

Eco-AlpsWater

Innovative Ecological Assessment and Water Management Strategy for the
Protection of Ecosystem Services in Alpine Lakes and Rivers

Priority 3: Liveable Alpine Space. SO3.2 - Enhance the protection, the
conservation and the ecological connectivity of Alpine Space

Project Eco-AlpsWater

Work Package WPT1

Activity A.T1.2

Deliverable D.T1.2.2

Version 1.0

Date June 2021

Coordination of the WP1: I. Domaizon

Deliverable D.T1.2.2

Protocol data storage, sharing, retrieval

Interreg Alpine Space - Eco-AlpsWater project – WP1

Marco Moretto¹, Alessandro Cestaro¹, Giulia Riccioni¹, and Claudio Donati¹

¹Edmund Mach Foundation(FEM), San Michele all'Adige, Trento, Italy

ABSTRACT

The aim of this document is to provide a concise description of the most common procedures to fill in the web-based platform. This step by step manual guide the users throughout the whole process of data and metadata upload/download.

1. Login to the web-based platform

In order to have access to the platform it is necessary to login to the EAW website (Figure 1).

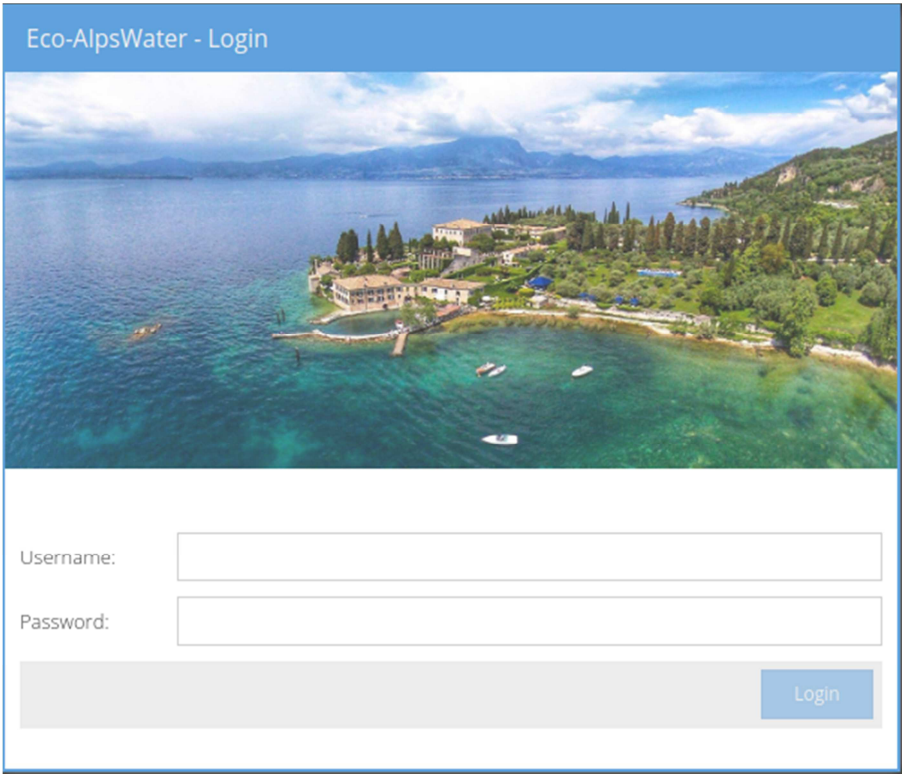


Figure 1. The EAW login page.

The Eco-AlpsWater (EAW) website is hosted at the Fondazione Edmund Mach at the URL **<https://eco-alpswater.fmach.it>**. For users accessing it from within the FEM network the private IP address **<https://10.234.110.141:8000>** (or **<https://docker02:8000>**) have to be used. A valid account, in order to login to the EAW website, was made available to the respective PPs responsible of the data input.

2. Create a new sample: Add new sampling station

In order to upload data for a new sample is mandatory to add the information about the sampling station the sample was collected from.

To create a new sampling station click on the **station** button on the left menu.

