

# Documentation and & Information Networking Group EPGRIS3- Activity 2-05

---

---

Implementing C&E data inclusion in EURISCO  
– preliminary concepts –

---

---

Sónia Dias  
EURISCO Coordinator

Bonn, Germany, 7 May 2009



# Based on the discussion paper Inclusion of C&E data in EURISCO

---

- Why
- When
- How

How  
Vision and concepts

# Content

---

- General principles of the proposed concept
- Data format / structure
- Data upload to EURISCO
- Data download options
- Next steps
- For this meeting

# General principles of the proposed concept

---

- **Dataset** – Is linked to the unique key of EURISCO (NICODE, INSTCODE, ACCENUMB and GENUS).
- **Experiment** – A brief description, in English, of relevant aspects of the experiment.
- **Descriptor (Trait)** – Accept the names used by the data providers and the English name of the **descriptor**.
- **Genotype** – **Genotype number** linked with the unique key of EURISCO (NICODE, INSTCODE, ACCENUMB and GENUS).
- **Method** – A brief description, in English, of the way the trait was scored: first of all the scale that was used and possibly additional info.
- **Score – GxExT**

**\*The field ...number should be mandatory.**

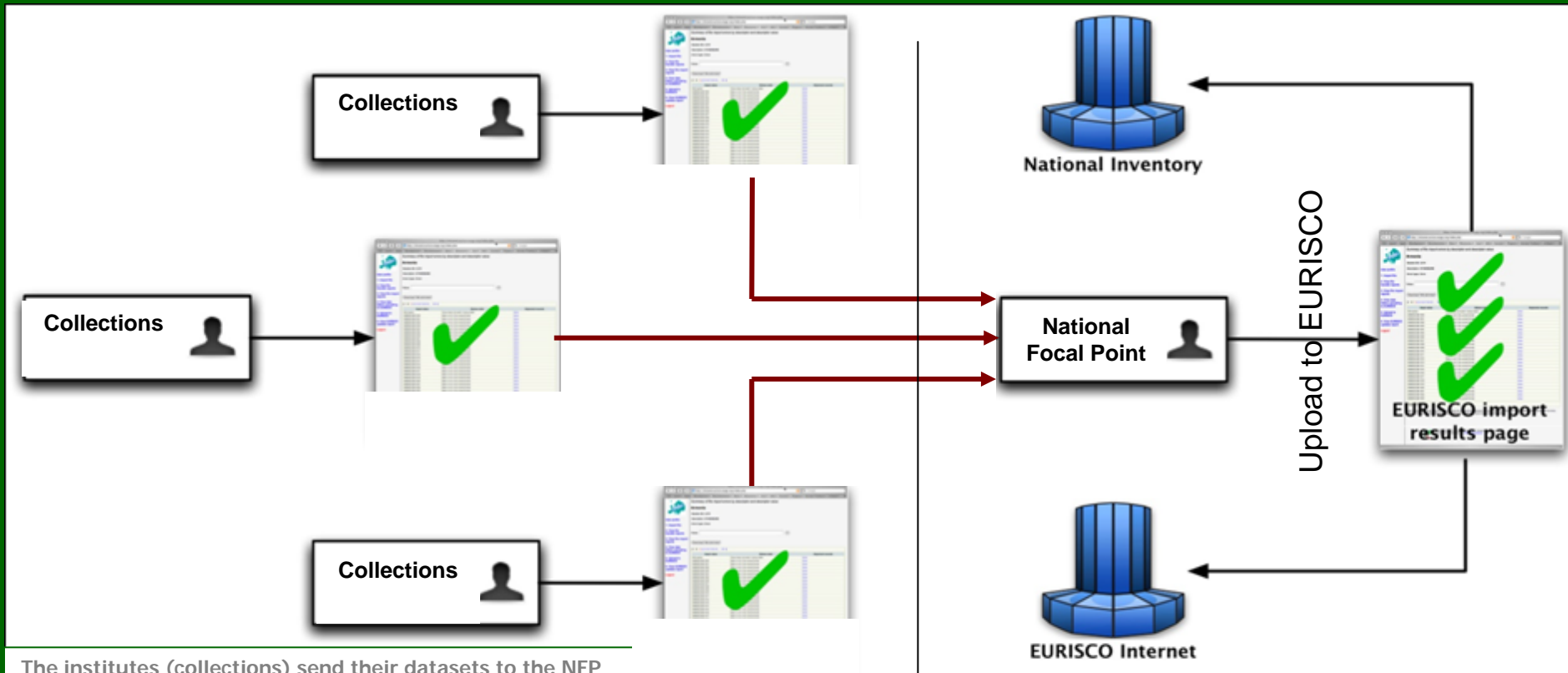
# Data format/structure

---

- The data exchange file type remains as Excel -Tab-delimited file (!?)
- MCPD +
- EURISCO descriptors (from 43-58):
  - DATASET
  - EXPERIMENT
  - TRAIT
  - GENOTYPE
  - SCORE

[EURISCO descriptors \(from 43-58\):](#)

