

Vol 4, Issue 5
November 2004



Feature article

Case study: Digital libraries for medical professionals

courtesy of Michael Reinke, RSWE, Germany

Michael Reinke, founder of Reinke Software Engineering (RSWE), recently brought to our attention Kinderumwelt gGmbH, a FirstClass site that uses the FC MEMEX RAD application to create and manage a series of digital libraries. Here is their story.

The environment

Kinderumwelt gGmbH is a non-profit company with limited liability, fully owned by the German Academy of Pediatrics. It runs a large FirstClass site that provides services for German medical professionals and public health departments. The system, named UMINFO (see www.uminfo.de), hosts several specialized groups on a single FirstClass server.

This is the UMINFO website:

 [UMINFO website](http://www.uminfo.de)




This site includes the following major areas:

Uminfo/Allinfo offers environmental health information through conferences. It is used by medical doctors, pharmacists, engineers, scientists, and medical insurance organizations.

ÖGD Intranet, supported by the famous Robert-Koch Institute (RKI), Berlin, is targeted to members of the German public health departments and offers specialized information in this area.

Pädinform is an intranet for pediatricians. This section offers a large collection of FirstClass conferences that provide pediatric information and allow different types of communication, including telecommunications.

Today, UMINFO is used by approximately 4000 professionals.

The challenge

Recently, Kinderumwelt gGmbH decided to expand its services by adding advanced digital libraries and more knowledge management functions not currently available in the FirstClass product. In order to achieve medical breakthroughs, UMINFO's users must be aware of others' projects and publications and have full access to published papers, PowerPoint presentations, journals, and digital images. These documents exist mainly as attachments to conference messages or as documents inside the file systems of dedicated file servers.

Dr. Otto, administrator of the UMINFO system, summarized the required functionality in the following three areas:

- Pädinform users must be able to access single issues out of a digital archive of "Kinder- und Jugendarzt" magazines. These periodicals are stored as PDF files and should be fully text indexed for speedy retrieval. Advanced search algorithms should also be supported to narrow the search criteria.

- Many conferences exist within the pädinform intranet. These conferences contain many messages with attachments that provide useful information (for example, PDF, Powerpoint, and JPEG files containing valuable EXIF metainformation). Pädinform users must be able to search and retrieve information from these conferences, including attachments, by the same methods used in the periodical digital archive.

- The same requirements exist for a third application area: the Robert Koch Institute supports a special conference about bioterrorism called IBBS. Inside this conference and all of its subconferences, many documents are available only as attachments to FirstClass messages. Users of the IBBS conferences must have a straightforward full text search and retrieval system for these attachments.

A short system analysis of the above requirements resulted in the following list of functions that needed to be added to the existing FirstClass system:

- full text search and retrieval of text and multimedia documents outside the FirstClass network store
- full text search and retrieval of documents attached to conference messages
- fast and accurate search results
- advanced search capabilities like boolean, phrase, and property search
- minimal administration overhead
- an intelligible search user interface
- seamless integration with FirstClass, because all users already know how to use a FirstClass client.

The solution

FC MEMEX: An off-the-shelf FirstClass extension

FC MEMEX is an example of the robust features a FirstClass Rapid Application Developer (RAD) application can add to an existing FirstClass site. In this case, a combination of RAD applications and Windows-specific services supports the integration of multiple Windows PCs as document search and retrieval servers. To FirstClass administrators and users this system looks like one large digital library.

Let's review the major features required for the UMINFO system:

- fast document search and retrieval

Searching a digital library consisting of thousands of mixed media files must be done in fractions of a second.

- fully automatic operation
- support of standard file formats with the ability to be expanded to support other formats
- ability to search within existing documents that reside on file servers.

Getting documents into the libraries

After installing the FC MEMEX application on the FirstClass server, Dr. Otto was ready to configure the first digital library inside the UMINFO system: the "Kinder- und Jugendarzt" periodical library.

In this case, all documents (a collection of PDF files) had already been copied to a dedicated FC MEMEX document server (a Windows 2000 workstation attached to the FirstClass server by a LAN), so the setup consisted of the following tasks:

- 1 Define a new library (called a catalog) inside the FC MEMEX management console.
- 2 Configure the library's root directory.