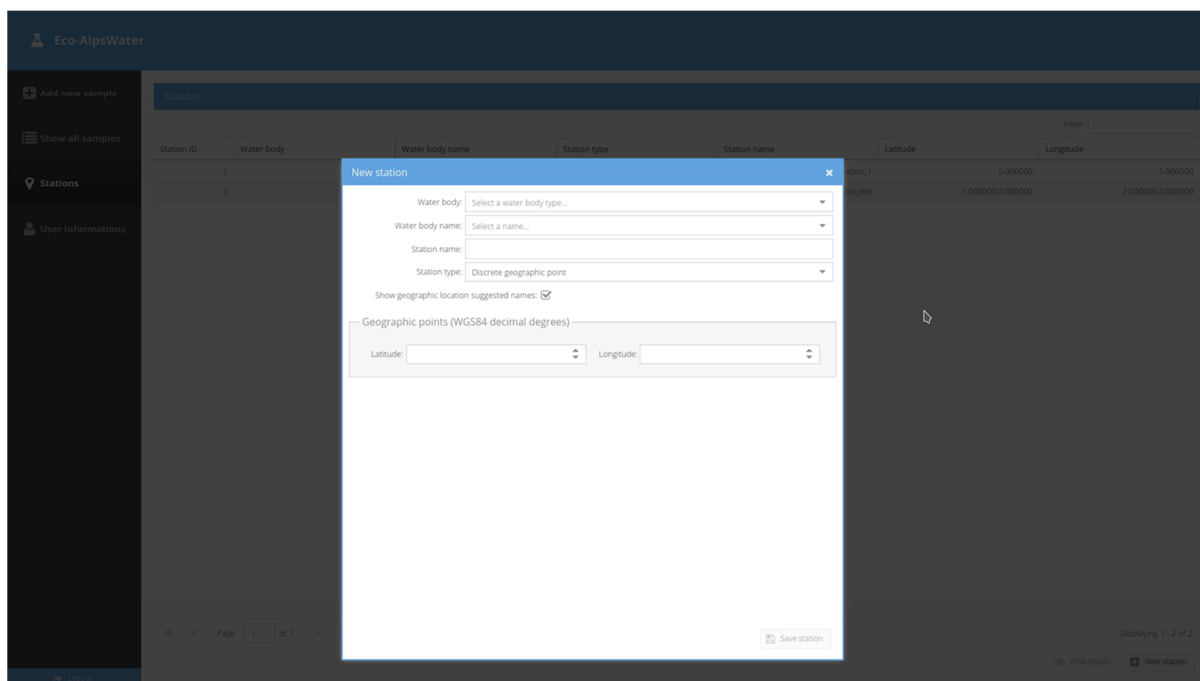


Figure 2. Add new sampling station panel.

The station geographic location can be composed by **single** or **multiple points**. Geographical point(s) are **WGS84 decimal degrees** with a precision of **6 decimal digits**. If the checkbox **show geographic location suggested names** is selected (Figure 2), when adding the first geographical point (latitude and longitude) a list of names based on OpenStreetMap (<https://www.openstreetmap.org/>) will appear.

3. Create a new sample: Add new sample

Once the sampling station have been created it is possible to add a new sample to the system.

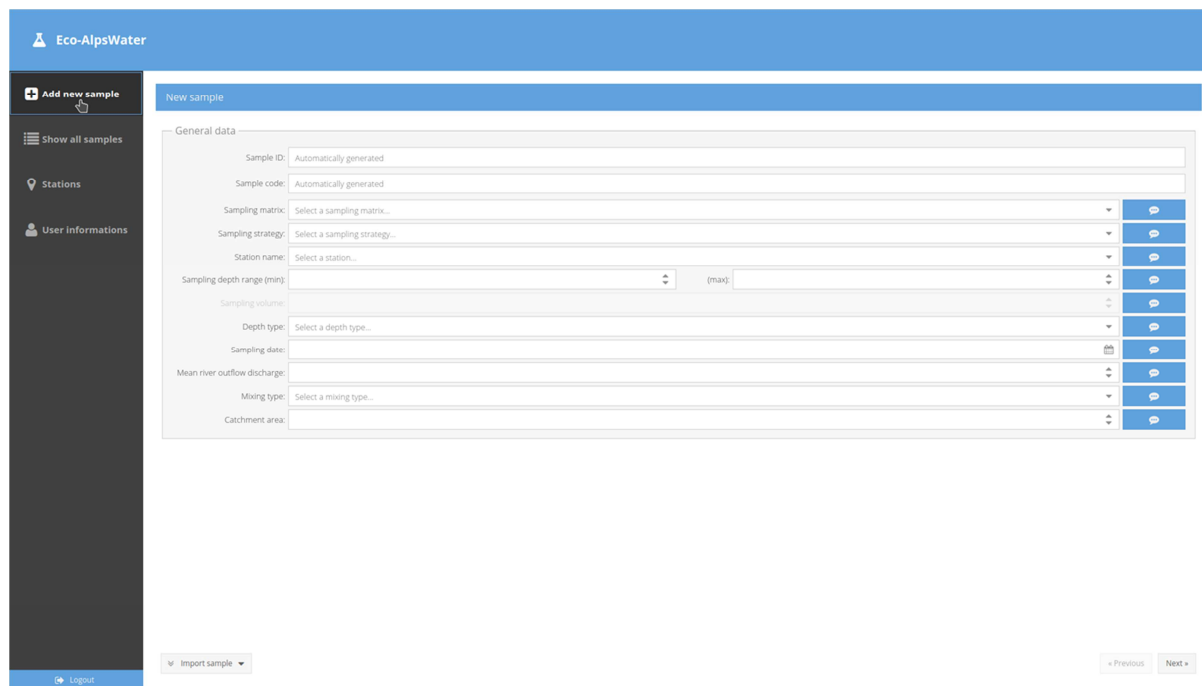


Figure 3. The New sample panel.

To add a new sample, click on the **Add new sample** button on the left menu. The first 2 fields will be automatically generated by the system when the **mandatory** fields are filled (Sample ID and Sample code, Figure 3). Mandatory fields are: Sampling matrix, Sampling strategy, Station name, Sampling depth range, Depth type and Sampling date. Once all the **mandatory** fields have been compiled the system create a customized Sample ID and Sample code. The **New sample panel** is composed by four pages that can be browsed using the **next** and **previous** buttons on the bottom of the panel. It is possible to associate a **comment** to any of the field by pressing the blue button on the right side of each of the them (Figure 3). A new sample could also be created by cloning an already existing one, by clicking the **import sample** button and selecting **clone from another sample** (the same functionality is available from the **show all samples** panel).

