



**Draft July 24, 2001**

**Technical paper**

**The EURISCO Institution Code List**

## **Introduction**

EURISCO contains a number of data elements that refer to institutions. There will be a field for the institution maintaining the accession, one for the institution that donated the accession to the one holding it, and there will probably be a field for the institution who bred or collected the accession. For reasons of integrity and accessibility these fields will have to contain codes referring to values in a separate decoding table, the EURISCO Institution Code List.

## **Institution Codes**

Compiling a list of institutions is a difficult job. There are a number of issues that make the task a complex one:

- *The unit, i.e., the definition of an institution.* Should it be possible to include a private individual, collaboration between institutes, projects such as collecting expeditions, sections of institutes, etc.?
- *The dynamics and relationships.* How to handle merging, splitting or moving institutions, institutions changing their names, acronym or contact details? In what cases should a new entry be made, and in what cases can the record be changed? Is it possible to include pointers to other entries, for example, in case of a merger, or in case of both the department and the main institute appearing?
- *The level of detail.* What balance between detail of information and accuracy should be chosen? The more detail the more management needed to keep an acceptable level of accuracy.
- *The curation.* Who can change what, how can anyone needing a code get it quickly, how can quality be assured?

The only common list of institution codes used in the PGR community is the FAO WIEWS Institute Table. The quality of this list is not high: there is incompleteness, duplication, and lack of accuracy. But it exists, some PGR institutions and Central Crop Databases use it, and it was recommended in the IPGRI/FAO Multi Crop Passport Descriptor List. This means that starting a new list would create a new standard and possibly confusion. That should obviously be avoided. Promising attempts to improve the list are underway, and EPGRIS should support these attempts. But also the possibility of the FAO list proving, despite attempts for improvements, not to be suitable for EURISCO should be considered. However, for the time being the FAO list has to be considered the standard; in this document the FAO codes will be referred to as the 'standard codes'.

## **The EURISCO Code List**

Converting locally used institution codes to standard codes is not always possible, for a number of reasons:

- The capacity needed for the conversion is not (yet) available. If the time needed to match all locally used codes with the standard codes is considered too much, and updating the National Inventory, and thus EURISCO, is considered more urgent, the locally used codes should somehow be processed by EURISCO.
- The locally used codes do not have a standard code equivalent, and the institution responsible for the conversion doesn't have the authority to add the institution to the standard list.
- The locally used codes provide a level of detail higher than the standard code. If it is considered important to refer to, for example, the person who donated a sample, the standard codes do not provide enough detail.

- The locally used codes do not have sufficient information to match with the standard codes. If, for example, the locally used code refers to 'RD, Brazil' or 'NGO from Germany', matching will not be possible. However, ignoring this information would be undesirable.

It can be concluded that there is a need for a mechanism to include the information connected to locally (at the National Inventory level) used codes in EURISCO. The EURISCO Institution Code list will therefore consist of both the standard codes and locally used codes. Since these locally used codes could be removed from the next version for the National Inventory, since they were matched to FAO codes or since the corresponding accessions were deleted from the inventory, locally used codes should not be used by others. The common part of the EURISCO Institution Code list is the standard list.

### Possible Mechanism

National Inventories are encouraged to use the standard codes as much as possible. However, it is also possible to send locally used codes to EURISCO. A file decoding these locally used codes should be sent with the file containing passport data.

The file decoding the locally used codes should have following structure:

- Local code. Key field, mandatory and unique in the list, starting with an asterisk, followed by the 3 letter ISO 3166 country code of the institution (if unknown 000 can be used), and an additional code (any combination of ciphers and uppercase letters).
- Corresponding standard code. Optional: to be used if the contact is part of, or closely related to the corresponding institution. Is used for searches.
- Interpretation of the code. Mandatory, any information about the institution coded by the local code.

Examples:

LOCALCODE	CORRCODE	INTERPRETATION
*NLDBAKKER *000HRF	*NLD037	Rob Bakker, consultant in 1987 to CGN, Wageningen, The Netherlands HRF
*FRA2304		Dr. Terlouw, France
*AUSFBAUER		Franz Bauer, Dorflein, Austria

When a dataset from a National Inventory is received by EURISCO, the validation program will check if all institution codes in the passport file are either in the standard code list or in the appended local code list. If not, this results in an error, which will be reported back to the National Inventory.

When the validation program accepts the dataset, all local codes previously donated by this National Inventory are deleted from the EURISCO Institution Code list. The new local codes are appended with a tag (for example an \* plus a two letter code for the country) both in the 'local code list' as in the passport data file. This makes the codes unique within EURISCO and allows EURISCO to recognize the source of the local codes. For example, \*AUSFBAUER supplied by the Dutch National Inventory will become \*AUSFBAUER\*NL.

Finally, the 'local code list' is loaded in the EURISCO Institution Code list.

In the detail pages of EURISCO, the 'local codes' should be decoded with the interpretation provided by the National Inventory. But in searches, EURISCO should accept standard codes, and display results with the standard code either in the code field (for standard codes) or in the 'corresponding standard code' field, for local codes. Also free text searches in the 'interpretation field' should be made possible.