

Project:	EURISCO	Date:	2025-03-21
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 2025 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: EVA infrastructure.....	2
1.2	WP2: Intranet support mechanisms for data providers.....	2
1.2.1	Maintenance of existing update mechanisms.....	2
1.2.2	Development of new update mechanism for phenotypic data	2
1.2.3	Implementation of SSD handling.....	2
1.2.4	Integration of EVA data	3
1.2.5	Update how-to documents	3
1.3	WP3: Public EURISCO application.....	3
1.3.1	Search interface for <i>ex situ</i> passport data.....	3
1.3.2	Search interface for <i>in situ</i> CWR passport data.....	4
1.3.3	Search interface for phenotypic data.....	4
1.4	WP4: Documentation and planning of the next steps	5
2	Schedule	5
3	Coordination tasks.....	6
3.1	Coordination task 1: EURISCO network development maintenance	6
3.2	Coordination task 2: Coordination of the EURISCO development	6
3.3	Coordination task 3: Participation in project proposals.....	7

1 Work plan for the EURISCO development

1.1 WP1: EVA infrastructure

The hosting of the intranet infrastructure of the European Evaluation Network (EVA) will be continued. Apart from ensuring ongoing operations, no further activities, in particular no further developments, are planned.

1.2 WP2: Intranet support mechanisms for data providers

1.2.1 Maintenance of existing update mechanisms

The established upload/update mechanisms for *ex situ* and *in situ* CWR passport data as well as for phenotypic data will continue to be maintained. 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 Development of new update mechanism for phenotypic data

A new standard for the submission of phenotypic data to EURISCO was already developed in 2024. On the one hand, this is intended to collect additional metadata for traits. On the other hand, the collection of the actual observation values has been simplified and takes into account the fact that phenotypic data is often compiled manually by the data providers. An Excel template has also already been created.

In order to be able to import data collected in the new template into EURISCO, a number of adjustments must be made in 2025:

- Extension of the database schema (both staging area and production schema)
- Development of PL/SQL packages for uploading, checking and integrating the data
- Extension of the intranet application

The aim is to enable the import of data using both the old and the new template until further notice.

1.2.3 Implementation of SSD handling

At last year's meeting of the ECPGR Documentation & Information Working Group (Doc&Info WG), 18-19 September 2024 in Tallinn, Estonia, an approach for handling single seed descent (SSD) lines derived from original genebank accessions was developed. The approach, which was subsequently also approved by the EURISCO Advisory Committee (AC), is to be further elaborated, verified and tested in 2025. It provides for:

- Extension of the database schema to create virtual entries for the SSD lines that are not part of the actual EURISCO dataset.
- Linking the virtual entries with the original accessions, but excluding them from the search for passport data.
- Linking of the phenotypic data with the virtual entries and presentation of the context in which the data was collected to the user.

In the current year, the necessary concept will be specified in more detail and arrangements will be made with Genesys. The latter is necessary in order to be able to synchronise phenotypic data with

Genesys in the future (currently only *ex situ* passport data). The necessary extensions to the database schema will then be carried out and the development of the software for importing and checking such data will begin.

1.2.4 Integration of EVA data

The data recorded in the database of the European Evaluation Network can only be accessed there by partners of the respective EVA consortia during an embargo period (three years by default). They will then be made available in EURISCO following approval by the responsible National Inventory Focal Points. By 30 June 2025, the embargo period will have expired for 185,279 phenotypic data points from 1,932 accessions. This first batch of data will then be transferred to EURISCO. The necessary procedures must be implemented and the transfer carried out.

1.2.5 Update how-to documents

Following the adaptation of the background mechanisms mentioned above, the corresponding guideline documents for data providers will be updated.

1.3 WP3: Public EURISCO application

A complete revision of the web application is planned for 2025 on the basis of the AC decision to strengthen cooperation between EURISCO and Genesys and the preliminary work already carried out last year.

A concept for the implementation of an ordering API was also discussed at the Doc&Info WG 2024 meeting. Realisation in the current year is unrealistic. However, planning for implementation and the necessary agreements with potential partners will begin.

1.3.1 Search interface for *ex situ* passport data

The *ex situ* passport data in EURISCO is regularly synchronised to Genesys, so the new development of the interface can be carried out using the Genesys REST APIs. Authentication and integration into the technology used for the web application (Oracle APEX framework) were already successfully tested in 2024. An initial mock-up was also created to discuss the design of the future interface.

Based on the basic design decision made in 2024, a more detailed mockup will be created in a first step in order to define the structure of the entire web application, including search interfaces and the presentation of results. The results of the last user survey will also be incorporated into this.

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).

The extent to which the synonym and fuzzy search implemented in EURISCO to date can be combined with the use of the Genesys APIs cannot yet be fully assessed. It may be necessary to make compromises here for the time being.

It may be necessary to revise the synchronisation of passport data in consultation with the Genesys coordination.

1.3.2 Search interface for *in situ* CWR passport data

Although the EURISCO backend has been fully extended for the management of *in situ* CWR passport data, the search capabilities of the public EURISCO web application are currently limited in this respect. 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.

In 2025, however, it is planned to implement a separate entry point for searching for *in situ* CWR data with more extensive features. As Genesys does not yet manage *in situ* CWR data, separate REST APIs must be developed for this.

The following activities are planned:

- Development of REST interfaces to access data on *in situ* CWR populations.
- 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.

1.3.3 Search interface for 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. EURISCO contains about 2.7 million records of phenotypic data points, comprising approximately 4,000 experiments and more than 9,000 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 so far only provides for a small amount of metadata (see above). 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.

It is also necessary to develop in-house EURISCO REST interfaces for this, as there is currently no synchronisation of phenotypic data with Genesys. The development will be based on the specification of the corresponding Genesys APIs so that it will be possible to switch to these APIs at a later date.

In addition, a feasibility study regarding the use of synonym searches and fuzzy searches for phenotypic data is also being considered; depending on the estimated effort involved, implementation will take place in the current year or be planned for the following year.

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 package 1 (hatched area). Should a need arise, operational plans will be made in the short term.

	Work package	2025			
		Q1	Q2	Q3	Q4
1	EVA infrastructure				
2	Intranet support mechanisms for data providers				
	Maintenance of existing update mechanisms				
	Development of new update mechanism for phenotypic data				
	Implementation of SSD handling				
	Integration of EVA data				
	Update how-to documents				
3	Public EURISCO application				
	Search interface for ex situ passport data				
	Search interface for in situ CWR passport data				
	Search interface for phenotypic data				
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 2025. 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.

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.

Until now, training workshops for providers of *ex situ* passport data and phenotypic data have been held every two years. No such event took place in 2024, so it is planned to organise a training course in the current year. It still needs to be discussed whether a joint training of providers of *ex situ* and *in situ* CWR data should take place or whether the focus should remain on *ex situ* as before.

Additionally, as in the past, smaller online training sessions can be organised at short notice if required.

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, especially in cooperation with the On-farm Working Group.

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.

At the meeting of the 'Extension of EURISCO for Crop Wild Relatives (CWR) *in situ* data and preparation of pilot countries' project that took place in Sadovo, Bulgaria, on 18–19 June 2024, it was agreed to develop principles for the assignment of DOIs to *in situ* CWR populations. A corresponding concept will be developed this year.

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) as well as in the Horizon Europe project PRO-GRACE (work package lead).

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.