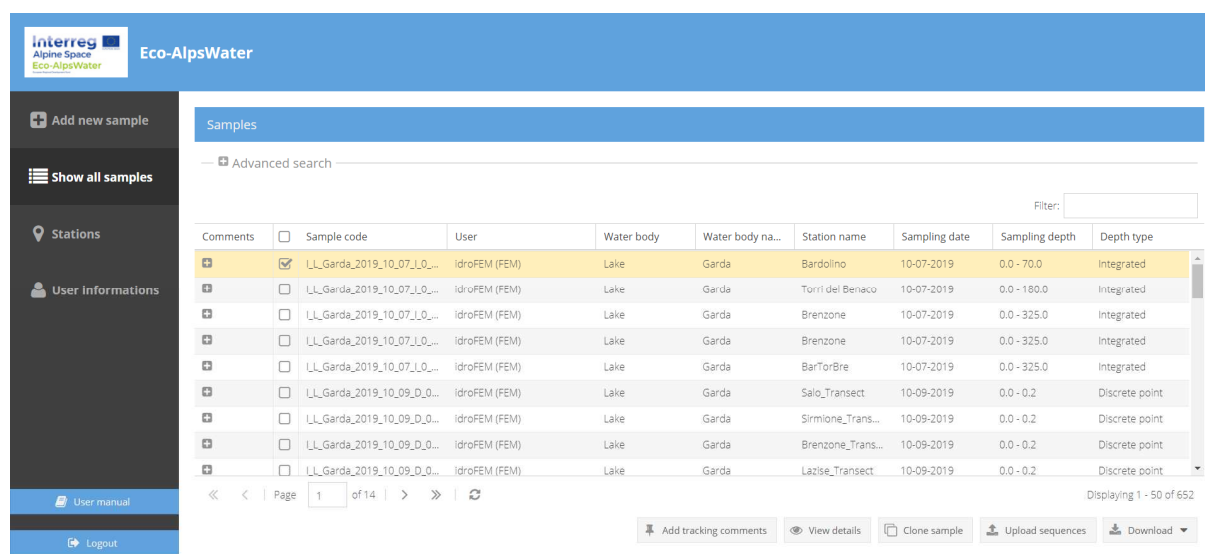
If all **mandatory** fields are filled, it is possible to add the new sample created into the database by clicking on the **save sample** button on the last page. The insertion of a new sample into the database is a **permanent** operation and thus the system will ask to **double-check** all the fields before continue. Once a new sample is added into the database the user **cannot modify or delete** it.

In case of need the user can request to **modify** or **delete** a sample by sending an e-mail to the **web-site platform managers**.

Once a sample has been successfully added into the database, a **confirmation e-mail** containing the generated **barcode** and an **Excel file (.xlsx)** with **environmental** and **meta-data** is sent to the user.

4. Overview of the samples created: show all samples

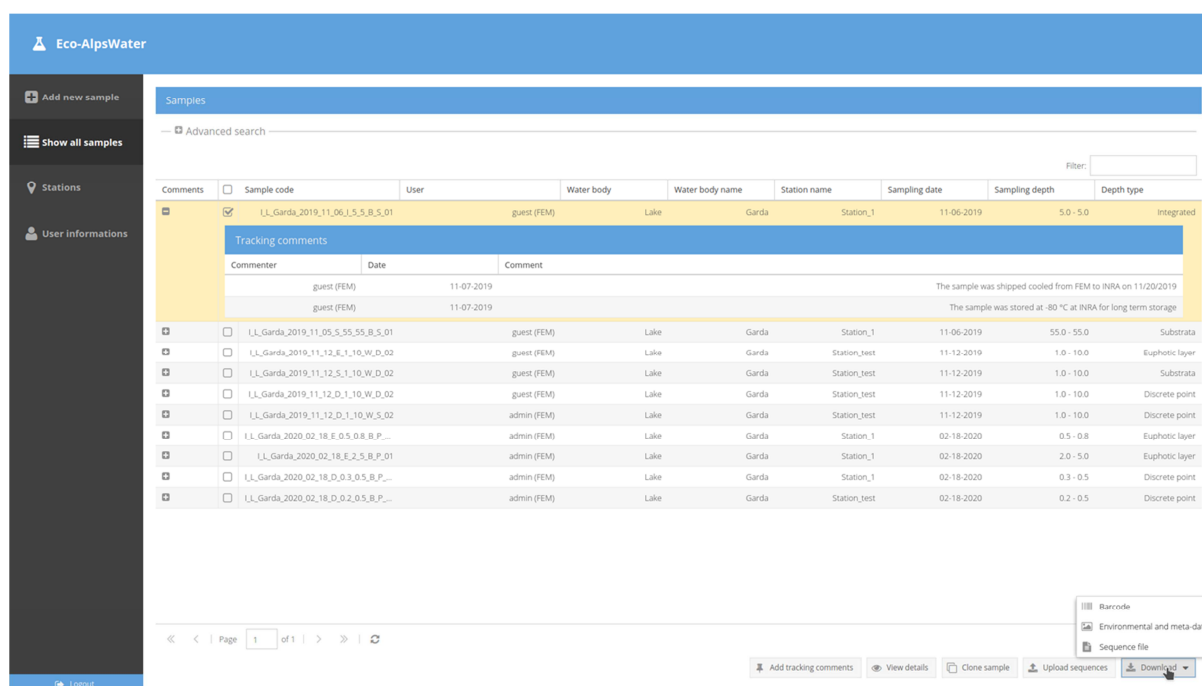
In order to have a general overview of the samples uploaded (and check if all the samples have been added), it is possible to display all the samples with the **Show all samples** option (Figure 4.1).



Comments	Sample code	User	Water body	Water body na...	Station name	Sampling date	Sampling depth	Depth type
<input checked="" type="checkbox"/>	L_L_Garda_2019_10_07_L_0...	IdroFEM (FEM)	Lake	Garda	Bardolino	10-07-2019	0.0 - 70.0	Integrated
<input type="checkbox"/>	L_L_Garda_2019_10_07_L_0...	IdroFEM (FEM)	Lake	Garda	Torri del Benaco	10-07-2019	0.0 - 180.0	Integrated
<input type="checkbox"/>	L_L_Garda_2019_10_07_L_0...	IdroFEM (FEM)	Lake	Garda	Brenzone	10-07-2019	0.0 - 325.0	Integrated
<input type="checkbox"/>	L_L_Garda_2019_10_07_L_0...	IdroFEM (FEM)	Lake	Garda	Brenzone	10-07-2019	0.0 - 325.0	Integrated
<input type="checkbox"/>	L_L_Garda_2019_10_07_L_0...	IdroFEM (FEM)	Lake	Garda	BarTorBre	10-07-2019	0.0 - 325.0	Integrated
<input type="checkbox"/>	L_L_Garda_2019_10_09_D_0...	IdroFEM (FEM)	Lake	Garda	Salo_Transect	10-09-2019	0.0 - 0.2	Discrete point
<input type="checkbox"/>	L_L_Garda_2019_10_09_D_0...	IdroFEM (FEM)	Lake	Garda	Sirmione_Trans...	10-09-2019	0.0 - 0.2	Discrete point
<input type="checkbox"/>	L_L_Garda_2019_10_09_D_0...	IdroFEM (FEM)	Lake	Garda	Brenzone_Trans...	10-09-2019	0.0 - 0.2	Discrete point
<input type="checkbox"/>	L_L_Garda_2019_10_09_D_0...	IdroFEM (FEM)	Lake	Garda	Lazise_Transect	10-09-2019	0.0 - 0.2	Discrete point

