

http://eurisco.ecpgr.org

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	Proc	cedure	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.

EURISCO in situ CWR intranet A Username Bassword Image: Compared and the semantic semant				
Remember username Sign In	EURISCO in s	situ CWR intrar	net	
Remember username Sign In	R Username			
Sign In	Remember username			
		Sign In		



Expand the navigation menu and click on

≓ CWR				େ ତ ∽ ୧ kumar ∽
≷≘ Passport Data Import — ∨	Upload file	Import File	Integrity Checks	Final Decision
1 Upload File		irst Sten - Unload File	ر	
 Integrity Checks 			÷	
🥬 Decision on Update		The first step of importing new or modified	data into EURISCO is to upload a file containing National must be formatted in accordance with the MCPD FIIPISCO	
$\beta_{\mathbf{z}}$ Administration		format.	must be formated in accordance with the mer b contaco	
	EX	CEL PILE		
		Drag and Drop Select a XLSX file or drop one here.	[]	
	Y	alidation Profile	~	
	1	IOTIFICATION EMAIL		
		Ua	oad . [↑] .	
		-		

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.

≓ CWR				ଢ଼େ ତେ ∨ ଥ kum
Ġ Home			0	
े 🗄 Passport Data Import 🗠	Upload file	Import File	Integrity Checks	Final Decision
🗘 Upload File				
 Integrity Checks 			-	
Decision on Update		The first step of importing new or modi Inventory data to the EURISCO server. This	ied data into EURISCO is to upload a file containing National file must be formatted in accordance with the MCPD EURISCO	
$\beta_{\mathbf{p}}$ Administration		format.		
		EV/DI EN E		
			~	
		Select a .XLSX file or drop one here.	لم	
		Validation Profile	~	
		NOTIFICATION EMAIL something@mail.com		
			Upload 🔿	

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.

Itome Itome <t< th=""><th></th></t<>	
Paraport Data Import Update file Import File Integrity Checks I Update File Import File Import File I Update File Import File Import File	
∴ Updade File	
Integrity Creats Processor on Update	
Decision on Update	
P _p Administration	
Success!	
The import setup for the new dataset has been created. The import itself as well as the integrity checks will be performed by a batch job, which runs once a day. You will be informed by email as soon as the import is finished.	
After finishing the import, different reports showing errors that occurred during the import (if any) will be provided.	
At last, you will be able to release the new data to EURISCO or to withdraw it.	
◆ Uplead Another File View Integrity Checks Status ◆	

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.

CWR			1 error has occurred	• • • • • • • • • • • • • • • • • • •
} Home	•		ORA-20100: Passport check column names	Data sheet does not conform to the EURISCO standard. Please and order.
Passport Data Import 🛛 🗸	Upload file	Import File	Integrity Checks	Final Decision
, Upload File		↑ First Step : Upland File	,	
Integrity Checks		E First Step : Opload File	¥.	
Decision on Update		The first step of importing new or modified data	into EURISCO is to upload a file containing National	
Administration		format.	t be formatted in accordance with the MCPD EURISCO	
		EXCEL FILE		
		Drag and Drop CWR_IN_SITUX#XX	<u>م</u>	
		Validation Profile Descriptors with errors are skipped	~	
		NOTIFICATION EMAIL kumar@ipk-gatersleben.de		
		Upload	Δ.	

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.

CWR			1 error has occurred	0
Home	•		ORA-20100: Passport D check column names an	lata sheet does not conform to the EURISCO standard. Please nd order. 🕕
Passport Data Import 🛛 🗸	Upload file	Import File	Integrity Checks	Final Decision
, Upload File			1	
Integrity Checks				
Decision on Update		The first step of importing new or modified data into Inventory data to the EURISCO server. This file must be	EURISCO is to upload a file containing National formatted in accordance with the MCPD EURISCO	
Administration		format.		
		INTE DUE		
		Drag and Drop	۵	
		Validation Profile Descriptors with errors are skipped	~	
		NOTIFICATION EMAIL kumar@ipic-gatersleben.de		
		Upload 🛧		

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 submenu item "Integrity checks".

