

Neue Wege für HyperText-Anwendungen auf CD-ROM

Ulrich Müller

Während bisher die Probleme der HyperText-Anwendung vor dem Hintergrund exakt strukturierter juristischer Informationen betrachtet wurden, soll es hier um erweiterte Anwendungsmöglichkeiten gehen und zugleich um praktische Lösungsansätze zur Behebung bestehender Mängel.

Die dargestellten Methoden sind selbstverständlich auch für juristische Informationen geeignet, beschränken sich aber nicht darauf.

Einleitend sollen die Besonderheiten der Arbeitsweise von HyperText, wie sie derzeit bereits praktisch als CD-ROM-Lösungen auf dem Markt sind, in aller Kürze und stichpunktartig charakterisiert werden:

- Textsprung vom Inhaltsverzeichnis an die relevante Stelle
- Querverweisimplementierung auf andere Textelemente und Bilder
- die Möglichkeit von Suchoperationen in Textabschnitten
- Anbringen von Markierungen und Bemerkungen
- vielfältige Ausgabevarianten für Textbereiche.

Erscheinungsformen von Hypertext

„Individuelles Textmanövrieren“

„Law is hypertext by nature.“

Bisher:
Ungenügende Informationsaufbereitung

„Text-Tagging“ mit SGML
(= Standard Generalized Markup Language)

Ulrich Müller studierte Informations-elektronik an der Hochschule für Verkehrswesen in Dresden. Danach arbeitete er mehrere Jahre als Patentassessor in einem großen Industrieunternehmen in Jena. Gegenwärtig beschäftigt er sich als Mitarbeiter der ER & P GmbH, Jena, mit praktischen Fragen der Informationsaufbereitung von Volltext auf CD-ROM.

HyperText vollzieht damit eine Arbeitsweise nach, wie sie für die Auswertung nicht nur juristischer, sondern wissenschaftlicher Literatur im allgemeinen seit langem manuell gehandhabt wird. Die Rechnerunterstützung schafft für den Nutzer den gewünschten Service einer Automatisierung dieses Prozesses. Die riesige Textmenge im Hintergrund bleibt meist unbemerkt, weil bei einfacher Textnavigation über das Inhaltsverzeichnis nur die relevanten Textelemente in individueller Zusammenstellung aneinandergereiht dargeboten werden. Die Querverweise sind dabei ein wichtiges Zusatzelement dieses individuellen Textmanövrierens.

Bei Gesetzestexten und juristischen Informationen sind diese Verkettungsstrukturen bereits im Ursprungstext enthalten, so daß es nicht schwer fällt, eine entsprechende CD-ROM-Aufbereitung vorzunehmen: Die Informationen sind streng hierarchisch gegliedert. Die Querverweise besitzen eindeutige Anfangs- und Zielmarkierungen.

Tatsächlich konzentrieren sich die bisherigen HyperText-Anwendungen vorwiegend auf juristische Informationen, die sich diesen Grundaufbau zunutze machen und rechnergestützt nachvollziehen.

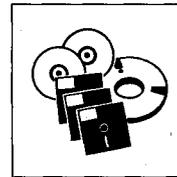
Die Zurückhaltung bei anderen Informationsangeboten dürfte demzufolge neben der normalen Anlaufphase für eine neue Technologie vor allem an der ungenügenden informationstechnologischen Aufbereitung der Informationspakete liegen.

Das heißt, so wie im Bereich der Spruchpraxis zwischen einem bibliographischen Informationsblock, den Leitsätzen und dem Textteil sachlich gegliedert werden kann, sind solche prinzipiellen Einteilungen zwingend notwendig, um einen akzeptablen Einstieg in ein Gesamtpaket zu finden, bevor eine hypertexttypische Strukturierung mit vielen parallelen Textelementen (Knoten) und Kanten als Verknüpfungswege vorgenommen wird. Üblicherweise ist das die interne Grundlage, um das schnelle Fortbewegen zwischen diesen Textelementen zu ermöglichen.

Das Ziel besteht aber nicht darin, diese Verknüpfungsstruktur zu optimieren oder eine Wichtung der Verknüpfungswege vorzunehmen, sondern in der Einführung zusätzlicher paralleler und hierarchischer Ebenen, um das Überangebot in der jetzigen flachehaften Verknüpfungsstruktur aufzubrechen.

Eine HyperText-Erschließung von Volltextinformationen erfordert neben der SGML-Strukturierung vor allem auch eine informationstechnologische Aufbereitung – einen speziell angepaßten Informationsaufbau.