Below this directory all file changes are observed by FC MEMEX.

- 3 Configure default new user rights.

Visiting users are added to the FC MEMEX user database automatically using this default set.

- 4 Define catalog-specific help text.

This text is displayed inside the search form to all visiting users and provides helpful information about how to search the catalog and what kinds of documents can be retrieved.

- 5 Install the associated FC MEMEX search application inside selected conferences.

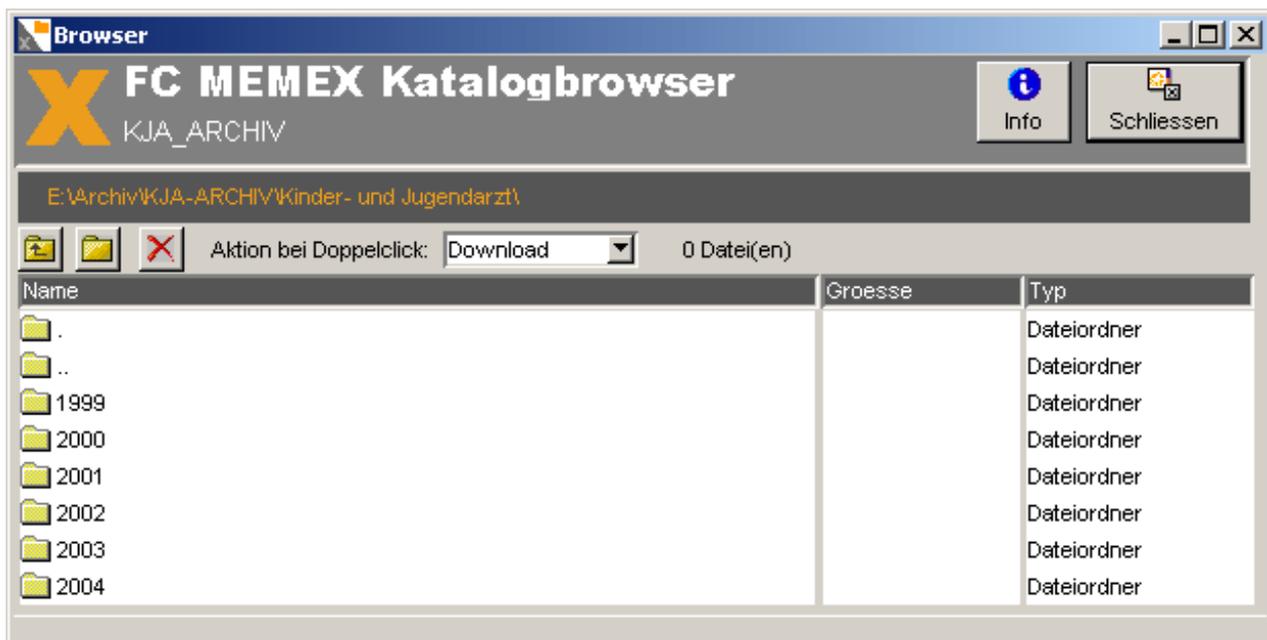
This is done through the standard RAD application installation procedure with a couple of mouse clicks.

That's it. FC MEMEX scans all documents inside this catalog, extracts full text and metadata, and updates a central index database. Now this catalog is ready for any search and retrieval operation.

A second way to add and organize documents inside the digital library is by using the provided FC MEMEX catalog browser. This application offers Windows Explorer-like remote access to the catalog server's file system and allows the addition of new documents using drag and drop onto the browser window. This tool also allows the deletion and creation of files and folders.

This is an FC MEMEX browser showing the PDF magazine catalog organized in annual folders:

[FC MEMEX browser](#)



Another way to get documents into a FirstClass MEMEX catalog is by monitoring existing FirstClass conferences with the archiving gateway FC ARCHIVE. The UMINFO team used this tool to set up the PaedArchiv catalog. The workflow used to populate this catalog was very simple:

1 Configure FC ARCHIVE to connect to the preselected conferences twice daily.

FC ARCHIVE runs on a dedicated FC MEMEX document server and connects to the remote FirstClass server using TCP/IP.

