



Ontwikkelingen bij het CBS

Inhoud en techniek gaan hand in hand

Dennis Ramondt

19-6-2019

Inhoud van de presentatie

- Terugblik afgelopen jaar
- Output visie
- Ontwikkelingen

Vragen tussendoor zijn welkom (beperkt) of aan het eind

Gebruikersbijeenkomst 29-3-2018:

Plannen 2^e half jaar 2018

3^e kwartaal 2018 (voorstel):

- Alle Nederlandstalige tabellen
- Engelstalige tabellen
- StatLine voor derden

4^e kwartaal 2018 (voorstel):

- Eenvoudige versie dataportaal
- Zoeken vernieuwen
- ...

Productie verstoringen zomer 2018

Probleem :

- 10 storingen open data services in 2 maanden tijd

Actie:

- Complete infra structuur vernieuwd
- Servers en programmatuur dubbel uitgevoerd inclusief upgrade

Open data draait vanaf 1-1-2019 stabiel



Gewijzigde inzichten

- Keuze voor productie omgeving in de Cloud
 - ✓ schaalbaarheid en zekerheid
- Prioriteit bij verbeteren meta data:
 - Sleutels (ID's) van onderwerpen wijzigen niet meer
 - ✓ toenemend gebruik open data
 - Bij onderwerpen worden dezelfde ID's gebruikt als bij dimensies
 - ✓ één keer een verandering voor de gebruikers open data
 - ✓ basis voor linked data
- Betere ondersteuning voor Data-analisten/Studenten
 - ✓ CSV download prioriteit

Bèta versie oData 4 (januari 2019)

Beta version



Open data v4

1. [Introductie](#)
2. [Catalogi](#)
3. [Datasets](#)
4. [Excel gebruik](#)
5. [Implementatie](#)
6. [Info](#)

Introductie

Het Centraal Bureau voor de Statistiek (CBS) ontwikkelt een nieuwe versie van haar open data API. Deze is gebaseerd op oData v4.0.

Catalogi

Het CBS beheert een aantal catalogi, maak uw keuze.

Catalogus van het CBS

Datasets



Nieuwe Datasets (geen StatLine)

Datasets

In de gekozen catalogus zitten de volgende datasets:

- [Regio-indelingen gekoppeld aan gemeente](#) (376.946 cellen)
- [Energielevering aan woningen en bedrijven; postcode 6, 2017](#) (1.219.456 cellen)
- [Omzetting identifiers van onderwerpen van Odata 3 naar Odata 4](#) (1.341 cellen)

Excel gebruik

Via een uitgebreide handleiding demonstreren wij het gebruik van onze [OData 4 service in Excel](#).

Implementatie

We hebben een overzicht gemaakt van de [geïmplementeerde OData functies](#).

Info

[Informatie pagina](#)



Belangrijkste wijzigingen oData 4: Van record naar Cel

- Nieuwe elementen:
 - Observations (waarden) en Measures (onderwerpen)
- Vervallen elementen:
 - TypedDataset, UnTypedDataset en Dataproperties
- Statistische symbolen vervangen door eigenschap van cel
 - Gebruik value attribute
 - Lege cellen worden niet meegeleverd (minder data)
- Services samengevoegd
 - oDataCatalog, oDataFeed en oDataAPI samengevoegd



Voorbeeld oud en nieuw

Oud, record

```
▼ 0:  
  WijkenEnBuurten: "GM1680"  
  AantalInwoners_5: 25243  
  Mannen_6: 12577  
  Vrouwen_7: 12666  
▼ 1:  
  WijkenEnBuurten: "WK168000"  
  AantalInwoners_5: 3555  
  Mannen_6: 1710  
  Vrouwen_7: 1840
```

Nieuw, cel

```
▼ 0:  
  Id: 110  
  Measure: "AantalInwoners_5"  
  ValueAttribute: "None"  
  IsNumeric: true  
  WijkenEnBuurten: "GM1680"  
  Value: 25243  
▼ 1:  
  Id: 216  
  Measure: "AantalInwoners_5"  
  ValueAttribute: "None"  
  IsNumeric: true  
  WijkenEnBuurten: "WK168000"  
  Value: 3555  
  Value: 12683  
  WijkenEnBuurten: "GM1680"
```

