

eGenVar web Interface tutorial

(Version beta 2.00, last updated 08 Feb 2014)

Authors: Sabry Razick & Pål Sætrom

Contents

eGenVar web Interface tutorial.....	1
1. Create an account	2
2. Login to the system	4
3. Search	5
3.1. Search query construction	6
3.2. Use case 1.....	8
3.3. Use case 2.....	9
4. Add records	10
5. Selecting records to be Updated or deleted	12
5.1. Update record	12
5.2. Delete records	13
6. Warnings and errors.....	14
6.1. Security warning.....	14

This tutorial is designed to introduce the eGenVar database management systems (EGDMS) web interface client. The EGDMS is a file metadata cataloguing system (please refer to the paper by Razick et al. for more details). There is an administrator's manual and a command line tool tutorial accompanying this document. The web interface described in this document mainly provides a search interface; however, it can be used to create and update records as well.

The following sections describe how use the web interface to create an account, to login to the system, to search, to add records, and to update or delete records. The final section describes common warnings and errors.

1. Create an account

To use the eGenVar data management system (EGDMS) an active account is required. The administrator can create accounts on behalf of the users (please refer the administrators manual), or the users can request accounts and the administrator can activate them. One user account is created during the installation; the user name for this account is the email of the administrator. The example below demonstrates the method to request an account using the web interface. (Note that during the first time visit to this page you may get a security warning; please refer to Section 7- Warnings and errors on how to handle this).



Figure 1: Create account. Click on the “New account” link to load the account request form

- Clicking on the “New account” link loads the account request form.
- Fill in the account request form with a valid email address and press submit.

- c. A confirmation code will be sent to the email and this should be used to confirm the account request.
- d. In the beta 2.x versions, the accounts will be immediately activated with minimum privileges. (Two factor authentication and activation by administrator may be required if you are using one of the beta 3.x versions).

2. Login to the system

Login is required to use the system and the user name is the email address, whereas the password will be emailed to the corresponding email address if the registration was a success. Click on the login link to start the login process and follow the instructions. After the login, more links will be available and you will get an interface similar to Figure 2. During the first time login you may get a security warning; please refer to Section 7 (Warnings and errors) on how to handle this.

The screenshot shows the eGenVar user interface for a registered user. At the top, there is a navigation menu with links: Home, Login/Logout, Search, Update, Delete, Create, Auto Create, and Help. Below the menu, a light blue box displays user information and links. The user is logged in as 'sabryr@gmail.com' with User ID '2'. It provides links for Logout, Delete account, Server address, Secure server address, WSDL location, Windows command line tool, and UNIX command line tool. Below this box, the section 'Project description' is visible, starting with the text: 'eGenVar data management system (EGDMS) is used for sharing information using a extended set of metadata, without compromising privacy or the security. This method could be used to create a catalogue of data for the biomedical research regardless of format, content or type and to describe'.

You are logged in as: sabryr@gmail.com	(Logout)
User ID: 2	(Delete account)
Server address	http://egenvar-VirtualBox.local:8085/eGenVar_web/
Secure server address	https://egenvar-VirtualBox.local:8185/eGenVar_web/
WSDL location	http://egenvar-VirtualBox.local:8085/eGenVar_WebService/Authenticate_service?wsdl
Windows. The download link to egenv command line tool	http://egenvar-VirtualBox.local:8085/egenv.zip
UNIX. The download link to egenv command line tool	http://egenvar-VirtualBox.local:8085/egenv.tar.gz

Project description

eGenVar data management system (EGDMS) is used for sharing information using a extended set of metadata, without compromising privacy or the security. This method could be used to create a catalogue of data for the biomedical research regardless of format, content or type and to describe

Figure 2: Interface for the registered user. The menu titles Search, Delete, Update and Create are used for just that (explained later). The Auto Create option (described below) is used to login to a remote server and select files to be reported to the system. The details box provides additional details generated dynamically. Here, the server address is the address of the server, which should be used when connecting from other devices. The WSDL location is the address that should be used with the command line tool. The link for the Windows and UNIX command line clients are also displayed.

3. Search

The Search menu loads the search page. Either the Basic (Figure 3) or the Advanced mode can be selected in this page.