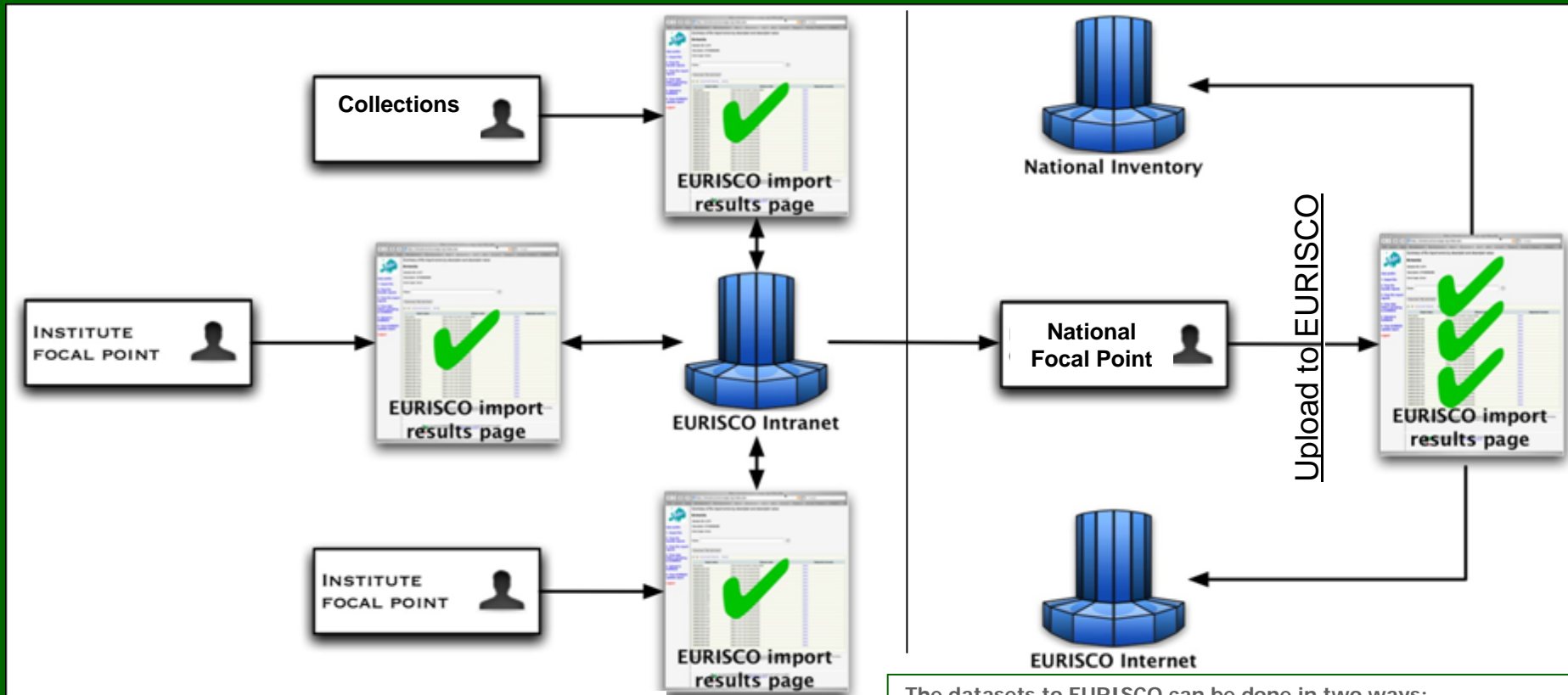
# Actual data format/upload mechanism



The institutes (collections) send their datasets to the NFP  
The NFP compiles them in the NI  
The NFP send the datasets to EURISCO  
The NFP receives error reports and send them to originator  
The institutes submit the data again with an "OK"  
The NFP submits the datasets to EURISCO

The National Focal Point (NFP) checks the overall results  
The NFP can curate the data at the national level  
The NFP decides when the datasets are "OK"  
The NFP updates the EURISCO portal  
The NFP uses the dataset for country NI

# Future data format/Upload mechanism



The institutes (collections) send their datasets to EURISCO  
 They receive error reports  
 They correct the error in their databases  
 They submit the datasets again  
 They flag their dataset as "OK"

The National Focal Point (NFP) checks the overall results  
 The NFP can curate the data at the national level  
 The NFP decides when the datasets are "OK"  
 The NFP updates the EURISCO portal  
 The NFP uses the dataset for country NI

The datasets to EURISCO can be done in two ways:

- 1- The actual process – NFP updates
- 2 – The NFP can give admin privileges to individual institutes to upload their data into the intranet.

The intranet will be updated to enable 1 and 2 to:

- delete individual accessions
- add individual accessions
- correct individual accessions

These can be either done by the NFP or by the curators if given admin permissions

For the upload ok the actual process applies

# Upload mechanism

---

As in the current mechanism for the passport data system:

- Upload mechanism should check the integrity of the files:
  - for format and codes.
- In case of errors, a report should be made available to the NFP and each of data providers (admin).