Figure 4.1. The Show all samples option and the list of samples.

From this panel a user can **download** all the information associated with a sample: the generated **barcode** (used to label sequencing sample), the **sequence** files, the **environmental** and **meta-data** as Excel files (.xlsx, Figure 4.2).



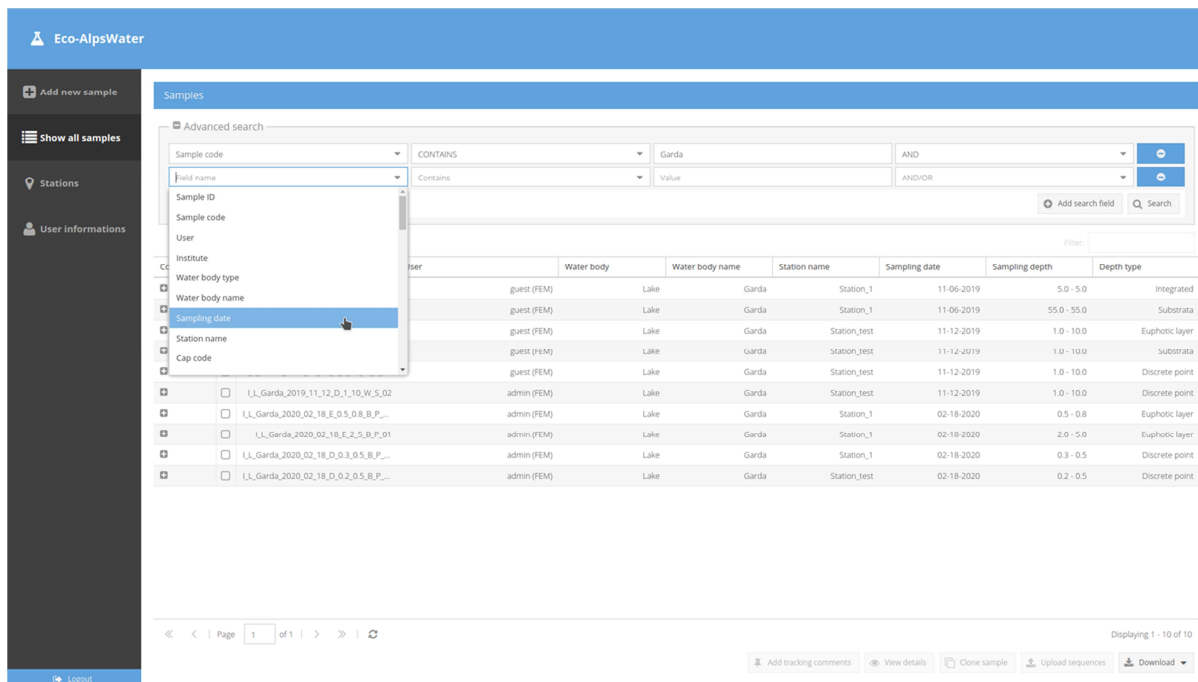
The screenshot shows the 'Eco-AlpsWater' interface. On the left is a sidebar with 'Add new sample', 'Show all samples', 'Stations', and 'User informations'. The main area is titled 'Samples' and contains an 'Advanced search' bar. Below this is a table of samples with columns: Comments, Sample code, User, Water body, Water body name, Station name, Sampling date, Sampling depth, and Depth type. The first sample is selected, and its details are shown in a pop-up window. This window includes 'Tracking comments' and a 'Download' button. The 'Download' button has a dropdown menu with options: 'Barcode', 'Environmental and meta-data', and 'Sequence file'.

