

## Handleiding GeoDataWizard

Deze handleiding is bedoeld als instructie voor GeoDataWizard en tevens als een hulpmiddel bij onduidelijkheden. Als u niet over een geschikte csv file beschikt, kunt u deze [hier](#) downloaden

## Doelstelling GeoDataWizard

De GeoDataWizard is een uitbreiding op de eerste versie van de LD Wizard. LD Wizard is een open source project, geïnitieerd door het Netwerk Digitaal Erfgoed, waarmee kleine tabulaire datasets (CSV bestanden) kunnen worden getransformeerd tot Linked Data. De uitbreiding, door Kadaster geïnitieerd, zorgt ervoor dat geografisch gerelateerde elementen in de dataset, zoals een coördinaat of een adres ook als correcte Geo Linked Data wordt getransformeerd. Deze uitbreiding is ook als open source software beschikbaar, en wordt op termijn ook terug gebracht naar de broncode van de LD Wizard.

De GeoDataWizard is een data transformatie tool, en geen data cleaning tool. Het is dan ook verstandig om de data eerst te cleanen (denk aan hoofdletter gebruik, lege waardes, etc) en vervolgens op te slaan als een CSV bestand. Hiervoor kunnen standaard tools zoals Microsoft Excel, Google Spreadsheets of Libre Office worden gebruikt.

**Linked data** is een digitale methode voor het publiceren van gestructureerde gegevens, zodanig dat deze beschikbaar gemaakt kunnen worden op het internet en daardoor ook beter bruikbaar zijn (bron: [https://nl.wikipedia.org/wiki/Linked\\_data](https://nl.wikipedia.org/wiki/Linked_data)). Hiermee wordt de data “live” bevroegbaar, en kunnen talloze toepassingen gebruik maken van de data.

Een unieke feature in de GeoDataWizard is dat naast het downloaden van de getransformeerde Linked Data, ook de data direct kan publiceren in de “store” van het Platform Linked Data Nederland. Hiermee wordt de data direct voor iedereen bruikbaar, bijvoorbeeld met het maken van zogenoemde SPARQL Queries. Voorbeelden van SPARQL queries worden door het Kadaster Data Science Team gepubliceerd op [labs.kadaster.nl](https://labs.kadaster.nl). Een mooi voorbeeld op basis van een CSV bestand met de Stolpersteine in de Gemeente Zutphen is hier te vinden:

<https://labs.kadaster.nl/stories/hack-a-lod-2019/index.html>

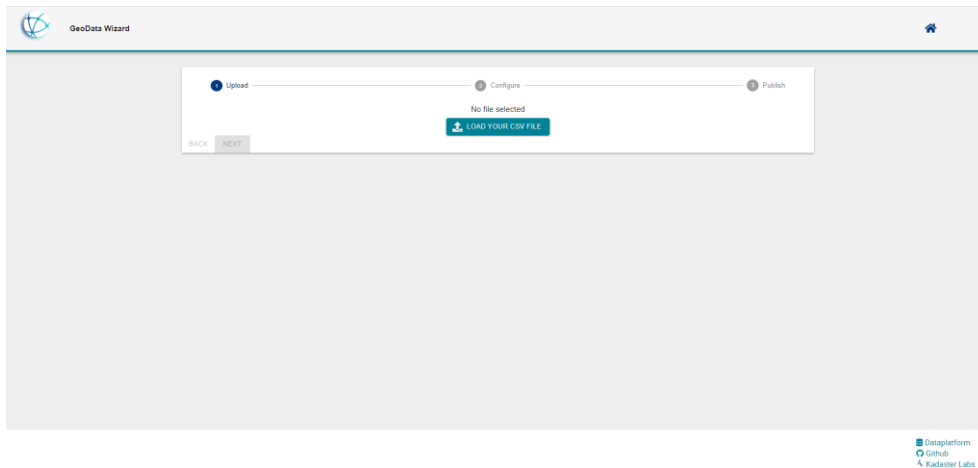
## Wat is er (nog) niet mogelijk?

