

## User guide:

# Uploading *in situ* CWR passport data to EURISCO

Author: Suman Kumar, Stephan Weise  
Date: 2024-01-08  
Version: 1.4

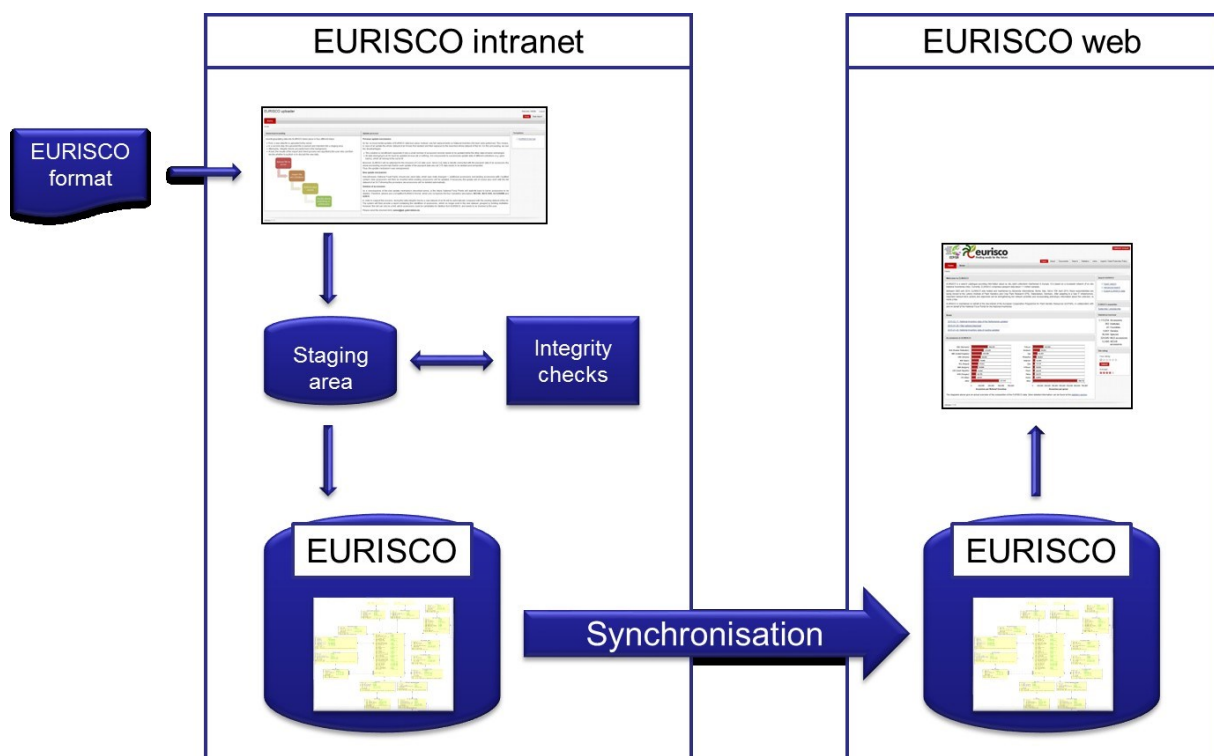
## Content

1	Aim and scope .....	2
2	Procedure .....	2
2.2.1	Steps to upload the Excel file .....	4
2.2.2	File validation errors.....	5
2.3.1	Check errors per descriptor.....	7
2.3.2	Check deletion candidates .....	8
2.3.3	Check taxonomy mappings .....	8
2.3.4	Check AEGIS status changes .....	9

# 1 Aim and scope

The aim of this document is to provide *in situ* National Focal Points (NFPs) with information about how to upload and update their *in situ* National Inventory (NI) datasets in EURISCO.

The general proceeding of an update is shown in the figure below. A National Focal Point formats data and then uploads them via the EURISCO *in situ* CWR intranet to a staging area. In the following steps, the data are checked for integrity and, after approval by the NFP, integrated in EURISCO. After final checks, the data are synchronised to a specific schema for the web frontend and are then available to the public.



The individual steps are described in the following sections.

