

# KHS — ein offenes Hypertext-System<sup>1</sup>

Rainer Hammwöhner / Marc Rittberger

Universität Konstanz Informationswissenschaft

## Zusammenfassung

Im vorliegenden Beitrag werden acht Kriterien für offene Hypertextsysteme beschrieben und anhand des Konstanzer-Hypertext-Systems diskutiert. Die Einbindung externer Informationsquellen und die sich daraus ergebenden Konsequenzen für das Konstanzer-Hypertext-System werden mit Hilfe von E-Mail und Online-Datenbanken ausführlicher dargestellt.

## 1 Einführung

"Offene Hypertext-Systeme" sind in neueren Arbeiten aus dem Bereich Hypertext zu einem vieldiskutierten Thema geworden. In den Proceedings der letzten ECHT '92<sup>2</sup> z.B. finden sich mehrere interessante Arbeiten zu diesem Thema, von denen eine [DHH+92] einen Kriterienkatalog für offene Hypertext-Systeme formuliert, der insbesondere auf die Unabhängigkeit von Systemplattformen und der Kommunikation mit externen Applikationsprogrammen abzielt. Wir halten diese Eigenschaften für wichtig<sup>3</sup>, sind aber der Meinung, daß Offenheit weitergehende Konsequenzen auch auf die innere Struktur des Hypertexts haben muß. Im folgenden werden wir einen eigenen Kriterienkatalog vorlegen, der dann an das Konstanzer Hypertext-System angelegt werden soll. Ausführlicher dargestellt werden zwei Beispiele für die Einbindung externer Informationsquellen in Hypertext, weil uns diese im Umfeld des Information Retrieval als besonders interessant erscheinen.

Nach unserem Dafürhalten zeichnen sich offene Hypertext-Systeme durch folgende Eigenschaften aus:

1. Jeder Leser ist im Prinzip auch ein Autor, seine Möglichkeiten, den Hypertext zu verändern, sind allein durch die Rechte anderer Hypertext-Nutzer eingeschränkt.
2. Das Hypertext-System erlaubt mehreren Benutzern gleichzeitig Zugang zum selben Hypertext.
3. Der Umfang der durch eine Vielzahl von Autoren eingebrachten Information erfordert eine Differenzierung der Ausdrucksmittel, um dem Leser eine Relevanzeinschätzung zu ermöglichen.
4. Offene Hypertexte sind konsistenzwährend.
5. Offene Hypertext unterstützen verschiedene Interaktionsstrategien.

---

<sup>1</sup> Dieser Beitrag wurde publiziert in: G. Knorz, J. Krause and C. Womser-Hacker (Hrsg.), *Information Retrieval '93. Von der Modellierung zur Anwendung*, pp. 208 - 222. Universitätsverlag Konstanz, 1993. Proceedings der 1. Tagung Information Retrieval'93.



This text is published under the following Creative Commons Licence: Attribution-NonCommercial-NoDerivs 2.0 Germany (<http://creativecommons.org/licenses/by-nc-nd/2.0/de/>).

<sup>2</sup> ACM Conference on Hypertext, Mailand, 30.11.-4.12. 1992

<sup>3</sup> KHS ist in Smalltalk-80 implementiert und somit auf Unix-Systemen, Apple McIntosh und Windows-Systemen verfügbar.

6. Offene Hypertexte sind aktive Medien.
7. Das System bietet bei aller Allgemeinheit Unterstützung für spezifische Teildiskurse.
8. Daten aus externen Informationsquellen können in den Hypertext integriert werden.

## 2 KHS als offenes Hypertext-System

Im folgenden wollen wir die oben eingeführten Kriterien offener Hypertext-Systeme detaillierter ausarbeiten und als Grundlage einer komprimierten Darstellung des Konstanzer Hypertext-Systems (KHS) verwenden.

### 2.1 Der Leser wird zum Autor

Die traditionellen Printmedien sind Instrumente öffentlicher bzw. veröffentlichter Schriftlichkeit [Gie91]. Ein Autor bzw. ein Autorenkollektiv verfaßt ein Werk, das — mit Hilfe eines Verlegers — einem Publikum von Lesern zur Verfügung gestellt wird. Letztere werden das Werk häufig als Grundlage für weitere Arbeit verwenden, die es erfordert, das Werk in rein privatschriftlicher Form durch Kommentare, ergänzende Notizen etc. zu erweitern. Diese Antinomien von Autor und Leser, sowie veröffentlichter und privater Schriftlichkeit werden von Hypertext entschärft, aber nicht vollständig aufgelöst, wie ersichtlich wird, wenn man den Gegenstand jeweils einmal aus Sicht des Lesers und der des Autors betrachtet.

**Der Leser** eines Hypertexts will Assoziationen, die sich bei der Lektüre einstellen, durch Links wiedergeben können, sowie Kommentare zu den gelesenen Inhalten notieren können. Als Leser hat er aber auch ein Interesse daran, daß die für ihn relevante Information nicht durch ein Übermaß an Verweisen überdeckt wird.

**Der Autor** einer speziellen Hypertext-Einheit hat ein Interesse daran, daß die von ihm für einen bestimmten Leserkreis geplante Didaktik einer Hypertext-Einheit nicht durch externe Eingriffe gestört wird.

Diesen Widerspruch versucht KHS wie folgt aufzulösen:

1. Der Autor einer Hypertext-Einheit kann festlegen, wer diese in Zukunft lesen bzw. ändern darf<sup>4</sup>. Das Einfügen eines global sichtbaren Links wird als eine Änderung der beteiligten Hypertext-Einheiten aufgefaßt.
2. Jeder Leser kann Links zwischen beliebigen Hypertext-Einheiten einfügen. Hat er keine Erlaubnis, diese Einheiten zu ändern, so sind die entstehenden Links privat, d.h. nur für ihn selbst sichtbar.
3. Wird eine Schreibsperre auf eine Unit aufgehoben, so werden auch die zugehörigen privaten Links automatisch veröffentlicht, falls ihr Autor dies nicht explizit unterbunden hat.

### 2.2 Mehrbenutzerzugang zum Hypertext

Ein offenes Hypertext-System erlaubt den parallelen Gebrauch durch mehrere Benutzer. Der Mehrbenutzerbetrieb von KHS erfolgt unter folgenden Randbedingungen:

---

<sup>4</sup> Wichtig scheint uns zu sein, daß die Präsentation des Hypertexts keinerlei Indizien auf Objekte geben darf, die dem Nutzer keinen lesenden Zugriff erlauben.

1. Wird eine Hypertext-Einheit geändert, erfolgt eine Schreib-, keine Lesesperre für andere Nutzer. Diese Schreibsperre schließt die Vergabe von Links und Anmerkungen ein, um eine inhaltliche Inkonsistenz von Links bzw. Anmerkungen zu vermeiden.
2. Erfolgreich abgeschlossene Änderungen führen zur Aktualisierung visueller Darstellungen der geänderten Units bei allen Lesern. Aus Effizienzgründen sind Ausnahmen von dieser Regel erlaubt, falls die jeweilige Unit nicht im Zentrum des Interesses ist, d.h. mit dem jeweiligen Tool nicht gearbeitet wird.

