

XML und RDF Dictionary

Austausch juristischer Informationen zwischen Computern

Die „Extensible Markup Language“ (XML) steht für die neue Generation von Websprachen. Das World Wide Web ist mit HTML (Hypertext Markup Language) groß geworden. HTML war jedoch nie für den Erfolg geschaffen, wodurch es die Welt erobert hat. HTML wurde entworfen, um Wissenschaftlern die Möglichkeit zu geben, ihre trockenen Ausführungen für den Online-Bereich mit etwas Layout aufzupeppen¹.

Das Internet ist bereits in den siebziger Jahren des vorigen Jahrhunderts entstanden. Damals diente das Internet vor allem der Online Kommunikation zwischen Universitäten². E-Mail war vom Anfang an auch dabei. Das World Wide Web wurde Anfang der neunziger Jahre von Tim Berners-Lee und anderen geplant und verwirklicht³. Da das WWW, neben E-Mail, die Hauptanwendung des Internets geworden ist, werden die zwei Begriffe oft als Synonym verwendet. Richtiger wäre es jedoch zu sagen, daß das WWW die Infrastruktur des Internets verwendet. Die Infrastruktur besteht aus der Vernetzung von vielen Servern und den Protokollen (wie TCP/IP), die die Kommunikation zwischen den Servern ermöglichen. Organisationen wie Internet Engineering Task Force ("IETF"⁴) und Internet Engineering Steering Group ("IESG"⁵) tragen für die Weiterentwicklung dieser Infrastruktur Sorge. Das World Wide Web Consortium („W3C“)⁶ ist die Organisation, die für die Weiterentwicklung des WWWs Sorge trägt. Im Grunde versucht das W3C im WWW die Initiative wieder zu gewinnen, nachdem HTML aus dem Ruder gelaufen ist. Das Hauptproblem ist, daß eine Zahl von großen Computerfirmen, jede auf ihre eigene Weise, die Fähigkeiten von HTML und vom Browser, der die HTML Dokumente liest, versucht zu erweitern. Die eigenhändigen Erweiterungen haben zu einem Wildwuchs geführt, der genau dem entgegenstrebt, was das WWW bieten will, nämlich eine offene Plattform für Informationsaustausch.

Im W3C sind alle namhaften Firmen der Computerbranche vertreten⁷. Im Forum des W3Cs strengt man sich tatsächlich gemeinsam an, die Weiterentwicklung des World Wide Webs in Bahnen zu lenken. Dazu ist man zurück zu den Wurzeln der HTML gegangen.

- Mag. Dr. Peter Jordan, Innsbruck, danke ich für seine Bemerkungen zum Entwurf dieses Aufsatzes

¹ für eine der ersten Versionen von HTML, siehe <http://www.w3.org/History/19921103-hypertext/hypertext/WWW/MarkUp/Tags.html>

² Für eine Geschichtsschreibung des Internets, siehe <http://www.isoc.org/internet/history/brief.html>.

³ Der Vorschlag an den CERN in Genf ist zu finden unter www.w3.org/History/19921103-hypertext/hypertext/WWW/Proposal.html

⁴ <http://www.ietf.org/>

⁵ <http://www.ietf.org/iesg.html>

⁶ <http://www.w3.org/>

⁷ <http://www.w3.org/Consortium/Member/List>

Die Wurzeln von HTML

Charles P. Goldfarb⁸ war in den sechziger Jahren Rechtsanwalt in Boston, USA, bis er für ein Projekt für Anwaltskanzleien bei IBM angestellt wurde. Extra für dieses Projekt wurde eine Sprache namens GML, **Generalised Markup Language** geschaffen, die eine IBM-interne Sprache blieb. In den siebziger Jahren entwickelte Charles Goldfarb, aus GML, SGML, die Standard Generalised Markup Language.⁹ SGML wurde also von einem Juristen eigentlich für Juristen geschaffen. SGML hat aber auf dem juristischen Markt kaum Einzug gehalten, mit Ausnahme von der Verwendung dieser Sprache bei größeren Dokumentverwaltung, wie in juristischen Verlagen.