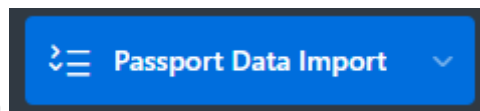
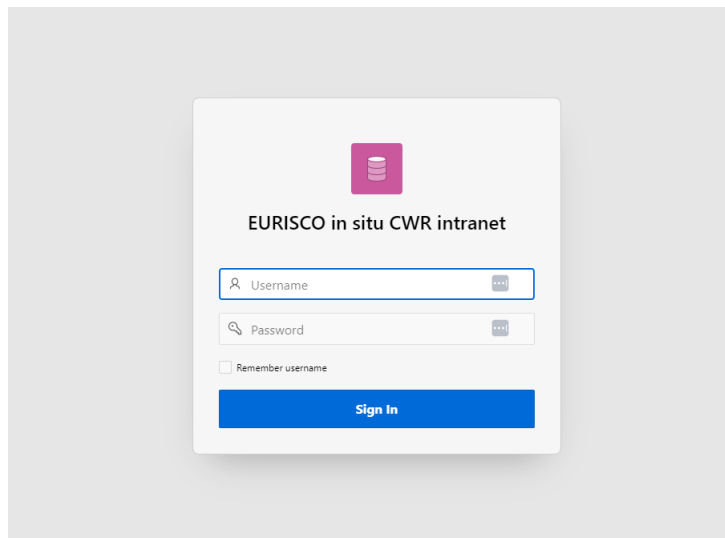
## 2 Procedure

### 2.1 DATA FORMATTING

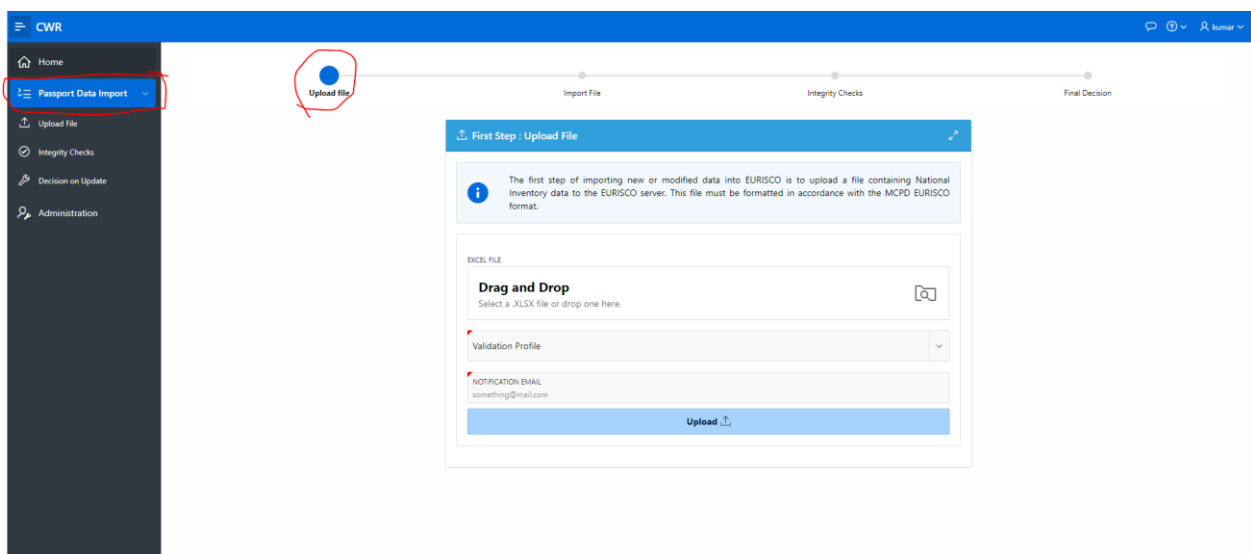
*In situ* CWR passport data to be uploaded to EURISCO need to be formatted in accordance with the *in situ* CWR data standard, which is available at the documents section of the EURISCO web interface as well as in the EURISCO *in situ* CWR intranet.

## 2.2 DATA UPLOADING

Please login to the EURISCO *in situ* CWR intranet for uploading data.