### 2.3 Differenzierung der Ausdrucksmittel

Die diskursive Funktion von Textpassagen ergibt sich bzgl. eines gedruckten Textes entweder aus dem Layout (Fußnoten, Glossare) oder durch die Position innerhalb einer fest vorgegebenen Lesefolge im Sinne einer Textgrammatik. Für Hypertexte greifen diese Konventionen des Schreibens und Lesens nicht mehr, da von einer vorgegebenen Lesefolge bzw. einem Layout eines Hypertexts nicht die Rede sein kann. Dieser Verlust an orientierender Metainformation muß in Hypertexten durch andere Ausdrucksmittel kompensiert werden. Im folgenden sollen die von KHS eingesetzten Strukturelemente kurz eingeführt werden:

1. **Aggregation:** Sogenannte "composite units" fassen mehrere andere Hypertext-Einheiten zu einer Gruppe zusammen. Die Einbettung einer terminalen, d.h. mediale Information enthaltenden Einheit, in eine Polyhierarchie aggregierter Hypertext-Knoten gibt dem Leser Information über seine augenblickliche Position im Hypertext und über Thematik bzw. Funktion der Einheit.
2. **Typisierung:** Durch die Definition von Typen lassen sich inhaltliche Restriktionen für Hypertext-Objekte definieren. Diese Informationen erlauben dem Leser, von dem aus der Darstellung erkannten Objekt-Typ auf den Inhalt zu antizipieren. Darüber hinaus kann die durch Typisierung verfügbare Strukturinformation zur Konsistenzwahrung im Hypertext herangezogen werden (s.u.).
3. **Unit-Typen** definieren die inhaltliche Struktur und die Präsentationsform von Hypertext-Einheiten. Terminale Einheiten werden differenziert, je nachdem, ob sie textuelle, tabellarische oder graphische Information enthalten (Ton ist kein Unterscheidungskriterium, da jede Einheit mit Ton verknüpft werden kann). **Typen nonterminaler Einheiten** definieren, welche Knotentypen für die Untereinheiten erlaubt sind, sowie ein Ordnungskriterium für eine mögliche Sequenzierung der Untereinheiten. **Link-Typen** definieren inhaltliche Eigenschaften der verknüpften Hypertext-Objekte (Units, Hotwords bzw. Hotareas in Units, Links), die Stelligkeit und Richtung ihrer Ausprägungen.

**Kontextualisierung von Verweisen:** Durch die Einbindung in die Polyhierarchie aggregierter Hypertext-Einheiten wird jede Unit in einen bestimmten von der Existenz von Verweisinformation unabhängigen Kontext gestellt. Durch polyhierarchische Einbindung kann eine Hypertext-Einheit in mehrere Kontexte integriert werden. Dabei kann die Gültigkeit von Links auf einen Teil dieser Kontexte eingeschränkt werden, so daß die während der Navigation verfügbare Verweisinformation einer Hypertext-Einheit vom Verwendungszusammenhang abhängig gemacht werden kann.

### 2.4 Konsistenzwahrung in Hypertexten

Der Umfang der in einem offenen Hypertext enthaltenen Information, die zudem noch ständiger Änderung unterworfen ist, bewirkt, daß Hypertext-Autoren die Konsequenzen einer Änderung in der Hypertext-Struktur nicht immer vollständig überblicken können.

Dementsprechend muß das System über Mechanismen zur Konsistenzwahrung verfügen, damit dem Leser keine semantisch inkorrekten oder gar "blind endende" Links etc. angeboten werden.

Andererseits ist eine der wichtigen Eigenschaften offener Hypertext-Systeme, daß sie verteilt und unter Einbeziehung externer Ressourcen definiert sein können, die nicht jederzeit verfügbar sein müssen, so daß Inkonsistenzen der Hypertext-Strukturen nicht immer zu vermeiden sind [Hew85]. KHS unterscheidet dementsprechend drei verschiedene Modi der Konsistenzprüfung:

1. Strukturelle Konsistenzwahrung zur Laufzeit betrifft alle Ressourcen, die ständig verfügbar sind. Die Konsistenz der im Hypertext enthaltenen Strukturen kann garantiert werden, obsoleete Objekte (z.B. Links, die auf eine gelöschtes Hypertext-Objekt verweisen) werden sofort aus sämtlichen Darstellungen und den internen Repräsentationen entfernt.
2. Strukturelle Konsistenzprüfung zur Ladezeit betrifft Ressourcen, die nur einem eingeschränkten Benutzerkreis ständig verfügbar ist. Hier wird z.B. überprüft, ob die von privaten, nur dem Leser verfügbaren Links referenzierten Hypertext-Objekte noch verfügbar sind, oder ob ein Link durch Aufheben der Schreibsperre eines Hypertext-Objekts von einem privaten in einen global verfügbaren Link verwandelt werden kann.
3. Konsistenz der Präsentation wird jeweils bei der Darstellung eines Hypertext-Objekts erzielt. Ist eine Ressource temporär nicht verfügbar, oder ist nicht entscheidbar, ob sie obsolet ist, so wird sie von der Darstellung ausgeschlossen, bleibt aber in den internen Repräsentationen enthalten.

## 2.5 Interaktionsstrategien

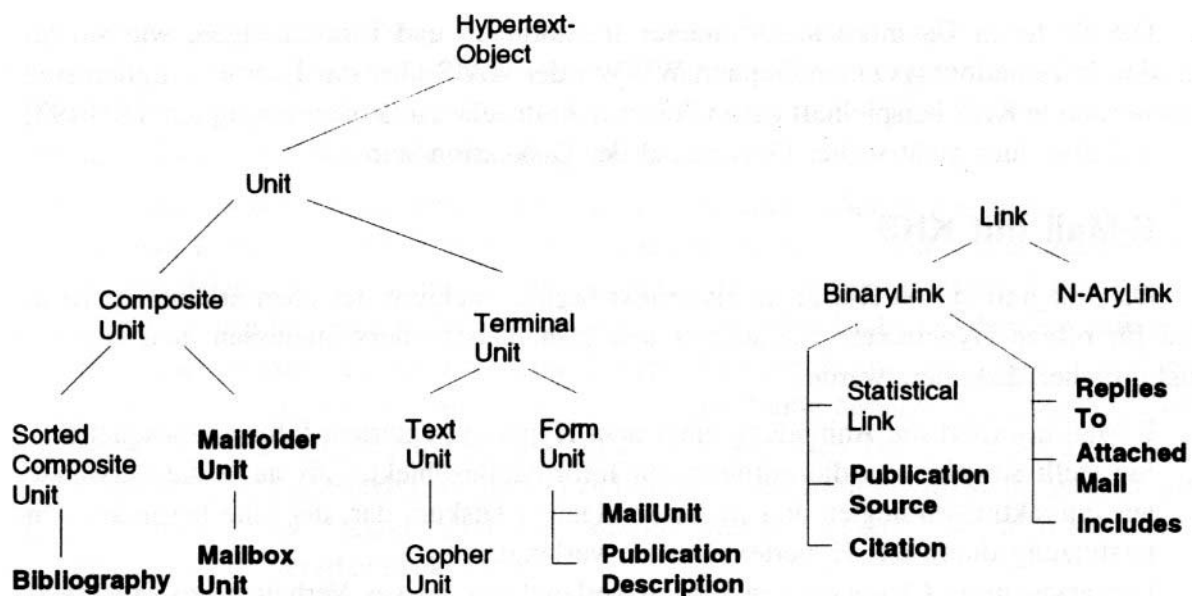
Die genuin hypertextorientierte Interaktionsstrategie ist die der Exploration und des Browsing. Offene Hypertexte mit ihrem großen Informationsangebot können allerdings nicht allein explorativ genutzt werden, die Chance, allein auf diese Weise auf die gesuchte Information zu stoßen, ist zu gering. Um geeignete Startpunkte für weitere Exploration zu finden, sind Verfahren des traditionellen Information Retrieval sehr geeignet, von denen KHS bisher folgende unterstützt:

1. Schlagwortsuche, jede Hypertext-Einheit kann mit Schlagworten deskribiert werden, die zur Suche herangezogen werden können,
2. Volltextsuche,
3. für strukturierte Hypertext-Einheiten durch Aufbau einer Beispiel-Hypertext-Unit mit den gewünschten Eigenschaften.

Auch der Vorgang der Exploration kann durch unterschiedliche Interaktionsstrategien unterstützt werden. Beispielhaft seien hier zwei einfache von KHS unterstützte Strategien erwähnt, die der Optimierung der dem Benutzer präsentierten Informationsmenge dienen:

1. Bei der Exploration hierarchischer Strukturen wird nur der Inhalt der aktuell im Fokus des Interesses befindlichen Unit dargestellt, sowie Kontextinformation, die den Fortschritt der Exploration anzeigt.
2. Bei Link-Navigation wird anhand des Link-Typs entschieden, ob der Inhalt der Ziel-Einheit interessant ist, insbesondere in Bezug auf die Ursprungs-Einheit, oder Relevanz aus sich selbst gewinnt. Im ersteren Fall (Glossar-Einträge, Annotationen) werden die Inhalte beider Einheiten komparativ dargestellt, während im zweiten Fall die Präsentation der Ursprungseinheit durch die der Zieleinheit ersetzt wird.

Zusätzlich eingeblendete Informationen werden bei weiterer Navigation automatisch wieder entfernt.



**Abbildung 1** Stark vereinfachter Ausschnitt aus der Hierarchie der im KHS definierten Objekttypen. Die allgemeingültigen Objekttypen sind mit normalem, die zum Bereich Mail und Online-Retrieval gehörigen mit fettem Druck gesetzt.

## 2.6 Offene Hypertexte sind aktive Medien

Im vorangegangenen Text wurde schon darauf hingewiesen, daß ständige Veränderung und Umfang der Information den Nutzer hindern, einen umfassenden Überblick über den Inhalt eines Hypertexts zu gewinnen. Eventuell eingebundene externe Ressourcen verschärfen dieses Problem noch. Dementsprechend ist der Nutzer hier mehr als bei anderen Software-Anwendungen auf aktive Leistungen des Systems angewiesen.

Bisher wurden im KHS folgende Formen aktiver Systemleistungen erprobt:

1. Beratung bei der Hypertext-Navigation. Basierend auf den im Hypertext vorkommenden Objekt-Typen wurde eine Dialoggrammatik spezifiziert, die es erlaubt, das Leseverhalten eines Nutzers zu klassifizieren und zu extrapolieren [Beh92].
2. Automatische Generierung von Links. KHS unterscheidet grundsätzlich zwei verschiedene Typen von Links. Autorenlinks werden von Hypertext-Nutzern vergeben, sie sind semantisch unterspezifiziert (d.h. das System kann allenfalls notwendige Bedingungen für ihr Bestehen benennen) und repräsentieren eine kommunikative Absicht des Autors. Systemlinks sind semantisch voll spezifiziert und können daher vom System (dialogbegleitend oder in Batch-Läufen) automatisch generiert werden. KHS unterstützt zur Zeit folgende Formen systemgenerierter Links:
  - a. Ähnlichkeit zwischen Units aufgrund von Worthäufigkeiten,
  - b. Glossarverweise,
  - c. Literaturreferenzen.
3. Automatische Auswertung externer Informationsquellen. Externe Informationsquellen, wie E-Mail, oder externe Datenbanken (s.u.) sind nur dann erfolgversprechend in ein Hypertext-System zu integrieren, wenn ihre Auswertung

nicht zwingend vom Nutzer angestoßen werden muß und die Integration der Ergebnisse in den Hypertext weitgehend unterstützt wird.

## **2.7 Unterstützung von Teildiskursen**

Der allgemeine Diskurs von Hypertext ist der des Stöberns, des Kommentierens, Annotierens usw. Offene Hypertexte sollten darüber hinaus auch spezifischere Teildiskurse unterstützen, die spezielle Interaktionsformen erfordern. Das hier kurz eingeführte Beispiel der E-Mail wird weiter unten im Text noch ausführlicher behandelt. Ein Benutzer von E-Mail will:

1. sich über neu eingetroffene Post informieren,
2. diese lesen,
3. archivieren, um später erneut darauf zugreifen zu können,
4. und beantworten.
5. Eventuell will er noch zusätzliche Briefe schreiben.

Diese Aufgaben erfordern eine spezielle Unterstützung hinsichtlich der Interaktions- und Präsentationsformen. KHS erlaubt eine solche Anpassung an spezielle Anwendungen durch die Definition spezifischer Unit- und Linktypen (s.u., sowie Abb. 1) mit den ihnen jeweils zugeordneten Strukturierungs-, Präsentations- und Interaktionsregeln.

## **2.8 Einbindung externer Informationsquellen (bzw. Ziele)**

Hypertext läßt sich in der Tat theoretisch als ein zusammenhängendes globales Netz von Informationselementen begreifen, das sämtliche archivierte Information umschließt und Ursprung und Ziel aller informationsverarbeitenden und -erzeugenden Prozesse ist. In der Praxis wird es häufig der Fall sein, daß die zur Erledigung einer Aufgabe benötigte Information nicht in dem konkret als Datenobjekt verfügbaren Hypertext enthalten ist, bzw. nicht die gesamte Aufgabe von dem Hypertext-System unterstützt wird. Im ersteren Fall bietet sich die Integration zusätzlicher Information in den Hypertext an, im zweiten die Externalisierung von Hypertext-Teilen, die dann mit anderen Werkzeugen weiterverarbeitet werden können.

Auf den Austausch von Daten mit anderen Hypertext- bzw. Textverarbeitungssystemen wollen wir in diesem Kontext nicht weiter eingehen. Hier bietet sich die Verwendung einer gemeinsamen Dokumentensprache an. In diesem Zusammenhang sind Vorarbeiten zur Integration eines SGML-Parsers in das KHS erfolgt [KH92].

