

Das CAMbase-Literaturdatenbankprojekt zur komplementärmedizinischen und naturheilkundlichen Literatur

1. Einleitung

CAM ist eine Abkürzung für „Complementary and Alternative Medicine“, die sich, aus dem amerikanischen kommend, auch im europäischen Raum etabliert hat. Zu den unter diesem Begriff subsumierten medizinischen Richtungen und Verfahren zählen u.a. Naturheilverfahren, Homöopathie, Phytotherapie, Traditionell chinesische Medizin, Anthroposophische Medizin. Im deutschsprachigen Raum werden für CAM auch andere Oberbegriffe wie Komplementärmedizin, Naturheilkunde oder Unkonventionelle Medizinische Richtungen (UMR) verwandt [2,4].

Die Tatsache, daß die unter dem Terminus Komplementärmedizin zusammengefaßten Ansätze und Richtungen in der Vergangenheit „out of the mainstream“ des etablierten medizinischen Wissenschaftsbetriebs angesiedelt waren und im wesentlichen von niedergelassenen praktisch tätigen Ärzten angewandt wurden, hatte in der Vergangenheit dazu geführt, daß sich auf diesem Felde nur unzureichend personelle und infrastrukturelle Ressourcen für wettbewerbsfähige Forschungsaktivitäten und eine wissenschaftliche Weiterentwicklung auszubilden vermochten. Im Verlauf der letzten zwei Jahrzehnte hat sich diese Situation jedoch deutlich geändert [3]. Nicht zuletzt vor dem Hintergrund einer stetig wachsenden Nachfrage nach komplementärmedizinischen Behandlungsangeboten in den westlichen Industrienationen greifen immer mehr Ärzte und Therapeuten auf komplementärmedizinische Behandlungsverfahren zurück. Nach wie vor ist jedoch der Zugang zur wissenschaftlichen Literatur auf diesem weiten und außerordentlich heterogenen Feld sehr erschwert und außerordentlich unübersichtlich [1]. Nur wenige relevante Artikel erscheinen in den etablierten, konventionellen Zeitschriften, der weitaus größte Teil wird in eigenen Publikationsorganen veröffentlicht. Einige dieser Zeitschriften haben eine langjährige Tradition und spielen eine wesentliche Rolle bei der Weitergabe von Informationen. Daneben gibt es neuere Zeitschriften, die im Rahmen aktueller Forschungsbemühungen vorwiegend wissenschaftliche Arbeiten veröffentlichen.

Insbesondere die wissenschaftlichen Arbeitsgruppen, die sich in den vergangenen Jahren um die Intensivierung der Forschung in diesem Bereich bemüht haben, mußten mit der Schwierigkeit umgehen, einen zuverlässigen Überblick über die jeweilige Veröffentlichungslage gewinnen zu können. In Ermangelung etablierter, klar strukturierter Zugriffsmöglichkeiten auf entsprechende Literatur haben diese Arbeitsgruppen eigene Literaturdatenbanken eingerichtet.

Insgesamt ist die wissenschaftliche Literatur damit auf unterschiedliche Quellen verstreut [9], und es existiert im deutschsprachigen Raum kein allgemein öffentlich frei zugängliches System, mit dessen Hilfe ein schneller Überblick über die relevanten Publikationen möglich ist. Konventionelle Datenbanken und Informationssysteme - wie z.B. MedLine - haben zwar einen Teil der Literatur des komplementärmedizinischen Bereichs in ihren Bestand aufgenommen. Zuverlässige Literaturrecherchen lassen sich mit den derzeitigen Möglichkeiten allerdings nur mit einem unverhältnismäßig hohen Zeitaufwand durchführen und erfordern darüber-

hinaus bereits umfangreiche Vorkenntnis der Publikationsstrukturen. Folglich besteht ein dringender Bedarf nach der Etablierung eines Datenbanksystems, das die komplementärmedizinische Literatur möglichst weitgehend erfaßt und allgemein zugänglich macht. Diese Tatsache ist besonders deutlich geworden im rahmen zweier gezielter staatlicher Forschungsförderprogramme „Unkonventionelle Medizinische Richtungen“ und „Unkonventionelle Methoden der Krebsbekämpfung“, die von gleichnamigen Arbeitsgruppen an der Universität Witten/Herdecke begleitet wurden [7,8]. Im Rahmen einer umfassenden Bestandsanalyse zum Stand der Forschung auf diesem Feld [2] hat sich gezeigt, wie schwierig und aufwendig es ist, einen Überblick über die publizierte Evidenz zu den komplementärmedizinischen Ansätzen und Verfahren zu gewinnen. Die beiden vorgenannten Arbeitsgruppen haben deshalb mit dem Aufbau einer eigenen Literaturdatenbank zur Komplementärmedizin begonnen. Vor diesem Hintergrund wurde 1999 am Lehrstuhl für Medizinteorie der Universität Witten/Herdecke das CAMbase-Projekt initiiert.