De huidige versie van de GeoDataWizard kent, net als elk ander programma, een aantal dingen die (nog) niet kunnen. Voor deze wizard zijn dat de volgende punten:

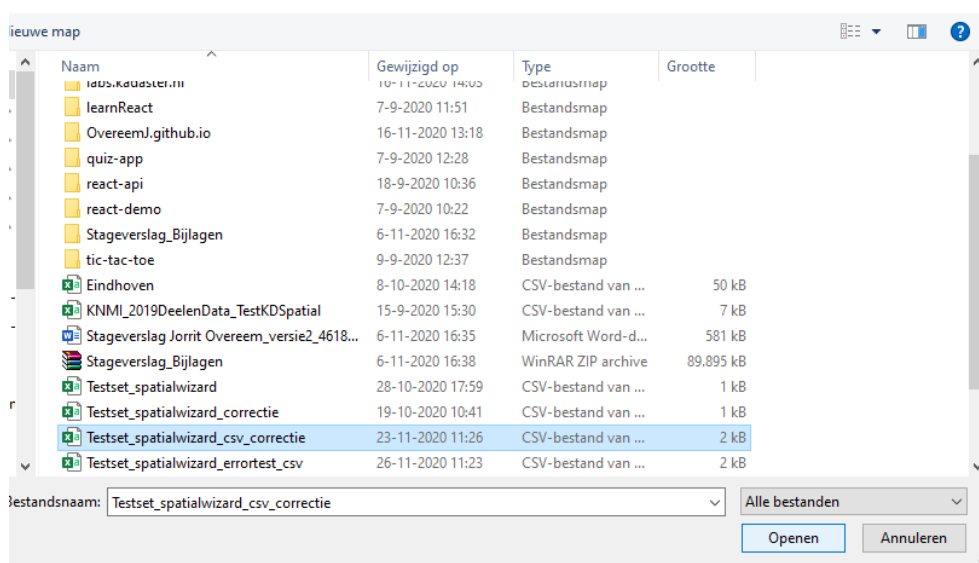
- Bestanden met een indeling anders dan csv kunnen niet worden geüpload.
- Correcties aan naamgevingen van adressen en woonplaatsen zult u zelf moeten doen in uw dataset. De wizard zal enkel aangeven bij welk adres of plaats het fout gaat, door middel van een melding tijdens het omzetten naar linked data.

## Stap 1: Uploaden van de dataset

De eerste stap voor het gebruik van de wizard is het uploaden van de dataset. De GeoDataWizard accepteert enkel csv bestanden. Dit zijn bestanden, meestal gemaakt met Excel, waarin de waarde met een komma gescheiden zijn. Om een bestand te uploaden klikt u in de hoofdpagina op de knop **Load your csv file** (afbeelding 1). Vervolgens verschijnt er een pop up waarin gevraagd wordt het bestand te kiezen die u wilt uploaden naar de GeoDataWizard (afbeelding 2).



Afbeelding 1 - Begin pagina met mogelijkheid tot uploaden van een CSV



Afbeelding 2 - Uploaden van het csv bestand

Vervolgens selecteert u het csv bestand van uw dataset en klikt op openen. Het venster verdwijnt en u wordt automatisch doorgestuurd naar de configuratie stap (Stap 2). Als de csv is ingeladen dan verschijnen de eerste 10 regels van uw dataset in een tabel (afbeelding 3)

