

Digitale Bibliotheken im medizinischen Umfeld

Die Kinderumwelt gmbH (gemeinnützige Kinderumwelt GmbH (vormals DISA/DISU) in Osnabrück in Zusammenarbeit mit dem Berufsverband der Kinder- und Jugendärzte in Köln) betreibt ein FirstClass Kommunikationsnetz. Sie bietet deutschen Medizinerinnen und öffentlichen Gesundheitsämtern umfangreiche Informationsdienste. Das System, genannt UMINFO (siehe www.uminfo.de) beherbergt die unterschiedlichsten Fachbereiche auf einem einzigen Server.



Abbildung 1 UMINFO Website

UMINFO/ALLINFO stellt Ärzten, Apothekern, Ingenieuren, Wissenschaftlern und Krankenkassen sachbezogene Informationen und Konferenzen zum Thema Umweltmedizin zur Verfügung.

ÖGD Intranet wurde für die Mitglieder der Deutschen Gesundheitsämter eingerichtet. Es bietet spezielle Informationen im Bereich des öffentlichen Gesundheitswesens an. Dieses Netzwerk wird vom Robert-Koch-Institut (RKI) in Berlin unterstützt.

Pädinform ist ein Intranet für Kinderärzte. Innerhalb dieses Netzwerkes wird Wissen auf dem Gebiet der Pädiatrie ausgetauscht. Eine umfangreiche Sammlung von Informationen liegt in schriftlicher, digitalisierter Form vor. Zusätzlich werden Telekommunikationsdienste angeboten.

Das UMINFO System wird derzeit von ca. 3.700 medizinischen Anwendern genutzt.

Die Herausforderung

Vor kurzem beschloss die Kinderumwelt gGmbH ihr Serviceangebot um eine leistungsfähige digitale Bibliothek mit Wissensmanagementfunktionen zu erweitern. Die geforderten Funktionen stehen im FirstClass Produkt zunächst nicht in der gewünschten Form zur Verfügung.

Um produktive Vorteile im medizinischen Umfeld zu erreichen, müssen die UMINFO Anwender über Publikationen anderer Teilnehmer informiert sein und vollen Zugang zu existierenden Projekten, Forschungsergebnissen, Powerpoint Präsentationen, Bildmaterial und weiteren Veröffentlichungen erhalten. Meist existieren diese Informationen bereits und liegen als Mail-Anhänge oder als Dokumente auf Fileservern vor.

Dr. Otto, UMINFO Administrator, fasste die geforderten neuen Funktionen in folgenden drei Punkten zusammen:

- Die **pädinform** Benutzer sollten Zugang zu einzelnen Ausgaben der Fachzeitschrift „Kinder- und Jugendarzt“ in einem digitalen Archiv erhalten. Dieses Magazin steht bereits als PDF-Datei zur Verfügung. Über eine Volltextindizierung soll ein schnelles Auffinden einzelner Beiträge ermöglicht werden. Ausgefeilte Suchalgorithmen müssen unterstützt werden, um die Suchergebnisse auch in umfangreichen Archiven zu optimieren.
- Innerhalb des **pädinform** Intranets existiert bereits eine Vielzahl von Konferenzen. Diese Konferenzen beinhalten Mails mit sehr hilfreichen Dateianhängen (z.B. PDF, Powerpoint Folien und digitales Bildmaterial als JPEG-Dateien mit umfangreichen EXIF Meta-Informationen). Die **pädinform** Benutzer sollten die Möglichkeit erhalten, Informationen aus den vorliegenden Mails und deren Anhängen zu gewinnen. Dabei sollte die Suche und der Dokumentenzugriff die gleichen Mechanismen wie im projektierten Zeitschriftenarchiv wiederverwenden.
- Dieselben Anforderungen bestehen für einen dritten Anwendungsbereich: Das Robert-Koch-Institut unterhält eine spezielle Konferenz über Bio-Terrorismus, genannt: **IBBS** (Informationsstelle des Bundes für Biologische Sicherheit). In dieser Konferenz und all ihren Unterkonferenzen sind viele Dokumente nur als Anhang einer FirstClass Mail vorhanden. Die Benutzer der IBBS Konferenz sollten daher ebenfalls mit einem leistungsfähigen Volltextsuchwerkzeug ausgestattet werden.

Eine kurze Systemanalyse der o.g. Anforderungen führte zu folgender Liste neuer Funktionen, die dem bestehenden FirstClass System hinzugefügt werden sollten:

- Volltextrecherche nach Text- und Multimedia-Dokumenten außerhalb des FirstClass Systems.
- Volltextrecherche nach Dateianhängen von existierenden FirstClass Mails.
- Schnelle und präzise Suchresultate.
- Ausgefeilte Suchkonzepte wie z.B. boolesche Suche, Phrasensuche und Eigenschaftssuche.
- Minimaler Administrationsaufwand.
- Eine verständliche Benutzeroberfläche. Die Anwender kennen bereits die Bedienung des FirstClass Clients, das neue Suchsystem sollte in dieser gewohnten Art und Weise bedient werden.

Die Lösung

Dr. Otto fand eine passende FirstClass Erweiterung: FC MEMEX. Er beschloss diese bereits bestehende Lösung zur Umsetzung der formulierten Anforderungen zu nutzen.

FC MEMEX – eine fertige FirstClass Erweiterung

FC MEMEX ist ein Beispiel für die eindrucksvollen Möglichkeiten einer FirstClass RAD Applikation. RAD (Rapid Application Developer) ist das Synonym für die FirstClass Applikationsservertechnologie, die eine serverseitige Erweiterung nach Kundenwünschen ermöglicht.

In diesem Fall erlaubt eine Kombination aus RAD Applikationen und windowsspezifischen Diensten die Integration von einem oder mehreren Windows PCs als Bibliotheksserver. Für den FirstClass Administrator und die Anwender präsentiert sich dieses System als einzige riesige digitale Bibliothek.