2 Move all conference content to an FC MEMEX observed catalog folder.

Body texts and all attachments are preserved and immediately fully text indexed.

3 Install a second FC MEMEX search application (PaedArchiv) in selected conferences for use by the medical professionals.

Through this search application all data fetched using FC ARCHIVE is available to PaedArchiv users.

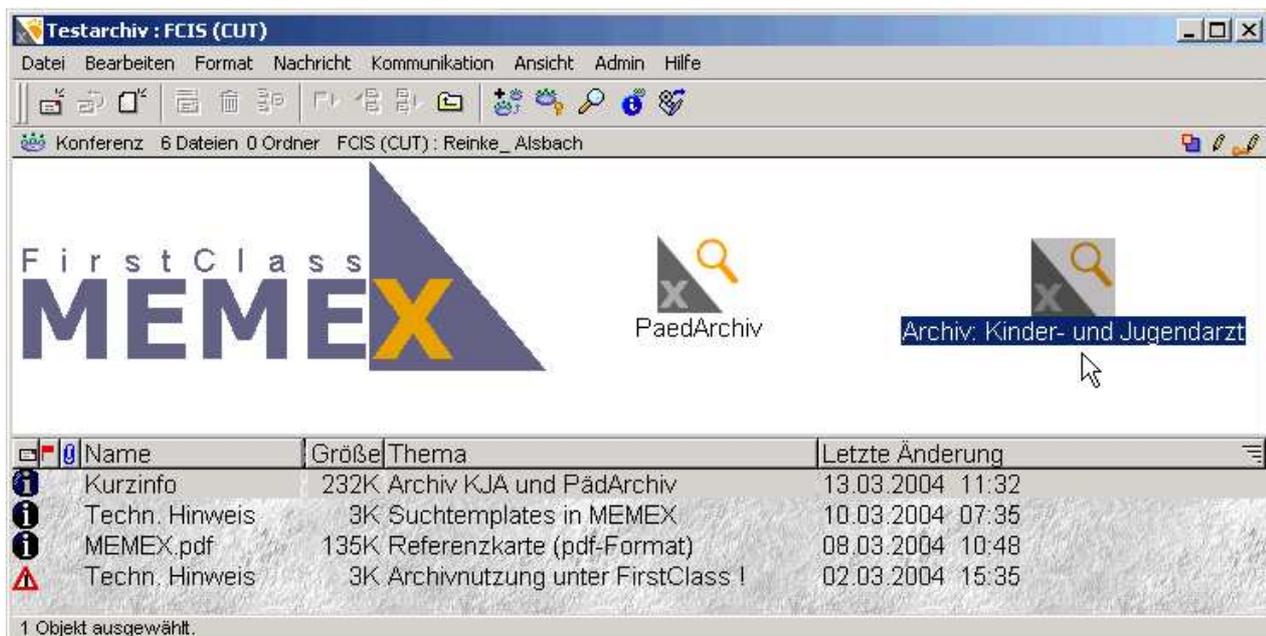
How medical professionals use FC MEMEX

The following images show a conference with the two above mentioned search applications. Some help documents and an FC MEMEX reference card are available for new users in the lower pane of the conference, while the upper pane contains the customized RAD application icons. Since users were already familiar with FirstClass conferences and mail objects, and launching an FC MEMEX search application is like opening a conference, the system was adopted quickly.

New FC MEMEX users find the second tab inside the search application especially helpful. The "Suchtemplates" form presents a list of predefined search queries in a domain-specific language that is understandable to all participants. After selecting a search template and pressing the search button, a search process is triggered and launched against the actual (and possibly updated) library content.

This is the UMINFO library conference:

[UMINFO conference](#)



If a user clicks Archiv:Kinder- und Jugendarzt, they are presented with the customized, catalog-specific interface to the "Kinder- und Jugendarzt" library. The user can then input a free text query and execute a search.

The advantage of this method is that highly complicated queries can be defined by search professionals and presented to all participants in a clear and easy-to-understand language.

Note

This concept is very similar to the new "smart folders" concept that will be introduced by Mac OS X Tiger (coming next year), but FirstClass offered this feature first with FC MEMEX -- and in a real client/server environment.