Upload Configure Publish

Key column  
Row number

Resource class IRI  
http://bag.basisregistraties.overheid.nl/def/bag#

Advanced

Naam	Adres	Geslacht	Woonplaats	Geboorteplaats	Sterfdatum	Geboortedatum	Sterfteplaats	Latitude woonplaats	Longitude woonplaats	Coördinaten
Akkermans, H.	Herengracht 13	M	Amsterdam	Utrecht	30-4-2016	2-1-1965	Amsterdam	52.37403	4.88969	POINT(4.88969 52.37403)
Bijleveld M.	Vondellaan 178	V	Arnhem	Groningen	17-8-1981	20-4-1951	Arnhem	51.9851	5.89873	POINT(5.89873 51.9851)
Plesman, C.	Mozartsingel 6	M	Den Bosch	Den Bosch	18-10-1944	15-11-1924	Heerenveen	51.69917	5.30417	POINT(5.30417 51.69917)
Fokker, H.	Achter de Toren 16	V	Nijmegen	Amsterdam	31-3-2018	13-4-1980	Nijmegen	51.8425	5.85278	POINT(5.85278 51.8425)
Roode - Akkermans, F.	Herengracht 13	V	Amsterdam	Nijmegen	18-11-2015	13-12-1960	Amsterdam	52.37403	4.88969	POINT(4.88969 52.37403)
Dolmans, J.	De Peel 15	M	Lelystad	Venray	2-5-1945	23-07-1871	Lelystad	52.50833	5.475	POINT(5.475 52.50833)
Boels, B.	Tangbroek 12	M	Venlo	Maastricht	18-9-2011	27-8-1980	Venray	51.37	6.16806	POINT(6.16806 51.37)
Martens, H.	Grote Markt 1	M	Nijmegen	Nijmegen	18-1-2004	12-2-1930	Nijmegen	51.8425	5.85278	POINT(5.85278 51.8425)

BACK NEXT

Afbeelding 3 - Dataset ingeladen voor configuratie

## Stap 2: Configuratie

In de configuratie stap heeft u een aantal mogelijkheden voor het configureren van de data. Er zijn met deze GeoDataWizard twee verschillen configuraties mogelijk. Configuratie als de **key column** en **resource class IRI** gelden voor de gehele tabel. Verder zijn er nog aparte configuraties te doen per kolom, deze zijn zichtbaar zodra er op een kolomkop wordt geklikt (afbeelding 4).

Bij de optie **key column** kunt u de sleutelkolom instellen. Deze kolomwaarden worden aan de resource class IRI toegevoegd met een id/. Dit moet echter wel een kolom zijn met unieke waarde zoals een serienummer. Bij **resource class IRI**, kunt u de resource IRI instellen die geldt als resource voor de in te stellen properties. Als u de **key column** en **resource class** leeg laat dan gaan de standaardwaarden gelden. Dit is voor mensen die niet bekend zijn met linked data zijn de standaardwaarde de beste optie.

### Stap 2.1: Kolom configuratie

Per kolom kan er een aparte configuratie worden uitgevoerd met een aantal instellingen. Deze instellingen zijn van belang om goede linked data te kunnen maken van uw dataset.

Column configuration (Woonplaats)

Datatype  
Select datatype

Property configuration  
property URI  
http://bag.basisregistraties.overheid.nl/def/bag/dev#Woonplaats

Value configuration  
Select config type

CONFIRM CANCEL

Afbeelding 4 - Kolom configuratie met de verschillende opties

## Datatype instellen

Stel u hebt een kolom met een aantal onderdelen van een historisch item. En deze wilt u graag niet als tekst configureren omdat er later nog een bepaalde berekening mee wilt doen. Dan kunt u ervoor kiezen om deze kolom als datatype int of float in stellen bij de instelling **Datatype**. Voor deze instelling zijn er verschillende mogelijkheden:

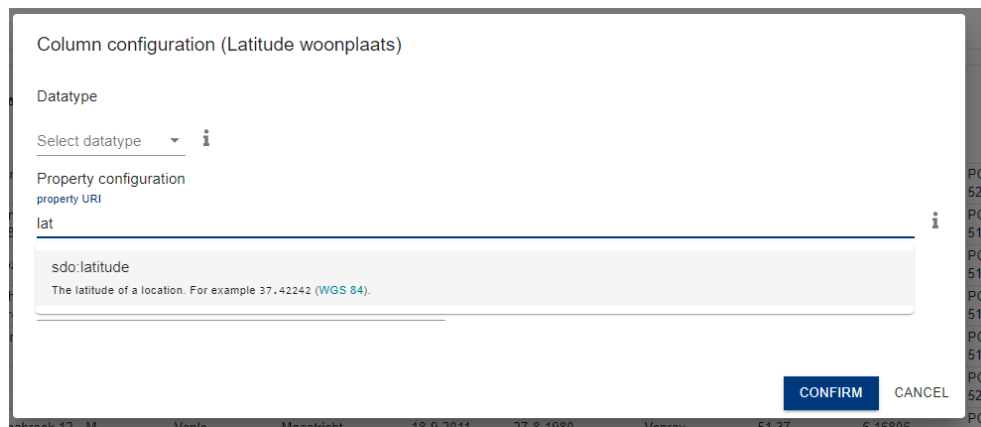
- String, dit is voornamelijk voor tekst. Dit is ook de standaardwaarde als datatype leeg gelaten is
- Integer (int), gebruik bij gehele getallen
- Float, gebruik bij decimalen getallen, zoals een coördinaatpunt
- WKLiteral, dit is een speciaal datatype die u kunt gebruiken als er eventueel coördinaten worden gebruikt in een aparte kolom. Let op: Deze waarden moeten als POINT(lat, long) zijn aangegeven in uw dataset. Zie voorbeeld in afbeelding 3.

## Property instellen

Elk kolomwaarde heeft in de link datastructuur een eigen property, of wel eigenschap. Het geeft het type aan van de waarde. Op die manier kan er specifiek per kolom een type toegewezen worden.

### Voorbeeld:

Stel u hebt de dataset geüpload zoals deze is te zien in afbeelding 3 en u wilt een geometrisch type (of property) toewijzen aan de kolom Latitude woonplaats. Dan klikt u op de kolom Latitude woonplaats dan verschijnt er een pop-up zoals in afbeelding 4 te zien is. Dan bij **Property configuration** typt u *latitude*, de GeoDataWizard zal dit dan herkennen als sdo:latitude en een suggestie geven.(afbeelding 5)



Afbeelding 5 – Zoek suggestie resultaten op basis ingevoerde zoektermen

Klik op de optie sdo:latitude en er zal **schema.org/latitude** verschijnen als property URI (afbeelding 6)


Column configuration (Latitude woonplaats)

Datatype

Select datatype 


Property configuration

property URI

<http://schema.org/latitude> 

sdo:latitude

Value configuration

Select config type 

**CONFIRM** CANCEL

Afbeelding 6 - Property URI verschijnt op basis van geselecteerde suggestie

Zodra je dan op **confirm** klikt dan zal deze property verschijnen onder de kolom naam (afbeelding 7)

Naam	Adres	Geslacht	Woonplaats	Geboorteplaats	Sterfdatum	Geboortedatum	Sterfteplaats	Latitude woonplaats sdo:latitude	Longitude woonplaats	Coördinaten
Akkermans, H.	Herengracht 13	M	Amsterdam	Utrecht	30-4-2016	2-1-1965	Amsterdam	52,37403	4,88969	POINT(4.88969 52.37403)
Bijleveld M.	Vondellaan 178	V	Arnhem	Groningen	17-8-1981	20-4-1951	Arnhem	51,9851	5,89873	POINT(5.89873 51.9851)
Plesman, C.	Mozartsingel 6	M	Den Bosch	Den Bosch	18-10-1944	15-11-1924	Heerenveen	51,69917	5,30417	POINT(5.30417 51.69917)
Fokker, H.	Achter de Toren 16	V	Nijmegen	Amsterdam	31-3-2018	13-4-1980	Nijmegen	51,8425	5,85278	POINT(5.85278 51.8425)
Roode - Akkermans, F.	Herengracht 13	V	Amsterdam	Nijmegen	18-11-2015	13-12-1960	Amsterdam	52,37403	4,88969	POINT(4.88969 51.37403)
Dolmans, J.	De Peel 15	M	Lelystad	Venray	2-5-1945	23-07-1871	Lelystad	52,50833	5,475	POINT(5.475 52.50833)

Afbeelding 7 - Ingestelde property verschijnt onder de kolomnaam


Dit kan voor de andere kolommen ook op dezelfde manier ingesteld worden.