Folgende UMINFO Hauptanforderungen konnten „out-of-the-Box“ realisiert werden:

- Schnelle Dokumentensuche. Ein Suchlauf in einer digitalen Bibliothek bestehend aus tausenden unterschiedlicher Dokumente wird in Sekundenbruchteilen ausgeführt.
- Vollautomatischer Betrieb.
- Unterstützt alle gängigen Standarddateiformate und ist erweiterbar.
- Erlaubt die Suche nach existierenden Dokumenten, die bereits auf Fileservern gespeichert sind.

Einstellung von Dokumenten in die digitalen Bibliotheken

Nach der serverseitigen Installation und Konfiguration konnte Dr. Otto sofort mit dem Einrichten der ersten digitalen Bibliothek des UMINFO Systems beginnen: Das „Kinder- und Jugendarzt“ Zeitschriftenarchiv.

Im diesem Fall standen bereits alle Dokumente als PDF-Dokumente auf einem dedizierten Fileserver zur Verfügung (Windows 2000 Workstation via LAN verbunden mit dem UMINFO FirstClass Server). Daher umfasste die Inbetriebnahme lediglich folgenden Aufgaben:

- Erstellung einer neuen Bibliothek mit der FC MEMEX Management Konsole, genannt „Katalog“.
- Konfiguration des Bibliotheksverzeichnis. Unter diesem Verzeichnis werden alle Dateiänderungen von FC MEMEX automatisch überwacht.
- Konfiguration der katalogspezifischen Benutzerrechte (neue Benutzer werden mit diesem Rechteset automatisch zur FC MEMEX Benutzerdatenbank hinzugefügt).
- Festlegung von katalogspezifischen Hilfstexten. Diese Texte erscheinen im FC MEMEX Suchformular und geben dem Benutzer hilfreiche Hinweise, wie man im vorliegenden Katalog sucht und welche möglichen Suchergebnisse erwartet werden können.
- Installation der FC MEMEX Suchapplikation in ausgewählten Konferenzen. Diese erfolgt über die reguläre RAD Applikationsinstallation mit wenigen Mausklicks.
- Das ist alles!

FC MEMEX durchsucht alle Dokumente im spezifizierten Katalog, extrahiert vorhandene Volltext- und Metadaten und erstellt bzw. aktualisiert eine zentrale Indexdatenbank. Danach ist der Katalog für alle folgenden Suchoperationen bereit und wird von FC MEMEX laufend aktualisiert.

Abb. 3 zeigt das endgültige Erscheinungsbild der UMINFO Konferenz nach der Installation von zwei FC MEMEX Suchapplikationen.

Eine zweite Möglichkeit zur Dokumentenablage in einer digitalen Bibliothek erfolgt über den FC MEMEX Browser. Diese Anwendung erlaubt einen Windows Explorer ähnlichen Zugriff auf das Dateisystem innerhalb eines FC MEMEX Dokumentenservers. Es erlaubt das Hinzufügen neuer Dokumente via Drag&Drop direkt im Browser Fenster. Dieses Werkzeug ermöglicht ebenfalls das Erstellen und Löschen von Dateien und Ordnern. Abb. 2 zeigt den Screenshot des PDF

Zeitschriftenkatalogs, geordnet nach Jahrgängen.

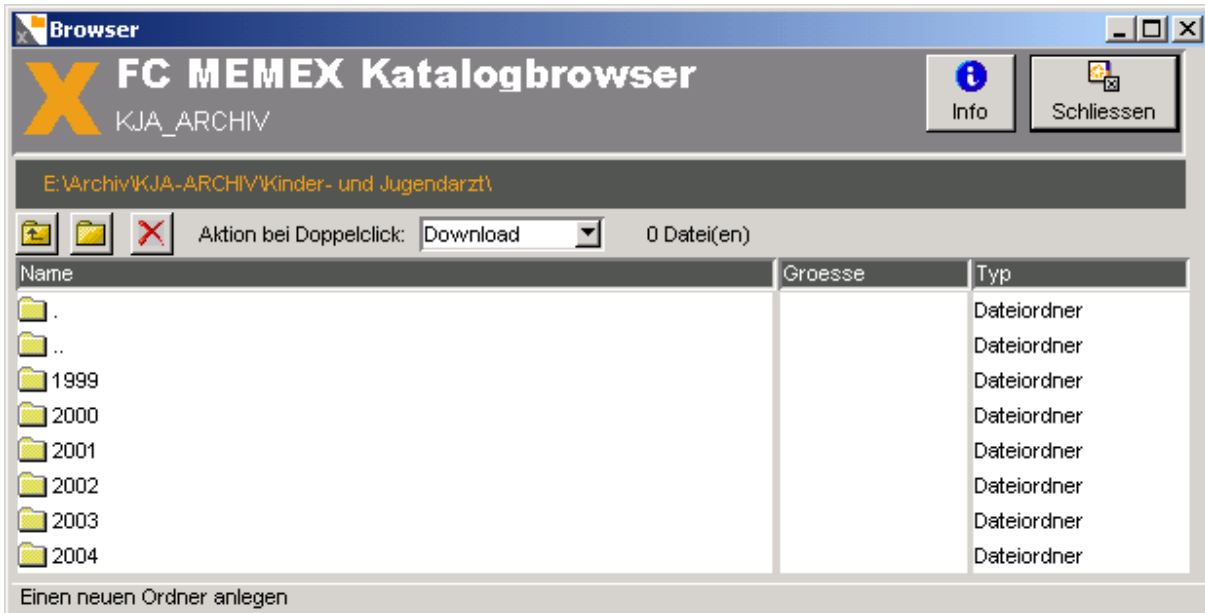


Abbildung 2 FC MEMEX Browser

Eine weitere Möglichkeit Dokumente in den FC MEMEX Katalog aufzunehmen, ist die Integration existierender FirstClass Konferenzen mit dem Programm FC ARCHIVE. Das UMINFO Team benutzt dieses Werkzeug, um den PaedArchiv Katalog mit Konferenzdaten zu füllen. Der Arbeitsablauf ist sehr einfach:

- FC ARCHIVE ist so konfiguriert, dass ausgewählte Konferenzen zweimal am Tag ausgelesen werden. FC ARCHIVE wird auf einem dedizierten FC MEMEX Dokumentenserver betrieben und verbindet sich zum entfernt liegenden FC Server via TCP/IP.
- Alle Konferenzinhalte werden nun in einem FC MEMEX überwachten Katalogordner abgelegt.
- Sowohl Mailtexte als auch alle Dateianhänge werden 1:1 ausgelesen und sofort volltextindiziert.
- Eine korrespondierende FC MEMEX Suchapplikation (PaedArchiv) ist in einer ausgewählten Konferenz installiert und wird vom medizinischem Anwenderkreis genutzt. Durch diese Suchapplikation werden alle von FC ARCHIVE gesicherten Daten für die PaedArchiv Benutzer verfügbar.

Wie wird FC MEMEX von den medizinischen Anwendern genutzt ?

Die folgenden Bildschirmaufnahmen zeigen eine Konferenz mit den erwähnten zwei FC MEMEX Suchanwendungen. Einige hilfreiche Dokumente und eine FC MEMEX Referenzkarte mit unterstützenden Hinweisen stehen für neue Benutzer im unteren Teil der Konferenz zur Verfügung. Im oberen Bereich befinden sich die Ikonen der installierten FC MEMEX Applikationen. Da der medizinische Anwenderkreis bereits über Erfahrungen im Umgang mit FirstClass Konferenzen und Nachrichten verfügen, ergaben sich keinerlei Anwendungsprobleme durch den Einsatz der neuen FC MEMEX Applikationen. Das System wurde schnell von den Benutzern akzeptiert. Die Anwendung der FC MEMEX Suche entspricht dem Öffnen einer Konferenz. Die UMINFO Benutzer mussten praktisch nichts neues hinzulernen.

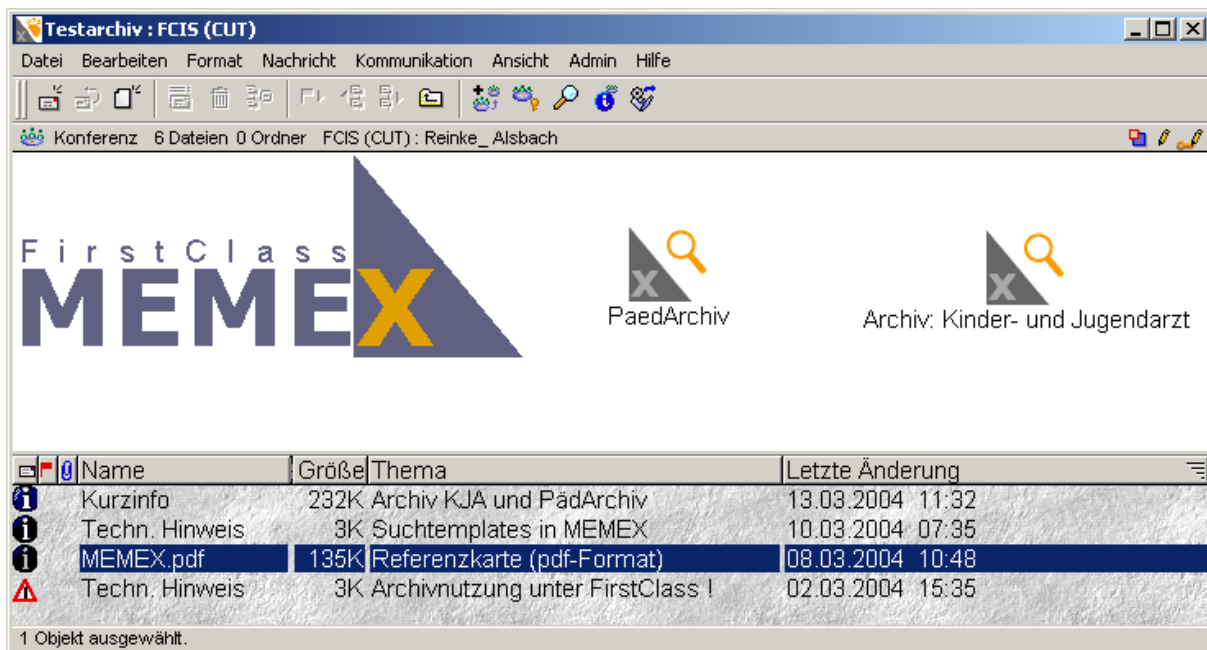


Abbildung 3 UMINFO Konferenz mit zwei FC MEMEX Suchapplikationen.

Nach dem Aktivieren der rechten Suchikone (Archiv: Kinder- und Jugendarzt) erscheint das FC MEMEX Suchformular. Es zeigt eine von der UMINFO Administration konfigurierte, katalogspezifische Benutzerschnittstelle zum Zeitschriftenarchiv „Kinder- und Jugendarzt“ (Abb. 4). Der Anwender kann in diesem Formular eine Freitextabfrage eingeben und die assoziierte Suche auslösen.

Speziell neue FC MEMEX Benutzer empfinden den zweiten Reiter in der Suchapplikation als sehr hilfreich: Die „Suchtemplates“ präsentieren eine Liste von bereits definierten Suchanfragen, die in einer domänenspezifischen Sprache formuliert und für alle Beteiligten leicht verständlich sind. Nach der Auswahl eines Suchtemplates und dem Anklicken des Such-Buttons wird eine korrespondierende Suche gestartet. Diese Suche erfolgt im aktuellen (und ggf. seit der Suchtemplatedefinition geänderten) Dokumentenbestand.

Der Vorteil dieser Methode ist, dass auch komplizierte Suchanfragen von Spezialisten vorab erstellt werden können und allen Benutzer in einer leicht verständlichen Sprache präsentiert werden.