Wijzigingen in de aanroep van de services

- Onderwerpen selecteren via Filter ipv Select
 - Was:
select=Mannen_6, Vrouwen_7
 - Wordt:
filter=Measure eq '3000' or Measure eq '4000'
- Informatie staat in Dataset omzetting Identifiers van onderwerpen van Odata 3 naar Odata 4
- Naamgeving van de subsets veranderd

Onderwerpen

Identifiers beginnend met M zijn variabelen
b.v. Totale bevolking

DatasetIdentifier	OData4Identifier	OData3Identifier	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	70200_2	k_5Tot10Jaar_14	5 tot 10 jaar
7379vt	X000222	TeZamenWonende_4	Te zamen wonende

Identifiers beginnend met cijfer of T zijn classificaties. _1 en _2 betekent dat het item van een classificatie dubbel voorkomt in dezelfde dataset

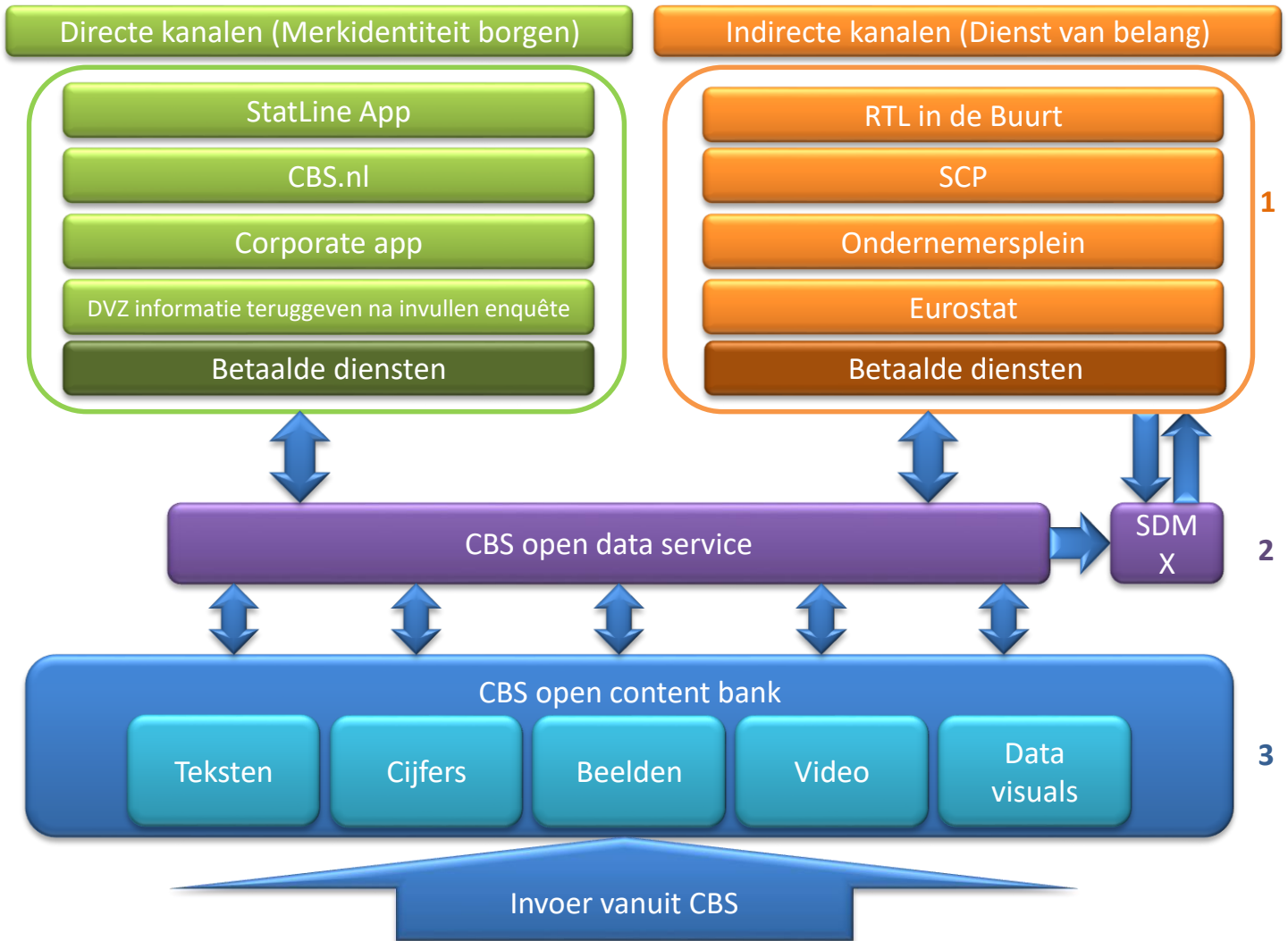
Identifiers beginnend met X staan in archief tabellen. Er is hier geen inspanning verricht om ze geschikt te maken voor linken (volgnummers).

