

Algemene Implementatie Richtlijnen codepakketten 1.4

Verwerking codelijsten via FTP

Bestandsnaam	: HZ21022V1.0 Algemene Implementatie Richtlijnen codepakketten 1.4.docx
Bestandsnummer	: HZ/21022
Datum laatste wijziging	: 28 April 2021
Documentversie/release	: 1.0/1.4
Documentstatus	: Definitief



Copyright

© Floricode, 2021

All Rights reserved. No part of this edition may be multiplied, stored in electronic database or published in any form or way, electronically, mechanically, by photocopy recordings or any other way, without prior written permission from „Stichting Floricode“. For obtaining permission, contact Stichting Floricode at P.O. Box 115, 2370 AC Roelofarendsveen.

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.

Inhoudsopgave

Wijzigingsblad	4
1. Inleiding	5
1.1 Doel van dit document	5
1.2 Bestemd voor...	5
1.3 Uitgangspunten	5
1.4 Afkortingen en begrippen	5
2. Gemeenschappelijke aspecten voor de implementatie van codelijsten	6
2.1 Codelijst identificatie	6
2.2 Updatevelden	6
2.3 Update detectie	6
2.4 Beschrijving codelijst	7
2.5 Bestandsformaat codelijst	7
2.6 Bestandsnaam codelijst	8
2.7 Compressie	8
2.8 Naamgeving distributieset	9
2.9 Controle bestand	9
2.10 Naamgeving conventie controlebestand	9
2.11 Inhoudconventie controlebestand	9
2.12 Code eigenaar/beheerder	9
2.13 Distributie	9

Wijzigingsblad

Versie/release	Datum	Auteur	Wijzigingen
1.0	26-06-2007	G. Heemskerk	Initiële versie
1.1	15-08-2007	G. Heemskerk	Na review CMG
1.2	05-11-2007	G. Heemskerk	Beschrijving MD-5 controle bestand is toegevoegd (deze ontbrak).
1.3	05-02-2014	M. Goedhart	Uitbreiding in paragraaf 1.5 in verband met gebruik van puntkomma in velden
1.4	16-08-2016	M. Goedhart	Tekstuele aanpassingen
1.5	26-06-2017	M. Goedhart	Uitbreiding bestandsnaam codelijsten: <ul style="list-style-type: none"> - Bestandsnamen nieuwe gegevenssets na 1 januari 2017 - GPC Vanwege deze wijzigingen is versie/release van 1.2 in 1.3 opgehoogd
1.6	4-9-2019	H. Zwinkels	Correctie in paragraaf 3.3
1.7/1.3	12-03-2020	M. Goedhart	Paragraaf 3.5 (pagina 7): <ol style="list-style-type: none"> 1. wanneer een puntkomma in een veld gebruikt wordt, dient de volledige inhoud van het veld tussenaanhalingstekens (double quotes) te komen. 2. Wanneer CR-LF in een veld gebruikt wordt, dient de volledige inhoud van het veld tussen aanhalingstekens te komen. 3. Wanneer een aanhalingsteken in een veld gebruikt wordt en de volledige inhoud van het veld staat tussen aanhalingstekens, dan krijgt elk aanhalingsteken in de tekst een extra aanhalingsteken.
1.0/1.4	28-04-2021	H. Zwinkels	Redactionele aanpassingen

	Datum	Functionaris	Paraaf
Akkoord (auteur)	28-04-2021	H. Zwinkels	HZ
Gecontroleerd		A. Zwanenburg (Floricode) L. Zandvliet (Floricode) E. Bakker (Floricode)	
Geaccepteerd namens Werkgroep standaarden		Voorzitter	

1. Inleiding

1.1 Doel van dit document

Dit document beschrijft coderingen die volgens het zogeheten 'Linnaeus formaat' worden gedistribueerd. Dit 'Linnaeus formaat' is eerder beschreven in 'De VBN Technische Blauwdruk Linnaeus 2.10' (publicatie 7 mei 2013). Dit had toentertijd alleen betrekking op de distributie van de productcodes. Omdat steeds meer gegevensbestanden met stamdata volgens het 'Linnaeus formaat' door Floricode en andere codebeheerders worden gedistribueerd zijn deze gemeenschappelijke aspecten nu in dit document opgenomen. Hiermee kunnen alle distributiebestanden met stamdata op een gestandaardiseerde wijze door de software worden verwerkt.

1.2 Bestemd voor...

Dit document is bestemd voor gegevensafnemers van codelijsten die door Floricode via FTP worden gedistribueerd. Hiervoor heeft Floricode abonnementen in de vorm van codepakketten beschikbaar. Voor gegevensafnemers van codelijsten via API wordt verwezen naar de betreffende API documentatie.

1.3 Uitgangspunten

Dit document is gebaseerd op:

- VBN Technische Blauwdruk Linnaeus 2.10, paragraaf 3.1 'Gemeenschappelijke aspecten van de codelijsten'¹.

De volgende uitgangspunten worden gehanteerd:

- RFC 2007-1

1.4 Afkortingen en begrippen

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	Vereniging van Bloemenveilingen in Nederland
ZIP	Eén van de bestandstypes voor gecomprimeerde bestanden.

