DRAFT MAPPING OF EURISCO DESCRIPTORS TO ABCD 2.06

This Document is based on the orginal table published under
www.ecpgr.cgiar.org/epgris/Tech_papers/EURISCO_Descriptors.doc

Only the table was annotated with the path and name of the respective ABCD schema element

W. Berendsohn, Jan. 31, 2006

EURISCO Descriptors
for uploading information from National Inventories to EURISCO

Introduction

This descriptor list is used for uploading information from the National Inventories to EURISCO, and thus purely a format of data exchange.
The list is an extension of the FAO/IPGRI multi-crop passport descriptors (MCPD) which were published December 2001, developed jointly by IPGRI and FAO, with input from many documentation specialists worldwide, to provide international standards to facilitate germplasm passport information exchange. All MCPD are included, without change and with the same format rules, in the current list. Six descriptors were added for the specific purposes of EURISCO: the first descriptor, identifying the National Inventory and the final five allowing the incorporation of information relevant to EURISCO, which otherwise would not fit in the MCPD.

General format rules

Following format rules, as copied from the MCPD-list, apply to all fields:

- If a field allows multiple values, these values should be separated by a semicolon (;) without space(s). (i.e. Accession name: “Rheinische Vorgebirgstrauben;Emma;Avlon”)
- A field for which no value is available should be left empty (i.e. Elevation). If data are exchanged in ASCII format for a field with a missing numeric value, it should be left empty. If data are exchanged in a database format, missing numeric values should be represented by generic NULL values.
- Dates are recorded as YYYYMMDD. If the month and/or day are missing this should be indicated with hyphens. Leading zeros are required (i.e. 197506--, or 1975----).
- Latitude and longitude are recorded in an alphanumeric format. If the minutes or seconds are missing, this should be indicated with hyphens. Leading zeros are required.
- For coding countries three-letter ISO 3166-1 codes are used (including the codes that are no longer in use in the ISO 3166-1, such as DDR).[1,2]

[1] The ISO 3166-1 Code List can be found at: http://www.un.org/Depts/unsd/methods/m49alpha.htm. Country or area numerical codes added or changed are not available on line, but can be obtained from IPGRI [t.metz@cgiar.org].
[2] cf. TDWG Geographical Standard. HK

- For coding institutes the FAO Institute Codes should be used as maintained by the FAO. The codes consist of the 3-letter ISO 3166 country code of the country where the institute is located plus a three-digit number.[3]
- The preferred language for free text fields is English (i.e. Location of collecting site and Remarks).

Descriptors

The descriptors are numbered according to the FAO/IPGRI multi-crop passport descriptors (MCPD); the first descriptor (numbered 0) and the last five (numbered 29-33) are additional, and specific to this EURISCO Descriptor List. Only the four fields identifying the accession are mandatory, all other fields are highly recommended. The mandatory fields are NICODE (0), INSTCODE (1), ACCENUMB (2) and GENUS (5). The combination of these fields has to be unique.

EURISCO DESCRIPTORS	ABCD path and comment
	Unit/* = Datasets/Dataset/Units/Unit/* PGRUnit/* = Unit/PlantGeneticResourcesUnit/* TaxonIdentified/* = Unit/Identifications/Identification/Result/TaxonIdentified/* ScientificName/* = TaxonIdentified/ScientificName/* NameBotanical/* = ScientificName/NameAtomised/Botanical/*
0. National Inventory code **(NICODE)** Code identifying the National Inventory; the code of the country preparing the National Inventory. Exceptions are possible, if agreed with EURISCO such as NGB. Example: `NLD`	PGRUnit/NationalInventoryCode
1. Institute code **(INSTCODE)** FAO Institute Code of the institute where the accession is maintained. Example: `NLD037`	Unit/SourceInstitutionID Mandatory (first part of globally unique identifier for an ABCD unit record)

[3] These codes are available from http://apps3.fao.org/wiews/ for registered WIEWS users. From the Main Menu select: 'PGR' and 'Download'. If new Institute Codes are required, they can be generated online by national WIEWS correspondents, or by the FAO WIEWS administrator [Stefano.Diulgheroff@fao.org].