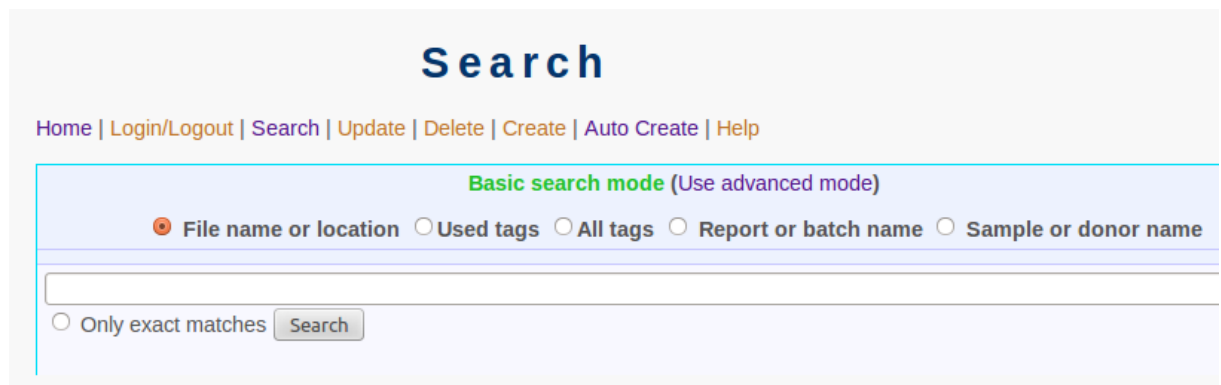


Figure 3: Basic Search. This has five search options. The “File name or location” option uses the file names, directory names, or the server name in the search. The “Used tags” option is used to find files, reports, report batches, samples and donors that have been tagged. The search is done in the tags and the results are the corresponding files that have been tagged. The “All tags” option searches through all tags; even the ones that have not been used. The “All tags” option could be used when the exact tag is not known and only a parent term is known. When the parent term is found then tags with all child terms can be loaded. For example, when the exact controlled vocabulary term for a particular cancer is not known, the term “cancer” can be searched in all tags and then specific sub types can be examined. The “Report or batch name” option is used to find reports and batches using their names. Once the report or the report batches are found, all files grouped under them can be loaded. The “Sample and donor name” option can be used to locate known samples or donors and then the files attached to those sample or donors can be loaded.

After the search is completed, links will be generated for each result. If the search returns large number of results, the results will be split on multiple pages. A history with the search term used is maintained for each user, which is convenient if the user needs to investigate the results later, and are synchronized across all logged in devices. When the search link is clicked on, a details page is generated as shown in Figure 4. Again, the results will be divided into pages if there are many results.

Results from FILES2TAGS

Filter: ID ▼ Apply filter History: 0,

Current_page=-1
 URL for this result: https://egenvar-VirtualBox.local:8185/eGenVar_web/Search/SearchResults3?TABLETOUSE_EGEN_DATAENTRY.FILES2TAGS=7415d68251b5815fcd4a946ad2b07fa00c2f43f6&reset=true

0
selected

[Clear all](#) | [Select all](#)