SGML ist die Mutter von HTML. HTML könnte als eine Anwendung von SGML betrachtet werden. Um sich ein Bild dieser Verwandtschaft zu machen, bedarf es einer Erläuterung von „markup languages“ oder „Auszeichnungssprachen“.

Jedem sind inzwischen die etwas ungastfreundlichen <spitzen_Klammern> bekannt, die die HTML Dokumente auszeichnen. <body>, <h1><p>,
, <a href=<http://www.lexml.de>> und dergleichen heißen „Tags“. Sie sind Anweisungen, wie der eigentliche Text im Dokument, also die Wörter die nicht zwischen den <spitzen_Klammern> stehen, im Ergebnis gezeigt werden soll. Die Tags bestimmen das Layout, das Aussehen des Dokuments. Welche Tags können verwendet werden? Eine sogenannte Data Type Definition („DTD“) beschreibt alle Tags, die zur Auszeichnung eines Dokuments verwendet werden dürfen. Es gibt für SGML Tausende von DTDs. Jeder kann seine eigene DTD schreiben. HTML hatte am Anfang nur eine SGML DTD¹⁰. Browserhersteller haben eigene Tags hinzugefügt oder bestehende Tags verändert und damit eigene DTDs geschaffen. Ein und dasselbe Dokument sieht dadurch mit verschiedenen Browsern unterschiedlich aus, oder ist mit dem einen oder anderen Browser nur teilweise lesbar.

Das W3C hätte nun die DTD von HTML so erweitern können, daß alle Varianten wieder unter einen Hut kommen. Es wurde jedoch ein anderer Weg eingeschlagen. Bei HTML, als layout-orientierte Sprache, wird man immer etwas vermissen. Das Layout beschreibt auf indirekte Weise die Struktur eines Dokumentes. Mit einer Überschrift wird der Anfang und die Kurzbezeichnung eines neuen Abschnitts angezeigt. Mit einem Paragraph wird ein neues Argument angefangen, mit **Fettschrift** wird etwas wichtiges hervorgehoben. Wenn man zunächst ein Dokument in Gedanken gliedert und dann auf dieser indirekten Weise die Gliederung zum Ausdruck bringt, geht bereits ein Teil dieser Gliederung, zumindest für den Computer, verloren. Webentwickler (aber nicht nur diese Art von Geschöpfen) bevorzugen eine explizite Weise, um die Struktur eines Dokumentes zu beschreiben. In einem Urteil würden Markierungen wie zum Beispiel <Parteibezeichnung>, <Datum>, <Tenor>, <Begründung> eine explizite Gliederung bilden.

Bis jetzt war SGML die Sprache, die diese explizite Beschreibung der Struktur von Dokumenten ermöglichte. Das W3C hätte also SGML als Websprache einführen können. Da SGML nicht gerade benutzerfreundlich ist, hat man ein einfachere Syntax geschaffen und damit ein neuer Sprachstandard für das Web geschaffen: XML.^{11 12}

⁸ <http://www.sgmlsource.com/>

⁹ the story of Charles Goldfarb, von ihm selbst aufgezeichnet, ist unter www.sgmlsource.com/history/roots.htm zu lesen

¹⁰ siehe die bereits genannte frühe Version von HTML: <http://www.w3.org/History/19921103-hypertext/hypertext/WWW/MarkUp/Tags.html>

¹¹ Februar 1998 als Standard verabschiedet, letzte Version: <http://www.w3.org/TR/REC-xml>

Mit XML ist jeder frei, seine eigenen DTD's zu kreieren. Somit hat das W3C dem WWW einen größeren Spielraum gegeben und zugleich die Möglichkeit geschaffen, untereinander wieder „kompatibel“ zu werden. Die Kompatibilität kann dadurch erreicht werden, in dem jedes XML Dokument auf die DTD¹³ verweist, aufgrund derer es ausgezeichnet wurde. Es gibt somit keine unbekannt Tags mehr.¹⁴