Besonders interessant erscheint uns jedoch die Einbindung von Informationsquellen in den Hypertext, die der Kontrolle des Nutzers nur bedingt unterliegen:

1. Das Einbinden von E-Mail scheint uns wichtig wegen der zahlreichen Verknüpfungen von E-Mail mit anderen Arbeitsabläufen. Durch das asynchrone Verhalten, E-Mail kann zu beliebigen Zeiten eintreffen, werden besondere Anforderungen an das Hypertext-System gestellt.
2. Ebenso wie E-Mail ist die Suche in Literaturdatenbanken in viele Tätigkeiten eingebunden, hier insbesondere das Publizieren, das zu den ausgewiesenen Zielanwendungen von Hypertext gehört. Eine besondere Anforderung ist hier, Datenbankauswahl und Queryformulierung in die normale Hypertext-Interaktion einzubinden, sowie die Rechercheergebnisse adäquat in den Hypertextkorpus einzufügen.
3. Die Suche in Datenbeständen anderer Institutionen und Einrichtungen, wie sie mit den Informationssystemen Gopher, WWW oder WAIS über das Internet möglich sind,

werden in KHS beispielhaft anhand einer Schnittstelle zu Gopher ermöglicht [AHR93], soll aber hier nicht weiter Gegenstand der Diskussion sein.

### **3 E-Mail mit KHS**

Die Integration von E-Mail in Hypertext berührt mehrere der oben erläuterten Kriterien für offene Hypertexte. Sie scheint uns deshalb besonders interessant und soll hier ausführlicher diskutiert werden.

1. E-Mail erfordert die Anbindung einer dem Hypertext externen Informationsquelle.
2. Sie stellt sowohl was die auftretenden Informationsobjekte, als auch die verwendeten Interaktionsstrategien angeht, einen eigenen Diskurs dar, der eine besondere Unterstützung durch das Hypertext-System verlangt.
3. Der asynchrone Charakter von E-Mail verlangt ein aktives Verhalten des Hypertext-Systems. Ankommende Mail sollte automatisch eingelesen, nach vergebenen Kriterien in eine Archivstruktur eingeordnet und dem Empfänger zur Lektüre angeboten werden.

Über die weiter oben im Text schon eingeführten, elementaren, im Umfeld von E-Mail anfallenden Interaktionsziele sind folgende weitere Aufgaben zu benennen, die durch ein Hypertext-System unterstützt werden können:

1. Verwalten von Adreßlisten. Diese sollten im Gegensatz zu reinen Alias-Listen auch für andere Fragestellungen zu benutzen sein.
2. Verwalten von Mailboxen. Hier unterscheidet KHS vollautomatische und redigierte Mailboxen. Erstere leitet die eintreffende Mail automatisch an einen gewissen Empfängerkreis weiter. Ein Interessent kann sich in die Empfängerliste eintragen (oder wieder aus ihr löschen), ohne daß ein Redakteur eingeschaltet werden muß. Redigierte Mailboxen leiten Nachrichten erst weiter, nachdem sie von einem Redakteur bestätigt wurden. Auch das An- bzw. Abmelden von der Interessentenliste bedarf der Bestätigung.

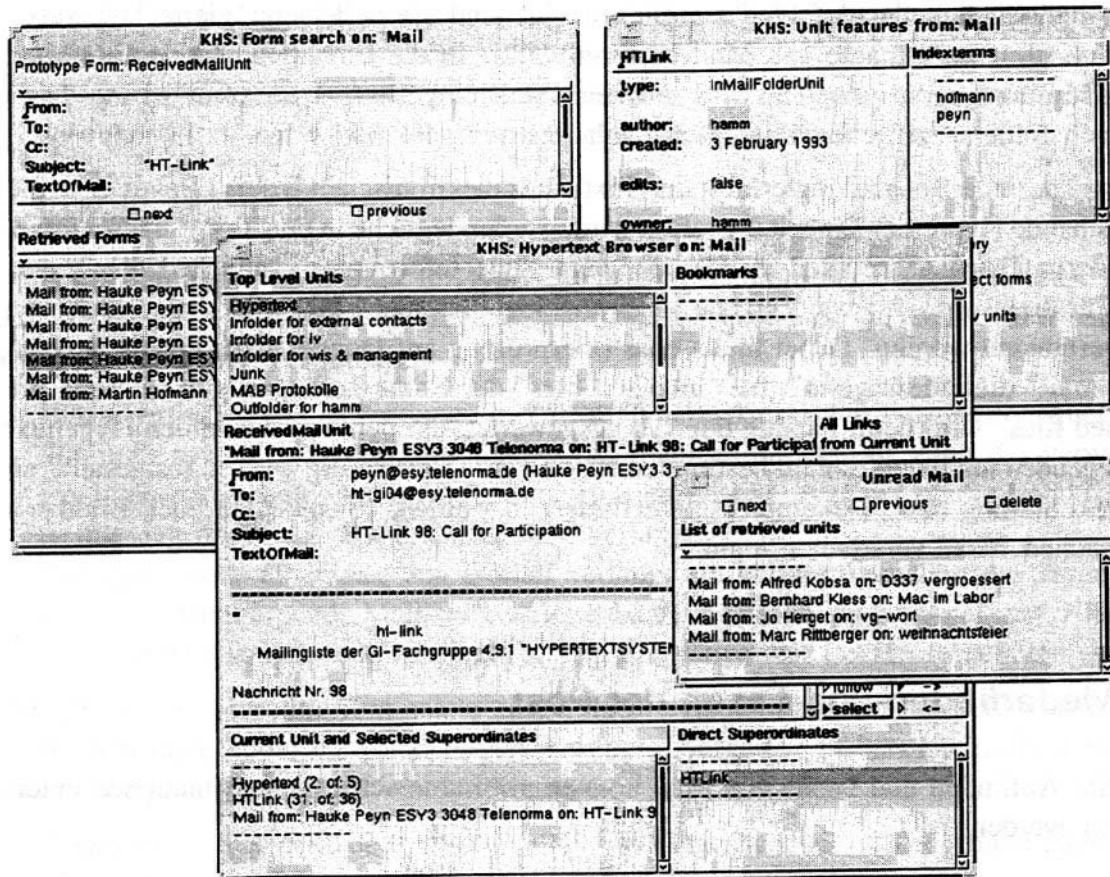
#### **3.1 Empfang von E-Mail**

Teil eines mail-orientierten Hypertexts ist ein aktiver Prozeß, der, wenn aktiviert, ständig auf das Eintreffen neuer Post achtet, diese einliest, parst und in den Hypertext einordnet. Elektronische Briefe sind sehr strukturiert, in besonderen Feldern sind Angaben über den Absender, die Adressaten und den Gegenstand des Briefs abgelegt, sowie ein identifizierendes Merkmal, die sogenannte "message-id". Diese Angaben sind für die spätere Verarbeitung von großer Bedeutung und müssen daher auch im Hypertext gezielt zugreifbar bleiben. Dementsprechend wird Mail in strukturierten Hypertext-Einheiten abgelegt, deren Felder Zugriff auf die oben genannten Merkmale und den Text der Mail erlauben.