„Mit SGML-Strukturierung (Standard Generalized Markup Language) ist die Belegung des Textes mit Strukturkennzeichnungsmarkierungen – sogenannten Tags – gemeint. Diese Markierungs-Tags sind notwendig, um eine interne Orientierungshilfe für den Text zu schaffen. Sie kennzeichnen durch Zusatzinformationen, die für den Nutzer nicht sichtbar



sind, Textabschnitte, Überschriften mit ihrer Hierarchiezuordnung, Fußnoten, Hervorhebungen usw. und ebenso die Absprungstellen und Zielpositionen der Querverweise. Der ursprüngliche Zweck dieser Markierungen ist es, Dokumente nur aufgrund ihres logischen Aufbaus zu strukturieren und Layout-Informationen erst später einzufügen. Das vereinfacht die Übergabekette vom Autor zum Verlag und dann weiter zur Druckerei ganz wesentlich.

Für HyperText-Anwendungen bietet es sich natürlich an, Vorhandenes zu übernehmen und darauf aufzusetzen.

```
<DOC>
<TITLE>.....</TITLE>
<CHAPTER1><FOREWORD>.....
.....</FOREWORD>
<ABSTRACT>.....
.....</ABSTRACT>
<BODY>.....
.....</BODY></CHAPTER1>
<TITLE>.....</TITLE>
<CHAPTER2><BODY>.....
.....</BODY></CHAPTER2>
</DOC>
```

*Beispiel eines
SGML-strukturierten Textes*

In diesem Beitrag sollen erste praktische Hinweise zur Problematik der Erschließung komplexer Informationspakete für eine HyperText-Anwendung gegeben werden.

Unter Informationslogistik sollen hier die Aufbereitung und die Darbietung eines Informationspaketes für eine optimale Auswertung verstanden werden.

Informationslogistik

1. Die richtige Strukturierung des Gesamttextes

Zwei Gesichtspunkte sind dafür wichtig:

Die Strukturierung soll eindeutig kennzeichnende Hierarchie-Tags aufweisen, die sich nicht wiederholen, um Querverweise an dieser Gliederung orientieren zu können.

Eindeutige „Hierarchie-Tags“

Die HyperText-Software nutzt mitunter die verbale Zieldefinition der Querverweise, um dem Nutzer eine Übersicht aller existierenden Verweisinformationen vorzulegen.

Die identische Markierung von Textbereichen für eine quasi feldweise Suchanfrage.

Gruppen- und Einteilungs-Tags

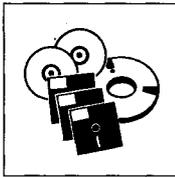
Das betrifft Bereiche mit vergleichbarem Informationsgehalt oder bibliographischen Angaben, die über den Gesamttext hinweg identisch markiert werden. Es kann sich dabei auch um Gruppen-Tags handeln, die andere Einteilungs-Tags umschließen und somit eine zweite unabhängige Zuordnung neben dem Hierarchieaufbau des Dokuments schaffen.

Beides kann bis zu einer Doppelstrukturierung des Gesamttextes nach eben diesen beiden Gesichtspunkten getrieben werden. Damit wird einerseits die hierarchische Ordnung beibehalten und andererseits fast perfekt der Einstieg über einen strukturierten Datenbankaufbau mit indextierten Stichworten angestrebt.

Allein das verdeutlicht schon, daß es bei weitem nicht ausreicht, gemeinhin eine SGML-Strukturierung für den Gesamttext vorzunehmen, um für eine CD-ROM-Produktion vorbereitet zu sein.

Ähnlich verhält es sich mit Querverweisen. Derzeit müssen alle Querverweisstellen gesondert gekennzeichnet werden. Bei eindeutigen Zielpositionen gibt es heute Werkzeuge, um diese Arbeiten zu automatisieren. Schwieriger wird es, wenn sinnverwandte Textbereiche aus unterschiedlichen Informationsblöcken über Kreuz verknüpft werden sollen, so daß von jeder Textstelle aus die anderen Texte erreichbar sind. Derartige Wanderungen innerhalb einer Ebene sind für den Nutzer sicher wünschenswert (schon weil die Suche nach Stichworten auch nicht alle relevanten Informationen ermitteln kann). Technisch läßt sich das nur sehr aufwendig realisieren, weil allein die Fülle der Verknüpfungsinformationen dann kaum noch beherrschbar wird.