Expand the navigation menu and click on



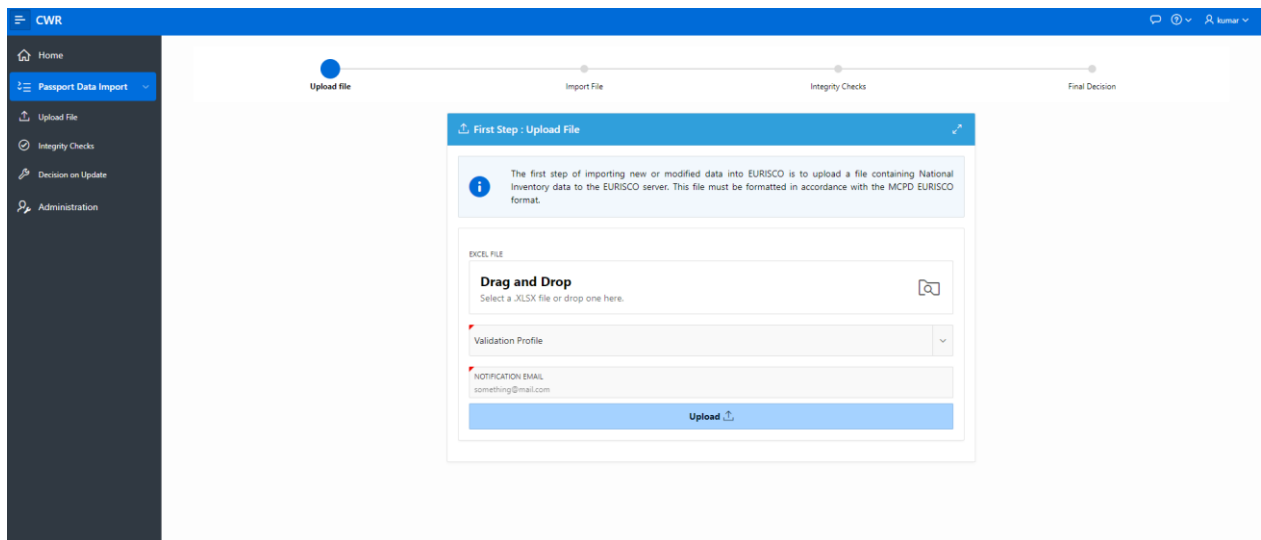
---

*The in situ CWR passport data uploader only supports files in the XLSX format (Excel 2007 and above).*