ID	LINK_TO_FEATURE	NAME	DESCRIPTION	URL	FILES_ID																		
4 ☐	BIOLOGICAL_ONTOLOGIES_TAGSOURCE=9966	_molecular_function_receptor activity_signaling receptor activity_natriuretic peptide receptor activity_2_3	This entry was reached following the path: biological_ontologies -> Biological_process.obo -> molecular_function -> receptor activity -> signaling receptor activity -> biological_ontologies.	NA	1																		
Filter using graph paths <table style="width: 100%; border-collapse: collapse;"> <tr> <td style="width: 50%;">Relevant DONORDETAILS</td> <td style="width: 50%; text-align: right;">Find</td> </tr> <tr> <td>Relevant REPORT_BATCH</td> <td style="text-align: right;">Find</td> </tr> <tr> <td>Relevant SAMPLEDETAILS</td> <td style="text-align: right;">Find</td> </tr> <tr> <td>Relevant REPORT</td> <td style="text-align: right;">Find</td> </tr> <tr> <td>Relevant FILES</td> <td style="text-align: right;">Find</td> </tr> </table>			Relevant DONORDETAILS	Find	Relevant REPORT_BATCH	Find	Relevant SAMPLEDETAILS	Find	Relevant REPORT	Find	Relevant FILES	Find	Details for FILES <table style="width: 100%; border-collapse: collapse;"> <tr> <td style="width: 50%;">NAME</td> <td style="width: 50%;">/home/egenvar/egenvar/cmd/test</td> </tr> <tr> <td></td> <td>/eGenVar_batch71/report71_3</td> </tr> <tr> <td></td> <td>/generated_8929_2707.txt</td> </tr> <tr> <td>Filter</td> <td>Get FILES</td> </tr> </table>			NAME	/home/egenvar/egenvar/cmd/test		/eGenVar_batch71/report71_3		/generated_8929_2707.txt	Filter	Get FILES
Relevant DONORDETAILS	Find																						
Relevant REPORT_BATCH	Find																						
Relevant SAMPLEDETAILS	Find																						
Relevant REPORT	Find																						
Relevant FILES	Find																						
NAME	/home/egenvar/egenvar/cmd/test																						
	/eGenVar_batch71/report71_3																						
	/generated_8929_2707.txt																						
Filter	Get FILES																						
5 ☐	BIOLOGICAL_ONTOLOGIES_TAGSOURCE=9966	_molecular_function_receptor activity_signaling receptor activity_natriuretic peptide receptor activity_2_3	This entry was reached following the path: biological_ontologies -> Biological_process.obo -> molecular_function -> receptor activity -> signaling receptor activity -> biological_ontologies.	NA	2																		
		_molecular_function_receptor	This entry was reached following the path: biological_ontologies ->																				

Figure 4: Search results. This page contains a static URL and the QR code for that URL, for reproducibility and communication purposes. Each result row is followed by a details row (displayed when the “+” sign is clicked). These details can also be used to expand the search; for example, to get all files for a found report.

Advanced mode requires an understanding of the schema (the names of the tables and their relationships). In the advanced mode, the user selects a table and a field to search and provides a search term.