Die Einordnung der Mail in den Hypertext geschieht anhand des Autorennamens und des Gegenstands (subject) der Mail. Aus dem Gegenstand der Mail wird, nach Entfernung der Stoppwörter, eine Menge von Schlagworten erzeugt, zu denen der Autorennamen hinzugefügt wird. Diese Schlagworte werden mit den Deskriptoren der im System befindlichen "mail folder" verglichen, das sind spezielle aggregierte Hypertext-Einheiten, die zur Aufnahme von Mail bestimmt sind. In die Einheiten, die den höchsten Übereinstimmungsgrad erzielen, wird die neue Mail eingefügt. Die Archivierungsdauer wird nach folderspezifischen Kassationsregeln ermittelt.

Eine spezielle Behandlung erfährt der Text eines elektronischen Briefs. Bevor er in die entsprechende Hypertext-Einheit eingefügt wird, werden jegliche in den Text eingeschlosse-

nen anderen Briefe ("included mail") oder beigefügte Information ("attached files") entfernt. Für diese werden jeweils eigene Units angelegt, die über Links mit dem eigentlichen Brief verbunden werden. Dabei findet eine Überprüfung auf Doppel statt, für die bei "included mail" die "message-id", bei "attached files" der Name herangezogen werden. Bei "attached files" wird beim Parsing der Typ ausgewertet, um den Typ der in den Hypertext einzufügenden medialen Einheit bestimmen zu können. Gleichzeitig wird so festgestellt, ob eine Dekodierung bzw. Dekompression erforderlich werden, so daß KHS auch Bilddaten als "attached files" verarbeiten kann.



**Abbildung 2** Die Abbildung zeigt eine typische Dialogsituation im Umgang mit E-Mail. Der Nutzer hat durch Formularsuche alle Briefe, die über HT-Link empfangen wurden, herausgesucht und einen davon in dem zentralen Browser zur Darstellung gebracht. Neben der textuellen Information enthält dieser Browser Angaben über die kontextuelle Einbindung der dargestellten Einheit in den Hypertext (In welchen Foldern ist sie enthalten, welche Links sind vorhanden). Ein Fenster zeigt Merkmale des geöffneten Folders, insbesondere die Schlagworte, die zur Einordnung der Mail führten. Ein weiteres Fenster gibt eine Liste der ungelesenen Briefe.

### 3.2 Wiederfinden und Lesen der Post

Beim Auffinden und Lesen von Mail können zwei unterschiedliche Situationen unterschieden werden:

- Ungelesene Mail (s.a. Abb. 2) kann aus einer Liste, die vom System angeboten wird, ausgewählt und sukzessiv gelesen werden. Mail, die nicht für eine dauerhafte Archivierung vorgesehen wird, kann dabei direkt gelöscht werden.
- Archivierte Mail kann anhand der polyhierarchischen Struktur der "mail folder", die es erlaubt, sich schrittweise der gewünschten Thematik anzunähern, aufgefunden werden, oder durch Formularsuche, nachdem die inhaltlichen Eigenschaften der gewünschten Hypertext-Einheit spezifiziert wurden. Darüber hinaus kann jedem



Folderelement ein Wiedervorlagedatum zugeordnet werden, zu dem es automatisch erneut der ungelesenen Mail beigelegt wird.

### 3.3 Versenden von Mail

Das Versenden von Mail geschieht durch Erzeugen einer neuen Hypertext-Einheit. Wird der Brief als Antwort auf einen im Hypertext befindlichen Brief verfaßt, werden die zugehörigen Hypertext-Einheiten durch ein Link verbunden. Die Adresse des Empfängers kann durch ein Kürzel spezifiziert werden, insofern im Hypertext eine Adreß-Einheit enthalten ist, die das jeweilige Kürzel als Merkmal enthält.

Zusätzlich können weitere Hypertext-Einheiten als "included mail" oder "attached files" beigelegt werden, indem sie mit entsprechenden Links an die zu versendende Einheit angeknüpft werden. Die in diesen Units enthaltene mediale Information wird beim Versand in den Text eingefügt. Bildinformation wird dabei kodiert und komprimiert.

## 4 Online Retrieval mit KHS

Um den in 2.8 genannten Aspekt der Offenheit von Hypertextsystemen zu vertiefen, wurde eine Komponente in das KHS eingefügt, die dem Benutzer eine Verbindung zu externen Datenbanken über elektronische Kommunikationsnetze erlaubt, wie sie in [PM89] vorgeschlagen und [Kuh91] diskutiert wurde. Betrachtet man den wissenschaftlichen Arbeitsplatz als einen Anwendungsbereich von KHS, so können als wesentliche elektronische Informationsquellen, neben der bereits beschriebenen E-Mail, drei Arten von externen Datenbanken genannt werden, die an einem wissenschaftlichen Arbeitsplatz von Nutzen sind.

1. Bibliothekskataloge: Die Bibliothek mit ihren Diensten ist als eine der Hauptquellen, die zur Deckung des Informationsbedarfs an einem wissenschaftlichen Arbeitsplatz dient, anzusehen. Zugangs- und Zugriffsmöglichkeiten der Bibliotheksbestände verbessern sich ständig. Nicht unwesentlich hat dazu sicherlich die zunehmende elektronische Verfügbarkeit der Kataloge öffentlicher Bibliotheken beigetragen. Für einen Bibliotheksbenutzer sind drei Gruppen von Bibliotheken von Interesse, die sich nach der physikalischen Nähe zum Benutzer unterscheiden lassen.
  - a. Zum einen die Bibliotheken, deren Bestand zwar abrufbar ist, aber nicht oder nur nach einer mehr oder weniger langen Wartezeit verfügbar ist (z.B. Fernleihe über einen Verbundkatalog).
  - b. Zum zweiten die Bibliotheken, deren Bestand abrufbar ist und auf den gegen eine Gebühr schnell zugegriffen werden kann (z.B. Ausleihe bei der Technischen Informationsbibliothek Hannover).
  - c. Zum dritten die Bibliotheken, die in unmittelbarer Reichweite des Benutzers liegen (z.B. die vor Ort liegende Universitäts- oder Fachbibliothek).
2. Online-Datenbanken: Seit Mitte der 60er Jahre werden Informationen über Online-Datenbanken angeboten. Während zunächst der Aufbau von bibliographischen Datenbanken und dann Faktendatenbanken unterstützt wurde [Hüg90], ist in den letzten Jahren besonders das Angebot an Wirtschaftsinformation überproportional gewachsen. Von allgemeinem Interesse für einen Wissenschaftler an seinem Arbeitsplatz sind jedoch eher die bibliographischen Datenbanken, welche die Verfügbarkeit internationaler Referateblätter garantieren. Daneben können noch fachgebietsspezifische Datenbanken mit bibliographischen, numerischen oder Fakteninformation eine Rolle spielen.