¹ Inmiddels is een geactualiseerde versie van dit document van toepassing waarin het betreffende hoofdstuk niet meer aanwezig is omdat dat in dit document is opgenomen.

2. Gemeenschappelijke aspecten voor de implementatie van codelijsten

De gemeenschappelijke aspecten van codelijsten die via FTP worden gedistribueerd volgens het 'Linnaeus formaat' worden in de onderstaande paragrafen beschreven.

2.1 Codelijst identificatie

Codelijsten worden gecomprimeerd aangeboden. Een gecomprimeerde distributieset (ZIP) bestaat uit één of meerdere codelijsten (een gegevensset). De naamgeving van de codelijst is volgens afspraken samengesteld (zie paragraaf 3.6). Ieder record in de codelijst begint met een bestandscode waarmee de codelijst wordt geïdentificeerd. Deze bestandscode is uniek binnen de gegevensset.

2.2 Updatevelden

Conform de VBN Blauwdruk Linnaeus zijn de volgende velden binnen ieder record van een codelijst opgenomen:

- Ingangsdatum (entry_date):
 - o datum waarop een nieuw item binnen de sector algemeen van kracht wordt of is geworden; dit kan een datum in de toekomst betreffen.
- Vervaldatum (expiry_date):
 - o datum waarop een item voor algemeen gebruik binnen de sector komt of is komen te vervallen; dit kan een toekomstige datum betreffen of leeg zijn.
- Mutatiedatum/tijd (change_date_time):
 - o Datum/tijd waarop een item voor het laatst is gemuteerd (opgevoerd, gewijzigd, vervallen).

2.3 Update detectie

Voor het vaststellen van nieuwe, vervallen of gewijzigde items dient de applicatie de publicatiedatum van de betreffende Floricode distributieset vast te houden en als volgt te bepalen welke items sedert die datum gemuteerd zijn:

- nieuw:
 - o change_date_time is later dan de publicatiedatum/tijd en entry_date is later dan of gelijk aan de change_date
- vervallen:
 - o change_date_time is later dan de publicatiedatum/tijd en expiry_date is later dan of gelijk aan de change_date
- gewijzigd:
 - o change_date_time is later dan de publicatiedatum/tijd en entry_date is eerder dan de change_date.

Voor de correcte werking van deze procedure zal het volgende worden gewaarborgd:

- vervallen items zullen in de lijsten opgenomen blijven, met uitzondering van items die zijn vervallen omdat ze foutief of ten onrechte waren opgevoerd; vervallen items behoeven niet per se permanent in de codelijsten gehandhaafd te blijven, maar kunnen na verloop van tijd (jaren) uit de lijsten verdwijnen.
- nieuwe items die kort na introductie worden gewijzigd kunnen door gebruikers met een lage updatefrequentie mogelijk niet als 'nieuw', maar als 'gewijzigd' worden gedetecteerd.
- nieuwe of vervallen items zullen niet met terugwerkende kracht worden opgevoerd respectievelijk komen te vervallen, maar slechts op de mutatiedatum of later.
- wijzigingen zullen niet met terugwerkende kracht worden aangebracht of vooraf worden aangekondigd, maar op de mutatiedatum zelf worden doorgevoerd (**Hierop bestaan thans uitzonderingen in de distributieset van de productcodes**).
- reeds vervallen items zullen niet opnieuw worden hergebruikt. Wel is het mogelijk items, bijvoorbeeld een productcode, te reactiveren. In dat geval verdwijnt een eerder gedistribueerde vervaldatum en de oorspronkelijke ingangsdatum blijft gehandhaafd.

- items welke langer dan 7 jaar geleden zijn vervallen worden niet meer gedistribueerd (N.B. Dit is thans alleen bij de distributie van de productcodes geïmplementeerd, bij de andere distributiesets nog niet).

2.4 Beschrijving codelijst

De beschrijving bestaat uit:

- de definitie van het betrokken entiteitstype,
- een beschrijving van de relaties,
- een voorbeeld van de invulling van de codelijst,
- een eventuele toelichting of opmerking.

Per veld in de codelijst wordt gespecificeerd:

- veldnummer: volgnummer van het veld,
- veldnaam: in het Engels,
- voorkomentye:
 - o M = mandatory (altijd gevuld)
 - o C = conditional (onder bepaalde voorwaarden gevuld),
- formaat:
 - o N = numeriek
 - o AN = alfanumeriek, aantal karakters vast: n, of variabel: ..n
- sleutelvelden:
 - o P#: primary key,
 - o F#: foreign key,
 - o PF#: zowel primary als foreign key.