EURISCO DESCRIPTORS	ABCD path and comment
2. Accession number (ACCENUMB) This number serves as a unique identifier for accessions within a genebank collection, and is assigned when a sample is entered into the genebank collection. Example: `CGN00254`	Two components: a) Unit/SourceID Mandatory (second part of globally unique identifier for an ABCD unit record) Equivalent to collection within an institution. b) Unit/UnitID Mandatory (third part of globally unique identifier for an ABCD unit record) Equivalent to an accession number which is unique inside the collection
3. Collecting number (COLLNUMB) Original number assigned by the collector(s) of the sample, normally composed of the name or initials of the collector(s) followed by a number. This number is essential for identifying duplicates held in different collections. Example: `FA90-110`	Two possibilities: 1. if a code is used: Unit/CollectorsFieldNumber 2. if a person or team name and a code are used: Unit/Gathering//GatheringAgent/AgentText + Unit/CollectorsFieldNumber
4. Collecting institute code (COLLCODE) Code of the Institute collecting the sample. If the holding institute has collected the material, the collecting institute code (COLLCODE) should be the same as the holding institute code (INSTCODE). Example: `NLD037`	Unit/Gathering/Agent/ Organisation/Name/Abbreviation
5. Genus (GENUS) Genus name for taxon, in latin. Initial uppercase letter required. Example: `Allium`	NameBotanical/GenusOrMonomial
6. Species (SPECIES) Specific epithet portion of the scientific name, in latin, in lowercase letters. Following abbreviation is allowed: 'sp.' Example: `paniculatum`	NameBotanical/FirstEpithet

EURISCO DESCRIPTORS	ABCD path and comment
7. Species authority (SPAUTHOR) The authority for the species name. Example: `L.`	IF /SecondEpithet is empty: NameBotanical/ /AuthorTeamParenthesis + NameBotanical/AuthorTeam
8. Subtaxa (SUBTAXA) Subtaxa can be used to store any additional taxonomic identifier, in latin. Following abbreviations are allowed: 'subsp.' (for subspecies); 'convar.' (for convariety); 'var.' (for variety); 'f.' (for form). Example: `subsp. fuscum`	NameBotanical/Rank + NameBotanical/SecondEpithet
9. Subtaxa authority (SUBTAUTHOR) The subtaxa authority at the most detailed taxonomic level. Example: `(Waldst. et Kit.) Arc.`	NameBotanical/ /AuthorTeamParenthesis + NameBotanical/AuthorTeam
10. Common crop name (CROPNAME) Name of the crop in colloquial language, preferably English. Example: `malting barley` Example: `cauliflower`	TaxonIdentified/InformalNameString [Language is expressed by the standard XML language attribute]
11. Accession name (ACCENAME) Either a registered or other formal designation given to the accession. First letter uppercase. Multiple names separated with semicolon without space. Example: `Rheinische Vorgebirgstrauben;Emma;Avlon`	NameBotanical/ CultivarGroupName + ";" + NameBotanicalCultivarName + ";" + NameBotanical/TradeDesignationName(s)
12. Acquisition date (ACQDATE) Date on which the accession entered the collection as YYYYMMDD. Missing data (MM or DD) should be indicated with hyphens. Leading zeros are required. Example: `1968----` Example: `20020620`	Unit/SpecimenUnit/Acquisition/AcquisitionDate [DateTimeType. Since the hyphens are not required, the ISODateTimeBegin should cover your usage.]
13. Country of origin (ORIGCTY) Code of the country in which the sample was originally collected. Example: `NLD`	Unit/Gathering /GatheringSite/Country/ISO3166Code
14. Location of collecting site (COLLSITE) Location information below the country level that describes where the accession was collected. This might include the distance in kilometres and direction from the nearest town, village or map grid reference point Example: `7 km south of Curitiba in the state of Parana`	Unit/Gathering/LocalityText