Hinweis: Dieses Konzept ist dem brandneuen „smart folders“ Konzept, dass im kommenden Jahr von MacOS X Tiger eingeführt wird, sehr ähnlich. FC MEMEX präsentiert diese Funktion bereits heute und dies in einer wirklichen Client/Server Umgebung namens FirstClass!

Suche

MEMEX SUCHE
KJA_Archiv

Info Schliessen

KINDER-UND JUGENDARZT Powered by FC MEMEX

Standardsuche | Suchtemplates | Erweiterte Suche | Einstellungen

Geben Sie bitte in diesem Feld Ihre Freitextsuche ein

In dieser Datenbank sind die Jahrgänge 1999 - 2003 der Zeitschrift "Kinder- und Jugendarzt" archiviert.

Der Zugang zum KJA-Archiv ist nur in Verbindung mit einer Freischaltung für PädInform möglich.

Suchtip: Die Trefferzahl kann eingegrenzt werden, indem die Suchbegriffe mit AND verknüpft werden.
Beispiel: Impfung AND Hyposensibilisierung

+ Diese Suche zu den Templates hinzufügen

Suche...

Abbildung 4 Standardsuchformular mit einem Eingabefeld zur Definition von Freitextabfragen. Die führende Bitmap „Kinder- und Jugendarzt“ und der dargestellte Text in der Formularmitte wurde von Herrn Dr. Otto mit Hilfe des Katalog Editors nach gruppenspezifischen Anforderungen angepasst.

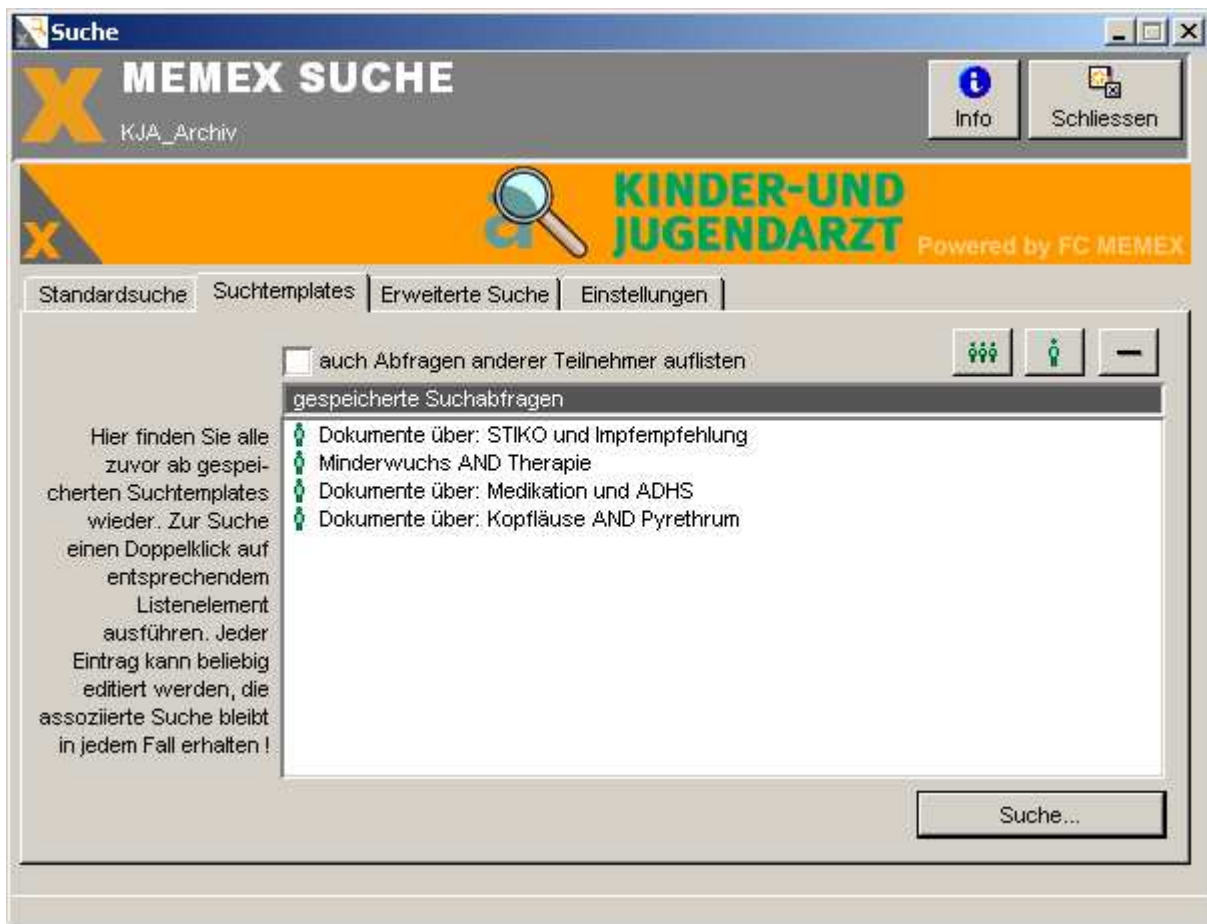


Abbildung 5 Das Suchtemplateformular zeigt eine Liste von bereits definierten Suchanfragen in einer domänenspezifischen Formulierung.

