorschaubilder», also in ihrem Informationsgehalt stark reduzierte Versionen, ausgelesen bzw. gespeichert. Umgekehrt kann jedoch nicht davon ausgegangen werden, dass eine Verpflichtung besteht, bestimmte Daten abzurufen (Datei robots.txt) oder auszuwerten (META-Tags), da laut der Entscheidung sogar individuelle direkt mitgeteilte Verbote irrelevant sind, sofern keine hinreichenden Schutzvorkehrungen (allgemein oder spezifisch) getroffen wurden — und sowohl robots.txt als auch META-Tags sind lediglich Erklärungen an die Allgemeinheit, aber keine Sicherungsmaßnahme.¹² Details hierzu sowie einige Urteile wurden schon in einem Artikel zu einer Vorversion der Software publiziert¹³.

2. Wenn Gegenmaßnahmen ergriffen werden, dürfen diese, bzw. welche davon, umgangen werden?

[Rz 21] Erforderlich ist laut einem Urteil des BGH lediglich, «... dass der Berechtigte überhaupt Schutzmaßnahmen getroffen hat, die für Dritte als solche erkennbar sind. Es reicht aus, dass die

¹¹ BGH 29. April 2010, I ZR 69/08, BGH 19. Oktober 2011, I ZR 140/10

¹² Siehe Leitsatz 2 sowie den Volltext (<http://www.suchmaschinen-und-recht.de/urteile/Oberlandesgericht-Jena-20080227.html>) zum Urteil im selben Fall in niedrigerer Instanz: OLG Jena 27. Februar 2008, 2 U 319/07. Vergleiche auch das zweite Urteil zur Bildersuche, wonach die Einwilligung selbst dann gilt, wenn das Bild von der Webseite eines Nicht-Berechtigten stammt: BGH 19. Oktober 2011, I ZR 140/10. Dies kann durch (auf der eigenen Webseite befindliche!) robots.txt Dateien niemals verhindert werden.

¹³ SONNTAG, Zur Urheberrechtlichen Zulässigkeit von Screen Scraping. In: *Schweighofer/Kummer*(Eds.): Europäische Projektkultur als Beitrag zur Rationalisierung des Rechts. Wien: OCG 2011, 141—148

Schutzmaßnahmen den Willen des Berechtigten erkennbar machen, den öffentlichen Zugang zu dem geschützten Werk nur mit den von ihm vorgesehenen Einschränkungen zu ermöglichen.»¹⁴. Wenn man diese, m.M. nach durchaus nicht so eindeutige, Meinung teilt, so bedeutet jede — sofern nur erkennbare — auch noch so kleine und unwirksame technische Maßnahme¹⁵, dass ein derartiger Zugriff nicht mehr erlaubt wäre, da bereits der Abruf der Webseite eine Vervielfältigung, und damit eine urheberrechtlich vorbehalten Handlung darstellt, die eben auf diese Weise vom Anbieter nicht vorgesehen ist.

4.1.3 Vor- und Nachteile

[Rz 22] Besonders vorteilhaft an Screen-Scraping ist, dass es, bei entsprechender Gestaltung des Programms, technisch nicht von einem Abruf durch einen Menschen unterscheidbar ist und keine Mitwirkung des Betreibers der Webseite mit den Quell-Daten erforderlich ist. Mit anderen Worten: Der Abruf von Informationen und die Bedienung sind technisch nicht an dessen Mitwirkung oder Zustimmung gebunden.

[Rz 23] Dies wird u.A. dazu genutzt, Metasuchmaschinen zu implementieren, etwa im Flugbereich: Die Webseiten mehrerer Fluggesellschaften werden automatisiert nach Flügen, die zu einer Anfrage eines Besuchers passen, abgefragt und die Ergebnisse direkt vergleichbar dargestellt. Fluggesellschaften wünschen dies oft gerade nicht, da hiermit fast ausschließlich ein Preisvergleich stattfindet, sodass Gegenmaßnahmen ergriffen werden. Eine Möglichkeit besteht etwa darin, Captchas einzusetzen um sicherzustellen, dass die (potentiell automatisiert) eingegebenen Daten auch tatsächlich von einem Menschen stammen. Andere Gegenmaßnahmen bestehen darin, eine Webseite möglichst komplex zu gestalten und häufig zu ändern (u.U. automatisierte Modifikation der Struktur ohne grafische Änderungen).

[Rz 24] Im Hinblick auf das RIS ist Screen Scraping durchaus eine Möglichkeit, da von aktiven Gegenmaßnahmen derzeit nicht (und in Zukunft wohl auch eher kaum) auszugehen ist und der Aufbau der Webseiten vergleichsweise einfach ist. So werden Metadaten explizit bezeichnet (z.B. «Geschäftszahl» im Fettdruck und darunter die konkrete Zahl). Schwierigkeiten ergeben sich daraus, dass z.B. Verweise als Links/JavaScript-Popups dargestellt werden. Auch der eigentliche Inhaltstext (Wortlaut eines Paragraphen/Entscheidung/...) ist potentiell problematisch im Hinblick auf eine Umformatierung: Absätze und Listen/Aufzählungen oder einzelne Paragraphen bei ganzen Gesetzen werden auf der Webseite in Text bzw. HTML umgesetzt, ohne dass eine logische Markierung existiert (eine Rekonstruktion über die Formatierungsanweisungen ist nur teilweise möglich: «» markiert etwa die Überschrift eines Paragraphen, sodass sich daraus ableiten lässt, dass hier ein neuer Paragraph in einem ganzen Gesetz beginnt).

¹⁴ BGH 29. April 2010, I ZR 39/08 «Session-ID»

¹⁵ Ein Beispiel könnte die Auswertung der Anfrage-Header sein, in denen normalerweise Browser-Typ und -Version mitgesendet werden. Automatische Abfragen müssten explizit vorgeben (technisch trivial), ein Browser zu sein, was hiermit schon ausreichen würde. Eine solche Sicherungsmaßnahme ist allerdings vom Client erst bei tatsächlichem Versuch erkennbar (Fehlermeldung).

4.2 Proxy

[Rz 25] Der Einbau fremder Elemente in die eigene Webseite per Proxy kann auf zweierlei Arten erfolgen: Einerseits kann die fremde Webseite direkt als Subframe eingebunden werden (kein echter Proxy, insb. ist auch eine Veränderung des bzw. ein Zugriff auf den fremden Inhalt wegen der SOP nicht möglich, sondern nur eine «gemeinsame» Präsentation), andererseits kann der Inhalt der Webseite tatsächlich über ein eigenes Serverprogramm geleitet werden. Dies ähnelt dem Screen Scraping darin, dass die Webseite von dem abgerufen wird, der sie verändern möchte, unterscheidet sich aber darin, dass der Inhalt nicht direkt ausgelesen oder bearbeitet wird, sondern typischerweise nur kleine Teile an JavaScript hinzugefügt sowie Links umgeschrieben werden. Der relevante «Inhalt» wird dann später über den Browser extrahiert. Dieser Ansatz eignet sich daher für die «Bearbeitung» im Rahmen eines interaktiven Besuchs, nicht jedoch für eine vollständige Automatisierung.

