

Wijzigingen NDFF Datamodel 2.0-->2.1

Versie: 1

Datum: 27 november 2017

Auteur: Willy-Bas Loos, Victor Mensing

1. Inleiding

In december 2017 wordt het datamodel van de database invoer ("telmee") aangepast. In dit document worden de wijzigingen beschreven die doorgevoerd worden.

Er worden wijzigingen doorgevoerd aan verschillende delen van het datamodel:

- Authenticatie
- Abundantie
- Flexibele velden ("Extra Info")
- Overige wijzigingen

2. Authenticatie

Tot nu toe waren er in de telmee database 2 velden voor wachtwoorden, in de tabel contacts.persons: webpassword en encrypted_password. In het eerste stonden nog steeds onversleutelde wachtwoorden, in het tweede wachtwoorden die met md5 versleuteld waren. In DM 2.1 vervalt het veld webpassword en het veld encrypted_password wordt grotendeels gevuld met random strings.

Daarvoor in de plaats komt de tabel public.auth_user, waarin de wachtwoorden versleuteld worden met het PBKDF2 algoritme, met een SHA256 hash en 36000 iteraties.

De reden dat het veld encrypted_password nog niet volledig verwijderd kan worden, is de telmee-loader. Die maakt nog gebruik van dit veld. Alleen voor de huidige telmee-loader-gebruikers (ongeveer 12 registraties op roltypid 85) zullen de wachtwoorden dus in tact blijven. In de volgende versie zal dit worden aangepast.

In auth_user staan ook velden voor email, first_name, last_name. Die worden **niet** gebruikt. Niet door de API en ook niet door andere applicaties. Wel is er op het moment van schrijven data ingevuld in veel gevallen, dit is mogelijk gebeurd tijdens de conversie van bestaande gebruikers.

Verder zijn er in auth_user de velden last_login en date_joined. Die worden door de API en ook door het invoerportaal ingevuld bij het aanmaken van een gebruiker en bij het inloggen door een gebruiker.

De werking van de velden "is_staff", "is_active" en "is_superuser" is niet relevant voor deze documentatie.

2.1 Synchroniteit met contacts.persons

Niet alle personen van contacts.persons staan in auth_user. Het is wel de bedoeling dat alle personen van auth_user in contacts.persons staan. Op het moment van schrijven is dat niet het geval en is er nog geen foreign key dat dat borgt.

De link kan gelegd worden met de gebruikersnaam.

2.2 Overzicht wijzigingen in tabelvorm

schema	tabel	veld	Wijziging, synchronisatie, gebruik
contacts	persons	webpassword	Wordt verwijderd
contacts	persons	encrypted_password	Wordt alleen nog door telmee loader gebruikt en na aanpassing daarvan verwijderd. Alleen gevuld voor gebruikers van de loader. Bij wijzigen wachtwoord wordt dit veld niet meer bijgewerkt!
contacts	persons	token	Token voor het onthouden van het wachtwoord , zodat het wachtwoord niet in de cookie hoeft te staan.
contacts	persons	token_date	tijdstip waarop de token is gezet of voor het laatst is aangepast. Na een in te stellen periode verloopt die token.
public	auth_user	* (tabel)	Wordt toegevoegd
public	auth_user	username	Behalve login naam ook de link met contacts.persons. Moet uniek zijn.
public	auth_user	first_name	Wordt niet gebruikt noch bijgehouden, is een standaard veld van django.
public	auth_user	last_name	Wordt niet gebruikt noch bijgehouden, is een standaard veld van django.
public	auth_user	email	Wordt niet gebruikt noch bijgehouden, is een standaard veld van django.
public	auth_user	password	Het wachtwoord, versleuteld met PBKDF2
public	auth_user	is_staff	systeemveld

public	auth_user	is_active	systeemveld
public	auth_user	is_superuser	systeemveld
public	auth_user	last_login	Tijdstip van laatste login. Wordt bijgewerkt door portaal en api.
public	auth_user	date_joined	Tijdstip waarop de gebruiker is aangemaakt.

3. Abundantie

Tot nu toe werd de waargenomen hoeveelheid opgeslagen in deze vier velden van de tabel survey.observations: countmin, countmax, count_method en originalabundance.

Die 4 velden komen in DM2.1 allemaal te vervallen en worden vervangen door 3 andere velden: scale, quantity en abundance_class.

Scale bevat de flditmid van een abundantieschaal, zoals “geschat aantal”, of “tansley”.

