

## Search Engine Deployment

### Requirements for Deployment:

1. At least 256MB RAM (Preferably 1GB or more)
2. Javascript supported browser (Preferably Firefox v3.6)
3. Operating System that supports Java
4. Java Run-Time Environment
5. Exist (Native XML Database) ver. 1.4.0 rev. 10440
6. Yui 2 module
7. Search Engine Application (e.g. MedEx)

### Installation for Deployment: (Windows)

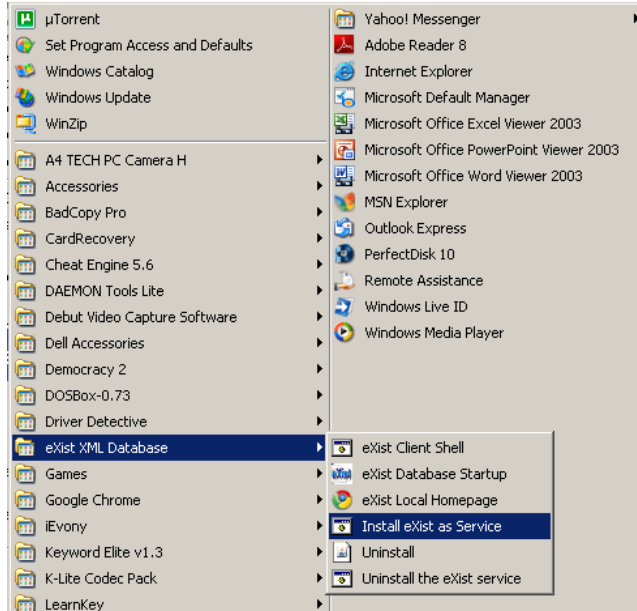
1. Install Firefox (download at [www.mozilla.com](http://www.mozilla.com))



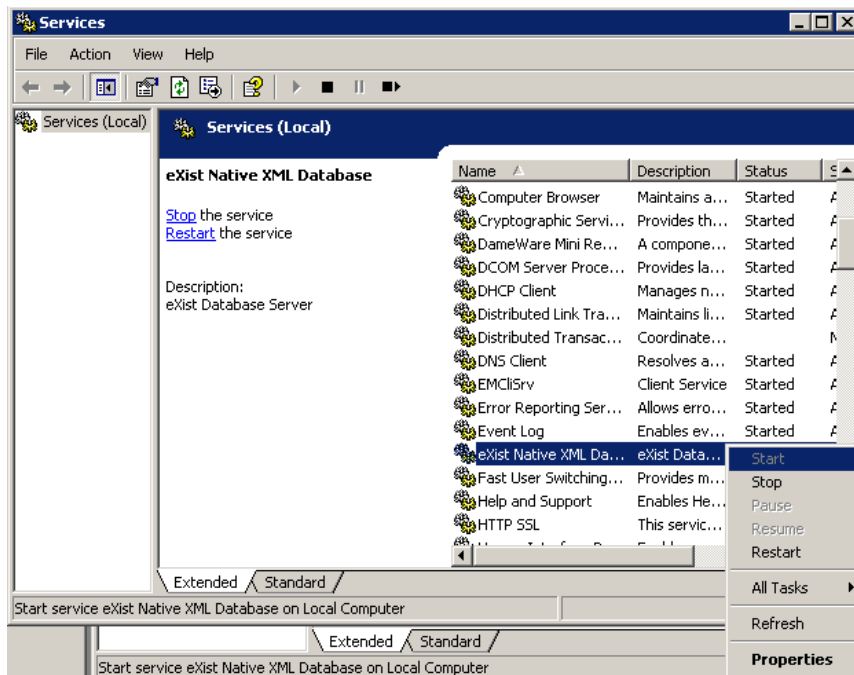
2. Install JRE for Java Support (download at [www.sun.com/java](http://www.sun.com/java))
3. Install Exist Native XML Database (download at [www.exist-db.org](http://www.exist-db.org))



4. Install Exist as Service (located at the start menu -> Exist)



5. Go to Control Panel -> Administrative Tools -> Services. Start the service Exist.



6. Open browser (Firefox)

7. Go to URL (<http://localhost:8080/exist>)



8. Go to Admin (located in the Left Tab)

Examples

XQuery Sandbox

XML Acronyms

Bibliographic

All Examples

Community

Mailing List/IRC

Development

Javadocs

ChangeLog

Bug Tracker

Subversion

Submit Patches

Administration

Admin

Server Status

Webstart Client

Launch

642.5K Lines

SOURCEFORGE.NET

Powered by eXist

- XML database specific: XMLDB, **XUpdate**, **XQ 1.0**

The **1.4** version adds a new full text index based on support for XProc. Most important, the XQuery engine is highly compliant with the **XQuery** standard (current XQTS score is 99.4%). The query engine is extensible and features a large collection of XQuery **Function Modules**.

eXist-db provides a powerful environment for the development of web applications. XQuery can be executed from the filesystem or stored in the database.

### Community Support

The eXist-db has a large community of users and developers. Getting help or posting bug reports or fixes is the preferred way to interact with the project.

Previous messages posted on the mailinglists are available in the archives. Additionally markmail.org provides a nice search interface.

The project has an IRC chat box #existdb which is accessible on Freenode. The channel is open for general discussions on eXist-db.

All eXist-db sources are freely available via a SourceForge repository and are continuously **monitored** and **built**.


### License

eXist is released to the public under the terms of the **GNU General Public License**.

9. Log in using your admin account and password

[illegible]

## 10. Go to Install Tools



```
<LINE>...all my ...you,</LINE>
<LINE>...so ...man ...together.</LINE>
<LINE>...the ...not .../LINE>
<LINE>...this is out of control!><br/>
<LINE>That ever I was born to set it right!</LINE>
<LINE>May ... come, let's go together.</LINE>
</SPECED>
```

## Select a Page

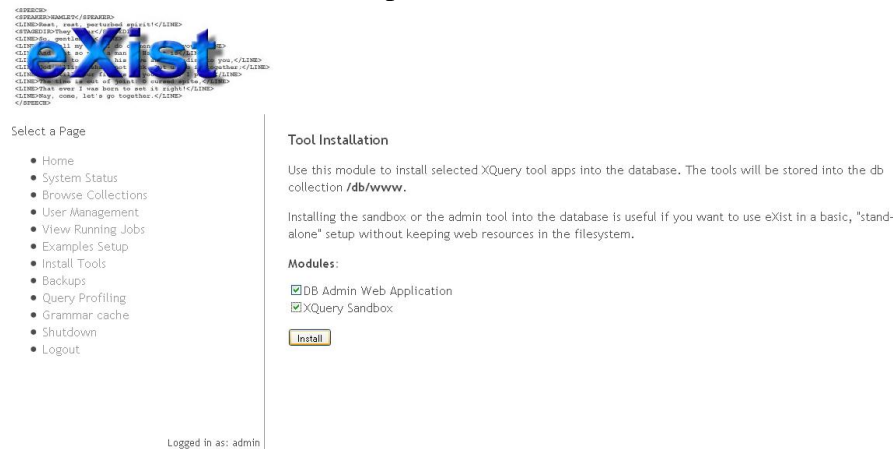
- [● Home](#)
- [● System Status](#)
- [● Browse Collections](#)
- [● User Management](#)
- [● View Running Jobs](#)
- [● Examples Setup](#)
- [● Install Tools](#)
- [● Backups](#)
- [● Query Profiling](#)
- [● Grammar cache](#)
- [● Shutdown](#)
- [● Logout](#)