4.2.1 Technische Umsetzung

[Rz 26] Realisiert wird diese Methode indem der Server der die Daten integrierenden Webseite ähnlich einem Web-Proxy die Zugriffe über sich selbst umlenkt, was über die Konfiguration des Browsers, Umschreiben der Links o.Ä. erfolgen kann. Gemeinsam ist diesen Methoden, dass sie ohne großen Aufwand und unabhängig von den zu integrierenden Webseiteninhalten sind. Allerdings erfordert z.B. die Browser-Konfiguration spezielle Berechtigungen oder eine Mitarbeit des Endkunden, weshalb meist das Umschreiben von Links eingesetzt wird. Dies bedeutet, dass der Endbenutzer eine Webseite abrufen, der Server diese selbst von der ursprünglichen Quelle abholt, alle enthaltenen Links entsprechend abändert (im Gegensatz zur Extraktion von Daten größtenteils trivial), sodass sie auf sich selbst und nicht die «echte» Quelle zeigen, und dann die Webseite an den Endbenutzer ausliefert. Klickt dieser dann auf einen Link, landet er nicht an der Quelle sondern beim Proxy, der die Seite wiederum selbst von der Quelle abrufen und nach Auswechseln der Links an den Benutzer weiterleitet.

[Rz 27] Hintergrund dieses Vorgehens ist, dass auch bei ansonsten völlig unveränderter Weitergabe der Seite diese für den Browser vom selben Server stammt wie andere Elemente (z.B. Frames), sodass ein Zugriff innerhalb des Browsers durch JavaScript aufgrund der Same-Origin-Policy nun nicht mehr verhindert wird. Damit kann etwa ein Script in einem «äußeren» Frame problemlos auf den Inhalt der in einem Subframe enthaltenen fremden Seite zugreifen, auch wenn dieser dynamisch per JavaScript angefordert, der Inhalt im Browser entschlüsselt oder er sehr komplex und in fehlerhaftem HTML formuliert wurde¹⁶.

[Rz 28] Auch eine quasi inverse Variante kommt in der Praxis vor: Teile einer Webseite werden an Dritte weitergegeben, damit diese dort eine Aktion vornehmen. Konkret bedeutet das, dass zum «Schutz» vor dem Download von Gratis-Daten (z.B. Filme, Musik, Bilder) ein sogenanntes Captcha zu lösen ist. Dieses wird jedoch nicht vom Betreiber erzeugt, sondern von einer Dritt-Webseite weitergereicht, ebenso wie die Antwort wieder retour. So ist es z.B. möglich «vollautomatisch» — wenn auch mit Hilfe menschlicher Dritter! — Gratis-E-Mail-Accounts o.Ä. zu registrieren, indem deren Schutz gegen automatische Bedienung per Proxy weitergereicht wird.

¹⁶ Hier wird daher die Fehlertoleranz des Browsers ausgenutzt, der die Seite interpretiert. Anschließend wird nur mehr das Ergebnis dieses Vorgangs, z.B. über das DOM-Modell, extrahiert.

4.2.2 Rechtliche Beurteilung

[Rz 29] Vom urheberrechtlichen Standpunkt aus wird vom Server eine Vervielfältigung erstellt und an einen einzelnen Client weitergereicht. Allerdings erfolgt dies vollkommen automatisiert und ausschließlich auf Anweisung eines echten Dritten, des Endbenutzers. Die Vervielfältigung kann daher diesem zugerechnet werden. Potentiell kommt auch eine flüchtige/begleitende Vervielfältigung in Frage, was jedoch an einer eigenen wirtschaftlichen Bedeutung der Vervielfältigung scheitern kann: Wird die Seite mit eigener Werbung umrahmt, oder Daten hierfür extrahiert (um Informationen über Interessen des Nutzers zu erlangen), so kann nicht mehr davon gesprochen werden, dass die Vervielfältigung, die eine unabdingbare Voraussetzung hierfür ist, keine eigenständige wirtschaftliche Bedeutung¹⁷ besäße. Ebenso ist der alleinige (!) Zweck nicht die Übertragung zwischen Dritten, sondern diese erfolgt, um eine inhaltliche Veränderung bzw. Auswertung vornehmen zu können.

[Rz 30] Ein weiterer Aspekt ist, dass die Webseite verändert wird, worin eine Werkbearbeitung oder Werkentstellung liegen könnte. Letzteres wird kaum vorkommen, da die Änderungen üblicherweise minimal und vielfach komplett unsichtbar sind. Es wird daher weder die optische Gestaltung (grafisches Werk) noch die Verbindungen zu anderen Webseiten (Datenbankwerk) verändert. Das veränderte HTML selbst genießt keinen Schutz als Computerprogramm, sodass auch dies nicht relevant ist. Allerdings können Eingriffe in JavaScript (eindeutig ein Computerprogramm, oft auch ein geschütztes) erfolgen, z.B. um Zugriffe umzuleiten, Buttons zu deaktivieren, etc., was auch Aussehen und Verhalten der Webseite stark verändern kann. Ebenso problematisch sind Änderungen in der Gesamtdarstellung, etwa wenn «vollständige» Webseiten als Unterframe eingebaut oder Teil-Frames aus ihrem Kontext gerissen werden. Es kommt daher hier auf die konkrete Ausgestaltung an, ob die Veränderungen der Webseite legal sind oder nicht¹⁸.

[Rz 31] Ebenso wie bei Screen Scraping darf hiermit kein Zugangsschutz umgangen werden, wobei das oben angeführte Beispiel des Weiterreichens von Captchas ein typisches Beispiel wäre: Die Sicherheitsmaßnahme wird nicht technisch beseitigt, aber es ist klar dass eine Umgehung stattfindet, und dass diese nicht im Sinne dessen ist, der die Sicherheitsmaßnahme einsetzt.

4.2.3 Vor- und Nachteile

[Rz 32] Ein Vorteil ist, dass die Webseite nur geringfügig geändert werden muss und damit in vielen Fällen eine Änderung des Formats oder des Inhalts an der Quelle für den Einbindenden kein oder nur ein geringeres Problem darstellt. Umgekehrt ist jedoch auch klar, dass damit gleichfalls nur eingeschränkter Zugriff auf die Daten erreicht wird, da wiederum der Inhalt erst interpretiert werden muss. Für einen Webseiten-Anbieter ist es äußerst schwer, den Einsatz solcher Techniken zu verhindern. Möglichkeiten bestehen z.B. darin, die Sitzung nachzuverfolgen und nur bestimmte Sequenzen an Webseitenaufrufen zu erlauben. Dies behindert jedoch auch Lesezeichen und erfordert zur Umgehung nur, genau eine solche Reihenfolge einzuhalten.