### Value configuration instellen

Met deze optie heeft u een aantal mogelijkheden, en wordt voornamelijk gebruikt om de kolomwaardes te transformeren naar IRI's waarmee de linked data wordt gevormd. Dit is eigenlijk de belangrijkste optie in deze stap. (afbeelding 8)


Column configuration (Latitude woonplaats)

Datatype

Select datatype 


Property configuration

property URI

<http://schema.org/latitude> 

sdo:latitude

Value configuration

Select config type 

- Value to IRI
- Link Woonplaats, Sterfteplaats or Geboorteplaats to BAG
- Link Geopoint
- Link Adres with BAG

**CONFIRM** CANCEL

Afbeelding 8 - Selecteren van de value configuratie

De mogelijkheden zijn:

- Value to IRI, deze optie zet de kolomwaardes om naar een IRI zoals deze is ingesteld is in de resource class in combinatie met de property. Dit is mogelijk voor alle kolommen
- Link Woonplaats Geboorteplaats or Sterfteplaats to BAG, Deze optie is alleen mogelijk voor kolommen die waardes bevatten met een stad of dorp. Dit geeft de applicatie ook aan als u de optie selecteert.
- Link Geopoint, Deze optie is alleen mogelijk als de kolom coördinaatpunten bevat met een POINT(.. ..) waarde. Een voorbeeld van zo een kolom is de kolom **Coördinaten** in afbeelding 7 of 3. Deze optie zal de punten linken met een BRT-gebied in de BRT door middel van een identificatie nummertje.
- Link Adres with BAG, deze optie is mogelijk als u een adres kolom hebt met straatnaam en huisnummer en daarnaast een woonplaatsen, sterfteplaatsen of geboorteplaatsen heeft staan in de dataset. Als u deze optie aanklikt zal er een nieuwe select box komen met keuze uit kolomnamen waaraan u het adres wilt koppelen. (afbeelding 9)

Column configuration (Latitude woonplaats)

Datatype

Select datatype

Property configuration

property URI

<http://schema.org/latitude>

sdo:latitude

Value configuration

Select config type

Link Adres with BAG

This function can only be used for addresses. Also be sure to make connection with a city

Select Column

Naam

Naam

Adres

Geslacht

Woonplaats

Geboorteplaats

Sterfdatum

Geboortedatum

Sterfteplaats

Latitude woonplaats

Longitude woonplaats

Coördinaten

CONFIRM CANCEL

Naam	Adres	Geboortedatum	Geboorteplaats	Latitude woonplaats	Longitude woonplaats	Coördinaten
...	...	...	...	...	...	...

Afbeelding 9 - Adres koppelen aan de BAG

Zodra de gewenste configuratie is ingesteld kunt u op **confirm** drukken om de instellingen te bevestigen.

Vervolgens klikt u op **Next** om door te gaan naar het publicatie scherm. De GeoDataWizard zal uw configuratie dan meenemen en uw dataset gereed maken voor publicatie. Bij een succesvolle afhandeling komt er bovenin in het groen een melding te staan dat de conversie succesvol is. (afbeelding 10)

Upload

Configure

3 Publish

Configuration succesfull completed

Download results

Download CSV

Download your tabular source data as standardized CSV.

DOWNLOAD CSV

Download RDF

Download the transformed Linked Data (RDF) to your local machine.

DOWNLOAD RDF

Download script

Download a script that you can use to run the transformation yourself. The following script languages are supported: RATT, CoW, RML.