Quantity bevat (indien gevuld) een getal, mogelijk met decimalen, dat de waargenomen hoeveelheid direct uitdrukt. Uit de bijbehorende schaal blijkt wat het getal uitdrukt, bijvoorbeeld bedekkingspercentage of minimaal aantal.

Abundance_class bevat (indien gevuld) de flditmid van een abundantie klasse, zoals “tansley lf”

Afhankelijk van de waarde van scale wordt OF het veld quantity gevuld, OF abundance_class; altijd precies 1 van beide.

Geldige waarden voor het veld Scale staan in de nieuwe tabel codes.scales . Die nieuwe tabel was nodig om een extra eigenschap aan de field-items toe te voegen: storage_type. In dat veld wordt aangegeven in welk veld van survey.observations de waarden bij deze schaal opgeslagen moeten worden:

- De waarde “QUANTITY” betekent dat waarden getallen zijn en dus opgeslagen worden in het veld “quantity”
- De waarde “NOMINAL” betekent dat de waarden **abundantieklassen** zijn en dus opgeslagen worden in het veld “abundance_class”.

3.1 Constraints

Door de verschillende datatypes is de opslag van de abundantie niet eenvoudig en ook de controle op foute input is daardoor ingewikkelder dan normaal.

De volgende constraints borgen de correcte abundantie inhoud van survey.observations:

- Check_1aantal_ingevoerd: zorgt ervoor dat minstens 1 van beide is ingevuld: quantity of abundance_class.
- Observations_check_abundance: Checkt of de waarden scale, abundance_class en quantity bij elkaar horen. Maar heeft geen effect op NULL waarden in abundance_class en quantity.

- Quantity_gt_0: Tegen harde nullen en negatieve aantallen in quantity. Heeft geen effect op NULL. Geef aub harde nullen door als abundantieklasse aanwezigheid: afwezig.

3.2 Overzicht wijzigingen in tabelvorm

schema	tabel	veld	Wijziging, synchronisatie, gebruik
survey	observations	countmin	Wordt verwijderd
survey	observations	countmax	Wordt verwijderd
survey	observations	count_method	Wordt verwijderd
survey	observations	originalabundance	Wordt verwijderd
survey	observations	scale	Wordt toegevoegd. Verplicht veld. Bevat de schaal waarin de waargenomen hoeveelheid uitgedrukt wordt.
survey	observations	quantity	Wordt toegevoegd. Verplicht indien codes.scales.storage_type = "QUANTITY". Anders leeg (NULL). Getswaarde voor getelde hoeveelheid.
survey	observations	abundance_class	Wordt toegevoegd. Verplicht indien codes.scales.storage_type = "NOMINAL". Anders leeg (NULL). Id van klasse voor getelde hoeveelheid.
codes	scales	* (tabel)	Wordt toegevoegd.
codes	scales	flditmid	De field-item id's van geldige schalen. Primary key+Foreign key naar field_items.
codes	scales	storage_type	Verplicht veld. Kan 2 waarden bevatten: "QUANTITY" of "NOMINAL". Bepaalt in welk veld de waarde bij de scale wordt opgeslagen.

4. Flexibele velden ("Extra Info")

De tabel extra_info.observation_info wordt verwijderd. Alle velden die daarin stonden, worden ondergebracht in extra_info.custom_fields .

De tabel extra_info.custom_fields wordt bovendien opgeschoond.

Boolean waarden werden tot nu toe opgeslagen in het veld itemid. In DM2.1 worden die net als de velden met andere datatypes opgeslagen in het veld "fieldvalue" van de tabel custom_fields.

De volgende velden worden niet meer gebruikt en worden verwijderd:

- fieldname
- itemid
- capid
- fldtype

4.1 Overzicht wijzigingen in tabelvorm

De velden in het overzicht zijn in dezelfde volgorde als in de database, in plaats van in alfabetische volgorde.

schema	tabel	veld	Wijziging, synchronisatie, gebruik
extra_info	custom_fields	field_name	Wordt verwijderd
extra_info	custom_fields	itemid	Wordt verwijderd
extra_info	custom_fields	capid	Wordt verwijderd
extra_info	custom_fields	fldtype	Wordt verwijderd
extra_info	observation_info	* (tabel)	Wordt verwijderd. Onderstaande veld URI's beginnen allemaal met "http://ndff-ecogrid.nl/codes/keys". Dit wordt aangeduid met "...".
extra_info	observation_info	obsinfid	Wordt verwijderd
extra_info	observation_info	obsid	Wordt verwijderd
extra_info	observation_info	created	Wordt verwijderd
extra_info	observation_info	modified	Wordt verwijderd
extra_info	observation_info	creator	Wordt verwijderd
extra_info	observation_info	modifier	Wordt verwijderd
extra_info	observation_info	status	Wordt verwijderd
extra_info	observation_info	exportstatus	Wordt verwijderd
extra_info	observation_info	proof	Wordt verwijderd
extra_info	observation_info	certainty	Wordt verwijderd
extra_info	observation_info	xml	Wordt verwijderd
extra_info	observation_info	f_lifestage	Wordt verwijderd