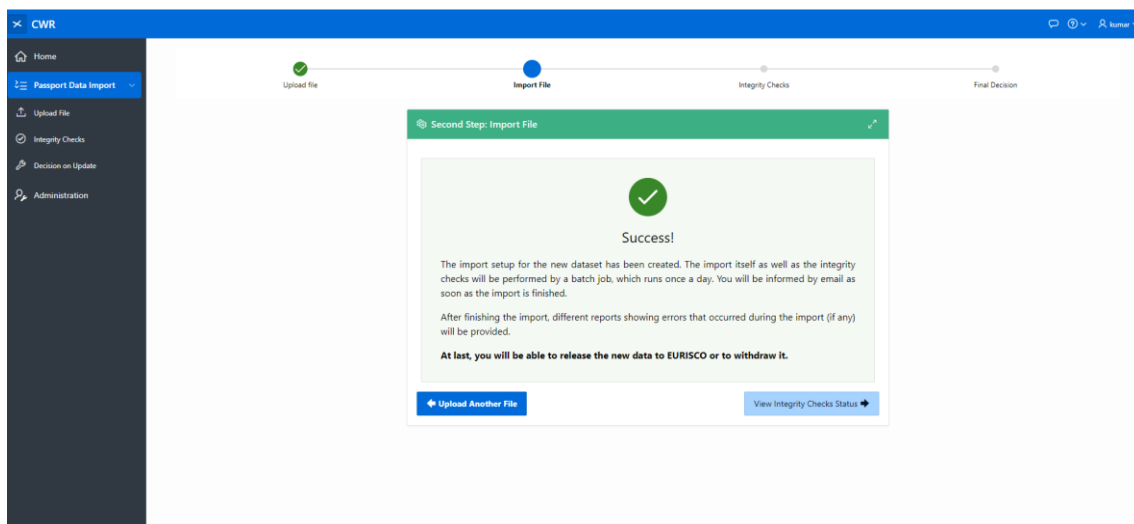
---

## 2.2.1 Steps to upload the Excel file

1. Please choose the Excel file to be uploaded (*as shown below*)
2. Select an appropriate validation profile from the select list:
  - **Errors abort session**  
If an error occurs, the whole import operation will be aborted.
  - **Records with errors are skipped**  
If an error occurs within a record, this record will be skipped. The remaining records will be imported.
  - **Descriptors with errors are skipped**  
If a descriptor value within a record contains an error, only this value will be skipped. The remaining descriptor values of the respective record will be imported. However, in case of an error within one of the mandatory descriptors, the whole record will be skipped.
3. Enter your email address for notification. Then, click on the **“Upload”** button.



During upload, the first step is to check if all column headings conform to the *in situ* CWR data standard. If this is the case, a success message is displayed to the user after a completed upload.



The uploaded Excel file is then parsed on the EURISCO server and the data is written to the EURISCO staging area. Subsequently, all necessary data integrity checks are carried out in the staging area.

This is done with a time delay by a batch job. You will be notified by email as soon as the integrity check results are available.

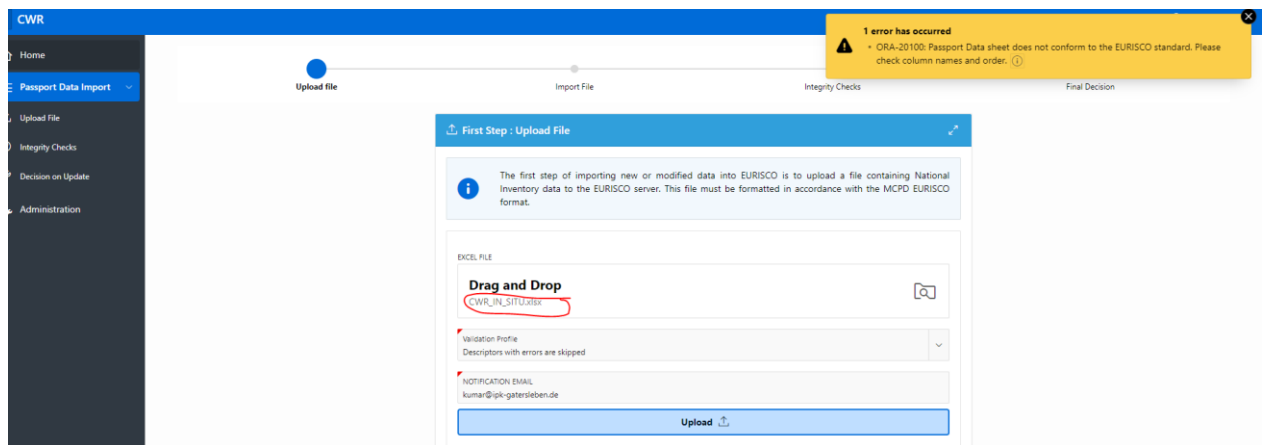
## 2.2.2 File validation errors

Allowed file extension is .xlsx:

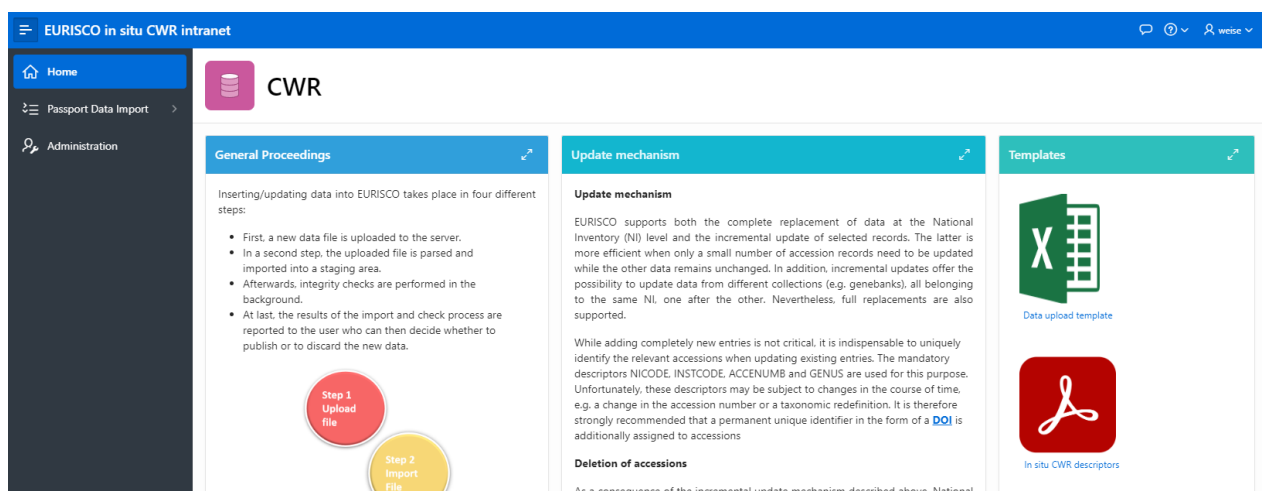
The uploader filters and allows only to select .xlsx files.