2. Voraussetzungen von CAMbase

CAMbase hat sich die Etablierung, den weiteren Ausbau und die Pflege einer über das Internet öffentlich und kostenfrei zugänglichen multizentrischen Literaturdatenbank zur Komplementärmedizin zu Aufgabe gemacht. Ausgangspunkt sind bereits vorhandene Datenbanken der wissenschaftlicher Arbeitsgruppen, die sich dieser Thematik widmen. In einem ersten Schritt wurden zunächst die vorhandenen Datenbanken dieser Arbeitsgruppen bezüglich des in ihnen enthaltenen Materials gesichtet. Obwohl diese Datenbanken hauptsächlich auf interne Forschungsaspekte ausgelegt waren, zeigte eine erste Analyse, daß dort über die Jahre so viel Material gesammelt worden ist, daß eine Abfragemöglichkeit auch für Außenstehende von großem Interesse ist. Aufgrund Ihrer Ausrichtung als Forschungsdatenbanken zeigte sich außerdem, daß diese Datenbanken bereits qualitativ selektiertes Material enthalten und sich durch unterschiedliche Schwerpunkte auszeichnen. Nachteilig wirken sich dabei die unterschiedlichen Systeme aus, mit denen die Archivierung der Literatur durchgeführt wurde. Neben den herkömmlichen Systemen wie z.B. EndNote oder ReferenceManager lagen bei der Bestandsaufnahme auch selbsterstellte Datenbankformate vor, die u.a. auf dBase- bzw. MS Access-Formaten beruhen.

3. Technische Realisierung

3.1 Probleme klassischer Retrievalsysteme

Thematische Abfragen in Datenbanken jeder Art (bibliographischen Datenbanken, Faktendatenbanken, Volltexten etc.) sind nach dem derzeitigen Stand der Technik ausnahmslos an der Verwendung sogenannter Boolescher Operatoren ('und', 'oder', 'nicht') orientiert, mit deren Hilfe komplexere Suchanfragen vom Benutzer strukturiert werden müssen. Die meist in 'natürlicher Sprache' präzise formulierbaren Fragen im Zusammenhang thematischer Suchen in Datenbanken, wie z.B.

- kleinzelliges Karzinom
- Misteltherapie des Krebses
- Laser in der Medizin
- das Gespräch in der Therapie

werden mit Hilfe Boolescher Operatoren auf Strukturen heruntergebrochen, die mit den ursprünglich syntaktisch-semantisch geordneten Konstruktionen nur noch wenig zu tun haben. Die vom Benutzer der Datenbank mit Ausdrücken wie „... in der Onkologie“, „... in den neuen Bundesländern“, „... mit den Mit-

teln der Spektralanalyse“, „... im Unterricht“, usw. explizit formulierte Modifikation eines Themas wird nicht als solche erkannt, sondern muß als eigenständiges Suchwort bei der Abfrage behandelt werden und würde ausgehend von den oben genannten Beispielen etwa zu folgenden Suchanfragen führen:

- kleinzellig **und** karzinom
- mistel\$ **und** therapie\$ **und** krebs\$
- laser\$ **und** mediziner
- gespräch **und** therapie

Die thematische Frage an die Datenbank wird also vom Retrievalsystem syntaktisch-semantisch darauf reduziert, ob (zufällig) Datensätze vorhanden sind, die die genannten Wörter gleichzeitig enthalten. Dabei kann der Sinn der Ausgangsfrage nicht mehr rekonstruiert werden. Die thematischen Fragen 'Krankenhaus im Unterricht' und 'Unterricht im Krankenhaus' können bspw. in einer Boole'schen Verknüpfungslogik nicht mehr unterschieden werden.