EURISCO DESCRIPTORS	ABCD path and comment
15. Latitude of collecting site (LATITUDE) Degree (2 digits) minutes (2 digits), and seconds (2 digits) followed by N (North) or S (South). Every missing digit (minutes or seconds) should be indicated with a hyphen. Leading zeros are required Example: `10----S` Example: `011530N` Example: `4531--S`	Unit/Gathering/ SiteCoordinateSet(s)/ CoordinatesLatLon/ LatitudeDecimal [ATTENTION: No provision for the format in ABCD, but transformation is easy.]
16. Longitude of collecting site (LONGITUDE) Degree (3 digits), minutes (2 digits), and seconds (2 digits) followed by E (East) or W (West). Every missing digit (minutes or seconds) should be indicated with a hyphen. Leading zeros are required. Example: `0762510W` Example: `076----W`	Unit/Gathering/SiteCoordinateSet(s)/ CoordinatesLatLon/ LongitudeDecimal [ATTENTION: No provision for the format in ABCD, but transformation is easy.]
17. Elevation of collecting site (ELEVATION) Elevation of collecting site expressed in meters above sea level. Negative values are allowed. Example: `763`	Unit/Gathering/Altitude/MeasurementAtomised/ MeasurementLowerValue [Unit/Gathering/Altitude/MeasurementAtomised/ MeasurementScale should be set to "m"]
18. Collecting date of sample (COLLDATE) Collecting date of the sample as YYYYMMDD. Missing data (MM or DD) should be indicated with hyphens. Leading zeros are required. Example: `1968----` Example: `20020620`	Unit/Gathering/DateTime/ISODateTimeBegin [DateTimeType. Since the hyphens are not required, the ISODateTimeBegin should cover your usage.]
19. Breeding institute code (BREDCODE) FAO Institute Code of the institute that has bred the material.	PGRUnit/BreedingInstitutionCode

EURISCO DESCRIPTORS	ABCD path and comment
20. Biological status of accession (SAMPSTAT) The coding scheme proposed can be used at 3 different levels of detail: either by using the general codes (in boldface) such as 100, 200, 300, 400 or by using the more specific codes such as 110, 120 etc. 100) Wild 110) Natural 120) Semi-natural/wild 200) Weedy 300) Traditional cultivar/landrace 400) Breeding/research material 410) Breeder's line 411) Synthetic population 412) Hybrid 413) Founder stock/base population 414) Inbred line (parent of hybrid cultivar) 415) Segregating population 420) Mutant/genetic stock 500) Advanced/improved cultivar 999) Other (Elaborate in REMARKS field)	PGRUnit/BiologicalStatus
21. Ancestral data (ANCEST) Information about either pedigree or other description of ancestral information (i.e. parent variety in case of mutant or selection). Example: `Hanna/7*Atlas//Turk/8*Atlas` Example: `mutation found in Hanna` Example: `selection from Irene` Example: `cross involving amongst others Hanna and Irene`	PGRUnit/AncestralData

EURISCO DESCRIPTORS	ABCD path and comment
22. Collecting/acquisition source (COLLSRC) The coding scheme proposed can be used at 2 different levels of detail: either by using the general codes (in boldface) such as 10, 20, 30, 40 or by using the more specific codes such as 11, 12 etc. 10) Wild habitat 11) Forest/woodland 12) Shrubland 13) Grassland 14) Desert/tundra 15) Aquatic habitat 20) Farm or cultivated habitat 21) Field 22) Orchard 23) Backyard, kitchen or home garden (urban, peri-urban or rural) 24) Fallow land 25) Pasture 26) Farm store 27) Threshing floor 28) Park 30) Market or shop 40) Institute, Experimental station, Research organization, Genebank 50) Seed company 60) Weedy, disturbed or ruderal habitat 61) Roadside 62) Field margin 99) Other (Elaborate in REMARKS field)	PGRUnit/CollectingAcquisitionSource (analogue to ProvenanceCategory in Bot. Gardens, but different controlled vocabulary)
23. Donor institute code (DONORCODE) FAO Institute Code for the donor institute.	Unit/SpecimenUnit/History/PreviousUnit(s)/PreviousSourceInstitutionID
24. Donor accession number (DONORNUMB) Number assigned to an accession by the donor. Example: `NGB1912`	Unit/SpecimenUnit/History/PreviousUnit(s)/PreviousUnitID
25. Other identification (numbers) associated with the accession (OTHERNUMB) Any other identification (numbers) known to exist in other collections for this accession. Use the following system: INSTCODE:ACCENUMB;INSTCODE:ACCENUMB;… INSTCODE and ACCENUMB follow the standard described above and are separated by a colon. Pairs of INSTCODE and ACCENUMB are separated by a semicolon without space. When the institute is not known, the number should be preceded by a colon. Example: `NLD037:CGN00254` Example: `SWE002:NGB1912;:Bra2343`	PGRUnit/OtherIdentification