Uploading the files without mandatory fields:

As per the *in situ* CWR data standard, the uploaded Excel file must contain the mandatory fields NICODE, INSTCODE, ACCENUMB and GENUS. The upload process fails when the user submits the Excel file with missing mandatory fields.



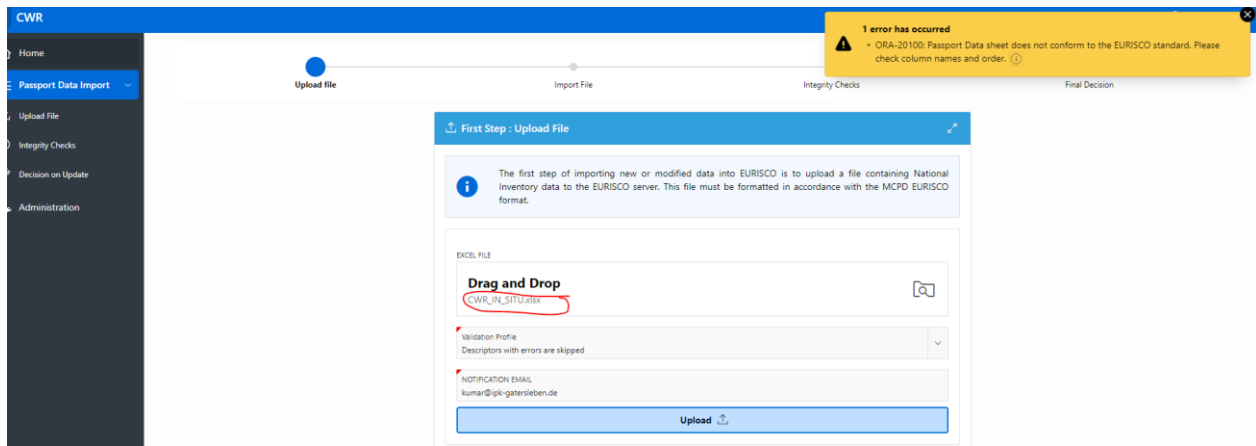
To fix the issue, please adhere to the file format available on the home page under **Templates**.