- Once the data file is through the automatic routine checks, and is ok, then can be uploaded by the NFP.
- Any number is linked through the mandatory fields (NICODE, INSTCODE, ACCENUMBER, GENUS)
- *E.g...*
  - the EXPERIMENT NUMBER is a unique number for each NI and INST,
  - When checked for duplication of the experiment number: the system should report it as a duplicate number and not uploading it

# Upload mechanism

---

- Decision on the correction of the number should be done by the NFP/(admin)-data originators
- ??? the other numbers - DESCRIPTOR NUMBER, GENOTYPE NUMBER - should **not** be replaced by the existing number, **but reported**, in case the corresponding data are already in the repository, **no** new internal number should be attributed???
- All (new) data should be loaded in the corresponding tables.
- A report about the **need to** replace and **other** actions should be **made available in the intranet for the NFP/(admin)**.

- **Import file data format**
- **Preliminary operations**

- **Data checks**  
**Data format, Descriptors, Required descriptors, Codes, Dates**

# How it could look like in EURISCO



A window to Europe's plant diversity

[About](#)

[Search](#)

[Meetings](#)

[Releases](#)

[Links](#)

[Home](#)

## About EURISCO



The EURISCO web catalogue receives data from the National inventories, and provides access to all *ex situ* PGR information in Europe.

## Search the database



### C&E Search

Search for information on experiments, descriptors, genotypes, methods and score

## Search the database



### Quick Search

Make a search using key criteria

### Advanced Search

Use various criteria to search db

[Contacts](#)

[Intranet](#)

[FAQs](#)

[Training](#)

[Tools](#)

[Documents](#)