DOWNLOAD RML

Publish to Kadaster or PLDN

Token account

Select Account

Token

LOAD TOKEN

PUBLISH

Remember

BACK

NEXT

Afbeelding 10 - Publicatie scherm als de configuratie met succes is uitgevoerd

### Stap 3: Publicatie

In deze stap kunt u er zelf voor kiezen om de geconfigureerde data te downloaden of te publiceren naar Kadaster of naar Platform Linked Data Nederland (PLDN). Voor de publicatie heeft u wel een account nodig bij Kadaster of PLDN om het token te kunnen aanvragen. Deze kunt u aanvragen via (contactpersoon / website / contactformulier) van Triply.

#### Het downloaden van de resultaten

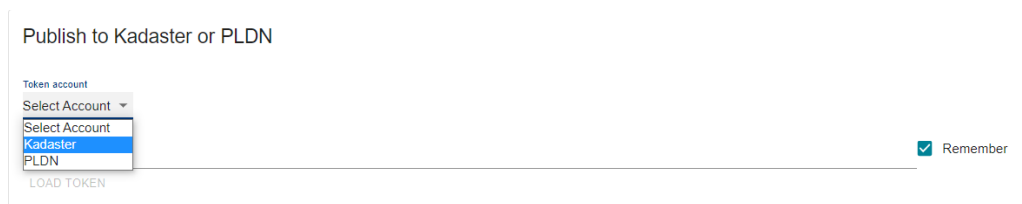
Voor het downloaden van de resultaten zijn er 3 mogelijkheden:

- Resultaten downloaden als **CSV**, deze optie downloadt de originele data terug als csv.
- Resultaten downloaden als **RDF**, deze optie downloadt de output van de configuratie als RDF-script.
- Resultaten downloaden als **Script**, met deze scripts kan de transformatie handmatig worden gedraaid. Deze optie vereist enige kennis van linked data en het uitvoeren van transformaties.

#### Het publiceren van de data

U kunt ook uw dataset publiceren naar de Kadaster en PLDN-dataplatformen. Onder **Token account** kunt u het account selecteren waarna u de dataset wilt publiceren. PLDN is een algemeen dataplatform die ook publiekelijk toegankelijk is. Het kadaster platform is meer gericht op geodata, deze bevat ook datasets van de BAG en BRT.

De eerste stap naar online publiceren van de data is het selecteren van het account. Als u op de selectbox klikt verschijnen twee keuzes, Kadaster en PLDN. (afbeelding 11)



Afbeelding 11 - Kiezen van het account

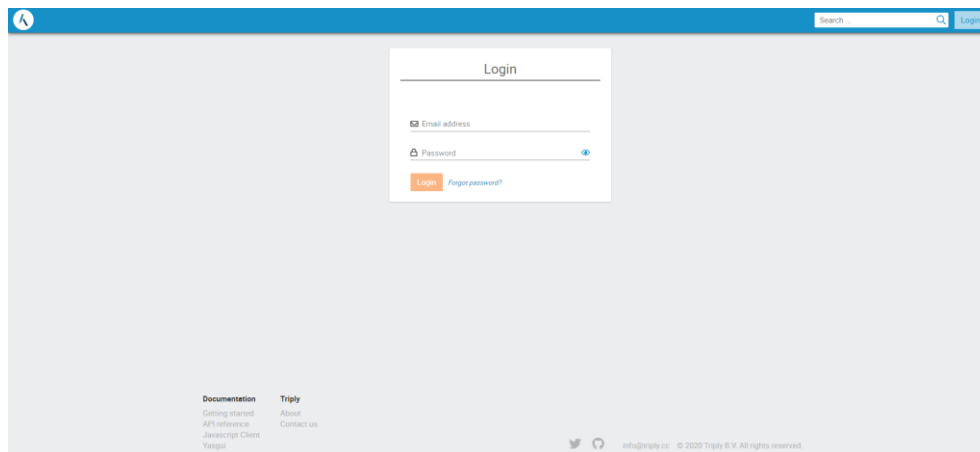
Voor deze voorbeeld stappen kiezen we voor Kadaster. Onder de selectbox verschijnt er dan een link om een token op te halen. (afbeelding 12)