Catalogi van Derden

Catalogus	StatLine nieuw/Widget	Dataportaal
MKB (Midden- en Kleinbedrijf)	https://mkbstatline.cbs.nl	https://mkbstatline.cbs.nl/portal.html
IV3 (Financiële gegevens decentrale overheden)	https://iv3statline.cbs.nl	https://iv3statline.cbs.nl/portal.html
Politie	https://www.politie.nl/politiecijfers	https://dataderden.cbs.nl/PolitieAngularWebApp/portal.html
RIVM	https://statline.rivm.nl	https://statline.rivm.nl/portal.html
Jeugdmonitor	https://jmopendata.cbs.nl	https://jmopendata.cbs.nl/portal.html
MLZ (Monitor langdurige zorg)	https://mlzopendata.cbs.nl	https://mlzopendata.cbs.nl/portal.html
AZW (Arbeid, zorg en welzijn)	https://azwstatline.cbs.nl	https://azwstatline.cbs.nl/portal.html
Mvstat (Databank van de Emancipatiemonitor)	https://mvstat.cbs.nl	https://mvstat.cbs.nl/portal.html
Interreg (Grens statistieken)	Najaar 2019	Najaar 2019

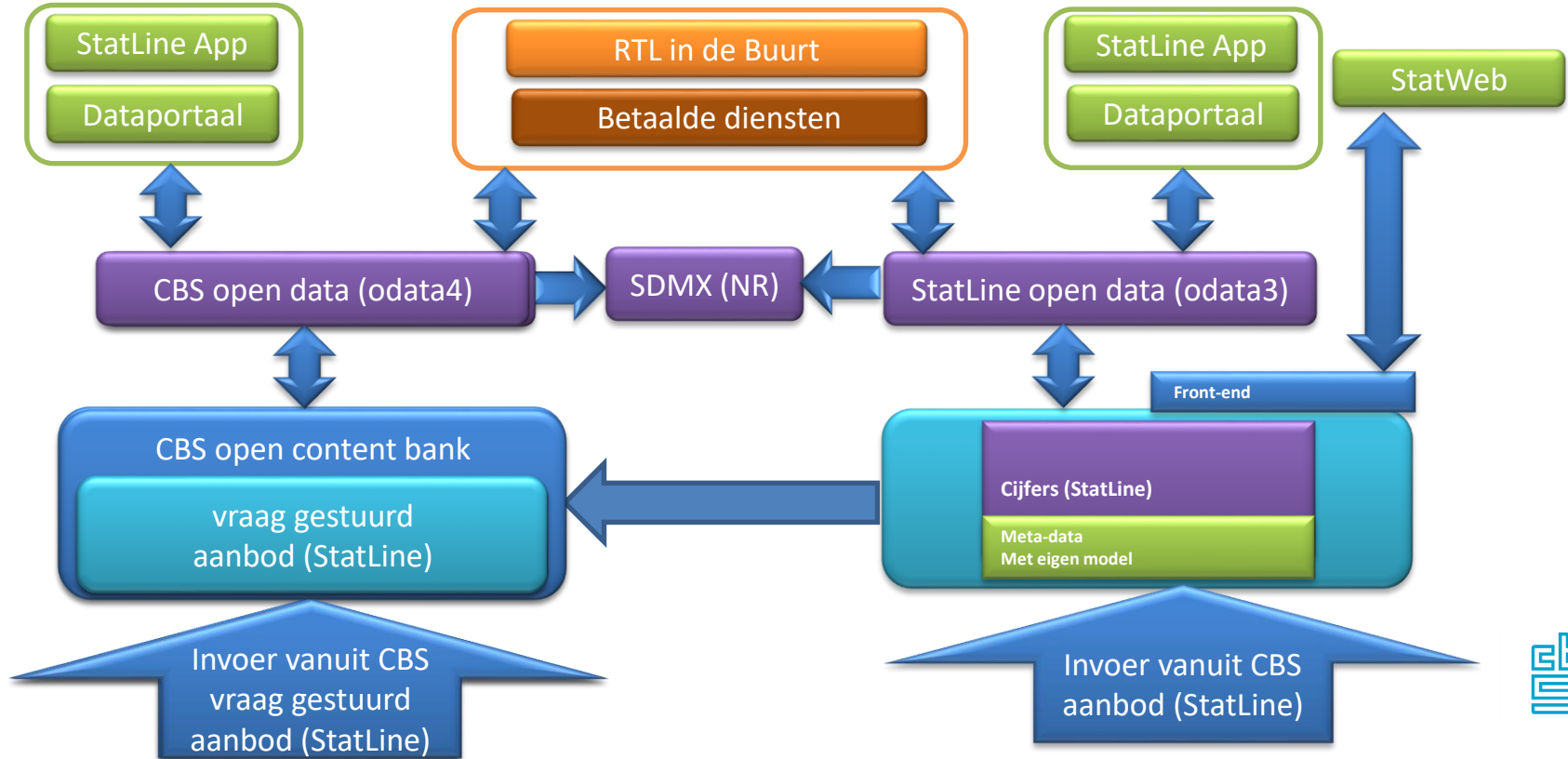
Output Visie





CCB (Content Cijferbank) Q2 2019

9 maanden draaien odata 3 en odata 4 naast elkaar voor gebruikers om over te kunnen stappen op nieuwe open data!!



Ontwikkelingen bij het CBS

Nieuwe Dataportaal (werk in uitvoering)

Home Datasets Informatie voor ▶ FAQ Zoeken...

Dataportaal

Introductie CBS publiceert de meeste cijfers in tabellen. Deze tabellen zijn toegankelijk via StatLine op de CBS-website. Om deze cijfers beter beschikbaar te maken, worden tabellen ook als open gegevens gepubliceerd. CBS

