

# Nieuwe versie van de CBS open data services

Presentatie van Dennis Ramondt,  
Product owner open data



# Nieuwe versie van de CBS open data services

Verhaal in vier delen:

1. Verbeteringen in de techniek
  - voor u van waarde
2. Verbeterde aansluiting bij open standaards
  - toekomst bestendig
3. Inhoudelijke veranderingen
  - verruiming en verbetering aanbod
4. Van oud naar nieuw
  - wat betekent dit voor u



# Datasets continue beschikbaar, sleutels veranderen niet

Datasets continue beschikbaar:

- Bij een update blijft de oude versie **actief** tot **moment** van **publicatie**

Sleutels veranderen niet, deze fout treed niet meer op:

- ❑ **Error** getting **UntypedDataSet** for **'70072ned'**: **Could not find a property named 'GftAfval\_238'** on type **'Cbs.OData.UData'**.



# Verbeterde aansluiting op Internationale standaards (1):

## DCAT:

- Nederlands profiel apnl 1.1 toegepast
- Alle informatie nog steeds beschikbaar

## Veranderingen:

- Naamgeving van velden gewijzigd
- Enkele velden toegevoegd
- Informatie computer leesbaar



# Voorbeeld van oud en nieuw:

## Oud:

ID: 328  
Identificer: "37230ned"  
Title: "Bevolkingsontwikkeling; regio"  
ShortTitle: "Bevolkingsontwikkeling; regio"  
ShortDescription: "\nDeze tabel bevat cijfe\_aande"  
Summary: "Bevolkingsontwikkeling: \_e, \n"  
Modified: "2020-04-29T02:00:00"  
MetaDataModified: "2020-04-29T02:00:00"  
ReasonDelivery: "Actualisering"  
ExplanatoryText: ""  
OutputStatus: "Regulier"  
Source: "CBS."  
Language: "nl"  
Catalog: "CBS"  
Frequency: "Permaand"  
Period: "Januari 2002 - maart 2020"

## Nieuw:

Identificer: "37230ned"  
Title: "Bevolkingsontwikkeling; regio"  
Description: "\nDeze tabel bevat cijfe\_aande"  
Language: "nl"  
Authority: "https://standaarden.overheid.?"  
Modified: "2020-04-29T02:00:00+02:00"  
TemporalCoverage: "Januari 2002 - maart 2020"  
Catalog: "CBS"  
Publisher: "https://standaarden.overheid.?"  
ContactPoint: "infoservice@cbs.nl"  
Version: "202004290200"  
VersionNotes: ""  
VersionReason: "Actualisering"  
Frequency: "Per maand"  
Status: "Regulier"  
ObservationCount: 2161384



# Verbeterde aansluiting op Internationale standaards (2):

## Data cube vocabulary:

Model voor “meer” dimensionale data

- Waar mogelijk gevolgd en eigen invulling hiërarchie

Wordt toegepast op alle datasets van oData3

- Wettelijke plicht voor CBS, eens gepubliceerd, altijd toegankelijk

Eerste stap naar Linked Open data

- Voordelen pas over enkel jaren



# Voorbeeld: van dataset 84296NED de subsets

oData3	↔	oData4	Toelichting:
TableInfos	1 : 1	Properties	Algemene informatie dataset
DataProperties	1 : 2-3	MeasureCodes	Onderwerpen
		MeasureGroups	Hiërarchie van onderwerpen (facultatief)
		Dimensions	Informatie over dimensies
TypedDataSet	2 : 1	Observations	De daadwerkelijke cijfers. In vorm van cellen, niet langer records.
UntypedDataset			
Gewassen	1 : 1	GewassenCodes	Items van de dimensie, met aanvulling "Codes", <b>facultatief bij niet StatLine datasets</b>
CategoryGroups	1 : 0-7	GewassenGroups	Hiërarchie van de dimensie (facultatief)
Perioden	1 : 1	PeriodenCodes	Items van dimensie, met aanvulling 'Codes'
	0 : 1	PeriodenGroups	<b>Nieuw, hiërarchie van periode dimensie (geldt ook bij regio dimensie)</b>



# Voorbeeld: van dataset 84296NED (data)

Oud:

▼ 0:

ID:	0
Gewassen:	"A042170"
Perioden:	"2010JJ00"
VoorlopigeOogstraming_1:	" 154013"
DefinitieveOogstraming_2:	" 154013"
Vershil_3:	" 0.0"
DefinitieveOogstraming_4:	" 153723"
VershilGeogstTOVBeteeld_5:	" -0.2"
VoorlopigeOogstraming_6:	" 9.4"
DefinitieveOogstraming_7:	" 8.9"
Vershil_8:	" -5.3"

Nieuw:

▼ 0:

Id:	0
Measure:	"M003444_1"
ValueAttribute:	"None"
Value:	154013
Gewassen:	"A042170"
Perioden:	"2010JJ00"

▼ 1:

Id:	1
Measure:	"M003172_1"
ValueAttribute:	"None"
Value:	154013
Gewassen:	"A042170"
Perioden:	"2010JJ00"