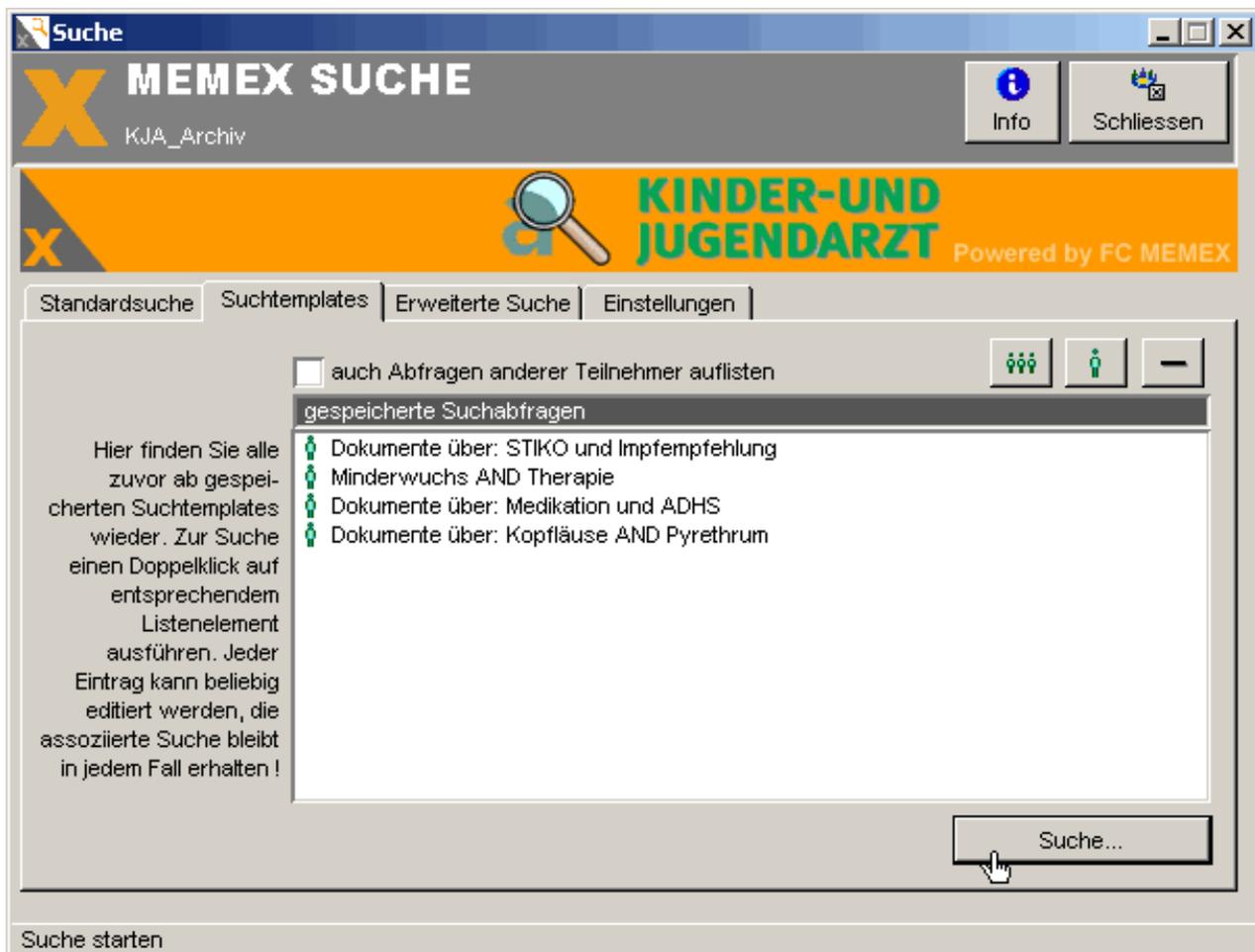
This is a standard search form with a single free text query input field on which Dr. Otto modified the leading text "Archiv: Kinder- und Jugendarzt" and the help text in the middle of the form using the FC MEMEX catalog editor:

[Search form](#)

The screenshot shows a web browser window titled "Suche". The main header area is orange and features a large "X" logo, the text "MEMEX SUCHE", and "KJA_Archiv". To the right are "Info" and "Schliessen" buttons. Below the header is a banner for "KINDER-UND JUGENDARZT" with a magnifying glass icon and "Powered by FC MEMEX". A navigation bar contains "Standardsuche", "Suchtemplates", "Erweiterte Suche", and "Einstellungen". The main content area has a search input field with the instruction "Geben Sie bitte in diesem Feld Ihre Freitextsuche ein". Below the field is explanatory text: "In dieser Datenbank sind die Jahrgänge 1999 - 2003 sowie die verfügbaren Ausgaben 2004 der Zeitschrift 'Kinder- und Jugendarzt' archiviert." and "Der Zugang zum KJA-Archiv ist nur in Verbindung mit einer Freischaltung für PädInform möglich." A search tip is provided: "Suchtip: Die Trefferzahl kann eingegrenzt werden, indem die Suchbegriffe mit AND verknüpft werden. Beispiel: Impfung AND Hyposensibilisierung". At the bottom are two buttons: "+ Diese Suche zu den Templates hinzufügen" and "Suche...". A status bar at the very bottom says "Suche starten".

This is a search templates form presenting a list of predefined queries in a domain-specific language:

[Search templates form](#)



A search produces a result list. The above form produced a ranked list of documents with the following options and information:

- a document-specific icon
- file name and date of last change
- automatically generated abstract text

This is extracted from all FC MEMEX supported document types (nearly every format, for example, Microsoft Office documents, Adobe Acrobat PDF, XML, TXT).

- the ranking value (1-1000) and optionally, the number of direct query term hits
- a selection of possible commands:
 - document preview
 - document download
 - document annotations

This is an FC MEMEX search result page:

[Search results](#)

Suchergebnis

MEMEX
Katalog: KJA_Archiv

Schliessen

Ihre Suche hat 6 Treffer erzielt

Typ	Dokumenteninfo	Ranking	Optionen
	seite438.pdf 02.05.03 13.50.19 PHARMA-INFO 438 KINDER-UND JUGENDARZT 34. Jg. (2003) Nr. 5 ADHS – aktuelle Prävalenzerhebungen zur Aufmerksamkeits-Defizit-Hyperaktivitäts-Störung gehen von zwei bis sechs Prozent betroffenen Kindern und Jugendlichen aus. Konsens herrscht über den Einsatz von Methylphenidat als empirisch gesichertes	105 1 Hits	  
	seite591.pdf 02.07.03 17.04.38 Prof. Dr. Aribert Rothenberger, Göttingen, fasste die unterschiedlichen pharmakologischen Therapieansätze bei ADHS zusammen. Zur Beurteilung einer erfolgreichen Therapie sind folgende Zielbereiche wichtig: • Kernsymptome wie Aufmerksamkeit, Motorik und Impulsivität • Assoziierte Symptome wie Störung	105 1 Hits	  
	s0810_0813.pdf 14.01.04 17.41.50 810 KINDER-UND JUGENDARZT 32. Jg. (2001) Nr. 10 FORUM Wie konnte dies geschehen? Zur Vorgeschichte: Seit April dieses Jahres tagt in Hamburg regelmäßig eine Arbeitsgruppe von 12 Personen, die sich mit der Verbesserung der Diagnostik und Behandlung des ADS/ADHS befasst. Auslöser waren die extrem lang	33 1 Hits	  
	seite370_373.pdf 02.05.03 13.26.40 370 KINDER-UND JUGENDARZT 34. Jg. (2003) Nr. 5 FORUM Einmal im Jahr, nämlich kurz vor den Versetzungszeugnissen, meldet sich über die Medien der schulpsychologische Dienst zu Wort. Eltern mögen doch bitte ihre Kinder nicht wegen schlechter Noten bestrafen. Was tut der schulpsychologische Dienst den	19 1 Hits	  
	s1082_1086.pdf 14.01.04 19.30.08 1082 KINDER-UND JUGENDARZT 32. Jg. (2001) Nr. 12 BERUFSFRAGEN Mitgliederbewegung Die Statistik des BVKJ weist zum 31.12.2000 einen Mitgliederbestand von 10.300 Mitgliedern aus; von diesen sind 5402 weiblich und 4889 männlich (9 außerordentliche Mitglieder) Bei 203 Austritten und 205 Eintritten blieb	16 1 Hits	  
	seite199_206.pdf 24.02.03 13.57.44 199 KINDER-UND JUGENDARZT 34. Jg. (2003) Nr. 3 AKTUELL Hierzu erklärt die Drogenbeauftragte der Bundesregierung und Parlamentarische Staatssekretärin im Bundesministerium für Gesundheit und Soziale Sicherung, Marion Caspers-Merk: „Ich begrüße es sehr, dass die unterschiedlichen Berufs- und Fachgruppe	10 1 Hits	  