Wenn die inhaltliche Markierung detailliert erfolgt, kann als zweiter Schritt an den Text zwischen den Markierungen, zwischen den Tags, eine bestimmte Layout-Anweisung verknüpft werden: Einspringen, Fett, unterstrichen usw. Die Layout-Anweisungen können in einem anderen Dokument enthalten sein, nämlich in einem sogenannten „Stylesheet“¹⁵

Und was ist dann bitte XHTML¹⁶? XHTML ist die neue, auf XML basierende, HTML mit erweiterbarer DTD, das entweder verwendet werden kann, wo sich die Mühe, Struktur und Layout strikt zu trennen, nicht lohnt, oder als Ergebnisdokument der Zusammenführung eines XML Dokumentes mit einem (XSL) Stylesheet.

Jetzt wollen wir uns nicht weiter mit Layout Angelegenheiten beschäftigen, sondern uns ganz und gar der Struktur von Dokumenten zuwenden. Wir werden dann auch sehen, welche anderen wunderbaren Möglichkeiten XML bietet.

Datenstrukturen

Eine DTD haben wir bis jetzt unter dem Aspekt einer Sammlung von möglichen Tags in einem Dokument betrachtet. Wenn man eine DTD macht, um den Inhalt von Dokumenten zu beschreiben, befasst man sich mit der Struktur von Dokumenten. So könnte man eine DTD machen um die Struktur des Dokumenttyps „Urteil“ zu beschreiben. Das wichtigste Dokument der Rechtspflege wurde für den deutschsprachigen Raum bereits in einer DTD gefasst und am 22. September 2000 während des EDV Gerichtstages unter dem Namen „Saarbrücker Standard“ abgesegnet. Die DTD mit Kommentar ist auf der Site <http://www.lexml.at/> zu finden. In dieser DTD sind Tags beschrieben die verwendet werden können um ein Urteil „auszuzeichnen“, um seine Struktur zu beschreiben. Ein Beispiel eines ausgezeichnetes Urteil ist ebenfalls auf <http://www.lexml.at/> zu finden .

Es werden zur Zeit vielerorts DTD's entwickelt, in der juristischen Welt, sowie in allen Branchen der Wirtschaft, Wissenschaft, Verwaltung, Technik. Diese DTD's werden verwendet um erst Tausende, später Millionen von Dokumente auszuzeichnen.

¹² Man könnte auch behaupten, SGML und XML seien überhaupt keine „Computersprachen“, sondern ein Syntax für die Schaffung von Sprachen. Insofern wäre das „L“ von Language etwas irreführend

¹³ ein XML Dokument ohne DTD ist auch möglich und heißt „wohlgeformt“, muß aber das Epitheton „gültig“ entbehren.

¹⁴ Dies gilt natürlich nur, insoweit das XML Dokument gültig ist, das heißt das Dokument muß den Syntaxregeln von XML und den Regeln der zugehörigen DTD entsprechen.

¹⁵ Auch für herkömmliche HTML Dokumente verwendet man bereits Stylesheets, insbesondere die Cascading Stylesheets („CSS“). Eigens für XML Dokumente wurde eine neue Stylesheet Sprache entwickelt, die Extensible Stylesheet Language („XSL“, <http://www.w3.org/Style/XSL/>). CSS kann man auch für XML Dokumente verwenden, schöpft die Möglichkeiten von XML jedoch nicht aus. Mit XML und XSL kann der im WWW lang gehegter Wunsch der Trennung von Inhalt und Aussehen erfüllt werden. XSL ist mit der XML Syntax geschrieben und ist somit eine Anwendung von XML.