Logged in as: admin

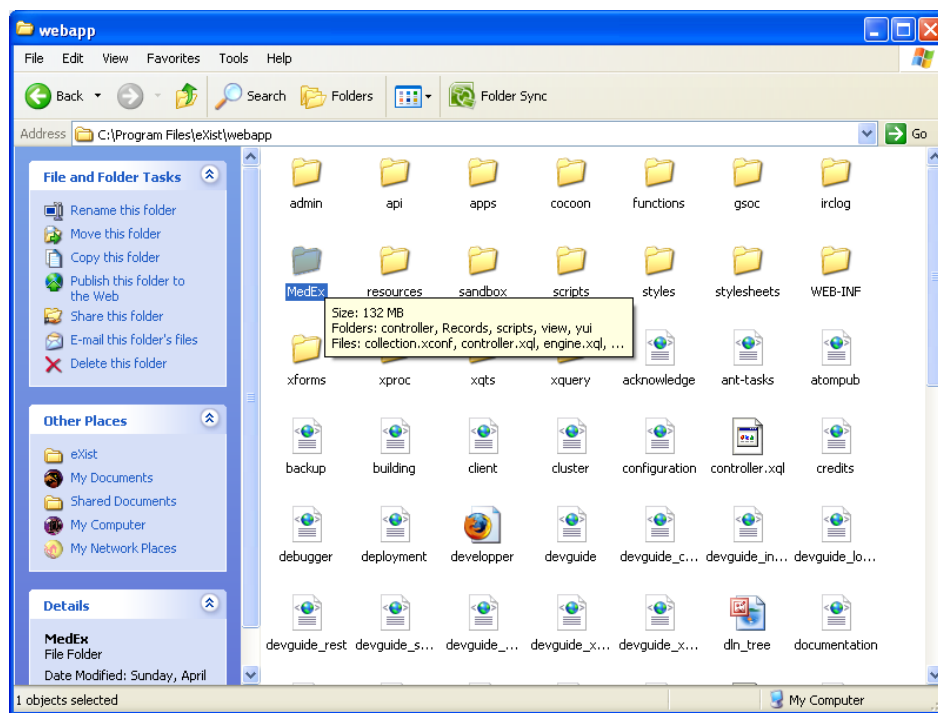
## System Status

General	
eXist Version:	1.4.0
eXist Build:	20091111
eXist Home:	C:\Program Files\eXist\tools\wrapper\bin\..\..
SVN Revision:	10440
Operating System:	Windows XP 5.1 x86
File encoding:	UTF-8
Java	
Vendor:	Sun Microsystems Inc.
Version:	1.6.0_18
Implementation:	Java HotSpot(TM) Client VM
Installation:	C:\Program Files\Java\jdk1.6.0_18\jre

## 11. Check both checkboxes then press Install



## 12. After installation, Go to the Home Folder of exist -> webapp -> Paste Deployment Folder here



## 13. Visit your Application Page (<http://localhost:8080/exist/FolderName>)

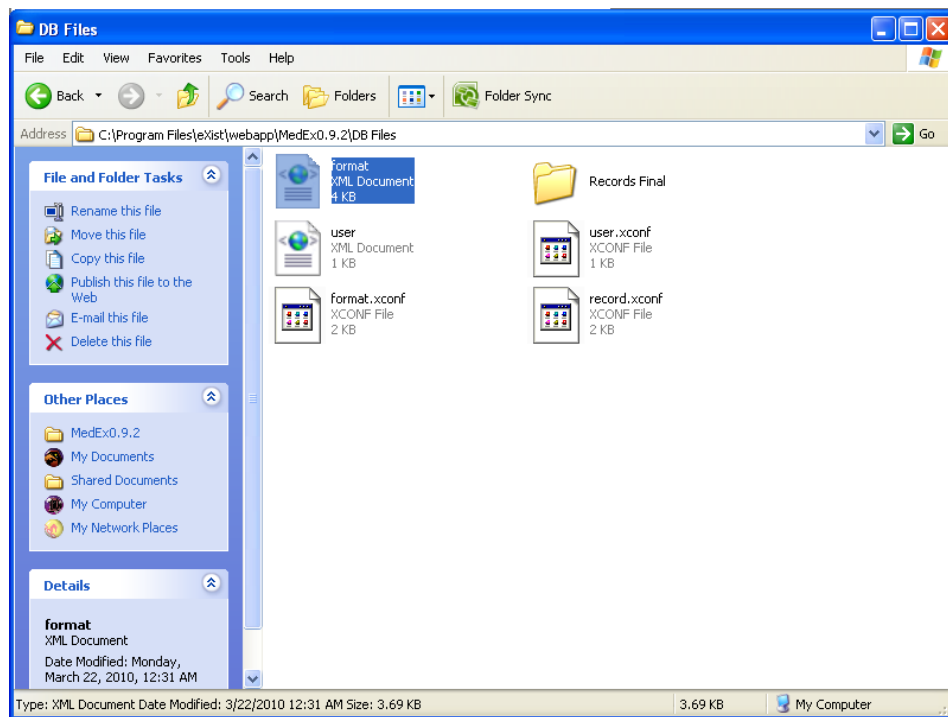
## Search Engine Application Modification:

### 1. Create a Format File for record format

-The format file is composed of sections where each section is an input field in the record (e.g. Name, Age)

Each Section is composed of Elements namely:

- a. name – name of the Field (string)
- b. display – if the field is going to be displayed in the results table (boolean)
- c. value – the value of the boost for the Lucene index (float)
- d. type – the data type of the field for the Range index (string-> “int”, “string”, etc)
- e. reqd – if the field is required to be filled up by the user. (boolean)
- f. single – if the field contains single or multiple values. This is for multiple values in an element (ex. <plans><plan>1</plan><plan>2</plan><plans>)  
- this is for creation of delimiters in the result table. (boolean)
- g. input – what is the input type in the form. Contains:
  - i. tag – Input type (e.g. text\_area, text, radio\_box)
  - ii. size – if input type is text\_area or text, the elements inside this element is <row> and <column> which contains the integer size. Else if the input type is a radio\_box, the element/s inside this element is <element> which contains the choices in the radio buttons.
- h. summary – how is the field summarized in the summary window. Contains:
  - i. summarize – set to true if the field is to be summarized after querying, false otherwise.
  - ii. level – if level 1, it is summarized based on the results. Else if level 2, it is summarized based on the referenced field. (int -> 1,2)
  - iii. reference – the referenced field (string). Its value is always a level 1 field. For sections with a level 1 summary, this is left blank.



