

---

# **NLExtract Documentatie**

***Release 1.1.3***

**NLExtract Team**

**Feb 12, 2020**



---

## Contents

---

<b>1</b>	<b>Algemeen</b>	<b>3</b>
1.1	Waarom NLExtract ? . . . . .	3
1.2	Meer dan Extraheren . . . . .	4
<b>2</b>	<b>Services</b>	<b>5</b>
2.1	NLExtract Downloads . . . . .	5
2.2	NLExtract Apps . . . . .	5
<b>3</b>	<b>BAG-Extract</b>	<b>7</b>
3.1	Handleiding BAG-Extract . . . . .	7
<b>4</b>	<b>TOP10-Extract</b>	<b>15</b>
4.1	Handleiding TOP10-Extract . . . . .	15
<b>5</b>	<b>BGT-extract</b>	<b>17</b>
5.1	Handleiding BGT-extract . . . . .	17
<b>6</b>	<b>BRK-extract</b>	<b>19</b>
6.1	Handleiding BRK-extract . . . . .	19
<b>7</b>	<b>Stetl-framework</b>	<b>21</b>
7.1	Afhankelijkheden . . . . .	21
7.2	Installatie . . . . .	22
7.3	Stetl uitvoeren . . . . .	23
<b>8</b>	<b>Downloads - Instructie</b>	<b>25</b>
8.1	Instructie (Windows - GUI) . . . . .	25
<b>9</b>	<b>Links</b>	<b>67</b>
<b>10</b>	<b>Contact</b>	<b>69</b>
<b>11</b>	<b>Indices and tables</b>	<b>71</b>



De laatste versie van deze documentatie vind je steeds op <http://docs.nlextract.nl>. Je kunt dit document in verschillende formaten downloaden zoals in PDF. Zie <https://readthedocs.org/projects/nlextract/downloads>.

NLExtract wordt ontwikkeld via de [NLExtract GitHub](#) .

Inhoud:



Hieronder staat algemene informatie over het hoe en waarom van NLExtract. Zie ook de website: <http://nlextract.nl> en de ontwikkeling [NLExtract GitHub](#).

Voor commerciële ondersteuning zoals gespecialiseerde extracties, downloads en andere services neem *Contact* met ons op.

### 1.1 Waarom NLExtract ?

Nederlandse overheidsinstellingen zoals Het Kadaster en Rijkswaterstaat stellen hun (geo) data meer en meer beschikbaar als Open Data. Bijvoorbeeld de BAG (Basisregistratie Adressen en Gebouwen, [www.kadaster.nl/BAG](http://www.kadaster.nl/BAG)) levert je alle adressen en gebouwen in Nederland met hun coördinaten. De Top10NL ([www.kadaster.nl/top10nl](http://www.kadaster.nl/top10nl)) bevat gegevens voor de gehele topografie van Nederland, veel gedetailleerder dan Google Maps dat biedt.

Dat is dus prachtig, want nu kan iedereen deze data zelf downloaden en naar eigen believen gebruiken. . . Ok, downloaden, maar dan ? Wat je op dit moment geleverd krijgt is niet bijvoorbeeld een digitale kaart maar de “ruwe data”, d.w.z. de vector-bestanden met alle punten, lijnen en vlakken tezamen met hun vele kenmerken. Bijvoorbeeld de straat van A naar B in geometrische lijn-coördinaten en vlak-coördinaten, maar ook met haar straatnaam en wegnummer. Dat is mooi, want dan kun je bijv. zelf bepalen welke kleur die straat op de kaart krijgt en met welk font de naam afgebeeld wordt. . . .

Maar. . . er moet nog flink wat gebeuren voor je een echte gedetailleerde kaart kunt zien of bij wijze van spreken de oppervlakte van je eigen huis (BAG). Het is alsof je een enorm spreadsheet met cijfers krijgt waarvan je eigenlijk de grafieken wilt zien. Wat is er aan de hand en wat moet er dan gebeuren ?

De gegevens en bestands formaten waarin deze overheids-datasets worden aangeleverd zijn dus ruwe data. Deze leveringen komen in XML (GML), CSV formaten en zelfs als MS Access database. Deze bestanden zijn dan ook bedoeld voor uitwisseling, zodat je zelf kunt bepalen wat en hoe je ze “op de kaart” wilt hebben. Er is ook vaak uitgebreide documentatie van deze formaten, maar voor de gemiddelde kaarten-maker kan dat (bijv. GML-schema’s) abacadabra zijn.

Om bijvoorbeeld een kaartdienst te maken met een open standaard zoals de Web Map Service (WMS) is het veel en veel handiger om met bijvoorbeeld een ruimtelijke database als PostGIS ([www.postgis.org](http://www.postgis.org)) te werken. Dan kun je

bijvoorbeeld een “query” maken om alleen zeg maar de rijkswegen of alle naaldbossen te extraheren uit de Top10NL gegevens. Of om de oppervlakte van je eigen huis te achterhalen uit de BAG.

Ook wil je vaak data combineren en/of afleiden uit verschillende data-sets. Een voorbeeld is het verrijken van BAG data met gemeente en provincie grenzen. Dit is in theorie allemaal binnen je bereik, echter er is nog net een stapje nodig: de aangeleverde bestanden omzetten naar een ruimtelijke database zodat je er echt mee aan de slag kunt. In de Open Source wereld is de op PostgreSQL gebaseerde geo-database PostGIS ([www.postgis.org](http://www.postgis.org)) de standaard. Ook kun je gemakkelijk een kaartdienst (via WMS) of datadienst (via WFS) met Open Source server-software als GeoServer of MapServer op een PostGIS database aansluiten. Via “Styled Layer Descriptors (SLD, soort CSS) kun je dan zelf je kaarten vormgeven en naar voren laten komen wat voor jou belangrijk is (voor mij bijvoorbeeld liever naaldbossen dan rijkswegen).

Een heel verhaal maar hoe zetten we die data dan om naar PostGIS en maken we een begin met die SLDs ? Daarom is er nu NLExtract !

NLExtract levert tools, recepten, voorbeelden om Nederlandse geodata sets te converteren en te ontsluiten. Het gehele traject van brondata conversie (naar PostGIS, later ook naar bijv Oracle en Shape) tot visualisatie (SLDs) wordt afgedekt. Er wordt binnen NLExtract zoveel mogelijk gebruik gemaakt van bestaande Open Source Geo tools zoals GDAL/OGR ([www.gdal.org](http://www.gdal.org)). Bij voorkeur wordt vector data geconverteerd naar PostGIS (en later ook rasterdata naar GeoTIFF).

De eerste datasets die gedaan zullen worden zijn BAG, Top10NL, IMGeo/BGT (Grootschalige Topografie) en NWB (Nationaal Wegen Bestand van Rijkswaterstaat). Ook zijn er plannen om raster datasets te doen zoals TopRaster en “Bonne Bladen” (historische topo-kaarten).

Iedere dataset heeft hieronder een eigen directory met aanwijzingen hoe te converteren en te visualiseren.

De GitHub hier bevat puur de NLExtract-broncode dus is in de eerste plaats gericht op ontwikkelaars. Wil je als ontwikkelaar meedoen ? Graag ! Laat ons weten, bijv. via [info@opengeogroep.nl](mailto:info@opengeogroep.nl). Vooral zoeken we nog Windows-experts zodat het op dat platform ook goed gaat werken. Wil je nieuwe zaken zien of vind je fouten ? Gebruik de “issues” link hierboven.

## 1.2 Meer dan Extraheren

Het installeren en uitvoeren van de NLExtract ETL (extracties) kan vaak tijdrovend zijn terwijl de meeste gebruikers vaak hetzelfde doel hebben: bijv maandelijks de BAG omzetten naar PostGIS of de laatste versie van Top10NL. Daarnaast kan het zijn dat voor andere datasets zoals [OpenTopo](#) je een web service zoals een TMS of WMS wil opzetten. Het kan ook zijn dat je bijv op smartphone/tablet eenvoudigweg de OpenTopo kaarten wil raadplegen.

Daartoe zijn in 2014 in NLExtract drie nieuwe NLExtract Services opgestart.

- NLExtract Downloads
- NLExtract Apps

In 2020 zijn de NLExtract Downloads te vinden via <https://geotoko.nl>. Veel informatie en voorbeeldbestanden vind je via <https://geocatalogus.nl>. De “Ga Naar Bron” links verwijzen naar [geotoko.nl](https://geotoko.nl).

Lees in het volgende hoofdstuk *Services* meer hierover.



Het installeren en uitvoeren van de NLExtract ETL (extracties) kan vaak tijdrovend zijn terwijl de meeste gebruikers vaak hetzelfde doel hebben: bijv maandelijks de BAG omzetten naar PostGIS of de laatste versie van Top10NL. Daarnaast kan het zijn dat voor andere datasets zoals [OpenTopo](#) je een web service zoals een TMS of WMS wil opzetten. Het kan ook zijn dat je bijv op smartphone/tablet eenvoudigweg de OpenTopo kaarten wil raadplegen.

Daartoe zijn in 2014 in NLExtract twee nieuwe NLExtract Services opgestart.

- NLExtract Downloads
- NLExtract Apps

Sinds februari 2020 zijn de NLExtract Downloads (betaald) te vinden via [geotoko.nl](#). Veel informatie en voorbeeldbestanden vind je op [geocatalogus.nl](#).

## 2.1 NLExtract Downloads

De *NLExtract Download Service* wordt verzorgd via [geotoko.nl](#).

Onder andere kunnen hier BAG, BRK, BGT en BRT-Top10NL PostGIS dumps worden gedownload. Deze worden zo snel mogelijk verversd wanneer nieuwe versies van betreffende datasets uitkomen. Normaal gesproken 12x per jaar voor de BAG en BRK-DKK. De BGT wordt maandelijks verversd. BRT plm 4x per jaar.

In de README bestanden bij de betreffende downloads vind je meer informatie.

Om gebruik te maken van de database dumps bieden we een overzicht om vlot tot een ‘up and running’ situatie te komen. Kijk hier voor een *instructie*.

Voor gespecialiseerde downloads, bijv voor andere databases of zeg de BAG per gemeente, neem [contact met geotoko.nl](#).

## 2.2 NLExtract Apps

In aanbouw. Basis link is <http://app.nlextract.nl>

Heb je zelf een mooie app op basis NLExtract? Misschien wil je die toevoegen. Dit loopt de [NLExtract GitHub](#).

Hieronder staat de handleiding voor het gebruik van de tools voor de BAG, NLEextract-BAG of BAG-Extract geheten.