Das Ergebnis des Suchprozesses erscheint in der FC MEMEX Ergebnisliste. Dieses Formular zeigt gefundene Dokumente in Form einer sortierten Liste und bietet folgende Optionen und Informationen:

- Eine dokumentenspezifische Ikone.
- Dateiname und Datum der letzten Änderung.
- Einen automatisch generierte Zusammenfassung. Diese wird aus allen unterstützten Dokumenttypen (nahezu alle Formate wie Microsoft Office Dokumente, Adobe Acrobat PDF, XML, TXT, ...) extrahiert.
- Die Suchtreffergüte (1 ... 1000) und optional die Anzahl der direkten Suchtreffer.
- Weiterführende Befehlsoptionen:
 - Dokumentenvorschau
 - Dokumentendownload
 - Dokumentenanmerkungen

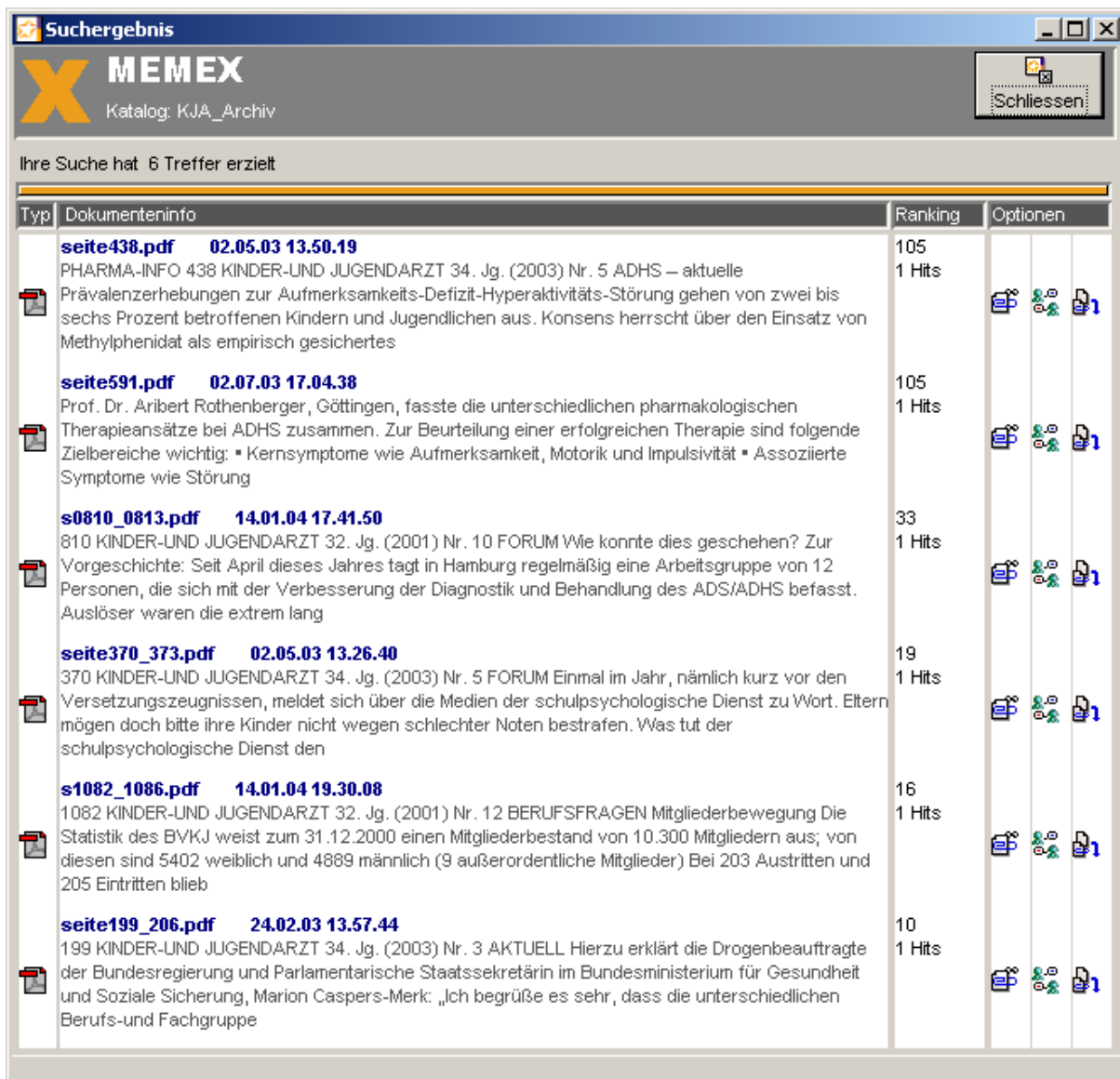


Abbildung 6 FC MEMEX Suchergebnisformular

Zusätzlich zu den bereits präsentierten Informationen innerhalb der Suchergebnisliste verfügt FC MEMEX über eine weitere Ebene zur Vorschau von serverseitigen Dokumenteninformationen. Das FC MEMEX Eigenschaftsformular erlaubt die umfassende Extraktion von Informationen aus dem ausgewählten Dokument, z.B.:

- Einen detaillierten Textauszug direkt aus dem Dokument. Die Länge des Textes kann in den FC MEMEX Benutzereinstellungen festgelegt werden. Mit dieser Textinformation können Benutzer den Dokumenteninhalt sichten und entscheiden, ob dies das gesuchte Dokument ist. Diese Bewertung findet vor einem möglichen Download des Originaldokumentes statt, welches möglicherweise eine Multi-Megabyte-Datei sein könnte.
- Im Falle von Bilddokumenten (JPG, PNG, GIF, TIFF, etc.) wird ein automatisch generierte, verkleinerte Darstellung (Thumbnail) angezeigt.
- Dokumentenattribute (Metainformation) wie Autor, Dateigröße, Datum der letzten Änderung, Titel, Schlagwörter, Kommentare, Betreff, etc. - tatsächlich alle Eigenschaften, welche im Microsoft Office Eigenschaften Dialog veränderbar sind. Nach diesen Eigenschaften kann selbstverständlich auch gesucht werden.
- Auch Bilddokumente mit EXIF Metainformationen werden unterstützt. Diese Informationen werden neben der Thumbnaildarstellung angezeigt. Dies erlaubt eine Suchanfrage wie z.B.: „zeige alle Bilddateien, die mit einer CANON Ixus Digitalkamera erstellt wurden“.
- Dokumentenbewertung. Anwender können einzelne Dokumente bewerten und bereits im Eigenschaftsformular alle bereits vorgenommenen Bewertungen sehen.