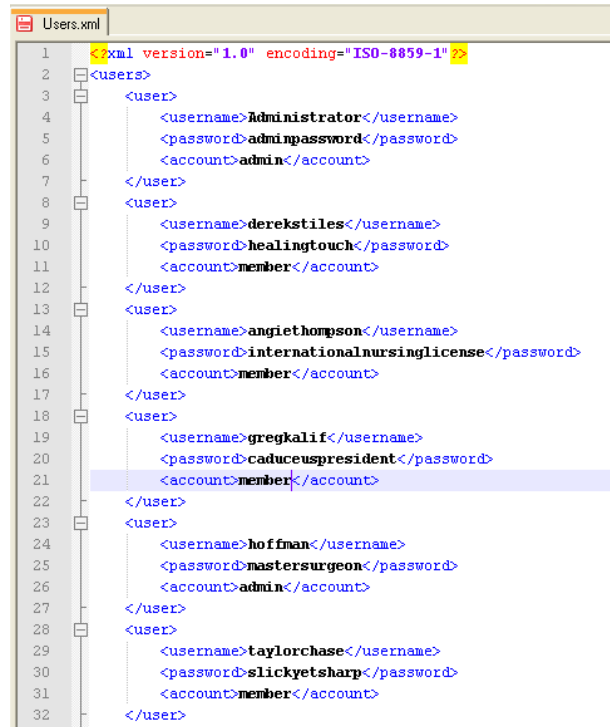
```
format.xml
1 <?xml version="1.0" encoding="ISO-8859-1" ?>
2 <Format>
3   <Section>
4     <name>id</name>
5     <display>true</display>
6     <value>1</value>
7     <type>int</type>
8     <reqd>false</reqd>
9     <single>true</single>
10    <input>
11      <tag>none</tag>
12    </input>
13    <summary>
14      <summarize>false</summarize>
15    </summary>
16  </Section>
17  <Section>
18    <name>case_number</name>
19    <display>true</display>
20    <value>1</value>
21    <type>int</type>
22    <reqd>false</reqd>
23    <single>true</single>
24    <input>
25      <tag>none</tag>
26    </input>
27    <summary>
28      <summarize>false</summarize>
29    </summary>
30  </Section>
31  <Section>
32    <name>service</name>
33    <display>true</display>
34    <value>1</value>
35    <type>string</type>
36    <reqd>false</reqd>
37    <single>true</single>
```

2. Create a User File for user authentication

-The file contains the information for each user in the application.

The file contains of <user> tags that is composed of:

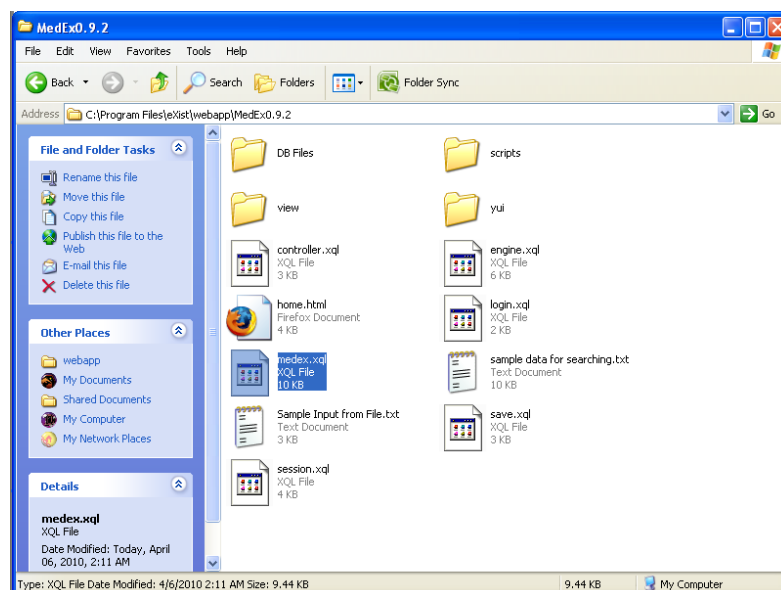
- username – the username of the user (string)
- password – the password of that account (string)
- account – the membership type of the account (string)



```
<?xml version="1.0" encoding="ISO-8859-1" ?>
<users>
  <user>
    <username>Administrator</username>
    <password>adminpassword</password>
    <account>admin</account>
  </user>
  <user>
    <username>derekstiles</username>
    <password>healingtouch</password>
    <account>member</account>
  </user>
  <user>
    <username>angie.thompson</username>
    <password>internationalnursinglicense</password>
    <account>member</account>
  </user>
  <user>
    <username>gregkalif</username>
    <password>caduceuspresident</password>
    <account>member</account>
  </user>
  <user>
    <username>hoffman</username>
    <password>mastersurgeon</password>
    <account>admin</account>
  </user>
  <user>
    <username>taylorchase</username>
    <password>slickyetsharp</password>
    <account>member</account>
  </user>

```

3. Edit the medex.xql line 24 for collection query restriction of the application (example: “/db/pgh”)





```

medex.xql
1 xquery version "1.0";
2
3 (:~
4     Displays the main page of the application.
5     This is also where the main collection that will be used is stored.
6     A function generates all the fields based from the format given
7 :~)
8
9 declare namespace MedEx="http://localhost:8080/exist/MedEx";
10 declare namespace system="http://exist-db.org/xquery/system";
11 import module namespace util="http://exist-db.org/xquery/util";
12 import module namespace request="http://exist-db.org/xquery/request";
13 import module namespace session="http://exist-db.org/xquery/session";
14 import module namespace xdb="http://exist-db.org/xquery/xmlldb";
15
16 declare option exist:serialize "method=xml media-type=text/xml omit-xml-declaration=no indent=no";
17
18 (:~
19     Generates all the fields in the form depending on the format given.
20     It supports text fields, text areas and radio buttons
21 :~)
22 declare function MedEx:get-inputs($form-number as xs:int?) as element(){
23
24     let $null := session:set-attribute('collection', "/db/pgh")
25     let $sequence := collection(session:get-attribute('collection'))//Format/Section
26     let $null := session:set-attribute("format", $sequence)
27     return
28     <form action="#" id="form{$form-number}" name="myform{$form-number}">
29     <div>
30         (if($form-number eq 2) then
31         <div>