### 3.1 Handleiding BAG-Extract

BAG-Extract is onderdeel van de NLEextract tools voor het inlezen en verrijken van Kadaster BAG (Basisregistratie Adressen en Gebouwen) GML leveringen in (voorlopig) een Postgres/Postgis database.

#### 3.1.1 BAG Bronbestanden downloaden

De BAG Leveringsbestanden (totaal plm 1.5 GB .zip) worden iedere maand ververs en zijn te downloaden via deze PDOK link: <http://geodata.nationaalgeoregister.nl/inspireadressen/atom/inspireadressen.xml> (Atom feed). Als je wilt testen met een kleiner bestand kun je via [https://data.nleextract.nl/bag/bron/BAG\\_Amstelveen\\_2011feb01.zip](https://data.nleextract.nl/bag/bron/BAG_Amstelveen_2011feb01.zip) ook de “BAG Amstelveen” (5.6 MB) downloaden. Let wel dat de bestandsstructuur van de Amstelveen-levering afwijkt van de tegenwoordige BAG-leveringen.

#### 3.1.2 BAG PostGIS dumps downloaden

Als je geen zin/tijd hebt om NLEextract-BAG zelf te installeren en te draaien, dan kun je ook de PostGIS database en CSV adressen dumps downloaden via <https://geotoko.nl>. Deze worden maandelijks ververs direct na uitkomen nieuwe BAG-levering. Een leveringsbijdrage is vereist.

#### 3.1.3 Wat doet BAG-Extract?

BAG-Extract biedt de volgende functionaliteiten:

- Laden van een Kadaster BAG Extract vanuit Kadaster (.zip GML) levering
- Toepassen van Kadaster BAG mutaties vanuit Kadaster (.zip GML) levering (v1.1.5 en hoger)

- Verrijken BAG met gemeenten en provincies
- Verrijken: tabel met volledige “ACN-achtige” adressen genereren
- Geocoderen: afgeleide tabellen en functies
- Stijlen (SLDs) om de ingelezen BAG data te visualiseren via een WMS
- Validatie van input vlak geometrie
- Database VIEWS om bagobjecten te selecteren die actueel, bestaand en valide geometrie hebben
- Checkpointing: bijhouden welke bestanden reeds verwerkt t.b.v. herstarts en mutatie-verwerking
- Logging in database zodat gechecked kan worden waar evt fouten zijn en welke bestanden verwerkt

### 3.1.4 BAG-Extract downloaden

- download NLEExtract, zie laatste versie op: <https://github.com/nlextract/NLEExtract/releases>. De nieuwste versie staat bovenaan: kies de “real-release” nl extract zip.
- of snapshot (huidige repo versie) via git: `git clone http://github.com/opengeogroep/NLEExtract.git`
- of snapshot als .zip: <https://github.com/nlextract/NLEExtract/zipball/master>

### 3.1.5 Afhankelijkheden

- PostgreSQL: relationele database, minimaal versie 9.1, optimaal is versie 9.5, Aangemaakte DB moet **UTF8 encoding** hebben. zie <http://www.postgresql.org>
- PostGIS: spatial extensie PostgreSQL, bijv. opslag geodata, minimaal versie 2.0, optimaal is versie 2.2, zie <http://postgis.org>
- Python: versie 2, minimaal versie 2.4.3, beste is 2.7 voor lxml, geen Python 3
- Python argparse package, voor argument parsing alleen indien Python < 2.7
- psycopg2: Python PostgreSQL client bibliotheek. Zie <http://initd.org/psycopg>
- lxml voor razendsnelle native XML parsing, Zie <http://lxml.de>
- GDAL/OGR tools en bibliotheek voor geodata manipulatie. Minimaal 1.8.1. Zie <http://gdal.org>
- GDAL/OGR Python bindings Zie [www.gdal.org](http://www.gdal.org) en <http://pypi.python.org/pypi/GDAL> voor Geometrie parsing/validatie en manipulatie

### 3.1.6 Installatie (Linux)

- Ubuntu: beste is om eerst UbuntuGIS PPA aan je package repo toe te voegen, voor laatste versie Geo-tools als GDAL en PostGIS.

```
apt-get install python-software-properties
add-apt-repository ppa:ubuntugis/ubuntugis-unstable
apt-get update
```

- Ubuntu: PostgreSQL+PostGIS . PostgreSQL is een OS relationele database (RDBMS). PostGIS is een extensie die van PostgreSQL een ruimtelijke (spatial) database maakt. Installatie PostgreSQL 9.1 + PostGIS 2.1

```

$ apt-get install postgis postgresql-9.1 postgresql-contrib

# Server Instrumentation, met admin pack.
$ sudo -u postgres psql
psql (9.1.10)
Type "help" for help.

postgres=# CREATE EXTENSION adminpack;
CREATE EXTENSION

# Installatie controleren met ::

$ psql -h localhost -U postgres template1

$ pg_lsclusters
Ver Cluster Port Status Owner    Data directory          Log file
9.1 main     5432 online postgres /var/lib/postgresql/9.1/main /var/log/postgresql/
↳postgresql-9.1-main.log

# Enablen locale connecties in ``/etc/postgresql/9.1/main/pg_hba.conf``.

# Database administrative login by Unix domain socket
local  all                postgres                    md5

# TYPE  DATABASE          USER            ADDRESS              METHOD

# "local" is for Unix domain socket connections only
local  all                all                    md5
# IPv4 local connections:
host   all                all                127.0.0.1/32        md5
# IPv6 local connections:
host   all                all                ::1/128              md5

# PostGIS en template opzetten. Ook dit nodig om Postgis extension aan te maken.
$ apt-get -s install postgresql-9.1-postgis-2.1

# Anders krijg je op ``CREATE EXTENSION postgis`` dit ::

# ERROR: could not open extension control file "/usr/share/postgresql/9.1/
↳extension/postgis.control": No such file or directory

# Template DB``postgis2`` opzetten. ::

$ su postgres
createdb postgis2
psql -h localhost postgis2
postgis2=# CREATE EXTENSION postgis;
# CREATE EXTENSION
postgis2=# CREATE EXTENSION postgis_topology;
# CREATE EXTENSION

# Nieuwe database "bag" aanmaken met template "postgis2"
# NB belangrijk is dat de bag DB de character-set UTF8 (-E UTF8) heeft!
createdb --owner postgres -T postgis2 -E UTF8 bag

```

- optioneel: Python package afhankelijkheden installeren bijv

```
apt-get of yum install python-setuptools (voor easy_install commando)
apt-get of yum install python-devel (tbv pycopg2 bibliotheek)
apt-get of yum install postgresql-devel (tbv pycopg2 bibliotheek)
```

- Onder Ubuntu zijn dat de volgende packages

```
sudo apt-get install python-setuptools
sudo apt-get install python-dev
sudo apt-get install libpq-dev
```

- razendsnelle native XML parsing met libxml2/libxslt libraries samen met Python lxml: kan meer dan een factor twee in snelheid schelen... Zie <http://lxml.de/installation.html>

```
apt-get of yum install libxml2
apt-get of yum install libxslt1.1
apt-get of yum install python-lxml
```

- GDAL ([www.gdal.org](http://www.gdal.org)) met Python bindings voor OGR geometrie-parsing en geometrie-validatie (NLX v1.1.0 en hoger)

```
apt-get of yum install gdal-bin
apt-get of yum install python-gdal
```

- de PostgreSQL python bibliotheek pycopg2

```
sudo easy_install pycopg2
```

- Python package “argparse”

```
sudo easy_install argparse
```

- NB als je een proxy gebruikt via http\_proxy doe dan easy\_install -E (exporteer huidige environment)

### 3.1.7 Installatie (Windows)

De installatie van BAG-Extract op Windows werd in het verleden gekenmerkt door lastige installaties, vanwege het feit dat open source ontwikkeling op Windows gefragmenteerd plaatsvindt. Tegenwoordig is het een stuk gemakkelijker om BAG-Extract aan de praat te krijgen. Als je zelf BAG-Extract wilt uitvoeren, voer dan onderstaande beschrijving uit. Voor het gebruiken van de PostGIS-dump, volg dan de instructie die door Geert Doornbos beschikbaar is gesteld ([work in progress](#)).

Benodigheden:

- PostgreSQL 9.x: <https://www.postgresql.org/download/>.
- PostGIS 2.x: wordt geïnstalleerd via de Stack Builder van PostgreSQL.
- Python 2.7: <https://www.python.org/downloads/windows/>, momenteel is versie 2.7.11 de meest recente versie. Neem de 64-bits versie als je een 64-bits machine hebt. Let op, als je Python via de hoofdpagina downloadt, krijg je de 32-bits versie.
- Recente GDAL-versie (1.11 of 2.x): te installeren via QGIS of via de OSGeo4W installer (niet getest).
- Python bindings: <http://www.lfd.uci.edu/~gohlke/pythonlibs/>. Nodig zijn:
  - lxml (alleen getest met lxml-3.4.4)
  - pycopg (dit is Pycopg2)

- gdal

Neem de cp27-versies. Neem de win\_amd64.whl-versie als je 64-bits Python gebruikt.

Let bij het downloaden van de software of je de 32-bits of de 64-bits versie gebruikt. De 64-bits versie werkt op de meeste recente computers. De 32-bits versie werkt op alle Windows-computers, maar issues met het geheugen zijn dan niet uitgesloten.

Installatie:

- PostgreSQL, PostGIS en het aanmaken van een spatial database: zie de instructie voor het terugzetten van de PostgreSQL dump. Kort gezegd komt het op het volgende neer:
  - PostgreSQL: voer de installer uit.
  - PostGIS: via de Stack Builder van PostgreSQL.
  - Aanmaken BAG-gebruiker en database: via pgAdmin III of via de commandline (niet beschreven).
- Python: voer de installer uit. Python 2.7.11 wordt helaas met een verouderde versie van Pip meegeleverd. Deze dient geüpgrade te worden naar versie 8. Dit is nodig voor het installeren van de Psycopyg-wheel. Commando:

```
python -m pip install -U pip
```

Je kunt ook pip rechtstreeks aanroepen. Voeg dan de Python scripts-directory eerst toe aan de PATH-variabele.

- Python dependencies:

```
python -m pip install <wheel>.whl
```

- GDAL: voer de installer van QGIS uit. Natuurlijk is niet altijd QGIS nodig, zeker op een server-omgeving. Op een desktop is het wel aan te bevelen, zodat je gelijk het resultaat in de database kunt controleren. Op een server kun je de OSGeo4W-installer gebruiken. Dit is niet getest met NLEExtract.

Zie *Instellingen* voor de configuratie en het gebruik van BAG-Extract.

