

XML in Judikaturdatenbanken

Felix Gantner

infolex

Glasauergasse 15, 1130 Wien

gantner@infolex.at

Schlagnworte: XML, Judikaturdokumentation, Judikaturdatenbank VwGH Wien, Saarbrücker Standard für Gerichtsentscheidungen

Abstract: Der vorliegende Beitrag erläutert den Einsatz von XML, einem Standard zur Beschreibung von Dokumentstrukturen, in der Judikaturdokumentation. Es werden die existierenden XML-Standards für den Dokumentenaustausch mit XML („Saarbrücker Standard“) und die Verwendung von XML in der Praxis am Beispiel des im VwGH Wien verwendeten Systems zur Geschäftsstellenverwaltung und Judikaturdokumentation dargestellt.

1. XML

XML (Extensible Markup Language) ist ein Standard für die Dokumentstrukturierung und -bearbeitung. Er ist seit dem 10.2.1998 durch das World Wide Web Consortium (W3C) als Recommendation (Standard) in Kraft gesetzt.¹ XML wird damit weltweit unterstützt und eingesetzt.

Auch wenn XML in erster Linie als Datenformat für das Internet entwickelt wurde, hat es sich inzwischen in fast allen Bereichen der Datenverarbeitung etabliert.

XML selbst ist kein gewöhnliches Datenformat, sondern legt eine Metasprache zur Definition und Formulierung von Dokumentstrukturen

¹ Vgl aktuell zum Einsatz von XML/SGML im Rechtsbereich: *Konzelmann*, Rechnergestützte Edition von Normtexten, JurPC Web-Dok. 66/2001 (<http://www.jurpc.de>); *Salminen/Lyytikäinen/Tiitinen/Mustajarvi*, Experiences of SGML Standardization: The Case of the Finnish Legislative Documents, Proceedings of the 34th Hawaii International Conference on System Sciences - 2001 (<http://dlib.computer.org/conferen/hicss/0981/pdf/09815003.pdf>); *Schweighofer*, Wissensrepräsentation in Information Retrieval-Systemen am Beispiel des EU-Rechts (2000); *Sjoeberg*, Critical Factors in Legal Document Management: A study of standardised markup languages. The Corpus Legis Project (1998).

fest. Es basiert – genauso wie HTML – auf SGML² und übernimmt damit auch das grundlegende Konzept der Trennung von Form und Inhalt.

Mit XML-DTD³ oder XML-Schemas wird die Struktur eines Dokumentes festgelegt und wie Inhalte in einem XML-Dokument markiert („ausgezeichnet“) werden.

Die Darstellung der Inhalte – z.B. ob ein bestimmter Textteil fett oder kursiv dargestellt wird – wird nicht über die XML-Strukturen, sondern über eigens definierte Regeln für die Anzeige bzw. Ausgabe festgelegt.⁴ Sie kann auch für dasselbe XML-Dokument in unterschiedlichen Medien oder Zusammenhängen variieren.

2. Datenformate und XML

XML ist eine *standardisierte* Metasprache zur Definition von Datenformaten. Wer XML einsetzt, kann Daten nach seinen Bedürfnissen strukturieren und aufbereiten, verwendet jedoch immer dieselben Grundbausteine, um die Dokumentstruktur zu definieren. Diese Grundbausteine können sowohl von einem Computer bearbeitet, als auch von Menschen gelesen und verstanden werden.

Wird das Dokument weitergegeben oder -bearbeitet, so sichert die Einhaltung des Standards, dass auch beim Empfänger die Dokumente mit geringem Aufwand gelesen und bearbeitet werden können. Voraussetzung dafür ist nur, dass das Know-how für die Bearbeitung von XML-Dokumenten vorhanden ist.