```

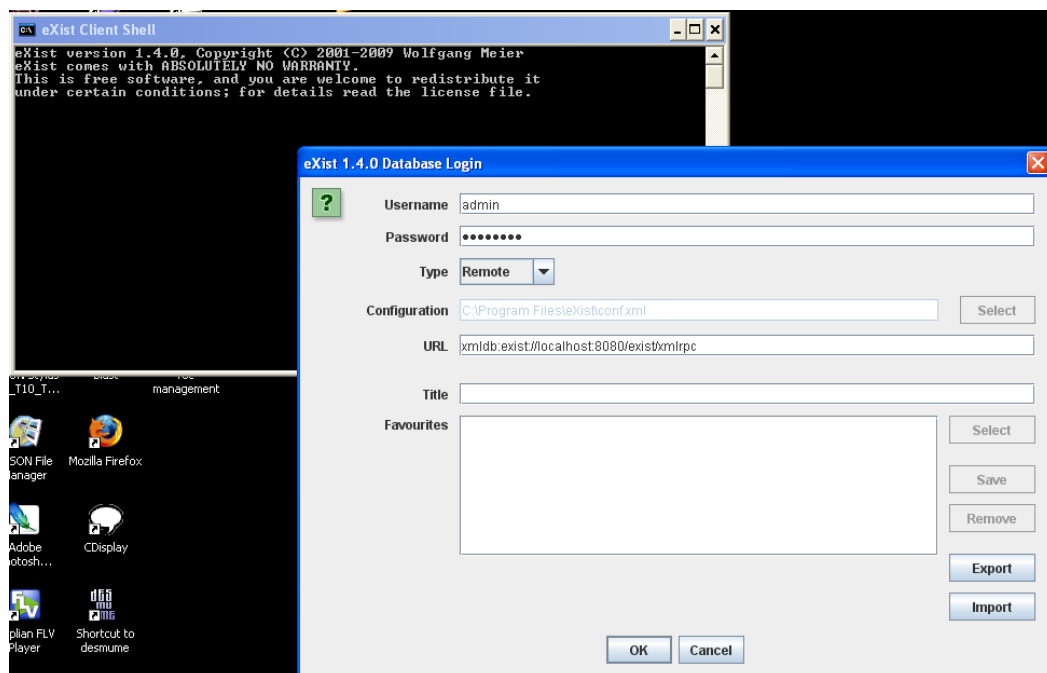
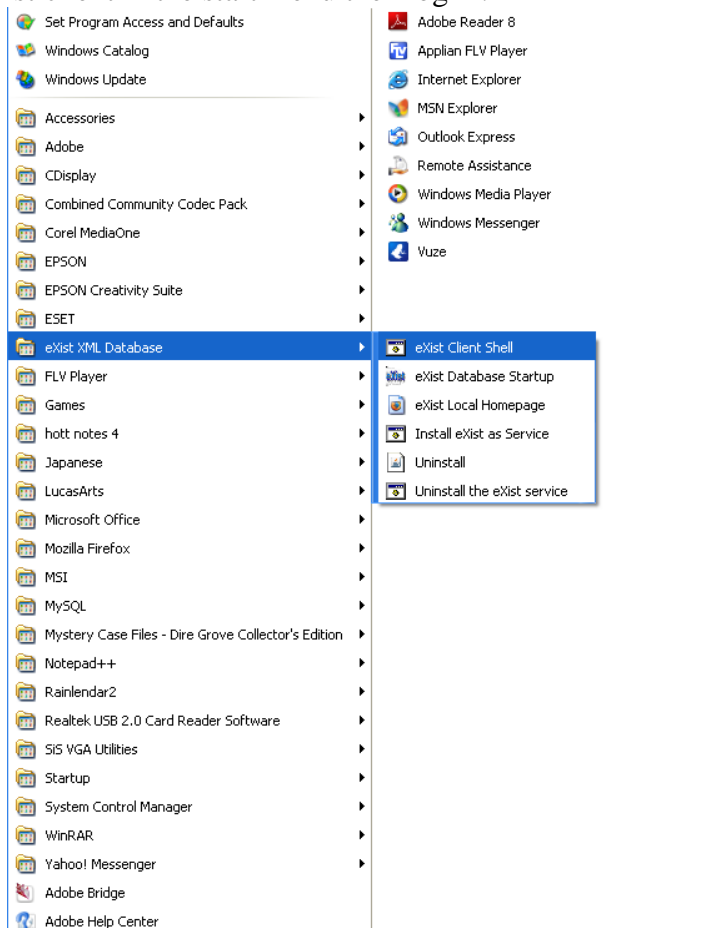
#### 4. Create an Index File for the Record based from the format (example: Medical Record)

```

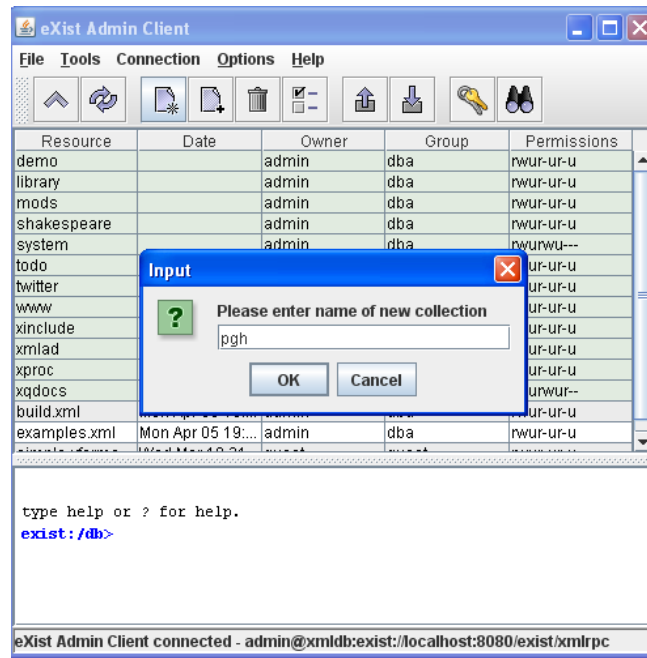
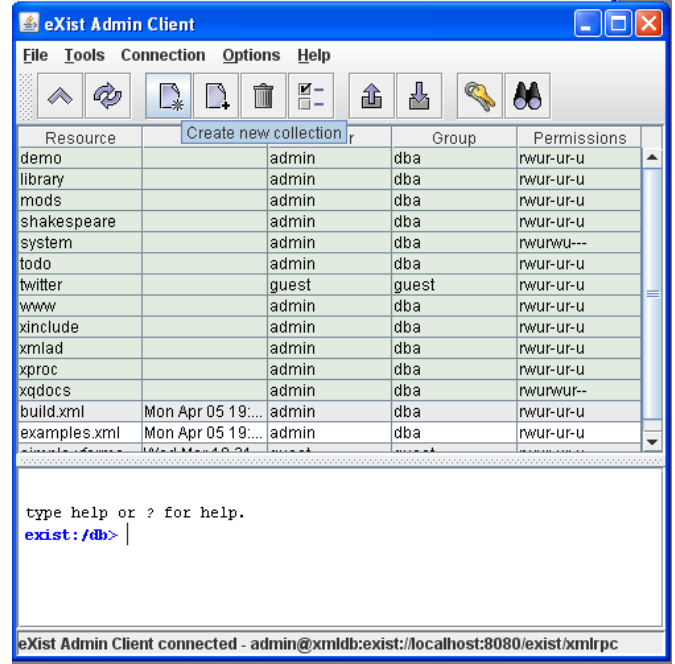
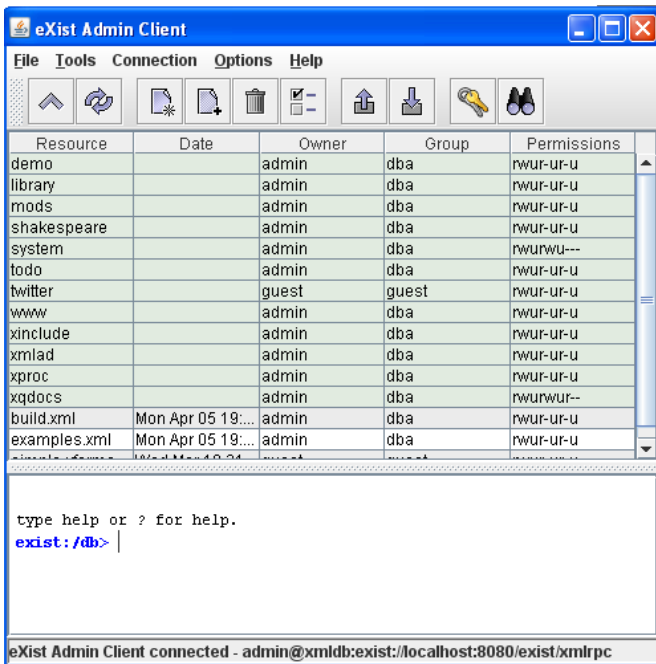
format.xconf
1 <collection xmlns="http://exist-db.org/collection-config/1.0">
2   <index xmlns:atom="http://www.w3.org/2005/Atom"
3     xmlns:html="http://www.w3.org/1999/xhtml"
4     xmlns:wiki="http://exist-db.org/xquery/wiki">
5
6     <fulltext default="none" attributes="no"/>
7     <!--Lucene Indexes-->
8     <.lucene>
9       <analyzer class="org.apache.lucene.analysis.standard.StandardAnalyzer"/>
10      <analyzer id="ws" class="org.apache.lucene.analysis.WhitespaceAnalyzer"/>
11      <text qname="name"/>
12      <text qname="display"/>
13      <text qname="value"/>
14      <text qname="type"/>
15      <text qname="reqd"/>
16      <text qname="single"/>
17      <text qname="tag"/>
18      <text qname="row"/>
19      <text qname="column"/>
20      <text qname="element"/>
21    </lucene>
22    <!--Range Indexes-->
23    <create qname="name" type="xs:string"/>
24    <create qname="display" type="xs:boolean"/>
25    <create qname="value" type="xs:int"/>
26    <create qname="type" type="xs:string"/>
27    <create qname="reqd" type="xs:boolean"/>
28    <create qname="single" type="xs:boolean"/>
29    <create qname="tag" type="xs:string"/>
30    <create qname="row" type="xs:int"/>
31    <create qname="column" type="xs:int"/>
32    <create qname="element" type="xs:string"/>
33  </index>
34  <triggers>
35    <trigger event="store,remove,update" class="org.exist.versioning.VersioningTrigger"/>
36  </triggers>
37 </collection>
38