- Dokumentenspezifische Befehlsoptionen:
 - Download der serverseitigen Originaldatei. Startet den FirstClass Download der betroffenen Datei.
 - Öffnen des Dokumentenanmerkungsformulars. In diesem Formular werden Anmerkungen und Links eingegeben und verwaltet.
 - Öffnen des Benachrichtigungsformulars. Dieses erlaubt die Festlegung von individuellen Benachrichtigungen zum vorliegenden Dokument, z.B. Downloadbenachrichtigung.

Bitte beachten Sie, dass bisher noch kein Dokumentendownload stattgefunden hat. Alle Informationen über serverseitige Dokumente, in nahezu allen Dateiformaten, wurden in einer sehr sicheren Art und Weise nur durch den FirstClass Client präsentiert. Kein Byte an Information wurde bislang auf der lokalen Festplatte gespeichert.

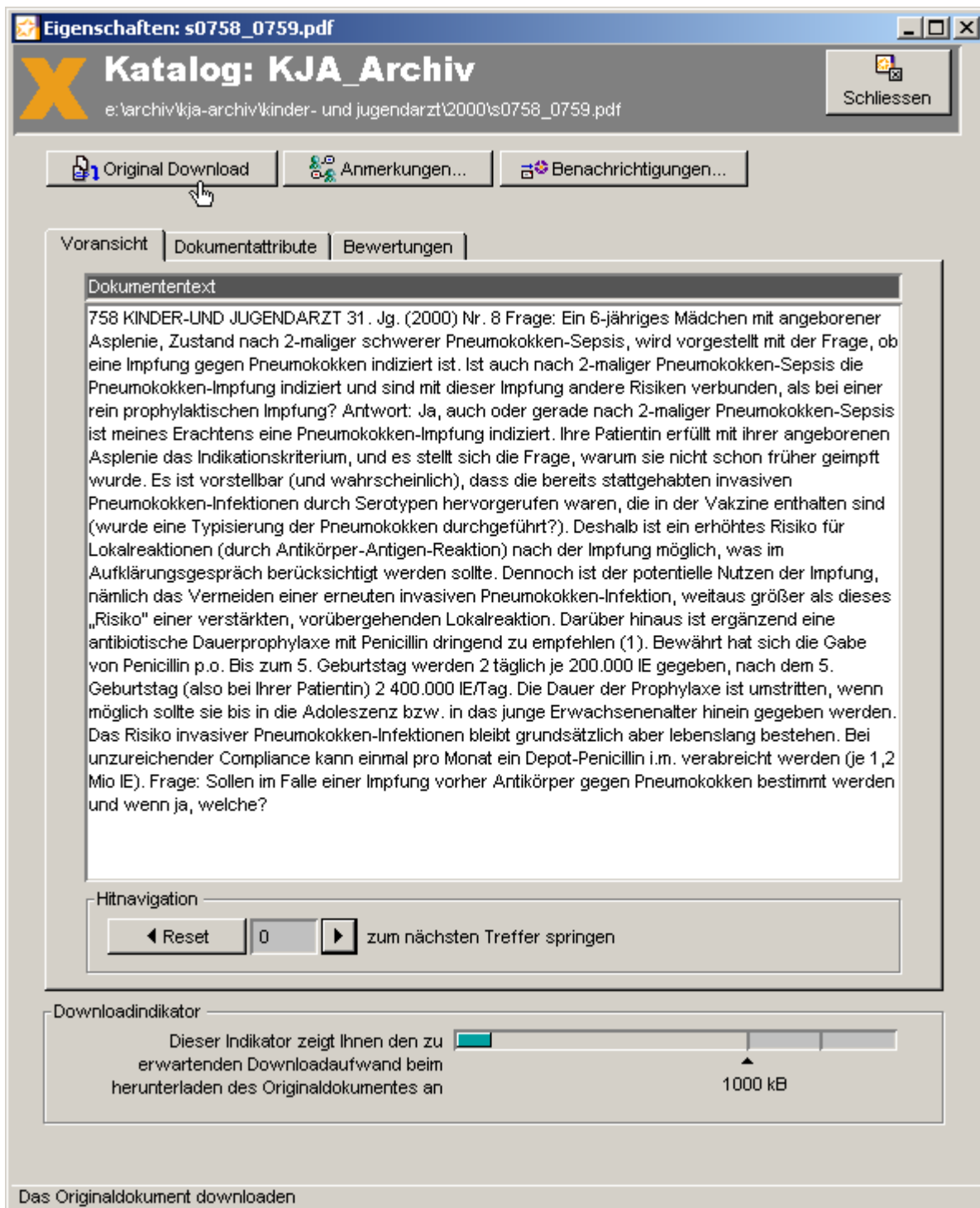


Abbildung 7 FC MEMEX Dokumentenvorschau eines zuvor gefundenen PDF Dokumentes. Die unterschiedlichen Reiter führen zu verschiedenen Vorschauabschnitten. „Voransicht“ zeigt den automatisch generierten Text aus dem Originaldokument. „Dokumentattribute“ enthält ein Formular mit weiteren dateispezifischen Eigenschaften. „Bewertungen“ erlaubt die Bewertung des vorliegenden Dokumentes.

Das Ergebnis

Seit der Einführung von FC MEMEX im laufenden UMINFO System haben Hunderte von Anwendern die Möglichkeiten der digitalen Bibliothek genutzt. Durch dieses System können sie Zeitschriftenartikel und ähnliche Informationen schnell und einfach finden. Sowohl Dateianhänge also auch Inhalte außerhalb des FirstClass Network Storage (früher: FirstClass

Post Office) befinden sich nicht länger in einer „Black-Box“ versteckter Informationen. Diese Daten rücken nun in die Reichweite von potentiell jedem FirstClass Anwender. Das war eine der wichtigsten UMINFO Anforderungen an eine digitale Bibliothek.