### 3.1.8 Installatie (Mac OSX)

Voor Mac OSX zijn meerdere mogelijkheden. Hieronder wordt uitgegaan van MacPorts <http://www.macports.org>, een Unix package manager waarmee je gemakkelijk tools en bibliotheken en hun afhankelijkheden kunt installeren. MacPorts is sowieso aan te bevelen als je meerdere Unix/Linux tools gaat gebruiken. Python is al aanwezig op de Mac en is bruikbaar, de versie van Python kan afhankelijk zijn van je OSX versie. Probeer te vermijden om Python te installeren tenzij je precies weet wat je doet. Ook het *easy\_install* Python programma zou al aanwezig moeten zijn. Al het onderstaande doe je in de Terminal.

Onder de manier die Just, een van de NLEExtract ontwikkelaars gebruikt. (NLEExtract werkt dus op de Mac!).

- Python, 2.6.1 of hoger, liefst 2.7+. 2.6.1 Mac-versie werkt.
- Python package “argparse” installeren (alleen nodig voor Python < 2.7)

```
sudo easy_install argparse
```

- libxml2 en libxslt: via MacPorts:

```
sudo port install libxml2
sudo port install libxslt
```

- lxml

```
sudo easy_install lxml
```

- GDAL: KyngChaos (indien MacPorts GDAL-versie < 1.8.1 is) : <http://www.kyngchaos.com/software/index> Download en install *GDAL Complete*.
- GDAL-Python bindings (zijn mogelijk al via GDAL beschikbaar?)
- Postgres client psycopg2

```
sudo python easy_install psycopg2
```

### 3.1.9 Commando

- direct via python “python src/bagextract.py”
- of (Unix, Linux, Mac) via shell script: “bin/BAG-Extract.sh”
- Windows: voorlopig alleen via “python src/bagextract.py”

Alle commando's werken onafhankelijk van de plek (directory) waar ze aangeroepen worden

### 3.1.10 Instellingen

- **extract.conf** Configuratiebestand dat nodig is bij het uitvoeren van de programma's. Dit bestand bevat de volgende instellingen: - database naam van de Postgres database - schema [optioneel] schemanaam of schema search path waar de tabellen worden aangemaakt (default “public”) - host host waar de Postgres database draait - user user voor toegang tot de Postgres database - password password van de user voor toegang tot de Postgres database

Deze settings kunnen via commandline opties of via `-f <mijn conf file>` overuled worden, bijv. `bagextract.py -H localhost -d bag -U postgres -W postgres -c bagextract.py -f mijn.conf -c`

### 3.1.11 Voorbeelden

0. Help en opties:

```
python src/bagextract.py -h of bin/BAG-Extract.sh -h
```

Alle commando's kunnen via Python of shell .sh script uitgevoerd vanaf elke directory.

1. Initialiseer de database en vul/verrijk met referentie-koppeldata (gemeenten/provincies) (-c):

```
python bagextract.py -c
of
bag-extract.sh -c

# -c vraagt gebruiker interactief voor bevestiging. Met -j (ja-optie) is er geen
↳prompt. Handig voor batch-situaties
python bagextract.py -cj
of
bag-extract.sh -cj
```

2. Importeer een extract in de database (-e):



```
python bagextract.py -e 9999STA01052011-000002.xml
python bagextract.py -e 9999STA01052011.zip

-e werkt op directory, file of .zip inclusief mutatie-bestanden
```

3. **Verrijken: genereren gemeente + provincie tabellen met geometrie uit woonplaatsen aggregeren** NB  
Doe altijd eerst stappen 1-2 anders blijft de tabel “gemeente” leeg. !

```
python bagextract.py -v -q ../db/script/gemeente-provincie-tabel.sql
```

Met de -q (query) optie kan elk SQL bestand worden uitgevoerd

4. Verrijken: aanmaken tabel met volledige “ACN-achtige” adressen uit BAG + gemeente + provincie tabellen (kan lang duren op gehele BAG, lijkt sneller te gaan via “psql” Postgres commando). NB Doe altijd eerst stappen 1-3!

```
psql -d bag < ../db/script/adres-tabel.sql
```

Gebruik het psql commando “set search\_path to <your schema>,public; ” als je de adres-tabel in een expliciet Postgres schema wilt. Bijv

```
# set search_path to bag,public;
# \i /opt/nlextract/git/bag/db/script/adres-tabel.sql
```

5. **Geocoding** [zie tabellen en functies onder db/script/geocode] De BAG is niet standaard geschikt om geocoding op uit te voeren. Daartoe dienen eerst afgeleide tabellen te worden aangemaakt en hulp functies voor met name “reverse geocoding” (vind adres voor x,y coördinaten).

### 3.1.12 Issues

Het is mogelijk de hele BAG .zip levering in te lezen vanuit de “hoofd” zip, maar dit kan soms geheugen-problemen opleveren. De voorlopige oplossing is om de hoofdzip uit te pakken in een enkele directory en dan de (7) individuele BAG .zip files te extraheren.

Het (geometrisch) aggregeren van woonplaatsen naar gemeenten en vervolgens naar provincies kent een probleem waarbij uit PostGIS de volgende melding komt: “NOTICE: TopologyException: found non-noded intersection between LINESTRING (...) at ... ERROR: GEOS union() threw an error!”. Dit is mogelijk een bug in “libgeos” (GEOS) een library gebruikt door PostGIS. Dit probleem trad op in GEOS v3.2.2 maar niet in versie 3.3.1.

Het script db/script/adres-tabel.sql vergt 20 minuten tot enkele uren. Vaak afhankelijk van je machine maar vooral ook je PostgreSQL instellingen. Beste is om deze met standaard PSQL uit te voeren.

Het resultaat van het genereren van gemeenten en provincies uit woonplaats geometrieën is nog “rommelig”: veel kleine polygonen. Die willen we nog uitfilteren.

Bij foutmeldingen als *COPY failed for table “nummeraanduiding”: ERROR: value too long for type character varying(20)* heeft je “bag” database niet de **UTF8 character encoding** (zie boven). Check bij aanmaken, vooral op Windows, of je DB de character-encoding UTF8 heeft. Is later aan te passen. Zie ook [dit issue](#).

Zie <http://docs.kademo.nl/project/bagextract.html> voor een installatie voorbeeld.



Hieronder staat de handleiding voor het gebruik van de tools om de BRT TOP10NL te extraheren. Deze tools heten kortweg TOP10-Extract of soms NLExtract-TOP10.

NB: als je alleen interesse hebt om een PostGIS versie van de laatste TOP10NL te hebben, kun je deze ook (betaald) downloaden als PostGIS dumpfile via de link <https://geotoko.nl/>. De dumpfile (.dump bestand) kun je direct inlezen in PostGIS, bijv met PGAdminIII. Dan hoef je alle zaken hieronder niet uit te voeren :-).

Om gespecialiseerde extracties bijv naar andere databases zoals Oracle te doen, neem contact op met het NLExtract-team, zie “Ondersteuning”: <http://www.nlextract.nl/issues>.

## 4.1 Handleiding TOP10-Extract

### 4.1.1 Algemeen

TOP10NL is onderdeel van de Kadaster Basisregistratie Topografie (BRT). Vind algemene info over TOP10NL op <http://www.kadaster.nl/web/artikel/productartikel/TOP10NL.htm>.

TOP10-Extract bevat de tools om de TOP10NL GML-bronbestanden, zoals geleverd door het Kadaster (bijv via PDOK), om te zetten naar hanteerbare formaten, zoals PostGIS. Tevens bevat TOP10-Extract visualisatie-bestanden (onder de map *style/* ) voor QGIS en SLDs om kaarten te maken. (NB deze zijn nu nog gebaseerd op TOP10NL 1.0!).

TOP10NL (v1.2) wordt geleverd door het Kadaster als een .zip file van plm 2 GB. Voor de landsdekkende versies zijn er 2 soorten .zip-bestanden, een op basis van kaartbladen en een .zip file op basis van “GML FileChunks” waarbij de totale GML is opgedeeld in files van 300 MB. Zie [Bestandswijzer\\_GML\\_TOP10NL\\_2012.pdf](#) voor de kaartbladindeling.

Er zijn 13 typen TOP10NL objecten. Zie voor de beschrijving van de structuur en verdere bijzonderheden voor de GML bestandsindeling in [BRT\\_Catalogus\\_Productspecificaties.pdf](#) (nog gebaseerd op versie 1.1.1).

### 4.1.2 TOP10NL downloaden

TOP10NL brondata in GML kun je via [PDOK TOP10NL Downloads](#) downloaden.

Er zijn twee download varianten: de “GML File Chunks” en “50D Kaartbladen”. De eerste is de totale verzameling opgesplitst in 300MB GML Files, de tweede bevat de GML bestanden per kaartblad. Download de Kaartbladen alleen als je bijv. een enkel gebied wilt inlezen of om te testen. Beide ZIP-bestanden zijn ca. 2 GB groot. Het is ook mogelijk om de TOP10NL via PDOK-services te downloaden, bijv. via WFS. Het inlezen van deze gegevens via NLExtract wordt niet ondersteund.

Als je heel Nederland wilt inlezen, kun je het beste de “GML File Chunks” gebruiken. De directe link is [http://geodata.nationaalgeoregister.nl/TOP10nlv2/extract/chunkdata/TOP10nl\\_gml\\_filechunks.zip?formaat=gml](http://geodata.nationaalgeoregister.nl/TOP10nlv2/extract/chunkdata/TOP10nl_gml_filechunks.zip?formaat=gml).

Voor de kaartbladen is dat: <http://geodata.nationaalgeoregister.nl/TOP10nlv2/extract/kaartbladtotaal/TOP10nl.zip?formaat=gml>.

*NB: het is heel belangrijk om de laatste versie van TOP10NL te gebruiken. Dit is versie 1.2. Deze wordt geleverd met ingang van november 2015. Alleen deze versie wordt ondersteund door de huidige versie TOP10-Extract. Met ingang van deze datum is ook het Kadaster volledig overgeschakeld. De oude versies van TOP10NL worden niet meer ondersteund. Mocht je toch de oude versie willen inlezen, gebruik dan een oude release van NLExtract.*

### 4.1.3 TOP10-Extract downloaden

Vind altijd de laatste versie op: <https://github.com/nlextract/NLExtract/releases>. De nieuwste versie staat bovenaan: kies de “real-release” nlextract zip.