¹⁶ Der Standard wurde am 26.1.2000 festgelegt: <http://www.w3.org/TR/xhtml1>

In der XML Sprache heißen Tags „Elemente“. Die Elemente stehen in einer bestimmten Reihenfolge und können ineinander verschachtelt sein:

```
<gericht>
  <gerichtstyp>
  Landgericht
  </gerichtstyp>
  <ort>
  Berlin
  </ort>
</gericht>
```

Die Elemente einer DTD weisen untereinander eine gewisse Struktur auf, sowohl bezüglich Reihenfolge als auch in hierarchischer Hinsicht (Subelement, sub-Subelement), die später in der Auszeichnung der Dokumente, in der Markierung mit Elementen, zum Ausdruck kommt.

Information und Daten

Es gibt viele Definitionen, um den Unterschied zwischen „Information“ und „Daten“ zu beschreiben. Wir wollen hier als Definition verwenden: Information ist etwas, was für den Mensch Bedeutung hat; Daten ist etwas, was für einen Computer Bedeutung hat. Für den Menschen beinhaltet der eigentliche Text, also nicht die Elemente, die Information. Für den Computer sind Daten die Elemente (Tags) zusammen mit dem eigentlichen Text.

Wiederverwerten und vergleichen von Dokumententeilen

Ein wichtiger Zweck der Auszeichnung ist, daß man Teile eines Dokumentes, die Daten, auf unterschiedlichster Weise wiederverwerten zu können. Die Teile können ganze Textpassagen sein, aber auch ein Wort, ein Datum, eine Zahl. XML hilft Daten wiederzuverwerten. XML ermöglicht eine Kreislaufwirtschaft für Daten.

Wenn man für zwei Dokumente dieselbe DTD für die Auszeichnung verwendet hat, kann ein Computer ohne weiteres Textteile, die Daten, vergleichen. Zum Beispiel kann festgestellt werden, daß Urteil A später als Urteil B verkündet wurde, die Klage aber früher eingegangen ist. Oder, daß in Urteil A 23% der Klagesumme und in Urteil B 98% zugewiesen wurde. Sinnlose Beispiele, aber sie sollen zeigen, daß die Möglichkeiten weitreichend sind.

Unterschiedliche DTDs

Das Vergleichen von Daten von Dokumenten, die mit derselben DTD ausgezeichnet sind, geht also glatt. Aber wie ist es mit Dokumenten, die mit unterschiedlichen DTDs ausgezeichnet wurden, zum Beispiel ein Schriftsatz und ein Urteil? Wenn die eine DTD das Element <Gericht> enthält und die andere DTD <gericht>, dann sind die Elemente für den Computer so unterschiedlich wie „Gericht“ und „Dessert“, weil XML zwischen Groß- und Kleinschreibung unterscheidet. Ein anderes Beispiel: in beiden DTDs kann das Element <ort> definiert sein. Auf erste Sicht kein Unterschied, aber in der einen taucht <ort> als Subelement von <gericht> und in der anderen DTD als Subelement von, zum Beispiel, <Beklagte> auf. Soll der Computer den Inhalt von <ort> in dem einen Dokument als vergleichbar mit dem Inhalt von <ort> in dem anderen Dokument betrachten?

Könnte man nicht versuchen, sich auf eine einheitliche Schreibweise und Verschachtelungsstruktur zu einigen, bevor man weitere DTDs schreibt? Schön wäre es! In den Vereinigten Staaten ist Legalxml¹⁷ seit über drei Jahren dabei, DTDs für die common law Welt zu kreieren. Es zeigt sich als nahezu unmöglich, die Vielfalt der juristischen Realität mit einheitlichen Elementnamen und –strukturen einzufangen. Und selbst wenn man es schaffen würde, wie vergleicht man Daten aus anderen Dokumenten, die teilweise Juristisches beinhalten, aber doch mit einer anderen DTD ausgezeichnet wurden, weil das Dokument an sich nicht als ein juristisches Dokument angesehen wird (zum Beispiel ein Bankauszug mit einigen AGB-Klauseln)?