² SGML (Standard Generalized Markup Language) ist ein Standard für die Dokumentenbeschreibung (ISO 8879; vgl. dazu *Goldfarb*, The SGML Handbook [1990]). SGML wird schon lange im juristischen Bereich eingesetzt. 1969 wurde GML (Generalized Markup Language) – die Ausgangsbasis für SGML – im Rahmen der Realisierung eines integrierten „law office information system“ entwickelt (*Goldfarb*, 568). Zum Einsatz von SGML im juristischen Umfeld vgl. z.B. auch *Burger/Quirchmayr/Reich/Tjoa*, Managing Structured Documents in Distributed Publishing Environments, Proceedings 5th International Conference DEXA 1994, 83 ff; *Poulin – Lavoie – Huard*, Supreme Court of Canada's cases on the Internet via SGML, E Law – Murdoch University Electronic Journal of Law, Vol 4, No 3 (September 1997), <http://www.murdoch.edu.au/elaw/issues/v4n3/poul43.html>; *Poulin/Huard/Lavoie*, The other formalization of Law: SGML modelling and tagging, in: Proceedings of The Sixth International Conference on Artificial Intelligence & Law - ICAIL 1997, 82 ff.

³ Document Type Definition.

⁴ Im Gegensatz dazu vermischen herkömmliche Textverarbeitungssysteme die Anzeige und die logische Struktur.

Damit ermöglicht XML eine wesentliche Vereinfachung des Datenaustauschs. Der Erzeuger von Dokumenten ist nicht mehr gezwungen, Dokumentformate neu zu erfinden, und der Empfänger muss nicht mehr für jedes neue Datenformat eigene Analyseprogramme entwickeln, wenn er die erhaltenen Daten mit dem Computer verarbeiten will.

2.1. Ein Beispiel: Datenformate für das RIS

Dass sich Datenformate mit geänderten technischen Möglichkeiten verändern und angepasst werden müssen, zeigt die Entwicklung beim RIS⁵ in den letzten Jahren. Die Bereitstellung der Daten im Intranet des Bundes bzw. im Internet war auch mit einem Wechsel des EDV-Systems verbunden. Dies hatte auch einen Wechsel der Einbringungsformate zur Folge.

Die folgende Tabelle stellt die unterschiedlichen Datenformate, die vor dem 1.1.2000 und jetzt verwendet werden, am Beispiel eines Dokumentteils gegenüber:

Alt	100 JWR/1984120002/19850114X01
Neu	-DOKNR- JWR/1984120002/19850114X01
Alt	110RS* Erkenntnis 1985/01/14 84/12/0002
Neu	-KINFO- RS* Erkenntnis 1985/01/14 84/12/0002
Alt	200VWGH Erkenntnis
Neu	-TYP- VWGH Erkenntnis
Alt	2501985/01/14
Neu	-DATUM- 1985/01/14
Alt	30084/12/0002
Neu	-KENNUNG- 84/12/0002

Das alte Datenformat kennzeichnete die Inhalte durch vorangestellte Zahlencodes, die zwar für Computer einfach bearbeitbar sind, doch für Menschen ohne Spezialwissen nicht interpretierbar sind. Diesen Mangel behebt das neue Datenformat.

⁵ Rechtsinformationssystem des Bundes (<http://www.ris.bka.gv.at>).

Beiden gemeinsam ist jedoch die Einschränkung, dass die Strukturtiefe der Untergliederung nur 1 Ebene beträgt. Untergruppen und Zusammenhänge zwischen Dokumentteilen können nicht oder nur unter Erweiterung des Datenformats dargestellt werden. Dies ist vor allem bei stark gegliederten Daten wie Rechtsvorschriftenzitate (z.B. Paragraph, Absatz, Ziffer, Rechtsvorschriftenbezeichnung, Fassung, ...) und bei der Strukturierung von Entscheidungstexten (z.B. Tabellen, Listen, ...) problematisch.

Weiters ist das Datenformat kein Standarddatenformat, so dass schon alleine eine Änderung der Form, in der die Daten formatiert werden müssen, auch einen wesentlichen Aufwand der Anpassung der EDV-Systeme verursacht.⁶

2.2. Der Saarbrücker Standard für Gerichtsentscheidungen

Diese Einschränkungen proprietärer Datenformate sollen mit XML überwunden werden. Die einheitliche Struktur von Datenformaten, die XML definiert, erleichtert den Austausch der Daten und reduziert den Aufwand bei Änderungen erheblich. Es ist nicht mehr die Form, in der die Daten abgelegt sind, sondern die inhaltliche Struktur der Daten, die analysiert werden muss.

Systeme, die XML-konforme Daten bearbeiten können, werden mit geringem Aufwand auf geänderte Datenformate umgestellt, wenn diese wieder XML-konform sind.

