Floricode.com
WORLD STANDARDS IN FLORICULTURE

General implementation Guidelines for Coding 1.4

Processing of code lists via FTP

File name	: HZ21024V1.0 General Implementation Guidelines for Coding 1.4.docx
File number	: HZ/21024
Date last update	: 28 April 2021
Document version/release	: 1.0/1.4
Document status	: Final version

Disclaimer

Floricode pays utmost attention to the information which it distributes. However, it cannot vouch for correctness, completeness and timeliness of the information. This also applies to the content or reliability of (references to) other web sites and hyperlinks. No right whatsoever can be derived from the information and Floricode is not liable for any damage or loss, direct or indirect, and of whatever kind, resulting from or in any way connected with use of the information or (temporary) inaccessibility of the web sites.

Floricode reserves all rights pertaining to the information on its web sites. No publication or modification of the information is allowed without prior written permission from Floricode.

Contents

Modifications

Version	Date	Author	Modifications
1.0	26-06-2007	G. Heemskerk	Initial version
1.1	15-08-2007	G. Heemskerk	After CMG review
1.2	05-11-2007	G. Heemskerk	Description of MD-5 control file has been added (it was missing).
1.3	05-02-2014	M. Goedhart	Additions to Section 1.5 in connection with the use of semicolons in fields
1.4	16-08-2016	M. Goedhart	Text adjustments
1.5	26-06-2017	M. Goedhart	Extension file name code list: - File names new data sets after 1 January 2017 - GPC Because of these changes the version/release is increased from 1.2 in 1.3
1.6	4-9-2019	H. Zwinkels	Text correction in paragraph 3.3
1.7	12-03-2020	M. Goedhart	Section 3.5 (page 7): 1. When a semicolon is used in a field, the entire content of the field must be enclosed in double quotes. 2. When using CR-LF in a field, the entire contents of the field should be enclosed in quotation marks. 3. When a quotation mark is used in a field and the entire contents of the field are enclosed in quotation marks, each quotation mark in the text is given an extra quotation mark.
1.0/1.4	28-04-2021	H. Zwinkels	Text corrections and new release

	Date	Officer	Initials
Agreement (author)	28-04-2021	H. Zwinkels	HZ
Checked		L. Zandvliet (Floricode) A. Zwanenburg (Floricode) E. Bakker (Floricode)	
Accepted on behalf of Working Group on Standards		chairman	

1. Introduction

1.1 Aim of this document

This document describes coding's that are distributed via FTP according to the 'Linnaeus format'. This 'Linnaeus format' was formally described in the 'VBN Linnaeus Technical Blueprint 2.10'. At that moment this model referred only to the distribution of product codes. Because more and more data files with master data from Floricode and from other code administrators are distributed by Floricode based on this LINNAEUS format it was decided to present these general implementation guidelines in this separate document. Now all distribution files with master data can be processes by the software in a standardized manner as described in this document.

1.2 Intended for...

This document is intended for data purchasers of the code lists distributed by Floricode via FTP. For this Floricode has subscriptions available in the form of code packages. Data purchasers of code lists via API are referred to the regarding API documentation.

1.3 Principles

This document is based on:

- VBN Linnaeus Technical Blueprint 2.10, Section 3.1 'Common aspects of code lists'[1].

The following principles apply:

- RFC 2007-1

1.4 Abbreviations and terms

CR-LF	Carriage Return and Line Feed
CSV	Coma Separated Values
FTP	File Transfer Protocol
GPC	Global Product Classification
ISO	International Standards Organisation
RFC	Request for Change
VBN	Association of the Dutch Flower Auctions
ZIP	One of the types for compressed files.

[1] Now an actual version of this document applies where this specific chapter is not present anymore because the content is included in this document.

2. Common aspects for the implementation of code lists

The common aspects of code lists that are distributed via FTP in the 'Linnaeus format' are described in the paragraphs here under.

2.1 Code list identification

Code lists are provided in a compressed format in a distribution set. A compressed distribution set (ZIP) consists of one or more code lists (a data set). The code list is named according to agreements made (see section 3.6). Each record in the code list begins with a file code that identifies the code list. This file code is unique within the dataset.

2.2 Update fields

In accordance with the VBN Blueprint Linnaeus, the following fields are included in each record in a code list:

- Entry Date (entry_date)
  - o date on which a new item became or will become generally effective within the sector; this may be a date in the future.
- Expiry date (expiry_date)
  - o date on which an item will no longer be applicable within the sector for general use or was withdrawn; this may be a future date or empty.
- Change date / time (change_date_time):
  - o date and time (change_date_time) when an item was last changed (enhanced, modified, deleted).

2.3 Detection of updates

To confirm new, expired or changed items, the application must retain the publication date of the specific Floricode distribution set and identify items which have been changed since that date as follows:

- New:
  - O change_date_time is later than the last publication date/time and entry_date is later than or equal to the change_date.
- Expired:
  - o change_date_time is later than the last publication date/time and expire_date is later than or equal to the change_date.
- Changed:
  - o change_date_time is later than the last publication date/time and entry_date is earlier than the change_date.