### Uploading the Excel file with incorrect header names:

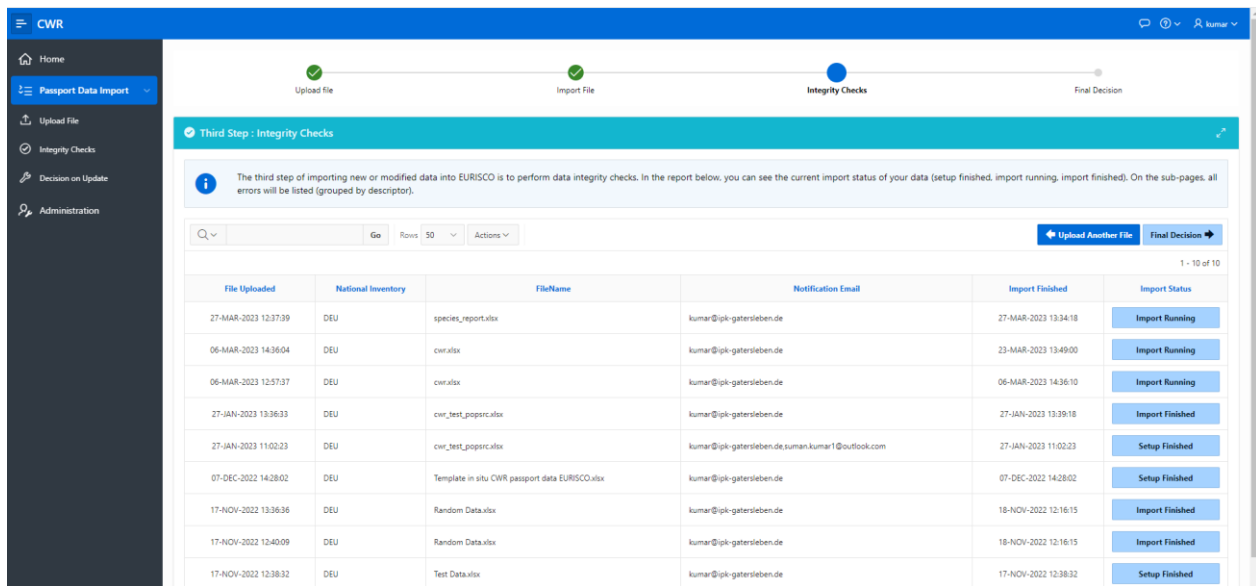
The file upload process fails when the user tries to upload the Excel file with incorrect header names, i.e. **HIST** instead of **HISTORIC** (error shown below).

To fix the issue, please don't change the template while editing.



### **2.3 REVIEW INTEGRITY CHECK RESULTS**

After you received a notification that the data integrity checks have been finished, please login to the EURISCO *in situ* CWR intranet again. Click on the menu item "Passport data import" and then on the sub-menu item "Integrity checks".



In the shown report, all uploaded datasets with their respective import status are listed. After the data integrity checks have been performed, the status will be "Import finished". The next steps now will be to review the errors that might have detected in your data and to decide whether to publish the data as they are (errors will be skipped according to your validation profile) or to rework them.

Please select a dataset and click on the "Import finished" button in order to see the error reports.

The screenshot shows the 'Errors per Descriptor' interface. At the top, there are four progress indicators: 'Upload File' (green checkmark), 'Import File' (green checkmark), 'Integrity Check' (blue circle), and 'Decision on update' (grey circle). Below these, there are three main sections:

- Errors per descriptor:** A table with columns 'Descriptor' and 'Error Count'. It shows 'LIAISONCODE' with 12 errors and 'STORAGE' with 12 errors. A 'Go' button and 'Rows 50' dropdown are visible.
- Deletion Candidates:** A section with a title bar and a large text area. The text explains that the number of accessions in an NI dataset may vary due to different reasons (e.g., removal from a genebank or changes in accession identifiers). It states that as a consequence of the new update mechanism, National Focal Points will explicitly have to name accessions to be deleted. It also mentions that in order to support this process, data integrity checks compare the new dataset with the existing one in EURISCO. A note says: 'Please send the checked list to [weise@ipk-gatersleben.de](mailto:weise@ipk-gatersleben.de)'.
- Taxonomy Check Results:** A table with columns 'Instcode' and 'Total Accessions'. It lists several instcodes and their corresponding total accessions: DEU005 (10), DEU022 (946), DEU078 (18), DEU098 (2857), DEU101 (12019), DEU146 (129748), DEU156 (286), and DEU159 (6242).

This screenshot shows the 'Errors per Descriptor' interface with the 'Taxonomy Check Results' section highlighted in green. The 'Integrity Check' progress indicator is now blue, indicating it is complete. The 'Taxonomy Check Results' section features a large green checkmark icon and the text: 'GRIN & MANSFELD. This dataset contains 5 distinct taxonomic names, which have been checked for correctness against the GRIN Taxonomy as well as against the Mansfeld's World Database of Agricultural and Horticultural Crops. The check results can be found [Ehere](#)'.