CWR						오 ⓒ ~ Զkums
Home		2				
Passport Data Import 🛛 🗸	Uplo	ad file	Import File	Integrity Checks	Final Decis	ion
lpload File	Third Step : Integrity Ch	necks				
legrity Checks						
ecision on Update	The third step of in errors will be listed	mporting new or modified o d (grouped by descriptor).	data into EURISCO is to perform data integrity checks. In the re	port below, you can see the current import status of your data (setup fir	ished, import running, import finish	ed). On the sub-pages, all
Administration						
	Qv	Go Rows	50 V Actions V		The Upload Anothe	Final Decision
	File Uploaded	National Inventory	FileName	Notification Email	Import Finished	Import Status
	27-MAR-2023 12:37:39	DEU	species_report.xlsx	kumar@ipk-gatersleben.de	27-MAR-2023 13:34:18	Import Running
	06-MAR-2023 14:36:04	DEU	cwr.xlsx	kumar@ipk-gatersleben.de	23-MAR-2023 13:49:00	Import Running
	06-MAR-2023 12:57:37	DEU	cwr.dsx	kumar@ipk-gatersleben.de	06-MAR-2023 14:36:10	Import Running
	27-JAN-2023 13:36:33	DEU	cwr_test_popsrc.xlsx	kumar@ipk-gatersleben.de	27-JAN-2023 13:39:18	Import Finished
	27-JAN-2023 11:02:23	DEU	cwr_test_popsrc.xlsx	kumar@ipk-gatersleben.de,suman.kumar1@outlook.com	27-JAN-2023 11:02:23	Setup Finished
	07-DEC-2022 14:28:02	DEU	Template in situ CWR passport data EURISCO.xlsx	kumar@ipk-gatersleben.de	07-DEC-2022 14:28:02	Setup Finished
	17-NOV-2022 13:36:36	DEU	Random Data.xlsx	kumar@ipk-gatersleben.de	18-NOV-2022 12:16:15	Import Finished
	17-NOV-2022 12:40:09	DEU	Random Data.xlsx	kumar@ipk-gatersleben.de	18-NOV-2022 12:16:15	Import Finished
	17-NOV-2022 12:38:32	DEU	Test Dataxisx	kumar@ipk-gatersleben.de	17-NOV-2022 12:38:32	Setup Finished

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.

	Errors per Descriptor						
נז \$≡ ב	Upload File		Import File	Integrity Check		Cecision on update	-
ی ور	Errors per descriptor	2	Deletion	Candidates	Taxor	nomy Check Results	
يو. موج	Go Rows 50) ~ Actions ~	Deletion Candidates			1. 1	
	Descriptor	Error Count	The number of accessions contained in an	NI dataset may vary due to different reasons.	G	io Rows 50 V Actions V	
	LIAISONCODE	12	identifiers (unique combination of INSTCC the new update mechanism, no accessions	DDE, GENUS, ACCENUMB) may change. With will be deleted automatically from EURISCO.	Instcode	Total Accessions	
	STORAGE	12	As a consequence of the new update me	echanism, National Focal Points will explicitly	DEU005	10	
		1 - 2 of 2	In order to support this process, during th	e data integrity checks your new dataset was	DEU022	946	
			automatically compared with the existing overview of accessions, which no longer e	dataset in EURISCO. This report provides an exist in the new dataset, grouped by holding	DEU078	18	
			However, this list can only be a hint, which	h accessions could be candidates for deletion	DEU098	2857	
			from EURISCO, and needs to be checked should be updated, this list may contain m	 Especially if only a part of the NI dataset any false positive entries. 	DEU101	12019	
			Please send the checked list to weise@ipk	r-gatersleben.de	DEU146	129748	
					DEU156	286	
					DEU159	6242	

CIA/D				
Errors per Descriptor				x
Upload File		Import File	Integrity Check	Decision on update
		Deletion (andidates	Taxonomy Check Results
Errors per descriptor	2			
Go Rows	50 V Actions V	Taxonomy Check Results		2
Descriptor	Error Count			
LIAISONCODE	12			
STORAGE	12		GRIN & N	/ANSFELD
	1 - 2 of 2	This dataset contains 5 distinct tax World Database of Agricultural and	onomic names, which have been checked Horticultural Crops. The check results can l	for correctness against the GRIN Taxonomy as well as against the Mansfeld's be found Elver

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.