Omdat NLExtract voortdurend in ontwikkeling is, kun je ook de actuele broncode, een *snapshot*, downloaden en op dezelfde manier gebruiken als een versie:

- snapshot via git: `git clone http://github.com/opeengeogroep/NLExtract.git`
- snapshot als .zip: <https://github.com/nlextract/NLExtract/archive/master.zip>

### 4.1.4 Ontwerp

In eerste instantie wordt de GML geconverteerd en geladen naar PostGIS. Dit gebeurt met de GDAL/OGR tool ogr2ogr. Echter, er zijn 2 belangrijke zaken die dit lastig maken:

- meerdere geometrieën per object, bijv een Waterdeel GML element kan een punt, een lijn of een vlak bevatten
- meerdere voorkomens van een attribuut (attribute multiplicity), bijv. een Wegdeel GML element kan meerdere element-attributen genaamd “nWegnummer” bevatten

Om het eerste probleem op te lossen worden middels een XSLT script (etl/xsl/TOP10-split\_v1\_2.xsl) de GML-elementen uitgesplitst naar geometrie, zodat ieder element een enkele geometrie bevat. Bijvoorbeeld Wegdeel kent maar liefst 5 geometrie-attributen. Dit wordt opgesplitst naar Wegdeel\_Lijn, Wegdeel\_Vlak etc. Een nieuw GML-bestand wordt hiermee opgebouwd. Vervolgens wordt via ogr2ogr dit uitgesplitste GML bestand in PostGIS geladen.

Met ingang van de november-release van de BRT (2015R11) wordt alleen het Stetl-framework ondersteund voor TOP10-Extract. Zie verder *Stetl-framework* voor de werking van TOP10-Extract.

Hieronder staat de handleiding voor het gebruik van de tools om BGT te extraheren. Deze tools heten kortweg `Bgt-extract` of soms `NLExtract-BGT`.

NB: als je alleen interesse hebt om een PostGIS versie van de laatste BGT te hebben, kun je deze ook (betaald) downloaden als PostGIS dumpfile via de link <https://geotoko.nl/>. De dump file (.dump bestand) kun je direct inlezen in PostGIS, bijv met `PGAdminIII`. Dan hoeft je alle zaken hieronder niet uit te voeren :-).

Om gespecialiseerde extracties bijv naar andere databases zoals Oracle te doen, neem contact op met het NLExtract-team, zie “Ondersteuning”: <http://www.nlextract.nl/issues>.

## 5.1 Handleiding BGT-extract

### 5.1.1 Algemeen

`Bgt-extract` is onderdeel van de `NLExtract` tools voor het inlezen en verrijken van de Basisregistratie Grootchalige Topografie (BGT). Deze open dataset bestaat uit een aantal GML-bestanden en wordt (voorlopig) ingelezen in een PostgreSQL/PostGIS database.

Er zijn 41 typen BGT-objecten (featureklassen). Iedere featureklasse heeft een groot aantal attributen en meestal ook meerdere geometrie-attributen. De BGT is gebaseerd op CityGML. Hierdoor biedt het datamodel ruimte voor toekomstige uitbreiding, bijv. de toevoeging van 3D-geometrie. Zie voor de beschrijving van de structuur en verdere bijzonderheden voor de GML-bestandsindeling de documentatie op [BGTweb](#).

### 5.1.2 BGT downloaden

De brondata van de BGT in GML kun je via [PDOK Download Basisregistratie Grootchalige Topografie](#) downloaden. Voor `NLExtract` zijn reeds downloadscripts gemaakt, voor zowel Linux als Windows.

De BGT wordt via PDOK geleverd in ZIP-bestanden. Het is mogelijk om zowel een landelijk bestand als deelbestanden te downloaden. De laatste zijn opgedeeld in een grid van 2x2, 4x4, 8x8, 16x16, 32x32 en 64x64 km. Omdat

de PDOK-download wel eens hapert, wordt aanbevolen om de 64x64 km-extracten te downloaden via een download-script van NLExtract. Het is mogelijk om de BGT via PDOK-services te downloaden, bijv. via WFS. Het inlezen van deze gegevens via NLExtract wordt niet ondersteund.

De BGT is beschikbaar in meerdere varianten:

- Met of zonder plaatsbepalingspunten: deze worden door NLExtract ondersteund. Vanwege de omvang van de data wordt sterk aanbevolen om de variant zonder plaatsbepalingspunten te gebruiken.
- GML of GML Light: alleen GML wordt ondersteund.

De BGT bestaat uit zowel een verplicht (BGT) als optioneel (IMGeo) deel. Dit is het zogenaamde “plus”-deel, welke ook door NLExtract wordt ondersteund. Dit laatste bestaat uit nieuwe featureklassen en extra attributen bij bestaande featureklassen.

Momenteel (juni 2016) zijn veel bronhouders bezig om de BGT te vullen. De omvang van de BGT ZIP-bestanden van heel Nederland, zonder plaatsbepalingspunten, was op 16 juni jl. 5,3 GB. Medio januari was het slechts 1,9 GB. De verwachting is dat het ZIP-bestand van heel Nederland uiteindelijk ca. 15 GB zal zijn.

Als je heel Nederland wilt inlezen, kun je het beste het bestand exclusief plaatsbepalingspunten, GML-variant, gebruiken. Zie PDOK voor de directe link. Deze verandert namelijk iedere dag.

### 5.1.3 BGT-extract downloaden

Vind altijd de laatste versie op: <https://github.com/nlextract/NLExtract/releases>. De nieuwste versie staat bovenaan: kies de “real-release” nlextract zip.

Omdat NLExtract voortdurend in ontwikkeling is, kun je ook de actuele broncode, een *snapshot*, downloaden en op dezelfde manier gebruiken als een versie:

- snapshot via git: `git clone http://github.com/opengeogroep/NLExtract.git`
- snapshot als .zip: <https://github.com/nlextract/NLExtract/archive/master.zip>

### 5.1.4 Ontwerp

In eerste instantie wordt de GML geconverteerd en geladen naar PostGIS. Dit gebeurt met de GDAL/OGR tool `ogr2ogr`. Echter, het feit dat er meerdere geometrieën per object kunnen voorkomen, maakt dit lastiger. De meeste objecten bevatten zowel een LoD 0-geometrie als een 2D-geometrie. Sommige objecten bevatten tevens een kruinlijn-geometrie. De BGT bevat geen attributen met multipliciteit, d.w.z. meerdere voorkomens van een attribuut.

Om het eerste probleem op te lossen worden middels een XSLT script (`etl/xsl/imgeo-split_v2.1.1.xsl`) de GML-elementen uitgesplitst naar geometrie, zodat ieder element een enkele geometrie bevat. Bijvoorbeeld het `TrafficArea`-element (Wegdeel) wordt opgesplitst naar `TrafficArea_2D` en `TrafficArea_kruinlijn`. Vervolgens wordt via `ogr2ogr` dit uitgesplitste GML bestand in PostGIS geladen. Hierbij vindt ook de uiteindelijke vertaling van de in CityGML gedefinieerde objecten, zoals `TrafficArea`, naar het Nederlands plaats, zoals `Wegdeel`.

Zie verder *Stetl-framework* voor de werking van `Bgt-extract`.

Hieronder staat de handleiding voor het gebruik van de tools om BRK te extraheren. Deze tools heten kortweg `Brk-extract` of soms `NLExtract-BRK`.

NB: als je alleen interesse hebt om een PostGIS versie van de laatste BRK te hebben, kun je deze ook (betaald) downloaden als PostGIS dumpfile via de link <https://geotoko.nl/>. De dump file (.dump bestand) kun je direct inlezen in PostGIS, bijv met `PGAdminIII`. Dan hoeft je alle zaken hieronder niet uit te voeren :-).

Om gespecialiseerde extracties bijv naar andere databases zoals Oracle te doen, neem contact op met het NLExtract-team, zie “Ondersteuning”: <http://www.nlextract.nl/issues>.

## 6.1 Handleiding BRK-extract

### 6.1.1 Algemeen

`Brk-extract` is onderdeel van de `NLExtract` tools voor het inlezen en verrijken van de Basisregistratie Kadaster (BRK). Deze open dataset bestaat uit een aantal GML-bestanden en wordt (voorlopig) ingelezen in een PostgreSQL/PostGIS database.

Er zijn vier typen BRK-objecten (featureklassen). Iedere featureklasse heeft een groot aantal attributen. Drie hiervan, Annotatie, Bebouwing en Kadastrale Grens, hebben één geometrie, maar Perceel heeft twee geometrieën (perceelgrens en labelpunt).

### 6.1.2 BRK downloaden

De brondata van de BRK in GML kun je via [PDOK Downloads Kadastrale Kaart](#) downloaden. Voor `NLExtract` zijn reeds downloadscripts gemaakt, voor zowel Linux als Windows.

De BRK wordt via PDOK geleverd in ZIP-bestanden. Deze worden per provincie beschikbaar gesteld. De bestanden bevatten geen overlappende gegevens. Ieder ZIP-bestand bevat vier GML-bestanden: één bestand per featureklasse.

Het is mogelijk om de kadastrale kaart via PDOK-services te downloaden, bijv. via WFS. Het inlezen van deze gegevens via NLExtract wordt niet ondersteund. De ZIP-bestanden zijn samen ca. 4 GB groot.

### 6.1.3 Brk-extract downloaden

Vind altijd de laatste versie op: <https://github.com/nlextract/NLExtract/releases>. De nieuwste versie staat bovenaan: kies de “real-release” nlextract zip.

Omdat NLExtract voortdurend in ontwikkeling is, kun je ook de actuele broncode, een *snapshot*, downloaden en op dezelfde manier gebruiken als een versie:

- snapshot via git: `git clone http://github.com/opengeogroep/NLExtract.git`
- snapshot als .zip: <https://github.com/nlextract/NLExtract/archive/master.zip>

### 6.1.4 Ontwerp

In eerste instantie wordt de GML geconverteerd en geladen naar PostGIS. Dit gebeurt met de GDAL/OGR tool `ogr2ogr`. GDAL/OGR versie 1.11 is de minimale versie. Hiermee kunnen ook de meerdere geometrieën van een perceel ingelezen worden. Het GFS-bestand, een stuurbestand voor het inlezen van GML-data via `ogr2ogr`, is hierop aangepast.

Zie verder *Stetl-framework* voor de werking van Brk-extract.



NLEextract gaat steeds meer gebruik maken van de ETL framework Stetl, zie <http://stetl.org>. Hierdoor hoeft niet meer per dataset een apart programma worden gemaakt. De volgende extract-tools maken gebruik van het Stetl-framework:

- *TOP10-Extract*
- *BGT-extract*
- *BRK-extract*

Stetl maakt i.h.a. gebruik van Python voor alle scripts. De Python-scripts voor de extract-tools die gebruik maken van het Stetl-framework roepen *native* tools aan:

- XML parsing via `libxml2`;
- XSLT processing via `libxslt`;
- GDAL/OGR via `ogr2ogr`.