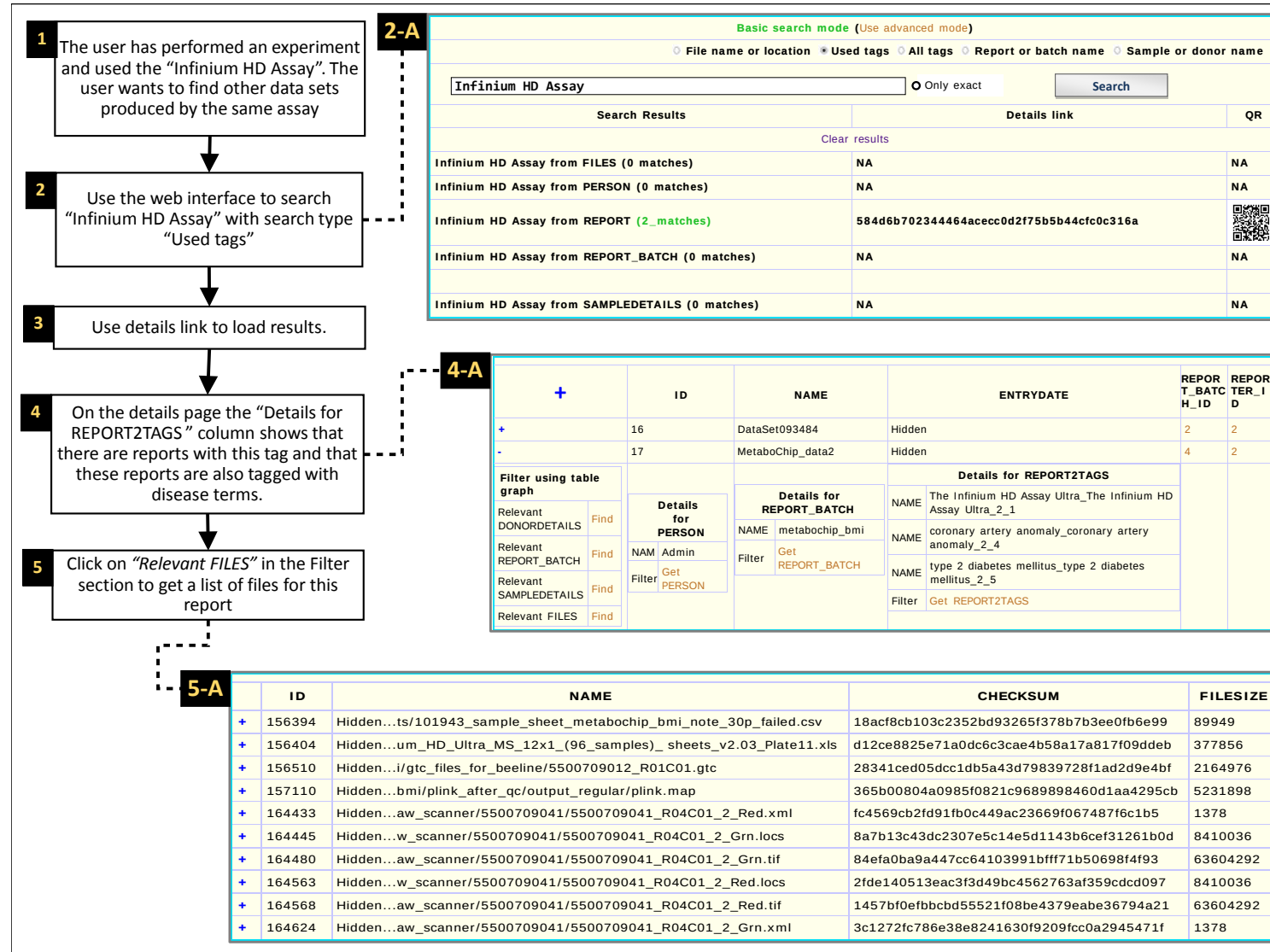
3.1. Search query construction

The following points provide step-by-step instructions on how the system can be used to handle different search types.

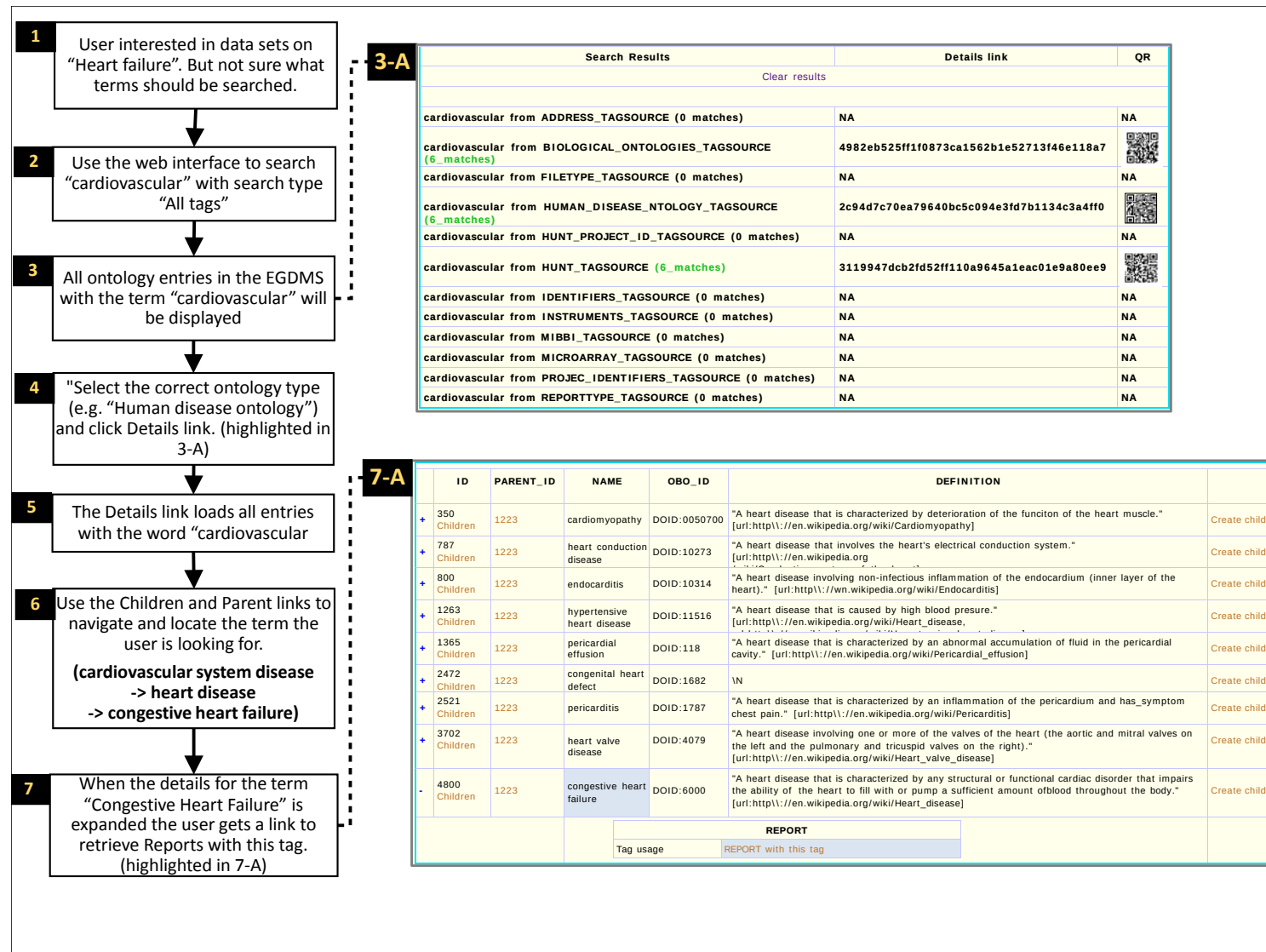
- Search for Reports and report batches tagged with the term “cardiovascular”
 - Use **Basic search mode**
 - Select **Used tags**
 - Type **cardiovascular** in search box
 - Leave **Only exact matches** box unchecked