Entsprechend gestaltet sich die Präsentation der Suchergebnisse zufällig, unscharf, unpräzise und sehr oft mit der Rückmeldung für den Benutzer 'Null Treffer' oder „zu viele Treffer'. Null-Treffer-Situationen ergeben sich für sehr oft aufgrund endungsmorphologischer Phänomene oder im Zusammenhang unregelmäßiger Pluralbildungen (Krankenhaus, Krankenhäuser). Hoffnungslos unübersichtliche Treffermengen entstehen massiv in großen Volltextsystemen, in denen aus dem Umstand, daß zwei Begriffe gleichzeitig (ohne Rücksicht auf ihren syntaktisch-semantischen Zusammenhang) in einem Dokument präsent sind, nahezu nichts mehr über die mögliche Relevanz des getroffenen Dokuments abgeleitet werden kann.

Weiterführende Retrievaltechniken wie Trunkierung, Additional Discriminating Information (ADI), Proximity oder die Verwendung von Relevance-Ranking-Techniken setzen, abgesehen davon, daß sie diese Probleme im allgemeinen auch nicht lösen können, darüberhinaus in der Regel detaillierte Retrievalkenntnisse des Benutzers voraus.

3.2 Das OSIRIS-Projekt

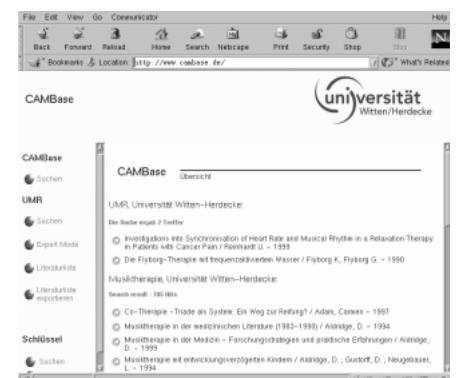
Das OSIRIS-Projekt ist ein gemeinsames Projekt der UB Osnabrück und des Instituts für Semantische Informationsverarbeitung (ISIV) der Universität Osnabrück, das von 1996-1999 von der Deutschen Forschungsgemeinschaft (DFG) im Rahmen des Bibliotheksförderprogrammes gefördert wurde [5,10].

OSIRIS ist ein intuitiv zu bedienendes Recherchesystem für Datenbanken, das Eingaben in natürlicher Sprache verarbeiten kann, z.B. für die Suche nach Literatur zu einem bestimmten Thema. Deshalb kann OSIRIS (im Gegensatz zu anderen Recherchesystemen) auch von Suchenden benutzt werden, die weder die speziell in einer Datenbank verwendete Terminologie kennen oder verwenden wollen noch etwas von klassischen Retrievaltechniken, speziellen Kommandos, Booleschen Verknüpfungen, Trunkierung etc. wissen oder wissen wollen. Dabei erzielt OSIRIS qualitativ wesentlich bessere Suchergebnisse im direkten Vergleich mit in konventioneller Indexierungstechnik aufbereiteten Literaturdatenbanken: Der Datenbank-Recall erhöht sich bei OSIRIS ganz wesentlich gegenüber konventionellen Abfragen, ohne im Rechercheergebnis an Präzision zu verlieren [6].

Die Einsatzmöglichkeiten der in OSIRIS entwickelten Lösungen sind nicht auf Literaturdatenbanken beschränkt, sondern bieten Perspektiven für Datenbankbenutzer allgemein. Einfacher, natürlichsprachlicher Zugang zusammen mit der Garantie hochwertiger Retrievalergebnisse kann das Dilemma heutiger Online-Datenbanken, im Hinblick auf die stark zunehmenden Massen abfragbarer Daten nur für geschulte Benutzer effektiv nutzbar zu sein, lösen helfen.

3.3 Die OSIRIS-Techniken

Unter Einsatz eines speziellen Verfahrens für syntaktisch-semantisches Retrieval werden im Osiris-System natürlichsprachliche Datenbankabfragen verarbeitet. Auf der Basis einer deklarativ programmierten Grammatik der natürlichen Sprache, die die Regeln des modellierten Sprachausschnitts berücksichtigt, werden Benutzereingaben syntaktisch und semantisch interpretiert. Dafür werden einschlägige Methoden der semantischen Informationsverarbeitung angewendet (fehlersensitives syntaktisches Parsing, Morphologie, Kompositazerlegung). Die Analyseer-



gebnisse werden unter Berücksichtigung erkannter syntaktisch-semantischer Zusammenhänge prozessiert, interpretiert und bewertet.