Der administrative Aufwand zur Pflege des laufenden FC MEMEX Systems ist sehr gering. Rückblickend fasst Dr. Otto zusammen, dass bereits mit der Einrichtung einer neuen FC MEMEX Suchapplikation die meisten administrativen Tätigkeiten erledigt sind. Alle neuen Katalogbenutzer erhalten im Augenblick der ersten Suchabfrage automatisch ihre standardisierten Zugangsberechtigungen. Alle Dokumentenänderungen, z.B. das Hinzufügen neuer Dateien in das Katalogverzeichnis, werden von FC MEMEX überwacht. Der FC MEMEX Volltextindex wird laufend, ohne weitere manuelle Eingriffe, aktualisiert.

Diese Erfahrungen mit FC MEMEX führten dazu, dass das UMINFO Team die neuen Dienste auch auf weitere Gebiete ausweiten konnte. Dr. Otto begann sofort nach Fertigstellung des Zeitschriftenarchivs mit der Einrichtung weiterer digitaler Bibliotheken innerhalb des UMINFO Systems. So können z.B. Anwender der RKI (Robert-Koch-Institut) - Bioterrorismuskonferenz in allen publizierten Veröffentlichungen zum Thema Bioterrorismus sowohl in Dokumenten als auch Anhängen (meist PDF und Powerpoint Dateien) via Volltextrecherche suchen. (Abb. 8).

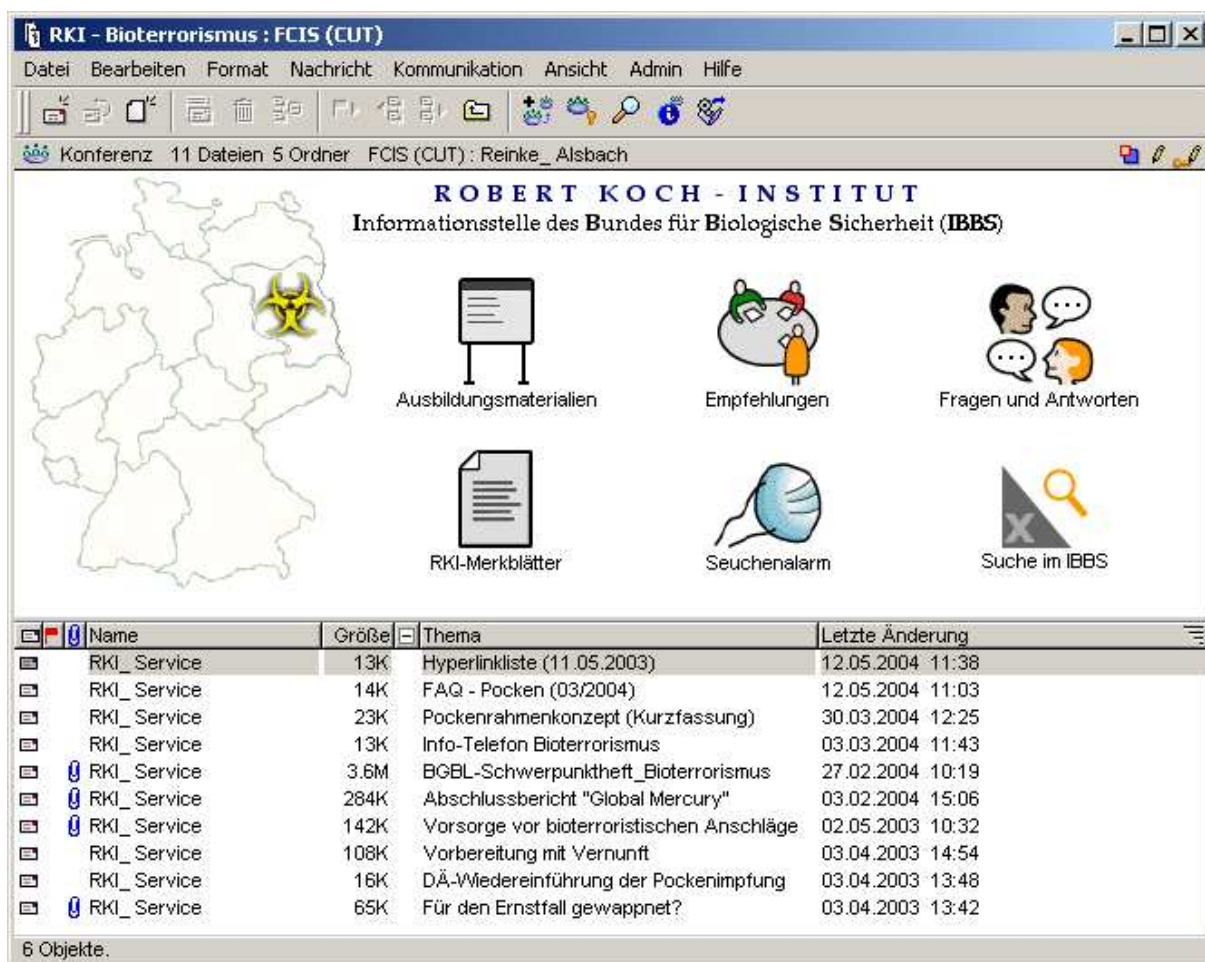


Abbildung 8 Nahtlose Integration der FC MEMEX Suche in der IBBS Bio-Terrorismus-Konferenz.