Comments	Sample code	User	Water body	Water body name	Station name	Sampling date	Sampling depth	Depth type									
<input checked="" type="checkbox"/>	I_L_Garda_2019_11_06_1_5_5_B_5_01	guest (FEM)	Lake	Garda	Station_1	11-06-2019	5.0 - 5.0	Integrated									
Tracking comments <table border="1"> <thead> <tr> <th>Commenter</th> <th>Date</th> <th>Comment</th> </tr> </thead> <tbody> <tr> <td>guest (FEM)</td> <td>11-07-2019</td> <td>The sample was shipped cooled from FEM to INRA on 11/20/2019</td> </tr> <tr> <td>guest (FEM)</td> <td>11-07-2019</td> <td>The sample was stored at -80 °C at INRA for long term storage</td> </tr> </tbody> </table>									Commenter	Date	Comment	guest (FEM)	11-07-2019	The sample was shipped cooled from FEM to INRA on 11/20/2019	guest (FEM)	11-07-2019	The sample was stored at -80 °C at INRA for long term storage
Commenter	Date	Comment															
guest (FEM)	11-07-2019	The sample was shipped cooled from FEM to INRA on 11/20/2019															
guest (FEM)	11-07-2019	The sample was stored at -80 °C at INRA for long term storage															
<input type="checkbox"/>	I_L_Garda_2019_11_05_5_55_5_B_5_01	guest (FEM)	Lake	Garda	Station_1	11-06-2019	55.0 - 55.0	Substrata									
<input type="checkbox"/>	I_L_Garda_2019_11_12_6_1_10_W_D_02	guest (FEM)	Lake	Garda	Station_test	11-12-2019	1.0 - 10.0	Euphotic layer									
<input type="checkbox"/>	I_L_Garda_2019_11_12_5_1_10_W_D_02	guest (FEM)	Lake	Garda	Station_test	11-12-2019	1.0 - 10.0	Substrata									
<input type="checkbox"/>	I_L_Garda_2019_11_12_D_1_10_W_D_02	guest (FEM)	Lake	Garda	Station_test	11-12-2019	1.0 - 10.0	Discrete point									
<input type="checkbox"/>	I_L_Garda_2019_11_12_D_1_10_W_5_02	admin (FEM)	Lake	Garda	Station_test	11-12-2019	1.0 - 10.0	Discrete point									
<input type="checkbox"/>	I_L_Garda_2020_02_18_E_0_5_0_8_B_P_...	admin (FEM)	Lake	Garda	Station_1	02-18-2020	0.5 - 0.8	Euphotic layer									
<input type="checkbox"/>	I_L_Garda_2020_02_18_E_2_5_5_B_P_01	admin (FEM)	Lake	Garda	Station_1	02-18-2020	2.0 - 5.0	Euphotic layer									
<input type="checkbox"/>	I_L_Garda_2020_02_18_D_0_3_0_5_B_P_...	admin (FEM)	Lake	Garda	Station_1	02-18-2020	0.3 - 0.5	Discrete point									
<input type="checkbox"/>	I_L_Garda_2020_02_18_D_0_2_0_5_B_P_...	admin (FEM)	Lake	Garda	Station_test	02-18-2020	0.2 - 0.5	Discrete point									

Figure 4.2. The Download option for a selected sample.

Sequence files cannot be downloaded directly from the website. When a user requests to download sequence files an **email is sent** with all the information to retrieve the files (see *Sequence file transfer*).

Selecting one sample from this panel, it is possible to either **view the sample details** (i.e. all the information related to environmental and meta-data) or use the sample as a template and **clone** it to add a new sample (the same functionality is available also from the **Add new sample panel**). To quickly see only specific samples user can use the **filter** textbox on the top right of the sample list panel (Figure 4.2). For more **advanced search**, there's a specific hidden panel on the top that can be extended (Figure 4.3).



Sample code	User	Water body	Water body name	Station name	Sampling date	Sampling depth	Depth type
I_L_Garda_2019_11_12_D_1_10_W_5_02	admin (FEM)	Lake	Garda	Station_1	11-06-2019	5.0 - 5.0	Integrated
I_L_Garda_2020_02_18_E_0_5_0_8_B_P_...	admin (FEM)	Lake	Garda	Station_1	11-06-2019	55.0 - 55.0	Substrata
I_L_Garda_2020_02_18_E_2_5_0_8_P_01	admin (FEM)	Lake	Garda	Station_test	11-12-2019	1.0 - 10.0	Euphotic layer
I_L_Garda_2020_02_18_D_0_3_0_5_B_P_...	admin (FEM)	Lake	Garda	Station_test	11-12-2019	1.0 - 10.0	Substrata
I_L_Garda_2020_02_18_D_0_2_0_5_B_P_...	admin (FEM)	Lake	Garda	Station_1	02-18-2020	0.5 - 0.8	Euphotic layer
I_L_Garda_2020_02_18_D_0_2_0_5_B_P_...	admin (FEM)	Lake	Garda	Station_1	02-18-2020	2.0 - 5.0	Euphotic layer
I_L_Garda_2020_02_18_D_0_2_0_5_B_P_...	admin (FEM)	Lake	Garda	Station_1	02-18-2020	0.3 - 0.5	Discrete point
I_L_Garda_2020_02_18_D_0_2_0_5_B_P_...	admin (FEM)	Lake	Garda	Station_1	02-18-2020	0.2 - 0.5	Discrete point

Figure 4.3. The Advanced search options panel.

The **Advanced search panel** allows to add more **search fields** and combine them using boolean logic operators (AND, OR).

5. Edit sample information

It is possible to edit **some** sample information. To edit sample information select a sample from the **Show all samples** panel and click on the **View details**. From the just opened **sample detail** panel click the **edit mode** button. In order to make the changes permanent click the **update** button.

6. Sample tracking comments

The EAW website allow to track information about samples exchanged between institutions implementing partially the Nagoya protocol (D.M.1.2 Annex2). To add a **New tracking comment**, select a sample from the **Show all samples panel** and click on the **Add comment**. The **New tracking comment** panel contains some dropdown selections to help the user to fill the text with basic information about sample. Users are free to add/delete/modify text before adding the comment.