De reden hiervoor is vooral de snelheid. Deze native libraries zijn beschikbaar binnen Python d.m.v. zogenaamde *bindings*.

## 7.1 Afhankelijkheden

De volgende software dient aanwezig te zijn om gebruik te maken van het Stetl-framework:

- Python 2.6 of hoger (niet Python 3!). Let op dat de ontwikkeling van Python 2.6 is stopgezet!
  - Op *Windows* is Python 2.7 de minimale versie. Deze versie wordt ook geïnstalleerd wanneer je OSGeo4W met de juiste opties, of QGIS hebt geïnstalleerd.
- GDAL/OGR v1.11 of hoger (voor `ogr2ogr`): <http://www.gdal.org>.
- PostGIS: PostgreSQL 8.x of 9.x database server met PostGIS 1.x of 2.x: <http://postgis.refractory.net>.
  - PostgreSQL 9.x met PostGIS 2.x wordt aanbevolen.

- Pip en Setuptools (o.a. voor package `pkg_resources`): <https://pip.pypa.io/en/latest/installing/#install-pip>. Dit is alleen nodig indien je Python-versie ouder dan 2.7.9 is.
- `argparse` voor argument parsing. Alleen nodig bij Python 2.6.
- `psycopg2` voor het maken van database connecties met PostgreSQL vanuit Python.
- `lxml` voor razendsnelle native XML parsing: <http://lxml.de/installation.html>.
- `libxml2` en `libxslt` bibliotheken (worden door `lxml` gebruikt).

NB: GDAL/OGR Python bindings zijn (voorlopig) *niet* nodig.

## 7.2 Installatie

Stetl werkt op de drie voornaamste platformen: Windows, Mac OSX, Linux. De bovengenoemde afhankelijkheden hebben ieder hun eigen handleiding voor installatie op desbetreffend platform. Raadpleeg deze als eerste. Hieronder volgt een aantal tips en bijzonderheden per platform.

### 7.2.1 Linux

Gebruik onder Ubuntu altijd *Ubuntu GIS*: <https://wiki.ubuntu.com/UbuntuGIS> om de laatste versies van veel packages, met name GDAL en PostGIS 1.x te verkrijgen!

- optioneel: Python package afhankelijkheden installeren bijv.

```
apt-get of yum install python-setuptools (voor easy_install commando)
apt-get of yum install python-devel (tbv psycopg2 bibliotheek)
apt-get of yum install postgresql-devel (tbv psycopg2 bibliotheek)
```

- `lxml`

```
apt-get of yum install libxml2
apt-get of yum install libxslt1.1
apt-get of yum install python-lxml
```

- GDAL

```
apt-get of yum install gdal-bin
```

- Python package “argparse” (alleen voor Python < 2.7)

```
sudo easy_install argparse
```

- NB: als je een proxy gebruikt via `http_proxy`, doe dan `easy_install -E` (exporteer huidige environment)

### 7.2.2 Windows

De Python scripts zijn ontwikkeld en getest op Windows 7 met Python 2.7.

Het is gebleken dat het lastig is om NLEextract goed op Windows aan de praat te krijgen. Een belangrijke reden is de OSGeo4W-installer, waar o.a. ook QGIS mee wordt geïnstalleerd. OSGeo4W levert een eigen Python-versie mee. Deze versie, 2.7.5, die o.a. met QGIS Essen (2.14) wordt geïnstalleerd, loopt een stuk achter bij de laatste Python 2.7-release, 2.7.11 (status juni 2016). De Python-executable bevindt zich zelfs in dezelfde directory als `ogr2ogr`, wat de zaak alleen gecompliceerder maakt.

Met onderstaande instructies is het mogelijk om NLEExtract werkend te krijgen op Windows. Hierbij maakt het niet uit of je de OSGeo4W-versie van Python gebruikt of je eigen Python-versie. In het laatste geval moet je wel zelf op de een of andere manier ogr2ogr op je machine krijgen en de locatie hiervan in de PATH-variabele zetten. Vanwege de vele mogelijkheden zijn niet alle situaties getest. Open s.v.p. een issue-report in Github bij problemen of meld het op de mailinglijst. De kans is het grootst dat er problemen zijn met de PATH-variabele of de PYTHONPATH-variabele.

Open een command prompt. Indien Python, QGIS of de OSGeo4W-software in C:\Program Files of C:\Program Files (x86) staat, dien je de command prompt als Administrator te openen. Indien je OSGeo4W/QGIS hebt, kun je eventueel ook de OSGeo4W-shell gebruiken.

In de instructies wordt gebruik gemaakt van Python wheels, ofwel WHL-bestanden. Voor Windows is een groot aantal van deze bestanden te vinden op de site van [Christian Gohlke](#). Kies de Python 2.7-versie en kies de 32- of 64-bits versie. Dit is afhankelijk van de Python-versie of OSGeo4W/QGIS-versie die je hebt.

- Installatie Pip, Setuptools en pkg\_resources: dit is alleen nodig indien je Python-versie ouder is dan 2.7.9, dus ook als je de OSGeo4W-versie gebruikt. Dit kan via het script `get-pip.py`. Zie <https://pip.pypa.io/en/latest/installing/#install-pip> voor verdere instructies. Hierbij krijg je tevens ondersteuning voor de installatie van Python-wheels (WHL-bestanden). Dit is nodig voor de vervolgstappen. Zorg ervoor dat de Scripts-directory van Python, waar `pip.exe` staat, in het pad is in het commando shell waarmee je de installaties uitvoert. Een WHL-bestand kan als volgt met Pip geïnstalleerd worden:

```
python -m pip install <package>.whl
```

- Installatie lxml: download en installeer het WHL-bestand. Je hoeft niet apart libxml2 of libxslt te installeren.
- Installatie psychopg (niet bij OSGeo4W): download en installeer het WHL-bestand.

Let bij Windows ook op het volgende: wanneer je op de command line met PostgreSQL wilt connecten, gebruik `chcp 1252` om de code page van de console bij te werken naar ANSI. Je krijgt anders een waarschuwing wanneer je in PostgreSQL inlogt. Dit komt omdat de code page standaard 437 is (extended ASCII) i.p.v. 1252 (ANSI).

### 7.2.3 Mac OSX

- Python, 2.6.1 of hoger, liefst 2.7+,
- Python package “argparse” (alleen voor Python < 2.7)

```
sudo easy_install argparse
```

- libxml2 en libxslt: via MacPorts: <http://www.macports.org/>
- lxml

```
sudo easy_install lxml
```

- GDAL: KyngChaos (MacPorts GDAL-versie is vaak outdated) : <http://www.kyngchaos.com/software/index> Download en install *GDAL Complete*. Om te zorgen dat de GDAL commando's, met name *ogr2ogr* kunnen worden gevonden, kun je het volgende wijzigen in */etc/profile*, die standaard Shell settings in het Terminal window bepaalt:

```
export PATH=/Library/Frameworks/GDAL.framework/Versions/Current/Programs:$PATH
```

## 7.3 Stetl uitvoeren

Uitvoeren: `./etl-<dataset>.sh` of `./etl-<dataset>.cmd` (Windows). Ga hiervoor met een prompt in de etl-directory staan van de desbetreffende dataset, dus in `<dataset>/etl`.

Opties zetten: maak hiertoe een eigen lokaal bestand in de options-directory, met de naam `<hostnaam>.args`. Dit kan door het kopiëren van het bestand `default.args`. Let op dat alle opties in je eigen optie-bestand gezet moet worden indien je geen gebruik maakt van `default.args`. Er is geen fallback-mogelijkheid, zoals voorheen wel het geval was.

De Stetl-configuratie in `etl-<dataset>-<versie>.cfg` hoeft niet te worden gewijzigd, alleen indien bijv. een andere output gewenst is.

### 7.3.1 Uitleg opties

De volgende opties worden door Stetl gebruikt bij het laden van data via NLEextract. Onder water worden ze gecombineerd tot het commando waarmee `ogr2ogr` wordt aangeroepen. De opties worden ingesteld door het meegeven van het juiste opties-bestand (args-bestand) aan het ETL-commando. Zie het SH- of het CMD-script voor meer informatie.

**input\_dir** Directory met inputbestanden. NB: ook op Windows kunnen forward slashes in paden worden gebruikt.

**zip\_files\_pattern** Bestandenfilter volgens Python `glob.glob` patronen.

**filename\_match** Filter op bestanden binnen de ZIP-bestanden. Meestal is `*.gml` voldoende. Kan gebruikt worden om bepaalde featuretypes uit te sluiten, indien de bestandsnaam hiervoor geschikt is.

**temp\_dir** Directory waar tijdelijke bestanden (bijv. opgesplitste GML-bestanden en kopieën GFS-bestanden) komen te staan.

**gfs\_template** Naam van het GFS template-bestand. Dit bevat de mapping naar de kolommen in PostgreSQL.

**host** Hostnaam van de server waarop de database staat.

**port** Poortnummer waarmee verbinding gemaakt kan worden met de database server.

**user** Gebruikersnaam van de PostgreSQL-gebruiker waarmee verbinding gemaakt moet worden.

**password** Wachtwoord van de PostgreSQL-gebruiker waarmee verbinding gemaakt moet worden.

**database** Naam van de database waarmee verbinding gemaakt moet worden.

**schema** Naam van het database schema die de datatabellen zal bevatten.

#### multi\_opts

**Wijze waarop omgegaan moet worden met multiattributen (ogr2ogr-opties). Varianten:**

- Eerstvoorkomende attribuutwaarde: `multi_opts=-splitlistfields -maxsubfields 1`
- Meerdere kolommen: `multi_opts=-splitlistfields`
- Stringlijst: `multi_opts=-fieldTypeToString StringList`
- Array (default): `multi_opts=`

**spatial\_extent** Definieert het in te lezen gebied. Formaat: `<minx> <miny> <maxx> <maxy>`. Wanneer dit leeggelaten wordt, wordt alle data ingelezen.

**max\_features** Aantal features (nog niet uitgesplitst) dat tegelijkertijd geladen zal worden. De waarde van 20000 wordt gebruikt, voldoet eigenlijk altijd. Hogere waarden kunnen op met name Windows tot geheugenproblemen leiden, maar dit heeft voor de verwerking geen voordelen.