EURISCO DESCRIPTORS	ABCD path and comment
26. Location of safety duplicates (DUPLSITE) FAO Institute Code of the institute where a safety duplicate of the accession is maintained. The codes consist of the 3-letter ISO 3166 country code of the country where the institute is located plus a number.	PGRUnit/LocationSafetyDuplicates
27. Type of germplasm storage (STORAGE) If germplasm is maintained under different types of storage, multiple choices are allowed (separated by a semicolon). (Refer to FAO/IPGRI Genebank Standards 1994 for details on storage type.) 10) Seed collection 11) Short term 12) Medium term 13) Long term 20) Field collection 30) In vitro collection (Slow growth) 40) Cryopreserved collection 99) Other (elaborate in REMARKS field)	PGRUnit/TypeGermplasmStorage
28. Remarks (REMARKS) The remarks field is used to add notes or to elaborate on descriptors with value 99 or 999 (=Other). Prefix remarks with the field name they refer to and a colon. Separate remarks referring to different fields are separated by semicolons without space. Example: `COLLSRC:roadside`	Unit/Notes
29. Decoded collecting institute (COLLDESCR) Brief name and location of the collecting institute. Only to be used if COLLCODE can not be used since the FAO Institution Code for this institute is not (yet) available. Example: `Tuinartikelen Jan van Zomeren, Arnhem, The Netherlands`	Unit/Gathering/Agent/Organisation/Name/Text
30. Decoded breeding institute (BREDDESCR) Brief name and location of the breeding institute. Only to be used if BREDCODE can not be used since the FAO Institution Code for this institute is not (yet) available. Example: `CFFR from Chile`	PGRUnit/DecodedBreedingInstitute
31. Decoded donor institute (DONORDESCR) Brief name and location of the donor institute. Only to be used if DONORCODE can not be used since the FAO Institution Code for this institute is not (yet) available. Example: `Nelly Goudwaard, Groningen, The Netherlands`	PGRUnit/DecodedDonorInstitute
32. Decoded safety duplication location (DUPLDESCR) Brief name and location of the institute maintaining the safety duplicate. Only to be used if DUPLSITE can not be used since the FAO Institution Code for this institute is not (yet) available. Example: `Pakhoed Freezers inc., Paramaribo, Surinam`	PGRUnit/DecodedSafetyDuplicationLocation

EURISCO DESCRIPTORS	ABCD path and comment
33. Accession URL (ACCEURL) URL linking to additional data about the accession either in the holding genebank or from another source. Example: `www.cgn.wageningen-ur.nl/pgr/collections/passdeta.asp?accenumb=CGN04848`	Unit/RecordURI

APPENDIX Differences between EURISCO, MCPDv2 and MCPDv1 descriptors

General changes from MCPDv1 to MCPDv2	
FAO Institution codes	Fields containing FAO institution codes should now use the codes with the format CCCNNN in which CCC is the country and NNN is the sequential number (in MCPDv1 acronyms and preliminary codes were acceptable).
Multiple values	Values in fields, which can contain multiple values, are separated by a semicolon without a space (;) (in MCPDv1 a semicolon with a space was used).

