

Project:	EURISCO	Date:	2024-01-30
From:	Stephan Weise (IPK)	To:	Lorenzo Maggioni (ECPGR Secretary) Theo van Hintum (Head, EURISCO Advisory Board)
CC:	Markus Oppermann (IPK)		

This document describes the work packages to be processed in 2024 for hosting and maintaining the European Search Catalogue for Plant Genetic Resources (EURISCO), and for coordinating the EURISCO network.

Content

1	Work plan for the EURISCO development.....	2
1.1	WP1: Hosting of the EVA infrastructure.....	2
1.2	WP2: Extension of the intranet support mechanisms for data providers (perm. activity)	2
1.2.1	EURISCO update mechanisms for <i>ex situ</i> passport data and phenotypic data	2
1.2.2	EURISCO backend for <i>in situ</i> CWR passport data	2
1.2.3	Update how-to documents	2
1.3	WP3: Extension of functionality of the public EURISCO application (perm. activity)	2
1.3.1	Extension of EURISCO frontend for <i>in situ</i> CWR passport data.....	2
1.3.2	Reorganisation of the search and display of phenotypic data	3
1.3.3	Continuous development of the existing public web interface	3
1.4	WP4: Documentation and planning of the next steps	4
2	Schedule	5
3	Coordination tasks.....	5
3.1	Coordination task 1: EURISCO network development maintenance	5
3.2	Coordination task 2: Coordination of the EURISCO development	6
3.3	Coordination task 3: Participation in project proposals.....	6

1 Work plan for the EURISCO development

1.1 WP1: Hosting of the EVA infrastructure

The hosting of the intranet infrastructure of the European Evaluation Network (EVA) will be continued. Apart from the upcoming integration of a DivBrowse instance for the visualisation of genotyping data, no further activities are planned for the time being. Necessary bug fixes are carried out independently of this.

1.2 WP2: Extension of the intranet support mechanisms for data providers (permanent activity)

1.2.1 EURISCO update mechanisms for *ex situ* passport data and phenotypic data

Both the upload/update mechanism for *ex situ* passport data and the upload/update mechanism for phenotypic data will be maintained and developed on an ongoing basis. Any errors that occur during operation will be rectified. Requests from users for modifications or extensions will be checked for feasibility and implemented promptly or included in longer-term planning.

1.2.2 EURISCO backend for *in situ* CWR passport data

The upload/update mechanism for *in situ* CWR data was implemented during the last year and has only been used by a small pilot group of data providers. It would be beneficial to wait for the acceptance of other users and incorporate their feedback into the continuous further development.

1.2.3 Update how-to documents

If it is necessary to adapt the background mechanisms mentioned above, the corresponding guideline documents for data providers will be adapted.

1.3 WP3: Extension of functionality of the public EURISCO application (permanent activity)

1.3.1 Extension of EURISCO frontend for *in situ* CWR passport data

EURISCO was extended for data on *in situ* CWR populations as part of the project 'Extension of EURISCO for Crop Wild Relatives (CWR) *in situ* data and preparation of pilot countries' data sets'. This mainly pertains to the backend for data uploading and management. However, the search capabilities of the public EURISCO web application for *in situ* CWR data are currently limited. At present, users can only search for all populations of a National Inventory, for taxonomic terms or a combination of both. Additional filters can be applied in a subsequent step using a faceted search. In addition, all previous features that search the entire EURISCO dataset, which includes the *in situ* CWR populations as well, can also be used. For this purpose, an additional filter has been implemented for each of the existing reports, which allows the number of hits to be limited to *in situ* material. The accession details pages in EURISCO have also been extended to display the additional information. Building on this preliminary work, the EURISCO application is to be extended specifically for *in situ* CWR data.

The following activities are planned:

- Implementation of a dedicated search page that specifically offers the descriptors of the *in situ* CWR data standard as search criteria.
- Development of a functionality that enables the download of all details of a group of populations as a result of a search.
- Revision of the statistics section on the EURISCO page; separation into *ex situ*, *in situ* and total.
- Carrying out a survey among the participants of the pilot project for additional use cases, evaluating them with regard to feasibility and implementation.