2.5 Bestandsformaat codelijst

De bestanden (codelijsten) zullen in CSV formaat (comma separated values) en gecomprimeerd als zip bestand worden aangeboden met:

- Puntkomma als veldscheidingstekens
 - o wanneer een puntkomma in een veld gebruikt wordt, dient de volledige inhoud van het veld tussenaanhalingstekens (double quotes) te komen.
Voorbeelden:
 - appels en peren
 - "appels; peren"
 - o Dit is gelijk aan de conventie die is beschreven voor het comma separated values format behalve dat het in Nederland gebruikelijk is een puntkomma te gebruiken in plaats van een komma. Dit omdat het decimale scheidingstekens in Nederland een komma is.
- CR-LF als recordscheidingstekens
 - o Wanneer CR-LF in een veld gebruikt wordt, dient de volledige inhoud van het veld tussen aanhalingstekens te komen.
Voorbeelden:
 - appels, peren
 - "appelsCR-LF
peren"
- Aanhalingstekens als tekstindicator
 - o Wanneer een aanhalingsteken in een veld gebruikt wordt en de volledige inhoud van het veld staat tussen aanhalingstekens, dan krijgt elk aanhalingsteken in de tekst een extra aanhalingsteken.
Voorbeeld:
 - "tekst" wordt ""tekst""

2.6 Bestandsnaam codelijst

De naamgeving van de codelijsten binnen een distributieset verschilt en heeft de extensie TXT.

1. De actuele bestanden
 - de bestandsnaam begint met de letter 'C'
2. De volledige (actuele en vervallen codes) bestanden
 - de bestandsnaam begint met de letter 'F'

Bestandsnamen in distributiesets voor het eerst gepubliceerd voor 1 januari 2017

Positie	Betekenis	format
1	Aanduiding actueel/volledig (C/F)	A1
2	Codelijst identificatie letter	A1
3 t/m 4	Dagnummer (DD)	N2
5 t/m 6	Maandnummer (MM)	N2
7 t/m 8	Jaarnummer (JJ)	N2

Bestandsnamen in distributiesets voor het eerst gepubliceerd vanaf 1 januari 2017

Positie	Betekenis	format
1	Aanduiding actueel/volledig (C/F)	A1
2 t/m 4	Codelijst identificatie nummer (is Code_list_id)	N3
5	Scheidingsteken (" _ ")	A1
6 t/m 7	Dagnummer (DD)	N2
8 t/m 9	Maandnummer (MM)	N2
10 t/m 11	Jaarnummer (JJ)	N2

Bestandsnamen GPC distributieset (afwijkend)

Positie	Betekenis	format
1	Aanduiding actueel/volledig (C)	A1
2 t/m 3	Codelijst identificatie nummer (is Code_list_id)	N2
4	Scheidingsteken (" _ ")	A1
5 t/m 8	Jaarnummer (CCYY)	N4
9 t/m 10	Maandnummer (MM)	N2
11 t/m 12	Dagnummer (DD)	N2

2.7 Compressie

De codelijsten worden als distributieset (= set van codelijsten) beschikbaar gesteld. De reden hiervoor is de onderlinge samenhang die tussen codelijsten bestaat. De codelijsten worden met de compressiemethodiek PKZIP gecomprimeerd.

2.8 Naamgeving distributieset

De naamgeving van de distributieset is als volgt opgebouwd.

Positie	Betekenis	format
1 t/m 3	Code eigenaar/beheerder	A3
4 t/m 5	Volgnummer set van gegevens (door eigenaar/beheerder toegekend)	A2
6 t/m 7	Releasenummer	N2
8 t/m 9	Versienummer	N2

Voorbeeld: VBN020101.ZIP

Het Floricode releasebeleid ketensoftware is van toepassing op de Floricode Coderingen.

2.9 Controle bestand

De set van gegevens wordt vergezeld van een bestand met een controlegetal.

2.10 Naamgeving conventie controlebestand

Idem naamgevingconventie gegevensset maar dan met extensie TXT

Voorbeeld: VBN020101.TXT

2.11 Inhoudconventie controlebestand

De inhoud van het controlebestand bestaat uit een MD-5 checksum waarde (er kunnen voor- of na loop teksten aanwezig zijn zoals het versienummer van de MD-5 routine opgenomen).

Voorbeeld inhoud controlebestand:

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 eigenaar/beheerder

Code	Betekenis
BCK	Boomkwekerij productcodes
CLE	Client Export (thans e-CertNL van de NVWA)
EBC	EdiBulb Codes
FEC	Florecom (thans Floricode)
FHC	FloraHolland Concern
FLC	Floricode
GS1	Global Standards One
ISO	International Standards Organisatie
MPS	Stichting MPS
VBN	Vereniging van Bloemenveilingen in Nederland

2.13 Distributie

De codelijsten worden als distributiesets (in de vorm van ZIP bestanden per gegevensbeheerder/eigenaar) door Floricode gedistribueerd.

De distributie van distributiesets vindt via FTP op de onderstaande wijze plaats:

- Benaderbaar via website Floricode: niet meer ondersteund vanaf april 2021
- via FTP Server: codes.Floricode.com, Initiële directory : codes, Port : 21
- FTP: upload flag: tijdens uploaden verschijnt het bestand: aFTPflagNowUploading
- FTP: einde flag: na einde uploaden verschijnt het bestand: aFinished.

Een abonnement op de beschikbare codepakketten via FTP is te verkrijgen middels het aanvraagformulier op de website van Floricode 'Beschikbaarstelling Codelijsten (FTP)'. Distributiesets die in testfase zijn zullen op een nader te bepalen locatie beschikbaar komen.