```

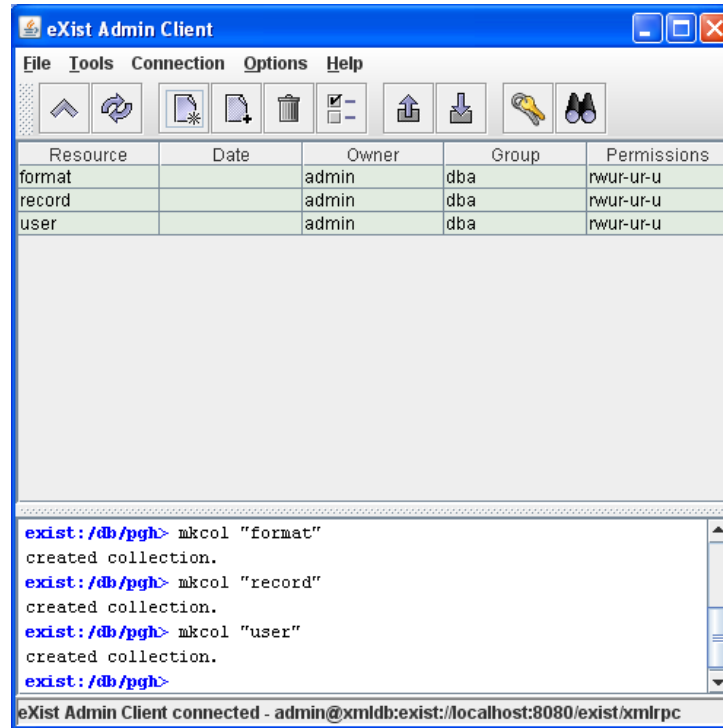
5. Open the Exist client in the start menu then log in.



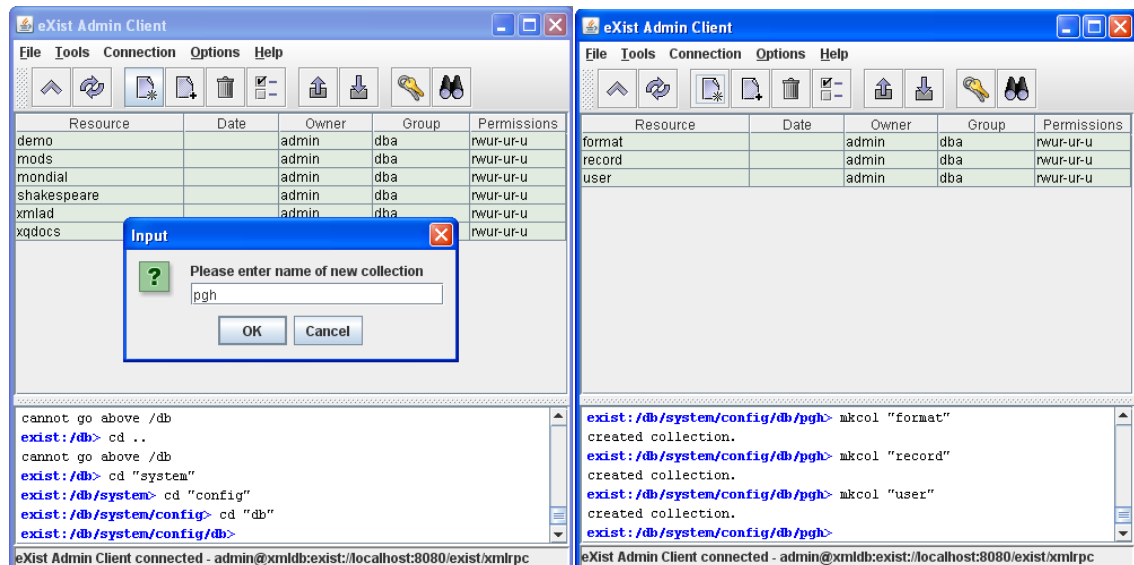
6. In the home (“/db”), create a collection according to the name in the edited medex.xml (“/db/CollectionName”)



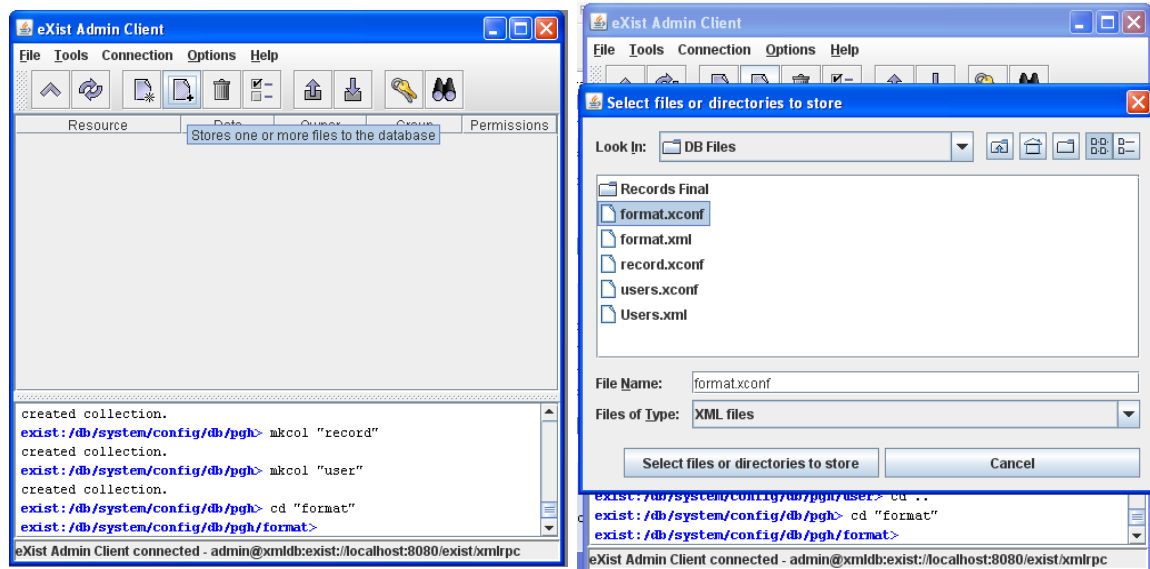
7. After creating the collection, create 3 collections inside that collection namely:
  - a. format
  - b. record
  - c. user