Wat is open data? Open Data is (overheids-)data die op eenduidige manier elektronisch beschikbaar wordt gesteld op een wijze dat programmeurs hier direct gebruik van kunnen maken voor allerlei toepassingen.

Webservices In deze service vindt u een overzicht van de beschikbare catalogi, inclusief enkele kenmerken van de catalogi. Op basis van de identifier, of kunt u de link naar de catalogus bouwen.

Catalogus	Aantal datasets
CBS	15
CBS-Maatwerk	1

Home Datasets Informatie voor ▶ FAQ Zoeken...

Dataportaal

Overheid; ontvangen belastingen en wettelijke premies

Laatste wijziging: 26-3-2019

Informatie Metadata Toegang

Statline nummer
84120NED

Omschrijving
Deze tabel bevat gegevens over de belastingopbrengsten en de ontvangsten van wettelijke sociale premies van de overheid. De gegevens zijn uitgesplitst naar soort belasting of wettelijke sociale premie. Onder overheid wordt hier verstaan de sector overheid volgens de definitie die in Europees Stelsel van Rekeningen (ESR 2010) wordt gebruikt. De gepresenteerde gegevens sluiten aan bij de publicaties over de Nationale rekeningen. De gebruikte begrippen sluiten aan bij de Nationale rekeningen. De Nationale rekeningen zijn gebaseerd op de internationale definities van het Europees Stelsel van Rekeningen (ESR 2010).