Afbeelding 12 - Gekozen account met onderstaande link om een token te verkrijgen

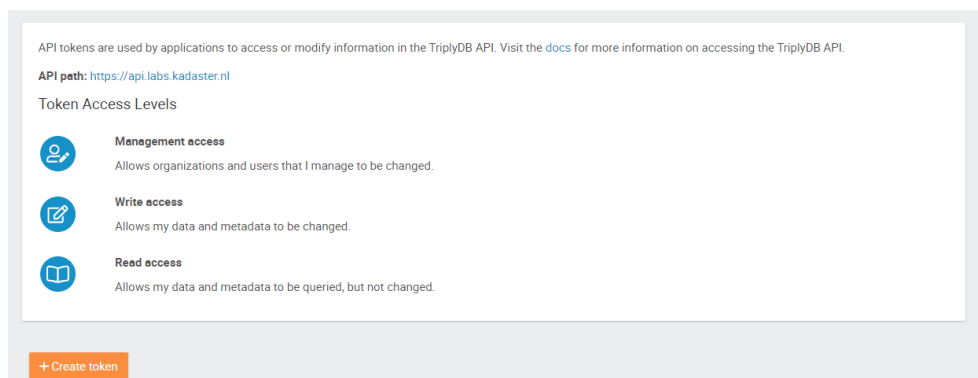
Als u op de link klikt wordt u doorverwezen naar een inlogpagina van het Kadaster dataplatform. Er zal gevraagd worden om een gebruikersnaam en wachtwoord. (afbeelding 13) Voor het voorbeeld wordt gebruik gemaakt van een voorbeeld account. U zal zelf een account moeten aanvragen.





Afbeelding 13 - Inlogscherf om het token op te halen

Log in met uw gebruikersnaam en wachtwoord en het volgende scherm verschijnt om een token aan te maken. (afbeelding 14)



Afbeelding 14 - Het aanmaken van een token, beginscherf

Een token werkt op 3 toegangsniveaus. Hiermee kunt u restricties leggen op uw gepubliceerde data voor andere gebruikers. Standaard zit er een **Read access** restrictie op, hiermee kan uw data alleen gelezen worden maar niet bewerkt door externe gebruikers. De **management access** houdt in dat organisaties en gebruikers kunnen worden aangepast, ook kan de data worden veranderd en gelezen. **Write access** wil zeggen dat de data kan worden aangepast en gelezen.

Klik op **+ Create token** om een nieuwe token aan te maken. Het volgende verschijnt:

## Create Token

Token name

Something that allows you to identify this token in the future (e.g., the name of the application that will use it).

### Access Level



Management access: Allows organizations and users that I manage to be changed.



Write access: Allows my data and metadata to be changed.



Read access: Allows my data and metadata to be queried, but not changed.

Create

Cancel

Afbeelding 15 - Aanmaken van een token, formulier met restrictie keuze en token naam

Onder **Token name** kan er een unieke naam worden gegeven aan de nieuwe token. Vervolgens kiest u welke access level uw wilt toekennen aan dit token. Als u dit niet doet is de access level standaard **Read access**. (afbeelding 15)

## Create Token

Token name

GeoDataToken

Something that allows you to identify this token in the future (e.g., the name of the application that will use it).

### Access Level



Management access: Allows organizations and users that I manage to be changed.



Write access: Allows my data and metadata to be changed.



Read access: Allows my data and metadata to be queried, but not changed.