---

## Downloads - Instructie

---

Om gebruik te maken van NLExtract data, gedownload via de [Geotoko.nl Website](#), is ondersteunende software nodig. Dit zijn de tools om met de data te kunnen werken. Hierbij is het uiteraard afhankelijk wat je wilt bereiken. In het algemeen kun je stellen dat de meeste gebruikers gewoon de data willen gebruiken. Dan volstaat een database dump en uiteraard de database software. Met behulp van afbeeldingen wordt dit proces stap voor stap uitgelegd voor Windows-gebruikers.

### 8.1 Instructie (Windows - GUI)

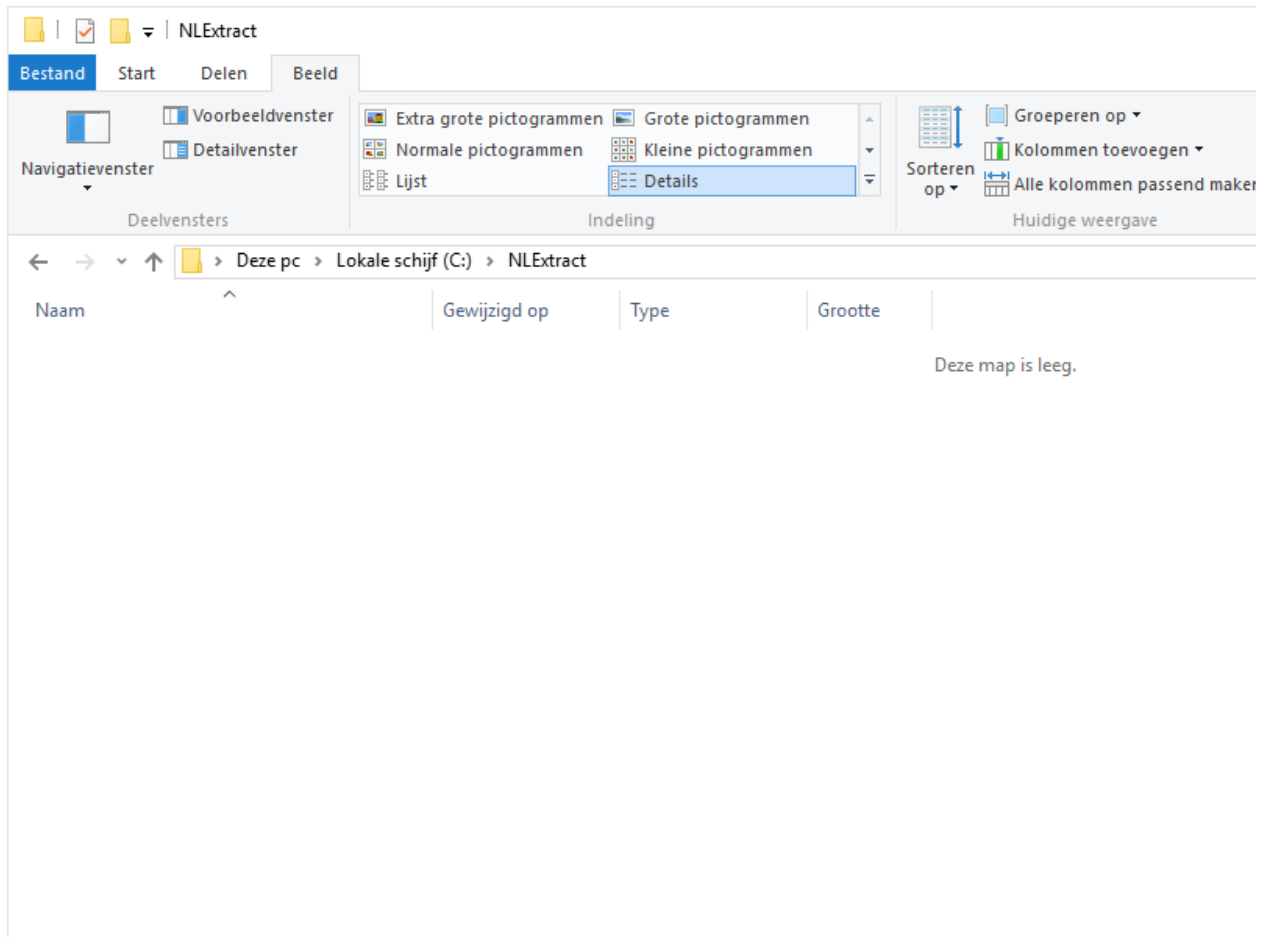
Deze beschrijving is een voorbeeldinstructie om de benodigde software te installeren en daarbij tevens een database dump terug te zetten (te restoren). Het voorbeeld hier beschreven is voor de BAG PostGIS download, maar zal hetzelfde zijn voor andere NLExtract downloads, zoals de BGT en de BRK. Met dank aan Geert Doornbos voor opstellen van deze instructie.

We beschrijven de volgende stappen:

- *Database Dump Downloaden*
- *PostgreSQL Installeren*
- *PostGIS Installeren*
- *PgAdmin III - BAG login toevoegen*
- *PgAdmin III - BAG database aanmaken*
- *PgAdmin III - BAG dump terugzetten*

## 8.1.1 Database Dump Downloaden

Maak een nieuwe map met bijvoorbeeld de naam 'NLEExtract'. Dit is gewoon een locatie om de dumps die gedownload worden te bewaren.



Ga naar de pagina (zoals beschreven in *NLEExtract Download Service*) waar de gewenste database dump gekozen kan worden. In deze instructie kiezen we de BAG database.

Index of /bag/postgis

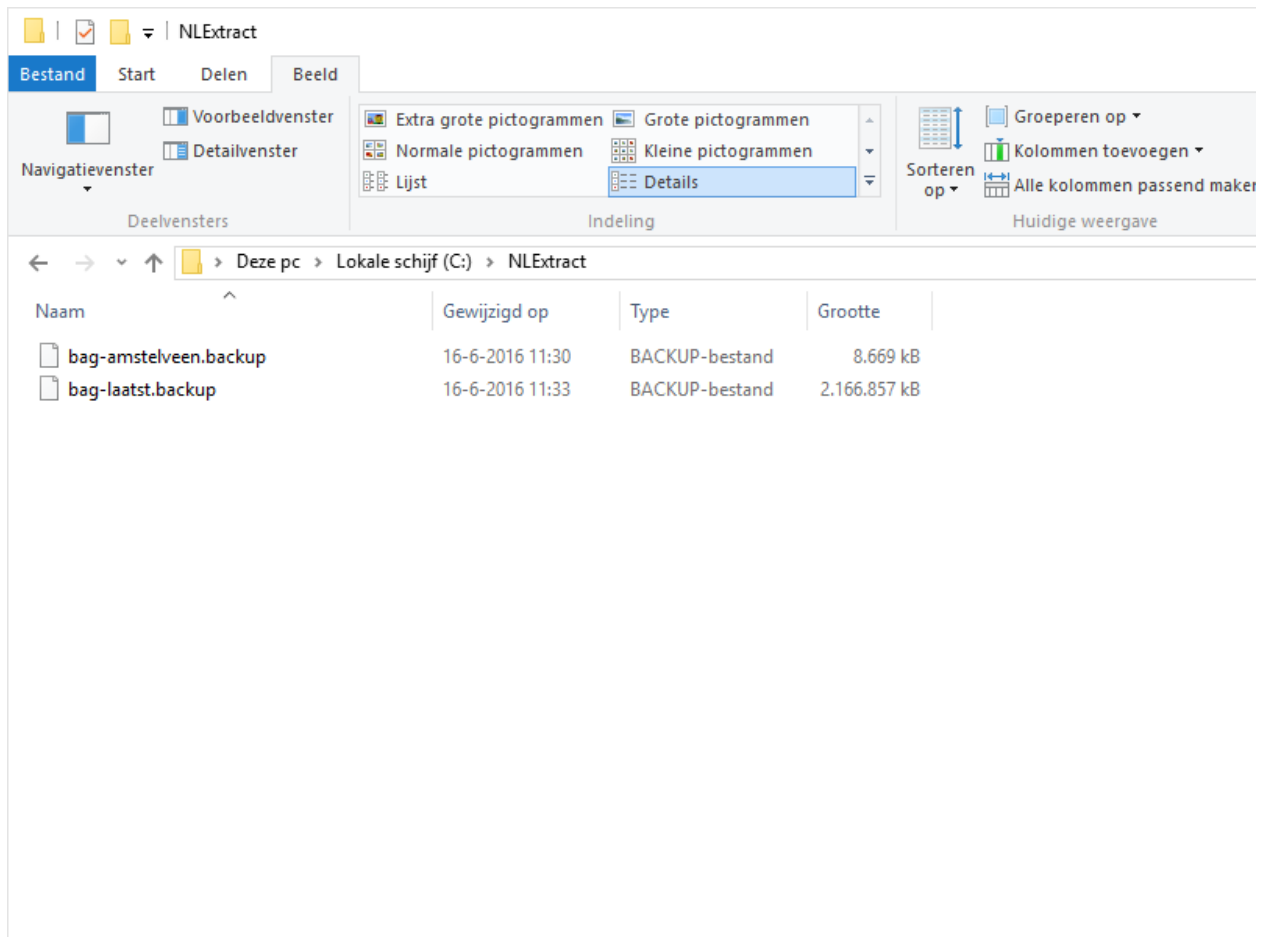
data.nlextract.nl/bag/postgis/

## Index of /bag/postgis

Name	Last modified	Size	Description
<a href="#">Parent Directory</a>		-	
<a href="#">README</a>	2016-03-05 11:53	2.3K	
<a href="#">bag-2016_01_11.backup</a>	2016-01-13 05:19	1.8G	
<a href="#">bag-2016_02_08.backup</a>	2016-02-13 20:52	1.9G	
<a href="#">bag-2016_03_09.backup</a>	2016-03-25 09:26	2.1G	
<a href="#">bag-2016_05_08.backup</a>	2016-05-29 01:14	1.9G	
<a href="#">bag-2016_06_08.backup</a>	2016-06-11 22:33	2.1G	
<a href="#">bag-amstelveen.backup</a>	2014-12-21 22:18	8.5M	
<a href="#">bag-laatst.backup</a>	2016-06-11 22:33	2.1G	
<a href="#">oud/</a>	2016-03-27 16:47	-	

Apache/2.4.7 Server at data.nlextract.nl Port 80

Kies 'link opslaan als' op 'bag-laatst.backup' via het contextmenu (rechter muisknop) en plaats het bestand in de zojuist gemaakte map 'NLEExtract'. Afhankelijk van de netwerkverbinding kan dit even duren, want het bestand is groter dan 2 GB.



Je kunt eventueel bag-amstelveen.backup downloaden om te testen met een kleiner bestand.