3. CD-ROM: Datenbanken auf CD-ROM finden zunehmend Absatzmärkte im Bereich der Wissenschaft. Die fest kalkulierbaren Kosten, das einfache Handling und die Verfügbarkeit multimedialer Daten sind die wichtigsten Faktoren, die für die Nutzung von CD-ROM sprechen. Im Interesse des Wissenschaftlers stehen, wie auch bei Online-Datenbanken, Referateblätter und fachgebietsspezifische Information.

Bei der Einbindung externer Datenbanken in KHS wurde das Hauptaugenmerk zunächst auf Online-Datenbanken gelegt, um einem potentiellen Nutzer den Zugang zu den Referateblättern zu ermöglichen. Der Ablauf der Online-Recherche aus einem Hypertext heraus entspricht der üblichen Vorgehensweise bei Recherchen in Online-Datenbanken, die sich in drei Teile gliedern läßt:

1. Die Vorbereitung der Recherche im KHS.
2. Die Durchführung der Recherche auf den externen Datenbanken.
3. Die Einbindung und Weiterverarbeitung der Ergebnisse in KHS.

#### **4.1 Die Recherchevorbereitung**

Wir möchten anhand eines Beispiels diese drei Punkte genauer beleuchten. Wir gehen von dem Szenario eines Informationswissenschaftlers aus, der mit dem KHS navigiert und nach Information aus dem Bereich 'Hypertext und Information Retrieval' sucht. Er gelangt an eine Texteinheit, für deren Inhalt er sich näher interessiert und beschließt, nach zusätzlicher Information zu den im Text angesprochenen Fragestellungen zu suchen, indem er eine Online-Recherche beginnt. Um das Retrieval vorzubereiten, werden drei Schritte durchlaufen:

1. Die Begriffe müssen ausgewählt werden, welche die Fragestellung repräsentieren. Um sich Suchworte zu definieren, hat der Benutzer mehrere Auswahlmöglichkeiten im Textteil einer Hypertext-Einheit:
  - a. Er kann ein oder mehrere Worte selektieren und die Selektion dann als einen Begriff für die Recherche verwenden. Somit stehen für die Recherche sowohl Einzelwortbegriffe, wie etwa 'Hypertext' als auch Mehrwortbegriffe (Bsp.: 'Information Retrieval') zur Verfügung.
  - b. Er kann einen Textbereich in der aktuellen Informationseinheit selektieren. Dieser wird zunächst in seine Einzelworte zerlegt und mit einer Stopwortliste verglichen, um sehr häufig vorkommende Worte zu eliminieren. Die verbliebenen Worte werden mithilfe eines Verfahrens nach [Kuh77] auf ihre Grundform reduziert. Jedes einzelne Wort kann dann als Begriff für die Recherche benutzt werden.
  - c. Die Schlagworte einer Hypertext-Einheit können für die Begriffsauswahl der Recherche verwendet werden, indem sie in der aktuellen Hypertext-Einheit selektiert und als Begriffe für die Recherche ausgewählt werden. Alle Schlagworte der Hypertext-Einheit stehen dann zur Weiterverarbeitung für die Recherche zur Verfügung.
2. In kommerziellen Datenbanksystemen wird fast ausschließlich mit Hilfe der Booleschen Logik gehorchenden Operatoren gesucht. Die ausgewählten Begriffe müssen daher so geordnet werden, daß eine Recherche mit diesen Operatoren möglich ist. Die in a. ausgewählten Begriffe werden zuerst an den Browser 'KHS: Online Recherche on: Konferenzen' (Abb.3) übergeben, welcher die Manipulation der Begriffe für die Boolesche Recherche ermöglicht. Sie werden dazu im linken Fenster des Browsers abgebildet. Dort können dann einer oder mehrere Begriffe selektiert werden, um sie zu einer Menge von Begriffen zu vereinigen, die durch 'or' verbunden

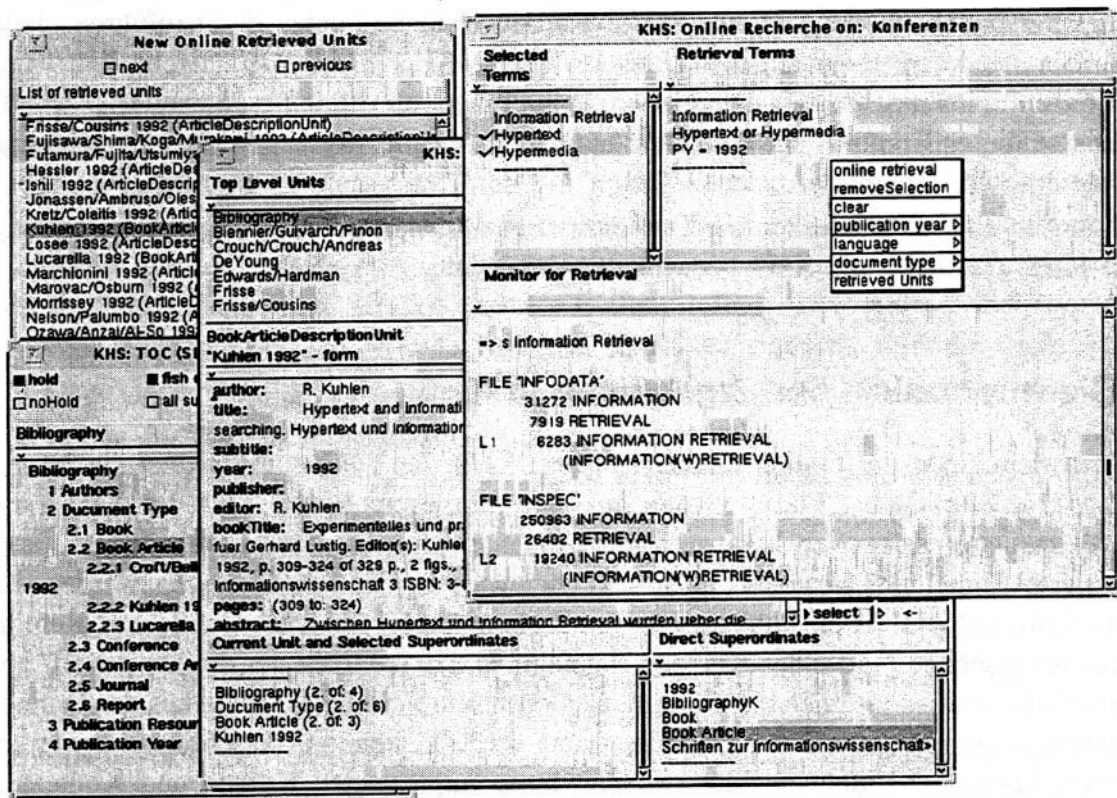
werden. Die entstehenden Begriffsblöcke werden im rechten Fenster dargestellt. Jede Zeile im oberen rechten Fenster des Browsers 'KHS: Online Recherche on: Konferenzen' repräsentiert somit eine Vereinigungsmenge mit einem oder mehreren Begriffen.