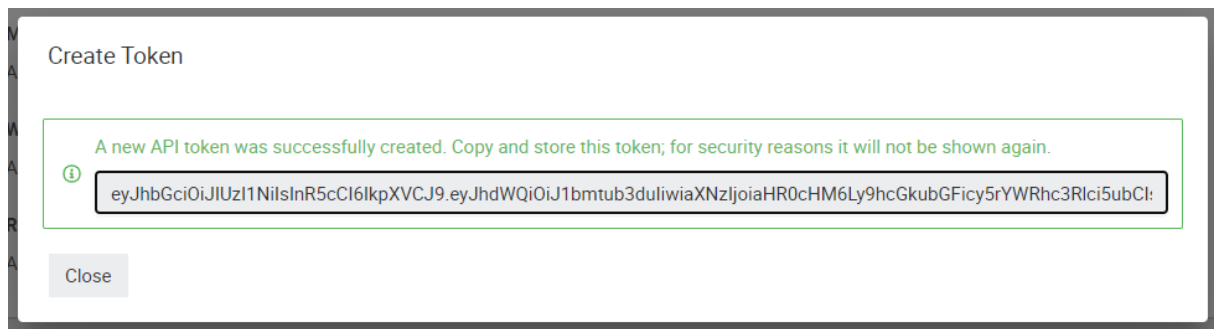
Create

Cancel

Afbeelding 16 - Ingevuld token formulier

In dit voorbeeld kiezen voor management access. Later in de publicatie stap is er nog een mogelijkheid om een nieuwe lege dataset aan te maken om uw dataset naar te publiceren. Om dit te laten zien is er voor management access gekozen in dit voorbeeld. (afbeelding 16)

Klik vervolgens op **Create** en uw nieuwe token verschijnt als pop in beeld. Kopieer dit token en bewaar deze goed. Dit token wordt eenmalig verstrekt. (afbeelding 17)



Afbeelding 17 - Aangemaakt token

Na kopiëren klik vervolgens op **Close**. En keer terug naar de GeoDataGeoDataWizard en plak deze in het **Token** veld. (afbeelding 18)

Afbeelding 18 - Volledig ingevuld publicatie formulier om het account op het halen

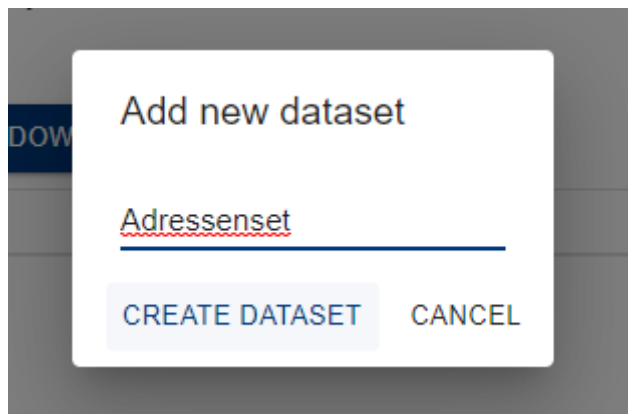
En klik op **LOAD TOKEN** om de token te laden. De GeoDataWizard zal, als het token juist is, uw account weergeven waar de dataset naar gepubliceerd wordt. (afbeelding 19)

Afbeelding 19 - Keuze van account en dataset op basis van token

Druk op **PUBLISH** om uw dataset te publiceren op kadaster. Als dat gedaan is het kunt u via de link **Click here to view your results** de resultaten bekijken op de kadaster dataplatform. (afbeelding 20)

Afbeelding 20 - Succesvolle publicatie

**Optioneel:** Vervolgens is er nog een mogelijkheid om een nieuwe lege dataset aan te maken, klik op het plusje naast **Dataset** om een nieuwe dataset aan te maken. Dan verschijnt er een klein formuliertje. (afbeelding 21)



Afbeelding 21 - Aanmaken van een nieuwe dataset

Daarin geeft u een naam op voor de nieuwe dataset en klikt vervolgens op **CREATE DATASET**. Deze nieuwe dataset zal dan verschijnen in de **Dataset** selectbox. (afbeelding 22)



Afbeelding 22 - Nieuwe dataset toegevoegd die gebruikt kan worden om naar te publiceren