Um auch diesen Aufwand zu reduzieren und einen noch einfacheren Austausch von Entscheidungstexten zu ermöglichen, wurde am 9. EDV-Gerichtstag in Saarbrücken (21.–22.9.2000) der „Saarbrücker Standard für Gerichtsentscheidungen“⁷ beschlossen. Es handelt sich dabei um eine

⁶ Im VwGH Wien wurde ein XML-basiertes Exportsystem realisiert, das eine Ausgabe in beliebigen Datenformaten ermöglicht (vgl.3.3. Export). Dies ermöglicht eine einfache Anpassung an geänderte Datenformate.

⁷ Der Standard ist zu finden auf dem Server des EDV-Gerichtstags (<http://edvgt.jura.uni-sb.de/>) und auf der Homepage des Österreichischen Forschungsinstituts für Rechtsinformatik (<http://www.rechtsinformatik.at>). Dort ist auch eine kommentierte Fassung des Saarbrücker Standards verfügbar, in der die Details der DTD erläutert werden und eine Musterentscheidung als standardkonformes XML-Dokument dargestellt wird.

XML-DTD⁸, die eine erweiterbare Grundstruktur für den Austausch von Entscheidungsdokumenten festlegt und erweitert werden kann. Der Standard soll eine einfache inhaltliche Erschließung von Gerichtsurteilen ermöglichen und Inkompatibilitäten beim Austausch von Daten vermeiden helfen.

Der Schwerpunkt des Standards liegt in der vorliegenden Fassung auf der Strukturierung und Definition von Metadaten zum Entscheidungstext, nicht jedoch auf der Strukturierung des Entscheidungstextes selbst.

3. Judikaturdokumentation im VwGH Wien

Im VwGH Wien wurde ein EDV-System realisiert, das die meistens getrennten Bereiche Geschäftsstellenverwaltung und der Judikaturdokumentation zusammenführt und eine einheitliche Verwaltung und Bearbeitung von Akten und den damit verbundenen Entscheidungsdokumenten ermöglicht.



Für die Judikaturdokumentation bedeutet dies, dass durch die Koppelung von Entscheidungstexten mit den Daten der Geschäftsstelle formale Daten der Akten, die auch in der Judikaturdokumentation enthalten sind, nicht mehr (fehlerträchtig) manuell erfasst werden müssen und auch immer im Dokument mit dem aktuellen Stand dargestellt werden.

Die Judikaturdokumentation selbst kann auf die inhaltliche Bearbeitung der Entscheidungen beschränkt werden. Die Verfügbarkeit des gesamten Bestands an Entscheidungsdaten auch für die Dokumentation ver-

⁸ Die DTD wurde von *Dipl.-Ing. Mag. Felix Gantner* (infolex) und *Mag. Peter Ebenhoch* (epico Informationssysteme GmbH) vorbereitet, zusammen mit *Dr. Kraft* (C.H.Beck-Verlag) und *Hr. Erbguth* (Gauss Interprise) sowie *Hr. Murk Müller* verfeinert.

bessert diese inhaltlich – z.B. durch die einfache und automatisierte Erstellung von Verweisen – wesentlich.

Geschäftsstellendaten und Metadaten von Entscheidungen werden in einem relationalen Datenbanksystem verwaltet. Die Entscheidungstexte sind als XML-Dokumente im System verfügbar.

In der Judikaturdokumentation sind 3 große Arbeits- und Systembereiche zu unterscheiden:

- Input: Dateneinbringung und Erzeugen der Metadaten
- Recherche: Recherchemöglichkeiten für Juristen in der Entscheidungsdokumentation und die Aufbereitung der Dokumente für die Anzeige
- Export: Automatisierte Aufbereitung der Daten für Dritte, z.B. das RIS

3.1. Input

Der Input-Bereich des Systems umfasst die Zuordnung eines Entscheidungstextes zu einem im System vorhandenen Akt, die Übernahme des Entscheidungstextes, die Bearbeitung des übernommenen XML-Dokuments und deren inhaltliche Auswertung.

3.1.1. Übernahme der Entscheidung

Am Anfang der Dokumentation eines Entscheidungstextes steht die Umwandlung des Textes in ein XML-Dokument, seine Übernahme in das System und die Zuordnung zu einem Akt der Geschäftsstelle.