In addition, the previous full-dump download feature has to be revised so that there are separate dumps (MS Access and CSV in each case) for *ex situ* and *in situ* data. This is necessary because the two data standards are not completely congruent.

1.3.2 Reorganisation of the search and display of phenotypic data

Together with the decision to expand EURISCO to include phenotypic data, the ECPGR Doc&Info WG decided in 2014 to standardise only the data exchange format, but not the data itself (especially traits and methods). The background to this was the desire to get a critical mass of data together first, without which further discussion did not seem expedient. This has been achieved, but new challenges have been looming for some time now. As of today, EURISCO contains 2,729,780 records of phenotypic data points from 74 datasets, comprising 3,920 experiments and 9,764 (!) traits.

The last point in particular, the large number of traits, poses major challenges for users of the system. These include many identical/similar traits, which differ due to different naming/spelling or differences in the definition of the methods used. Mapping the different traits to each other cannot really be automated (e.g. with the help of ontologies), but requires a substantial amount of time on the part of the data providers, which cannot be realised. Although the current search options in EURISCO allow some filters to be set, they search across the entire database. Due to the large number of experiments and traits, this is confusing for users and leads to dissatisfaction. Added to this is the fact that the data standard used for phenotypic data only provides for a small amount of metadata. This makes it almost impossible to compare data across datasets.

It therefore seems sensible to reorganise the search for phenotypic data so that it is only possible within the experiments of a dataset. Before implementation, various options will be examined as to how this can be realised in the most efficient and user-friendly way.

1.3.3 Continuous development of the existing public web interface

In addition to necessary maintenance and performance tuning, the public web interface of EURISCO is continuously being expanded. This is done on the basis of feedback/wishes from the users of the system.

To this end, a structured user survey was conducted as part of last year's EURISCO training workshop. While the majority of respondents were positive about the design, content and search options of the web application, there were also some requests for improvement/expansion:

- General visual overhaul (user preferences change over time) and, if possible, more intuitive operation.
- Clearer structuring of the information provided.

- Display of photos.
- Additional fields for the standard search in order to better limit the number of hits.
- Request for a discussion forum.
- Desire for reference to genomic data and corresponding search options
- Additional display of the plant family name.
- If possible, support synonym search for phenotypic data (currently only for passport data).

The challenges regarding the search for phenotypic data have already been mentioned above.

There was also a request for additional information regarding collection strategy and availability of material. EURISCO is not the right place for the first point; this is an autonomous decision of the respective collecting institution and EURISCO does not have any further information on this. The availability of material, in contrast, is not part of the EURISCO-MCPD standard, but has been the subject of controversial debate for years. This could be a discussion point for this year's Advisory Board meeting.

The usefulness and feasibility of the other requests were discussed. The idea of a discussion forum should not be pursued further for the time being. Discussion forums have to be actively moderated and appropriate guidelines for their use must be drawn up. There is currently no capacity for either. Linking with genomic data is currently being worked on as part of the AGENT project. Following a positive evaluation of the solution developed, it should be adopted for EURISCO in the future rather than developing our own.

The following work is planned:

- Visual revision of the design of the web application with a focus on more intuitive design and clearer structuring of content.
- Extension to display photos at accession level (max. 5), but not for phenotypic data.
- Revision of the standard searches with provision of additional filter options (possibly again as part of an advanced search, as it already existed in an older version).
- Display of the plant family name (this could be done via an internal mapping without having to change the data standard).
- Feasibility test regarding the use of the synonym search and fuzzy search for phenotypic data as well; depending on the estimated effort, implementation in the current year or planning for the following year.

In addition, changes and additions will be made on an ongoing basis.

1.4 WP4: Documentation and planning of the next steps

In parallel to the implementation processes, all developments will be documented.

Furthermore, the next development steps (WP1-3) will be specified in detail.

2 Schedule

Apart from continued hosting, no further specific activities are planned for work packages 1 and 2 (hatched areas). Should a need arise, operational plans will be made in the short term.