# oData, van versie 3 naar versie 4

Van 2 x 3 services in oData3:

<https://opendata.cbs.nl/odatacatalog>

<https://opendata.cbs.nl/odataapi>

<https://opendata.cbs.nl/odatafeed>

<https://dataderden.cbs.nl/odatacatalog>

<https://dataderden.cbs.nl/odataapi>

<https://dataderden.cbs.nl/odatafeed>

Naar 1 service in oData4:

<https://beta-odata4.cbs.nl>



# 1) Slimme Identifiers bij Measures (onderwerpen)

Identifiers beginnend met M zijn variabelen

DatasetIdentifier	OData4Identifier	OData3Identifier	Title
70072ned	M0000352	TotaleBevolking_1	Totale bevolking
70072ned	4000	Vrouwen_3	Vrouwen
70072ned	70200_1	k_5Tot10Jaar_5	5 tot 10 jaar
70072ned	70200_2	k_5Tot10Jaar_14	5 tot 10 jaar
7379vt	X000222	TeZamenWonende_4	Te samen wonende

Identifiers **niet** beginnend met M zijn items uit een dimensie. Bij andere tabellen kun je deze identifiers terug vinden in de dimensie.

Identifiers beginnend met X staan in archief tabellen. Er is geen inspanning verricht om ze geschikt te maken voor koppelen

\_1 betekent dat het item dubbel voorkomt in dezelfde dataset.  
Volgnummer verandert niet in de tijd



# 1) Let op (de kleine letters, tot eind 2021):

DatasetIdentificer	OData4Identificer	OData3Identificer	Title
70072ned	M0000352	TotaleBevolking_1	Totale bevolking
70072ned	3000	Mannen_2	Mannen
70072ned	4000	Vrouwen_3	Vrouwen
70072ned	70200_1	k_5Tot10Jaar_5	5 tot 10 jaar
70072ned	X000222	k_5Tot10Jaar_15	5 tot 10 jaar

Als er bij een update een X verschijnt in een actieve tabel (geen archief), heeft het CBS een fout gemaakt. Dit proberen we zo snel mogelijk te herstellen en de waarde 70200\_2 terug te zetten, maar we kunnen niet garanderen dat we nooit een fout maken.

Om garanties te geven is het nodig dat we afscheid genomen hebben van oData3. We controleren op de identificers, maar deze controle is niet waterdicht. Onze inzet is erop gericht dat het niet voorkomt, maar we geven geen garanties.



## 2) Verbetering voor Excel gebruikers

Oud:

- Onderwerp**
- Regioaanduiding**
- Gemeentenaam
- Soort regio
- Codering
- Indelingswijziging wijken en buurten
- Bevolking**
- Aantal inwoners

Nieuw:

- Bevolking
  - Aantal inwoners
  - Bevolkingsdichtheid
  - Leeftijdsgroepen
    - 0 tot 15 jaar
    - 15 tot 25 jaar
    - 25 tot 45 jaar
    - 45 tot 65 jaar
    - 65 jaar of ouder
  - Personen met een migratieachtergrond
    - Westers totaal
    - Niet-westers
      - Niet-westers totaal

- ✓ Datasets bevatten alleen numerieke of string gegevens
- ✓ Informatie over gemeentenaam, soort regio opgenomen in nieuwe dataset
- ✓ Is niet van toepassing bij enkele oude datasets



# Deel 2 van de demo van het product

<https://beta-dataportal.cbs.nl>



# Overstap oData v3 naar oData v4



Dataportaal

Home

Datasets

Informatie voor

FAQ

Zoeken...



## Frequently asked questions

+	Heeft het CBS een helpdesk voor vragen over open data?
+	Waarom krijg ik een waarschuwing dat de browser de CSV download misschien niet kan verwerken?
+	Hoe kan ik datasets vinden in het dataportaal?
+	Welke OData 4 commando's zijn beschikbaar?
-	Wat zijn de belangrijkste verschillen tussen OData3 en OData4?

Het CBS maakt bij de nieuwe oplossing gebruik van de meest recente versie van OData (4.01). Het model is op een aantal plaatsen aangepast waarbij de grootste verandering is dat we van (Un)TypedDataSet (record oriëntatie) overstappen naar Observations (cel oriëntatie). Dit betekent een andere manier om onderwerpen te filteren. De onderwerpen hebben identifiers die niet veranderen in de tijd en die identiek zijn over datasets heen. Tot slot zijn de datasets continue beschikbaar. In de [handleiding](#) wordt uitleg gegeven wat de verschillen zijn.



Centraal Bureau  
voor de Statistiek

Handleiding

CBS Open data overstap van OData v3  
naar OData v4

April 2020

# Roadmap CBS

- Juni 2020, productie versie CCB met 53 StatLine + 3 nieuwe datasets
- 3<sup>e</sup> kwartaal uitbreiding naar alle datasets
- 4<sup>e</sup> kwartaal uitbreiding naar alle catalogi
- 1 juli 2021 Odata3 komt te vervallen

Voorbehoud:

- ✓ Intern onderhoud van de identifiers van de onderwerpen goed geregeld.