[Rz 33] Am Beispiel des RIS sieht diese Methode wie folgt aus: Die komplette Analyse der Web-

¹⁷ Siehe dazu VOGELIN *Kucsko*, urheber.recht § 41a 4.4

¹⁸ Die Bearbeitung selbst ist erlaubt, aber das veränderte Werk wird hier wiederum der Öffentlichkeit zugänglich gemacht. Dass diese aus jeweils nur einem einzigen Benutzer besteht, spielt hier keine Rolle, da die gleichen (oder zumindest ähnliche) Änderungen für jeden Besucher erfolgen und dadurch eine sukzessive Öffentlichkeit gegeben ist — ähnlich den Exemplaren eines übersetzten Buches.

seiten des RIS und der Nachbau aller Suchformulare wäre sehr komplex. Wird die Webseite hingegen als Unterframe eingebunden, so ist kein Zugriff auf deren Inhalte möglich und somit ein Auslesen der Daten für die Erstellung eines individuellen Gesetzbuches unmöglich. Eine mögliche Lösung hierfür ist, ein Programm als Proxy zu verwenden und die RIS-Webseiten direkt «durchzuschleusen» und als Subframe einzubetten. Dann kann mittels JavaScript ein Zugriff auf die Dokumentennummer erfolgen (ohne jede Veränderung der RIS-Webseiten¹⁹) und daraus die gewünschte Sammlung des Benutzers aufgebaut werden, vollständig unabhängig davon, um welche Datenbank (Bundesrecht, Judikatur etc.) es sich handelt, und ohne eine eigene Suchfunktionalität implementieren zu müssen. Über diese Dokumentennummer ist später ein Abruf der Webseite für die Erzeugung einer Sammlung möglich, wobei allerdings die HTML-Datei umgewandelt werden muss: Entfernen von Kopf- und Fußzeilen usw., sodass nur mehr der eigentliche Inhalt übrig bleibt, auch wenn dieser nicht analysiert wird.

4.3 Webservices

[Rz 34] Eine dritte und relativ neue Möglichkeit besteht darin, vom Anbieter der Webseite bereitgestellte Webservices einzusetzen. Der wichtigste Unterschied hierbei ist, dass ein Webservice nur dann genützt werden kann, wenn — und insoweit — er vom Anbieter zur Verfügung gestellt wird, während Screen Scraping und Proxies auch ohne dessen Einverständnis betrieben werden können.

[Rz 35] Dies ist der typische Fall von Open Government Data, da die Alternative, Download der jeweils «aktuellen» Gesamtdatenbank, u.U. sehr viel mehr Datenverkehr erfordert und immer zu - mehr oder weniger - veralteten Daten führt. Allerdings bedeutet dies auch, permanent eine entsprechend leistungsfähige und einsatzbereite Infrastruktur vorzuhalten, unabhängig von der konkreten Nutzung²⁰. Es existieren 10 Prinzipien von OGD²¹, welche auch sehr konkrete technische Auswirkungen besitzen können.

4.3.1 Technische Umsetzung

[Rz 36] Webservices sind von der Idee her nichts Neues: Aufrufe von auf anderen Rechnern befindlichen Prozeduren bzw. gelagerten Daten existierten schon vorher, z.B. CORBA, COM oder Datenbankzugriff über Sockets. Die Besonderheit von Webservices liegt einerseits in der Konzeption, jeder Aufruf sollte nach Möglichkeit eigenständig und sehr einfach sein (vergleiche auch Serviceorientierte Architektur), wie der Umsetzung. Letztere erfolgt nicht über neue oder komplexe Protokolle sondern es werden «Webseiten» abgerufen (à z.B. keine Probleme mit Firewalls), die in einem einfachen maschinenlesbaren Format strukturiert sind (à leichte Programmierung und schnelles Parsen). Typischerweise wird XML verwendet (hier ist die Verarbeitungsgeschwindigkeit potentiell ein Problem), doch auch JSON (JavaScript Object Notation) findet häufig Ein-

¹⁹ Die konkrete Implementierung (ein Vorgänger der hier vorgestellten Software), fügte der Einfachheit halber jedoch ein minimales Script zur Extraktion der Dokumentennummer ein.

²⁰ Sofern nicht Cloud-Services o.Ä. eingesetzt werden bedeutet dies, dass eine Auslegung auf die maximale Nutzung erfolgen muss, will man nicht eine Reduktion der Qualität (z.B. verlängerte Antwortzeiten) akzeptieren.

²¹ Abschnitt 2 Open Government Data Prinzipien. In: *Eibl et al*, Rahmenbedingungen für Open Government Data Plattformen (1.1.0) http://reference.e-government.gv.at/uploads/media/OGD-1-1-0_20120730.pdf (29. Januar 2013)

satz.

[Rz 37] Eine spezifisches Frage stellt bei Webservices die Definition des Inhalts sowie des Vorgehensmodelles dar: Wie werden die Daten formatiert und welche Aufrufe sind wann möglich? Auch hierfür existieren Standards²², diese sind jedoch nicht so weit verbreitet wie Webservices selbst. Sind sie vorhanden besteht oft die Möglichkeit, notwendigen Schnittstellen-Programmcode vollautomatisch daraus zu erzeugen.

[Rz 38] Auf der Seite des Anbieters erfordern Webservices meist eine Umsetzung vom internen Datenformat (z.B. SQL-Datenbank) auf das Externe (z.B. XML), was einen zusätzlichen Aufwand bedeutet, aber auch eine Abstraktionsschicht einzieht: Ändert sich das interne Format, kann das Externe dennoch gleich bleiben. Darüber hinaus erlaubt dies, Daten nur eingeschränkt zu veröffentlichen, d.h. nur bestimmte Datenkategorien oder lediglich Untermengen sind öffentlich/gratis zugänglich.

[Rz 39] In vielen Fällen sind Webservices reine Datenabfragen, es werden also Informationen angeboten und von Dritten genutzt. Dies ist jedoch keine technische Einschränkung und es existiert eine Vielzahl an Webservices mit denen ebenso Aktionen ausgelöst werden können, z.B. Buchungen/Kauf von Waren oder Dienstleistungen oder das Einfügen von Daten auf dem Server im weitesten Sinne.