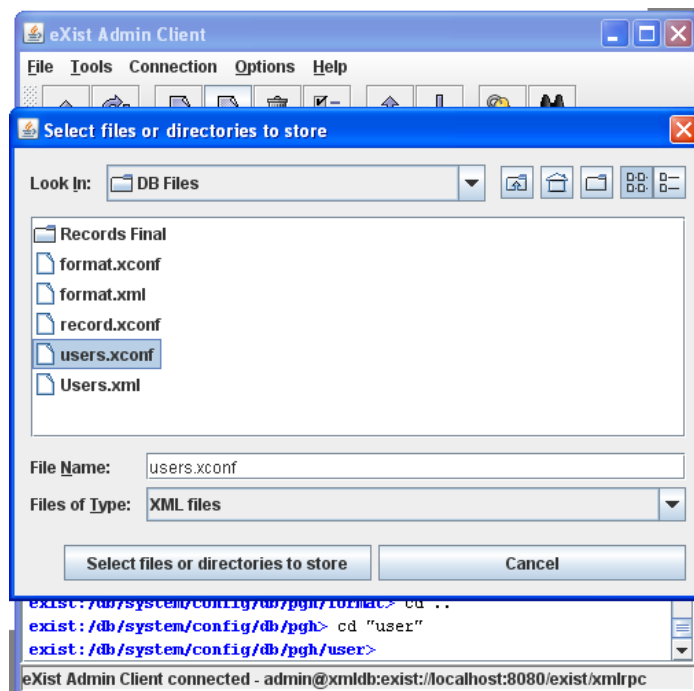
8. Go back to home ("/db") then go to system->config->db. Inside this collection, repeat step 6 and 7.



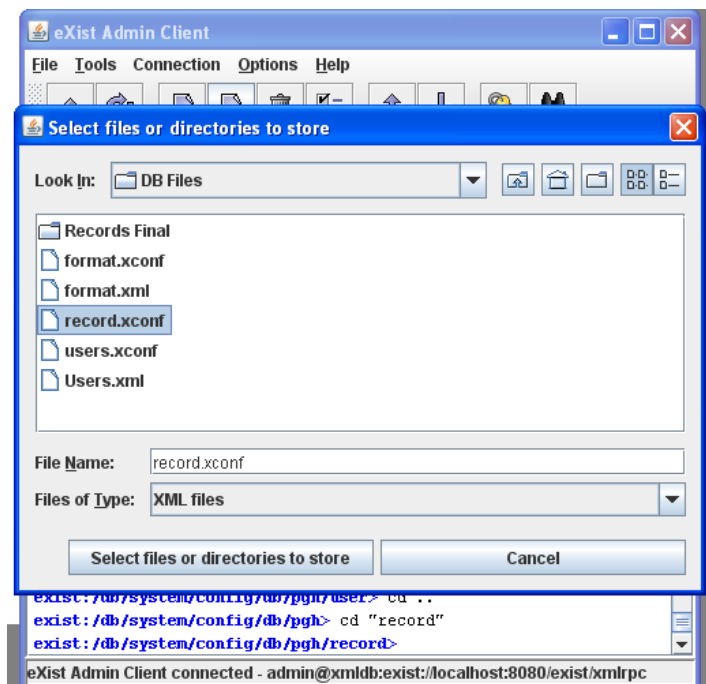
9. In the system->config->db->CollectionName->format, store the index file for the format that is found in your Application Folder->DB files->format.xconf



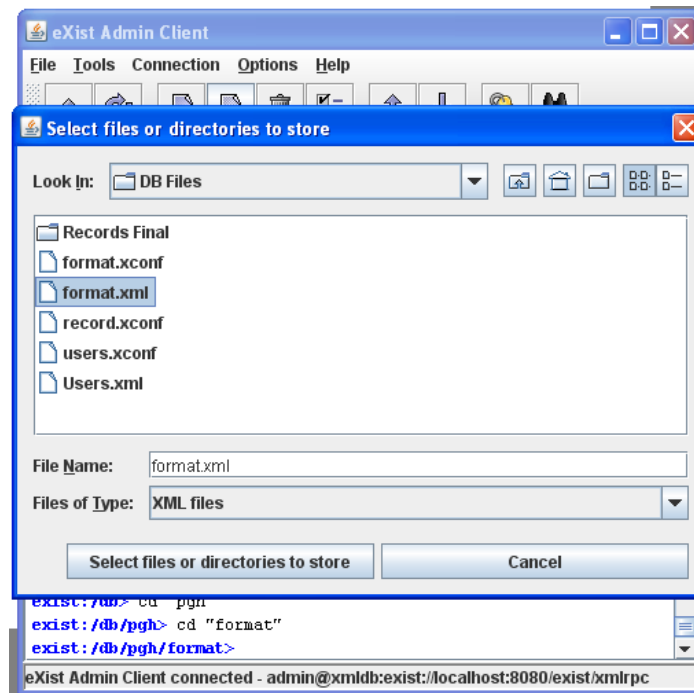
10. In the system->config->db->CollectionName->user, store the index file for the user that is found in your Application Folder->DB files->user.xconf



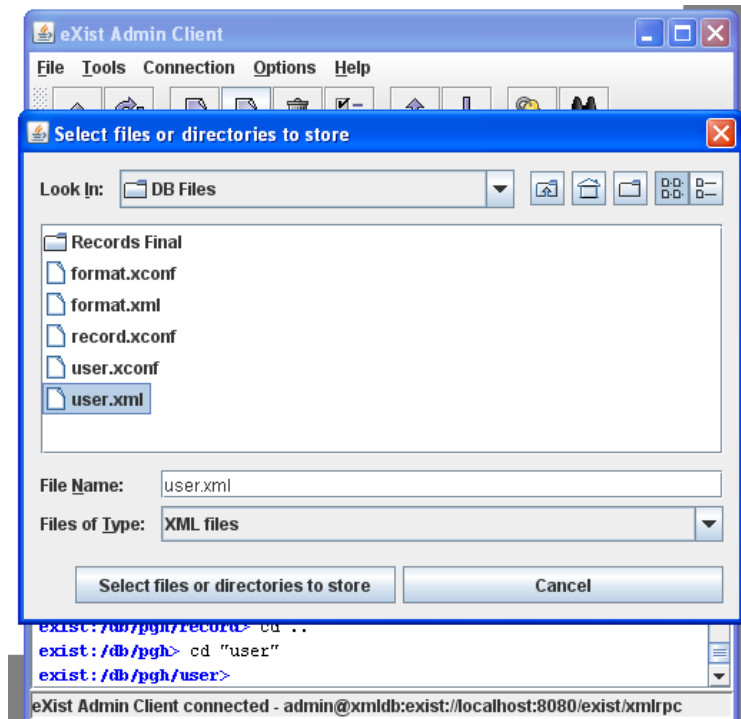
11. In the system->config->db->CollectionName->record, store the index file that you created for the records. Example of this file is located in the Application Folder->DB files->record.xconf