### 2.3.1 Check errors per descriptor

At the left side of the browser window, the number of errors is shown, grouped by descriptor. Click on the numbers next to the descriptor names to see all errors for the respective descriptors, e.g. STORAGE. In this example, several STORAGE values do not belong to pre-defined list of values.

The 'Errors Details' interface shows a table with 12 rows of error details. The table has columns for 'Nicode', 'Instcode', 'Genus', 'Accenumb', 'Descriptor', and 'Error Description'. The 'Error Description' column indicates that the 'STORAGE' value is invalid for each row.

Nicode	Instcode	Genus	Accenumb	Descriptor	Error Description
GBR	GBR004	Acis	1033986	STORAGE	Line 2: STORAGE 10 invalid.
GBR	GBR004	Acis	1033986	STORAGE	Line 2: STORAGE 10 invalid.
GBR	GBR004	Acis	666361	STORAGE	Line 3: STORAGE 10 invalid.
GBR	GBR004	Acis	666361	STORAGE	Line 3: STORAGE 10 invalid.
GBR	GBR004	Acis	666350	STORAGE	Line 4: STORAGE 10 invalid.
GBR	GBR004	Acis	666350	STORAGE	Line 4: STORAGE 10 invalid.
GBR	GBR004	Acis	877170	STORAGE	Line 5: STORAGE 10 invalid.
GBR	GBR004	Acis	877170	STORAGE	Line 5: STORAGE 10 invalid.
GBR	GBR004	Acis	791292	STORAGE	Line 6: STORAGE 10 invalid.

Please go through all errors listed for the particular descriptors and check/correct the data file you uploaded. The *in situ* CWR data standard description is available from the start page of the intranet application as well as from the documents section of the EURISCO web application.

### 2.3.2 Check deletion candidates

Updating data in EURISCO does not mean that the whole dataset of a National Inventory will be replaced by the new dataset. Instead, incremental updates will be performed. This means that new accessions will be inserted while existing accessions will be updated. Thus, it is possible to successively update data of different collections (e.g. genebanks), which all belong to the same NI. If necessary, the update will of course also work with the full dataset of an NI. Following this procedure, no accessions will be deleted automatically.

In order to distinguish between new and existing accessions, the four mandatory descriptors NICODE, INSTCODE, GENUS, ACCENUMB are used. Together, they form an identifier, which should be unique within a National Inventory dataset.

Nevertheless, values of these four descriptors may change over time, e.g. a typo in the genus name may be corrected in the new dataset. In such a case, it is not possible for the import system to make a connection between the old record and the new record (with the corrected genus name). Thus, National Focal Points will explicitly have to name accessions to be deleted.

In order to support this process, during the data integrity checks a new dataset of an NI will be automatically compared with the existing dataset of this NI. The system will then provide a report containing the identifiers of accessions, which no longer exist in the new dataset, grouped by holding institution.

---

*This list can only be a hint as to which accessions could be candidates for deletion from EURISCO, and needs to be checked by the NFP.*

---

The screenshot displays the 'Errors per Descriptor' and 'Deletion Candidates' sections of the EURISCO interface. The 'Deletion Candidates' section is highlighted with a red circle and arrows. The interface shows progress bars for 'Upload File', 'Import File', 'Integrity Check', and 'Decision on update'. The 'Deletion Candidates' table lists holding institutions (Instcode) and their total accessions.

Instcode	Total Accessions
DEU005	10
DEU022	946
DEU078	18
DEU098	2857
DEU101	12019
DEU146	129748
DEU156	286
DEU159	6242
DEU271	15049
DEU363	235

### 2.3.3 Check taxonomy mappings

Since the correctness of scientific plant names poses one of the most important challenges, the taxonomic terms provided to EURISCO are checked automatically against taxonomic repositories, currently against GRIN-Taxonomy and Mansfeld's World Database of Agricultural and Horticultural Crops, which represent opposite taxonomic schools.



Reports of matching/non-matching results of the provided terms and suggestions of closest existing matches are generated. These check results should only be considered as suggestions, but may help to improve the dataset.