Beispiel:

Ich suche Literatur zum Thema:
placebokontrollierte Doppelblindstudien in der Homöopathie

Die im OSIRIS-Projekt entwickelten Techniken führen in Datenbanken, die nur über knappe textuelle Beschreibungen ihrer Objekte verfügen (z.B. in Bibliothekskatalogen, Branchendatenbanken etc.), zu einer dramatischen Verbesserung der Recalls für den Datenbankbenutzer. Er profitiert stark von den Möglichkeiten natürlichsprachlicher Eingaben, sowie von den syntaktisch-semantischen Analysen und Bewertungen der ausgewählten Treffermengen, die relativ unabhängig von endungsmorphologischen Phänomenen oder zufälligen Kompositakonstruktionen etc. gebildet werden können.

Die natürlichsprachlichen Komponenten des OSIRIS-Systems sind zur Zeit für die deutsche und die englische Sprache optimiert. Die entwickelten Verfahren sind jederzeit für andere Sprachen adaptierbar.

3.4 Der virtuelle CAMBase-Verbund

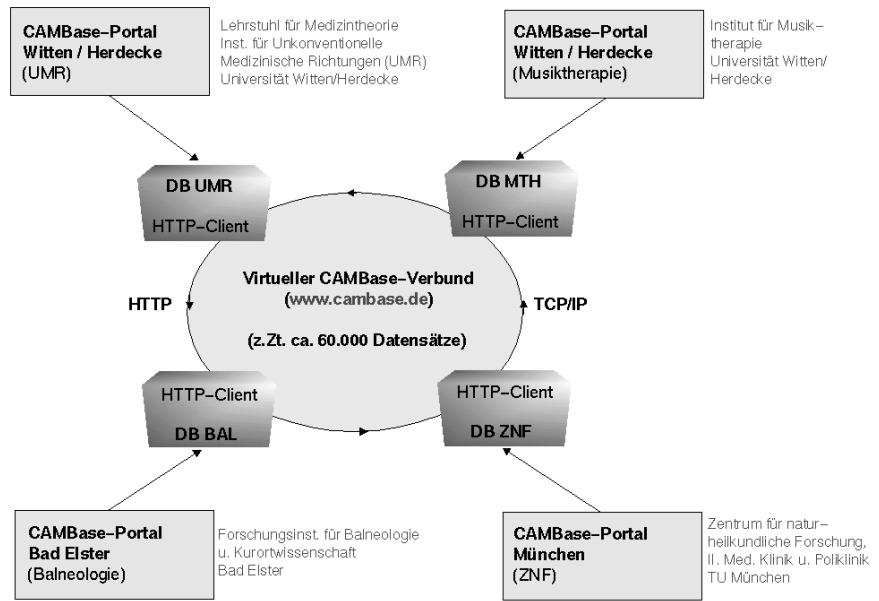
Der CAMBase-Verbund ist als virtueller Verbund realisiert, d.h. die Verbundabfrage erfolgt auf dem Protokoll-Level HTTP auf physikalisch getrennten und durchaus heterogenen Datenbanken (unabhängig von der Hardware- und Datenbankplattform und natürlich standortunabhängig). Jeder Partner des CAMBase-Verbundes bleibt in dieser Architektur völlig eigenständig bis hin zum Layout seiner Datenbank. Anders als in klassisch organisierten Verbundsystemen entfallen in dieser Architektur die aufwendigen Absprachen über Datenformate, Erfassungsregeln etc. Die dauerhaft garantierte Gestaltungsfreiheit auch künftiger Verbundpartner macht die Entscheidung leicht, sich an den gemeinsamen Aktivitäten in CAMBase zu beteiligen.

4. Aktueller Stand

Der CAMbase-Verbund umfaßt derzeit ca. 60.000 Datensätze, verteilt auf die Datenbanken der kooperierenden Partner:

Institut für Komplementärmedizin, Lehrstuhl für Medizintheorie, Universität Witten-Herdecke,

Abb.: Derzeitige Struktur des Cambase Verbundes



Institut für Musiktherapie, Lehrstuhl für qualitative Forschung, Universität Witten-Herdecke,

Forschungsinstitut für Balneologie und Kurortwissenschaft, Bad Elster,

Zentrum für naturheilkundliche Forschung der II. Medizinischen Klinik und Poliklinik der Technischen Universität München.