				×
Uplaaf Fie Import Fie		Integrity Check	Decision on update	
Go Rows 50 V A	Actions ~			
Instcode	Genus	Accenumb	Descriptor	Error Description
GBR004	Acis	1033986	STORAGE	Line 2: STORAGE 10 invalid.
GBR004	Acis	1033986	STORAGE	Line 2: STORAGE 10 invalid.
GBR004	Acis	666361	STORAGE	Line 3: STORAGE 10 invalid.
GBR004	Acis	666361	STORAGE	Line 3: STORAGE 10 invalid.
GBR004	Acis	666350	STORAGE	Line 4: STORAGE 10 invalid.
GBR004	Acis	666350	STORAGE	Line 4: STORAGE 10 invalid.
GBR004	Acis	877170	STORAGE	Line 5: STORAGE 10 invalid.
GBR004	Acis	877170	STORAGE	Line 5: STORAGE 10 invalid.
GBR004	Acis	791292	STORAGE	Line 6: STORAGE 10 invalid.
	File	Bit Bit S0 Actions Instance Genus GBR004 Actis GBR004 Actis	Import File Import File 6 o Ross 50 × Actions × Matcode Actions × GBR004 Actions ×	File Import File Import File 60 Rons 20 × Actions 20 × Actions 100 Rons 20 × Actions 20 × Actions 100 Rons 20 × Actions 20 × Actions 101 Rons 20 × Actions 20 × Actions 102 Rons Acis 103386 510RAGE 102 Rons Acis 103386 510RAGE 102 Rons Acis 666581 510RAGE 102 Rons Acis 666530 510RAGE 102 Rons Acis 666530 510RAGE 102 Rons 877170 510RAGE 103086 102 Rons 877170 510RAGE 1030RAGE 102 Rons 877170 510RAGE 1030RAGE 102 Rons 877170 510RAGE 10408 102 Rons 877170 510RAGE 10408 102 Rons 877170 510RAGE 104

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.

Errors per Descriptor					X
Upload File		Import File	Integrity Check		Decision on update
Errors per descriptor	e.	Det	ation Candidates	Is	axonomy Check Results
Go Rows Descriptor	50 V Actions V Error Count	Deletion Candidates The number of accessions contained For example, accessions could be re	in an NI dataset may vary due to different reasons. moved from a certain genebank or the accession		Ge Rows 50 V Actions V
LIAISONCODE	12	identifiers (unique combination of II the new update mechanism, no access	ISTCODE, GENUS, ACCENUMB) may change. With sions will be deleted automatically from EURISCO.	Instcode	Total Accessions
STORAGE	12 1 - 2 of 2	As a consequence of the new upda have to name accessions to be delete	te mechanism. National Focal Points will explicitly d from the system.	DEU005	946
		In order to support this process, duri automatically compared with the exi- overview of accessions, which no lor	ng the data integrity checks your new dataset was sting dataset in EURISCO. This report provides an ager exist in the new dataset, grouped by holding	DEU078	18
		institution. However, this list can only be a hint,	which accessions could be candidates for deletion	DEU098	2857
		from EURISCO, and needs to be ch should be updated, this list may cont	ecked. Especially if only a part of the NI dataset ain many false positive entries.	DEU101	12019
		Please send the checked list to weis	e@ipk-gatersleben.de	DEU146	129748
				DEU156	286
				DEU159	6242
				DEU271	15049
				DEU363	235

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.

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 De	escriptor							×
	Upload	File	Import File		Integrity Check		Decision on updat	2
				Deletion C	andidates		Taxonomy Check Results	
Errors per	r descriptor	2						
	Go	Rows 50 V Actions V	Taxonomy C	heck Results				2
	Descriptor	Error Count						
LIAISONCODE		12]		\rightarrow	\checkmark		
STORAGE		12]		GRIN 8	& MANSFELD		
		1 - 2 of 2	This datase	t contains 5 distinct tax	onomic names, which have been cheo	ked for correctness again	st the GRIN Taxonomy as well as a	against the Mansfeld's
			World Data	base of Agricultural and	Horticultural Crops. The check results	can be found Bhere		
Taxonomy Re	esults							\times
	Upload	File	Import File		Integrity Check		Decision on upd	ate
	opicau	1 110	mportric		integrity enco		occourt on ope	
Taxonomy	check results							
	The distinct taxonomic i	names found in the new dataset have been checked	I against two taxonomic	: repositories - the GRIN Tax	conomy as well as the Mansfeld's World Da	tabase of Agricultural and Ho	rticultural Crops. If the provided taxon	could not be matched, the
0	closest match to the resp	pective repository - GRIN or Mansfeld - is shown.	whelp to successively in	more the dataset				
	The results of these chec	its should only be considered as suggestions, but his	ay help to successively in	nprove the dataset.				
Qv		Go Rows 50 \checkmark Actions \checkmark						
								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	(Quézel & Girerd) Lledó, A.P.Davis & M.B	Crespo	no exact match	Leucojum	no exact match		1
Asia	airssaarir	(Anderine) Handé A.D.Douris & M.D.Con		an over metric	Lauraium	an nort match		· · · · · · · · · · · · · · · · · · ·