Zur Zeit werden die Entscheidungstexte mit einem Textverarbeitungsprogramm (MS Word) erstellt, wobei die Grobstruktur des Dokuments mit Formatvorlagen gekennzeichnet wird.

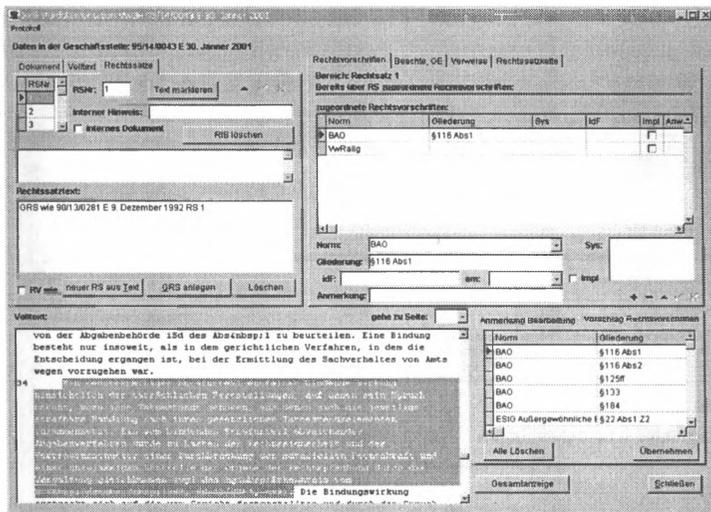
Um Probleme mit Versionen der Software, die sich laufend ändern, und damit verbundenen Inkompatibilitäten der Daten und der eingesetzten Makro-Programme zu vermeiden oder zumindest so gering wie möglich zu halten, wurde der Weg gewählt, das Dokument gleich am Anfang der Bearbeitung aus dem proprietären Textverarbeitungsformat in ein Standard-XML-Dokument umzuwandeln. Dies geschieht unter Verwendung von Makros, die aus Formatvorlagen XML-Auszeichnungen generieren.

Das so erhaltene XML-Dokument kann dann in die Judikaturdokumentation übernommen und juristisch ausgewertet werden. Jedes Dokument wird einem Akt zugeordnet, dessen formale Daten (Aktenzahl, Entscheidungsdatum, ...) immer verwendet werden.

3.1.2. Juristische Bearbeitung

Die Bearbeitungsoberfläche für die juristische Auswertung des Dokuments erfolgt in einer eigenen Teilanwendung des Systems. Diese ermöglicht:

- die Überarbeitung der XML-Entscheidungstexte (z.B. Bearbeitung von Tabellen)
- das Erzeugen von Rechtssätzen
- das Erzeugen von Rechtssatzketten
- die Zuordnung von Rechtsvorschriften zum Entscheidungs- oder zu einem Rechtssatztext
- die Zuordnung von frei formulierten Metadaten zum Entscheidungs- oder zu einem Rechtssatztext
- die Erzeugung von Verweisen zu Akten, Dokumenten oder Rechtssätzen. Die Verweise werden für die Anzeige mit vordefinierten Verweis- bzw. Rückverweistexten aufbereitet.



Nach der Freigabe des ausgewerteten Dokuments kann die Entscheidung in der internen Judikaturdokumentation gesucht und angezeigt werden. Das Dokument wird auch über die Exportfunktionen an Dritte weitergegeben.