For the proper functioning of this procedure, the following will be ensured:

- expired items will remain in the list, with the exception of items that have expired because they were entered wrongly or in error; these items do not necessarily need to be retained permanently in the code lists and may disappear from the lists after some time (years).
- new items that are changed shortly after their inclusion may not appear as 'new' to users with low update frequency, but as 'modified'.
- new or expired items will not be performed or expire retroactively, but only on the change date or later.
- changes will not be made retroactively or announced in advance, but are implemented on the change date itself***. (There are some exceptions on this with items in the distribution set of product codes).***

- items that have already expired will not be reused. But it is possible, e.q. for a product code, to reactivate the code. In this case a former distributed expire date will disappear and the original entry date will be maintained.
- Items that expired more than 7 years ago are not distributed anymore. (but at the moment this is only implemented at the distribution of the data of product codes and not yet at the distribution of the other data sets.

2.4 Description of code list

The description consists of:
- the definition of the entity type involved,
- a description of the clients,
- an example of the implementation of the code list,
- any relevant explanation or comments.

For each field in the code list, the following is specified:
- field number: sequence number of the field,
- field name: in English,
- occurence type:
  - M = mandatory (always completed)
  - C = conditional (under certain conditions),
- format:
  - N = numeric
  - AN = alphanumeric characters fixed number n, or variable: ..n,
- key fields:
  - P#: primary key,
  - F#: foreign key,
  - PF#: both primary as foreign key.

2.5 File format code list

The files (code lists) will be offered in CSV format (comma separated values) and compressed as a zip file with:

- Semicolon as field separator
  - when a semicolon is used in a field, the entire contents of the field must be enclosed in double quotes.
  - Examples:
    - apples and pears
    - “apples; pears”
  - This is similar to the convention described for the comma separated values format except that it is common practice in the Netherlands to use a semicolon instead of a comma. This is because the decimal separator in the Netherlands is a comma.
- CR-LF as a record separator
  - When CR-LF is used in a field, the entire contents of the field must be enclosed in quotation marks.
    Examples:
    - apples, pears
    - “applesCR-LF
      pears”
- Quotation marks as text indicator
  - When a quotation mark is used in a field and the entire content of the field is enclosed in quotation marks, each quotation mark in the text will receive an additional one. Example:
    - “text” becomes ““text””.

2.6 File name code list

The name of the database may be up to 8 characters (including first letter) with the file extension TXT.

1. The current files
   - the file name begins with the letter 'C'
2. The full (current and expired codes) files
   - the file name begins with the letter 'F'

File names in data sets first published before 1 January 2017

Position	Meaning	format
1 to 1	Designation current/full (C/F)	A1
2 to 2	Code list identification letter	A1
3 to 4	Day number (DD)	N2
5 to 6	Month number (MM)	N2
7 to 8	Year Number (YY)	N2

File names in data sets first published from 1 January 2017

Position	Meaning	format
1	Designation current/full (C/F)	A1
2 t/m 4	Code list identification number (is Code_list_id)	N3
5	Separator (“_”)	A1
6 t/m 7	Day number (DD)	N2
8 t/m 9	Month number (MM)	N2
10 t/m 11	Year Number (YY)	N2

File names GPC data sets (different)

Position	Meaning	format
1	Designation current/full (C/F)	A1
2 t/m 3	Code list identification number (is Code_list_id)	N2
4	Separator (“_”)	A1
5 t/m 8	Year Number (CCYY)	N4
9 t/m 10	Month number (MM)	N2
11 t/m 12	Day number (DD)	N2

2.7 Compression

The code lists are made available as a distribution set (data sets of code lists). The reason for this is so that all the code lists are coherent. The code lists are compressed using the PKZIP compression method.

2.8 Naming the distribution set

The name of the distribution set is made up as follows.

Position	Meaning	format
1 to 3	Code owner/administrator	A3
4 to 5	Sequence number of data set (supplied by owner/administrator)	A2
6 to 7	Release number	N2
8 to 9	Version number	N2

Example: VBN020101.ZIP
The Floricode release policy chain software is applicable to the Floricode Coding’s.

2.9 Control file

The data set is accompanied by a file with a checksum.

2.10 Conventions for naming control file

The same naming convention as applies to the distribution set but with extension TXT.

Example: VBN020101.TXT

2.11 Conventions for content of control file

The content of the control file consists of an MD-5 checksum value (there may be a prefix or suffix, such as the version number of the MD-5 routine included).

Example of content of control file:

MD5sums 1.1 freeware for Win9x/ME/NT/2000/XP+
Copyright (C) 2001-2002 Jem Berkes - http://www.pc-tools.net/

569ada27e5d62fdc0243ff20962860d2 VBN020101.zip

2.12 Code owner/administrator

Code	Meaning
BCK	Product codes for perennials
CLE	Client Export (now known as e-CertNL)
EBC	EdiBulb Codes
FEC	Floricode (Florecom)
FHC	FloraHolland Concern
FLC	Floricode
GS1	Global Standards One
ISO	International Standards Organization
VBN	Association of the Dutch Flower Auctions

2.13 Distribution

The code lists are distributed by Floricode as distribution sets (in the form of ZIP files by data administrator/owner).

The distribution of data sets occurs in the following ways:

- via the Floricode website: not possible anymore since April 2021
- via the FTP Server: codes.Floricode.com, Initial directory: codes, Port: 21
- FTP: upload flag: during uploading, the file appears: aFTPflagNowUploading
- FTP: end flag: after uploading, the file appears: aFinished.

A subscription to get access to the code lists via FTP can be obtained using the form ‘Access to code lists (FTP)’.

Datasets that are still in the testing phase will be available at locations that are yet to be confirmed.