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												3
Upload File		Imp	Sort File Deletion Can	didates			Integrity	Check	k Results	(Decision on update	
Errors per descriptor	2	AEGI	S Status Change									2
Descriptor	Error Count		During the data integ existing in EURISCO,	grity check which wer	s, the AEG e labelled	IS status of as part of A	the accessions o EGIS. These acce	f the new datase ssions are either	t was compare missing from	ed with the current dataset in the new dataset or the AEGIS	n EURISCO. The report below shows SSTAT value changed.	accessions
GENUS	1	-	In accordance with automatically.	the AEGIS	5 principle	s, please e	nter a reason fi	or removing the	ese accessions	from AEGIS. These inform	ation will be sent to the ECPGR	Secretariat
INSTCODE	1	Qv	Search: All Text Columns		Go	Actions \vee	Edit Save					🕞 Reset
INSTNAME	1		Approval Timestamp	Nicode	↑L	Genus	Accenumb	Old Status	New Status	Change Reason	Aş	proved
LIAISONCODE	1							Q				
NICODE	1							IVO Gata foi	and			
POPSRC	6											
STORAGE	2											
	1 - 8 of 8											

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".

× cwr								두 ⑨ ~ 옷 kumar				
යි Home						0						
$\xi \equiv $ Passport Data Import $ \sim $		Upload file		Import File		ntegrity Checks	Final Decision					
Upload File Integrity Checks	Fourth Step : Publish	h or Discard new data						2				
Decision on Update	After you have reviewed the errors which occurred during the data integrity checks, the fourth step of importing new or modified data into EURISCO is now to decide either to publish the new data to the EURISCO web frontend or to discard the imported data. In the latter case, the whole update procedure should be repeated with a debuaded data set.											
$ ho_{\mu}$ Administration												
	QŸ	Go Row	s 50 V Actions V					Upload Another File				
	National Inventory	Filename	File Uploaded	Notification Email	End Of Import	Validation Profile	Import Status	Decision				
	DEU	DEU Random Datautax 2222-11-17 124009 kumar@ipik-gatersiben.de DEU Random Datautax 2222-11-17 133636 kumar@ipik-gatersiben.de		kumar@ipk-gatersleben.de	2022-11-18 12:16:15	Descriptors with errors are skipped	Import finished	Update/Discard				
	DEU			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				
								1 - 3				

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

Cint											
යි Home			0		•						
2≘ Passport Data Import ~			Upload file		Import File		Integ	rity Check	3	Final Deci	sion
1 Upload File	A Fourth	Sten · Public	or Discard new data					_			
Integrity Checks		otep : Publisi		Final Decision		\times					
Decision on Update	0	After you have	reviewed the errors which In the latter case, the wh						to decide either to publish th	e new data to the EURISCO web	o frontend or to discard
مر Administration				Upload File	Import File	Integrity Check	Decision on update				
	Qv		Go								🔶 Upload Another
	Nationa	al Inventory	Filename			R			Validation Profile	Import Status	Decision
	DEU		Random Data.xlsx		0				with errors are skipped	Import finished	Update/Discard
	DEU		Random Dataxisx		DISCARD	or UPDATE			h errors are skipped	Import finished	Update/Discard
	DEU		cwr_test_popsrc.xlsx	Your uploaded f	ile (Random Data.xlsx)	has been checked for	integrity and can		t session	Import finished	Update/Discard
				accesssions liste	d in your file will be	updated. All other a	ccessions of your				
				National Invento remain untouch	ory in EURISCO, which ed in EURISCO.	are not covered by	the input file, will				
				The final update	will run as a batch job	in the background.					
				Discard data		Update	EURISCO data 🗹				

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).