Nein, Vereinheitlichung auf der vollen Breite führt nicht zum Erfolg. Vereinheitlichung ist nicht möglich, nicht gewünscht und, so stellt sich bei genauerer Betrachtung heraus, auch gar nicht für die Verwendung von XML erforderlich. Es ist nicht die Vereinheitlichung, die man haben will, es ist die Kompatibilität, die man braucht. Man will erreichen, daß Computer einander so verstehen, daß sie Daten austauschen und richtig zuordnen können.

Die Lösung liegt eigentlich auf der Hand: Schaffe eine Ebene über den DTDs, worüber die verschiedene DTDs sich mit einander verständigen können. Auf der höheren Ebene könnte zum Beispiel ein Begriff „Spruchkörper“ definiert sein. Die eine DTD teilt mit: das was bei mir <gericht> heißt, ist auf der höheren Ebene mit „Spruchkörper“ gleichzustellen. Die andere DTD teilt mit, daß <Gericht> mit „Spruchkörper“ gleichzustellen ist, und schon können die zwei DTD´s mit einander kommunizieren. Die Komplexität eines Vergleichs von Verschachtelungsstrukturen ist natürlich größer, kann aber auf ähnliche Weise erfolgen.

Es ist sogar möglich, eine jurisdiktions- und sprachenübergreifende Ebene zu schaffen, die einen internationalen Austausch von juristischen Daten zwischen Computern ermöglicht.

Wozu Datenstrukturen für Computer?

Nun ist eine Frage weitgehend offen geblieben: warum sollten wir uns für die Schaffung von juristischen Datenstrukturen Mühe geben und was nutzt uns Datenaustausch zwischen Computern? Zwei Gründe wurden bereits genannt:

1. Kreislaufwirtschaft für Daten
2. das Vergleichen des Inhalts, der Daten, von zwei Dokumenten durch den Computer

Hinzuzufügen wäre:

3. Ausgabe desselben Inhalts in verschiedenen Formen (Layoutwechsel)
4. Suchen und Finden von Dokumenten

Ad 1. Kreislaufwirtschaft

Ein Richter wird wenig Gefallen daran finden, seine Urteile mit dem Saarbrücker Standard auszuzeichnen, wenn er nicht zumindest Teile von anderen XML Dokumenten für sein Urteil verwendet. Diese anderen Dokumente können andere Urteile sein, interne Textbausteine, aber auch Teile der Schriftsätze der Parteien. Ein Anwalt wird wenig Gefallen daran finden, seine Schriftsätze in XML Form einzureichen, wenn er nicht zumindest teilweise von anderen XML

¹⁷ <http://www.legalxml.org>

Dokumenten Gebrauch machen kann, die ihm als Grundlage für seinen Schriftsatz dienen: Unterlagen des Mandanten, Handelsregisterdaten, Rechtsprechung. So kann man die Kette in viele Richtungen verfolgen. Deutlich ist, daß man mit dem Huhn und Ei Problem konfrontiert wird, das nur durch die Entstehung einer „kritischen Masse“ zu bewältigen ist. Erst wenn man einen klaren Vorteil von XML hat, wird man es für die Erstellung seiner Dokumente verwenden. Zudem braucht man ein Programm, das die Auszeichnung von Dokumenten mit XML so leicht macht, wie heute das Schreiben mit einem Textverarbeitungsprogramm ist. Es gibt Bestrebungen ein „office tool“ für die juristische Welt open source zu entwickeln; ein Programm, das XML versteht und produziert.¹⁸

Ad 2. Vergleichen von Daten

Die sinnlosen Beispiele über den Prozentsatz der zugesprochenen Klagesummen können durch sinnvollere ersetzt werden. In jeder größeren Organisation, in der die Papierproduktion nicht mehr von einer Person überblickt werden kann, will man früher oder später Statistiken über die produzierte Arbeit haben. Mit XML kann man Inhalte von Dokumenten, Daten, automatisiert vergleichen, und zwar auf einer Weise, die bisher kaum möglich war.