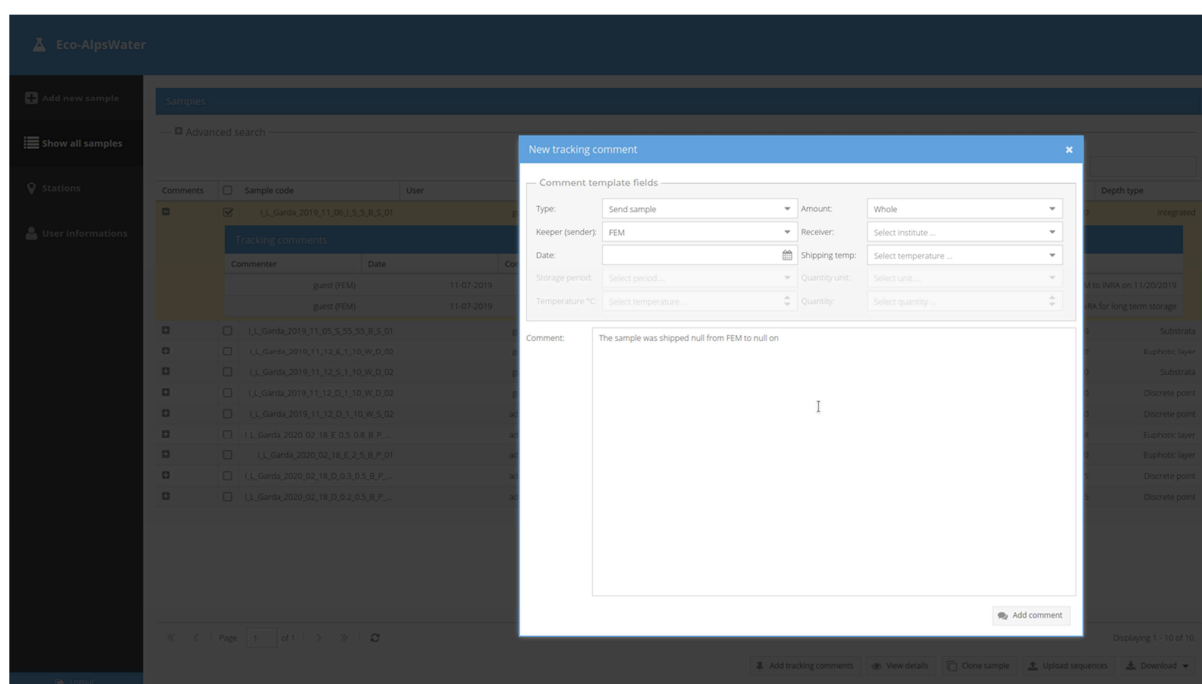
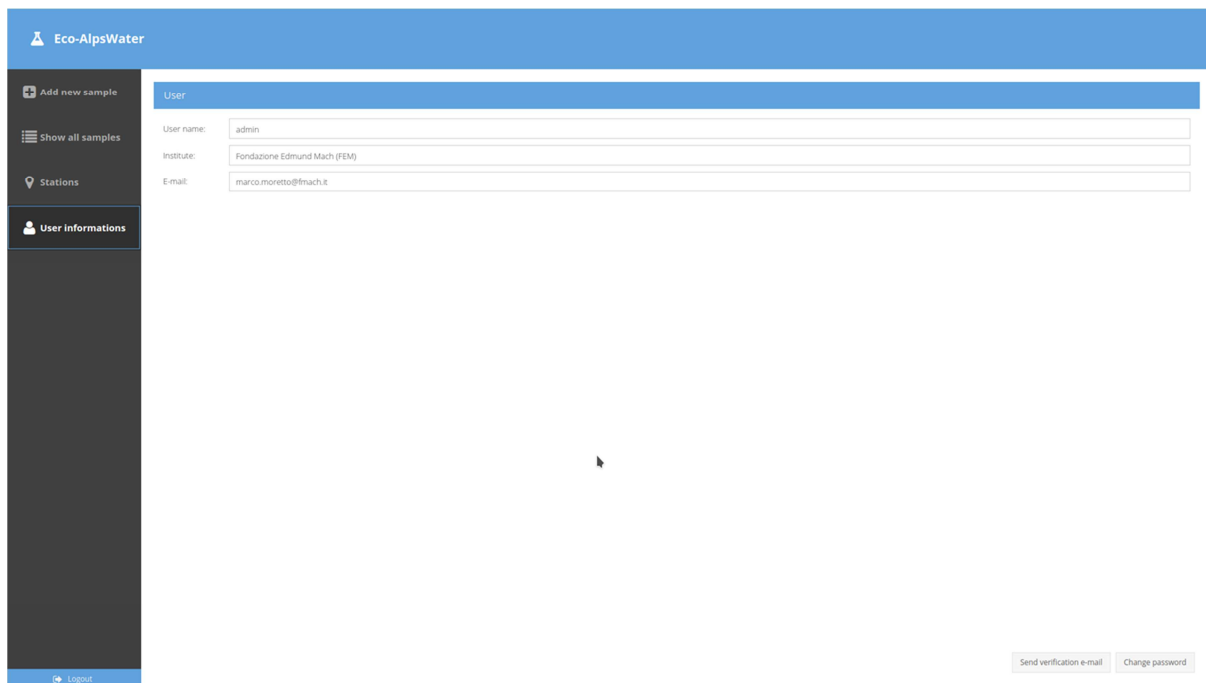


Figure 6. Add New tracking comment to sample.

7. User information

The **User information** panel display few basic information about the user. From this panel it is possible to **send** a simple **verification e-mail** to the user's email address or to **change** the **current password** (Figure 7).



Eco-AlpsWater

User

User name: admin

Institute: Fondazione Edmund Mach (FEM)

E-mail: marco.moretto@femach.it

Send verification e-mail Change password

Figure 7. The User information panel.

8. Sequence file upload

In order to upload sequence files, **fastq** files have to be linked to the respective **MD5 checksum files**. An **MD5 checksum** is a 32-character hexadecimal number that is computed on a file to verify the integrity of files and allow to check **fastq** files consistency (see <https://www.lifewire.com/validate-md5-checksum-file-4037391> for further information).

To upload sequence files into the EAW server, select the corresponding sample from the **Show all samples** panel and click the **Upload sequence** button (Figure 8).