### 8.1.2 PostgreSQL Installeren

Ga naar de [PostgreSQL website](#) en klik door naar de [download](#) pagina van EnterpriseDB. We kiezen hier het installatieprogramma (32 of 64 bit - zie <Win>+<Break>) voor de meest recente en stabiele versie, in dit geval 9.5.3.



Index of /bag/postgis | Download PostgreSQL | EnterpriseDB

www.enterprisedb.com/products-services-training/pgdownload#windows

# Download PostgreSQL

*Please Note: Cookies should be enabled for the download process to function correctly*

Installer version **Version 9.6 Beta 1** [ This is a beta release for testing purposes, and should not be used in p

Linux x86-32 Linux x86-64 Win x86-32 Win x86-64 Mac OS

Installer version **Version 9.5.3** [Readme file for customers interested in using PL/Perl, PL/Python or PL/Tcl]

Linux x86-32 Linux x86-64 Win x86-32 Win x86-64 Mac OS

Installer version **Version 9.4.8** [Readme file for customers interested in using PL/Perl, PL/Python or PL/Tcl]

Linux x86-32 Linux x86-64 Win x86-32 Win x86-64 Mac OS

Installer version **Version 9.3.13**

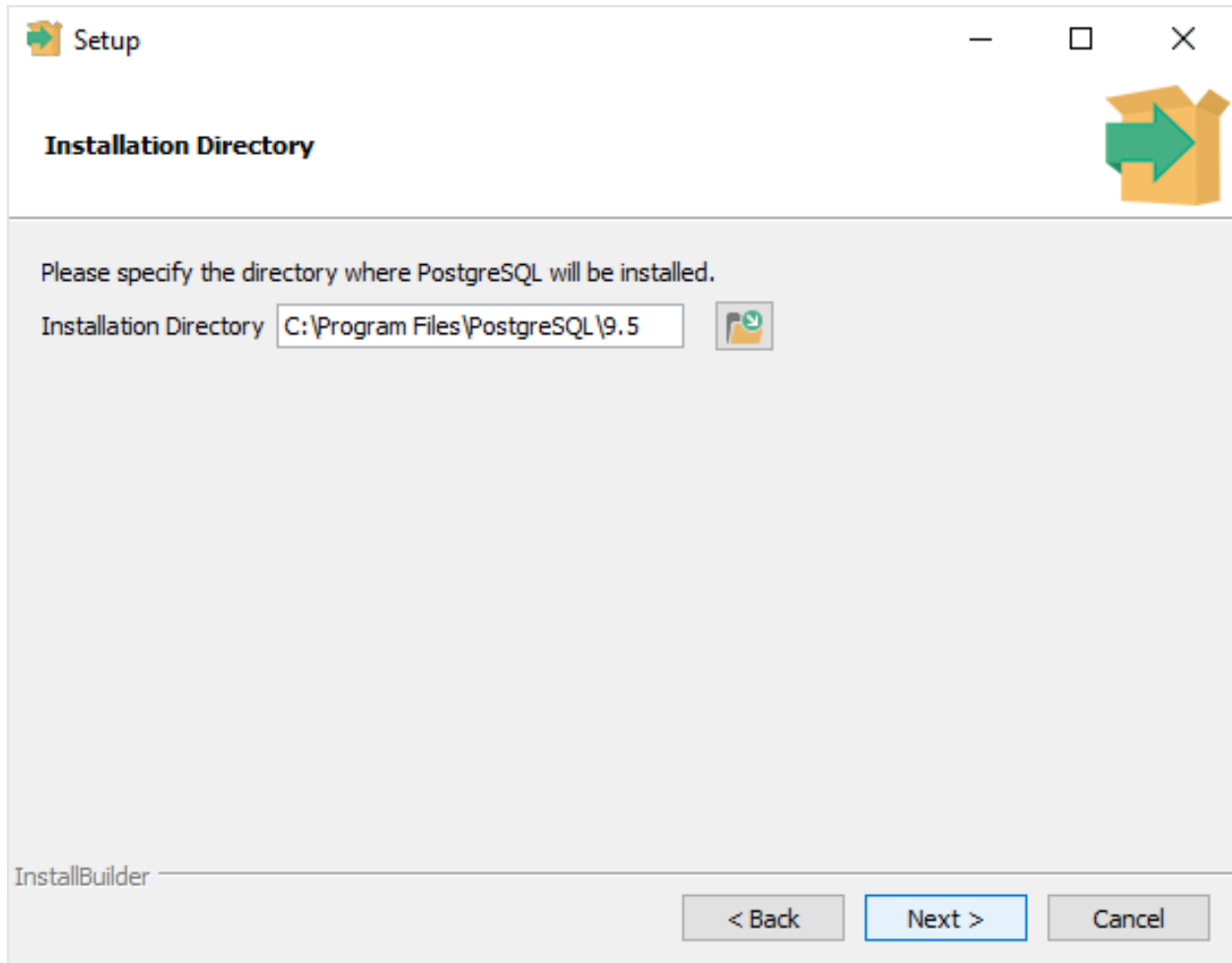
Linux x86-32 Linux x86-64 Win x86-32 Win x86-64 Mac OS

www.enterprisedb.com/postgresql-953-installers-win64?ls=Crossover&type=Crossover

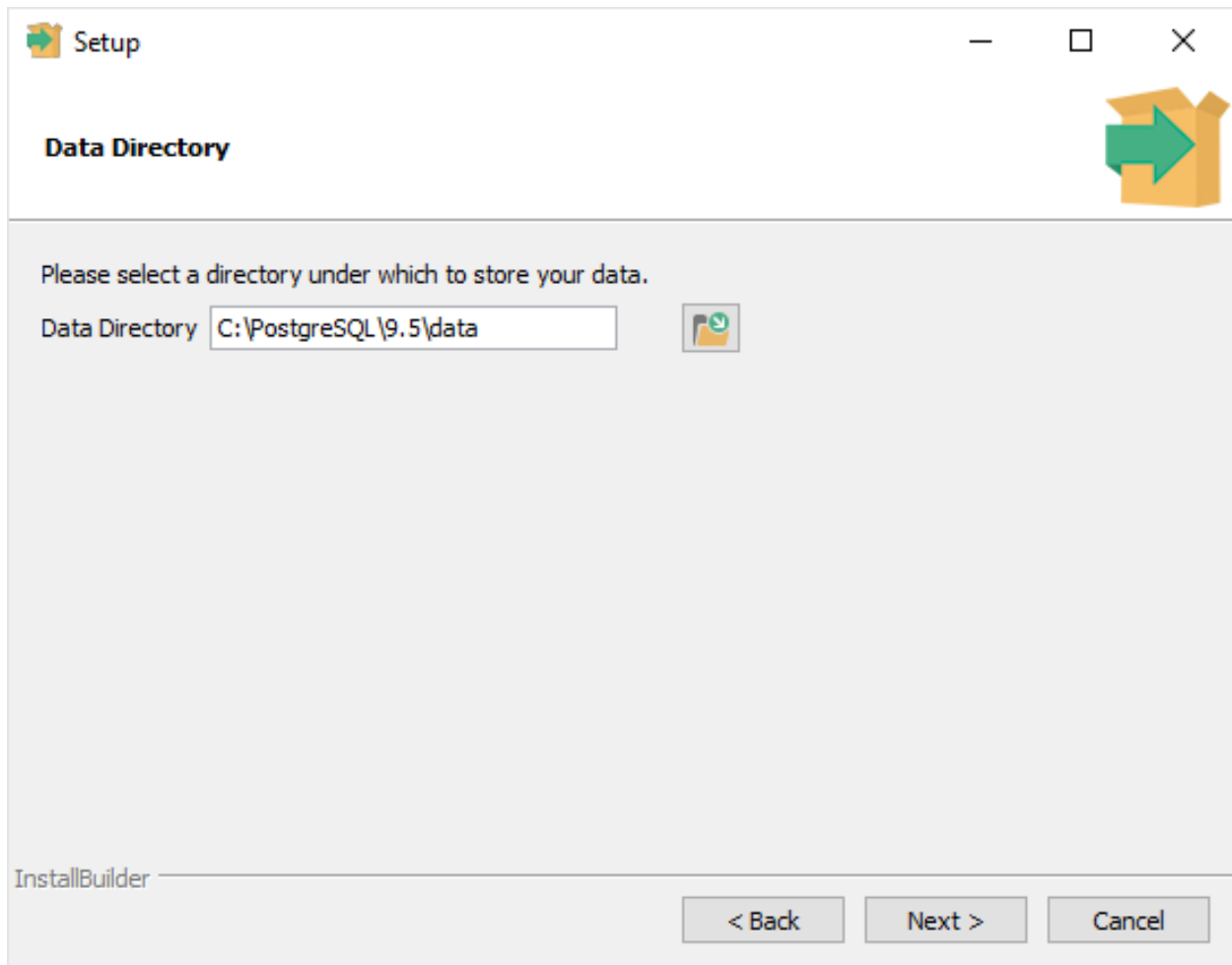
Voer het gedownloade bestand uit en de installatie begint.