Changes per EURISCO descriptor		
EURISCO Descriptor		**Remark concerning change**
0	NICODE	Field specific for EURISCO, identifying the National Inventory. Use the country codes as specified by the ISO 3166-1 standard. Exceptions are possible, if agreed with EURISCO (such as NGB).
1	INSTCODE	See remark on FAO Institution codes.
2	ACCENUMB	Same as in MCPDv1.
3	COLLNUMB	Same as in MCPDv1.
4	COLLCODE	New descriptor in MCPDv2.
5	GENUS	Same as in MCPDv1.
6	SPECIES	Original MCPDv1 field split into two separate fields: SPECIES and SPAUTHOR.
7	SPAUTHOR	New descriptor in MCPDv2.
8	SUBTAXA	Original MCPDv1 field split into two separate fields: SUBTAXA and SUBTAUTHOR.
9	SUBTAUTHOR	New descriptor in MCPDv2.
10	CROPNAME	New descriptor in MCPDv2.

11	ACCENAME	Field name has changed (was ACCNAME). See remark on multiple values.
12	ACQDATE	New descriptor in MCPDv2.
13	ORIGCTY	Same as in MCPDv1.
14	COLLSITE	Same as in MCPDv1.
15	LATITUDE	Format changed. Now seconds are also required (or hyphens if missing), so all values will be exactly two positions longer.
16	LONGITUDE	Format changed. Now seconds are also required (or hyphens if missing), so all values will be exactly two positions longer.
17	ELEVATION	Same as in MCPDv1.
18	COLLDATE	Same as in MCPDv1.
19	BREDCODE	New descriptor in MCPDv2. See remark on FAO Institution codes.
20	SAMPSTAT	Coding system changed, all codes are now three digits long. Following list first gives the old MCPDv1 code, followed by the new MCPDv2 two digit code: 1►100, 2►200, 3►300, 4►410, 5►500, 99►999, 0►*null.*
21	ANCEST	New descriptor in MCPDv2.
22	COLLSRC	Coding system changed, all codes are now two digits long. Following list first gives the old code, followed by the new two digit code: 1►10, 1.1►11, 1.2►12, 1.3►13, 1.4►14, 2►20, 2.1►21, 2.2►22, 2.3►23, 2.4►24, 2.5►26, 3►30, 3.1►30, 3.2►30, 3.3►30, 3.4►30, 4►40, 99►99, 0►*null.*
23	DONORCODE	See remark on FAO Institution codes.
24	DONORNUMB	Same as in MCPDv1.
25	OTHERNUMB	Format changed. Now following format is used: INSTCODE:ACCENUMB. When the institute or its code is not known, the number should be preceded by a colon only. See also remark on FAO Institution codes and on multiple values.
26	DUPLSITE	See remark on FAO Institution codes.
27	STORAGE	Coding system changed, all codes are now two digits long. Following list first gives the old code, followed by the new two digit code: 1►11, 2►12, 3►13, 4►30, 5►20, 6►40, 99►99. See also remark on multiple values.
28	REMARKS	See remark on multiple values.
29	COLLDESCR	Field specific for EURISCO, free text field for solving institution code problems. Only to be used if the corresponding field COLLCODE can not be used since the FAO Institution Code for this institute is not (yet) available. The fields can contain a brief name and location of the institute, but can also contain for example the locally used acronym if this code doesn't have any corresponding additional information.
30	BREDDESCR	Field specific for EURISCO, free text, only to be used if the corresponding field BREDCODE can not be used. (see remarks COLLDESCR)

31	DONORDESCR	Field specific for EURISCO, free text, only to be used if the corresponding field DONORCODE can not be used. (see remarks COLLDESCR)
32	DUPLDESCR	Field specific for EURISCO, free text, only to be used if the corresponding field DUPLSITE can not be used. (see remarks COLLDESCR)
33	ACCEURL	Field specific for EURISCO, providing a link to additional information about the accession maintained elsewhere. Should contain a valid URL pointing to details about the accession either in the holding genebank or from another source.