Gegevens beschikbaar vanaf:
Jaarcijfers vanaf 1995, kwartaalcijfers vanaf 2008 voor de belastingen en 1999 voor de wettelijke sociale premies.

Home Datasets Informatie voor ▶ FAQ Zoeken...

Dataportaal

Catalogi en Datasets

—	CBS	
81589NED	Bedrijven; bedrijfstak	937.545
82245NED	Bevolking en huishoudens; viercijferige postcode, 1 Januari 2013	326.174
37230ned	Bevolkingsontwikkeling; regio per maand	2.051.944
85136NED	Consumentenprijzen; basisjaren vanaf 1969	5.886
83131NED	Consumentenprijzen; prijsindex 2015=100	861.126
83435NED	Consumentenprijzen; werknemers laag, alle basisjaren 1969-1995	2.580
82931NED	Kerncijfers wijken en buurten 2014	1.348.572
83220NED	Kerncijfers wijken en buurten 2015	1.401.105

Home Datasets Informatie voor ▶ FAQ Zoeken...

Dataportaal

Overheid; ontvangen belastingen en wettelijke premies

Laatste wijziging: 26-3-2019

Informatie Metadata Toegang

Frequentie
Per kwartaal

Verslagperiode
1995 - 2018; 1999 kw I - 2018 kw IV

Reden levering
Actualisering

Laatste wijziging
26-3-2019



CSV download vanuit Dataportaal (1)

- Geen limieten op omvang CSV bestand
- Optie om te kiezen voor de code of voor de titel van een element
- Op basis van record
 - ✓ Keuze mogelijkheid welke dimensie in record wordt uitgesplit
- Preview van CSV download in Dataportaal

CSV download vanuit Dataportaal (2)

- Selectie opties voor onderwerpen en dimensies
- Opties om Postcode 6 te selecteren:
 - ✓ Gebruik Startswith functie
 - ✓ Met uitbreidingen (b.v. OR functie)
- Opties om alle gemeenten van 1 jaar te selecteren
- Opgebouwde link te kopiëren
- CSV download service die CSV aanmaakt die vervolgens gedownload wordt.

Uitbreiding model CCB (1)

Identifier	Title	Description	Kind	MapYear	ReleasePolicy	ClassificationIdentifier
WijkenEnBuurten	Wijken en buurten		GeoDetail	2017		D1234

- Toevoegen van classification identifier aan Dimensions.
- Is identifier van centrale lijst met classificaties.
- Basis voor meerdere talen

Uitbreiding model CCB (2)

MeasureCodes						
Identifier	Index	Title	MeasureGroupID	DataType	Unit	ClassificationIdentifier
M0000370	7	Aantal inwoners	M0000351	Long	aantal	
3000	9	Mannen	T001038	Long	aantal	D1200
4000	10	Vrouwen	T001038	Long	aantal	D1200
10680	12	0 tot 15 jaar	10000	Long	aantal	D1400

MeasureGroups				
ID	Index	Title	ParentID	ClassificationIdentifier
M0000351	6	Bevolking		
T001038	8	Geslacht	M0000351	D1400

- Toevoegen van classification identifier aan MeasureCodes en MeasureGroups.
- Is identifier van centrale lijst met classificaties.
- Basis voor meerdere talen

Classificaties en Variabelen beschikbaar als open data (proef)

- Dataset zonder observations
- Variabelen opgenomen in MeasureCodes en MeasureGroups
- Classifications opgenomen als Dimensions en afzonderlijke dimensions

Nieuwe zoek mogelijkheden

Keuzes:

- ✓ Zoeken via variabele
- ✓ Zoeken via classificatie
- ✓ In alle levende datasets

Alle datasets beschikbaar als open data

- Traject start vanaf september 2019
- Gebruik CloudFoundry
- Verwachting 1-2 kwartalen werk
- Als volledig start traject van dubbel draaien van oData 3 service en oData 4 service
- Periode van 9 maanden (tenzij anders advies van gebruikers)

Vragen?

??