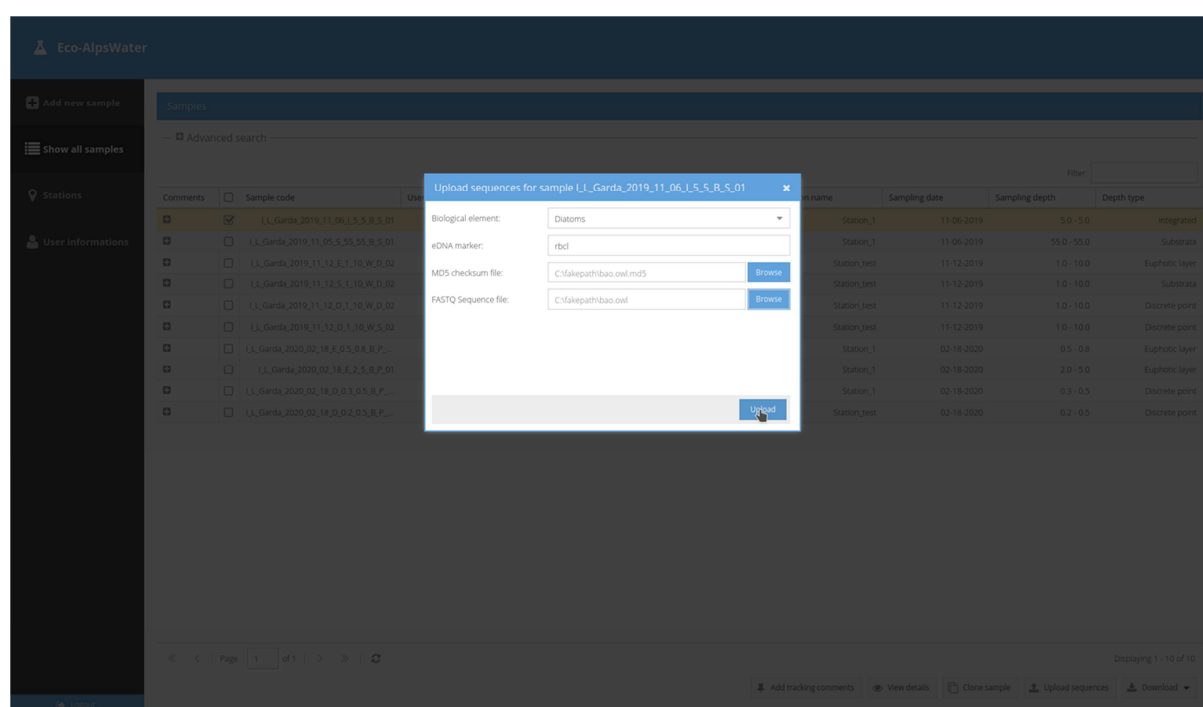


Figure 8. The Upload sequence file form.

The user will need to provide the **Biological Element**, the **MD5 checksum file** and the related **fastq file**. MD5 checksum file are usually provided together with fastq files.

Users are free to upload whatever file they want as long as the corrected MD5 file are provided, thus one could upload one single ZIP file instead of multiple fastq files. In this case the MD5 checksum file should be calculated from the user as well. For Windows see <http://www.md5summer.org/> and/or <https://support.microsoft.com/it->

[it/help/889768/how-to-compute-the-md5-or-sha-1-cryptographic-hash-values-for-a-file,](https://www.geeksforgeeks.org/md5sum-linux-command/)

while Linux users might simply use the md5sum command directly from the command line <https://www.geeksforgeeks.org/md5sum-linux-command/>.

Please note that the upload system is capped for upload speed and will timeout after 20 minutes, thus **each uploaded file should be smaller than 15Gb in size**.

Once a file is successfully uploaded the user will receive a **confirmation** via **email** and the file will be **automatically moved** into the vault (/vault) directory.

The vault directory is **inaccessible**, i.e. the user cannot enter the directory and see its content. This means that once a file is successfully copied into the FTP server **it is not possible for users to read, modify or delete it**.

9. Sequence file download

Sequence files can be **downloaded** from the EAW server directly from the EAW website, selecting the **Show all samples** panel.

From the **Show all samples** panel the user can select the sample(s) to download and click on **Download sequence** file button. The system will create a **temporary** directory into the **download** (/download) directory with all the requested files and notify the user with an email on how to access and download it. **The download directory is erased every day at midnight UTC**.

10. Software architecture

The EAW data management website is structured as shown in **figure 9**. The user interacts with the Data Management System through a **web interface** written in Javascript and based on the ExtJS Framework. The web interface's purpose is to **collect user's input**, check that the correct data types and format are used and then **pass the requests to the backend**. The

business logic is implemented using the Python programming language and the Django Web Framework. It's role is to **process user requests, query the database and return the results** back to the Javascript web interface to be displayed for the user. The last part is composed by the PostgreSQL **database** that holds all the data related to samples, stations, comments and the samples' sequences location (see **figure 10**), while **samples' raw sequences** are stored on a dockerized file server in a secured directory.