4.3.2 Rechtliche Beurteilung

[Rz 40] Bei Webservices stellen sich viel weniger rechtliche Probleme, da diese an die Mitarbeit, und damit die Zustimmung, des Bereitstellenden gebunden sind. Dieser kann daher «beliebige» Bedingungen voraussetzen und diese auch technisch durchsetzen (z.B. trivial über Benutzername und Passwort, die bei einer Anfrage mitzusenden sind). Zu beachten ist allerdings, dass es sich hierbei naturgemäß um AGBs handelt, sodass sie entsprechenden Anforderungen genügen müssen.

[Rz 41] Ein Aspekt ist jedoch eher ungewöhnlich: Hier sind mehrere Personen aktiv und in rechtlichen Beziehungen involviert. Dies steht im Gegensatz zum Screen Scraping bzw. Proxies, bei denen nur eine Person, der unmittelbare Nutzer, involviert ist und für diesen der eigentliche Anbieter der Daten (mehr bei Screen Scraping bzw. weniger bei Proxies) verborgen bleibt. Ebenso liegt üblicherweise kein Vertrag zwischen dem Nutzer der vorgestellten Techniken und dem «Datenlieferanten» vor.

[Rz 42] Bei Webservices sind hingegen einerseits der Entwickler der Applikation, der Zugang zum Testen benötigt sowie die genaue Definition der Daten und genau weiß wie und von wem die Daten abgerufen werden, andererseits der Endnutzer, welcher keinerlei technische Details kennt, nur ungefähr weiß, wo die Daten tatsächlich herkommen (z.B. RIS: «irgendwo vom Staat») vertraglich verbunden. Dieses Streckenverhältnis ist insb. bei OGD auch in einer weiteren Hinsicht wichtig. So kann etwa eine App kostenpflichtig angeboten werden. Was passiert jedoch, wenn die Datenquelle plötzlich versiegt oder das Datenformat ändert? Besteht dann für den Endnutzer ein Anspruch auf «Reparatur» der Anwendung? Wie lange (nächster Tag oder 1 Jahr später!)? Wer haftet für falsche/unvollständige Daten? Für den Anbieter von OGD ist es daher sehr wichtig,

²² Ein Beispiel hierfür ist WSDL (Web Services Description Language), das insb. Schnittstelle und Protokoll spezifizieren kann.

eine stabile und langfristige Schnittstelle unter genauen und dauerhaften rechtlichen Regelungen anzubieten, da er ansonsten Probleme haben dürfte, «Mittler» zu finden, welche mit ihren Anwendungen die Verbindung zum Endkunden herstellen.

[Rz 43] Ein weiterer Punkt ist der Datenschutz, da es sich u.U. um personenbezogene Daten handeln kann. Dies betrifft etwa das RIS, wo bei Entscheidungen naturgemäß auch Namen enthalten sind. In manchen Ländern erfolgt die Veröffentlichung unverändert, in Österreich jedoch anonymisiert. Und hier besteht durchaus ein Unterschied zwischen den «normalen» RIS-Webseiten und einem Webservice: Eine automatisierte Suche nach bestimmten Personen ist über Webseiten schwierig, bei einem (inhaltlich nicht-anonymisierten) Webservice jedoch deutlich leichter, auch ohne spezifische Suchmöglichkeiten, da die Daten sofort automatisiert weiterbehandelt werden können und eine Enumeration aller Datensätze leichter fällt (beim RIS könnten etwa alle Dokumentennummern problemlos durchprobiert werden!).

4.3.3 Vor- und Nachteile

[Rz 44] Ein besonderer Vorteil von Webservices ist naturgemäß die Einfachheit der Verarbeitung und die (potentiell: je nach vertraglicher Beziehung) vorhandenen Garantien. Dem steht gegenüber, dass in der Praxis vielfach keinerlei Zusage über Dauerhaftigkeit besteht: Die Schnittstelle kann jederzeit geändert oder eingestellt werden. Darüber hinaus ermöglicht es die gesonderte Implementierung von Webservices, diese «schlechter», z.B. langsamer, Anfragebegrenzung pro Zeiteinheit etc., zu behandeln als normale Webseitenaufrufe. Bei Screen-Scraping und Proxies erscheinen die Abrufe dem Server hingegen wie solche von Benutzern und können daher nicht oder nur sehr schwer diskriminiert werden.

[Rz 45] Dies ist z.B. beim RIS der Fall, wo der Abruf per Webservice deutlich langsamer ist als über Webseiten — dies könnte jedoch ganz trivial auf der Leistungsfähigkeit der eingesetzten Server beruhen: Der Dienst ist kostenlos, sodass die Webservice-Schnittstelle echten Zusatzaufwand darstellt. Ein sekundäres Problem beim RIS ist, dass derzeit nur das Bundesrecht als Webservice zugänglich ist, was technisch zwar vermutlich leicht erweiterbar wäre, jedoch zusätzlichen Aufwand für Erstellung (eher gering) sowie dauerhaften Betrieb erfordern würde. Der Grund dürfte rechtlicher Natur sein, da etwa das Landesrecht «Eigentum» der Länder ist und auch Entscheidungen nur mit Zustimmung des jeweiligen Gerichts verwendet, und damit auch auf eine neue Art veröffentlicht, werden dürfen (siehe auch oben).

4.4 Gesetzbücher aus dem RIS

[Rz 46] Um die Erstellung individueller Gesetzbücher zu ermöglichen wurde eine Anwendung basierend auf einem Proxy implementiert (siehe Abbildung 1 für die Benutzerschnittstelle, sowie Abbildung 2 für ein Beispiels-Ergebnis). Dies erfolgte noch vor der Verfügbarkeit der OGD-Schnittstelle. Hierbei konnten zwar beliebige Paragraphen und Entscheidungen aufgenommen werden (Zugriff durch Injektion eines kleinen Scripts, welches die Dokumentennummer ausliest), doch besaß diese Lösung mehrere Nachteile:

- Das Gesetzbuch wurde derart zusammengestellt, dass die entsprechenden Webseiten abgerufen, zu einer einzigen langen HTML-Webseite zusammengefasst, und diese anschließend nach PDF konvertiert wurde. Daraus ergab sich ein suboptimales Erscheinungsbild, das stark von der Webseitendarstellung abhängt und nur beschränkt umformatiert werden konnte.

- Da das System auf der Dokumentennummer basiert, konnten nur einzelne Paragraphen aber keine ganzen Gesetze integriert werden. Dies hätte jedoch durch eine Erweiterung der Umsetzung realisiert werden können. Problematisch ist eine solche Sonderbehandlung jedoch jedenfalls.
- Die Konvertierung von HTML nach PDF ist sehr ressourcenaufwändig und benötigte viel Zeit. Dies insb. deshalb, da HTML zwar genau definiert ist, aber sich nur äußerst wenige Webseiten exakt an diesen Standard halten. Die Software ist daher gezwungen, auch extrem fehlerhaftes HTML korrekt zu interpretieren (so wie es auch Browser tun). Dass dies bei der RIS Webseite nicht nötig ist reduzierte leider nicht die Komplexität der eingesetzten Bibliotheken.