Ad 3. Layoutwechsel

Vor allem Verlage schwärmen Zeit davon, daß XML durch die Trennung von Struktur und Layout die verschiedensten Herausgabeformen auf Basis eines Basisdokuments ermöglicht. Zwar bot SGML diese Möglichkeit bereits auch, aber nicht in dem Maße und mit der Bequemlichkeit wie es mit XML möglich ist. Man kann durch Verwendung von unterschiedlichen Stylesheets zum Beispiel den Inhalt automatisch kürzen, um WAP Seiten zu veröffentlichen, etwas länger machen für (X)HTML Seiten und noch länger für Aufsätze oder Bücher. Auch kann das XML-Dokument in ein anderes Daten- und Textformat umgewandelt werden, zum Beispiel RTF oder PDF. Zudem ermöglicht XML es, auf einfache Weise in einem Dokument automatisch Teile eines anderen Dokumentes einzufügen, das ganz woanders auf der Welt gespeichert sein kann.

Ad 4. Suchen und Finden

Die Suchmöglichkeiten im Internet zu verbessern, liegt mir persönlich am meisten am Herzen. Finden statt Surfen! XML wird von einer Reihe Begleitsprachen unterstützt. Die Sprachen Xlink, Xpath und Xpointer bieten die Möglichkeit, sehr gezielt nach Dokumenten mit genau umschriebenen Inhalten, Daten, zu suchen. Man ist dann nicht mehr abhängig von Suchmaschinen, nein, man kann seine Anfrage an Tausende Server gleichzeitig abschicken, die dann, jeder für sich, die gespeicherte Dokumente mit der Abfrage vergleichen und das Ergebnis zurückgeben. Das Internet kommt damit einer Datenbank gleich.

Schöne Zukunftsmusik, die Zukunft bleiben wird, wenn wir uns nicht zuerst an die Basisarbeit setzen. Das heißt in concreto: die „höhere Ebene“ muß geschaffen werden. Wie? Durch die Erschaffung eines RDF dictionary!

RDF Dictionary

RDF („Resource Description Framework“) ist in der XML Syntax geschrieben. RDF ist eine Anwendung von XML und ein vom W3C verabschiedeter Standard.¹⁹

¹⁸ siehe <http://www.lexml.de/> und <http://opt.lexml.de>

¹⁹ <http://www.w3.org/TR/REC-rdf-syntax>

RDF in seiner einfachsten Form ist eine Sprache, womit Metadaten von „Webresources“ beschrieben werden können. „Webresource“ ist, vereinfacht gesagt, alles was mittels einer Hyperlink oder E-Mail Adresse lokalisiert werden kann²⁰. Metadaten sind in Rahmen von RDF Daten, die nicht unbedingt in einem HTML oder XML Dokument stehen, sondern eher die Information über dieses Dokument. Ein gutes Beispiel für das Urteil wäre: Berufung eingelegt ja oder nein. Diese Information steht natürlich nicht im Dokument selbst, sondern ist eine Information, die später, sozusagen am Rande, hinzugefügt wurde. Bei Metadaten konzentrieren wir uns also auf Information am Rande des Dokuments. Das heißt nicht, daß diese Information nicht auch im Dokument selbst vorkommen kann, wie das Datum oder der Verfasser. Information, die man immer wieder braucht um ein Dokument zu finden, weiterzugeben oder zu beschreiben, wird extrahiert und am Rande nochmals erwähnt.