In addition to the information provided in the search result list, FC MEMEX adds another layer of preview information. A document property form allows the user to extract as much information from the the selected document as possible, for example:

- a more detailed text summary retrieved directly from the document

The length of this summary can be modified on the user preference form. Users can decide if this is the document they are looking for before downloading a potentially large document.

- automatically generated thumbnails of any image documents (JPG, PNG, GIF, TIFF, etc.)
- document attributes (metainformation) such as author, file size, date of last change, title, keywords, comments, subject, etc. (all properties that are editable within the Microsoft Office properties dialogs, as these properties are also searchable)
- EXIF metainformation, if it exists, for image documents

This allows queries like: "give me all images made with a digital CANON Ixus camera".

- document rating

A user can rate this preview document or see all ratings already done.

- document-related commands:

- download original file

This starts the FirstClass download process.

- open the document annotation form

You can enter or retrieve annotations and links to other references belonging to this document.

- open the document notification form.

This form lists events that occur with the document, such as annotations and downloads.



Note

Notice that to this point no download action has been triggered; all information about any document in nearly any file format has been presented in a secure way through the FirstClass client. Not a single byte of information has been saved on the local hard disk.

This is an FC MEMEX document preview of a PDF document:

 [MEMEX preview](#)

Eigenschaften: seite438.pdf

X Katalog: KJA_Archiv Schliessen

e:\archiv\kja-archiv\kinder- und jugendarzt\2003\pdfdateien_kja_05_03\bogen6\seite438.p

Original Download Anmerkungen... Benachrichtigungen...

Voransicht Dokumentattribute Bewertungen

Dokumententext

PHARMA-INFO 438 KINDER-UND JUGENDARZT 34. Jg. (2003) Nr. 5 <<ADHS>> – aktuelle Prävalenzerhebungen zur Aufmerksamkeits-Defizit-Hyperaktivitäts-Störung gehen von zwei bis sechs Prozent betroffenen Kindern und Jugendlichen aus. Konsens herrscht über den Einsatz von Methylphenidat als empirisch gesichertes Medikament zur Behandlung der <<ADHS>>. Das zur Zulassung eingereichte mittellangwirksame Methylphenidat-Präparat der Firma Medice (Medikinet® Retard) ist dank einer Wirkdauer von ca. 8 Stunden optimal an deutsche Schulverhältnisse angepasst: Die zweistufige Freisetzungsdynamik erlaubt eine praxistaugliche morgendliche Einmalgabe, durch das Nachlassen der Wirkung in den Nachmittagsstunden sind Probleme wie Appetitlosigkeit oder Schlafstörungen nicht zu erwarten. Plazebokontrollierte multizentrische Studien belegen eine sichere Wirkung bei sehr guter Verträglichkeit. Die Zulassung von Medikinet Retard wurde kürzlich eingereicht. In Berlin fand im Oktober 2002 eine Konsensuskonferenz zur „Verbesserung der Versorgung von Kindern, Jugendlichen und Erwachsenen mit <<ADHS>>“ statt, an der neben dem einladenden Bundesministerium für Gesundheit und Soziale Sicherung auch alle betroffenen Arztgruppen sowie Lehrer- und Elternverbände teilnahmen. Das Fazit der Konferenz: Die Behandlung mit Methylphenidat gilt demnach als gesicherte empirische Therapie, die – eingebunden in ein multimodales Konzept – für die Betroffenen größtmögliche Therapieerfolge bringt. Bisher wurde Methylphenidat in Deutschland meist zweimal täglich eingenommen, wobei die zweite Einnahme mittags in der Schule oftmals vergessen wurde. Neue galenische Zubereitungen erlauben es, langzeitwirksame Methylphenidatpräparate durch die zweistufige Freisetzungsdynamik nur noch einmal am Tag zu geben. Eine Kapsel enthält 5 bzw. 10 mg schnell freisetzendes und 5 bzw. 10 mg langsam freisetzendes Methylphenidat in magensaftresistenten Pellets. Diese praxistaugliche morgendliche Einmalgabe fördert die Compliance der Patienten und betroffenen Eltern. Erfolgreiche Symptom-Reduktion, gute Verträglichkeit Das in zwei Dosierungen – 10 mg bzw. 20 mg