[Rz 47] Aufgrund dieser Unzulänglichkeiten und der Einführung der OGD-Schnittstelle wurde eine Neuentwicklung gestartet. Diese ist eine eigene Anwendung, d.h. es wird nicht die RIS-Benutzeroberfläche übernommen, was besonders bei der Suche deutlich wird, die hier nicht mehr über die RIS-Formulare erfolgt.

Sie arbeiten aktuell mit dieser Sammlung: **Urheberrecht** (2 Dokumente)

Navigation
Sammlungen
Bearbeiten
Dokumente
RIS durchsuchen

Dokument hinzufügen

Anmelden oder Registrieren

Impressum

Optimized for mobile

Home | Kontakt | English

BUNDESKANZLERAMT RECHTSINFORMATIONSSYSTEM RIS

Bundesrecht Landesrecht Gemeinderecht EU-Recht Judikatur Erlässe Gesamtabfrage Übersicht

Suchbegriff

Webseiten
Bundeskanzleramt
HELP.gv.at
Parlament
Fundstellennachweis zu den Materialien
AVSV
SozDok

Informationen
Zum RIS
Open Government Data
RIS-App
Links auf Dokumente im RIS setzen
RIS Recherche für Microsoft Word
Linkliste

Bundesrecht konsolidiert
Hier finden Sie das österreichische Bundesrecht in konsolidierter Fassung, wobei Konsolidierung bedeutet, dass in einer Rechtsvorschrift sämtliche später kundgemachten Änderungen und Berichtigungen eingearbeitet wurden. Diese Dokumente dienen lediglich der Information, sind also rechtlich unverbindlich. Die Dokumentationseinheit (ein Dokument) ist ein Paragraph oder ein Artikel oder eine Anlage einer Rechtsvorschrift.

Bundesgesetzblatt authentisch ab 2004
Hier finden Sie das Bundesgesetzblatt in seiner rechtlich verbindlichen Form. Seit 1. Jänner 2004 ist die hier kundgemachte elektronische Fassung die authentische. Es ist zu beachten, dass nur die signierte Fassung rechtlich verbindlich ist.

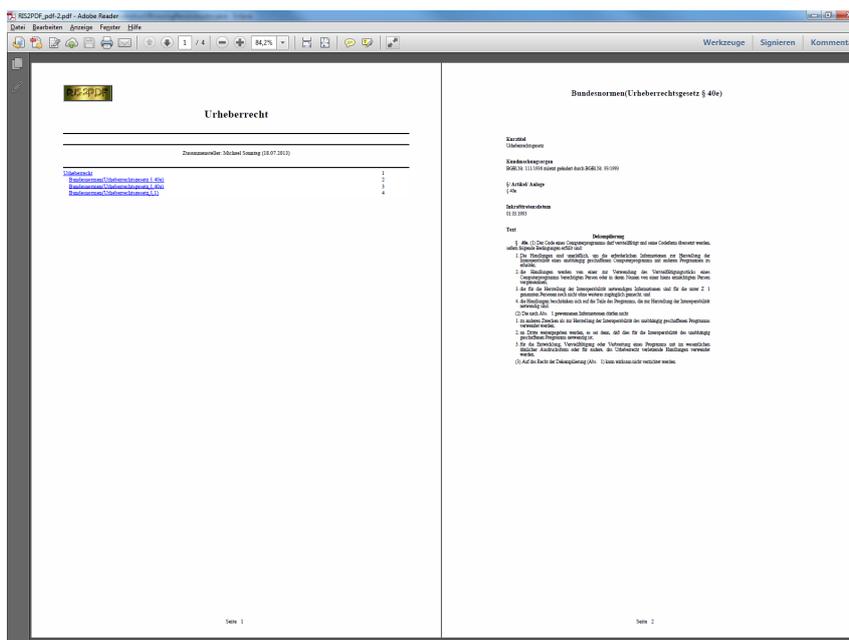
Bundesgesetzblatt 1945 - 2003
Hier sind die Jahrgänge 1945 bis einschließlich 2003 des Bundesgesetzblattes und das Staatsgesetzblatt 1945 in rechtlich unverbindlicher Form verfügbar.

Reichs-, Staats- und Bundesgesetzblatt 1848 - 1940
Hier finden Sie das Reichsgesetzblatt 1849 bis 1918, das Staatsgesetzblatt 1910 bis 1920, das Bundesgesetzblatt 1920 bis 1938 und das Gesetzblatt für das Land Österreich 1938 bis 1940 in rechtlich unverbindlicher Form.

Gesetzes- und Verordnungssammlungen 1740 - 1848 (extern)
Dieser Link verweist auf das Internetangebot der Österreichischen Nationalbibliothek, wo Sie insbesondere die Justizgesetzsammlung (1780 - 1848) und die Politische Gesetzessammlung (1792 - 1848) finden.

Deutsches Reichsgesetzblatt 1919 - 1945 (extern)
Dieser Link verweist auf das Internetangebot der Österreichischen Nationalbibliothek, wo Sie das Deutsche Reichsgesetzblatt von 1919 bis 1945 finden.

Begutachtungsentwürfe
Hier finden Sie Begutachtungsentwürfe, soweit diese verfügbar sind.



4.4.1 Umsetzung der neuen Version

[Rz 48] Die Umsetzung erfolgte in Java unter Einbeziehung mehrerer Bibliotheken durch Josef Schaitl in seiner Masterarbeit²³ und ist unter <http://opencodex.fim.uni-linz.ac.at/> erreichbar. Die konkrete Formatierung erfolgt in Form einer Umwandlung der Webservice-Ergebnisse (XML-Format) nach LaTeX, welches anschließend nach PDF konvertiert wird (wobei natürlich auch andere Formate möglich wären). Als problematisch stellten sich bei der Umsetzung heraus:

- Die Geschwindigkeit des Datenabrufs ist gering (siehe oben). Darum wurde ein Caching-System eingeführt, in dem heruntergeladene Dokumente gespeichert werden. Dies stellt beim RIS kein technisches Problem dar, da Dokumente unveränderlich sind (Novellierung des Paragraph à Neues Dokument mit neuer Dokumentennummer). Potentiell könnte diese Zwischenspeicherung jedoch rechtliche Schwierigkeiten bereiten. Da aber jede Ausgabe direkt auf eine Benutzeranforderung zurückgeht, ist dies legal. Fraglich könnte noch der Datenschutz sein, falls das RIS als geschützte Datenbank (§ 76c ff. UrhG) nach dem Urheberrecht anzusehen wäre, denn es entsteht kontinuierlich eine fast vollständige Kopie des RIS. Allerdings steht dies der normalen Verwertung der Datenbank nicht entgegen²⁴ und beeinträchtigt die berechtigten Interessen des Herstellers auch nicht unzumutbar (ganz im Gegenteil: kostenlos, Bandbreitensparnis etc.). Darüber hinaus wurde der Abruf parallelisiert, sodass mehrere Dokumente überlappend abgerufen werden können²⁵.
- Lediglich das Bundesrecht ist derzeit über die OGD-Schnittstelle abfrag- und zugreifbar. Eine Kombination mit der vorhergehenden Applikation ist jedoch nicht möglich, sodass der-

²³ Derzeit in Fertigstellung befindlich: *Schaitl*, OpenCodex — Erzeugung individualisierter Gesetzbücher aus den Daten des RIS. Master Thesis an der JKU.

²⁴ Es werden weiterhin nur Gesetzbücher auf individuelle Anfrage erstellt. Anders wäre dies ev., wenn der Cache auch so als Ganzes zugänglich gemacht würde.

²⁵ Denn das Problem ist nicht die Übermittlungsgeschwindigkeit sondern die Verzögerung, bis mit der Übermittlung begonnen wird (Latenz).

zeit weder Landesrecht noch Judikatur enthalten sind. Eine Einbindung derselben in das Programm wäre jedoch vergleichsweise trivial, da lediglich eine entsprechende Suchmaske eingebaut werden müsste: Es handelt sich um gleichartige Dokumente mit anderem Inhalt (wahrscheinlich wären zusätzlich Anpassungen der Formatierungsanweisungen erforderlich).

- Das zurückgegebene Format der Nutzdaten ist XML, welches selbst in XML (der Webservice-Antwort) eingebettet ist, wobei aber auch *character entities* und UTF-8 Sonderzeichen eingesetzt werden. Diese müssen zuerst umgewandelt werden, bevor der Text mittels LaTeX weiterverarbeitet werden kann. Der Grund für diese Kodierung ist nicht bekannt, da der Text selbst, genau wie die Webservice-Antwort, in UTF-8 Kodierung übermittelt wird und damit fast alle *character entities* überflüssig wären. Allerdings wird in der OGD-Spezifikation erklärt, dass das «Element <Dokumenteninhalt> [] die *formatierten* Nutzdaten [enthält]» (Hervorhebung durch Verf.). Mit anderen Worten, es werden keine strukturierten Daten übermittelt, sondern solche in einer bestimmten Vorformatierung. Dies ist nachteilig, aber wohl in der langen Historie des RIS zu begründen.
- Manche Gesetze enthalten eingebettete weitere Elemente, typischerweise Bilder (z.B. die Straßenverkehrsordnung; Schilder). Diese sind als Binärdateien direkt in das Ergebnisdokument eingebettet. Damit wird ein zusätzlicher Abruf erspart, allerdings die Datenmenge auch stark vergrößert. Die Verarbeitung dieser Elemente stellt potentiell Probleme dar, je nach Datentyp. Möglich sind laut Spezifikation XML, PDF, sowie GIF/JPG/TIFF/PNG²⁶. Bilder werden einfach in das Ergebnisdokument eingebettet, ebenso wie PDF, doch bei den Dokumentenformaten (z.B. DOC) ist eine standardisierte Behandlung nicht möglich, daher wird in der Ausgabe eine Warnung eingefügt. XML-Dateien werden im Quelltext dargestellt — was zwar eine Einbindung erlaubt, aber in vielen Fällen wohl nicht das eigentlich gewünschte Ergebnis darstellen wird. Praktisch wurden bei Tests (nur Bundesrecht!) bisher ausschließlich eingebettete Bilder, aber keine sonstigen Formate, angetroffen.

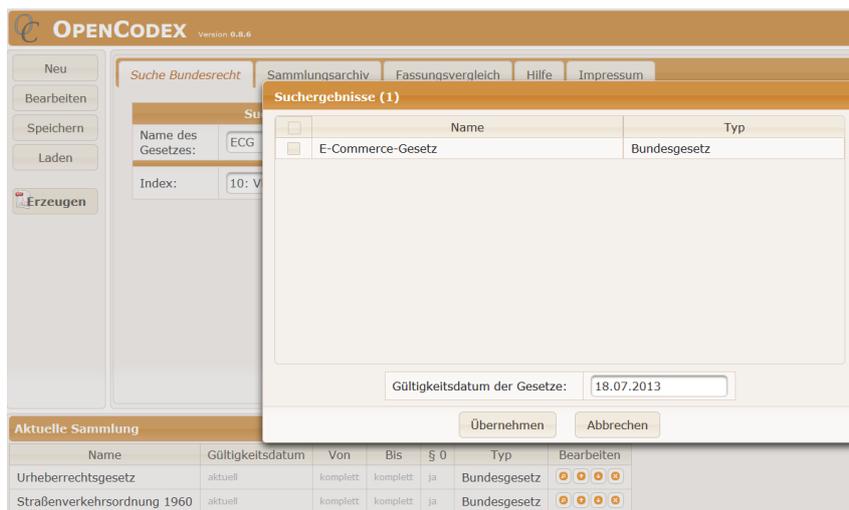
[Rz 49] Ein erstelltes Gesetzbuch besteht aus mehreren Elementen. Erstens, einer Titelseite mit einem konfigurierbaren Titel und Untertitel. Hier ist auch das Datum der Erzeugung ersichtlich. Anschließend folgt ein Inhaltsverzeichnis, das aus den in das Dokument eingebundenen Elementen generiert wird, d.h. bei Einfügen eines Gesetzes nur dessen Titel, bei einzelnen Paragraphen jedoch deren jeweilige Titel (natürlich jeweils inkl. Seitennummer). Danach folgt der eigentliche Text, wobei eine grafische Gestaltung ähnlich wie in bekannten Gesetzbüchern Einsatz findet: Zweispaltiger Text in kleiner Schrift (Schriftgröße ist einstellbar) mit «Marken» am äußeren rechten Rand, welche den Gesetzes-Kurztitel enthalten. Optional ist auch eine ein- oder dreispaltige Ausgabe möglich. Abschließend findet sich noch ein Schlagwortverzeichnis aus den bei den einzelnen Paragraphen annotierten Begriffen.