3. Zusätzlich zur inhaltlichen Suche können im rechten Fenster formale Parameter ausgewählt werden. Bei einer Einschränkung bezüglich des Publikationsjahres kann ein einzelnes Jahr oder mehrere Jahre spezifiziert werden. Der Benutzer kann die Originalsprache der Dokumentationseinheiten benennen und dabei zwischen verschiedenen Vorgaben, wie nur deutsch, nur englisch, englisch und französisch, etc. wählen oder die Sprache, die er wünscht, direkt eingeben. Weiterhin kann die Art der Publikation ausgewählt werden, so daß die Recherche beispielsweise auf Dissertationen und Reports eingeschränkt werden kann.

#### **4.2 Die Recherchedurchführung**

Nach der Auswahl und dem Ordnen der Suchbegriffe kann der Benutzer die Online-Recherche auf dem externen Host beginnen. Über den Verlauf der Recherche wird er im unteren Fenster des Browsers 'KHS: Online Recherche on: Konferenzen' informiert. Als erstes werden die vom Benutzer definierten Suchbegriffe in Retrievalbefehle konvertiert, die auf der Syntax der Retrievalsprache des verwendeten Hosts beruhen. KHS stellt dann eine Verbindung zum Online-Datenbankanbieter über öffentliche oder private Kommunikationsnetze her. Nach der Eingabe von Benutzernummer und Paßwort beginnt KHS mit der Datenbankauswahl und der Ausführung der Retrievalbefehle. Die Recherche wird auf dem Host STN in Karlsruhe durchgeführt, der über das Wissenschaftsnetz der Deutschen Forschungsgemeinschaft angewählt wird. Die Recherchedurchführung auf dem Host wird durch die Auswahl der Datenbanken, dem Bearbeiten der Fragestellung und dem Ausdruck der Ergebnisse beschrieben:

1. Für die Auswahl der relevanten Datenbanken sind verschiedene Verfahren denkbar. Für unser Beispiel wurden alle bibliographischen Datenbanken (ca. 90) auf dem Host STN zur Fragestellung 'hypertext and information retrieval' untersucht, um eine intellektuelle Auswahl zu unterstützen. Diese Fragestellung ergab eine Gesamtsumme von 832, nach Abzug der Dubletten 668 Treffern für alle Datenbanken. Für die Recherchen mit KHS wurden die bibliographischen Online-Datenbanken Compendex, Compuscience, Infodata und Inspec gewählt, die Nachweise aus dem Bereich Informatik, Informationswissenschaft und Ingenieurwissenschaften anbieten. Diese vier Datenbanken erzielten 699, nach der Dublettenkontrolle noch 559 Treffer, was in beiden Fällen 84% der maximalen Trefferanzahl entspricht. Somit ist durch diese Datenbanken eine ausreichende Abdeckung des Themengebiets im Bereich Hypertext und Information Retrieval gewährleistet.
2. Nach dem Aufbau der Verbindung werden die 4 Datenbanken gemeinsam angewählt, um eine datenbankübergreifende Suche auf den Datenbeständen durchzuführen. Jede einzelne Zeile im rechten Fenster des Browsers 'KHS: Online Recherche on: Konferenzen' wird bearbeitet. Die erzielten Antwortmengen werden miteinander geschnitten und dieses Ergebnis auf Dubletten untersucht. Damit werden die in verschiedenen Datenbanken doppelt gefundenen Treffer aus der Treffermenge eliminiert.
3. Zum Schluß werden die um die Dubletten reduzierten Treffer zur Weiterverarbeitung an KHS geschickt und KHS beendet die Verbindung zum Host.



**Abbildung 3** Die Abbildung zeigt die Situation nach Beendigung der Recherche. Im rechten Teil sieht man den Browser, der zur Steuerung der Recherche dient. Im linken Fenster sind die selektierten Begriffe, im rechten die Begriffsmengen und Optionen für die Recherche aufgelistet. Im unteren Fenster ist ein Mitschnitt aus der Online-Recherche zu sehen. Ein anderer Browser listet die recherchierten Einheiten auf, ein weiterer zeigt die inhaltliche Einordnung des Artikels als Buchartikel im Inhaltsverzeichnis der Einheit Bibliography. Der mittlere Browser enthält den Textteil einer neuen strukturierten Hypertext-Einheit. Des weiteren zeigt dieser Browser, wo sich die Einheit befindet und welchen Kontexten sie zugeordnet ist.

### 4.3 Die Integration der Ergebnisse in KHS

Nach dem Ende der Online-Recherche parst KHS die Treffer und ordnet sie in den aktuellen Hypertext ein. Dazu werden aus dem mitgespeicherten Text der Recherche die Dokumentationseinheiten herausgesucht und jeder Dokumentationseinheit eine neue, strukturierte Einheit im aktuellen Hypertext zugewiesen. In den Einheiten werden die Angaben für Autoren, Titel, Publikationsjahr, Dokumenttyp, Sprache, Quelle und Referat in dafür vorgesehene Felder des Knotens abgelegt, so daß jederzeit der direkte Zugriff auf die Einzelinformation möglich ist. Sofern vorhanden, wird die Indexierung aus der Online-Datenbank übernommen und für die Indexierung im KHS benutzt. Anschließend werden die neuen Hypertext-Einheiten unter verschiedenen aggregierten Knoten und Kontexten abgelegt:

1. Sie werden in eine alphabetische Liste der Autoren eingeordnet, wobei eine Hypertext-Einheit unter jedem Autorennachnamen wiederzufinden ist.
2. Sie werden nach Jahren geordnet, so daß alle Publikationen aus einem Jahr unter einem Knoten zusammengefaßt sind.
3. Hypertext-Einheiten mit der gleichen Sprache des Originaldokuments, auf das in der Referenz hingewiesen wird, werden aggregiert.
4. Jede recherchierte Hypertext-Einheit wird unter dem Dokumenttyp des Originaldokuments gespeichert, so daß beispielsweise alle Einheiten, die auf Dissertationen verweisen, in einem Knoten stehen.

5. Der Benutzer kann eine Liste mit den für ihn wichtigsten Publikationsorganen erstellen. Wurde in der Recherche eine Einheit gefunden, die einer der in der Liste der Publikationsorgane beschriebenen Quellen entstammt, so wird sie unter dem Namen der Quelle abgelegt. Werden etwa die 'Schriften zur Informationswissenschaft' des Universitätsverlags Konstanz aufgenommen, so wird bei den recherchierten Informationseinheiten geprüft, ob die dort nachgewiesene ISSN-Nummer derjenigen der 'Schriften zur Informationswissenschaft' entspricht. Ist das der Fall, so wird die recherchierte Informationseinheit in einen Knoten 'Schriften zur Informationswissenschaft' eingeordnet.
6. Jede Recherche wird nochmals extra abgespeichert. Die Frageformulierung wird dabei gemeinsam mit dem Datum der Recherche und dem Namen des Benutzers einer Hypertext-Einheit zugewiesen. Zusätzlich werden alle aktuell recherchierten Hypertext-Einheiten in diesem Kontext abgelegt, um somit jede Recherche dokumentieren zu können.