Hitnavigation

◀ Reset 335 ▶ zum nächsten Treffer springen

Downloadindikator

Dieser Indikator zeigt Ihnen den zu erwartenden Downloadaufwand beim herunterladen des Originaldokumentes an 1000 kB

The form tabs lead to different preview sections:

- "Voransicht" displays the automatically retrieved text from the original document.
- "Dokumentattribute" displays a form with many file-specific properties.
- "Bewertungen" displays the rating of this document.

End result

Several hundred medical professionals are now using this digital library. The system enables users to easily find periodical articles or similar information using a broad range of search capabilities. Attachment files and content outside the FirstClass network store are no longer

"black boxes" of hidden information. These items are now within reach of any FirstClass client, which was one of the most important UMINFO digital library requirements.

Administration overhead to keep the digital libraries running has been very low. Dr. Otto pointed out that most FC MEMEX administrative tasks were completed right after publishing the search applications inside the selected conferences. All new catalog users receive their default permission set automatically the first time they search these catalogs. All document changes (for example, the addition of new files into the catalog file system), are monitored by FC MEMEX and the text index database is updated without manual intervention.

The UMINFO team's experience with FC MEMEX has encouraged them to broaden their services to other areas. Dr. Otto started the creation of additional digital libraries inside the UMINFO system very soon after the initial periodical library. Now members of the Robert Koch Institut (RKI) bioterrorism conference can search and retrieve every attached document (mostly PDF and PowerPoint files) published in this area.

In comparison to other available digital library applications (for example, the complete range of Microsoft servers like Sharepoint Portal Server), the advantages of using the RAD-based FC MEMEX with FirstClass are:

- no "rip and replace" solution

FirstClass is still fully operable and does not require any replacement software. This is very important because the investment to get the new functionality is only focused on the new features.

- no client changes

All FC MEMEX applications running on the UMINFO system are accessible with the FirstClass client. Most of the 4000 users use the FirstClass client, so no effort is wasted to set up an application-specific client infrastructure.

- scalability of FC MEMEX digital libraries is very compelling

It's like building libraries with LEGO blocks: new document servers can easily be attached to a FirstClass server as needed. If future library requirements force more search-and-retrieval power, new document servers can be added anytime. Theoretically, a cluster consisting of hundreds of document servers can work inside a single FC MEMEX system.

- compatibility to common document formats covers most library applications

FC MEMEX is compatible with many file formats (.doc, .wri, .rtf, .xls, .txt, .html, .xml, .vsd, .pdf, .mp3, .jpg, .tif, .png, and so on). New file formats can be added using plugin architecture.

- query language is very powerful.

FC MEMEX supports a complete range of searches, from simple text searches to advanced property searches and metasearches. New users get instant results, just like doing a search using the well-known Google service. More advanced users can step ahead and use highly advanced query definitions.

After several months of working with the FC MEMEX system, Dr. Otto concluded, "The introduction of FC MEMEX brought a great added value to our UMINFO system. It extends the FirstClass base functionality with very compelling digital library and knowledge management functions without the need to increase the administration staff".

Finally, the integration of the FC MEMEX application underlines the suitability of FirstClass as a powerful and extendable communications platform.