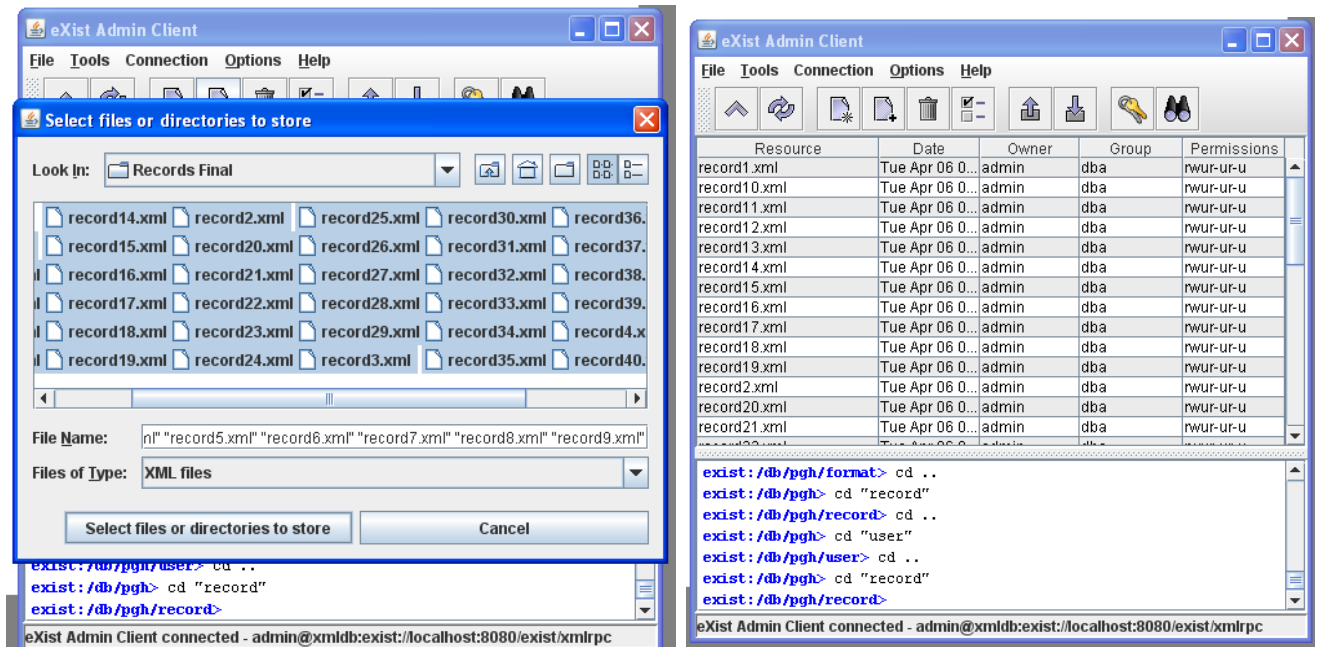
12. Go back to the home collection ("/db"). Go to CollectionName->format and import the format.xml you have created.



13. Next is going to the CollectionName->user and import the user.xml you have created.



14. Next is going to the CollectionName->record and store all the records in xml format.



15. After storing all the data, you can now use the whole application.

## Tuning Exist

### 1. Memory Settings:

By default, Java limits the amount of memory that it can access, thus eXist will not automatically exploit the machines available memory. Tuning the memory can optimize searches on large databases. After installing eXist as a service, users may want to edit *EXIST\_HOME/tools/wrapper/conf/wrapper.conf* to adjust Java Heap Size to their preference. Recommended setting is at least 512 MB of memory.

```
# Maximum Java Heap Size (in MB)
wrapper.java.maxmemory=512
```

### 2. Cache Settings:

All of the core database files and indexes has a page cache. It is used to load frequently used pages in the db faster. If the cache is too small, eXist may unload pages then reload them a few moments later, which is called “trashing effect”. Users may want to tune eXist to have a larger cache size. This can be done by editing *EXIST\_HOME/conf.xml*. Recommended setting is at least 128M of memory.

```
<db-connection cacheSize="128M" collectionCache="24M" database="native"
  files="webapp/WEB-INF/data" pageSize="4096" nodesBuffer="-1">
```

## Using the Application (MedEx Deployment):

### 1. Tabs:

- a. Home – contains the information of the Application
- b. Record Management – new, edit, delete records. This can only be accessed by the admin and members only.
  - i. New
  - ii. Edit
  - iii. Delete
- c. Search – contains the form for searching
  - i. Text Input
  - ii. Import File
- d. Results – contains a table where results are stored
  - i. Column Sorting
  - ii. Column Editing
  - iii. Page
  - iv. Row Click
- e. About – contains information about the creators of the program.

### 2. Record View

-contains the whole information of a record clicked.

### 3. Summary View

-contains the summary of the whole results found.





Fig 1-a. Home

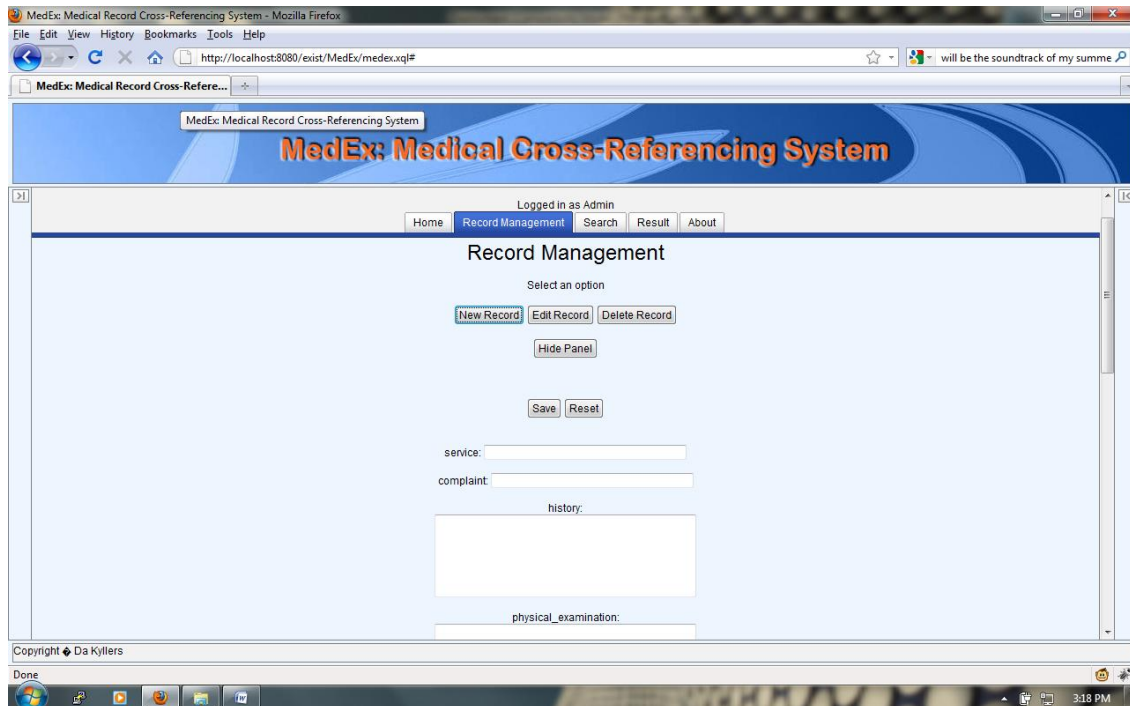


Fig 1-b. Record Management



Fig 1-c. Search

MedEx: Medical Record Cross-Referencing System - Mozilla Firefox

http://localhost:8080/evist/MedEx/9.2/index.cgi

MedEx: Medical Cross-Referencing System

Logged in as guest [Log in](#)