- Click on **Search**
- Search Reports which are tagged with “type 2 diabetes” from the disease ontology and “The Infinium HD Assay” from MAIME (minimum information about a microarray experiment. In other words, find reports about files containing data on type 2 diabetes and where the experiment leading to the file used microarrays and a proprietary protocol called Infinium HD assay.
 - Use **Basic search mode**
 - Select **Used tags**
 - Type **type 2 diabetes && The Infinium HD Assay** in search box
 - Leave **Only exact matches** box unchecked
 - Click on **Search**
- Find all the files with **SampleSheet** in their name
 - Use **Basic search mode**
 - Select **File name or location**
 - Type **SampleSheet** in search box
 - Leave **Only exact matches** box unchecked
 - Click on **Search**
- Find all the files connected with the sample named **CVD_1045860123383**
 - Use **Basic search mode**
 - Select **Sample or donor name**
 - Type **CVD_1045860123383** in search box
 - Select **Only exact matches** box.
 - Click on **Search**
 - Click on the **Details link** to load details
 - Expand details by clicking on the plus sign
 - Click on **Find** link of the **Relevant FILES** in the **Filter using table graph** section.

3.2. Use case 1



3.3. Use case 2



4. Add records

The files are organized into Reports and Reports are grouped in to Report batches (please refer to the command line tutorial for a detailed description of this). This structure is automatically generated using folder names when creating records using the web interface. The Auto create option can be used to connect to any computer that allows SSH connection via port 22 and acquire file details remotely. These details can be used to create reports.

Home | [Login/Logout](#) | [Search](#) | [Update](#) | [Delete](#) | [Create](#) | [Auto Create](#) | [Help](#)

Attribute	Value	Comments
Server:		Use the full name of the file hosting server. E.g. egenvar-VirtualBox.local
Username:		
Password:		

Please note that you can only store files located on a server accessible via Internet or an intranet

Figure 5: Auto create 1. The login screen for the auto create option. At the moment only computers configured for SSH connection via port 22 is allowed.

Attribute Value Comments

Server: Use the full name of the file hosting server.
E.g. egenvar-VirtualBox.local

Username:

Password:

Please note that you can only store files located on a server accessible via Internet or an intranet. Use the full name of the file hosting server. Can not use localhost as server name here as the files will not be accessible from outside.