# Could be searched like this

The image shows a screenshot of the eurisco website interface. The top navigation bar includes 'About', 'Search', and 'Meetings'. Below the navigation bar, there are three main search panels. The left panel shows 'Quick Search' with options for Taxonomy, Accession, Status, and Site, and an 'Advanced Search' link. The middle panel shows 'Advanced Search' with categories: 'Institute holdings' (Holding institute, National inventory name), 'Accession identifiers' (Accession number, Accession name(s), Common crop name, Biological status, Country of origin), 'Taxonomy information' (Genus, Species, Subtaxa), 'Acquisition/Storage information' (Acquisition date, Collecting/acquisition source, Storage type, MLS status, AEGIS status), and 'Collecting mission information' (Collecting number, Collecting institute code). The right panel shows 'Quick Search' with the same options as the left, and a 'C&E Search' section with 'Experiment', 'Descriptor', 'Genotype', and 'Score'. Below the search panels, there are links for 'Contacts', 'Intranet', 'FAQs', 'Training', 'Tools', and 'Documents'. On the right side, there are three highlighted search result boxes: 'Experiment' with options for Experiment number, Experiment description, Experiment year, and Experiment report; 'Descriptor' with options for Descriptor number, Descriptor name, Descriptor remark, and Descriptor method; and 'Genotype' with an option for Genotype number.

**eurisco**  
Finding seeds for the future

**eurisco**  
Finding seeds for the future

*A window to Europe's plant diversity*

Home → Advanced Search

Home → Advanced Search

Home → Advanced Search

**Quick Search**

- Taxonomy
- Accession
- Status
- Site

Advanced Search

**Advanced Search**

**Institute holdings**

- Holding institute
- National inventory name

**Accession identifiers**

- Accession number
- Accession name(s)
- Common crop name
- Biological status
- Country of origin

**Taxonomy information**

- Genus
- Species
- Subtaxa

**Acquisition/Storage information**

- Acquisition date
- Collecting/acquisition source
- Storage type
- MLS status
- AEGIS status

**Collecting mission information**

- Collecting number
- Collecting institute code

**Quick Search**

- Taxonomy
- Accession
- Status
- Site

Advanced Search

**C&E Search**

- Experiment
- Descriptor
- Genotype
- Score

**Experiment**

- Experiment number
- Experiment description
- Experiment year
- Experiment report

**Descriptor**

- Descriptor number
- Descriptor name
- Descriptor remark
- Descriptor method

**Genotype**

- Genotype number

Contacts

Intranet

FAQs

Training

Tools

Documents

Tools

Contacts

Intranet

FAQs

Training

Tools

Documents

Tools

Quick Search

- [Taxonomy](#)
- [Accession](#)
- [Status](#)
- [Site](#)

Advanced Search

**C&E Search**

**View by:**

- [Experiment](#)
- [Descriptor](#)
- [Genotype](#)
- [Score](#)

- [Contacts](#)
- [Intranet](#)
- [FAQs](#)
- [Training](#)
- [Tools](#)
- [Documents](#)

**Your selection (97169 Accessions)**

**Genus**  
hordeum (Contains)

**Number of lines returned: 97169**

**1** Results per Page  **20 >>>**

**View summary by Score**

Accession number	Genus	Experiment number	Descrip number	Score	Details
00014	anni				
00015	elo				
00016	esme				
00017	jogeva 453				
00018	jogeva 1104				
00019	leelo				
00020	liisa				
<b>00021</b>	miina				
00022	teele				
00023	toomas				
00142	gandvits				
00143	vobor				
00144	bios-1				
00145	elf				

Download all results in EXCEL, CSV, XML

Search within results

Hyperlinked to the passport data and vice-versa

# Data download options

---

Like the current download options:

- Download the selection you make
- Download the full data set (58 standards)
- Download the selection in Excel, CSV and XML format
- Download with one click

# Next steps

---

- To agree that C&E data flow through EURISCO
- To propose this to the network
- To agree on the standards to use
- To add these C&E standards to the EURISCO standards
- Adoption endorsed as part of the MoU
- Implement the necessary changes in the intranet
  - Passport data
  - C&E data
- Implement the necessary changes on the front end
- Update upload mechanism: update routine checks
- Test with few data providers
- Case show it
- Feedback
- Deployment to the network (training, helpdesk, etc)
- EURISCO linked with ALIS

# For this meeting

---

- Is this the way you see it?
- What do you have in mind, tell us?
- Feedback?
- How is the network going to be involved?
- Inclusion on the MoU?
- Who does what?

---

Thank you for your attention