extra_info	observation_info	f_sex	Wordt verwijderd
extra_info	observation_info	f_habitatuse	Wordt verwijderd
extra_info	observation_info	f_observationtype	Wordt verwijderd
extra_info	observation_info	perid	Wordt verwijderd
extra_info	observation_info	comment	Wordt verwijderd
extra_info	observation_info	locationname	Wordt verwijderd
extra_info	observation_info	f_tracks	Wordt verwijderd
extra_info	observation_info	taxgroup	Wordt verwijderd
extra_info	observation_info	locorsci	Wordt verwijderd
extra_info	observation_info	origcount	Wordt verwijderd
extra_info	observation_info	origcounttype	Wordt verwijderd
extra_info	observation_info	origdate	Wordt verwijderd
extra_info	observation_info	orignumber	Wordt verwijderd
extra_info	observation_info	formid	Wordt verwijderd
extra_info	observation_info	loctypid	Wordt verwijderd
extra_info	observation_info	error_status	Wordt verwijderd
extra_info	observation_info	recountgridcells	Wordt verwijderd
extra_info	observation_info	bronhouder	Wordt verwijderd
extra_info	observation_info	comnamid	Wordt verwijderd
extra_info	observation_info	evidence	Wordt verwijderd. Inhoud wordt verplaatst naar ../observation/evidence
extra_info	observation_info	dwelling	Wordt verwijderd. Inhoud wordt verplaatst naar ../observation/dwellings
extra_info	observation_info	cause_of_death	Wordt verwijderd. Inhoud wordt verplaatst naar ../observation/causeofdeath
extra_info	observation_info	substrate	Wordt verwijderd. Inhoud wordt verplaatst naar ../observation/substrate

extra_info	observation_info	origin	Wordt verwijderd. Inhoud wordt verplaatst naar ../observation/origin
extra_info	observation_info	equipment	Wordt verwijderd. Inhoud wordt verplaatst naar ../survey/equipment
extra_info	observation_info	broedcode	Wordt verwijderd. Inhoud wordt verplaatst naar ../observation/breedingcode
extra_info	observation_info	error_message	Wordt verwijderd.
extra_info	observation_info	remarks	Wordt verwijderd. Inhoud wordt verplaatst naar ../observation/remarks
extra_info	observation_info	orig_created	Wordt verwijderd
extra_info	observation_info	uncertain	Wordt verwijderd. Inhoud wordt verplaatst naar ../observation/certainty
extra_info	observation_info	specimen_code	Wordt verwijderd. Inhoud wordt verplaatst naar ../observation/collectionnr
extra_info	observation_info	webid	Wordt verwijderd
extra_info	observation_info	frmid	Wordt verwijderd

5. Overige wijzigingen

5.1 survey.observations.counttype

Het veld counttype van survey.observations heeft geen functie meer, het wordt verwijderd uit het datamodel. Het is geen deel van de abundantie, maar de wijziging wordt meegenomen in de slipstream.

Helaas heeft deze wijziging het niet gehaald in de release van van de telmee-loader. Daarom wordt het veld op acceptatie verwijderd na de acceptatie van de telmee loader, maar kan het op productie pas verwijderd worden wanneer de telmee loader alsnog aangepast wordt.

De acceptatie tests houden rekening met het verwijderen van dit veld.

5.2 folders.folders.locid

Er waren 2 manieren om de locatie bij een map aan te geven. Daarvan was folders.folders.locid de minst gebruikte. Dat veld wordt verwijderd uit het datamodel.

De data uit folders.folders.locid is overgezet naar folders.survey_info.locid.

5.3 folders.folders.created

Dit veld is nodig voor de administratie van het proces, de datastroom.

De inhoud wordt gevuld vanuit de default instelling, met now().

5.4 Overzicht wijzigingen in tabelvorm

schema	tabel	veld	Wijziging, synchronisatie, gebruik
survey	observations	counttype	Wordt verwijderd zodra de telmee loader is aangepast
folders	folders	locid	Wordt verwijderd, inhoud wordt verplaatst naar folders.survey_info.locid
folders	folders	created	Wordt toegevoegd. Default=now()