Select directory

[Go up one level](#)

- [/home/sabryr/tmp/rand/u/\(1 files\) \(Use this location | Expand\)\)](#)
- [/home/sabryr/tmp/rand/batch_7/\(1 files\) \(Use this location | Expand\)\)](#)
- [/home/sabryr/tmp/rand/preps/\(0 files\) \(Use this location | Expand\)\)](#)
- [/home/sabryr/tmp/rand/batch_3/\(6 files\) \(Use this location | Expand\)\)](#)

Figure 6: Auto create 2. When the login is successful, the directories at the remote server will be listed as shown. The user can expand the directories to display sub directories or use the location. When the “Use this location” is selected the file metadata for that location is extracted and the page shown in Figure 7 is generated. (Warning: when recording directories with a large number of files (> 200) the command line tool should be used instead; otherwise the system may hang for a long time.)

Directory Lister

Processing completed

/home/sabryr							
☒ /tmp/rand							
☒ /batch_3	2013-10-17 15:49:43	419Kb	a66a90e5c37de36c82ee41456b75111a16971d3f	New	New		
/report_11							
/genrated_118.txt							
/home/sabryr							
☒ /tmp/rand							
☒ /batch_3	2013-02-23 14:16:37	128 bytes	9756dd4273f9d6398b34b153e8c2a8d579286230	New	New		
/report_11							
/file_list.txt							
/home/sabryr							

Figure 7: Auto create 3. When the location is selected (Figure 6) and when the metadata extraction is finished, a view similar to this will be generated. When the metadata extraction is going on, a “still hashing” message will be displayed and the page will be automatically refreshed. This view also tells whether the file is new, one already recorded in the system, or a modified version of a file recorded in the system. This is determined by using the SHA1 checksum value and the full file path. After selecting the files to be recorded, the Submit selection button should be clicked to send the details to the database.

Started committing updates....

All updates successful

Following details were added to the database or they were already there

server_name	Name	Checksum	batch_name	Filepath	FileType	Group	Timestamp	Report_name	
ans-180228.stolav.ntnu.no	/home/sabryr /tmp/rand /batch_3 /report_11 /file_list.txt	9756dd4273f9d6398b34b153e8c2a8d579286230	batch_3	/home/sabryr /tmp/rand /batch_3 /report_11 /file_list.txt	NA	1001	1361625397	report_11	1
ans-180228.stolav.ntnu.no	/home/sabryr /tmp/rand /batch_3 /report_11 /genrated_118.qtt	5f487fa9802d0ec0172c1ab215f5a5ba68e1e629	batch_3	/home/sabryr /tmp/rand /batch_3 /report_11 /genrated_118.qtt	NA	1001	1382017720	report_11	4
ans-180228.stolav.ntnu.no	/home/sabryr /tmp/rand /batch_3 /report_11 /genrated_118.txt	a66a90e5c37de36c82ee41456b75111a16971d3f	batch_3	/home/sabryr /tmp/rand /batch_3 /report_11 /genrated_118.txt	NA	1001	1382017783	report_11	4

Modify values: [Files](#) | [Report](#) | [File Tags](#)

Figure 8: Auto create 4. After the files selected have been submitted (as in Figure 7), a view similar to this will be seen. The report name for the files is the folder name and the report batch is the parent folder name (these values can be modified using Update). The “File Tags” link at the bottom can be used to add tags common to all the files in the list. When tagging individual files or a subset of files, the Update option should be used instead.

5. Selecting records to be Updated or deleted

The “Select” option is used to select records to be changed, or to delete values or records. This page is loaded when the Delete or Update menu is selected. One record at a time or a set of records at once can be selected this way.