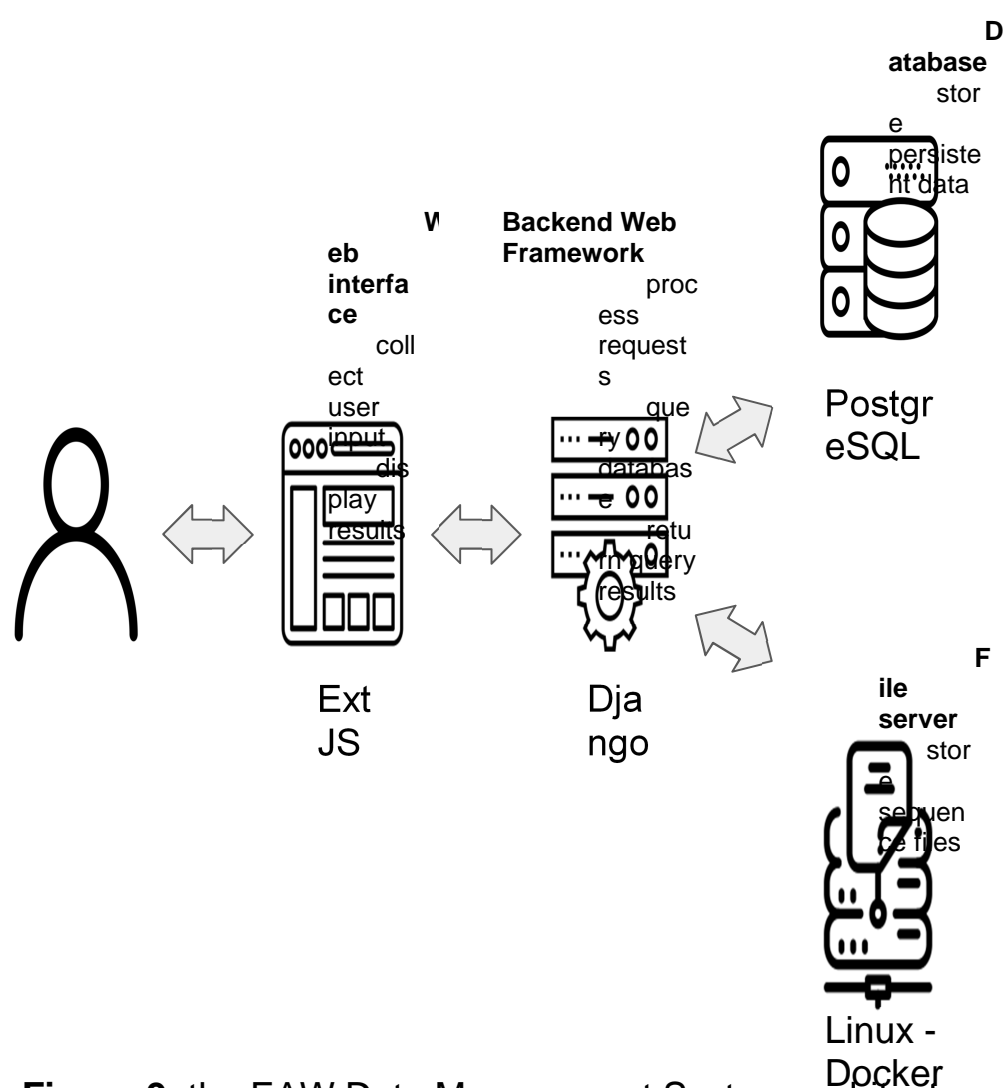


Figure 9: the EAW Data Management System architecture

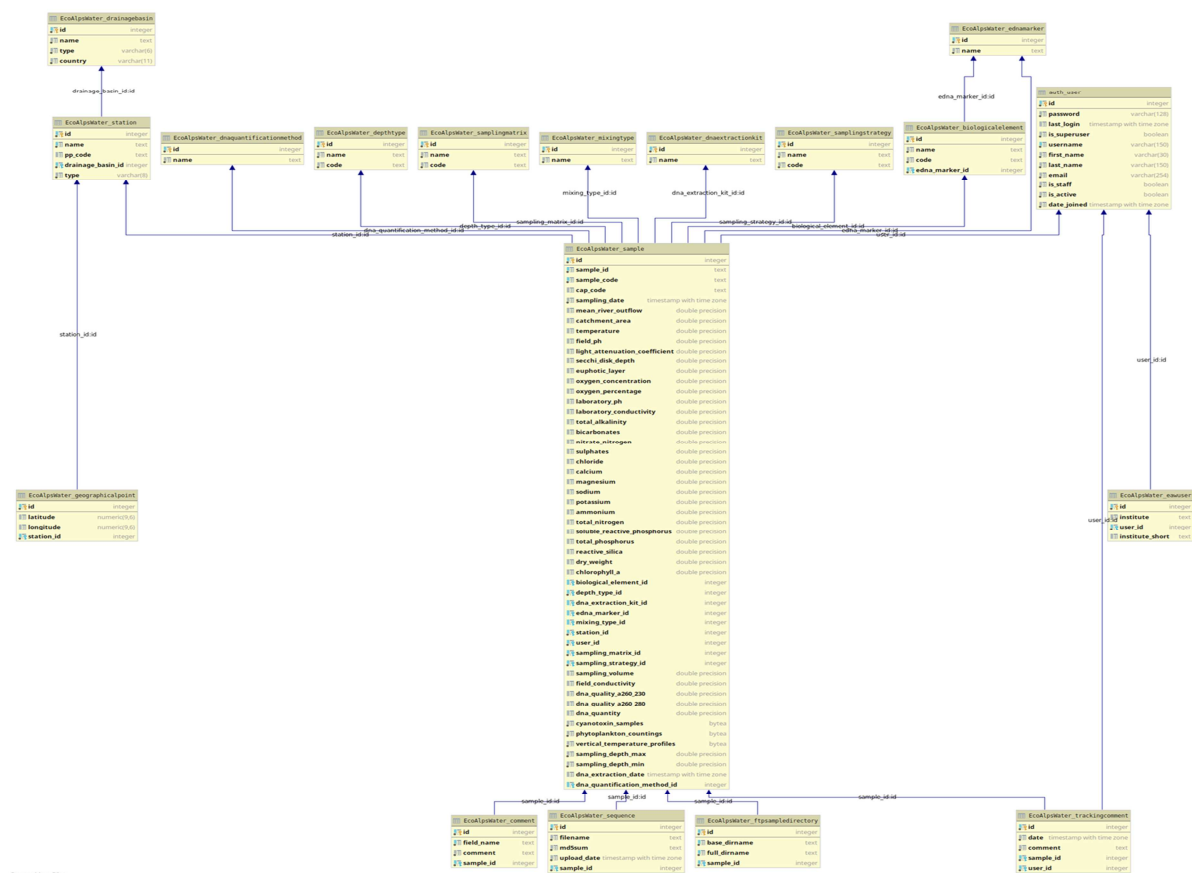


Figure 10: the EAW PostgreSQL database structure