Erfasst werden die herkömmlichen bibliografischen Daten aus den Publikationen (Autor, Titel, Quelle, Seitenzahl, Keywords,...). Sofern ein Abstract oder eine Zusammenfassung vorliegt, werden diese seit 1998 miterfaßt. Rückwirkend werden seit kurzem auch die älteren Datensätze durch Abstracts ergänzt, falls diese vorhanden sind.

Das Verhältnis von Büchern/Buchbeiträgen zu Artikeln ist je nach Datenbank der Partner unterschiedlich, liegt jedoch im Mittel bei etwa 1:4. Im Verbund sind mehr als 80 Zeitschriften aus dem komplementärmedizinischen und naturheilkundlichen Bereich erfaßt, wobei die folgenden Zeitschriften schwerpunktmäßig von den CAMbase-Partnern erfaßt werden:

Ärztezeitschrift für Naturheilverfahren, Allgemeine Homöopathische Zeitschrift, Alternative Therapies, Biologische Medizin, British Homeopathic Journal, British Journal of Music Therapy, Complementary Therapies in Medicine, Deutsche Zeitschrift für Akupunktur,

Erfahrungsheilkunde, Forschende Komplementärmedizin, Journal of Alternative and Complementary Medicine, Journal of Music Therapy, Musiktherapeutische Umschau, Physikalische Medizin, Phytomedicine, Zeitschrift für Klassische Homöopathie, Zeitschrift für Phytotherapie,

Neben diesen Zeitschriften werden auch "konventionelle" Periodika auf komplementärmedizinischen Inhalt hin durchsucht und entsprechende Datensätze aufgenommen.

Während der Suche im CAMbase-Verbund besteht für den Benutzer die Möglichkeit, eine Literaturliste anzulegen, in der die für den Benutzer relevanten Datensätze abgespeichert werden können. Neben der konventionellen Darstellung als html-file, besteht auch die Möglichkeit, diese als Endnote-kompatible tagged-ascii-Textfile zu generieren. Damit können die Suchergebnisse in die persönlichen Literaturverzeichnisse der Benutzer integriert werden.

5. Diskussion und Ausblick

Ein großes Problem ist die qualitative Heterogenität der Publikationen. Es stellt nach wie vor eine Hauptschwierigkeit dar, aus der Fülle an Veröffentlichungen jeweils inhaltlich relevante und methodisch gute, aussagekräftige Arbeiten herauszufinden. Speziell die

Identifikation von qualitativ hochwertigen wissenschaftlichen Arbeiten, für die es international akzeptierte Kriterien gibt, ist besonders erschwert. Auch die Identifikation von guten Kasuistiken, Erfahrungsberichten oder theoretischen Aspekten gewidmeten Darstellungen ist weiterhin schwierig. Dies betrifft auch die Buchpublikationen. Inwieweit eine inhaltliche Bewertung der erfassten Datensätze neben dem zu leistenden Zeitaufwand für die potentiellen Nutzer des Systems von Vorteil ist, wird sich in der Zukunft herausstellen. Zumindest ist dies eine Richtung, in der CAMbase zukünftig weiter ausgebaut werden kann.

Langfristig soll auch der Zugang zum Volltext der Originalpublikationen verbessert werden. Eine praktikable und sinnvolle Möglichkeit bestünde darin, eine zentrale Sammelstelle für Zeitschriften und Buchpublikationen der UMR - z.B. in einer Universitätsbibliothek - zu erstellen. Eine Kombination einer solchen Sammelstelle mit dem geplanten Datenbanksystem wäre folgerichtig und sinnvoll. Ein entsprechendes Projekt aus den USA, das Office of Alternative Medicine (OAM) am National Institute of Health (NIH) ist dafür ein gutes Beispiel. Entsprechende Kontakte werden derzeit geknüpft, um neben der europäischen Ausweitung CAMbase auch in den USA bekannt zu machen.

Thomas Ostermann¹

Dr. Hartmut Zillmann²,
Alte Münze 16, 49074 Osnabrück

Peter F. Matthiessen¹

¹Lehrstuhl für Medizintechnik, Universität Witten/Herdecke,

²Universitätsbibliothek Osnabrück

Internet-Adresse von CAMbase:
www.cambase.de

Literatur:

Barnes J; Abbot NC; Harkness EF; Ernst E: Articles on complementary medicine in the mainstream medical literature: an investigation of MEDLINE, 1966 through 1996. Arch Intern Med 1999; 159(15):1721-5