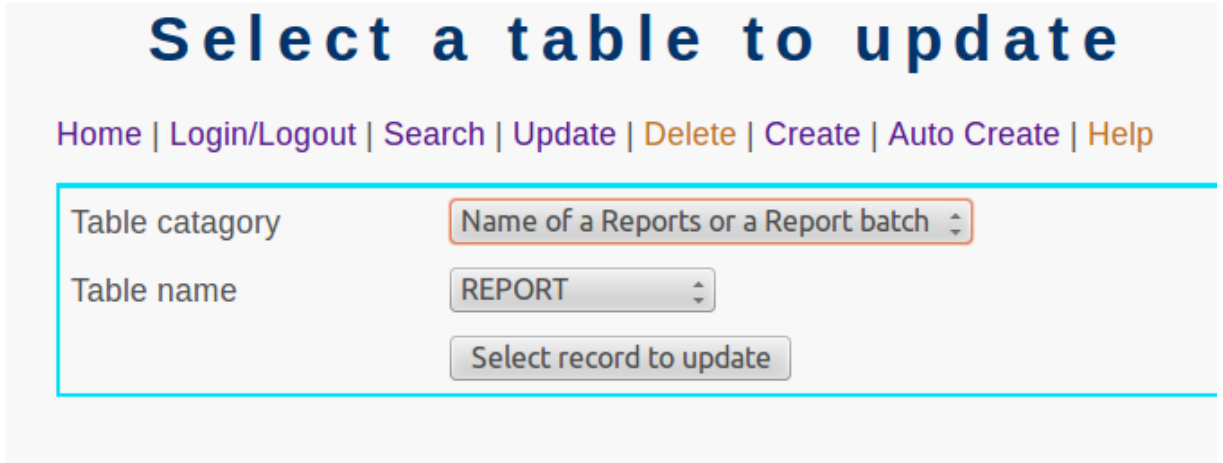


Figure 9: Select to Update or Delete. This screen shot shows how to select a Report to be updated. The category and table name drop down menus are used to perform this selection. When the category is updated, the table name list is refreshed automatically.

5.1. Update record

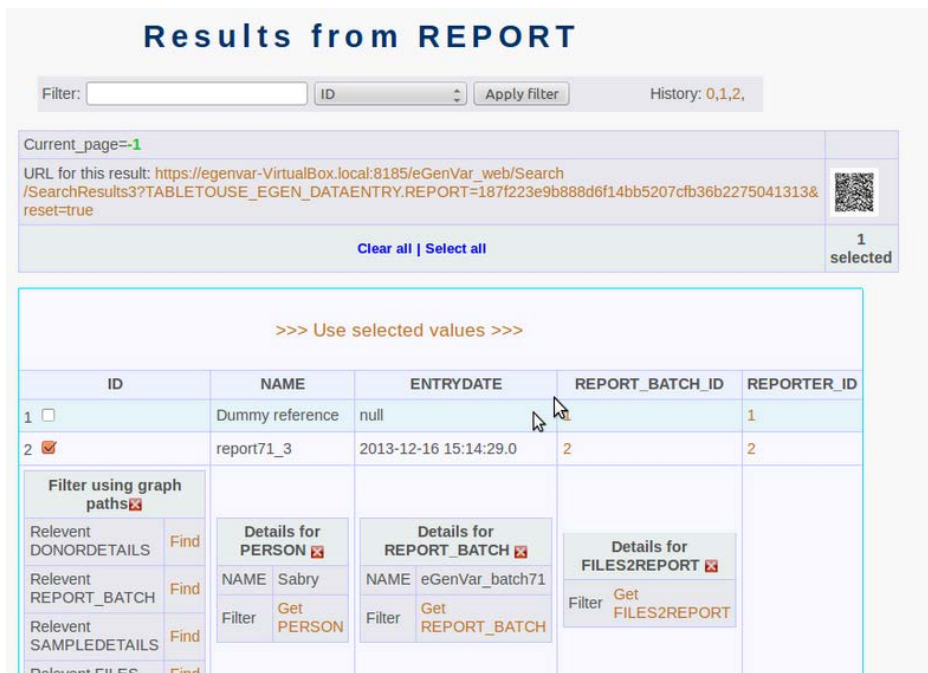


Figure 10: Update. The “select” option shown in Figure 9 will load this page. Here the records to be updated can be selected and the “>>> Use selected values >>>” link is used to go to the next stage.

5.2. Delete records

Records can be deleted by privileged users. When performing a deletion, the user must first select the records to be deleted, as for the Update operation, and then request deletion. When performing a deletion, a recursive check is performed to find other records affected by the deletion and they are also modified after informing the user.