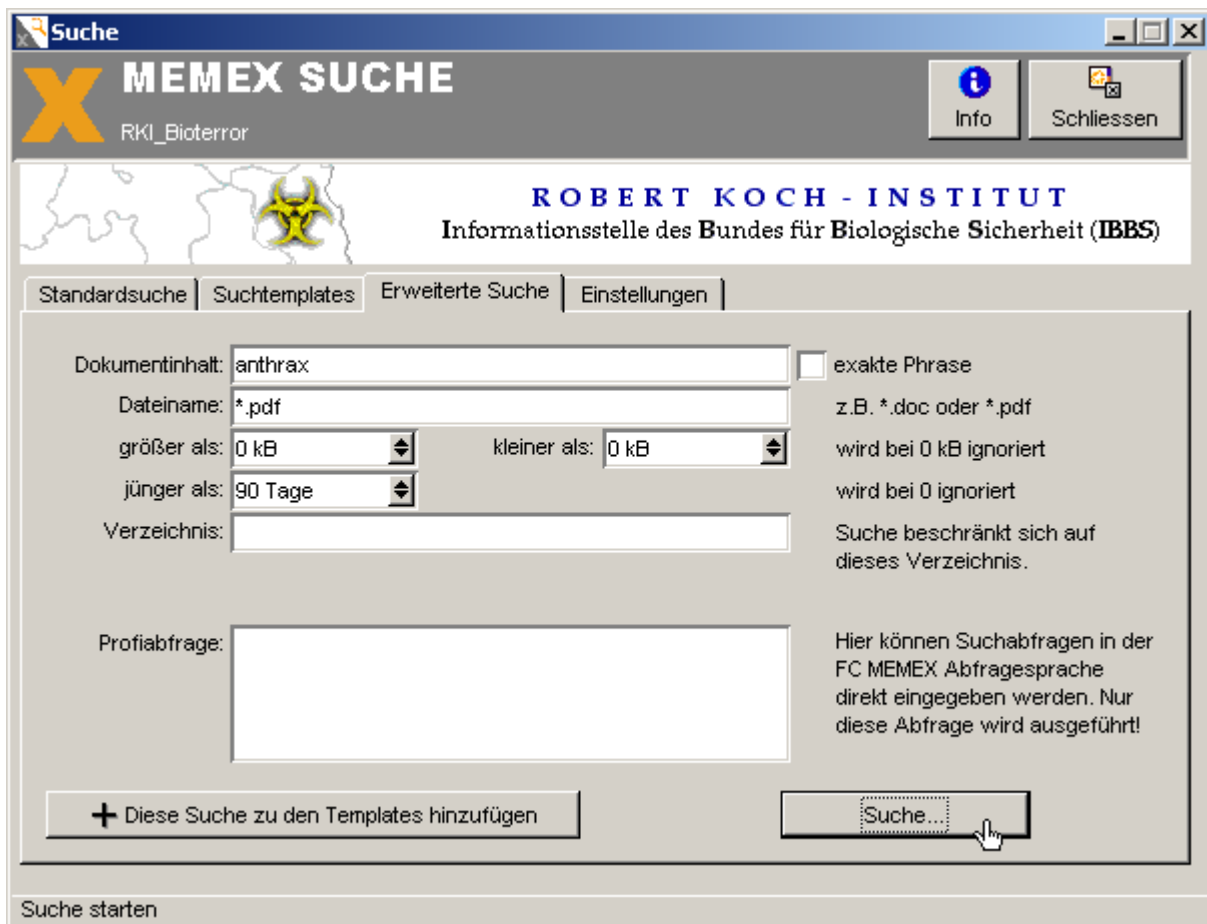


Abbildung 9 Komplexe Suchabfrage innerhalb der „Erweiterten Suche“. Diese Abfrage sucht nach allen PDF Dokumenten die in den vergangenen 90 Tagen hinzugefügt wurden und das Wort „Anthrax“ enthalten.

Im Vergleich zu anderen digitalen Bibliotheken (z.B. die komplette Bandbreite der Microsoft Server wie z.B. Sharepoint Portal Server) bietet FC MEMEX folgende Vorteile:

- Keine „Wegwerfen und Neuanfangen“ Lösung. UMINFO konnte das existierende FirstClass System ohne Änderungen weiterbetreiben. Es wurden keine FirstClass Softwarekomponenten ausgetauscht oder ersetzt. Dies ist ein wichtiger Aspekt, denn die Investition in neue Funktionalitäten ist nur auf die wirklich benötigten Module beschränkt.
- Keine Clientänderungen. Alle FC MEMEX Applikationen arbeiten im UMINFO System mit dem existierenden FirstClass Client. Mehr als 3.500 Benutzer verfügen über einen bereits installierten Client. Es wird keinerlei Aufwand zur Modifikation der existierenden Clientinfrastruktur verschwendet.
- Die Skalierbarkeit der FC MEMEX Bibliothek ist sehr überzeugend. Es funktioniert wie das Konstruieren mit Legosteinen: weitere Dokumentenserver können bei Bedarf an den zentralen FirstClass Server angekoppelt werden. Falls zukünftige Anforderungen mehr Such- und Zugriffperformance benötigen, können jederzeit neue Dokumentenserver hinzugefügt werden. Theoretisch ist es möglich, hunderte Dokumentenserver in einem einzige FC MEMEX System arbeiten zu lassen.
- Kompatibilität mit allen gängigen Dokumentformaten. FC MEMEX ist kompatibel mit vielen Dateiformaten: z.B. .doc, .wri, .rtf, .xls, .vsd, .pdf, .mp3, .jpg, .tif, .png, etc. Neue Dateiformate können auf der Basis einer „Plugin“-Architektur hinzugefügt werden.
- Die Abfragesprache ist sehr umfangreich. FC MEMEX unterstützt die komplette Bandbreite von der einfachen Volltextsuche bis hin zu komplizierten Eigenschafts- und Metasuchen. Neue Benutzer erzielen schnelle Sucherfolge durch Freitextsuchen, wie sie

bereits beispielsweise von Google bekannt sind. Erfahrene Benutzer optimieren die Treffergüte mit Hilfe komplexerer Abfragen.

Nach einigen Monaten Arbeiten mit FC MEMEX fasste Dr. Otto seine Erfahrungen wie folgt zusammen:

„Die Einführung von FC MEMEX war von großem Nutzen für unser UMINFO System. Es erweiterte und ergänzte die FirstClass Funktionalität mit einer sehr überzeugenden digitalen Bibliothek inklusive Wissensmanagementfunktionen ohne einen höheren administrativen Aufwand zu verursachen.“