Querverweis-Markierung



2. Sternverknüpfungen

Anders sieht es aus, wenn man anstelle der netzartigen Verknüpfung der Textelemente zu einer Sternverknüpfung übergeht (Abb. 1).

Vorteile der Sternverknüpfung

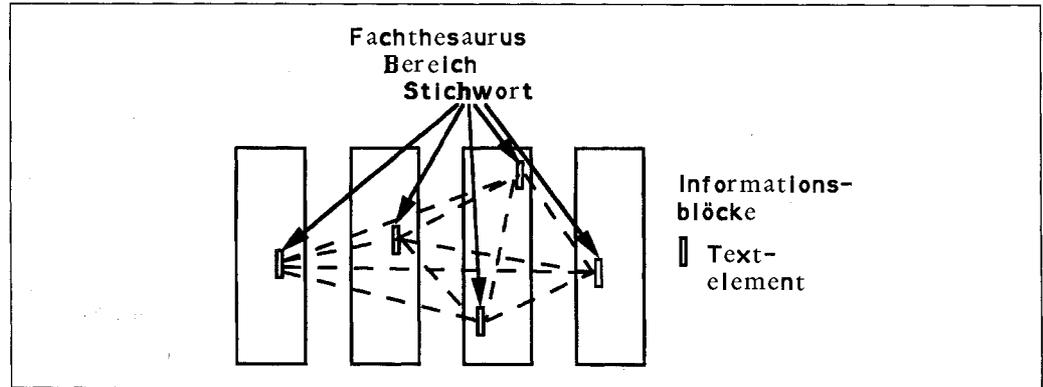


Abb. 1: Sternverknüpfung von Textelementen

Das hat mehrere Vorteile:

- Die Anzahl der Verknüpfungen wird drastisch reduziert.
- Die Textelemente können auch mehrfach zugeordnet werden.
- Die Textstellen selbst sind durch ihre sinngemäße hierarchische Zuordnung leichter auffindbar.
- Die Trefferwahrscheinlichkeit ist größer, weil Synonymworte sofort erkannt werden.

Aufwendig: Fachthesaurus-Erstellung

Nachteilig ist zweifellos, daß ein zusätzlicher manueller Aufwand notwendig ist, um den Fachthesaurus zu erstellen. Vieles wird aber allein dadurch wettgemacht, daß die Informationsflut weit übersichtlicher beherrscht werden kann:

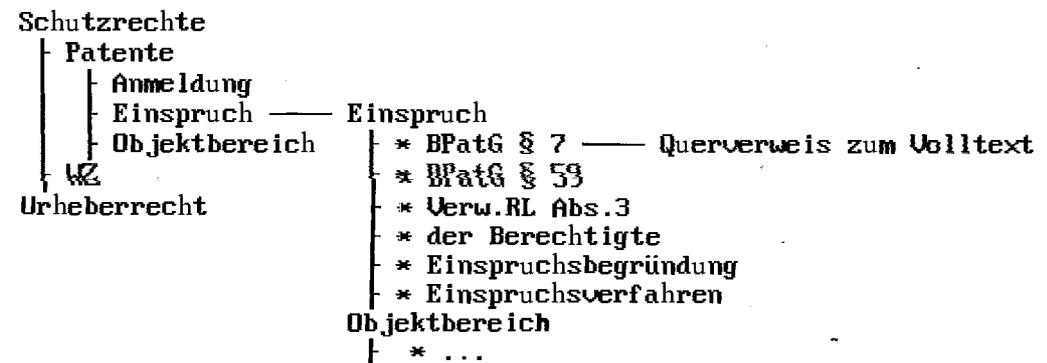
- Die Querverweise können nach einer Kurzkennzeichnung automatisiert angebracht werden.
- Dem Nutzer wird eine Übersicht über die relevanten Sprungstellen vorgelegt. Soweit der Textbereich ohnehin eine Überschrift trägt, kann auch diese Eintragung in die Übersicht automatisiert ablaufen.

Dreistufiger Textzugang

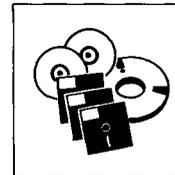
Letztlich kann diese Herangehensweise so ausgebaut werden, daß der Nutzer in drei Etappen an den Text herangeführt wird:

1. Einstieg über einen Thesaurus mit einer beliebigen Anzahl von Hierarchieebenen
2. Sprung in eine vorangestellte Übersichtsliste als Pseudotext mit einem Angebot von Querverweismöglichkeiten
3. Querverweis in den Volltext.