Das Geniale von RDF ist, daß die Beschreibung einer Webresource in Prinzip durch das Heranziehen von anderen Webresources erfolgt, was ein geschlossenes System ermöglicht. Murk Muller (mm@mmrecht.com) ist der Autor (<http://purl.org/DC/documents/dces/#Creator>) dieses Aufsatzes (<http://xml-rdf.lexml.de>). Der Term „Creator“ (Autor) stammt aus der Sammlung von Metadatenbeschreibungen „Dublin Core Metadata“. Diese Sammlung findet man unter der angegebenen Hyperlink. Das Zeichen „#“ wird verwendet, um auf eine bestimmte Stelle in einem HTML Dokument zu verweisen, in diesem Fall auf die Stelle, wo der Term „Creator“ genauer beschrieben steht. Dieser Aufsatz ist auch im Internet verfügbar, nämlich unter der angegebenen Hyperlinkadresse. Die Person Murk Muller ist mit ihrer E-Mail Adresse eindeutig im Internet zu lokalisieren.

Drei Webresources, die man nennt: Objekt, Prädikat und Subjekt. Das Subjekt ist die Webresource, die beschrieben werden soll. Das Objekt ist die Webresource, die mit dem Subjekt in Bezug gesetzt werden soll. Das Prädikat stellt diesen Bezug her.

Objekt ist das	Prädikat von	Subjekt
Murk Muller ist	der Autor	dieses Aufsatzes
mm@mmrecht.com	http://purl.org/DC/documents/dces/#Creator	http://xml-rdf.lexml.de

In RDF sieht das dann so aus:

```
<Description about = "http://xml-rdf.lexml.de" >
  <Creator xmlns:DC=( "http://purl.org/DC/documents/dces/" )>
    mm@mmrecht.com21
  </Creator>
</Description>22
```

²⁰ das Konzept von URI (Uniform Resource Identifier) lasse ich hier außer Betracht.

²¹ „Murk Muller“ wäre auch zulässig. Wenn keine Hyperlink oder E-Mail Adresse für das Subjekt verwendet wird, heißt es ein „literal“

²² Das vollständige XML/RDF Dokument sieht dann so aus:

```
<? xml version=„1.0“? >
<rdf:RDF
xmlns:rdf=“http://www.w3.org/1999/02/22-rdf-syntax-ns#“
xmlns:dc=“http://purl.org/DC/documents/dces/“>

  <rdf:Description about = “http://xml-rdf.lexml.de“>
    <dc:Creator>
      mm@mmrecht.com
    </dc:Creator>
```

XMLNS:DC heißt "XML Namespace, Dublin Core Metadata". Anders gesagt: „Creator“ ist so zu verstehen, wie es in <http://purl.org/DC/documents/dces/> beschrieben worden ist. „Namespace“, oder Namensraum, gibt somit den Kontext an, wie ein gewisser Ausdruck zu interpretieren ist.

RDF Schema ist „Candidate Standard“²³ des W3Cs, der die Beschreibungsmöglichkeiten von RDF erweitert, mit einem System, das inspiriert wurde durch die Prädikatenlogik, semantische Netze, Datenbanktheorie und mehr im Allgemeinen von der Erfahrung, die man mit dem Aufbau von Wissensrepräsentationssystemen gesammelt hat. Mit RDF Schema kann man, genauso wie mit einer DTD, Datenstrukturen beschreiben. Anders als eine DTD, die von der SGML Sprache geerbt wurde und nur die Beschreibung einer hierarchischen (Baum-)Struktur zulässt, ist RDF Schema objektorientiert und ermöglicht sogar Mehrfachvererbung.²⁴ Durch die Verwendung von Namensräumen (namespaces) können verschiedene Datenstrukturen in eine RDF Datenstruktur, RDF Schema, zusammengeführt werden.

Der Namensraum und RDF Schema „Dublin Core Metadata“ haben wir bereits kennengelernt. Stellen Sie sich vor, wir würden ein RDF Schema „recht“ erstellen, darin den Begriff „Spruchkoerper“ und „Instanz“ beschreiben und das Schema unter <http://namespace.lexml.de> veröffentlichen. Damit haben wir die Möglichkeit geschaffen, daß jeder den Elementnamen „Spruchkoerper“, mit dem Ausdruck *xmlns: recht = „<http://namespace.lexml.de>“*, den Elementnamen *<recht:Spruchkoerper>* auf unmissverständliche Weise verwenden kann.