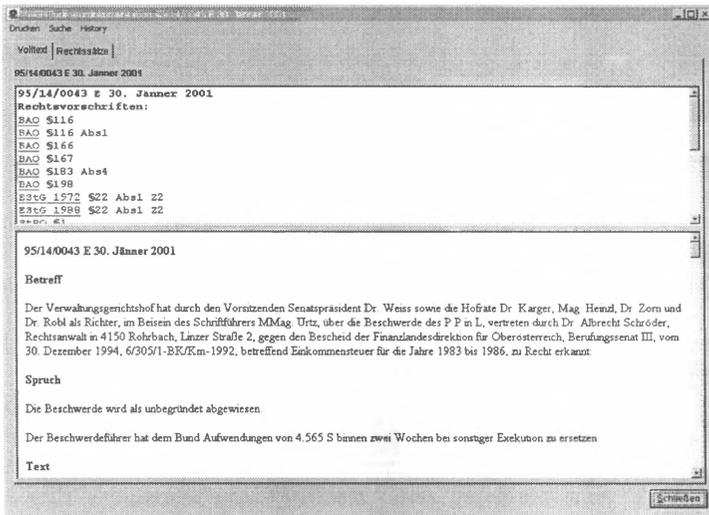
3.2. Recherche

In der internen Judikaturdokumentation können Richter und juristische Bearbeiter am VwGH direkt die Entscheidungstexte des VwGH suchen und auswerten. Dabei ist es möglich, sowohl nach Daten, die in der Geschäftsstelle verwaltet werden (Aktenzahl, Entscheidungsdatum, usw.), als auch nach den im Rahmen der juristischen Auswertung zugeordneten Daten (Rechtsvorschriften, Schlagworte, usw.) zu suchen.

Da die Dokumente als XML-Dokumente gespeichert und ausgewertet sind, ist natürlich auch eine Volltextsuche über das ganze Dokument oder nur in bestimmten Teilen möglich.

In den Metadaten werden Verweise auf andere Dokumente oder auf Rechtsvorschriften in Hyperlinks aufbereitet. Durch Anklicken der Links kann die entsprechende Information oder das gewünschte Dokument angezeigt werden.

Die Darstellung von Dokumenten und deren Metadaten erfolgt in folgender Weise:



Zugeordnete Rechtssätze und Rechtssatzketten werden in ähnlicher dargestellt.

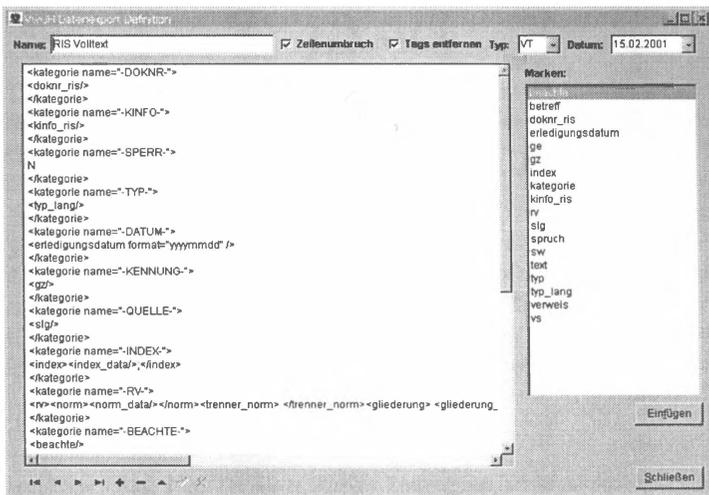
3.3. Export

Beim Export stellte sich das Problem, dass die im RIS verwendeten Datenformate (vgl. 2.1. Ein Beispiel: Datenformate für das RIS) während der Systementwicklung geändert wurden und u.U. auch in Zukunft geändert werden. Weiters müssen u.U. auch andere Empfänger von Daten (z.B. Verlage) in Zukunft mit Entscheidungsdaten des VwGH direkt beliefert werden.

Um nicht an ein bestimmtes Datenformat gebunden zu sein und um bei Änderungen von Datenformaten nicht mit dadurch notwendigen Entwicklungskosten belastet zu werden, wurde ein XML-basiertes Exportsystem entwickelt.

In einer Benutzeroberfläche kann ein XML-Musterdokument entwickelt werden, das als Vorlage für den Export dient. Die aus der Datenbank ausgelesenen Daten werden bei der Dokumentaufbereitung in XML-Marken eingesetzt.

Das folgende Bild zeigt das XML-Musterdokument für das in „2.1. Ein Beispiel: Datenformate für das RIS“ dargestellte neue Einbringungsformat für Entscheidungsdokumente:



Änderungen in der Vorlage führen unmittelbar zu geänderten Ausgaben. Es ist damit problemlos möglich, den Export an geänderte Datenformate anzupassen oder neue Datenformate zu definieren.

Die Datenaufbereitung und der Export erfolgen automatisiert und zeitgesteuert. Fertig aufbereitete Daten werden über das Internet an den Empfänger übertragen. Geänderte Dokumente (z.B. nach Korrekturen der juristischen Metadaten) werden erneut aufbereitet und übertragen.