Thesaurus



Punkt 2 stellt den aufgefüllten Thesaurus mit der Übersichtsliste aller zugeordneten Textelemente für eben diesen ausgewählten Fachbegriff dar. Der Thesaurus selbst wird wie ein Inhaltsverzeichnis gehandhabt. Die Entscheidung für einen Fachbegriff bewirkt zunächst lediglich den Sprung in die Übersichtsliste mit einer Markierung (Highlighting).'



Erst nach einer erneuten Entscheidung erfolgt von hier aus und ausschließlich sternförmig der Sprung in den Volltext. Da es sich bei Punkt 2 um einen quasi Textteil handelt, kann die gefundene Textübersicht ständig auf einem Bildschirmteil beibehalten werden.

Damit übernimmt der aufgefüllte Thesaurus für diesen Informationsblock de facto die Funktion des Inhaltsverzeichnisses, so daß bei großen Informationsmengen – wie sie die CD-ROM nun einmal speichert – der Thesaurus eine vernünftigen Alternative zum Inhaltsverzeichnis darstellen kann.

3. Strukturierte Handlungsabläufe

Eine weitere Einstiegsmöglichkeit in ein Informationspaket stellen übersichtlich strukturierte Handlungsabläufe dar, die für den jeweiligen Sachverhalt vom Handlungsstrang aus den Sprung in das Informationspaket organisieren (Abb. 2). Das erfolgt sternförmig und ist im Grunde eine Informationseinschränkung, die nur für diesen Handlungsstrang Gültigkeit besitzt.

Der Sprung vom Handlungsstrang zum Informationspaket

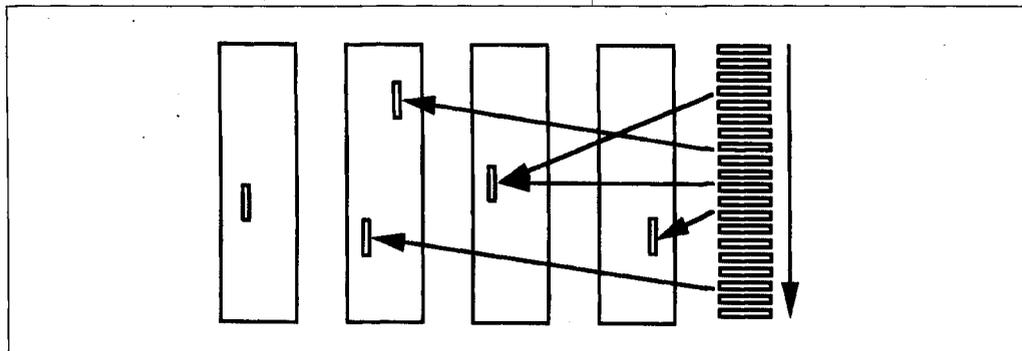


Abb. 2:
Strukturierte Handlungsabläufe

Bei einer ausreichenden Menge unabhängiger Handlungsangebote ist der Informationshintergrund tatsächlich nur noch Beiwerk und muß selbständig kaum noch in Anspruch genommen werden.

Für den Nutzer ist nur noch die Handlungschronologie wichtig. Er wird in der höheren Ebene einer Informationsaufbereitung geführt. Das Textelement selbst wird zur Bestätigung und Untermauerung der vorgelegten Abstraktion nur auf Wunsch noch herangezogen.

Hinterlegung einer Patentanmeldung beim DPA

- **Definition AN-Erfindung oder freie Erfindung** —————> **Hinweise Rechtsprechung**
- **Definition Erfindung** —————> **BPatG §**
- **Objektbereich**
- **Anforderungen**
- **Anmeldeunterlagen** —————> **Merkblatt DPA**
- **Gebühren**
- **Fristen**
- **Hinweise**

Beispiel „strukturierte Handlungsabläufe“

In der Praxis hat jedes Informationspaket seinen individuellen Charakter.

Technische Feinassen müssen sich dem Ziel unterordnen, daß alle relevanten Informationen möglichst schnell auffindbar sind und dabei ein Mindestmaß an Überblick gewährleistet wird. Welche Aufbereitungsform dafür am besten geeignet ist, kann nur der Nutzer selbst entscheiden. Es steht aber außer Frage:

Für zufriedenstellende CD-ROM-Anwendungen unter HyperText sind immer auch Überlegungen zur Informationslogistik erforderlich.

- Das trifft um so mehr zu, je weniger das Informationspaket von seiner Grundstruktur her für eine solche Anwendung aufbereitet ist.