Ergebnisse aus Online-Recherchen entsprechen nicht immer in vollem Umfang den Erwartungen. Häufig werden auch Dokumentationseinheiten recherchiert, die der Fragestellung nicht gerecht werden. Hauptursache für diese Fehltreffer sind die ungenaue Vorstellung des Benutzers über sein Wissen [BOB82] und damit seine Fragestellung, die ungenaue Formulierung der Frage und die schlechte Repräsentation der Inhalte der Dokumente durch die Dokumentationseinheiten [Fuh90]. Um die nicht relevanten Ergebnisse löschen zu können, erhält der Benutzer nach Abschluß der Recherche eine Liste der durch die Recherche neu hinzugekommenen Hypertext-Einheiten. Er kann diese durchsehen und eine Relevanzentscheidung für jede einzelne Einheit treffen. Ist eine der neuen Hypertext-Einheiten nicht relevant, so wird sie aus dem Hypertext gelöscht.

Auch wenn Volltexte in zunehmendem Maße elektronisch lesbar zur Verfügung stehen, so ist die Menge der elektronisch verfügbaren Volltexte im Vergleich zur Menge der publizierten Volltexte doch verschwindend gering. Dem Benutzer stehen die Bibliotheken, kommerzielle Dokumentlieferanten, die sich häufig auf ein Gebiet spezialisieren, Buchhändler und Bibliotheksagenturen bei der Beschaffung von Originaldokumenten zur Auswahl [Oßw92 S.139 ff.]. Nach einer erfolgreichen Recherche in der Online-Datenbank kann KHS für den Benutzer den Standort einer Publikation in der aktuellen Bibliothek herausuchen, in unserem Falle der Universitätsbibliothek Konstanz, und den Standortnachweis im KHS eintragen. Ist die Publikationsquelle einer Veröffentlichung in der Bibliothek vorhanden (beispielsweise ein Proceedingsband), so wird sie in den Hypertext eingebunden und mit der Publikation (aus dem Proceedingsband) über einen Link (s. Abb.1 'Publication Source') verknüpft.

## 5 Diskussion und Ausblick

Ausgehend von den Anforderungen an ein offenes Hypertextsystem haben wir die Kriterien für die Offenheit anhand des Konstanzer-Hypertext-Systems erläutert. Dabei sind wir ausführlich auf die Konsequenzen für die Entwicklung eines Hypertextsystems, die sich aus dem Kriterienkatalog ergeben, eingegangen. Unter den Aspekten der Organisation und des Wiederfindens von Information ist die Einbindung von externen Informationsquellen in das KHS von besonderem Interesse. Wir haben in Abschnitt 3 und Abschnitt 4 die besonderen diskursiven Anforderungen und Interaktionsstrategien, die sich bei der Nutzung von E-Mail im KHS ergeben, kennengelernt und die Differenzierung der Ausdrucksmittel sowie die Anforderungen an das KHS als aktives Medium bei der Suche in externen Datenbanken betrachtet. Um den Mehrbenutzerzugang zu optimieren, Inkonsistenzen zu vermeiden und eine bessere Ausschöpfung der Ressourcen zu erreichen, wird in der näheren Zukunft die Datenverwaltungskomponente von KHS auf ein objektorientiertes Datenbanksystem

übertragen werden. Um die Interaktionsstrategien und aktiven Systemleistungen zu verbessern, ist der Aufbau einer auf einem Thesaurus basierten Wissensbasis geplant, die, bei Verfahren des traditionellen Information Retrieval, automatisch Suchhilfen für den Benutzer generiert und das KHS bei dem Aufbau automatischer Verknüpfungen zwischen Hypertext-Einheiten unterstützt. Nicht zuletzt werden wir die Einbindung externer Informationsquellen auf Grundlage der erzielten Ergebnisse im Hinblick auf weitere mögliche Anwendungen verbessern.

## Literatur

- [AHR93] R. Aßfalg, R. Hammwöhner und M. Rittberger: Linking Internet Capabilities within Hypertext. Eingereicht zum 'Seventeenth international online information meeting', London, 1993.
- [Beh92] M. Behrens: Navigation in Hypertext. Diplomarbeit, Informationswissenschaft, Universität Konstanz, 1992.
- [BOB82] N.J. Belkin, R.N. Oddy und H.M. Brooks: ASK for Information retrieval: Part I. Background and theory. *Journal of Documentation*, Vol. 38(2), Seite 61-71, 1982.
- [DHH+92] H. Davis, W. Hall, L. Heath, G. Hill und R. Wilkins: Towards an integrated Information environment with open hypermedia Systems. In: P. Lucarella, J. Nanard, M. Nanard und P. Paolini, Hrsg., *Proceedings of the ACM Conference on Hypertext*, Milano, Italy, 1992.
- [Fuh90] N. Fuhr: Hypertext und Information Retrieval. In: P.A. Gloor und N.A. Streitz, Hrsg., *Hypertext und Hypermedia. Von theoretischen Konzepten zur praktischen Anwendung*, Aus der Serie: *Informatik Fachberichte 249*, Seite 101-111. Berlin: Springer, 1990.
- [Gie91] M. Giesecke: *Der Buchdruck in der frühen Neuzeit. Eine historische Fallstudie über die Durchsetzung neuer Informations- und Kommunikationstechnologien*. Frankfurt a.M.: Suhrkamp, 1991.
- [Hew85] C. Hewitt: The challenge of open Systems. *Byte*, 5, Seite 223-242, 1985.
- [Hüg90] R. Hügel: *Der internationale Markt für Online-Datenbanken*. Aus der Serie: *Europäische Hochschulschriften. Reihe V. Volks- und Betriebswirtschaft. Bd. 1138*. Frankfurt a.M.: Peter Lang, 1990.
- [Kli92] P. Klingele: *Transformation von SGML in Hypertextstrukturen*. Diplomarbeit, Informationswissenschaft, Universität Konstanz, 1992.
- [Kuh77] R. Kuhlen: *Experimentelle Morphologie in der Informationswissenschaft*. Aus der Serie: *DGD-Schriftenreihe Band 5*. München: Dokumentation, 1977.
- [Kuh91] R. Kuhlen: *Aktivierung von Online-Informationsbanken aus Hypertextbasen*. In: S. Sorg, Hrsg., *Online '91. 14. Europäische Congressmesse für Technische Kommunikation. Congress IV. Bürokommunikation: Konzepte und Strategien zur Unterstützung der Büroarbeit*, Seite IV.01.01-IV.01.14, 1991.
- [Oßw92] A. Oßwald: *Dokumentlieferung im Zeitalter elektronischen Publizierens*. Aus der Serie: *Schriften zur Informationswissenschaft 5*. Universitätsverlag Konstanz, 1992.
- [PM89] M. Percival und N. Mac Morrow: *Evaluating the feasibility of using hypercard as an interface prototyping tool with reference to online Services: the impact of ISDN*. In: *Online Information 89. 13th International Online Information Meeting Proceedings*, Seite 265-276. Oxford: Learned Information, 1989.