Deletion confirmation

ID	NAME	ENTRYDATE	REPORT_BATCH_ID	REPORTER_ID	info
2	report71_3	2013-12-16 15:14:29.0	2	2	Marked for deletion

Cancel | Delete

When this record is deleted, the data listed below would also be deleted

Following rows from the table "EGEN_DATAENTRY.FILES2REPORT" will be deleted

ID	REPORT_ID	FILES_ID	info
1	2	1	Marked for deletion
2	2	2	Marked for deletion

11 rows will be deleted and only the first 2 rows are shown here. To view all [click here](#)

Figure 11:Delete. The delete is requested for the report "report71_3". This also shows the modifications required for the files2report table (two additional deletions) due to the deletion.

6. Warnings and errors

6.1. Security warning

A self-signed certificate is included with the EGDMS system. If administrators do not include or set up their own certificates, this self-signed certificate will be used by default. As this certificate is not certified by a certifying authority, most currently used web-browsers will complain that the certificate and site cannot be trusted. To proceed, the user needs to tell the browser that it is safe to trust this certificate and site.

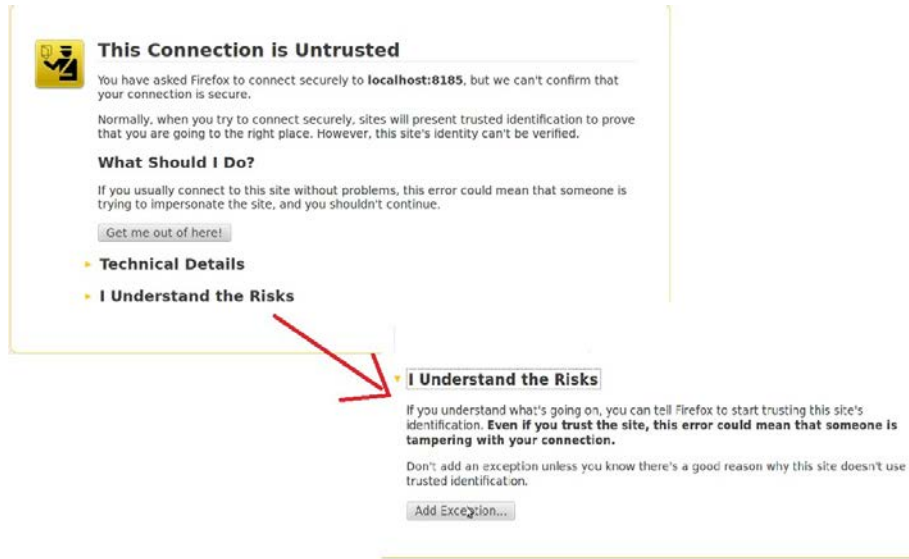


Figure 12: Security warning 1. If you get this warning, click on “I understand the Risks”. Then click on the “Add Exception” button. This will load a message similar to Figure 13.

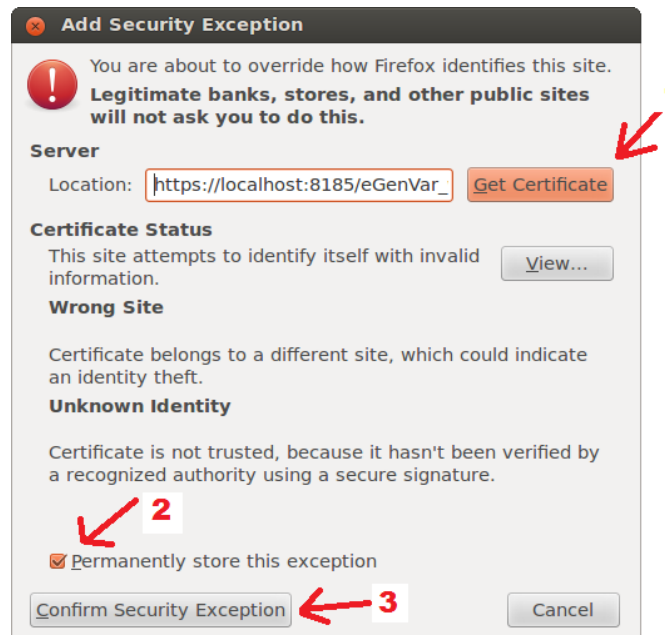


Figure 13: Security warning 2. When you get this, first click on “Get Certificate” (1), then tick the “permanently store this exception” (2) and then click the button “Confirm Security Exception” (3).