[Rz 50] Eine einmal erstellte Sammlung kann gespeichert werden, wobei als Format in XML serialisierte Objekte dienen. Diese könnten bei Bedarf auch durch andere Programme verarbeitet oder erstellt werden. Sie enthalten keine personen- oder sitzungsbezogenen Daten (ausgenommen ev. Titel/Untertitel) und können daher problem- und gefahrlos zwischen Nutzern ausgetauscht werden. Ein (optionales) öffentliches Verzeichnis von fertigen Zusammenstellungen ist damit möglich, aber derzeit noch nicht implementiert.

²⁶ Nach Auskunft des RIS zusätzlich auch RTF, DOC, DOCX und ODT.

4.4.2 Benutzerschnittstelle und Ergebnisdokument

[Rz 51] Als Beispiel wird hier eine Kombination mehrerer Paragraphen/Gesetze verwendet. In Abbildung 3 sieht man die (derzeitige und noch in Entwicklung befindliche) Benutzerschnittstelle, wobei bereits einige Elemente ausgewählt wurden (siehe unten «Aktuelle Sammlung») sowie das Suchergebnis (Suche nach «ECG» im Index des Bundesrechts) für weitere hinzuzufügende Elemente. In Abbildung 4 wird ein Ergebnisdokument dargestellt (normale Inhaltseite).



- 31 -

**1 StVO
1960**

2. für die Anbringung der Hinweiszeichen „Wegweiser“ sowie

3. für die Anbringung von Straßenverkehrszeichen, deren Inhalt miteinander in Zusammenhang steht.

Die Anbringung sonstiger Beschriftungen, bildlicher Darstellungen, Tafeln oder dgl. auf derselben Anbringungsrichtung bewirkt - unbeschadet der §§ 31 Abs. 2 und 53 Abs. 1 Z 17a - nicht die Unwirksamkeit der Kundmachung einer Verordnung (§ 44 Abs. 1).

(5) Der Abstand zwischen dem unteren Rand eines Straßenverkehrszeichens und der Fahrbahn darf bei seitlicher Anbringung nicht weniger als 0,60 m und nur in Ausnahmefällen mehr als 2,50 m, bei Anbringung oberhalb der Fahrbahn nicht weniger als 4,50 m und nur in Ausnahmefällen oder bei Verwendung beleuchteter Straßenverkehrszeichen mehr als 5,50 m betragen, sofern sich aus den Bestimmungen dieses Bundesgesetzes bei einzelnen Straßenverkehrszeichen nichts anderes ergibt. Bei seitlicher Anbringung darf der seitliche Abstand zwischen dem der Fahrbahn zunächst liegenden Rand eines Straßenverkehrszeichens und dem Fahrbahnrand im Ortsgebiet nicht weniger als 0,30 m und nur in Ausnahmefällen mehr als 2 m, auf Freilandstraßen nur in Ausnahmefällen weniger als 1 m und mehr als 2,50 m betragen. Sind auf einer Anbringungsrichtung mehr als ein Straßenverkehrszeichen angebracht, so gelten bei untereinander angebrachten Zeichen die Maßangaben bezüglich des Höhenabstandes für das untere Zeichen, bei nebeneinander angebrachten Zeichen die Maßangaben bezüglich des Seitenabstandes für das näher der Fahrbahn angebrachte Zeichen. Die weiteren Zeichen sind in einem solchen Fall entsprechend den Größenverhältnissen anzubringen.

(6) Wenn es die Verkehrssicherheit erfordert, insbesondere bei unübersichtlichem Straßenverlauf, sind in angemessener Entfernung vor einem nach den Bestimmungen der §§ 49, 50, 52 oder 53 angebrachten Straßenverkehrszeichen ein oder mehrere gleiche Zeichen - ausgenommen beim Zeichen „Halt“ - anzubringen, unter denen auf einer Zusatztafel nach § 54 Abs. 5 lit. a die Entfernung bis zu der Straßenstelle anzugeben ist, auf die sich das Straßenverkehrszeichen bezieht. Dies gilt insbesondere für die Gefahrenzeichen „Fußgängerübergang“ und „Voranzeiger für Verkehrsampel“, für die Vorschriftszeichen „Vorrang geben“ und „Wartepflicht bei Gegenverkehr“ sowie für das Hinweiszeichen „Krankenhaus“. Wird das Vorschriftszeichen „Halt“ vorangekündigt, so ist hierfür das Vorschriftszeichen „Vorrang geben“ mit einer Zusatztafel nach § 54 Abs. 5 lit. c zu verwenden.

§ 49. Allgemeines über Gefahrenzeichen.

(1) Die Gefahrenzeichen kündigen an, daß sich in der Fahrtrichtung auf der Fahrbahn Gefahrenstellen befinden. Die Lenker von Fahrzeugen haben sich in geeigneter Weise, erforderlichenfalls durch Verminderung der Geschwindigkeit, der angekündigten Gefahr entsprechend zu verhalten.

(2) Auf Autobahnen sind die Gefahrenzeichen 250 m bis 400 m, auf anderen Straßen 150 m bis 250 m vor der Gefahrenstelle anzubringen, sofern sich aus § 50 nichts anderes ergibt.

(3) Wenn es jedoch der Verkehrssicherheit besser entspricht, sind die Gefahrenzeichen in einer anderen als im Abs. 2 bezeichneten Entfernung anzubringen. In einem solchen Fall ist auf Freilandstraßen unter dem Zeichen auf einer Zusatztafel nach § 54 Abs. 5 lit. a die Entfernung bis zur Gefahrenstelle anzugeben.

(4) Wenn sich Gefahrenstellen über einen längeren Straßenabschnitt erstrecken (wie etwa Gefälle, Schleudergefahr, Steinschlag, Wildwechsel u. a.) und dies nicht erkennbar ist oder nicht vermutet werden kann, so ist auf einer Zusatztafel nach § 54 Abs. 5 lit. b die Länge der Gefahrenstelle

anzugeben. Innerhalb einer längeren Gefahrenstelle ist das betreffende Gefahrenzeichen zu wiederholen, wenn dies die Verkehrssicherheit erfordert.

§ 50. Die Gefahrenzeichen.

Die Gefahrenzeichen sind

1. „QUERRINNE“ oder „AUFWÖLBUNG“



Dieses Zeichen zeigt Hindernisse, wie Querrinnen, Aufwölbungen oder aufgewölbte Brücken, an.

2. „GEFÄHRLICHE KURVEN“ oder „GEFÄHRLICHE KURVE“



Diese Zeichen zeigen an:

a) eine Rechtskurve,
b) eine Linkskurve,
c) eine Doppelkurve rechts beginnend,
d) eine Doppelkurve links beginnend;

sie sind vor Kurven, die wegen ihrer Beschaffenheit oder Unübersichtlichkeit gefährlich sind, dem Verlauf der Kurve entsprechend anzubringen. Sind auf einer kurvenreichen Strecke die Kurven nicht mehr als 240 m voneinander entfernt, so genügt die Anbringung eines der ersten Kurve entsprechenden Zeichens nach lit. c oder d, unter dem auf einer Zusatztafel nach § 54 Abs. 5 lit. b die Länge der kurvenreichen Strecke anzugeben ist.