Results Retrieved

Found 22697 records in 0 seconds

Home Search Result About

RANK	ID	case_number	service	diagnoses	pathological_diagnoses	course	plan	outcome	SCORE
1	70	3,229,060	Pediatrics, Urology	Superior pole abscess probably secondary to infundibular stenosis secondary to OUTB, VUR B or IV, PC-neurogenic voiding dysfunction, superior pole abscess 2 to OUTB, VUR B Grade IV	final	patient is currently admitted for elective surgery. She was advised to gain weight and for nutritional upbuilding (from Dec to present) prior to performing the surgery. 7/20: admitted in Ward 6 Bed 25/731: given Mebendazole 500 mg single dose BID; video CMO, VCUG done. CMO: small capacity bladder (150 cc), leak point pressure 25 cm H2O, post void residual 40 cc, VCUG: (+) B reflux at 10 cm H2O after infusion of 100 cc, (+) vesicoureteral reflux, R grade IV, L grade IV	may go home nutritional upbuilding on opd basis for reexamination of urology for superior pole partial nephrectomy r for urodynamics, ureteral reimplantation; MOH Home Meds Cefuroxime 250mg/tab 2x a day PR Pedis kit CPO RUP Sept 4, 2008 7am URO clinic for uroflowmetry on RUP	improved, recovered	5.616546
2	9,839	3,261,805	Urology	staghorn calculus L	final	7/29: admitted in the ward 731: SIP nephrectomy, L	Home Meds: Cefuroxime 500 mg/tab BID x 5 more days Dolcet 1 tab TID prn for pain Follow-up: 8/14/08 at URO OPD (2:04) 7 AM, with repeat crea and urinalysis	recovered	0.04092873

Copyright © Da Kyllers

Done

Fig 1-d. Results

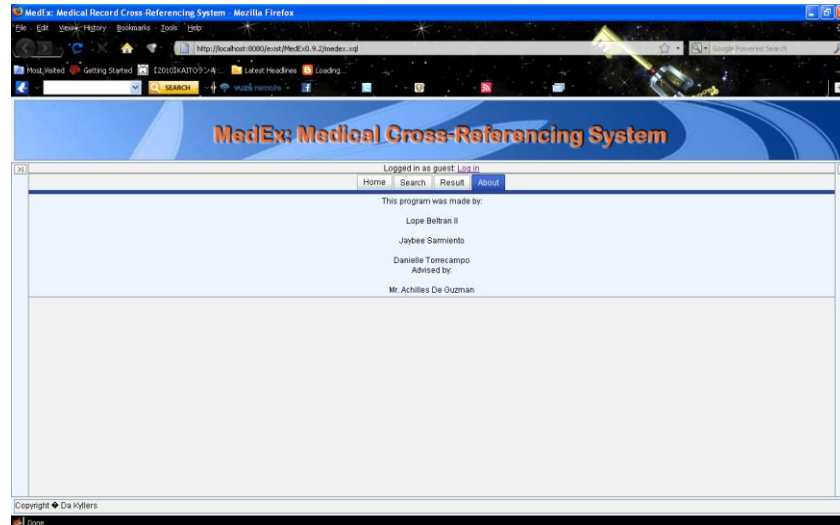


Fig 1-e About

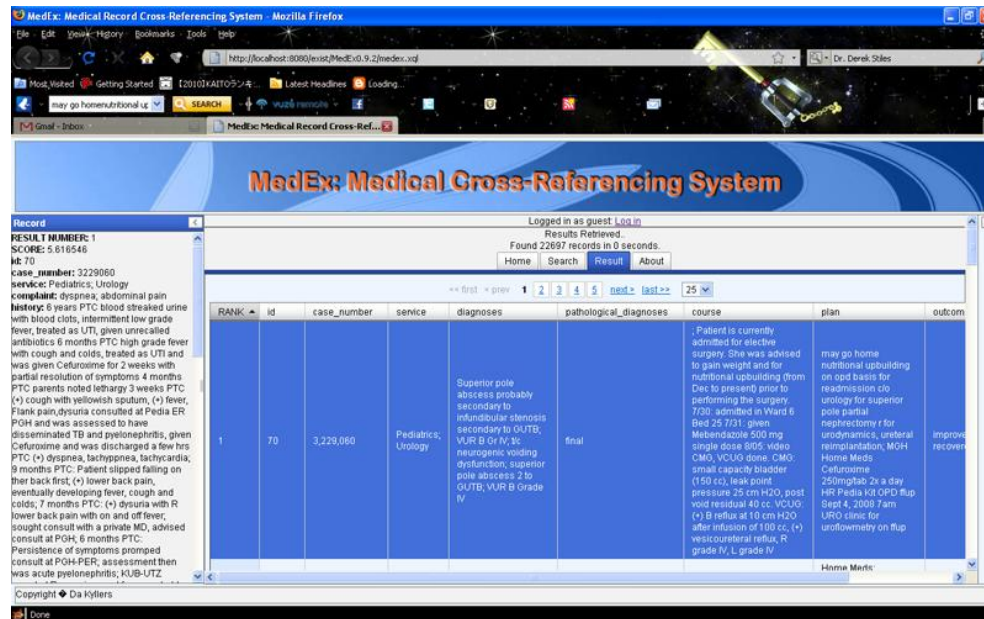


Fig 2. Result screen with Record Window

Summary	
<b>AGE:</b>	
19 :	46.153847%
20 :	15.384615%
50 :	7.6923075%
22 :	7.6923075%
47 :	7.6923075%
23 :	7.6923075%
29 :	7.6923075%
<b>GENDER:</b>	
male :	100%
<b>SYMPTOMS:</b>	
fever :	16.071428%
nosebleeds :	1.7857143%
loss of appetite :	1.7857143%
iredness :	1.7857143%
jaundice :	1.7857143%
R Flank Pain :	1.7857143%
edema :	1.7857143%
exertional dyspnea :	1.7857143%
chestpain :	1.7857143%
generalized aches and pains :	14.285714%
cough :	12.5%
papitations :	1.7857143%
hematochezia :	1.7857143%
body weakening :	1.7857143%
poor appetite :	1.7857143%
headaches :	1.7857143%
lethargy :	1.7857143%
diarrhea :	1.7857143%
sore throat :	1.7857143%
stuffy nose :	1.7857143%
headache :	1.7857143%

Fig 3. Summary of Results