	Work package	2024			
		Q1	Q2	Q3	Q4
1	Hosting of the EVA infrastructure				
	Continued hosting of the EVA infrastructure				
2	Intranet support mechanisms for data providers				
	Maintenance of the EURISCO backend for <i>ex situ</i> passport data and phenotypic data				
	Maintenance of the EURISCO backend for <i>in situ</i> CWR passport data				
	Update how-to documents				
3	Extension of functionality of public application				
	Extension of EURISCO frontend for <i>in situ</i> CWR passport data				
	Reorganisation of the search and display of phenotypic data				
	Continuous development of the existing public web interface				
4	Documentation and planning				

3 Coordination tasks

3.1 Coordination task 1: EURISCO network development maintenance

As in previous years, we will continue to engage with (potential) EURISCO stakeholders throughout 2024. The aim is to raise awareness of EURISCO among breeders, researchers and also at the political level by demonstrating the potential of this common European approach. In this context, public awareness materials will be continuously updated. Opportunities to present EURISCO (posters, lectures) will be exploited.

EURISCO users will receive feedback through the regular publication of the EURISCO newsletter (e-bulletin). We intend to again publish the newsletter twice a year (June/December).

This will be complemented by direct, personal communication with users, e.g. for supporting updates, providing specific database queries and special data export formats, discussing future developments, etc. An important task in this context is to provide data sets that have been missing or incomplete in EURISCO so far. We will work on identifying gaps and developing ideas on how to close them.

Various activities are currently being carried out as part of the PRO-GRACE project to identify data gaps and determine the willingness of collection holders to close them. In addition to EURISCO coordination, various other partners active in the ECPGR are also involved in this project so that the results can be utilised for EURISCO.

In addition, an update of the existing passport data is aimed for at least once a year. Those National Inventory Focal Points whose last update was longer ago will therefore be specifically approached and supported in a targeted manner.

Training workshops for providers of *ex situ* passport data and phenotypic data take place every two years; no such event is planned for 2024. Nevertheless, as in the past, smaller online training sessions can be organised at short notice if required.

Training for data providers has not yet taken place for EURISCO's extension for *in situ* CWR data, which was finalised at the end of 2023. An online event will be organised this year.

3.2 Coordination task 2: Coordination of the EURISCO development

The coordination of the future EURISCO development mainly comprises (i) the definition of new services, (ii) the advancement of current standards and (iii) the discussion and definition of new standards.

To achieve shared objectives, EURISCO coordination will investigate opportunities for collaboration with initiatives like Genesys and the GLIS under the International Treaty. In order to improve the coverage of EURISCO, we will focus on bilateral communication with the respective National Focal Points / National Coordinators. This will be a permanent activity. In this context, the EURISCO coordination will participate in activities under the ECPGR Grant Scheme.

To plan the further development of EURISCO, we will specifically collect user requirements. On the one hand, this will be done through a standardised survey. On the other hand, we will conduct one or two structured interviews with power users.

A meeting of the EURISCO Advisory Committee will be organised. In addition to the above-mentioned request for an extension to include information on the availability of material, a revision of the exchange format for phenotypic data should also be discussed. The current format follows the Single-Observation concept and is easy to use for data providers that employ a relational database. However, the vast majority of phenotypic data is collected in a project context and mainly managed in spreadsheets. It would also be desirable to capture some additional metadata. A first draft for a modified format already exists and should be discussed with the Advisory Committee.

3.3 Coordination task 3: Participation in project proposals

An indispensable task of the EURISCO coordination is to participate in the preparation of project proposals in order to acquire additional funding for the future development of EURISCO.

Support and supervision will be provided for ongoing projects. Currently, the EURISCO coordination participates in the Horizon 2020 project "AGENT" (deputy work package lead, task lead).

The EURISCO coordination is involved in the Horizon Europe project "Promoting a Plant Genetic Resource Community for Europe" (PRO-GRACE) from January 2023 (work package lead). This project aims to develop a concept and proof-of-concept actions for the establishment of a large European

Research Infrastructure for plant genetic resources. In this context, the EURISCO coordination will focus on information standards.

A considerable amount of time will be involved in the management and supervision of the activities within the context of the third-party funds raised.