Matthiessen PF, Roßlenbroich B, Schmidt S. Unkonventionelle Medizinische Richtungen: Bestandsaufnahme zur Forschungssituation. Materialien zur Gesundheitsforschung. Bd.21 Bonn 1992

Ostermann T., Brinkhaus, B. Melchart, C: Das Forum universitärer Arbeitsgruppen für

Naturheilverfahren und Komplementärmedizin. Forsch Komplementärmed 1999; 6:41-42

Panel on Definition and Description, CAM Research Methodology Conference, April 1995: Defining and Describing Complementary and Alternative Medicine. Alternative Therapies 1997;3(2):49-57

Ronthaler, M.: Osiris: Qualitative Fortschritte bei der Literaturrecherche In: Dassow, J./ Kruse, R. (Hrsg.): Informatik '98 : Informatik zwischen Bild und Sprache ; 28. Jahrestagung der Gesellschaft fuer Informatik; Magdeburg, 21. bis 25. September 1998. Berlin [u.a.] : Springer, Informatik aktuell, 1998

Ronthaler, M., Zillmann, H.: Literaturrecherche mit OSIRIS: ein Test der OSIRIS-Retrievalkomponente. In: Bibliotheksdienst, 32(7):1203-1212, 7 1998

Roslenbroich, B., Teichert, J., Schulze-Pillot, Th., Matthiessen, P.F.(1997): Erste Etappen der Forschung in der Unkonventionellen Medizin und die staatliche Forschungsförderung. Forschende Komplementärmedizin, 4: 52-57

Teichert, J., Schulze-Pillot, Th., Matthiessen, P.F. (1994): Unkonventionelle Methoden der Krebsbekämpfung. Deutsches Ärzteblatt, 91. Jahrg., Heft 48

Wootton JC: Directory of Databases for research into alternative and complementary medicine. J Altern Complement Med 1997;3(2):179-90

Zillmann, H.: The OSIRIS project: Osnabrück Intelligent Research Information System Vortrag auf der Anglo-German conference: Building the Electronic Library - British and German experiences, 09.,10.,11. Januar 1998, Worcester College, Oxford

Kluwer-Verlagsdatenbank für Volltexte

Seit 29.03.2001 bietet DIMDI in Kooperation mit der Deutschen Zentralbibliothek für Medizin (ZBMed) in Köln die englischsprachige Kluwer-Verlagsdatenbank sowie die zugehörigen Volltexte im Portable Document Format (PDF) an. Die Literaturdatendank beinhaltet bibliographische Angaben, nicht kontrollierte Schlagworte (UT) und häufig eine englische Kurzzusammenfassung zu Dokumenten aus ca. 600 internationalen Zeitschriften des Kluwer-Verlages. Diese kommen aus den Bereichen Naturwissenschaften, Medizin, Geisteswissenschaften, Mathematik und Informatik sowie Rechtswissenschaften. Der jährlich um ca. 25.000 Dokumente wachsende Datenbestand umfasst ca. 115.000 Dokumente (Stand März 2001), die seit 1997 veröffentlicht wurden.

Die Datenbank wird ausschließlich über grips-WebSearch sowie Free grips-WebSearch angeboten und ist entgelt- und lizenzfrei. Der Zugriff auf elektronische Volltexte ist kostenpflichtig und nur mit User Code möglich. Durch Anklicken des Buchsymbols rechts neben dem entsprechenden Titel kann der Volltext im PDF-Format angefordert werden. Vor dem endgültigen Öffnen des Volltexts wird auf entstehende Kosten hingewiesen.

Ein besonderes Highlight stellt die Verknüpfung der Kluwer-Volltexte in grips-WebSearch mit Publikationshinweisen aus anderen Literaturdatenbanken z.B. Medline oder Embase dar. Selbst wenn die Kluwer-Verlagsdatenbank nicht explizit mit aufgerufen wurde, weist das Buchsymbol auf einen verfügbaren Volltext hin.

Durch Vertragsabschluss mit weiteren Verlagen wird in Zukunft eine ständig wachsende Anzahl Volltexte online durch einfaches Anklicken verfügbar sein. Die Abrechnung erfolgt auf der Basis von „pay per view“ im Rahmen der vertraglichen Konditionen von DIMDI

(s.a.: <http://www.dimdi.de/germ/vertrag/fr-vertrag.htm>). Ansprechpartnerin: Sylvia Schlutius, Tel.: 0221 47 24 229, E-Mail: schlutius@dimdi.de (aus den grips-News April 2001)