Errors per Descriptor

Upload File Import File Integrity Check Decision on update

Deletion Candidates Taxonomy Check Results

**Errors per descriptor**

Descriptor	Error Count
LIAISONCODE	12
STORAGE	12

1 - 2 of 2

**Taxonomy Check Results**

GRIN & MANSFELD

This dataset contains 5 distinct taxonomic names, which have been checked for correctness against the GRIN Taxonomy as well as against the Mansfeld's World Database of Agricultural and Horticultural Crops. The check results can be found [Here](#)

Taxonomy Results

Upload File Import File Integrity Check Decision on update

**Taxonomy check results**

The distinct taxonomic names found in the new dataset have been checked against two taxonomic repositories - the GRIN Taxonomy as well as the Mansfeld's World Database of Agricultural and Horticultural Crops. If the provided taxon could not be matched, the closest match to the respective repository - GRIN or Mansfeld - is shown. The results of these checks should only be considered as suggestions, but may help to successively improve the dataset.

1 - 5 of 5

Genus	Species	Spauthor	Grin Match	Closest Grin Match	Mansfeld Match	Closest Mansfeld Match	Total Accessions
Acis	autumnalis	(L.) Sweet	no exact match	Acis autumnalis (L.) Herb.	no exact match		1
Acis	fabrei	(Quétel & Girard) Uedé, A.P.Davis & M.B.Crespo	no exact match	Leucojum	no exact match		1

### 2.3.4 Check AEGIS status changes

The AEGIS status of EURISCO accessions is checked automatically. According to the AEGIS principles, the reasons for de-flagging of accessions need to be documented. In case of de-flagging an accession previously flagged for AEGIS, the user will be asked for providing a change reason.

Errors per Descriptor

Upload File Import File Integrity Check Decision on update

Deletion Candidates Taxonomy Check Results AEGIS Status Change

**Errors per descriptor**

Descriptor	Error Count
ACCENUMB	1
GENUS	1
INSTCODE	1
INSTNAME	1
LIAISONCODE	1
NICODE	1
POPSRC	6
STORAGE	2

1 - 8 of 8

**AEGIS Status Change**

During the data integrity checks, the AEGIS status of the accessions of the new dataset was compared with the current dataset in EURISCO. The report below shows accessions existing in EURISCO, which were labelled as part of AEGIS. These accessions are either missing from the new dataset or the AEGISSTAT value changed. In accordance with the AEGIS principles, please enter a reason for removing these accessions from AEGIS. This information will be sent to the ECPGR Secretariat automatically.

Search: All Text Columns Go Actions Edit Save

Approval Timestamp	Nicode	Instcode	Genus	Accenumb	Old Status	New Status	Change Reason	Approved
No data found								

## 2.4 FINAL DECISION ABOUT UPDATE

After error checking, the final step is to decide whether the new data should be published in EURISCO or not. Please click on the sub-menu item “Decision on update”.

National Inventory	Filename	File Uploaded	Notification Email	End Of Import	Validation Profile	Import Status	Decision
DEU	Random Data.xlsx	2022-11-17 12:40:09	kumar@ipk-gatersleben.de	2022-11-18 12:16:15	Descriptors with errors are skipped	Import finished	Update/Discard
DEU	Random Data.xlsx	2022-11-17 13:36:36	kumar@ipk-gatersleben.de	2022-11-18 12:16:15	Records with errors are skipped	Import finished	Update/Discard
DEU	cwr_test_popsrc.xlsx	2023-01-27 13:36:33	kumar@ipk-gatersleben.de	2023-01-27 13:39:18	Errors abort session	Import finished	Update/Discard

Here, you will find a report showing all uploaded datasets that have been checked for data integrity. Please click on the “update/discard” button.

DISCARD or UPDATE

Your uploaded file (Random Data.xlsx) has been checked for integrity and can now be used to update the data of your National Inventory in EURISCO. Only accessions listed in your file will be updated. All other accessions of your National Inventory in EURISCO, which are not covered by the input file, will remain untouched in EURISCO.

The final update will run as a batch job in the background.

If your dataset needs revision, please click on “Discard data” and repeat the upload and check process with a reworked file. If your dataset does not contain errors or if you wish to skip the errors during the update, please click on “Update EURISCO data.” Your data, which are still stored in a staging area, will then be used for updating EURISCO. Afterwards, the data will be synchronised to the EURISCO web frontend as described in section 1 (time lag; not in main business hours).