Wenn jemand es bevorzugt in seiner DTD als Elementname *<Gericht>* zu verwenden, braucht er lediglich in einem RDF Schema festzulegen, daß *<Gericht>* in seiner DTD ein Unterbegriff von *<recht:Spruchkoerper>* ist. Wenn *<Gericht>* in einer DTD zudem als Unterbegriff von „Instanz“ zu verstehen ist, dann kann dieser Bezug mit *<recht:Instanz>* hergestellt werden. Mit diesen Beispielen bekommt man ein Vorgeschmack davon, wie mit RDF Schema Verschachtelungsstrukturen mit einander in Bezug gesetzt werden können.

RDF Schema ist die angewiesene Sprache, um die höhere Ebene zu schaffen. Ein RDF Schema kann die verschiedenen Datenstrukturen integrieren, als Schnittstelle fungieren und ein Orientierungspunkt bei der Erschaffung von neuen Datenstrukturen sein. Eine derartige Verwendung von RDF wird vom W3C gefördert und hat einen Namen: RDF Vocabulary oder RDF Dictionary²⁵.

Weiteres über RDF und RDF Schema kann man auf den Sites <http://www.lexml.de> und <http://www.lexml.at> finden. Ein Experiment für die Schaffung eines RDF Dictionary auf internationaler Ebene ist auf <http://www.lexml.de>, unter „RDF, RDF Schema, Metadata, RDF Dictionary“ zu finden.

</rdf:Description>

</rdf:RDF>

²³ <http://www.w3.org/TR/2000/CR-rdf-schema-20000327/>

²⁴ In diesem Zusammenhang ist auch der „Candidate Standard“ XML Schema zu erwähnen (<http://www.w3.org/XML/Schema>). Obwohl XML Schema schon einiger Zeit Anwärter ist die DTD abzulösen, hat es noch immer nicht den Status von „Recommended Standard“ erreicht. März 2001 hat es den Status von „Proposed Standard“ errungen. XML Schema ist auch objektorientierter als die DTD, lässt aber keine Mehrfachvererbung zu.

²⁵ Offizieller Name ist „Vocabulary“. Wir halten jedoch „RDF Dictionary“ für aussagekräftiger

Zur Zeit wird die Infrastruktur für das Schreiben von RDF Dictionaries für die juristische Welt auf nationaler und internationaler Ebene aufgebaut. Die Infrastruktur besteht aus einer Mailingliste und einem Programm aus der Open Source Software Entwicklung, das die Bearbeitung eines Dokumentes durch viele Personen gleichzeitig ermöglicht (CVS, concurrent Versions System). Sollten Sie Interesse haben, mitzuarbeiten, schreiben Sie eine E-Mail an mm@mmrecht.com oder schauen Sie auf <http://www.lexml.de> nach dem letzten Stand.

Gefragt sind nicht so sehr technische Personen. Es geht um juristische Systematik. Eine neue Weise der Betrachtung des juristischen Systems auf eine Weise, daß die neue Betrachtung verwendet werden kann, die Möglichkeiten der Kommunikation zwischen Juristen durch den Einsatz von Computern um einen Quantumssprung zu verbessern. Auf internationaler Ebene ist zum Beispiel die rechtsvergleichende Wissenschaft gefragt. Es wird andere Bereiche geben, wo man sich bis jetzt nicht bewußt war, daß ihr Wissen beitragen kann, um von den Möglichkeiten, welche die internationale Vernetzung bietet, Gebrauch zu machen. Wenn Sie beim Lesen dieser Seiten ein Aha-Erlebnis haben, machen Sie mit!

Murk Muller