4.4.3 Erweiterungsmöglichkeiten

[Rz 52] Erweiterungen des Projektes sind in verschiedene Richtungen möglich:

- Wie schon erwähnt, könnte eine Möglichkeit eingebaut werden, Sammlungen zu teilen. Dies könnte durch eine einfache Schaltfläche erfolgen, welche die Konfiguration in einem öffentlichen Bereich speichert. Dieser könnte als Liste dargestellt bzw. durchsuchbar sein. Da dies auch ohne Benutzerregistrierung möglich ist, wäre jedoch eine gewisse Qualitätskontrolle erforderlich: Redaktionell oder über Bewertungen von Benutzern oder Statistiken über die Häufigkeit des Zugriffs darauf. Auch eine Möglichkeit zur Entfernung von schlechten/inaktuellen Sammlungen müsste eingebaut werden. Mit anderen Worten: Dies stellt nur geringe technische Probleme aber einen deutlich höheren Aufwand für Organisation und Implementation einer Benutzerschnittstelle dar.

- Aktualisierung einer Sammlung: Wird eine «alte» Sammlung geladen, so könnte eine Abfrage erfolgen, ob diese noch dem aktuellen Stand entspricht bzw. gegebenenfalls aktualisiert werden. Hierzu wäre es erforderlich, zu einer Dokumentennummer die aktuellste Version zu bestimmen.
- Um die Benutzbarkeit zu verbessern könnten interne Referenzen auf andere Paragraphen (ev. auch Absätze; schwieriger da meist keine vollständige Referenz) als klickbare Querverweise integriert werden. Dies erfordert jedoch das Parsen des Textes, da derartige Referenzen im Gesetzestext lediglich als normaler Text enthalten sind und keine Metainformationen (z.B. Dokumentennummer des Ziels oder die offizielle Abkürzung des betreffenden Gesetzes) enthalten.
- Eine größere Erweiterung wäre die Möglichkeit von Gegenüberstellungen: Die Differenz zwischen einem Gesetz zu zwei unterschiedlichen Zeitpunkten. Basierend auf der Dokumentennummer kann erkannt werden, ob es sich um den gleichen Inhalt handelt. Schwieriger ist jedoch das Einfügen bzw. Aufheben von Paragraphen/Artikeln. Innerhalb der Paragraphen, d.h. für den Text an sich, könnten klassische Tools zur Anzeige von Textdifferenzen eingesetzt werden.
- Benachrichtigungen über Aktualisierungen für ein individuelles Gesetzbuch könnten eingerichtet werden. Dies wäre technisch einfach zu realisieren, bedingte jedoch eine große konzeptionelle Umstellung in der Benutzeroberfläche. Derzeit ist die Nutzung anonym möglich. Sollten individuelle Benachrichtigungen erfolgen, müsste eine Benutzerregistrierung eingeführt werden — mit allem was daraus folgt (Passwort-Rücksetzung, Auskunftsrechte etc.). Alternativ könnte für jedes öffentliche (Anderen zur Verfügung gestellte) Gesetzbuch z.B. ein RSS-Feed eingerichtet werden, in dem Aktualisierungen dieser Sammlung mitgeteilt werden.

5 Zusammenfassung

[Rz 53] Die Einführung des Rechtsinformationssystems des Bundes war ein großer und international beachteter Fortschritt, der maßgeblich auf Friedrich Lachmayer zurückging. Doch dieser blieb dort nicht stehen sondern arbeitete weiter daran, das RIS zu verbessern und für andere Nutzer zu öffnen. Der derzeitige Stand ist bereits eine starke Verbesserung im Vergleich zur Initialversion: Vor dem RIS war es praktisch ausgeschlossen, individuelle Gesetzbücher zu erstellen, mit der Öffnung zum Internet wurde dies erstmals, wenn auch auf niedriger Ebene, möglich, und mit der Einführung von Open Government Data ist es nunmehr möglich, optisch ansprechende und einfach nutzbare Sammlungen zu erstellen. Diese können auch als Ausgangspunkt von wissenschaftlichen Analysen dienen, wie z.B. im Hinblick auf die Querverweise angedeutet, oder mit anderen Daten verbunden werden (etwa aus dem Parlament).

[Rz 54] Wünschenswert wäre eine Ausweitung auf die anderen Elemente des RIS. Dies bedeutet einerseits einen vermutlich vergleichsweise geringen technischen Aufwand, andererseits die Definition einer neuen Schnittstelle: Jede Art von Entscheidung besitzt zwar ähnliche Elemente, aber auch Eigenheiten, und unterscheidet sich naturgemäß von Gesetzen. Etwas einfacher sollte eine Integration der Gesetzblätter sein, zumindest was die authentische Verlautbarung betrifft, da diese ohnehin bereits im XML-Format vorliegen. Für die vorgestellte Applikation sind letztere allerdings weniger interessant.

[Rz 55] Wie die Untersuchung und die beiden Beispiele gezeigt haben, ist es zwar möglich (und

vielfach auch legal) Informationen von Webseiten zu extrahieren und weiterzuverwenden, doch erst eine Zusammenarbeit mit der Datenanbieter ermöglicht es, komplexe und dennoch zuverlässige Lösungen zu erstellen. Dies enthält einerseits technische Aspekte — wie ermögli- che ich es Dritten auf einfache Weise auf meine Daten zuzugreifen, ohne diese dadurch zu gefährden? — andererseits rechtliche — unter welcher Lizenz erfolgt das Angebot bzw. welche Aktionen dürfen auf Webseiten in welchem Kontext ausgeführt werden? — beides Elemente, die Friedrich Lachmayer immer zu verbinden wusste. Schlussendlich ist noch auf seine vielfältigen und erfolgreichen Bemühungen hinzuweisen, Zusammenhänge graphisch darzustellen. Automatisiert ist dies weiterhin schwer möglich, doch Ideen wie die Gegenüberstellung von Gesetzen zu verschiedenen Zeitpunkten bzw. interne Verlinkungen nehmen hier Anleihen, um Syntax mit Semantik und Visualisierung zu verbinden, wie es stets Friedrich Lachmayers Anliegen war.

Michael Sonntag, Institut für Informationsverarbeitung und Mikroprozessortechnik (FIM), Johannes Kepler Universität Linz, Österreich. Josef Schaitl, Institut für Informationsverarbeitung und Mikroprozessortechnik (FIM), Johannes Kepler Universität Linz, Österreich.