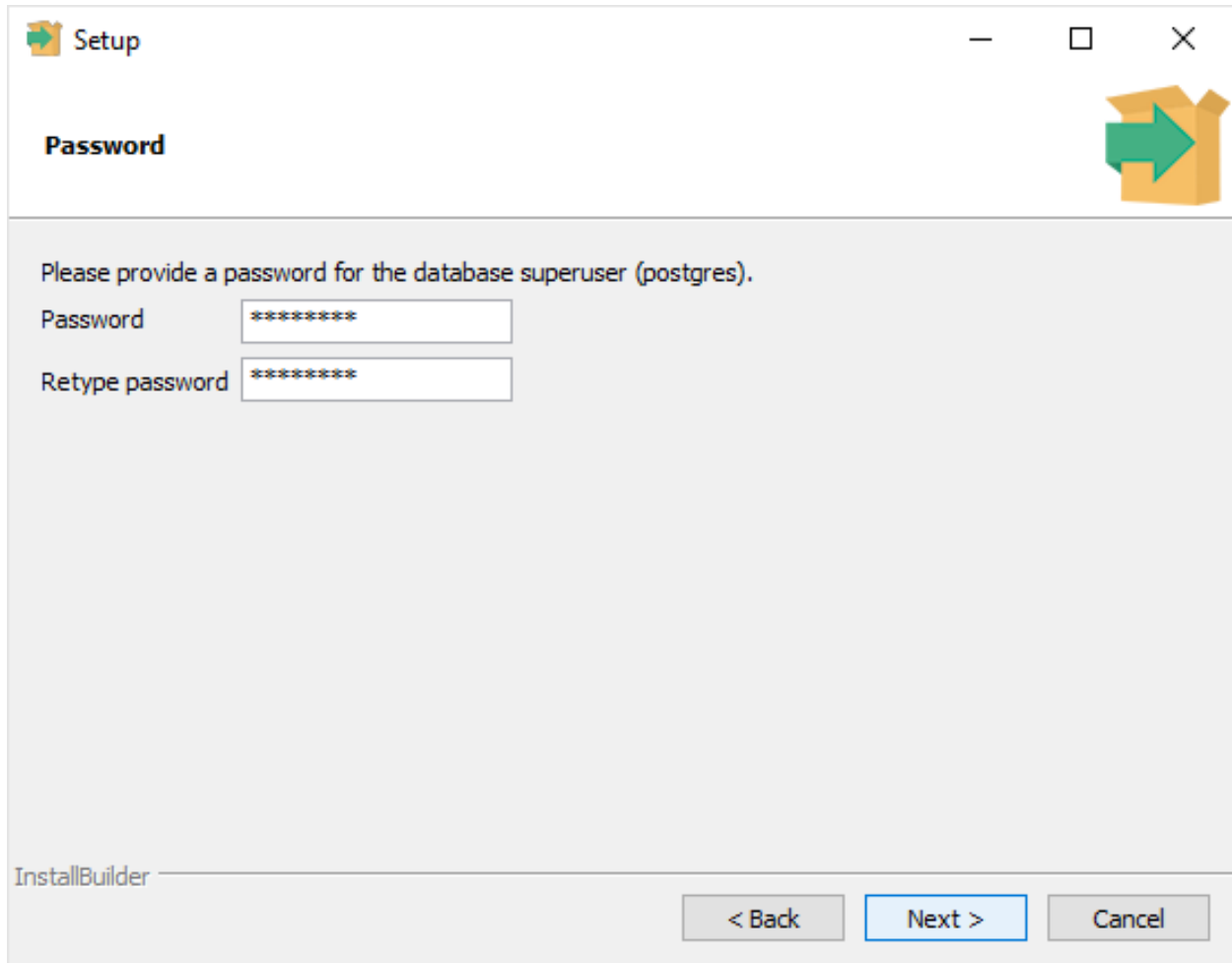
Er wordt gevraagd waar het programma geïnstalleerd moet worden.



Daarna wordt gevraagd welke map gekozen kan worden voor data (databases). Er wordt voorgesteld om een data map in de 'Program Files' map te gebruiken voor opslag van data. Dat is ongebruikelijk; beter is het om data van programmabestanden te scheiden. We kiezen een eigen map.



Kies een wachtwoord voor de hoofdgebruiker 'postgres' en onthoud deze. De gebruikersnaam 'postgres' en het wachtwoord vormen de combinatie om op een later tijdstip toegang tot PostgreSQL te krijgen.



Setup

**Password**

Please provide a password for the database superuser (postgres).

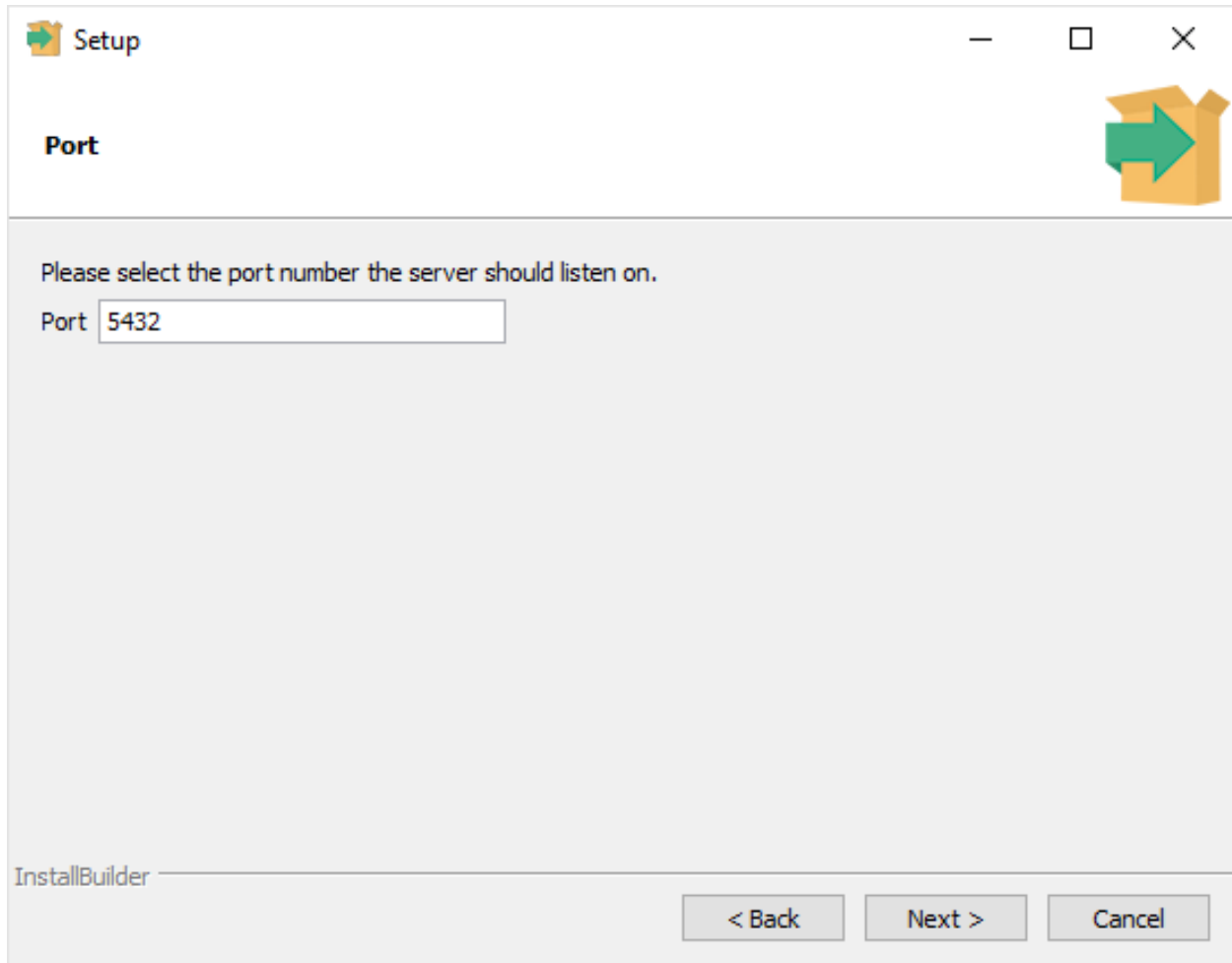
Password

Retype password

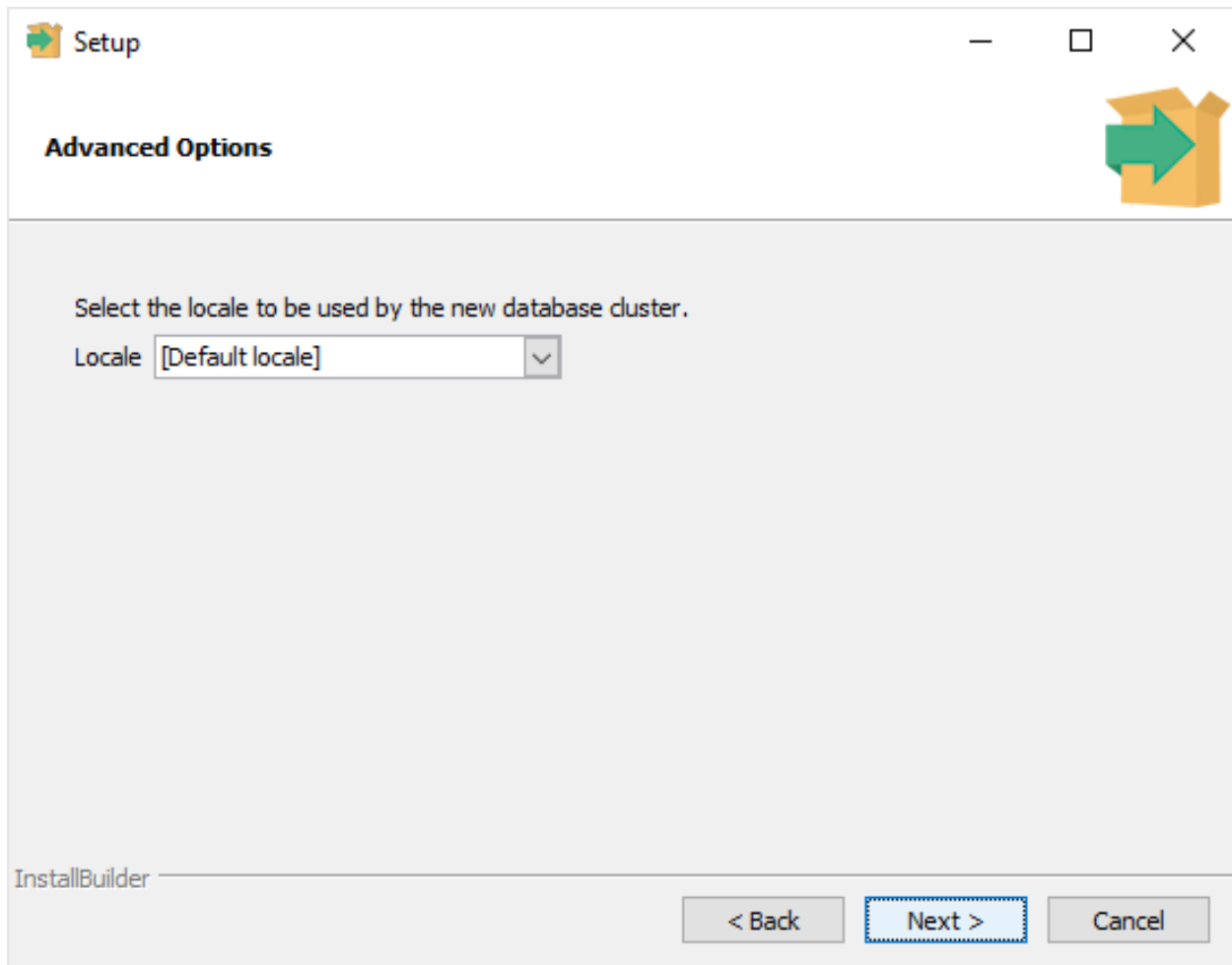
InstallBuilder

< Back   Next >   Cancel

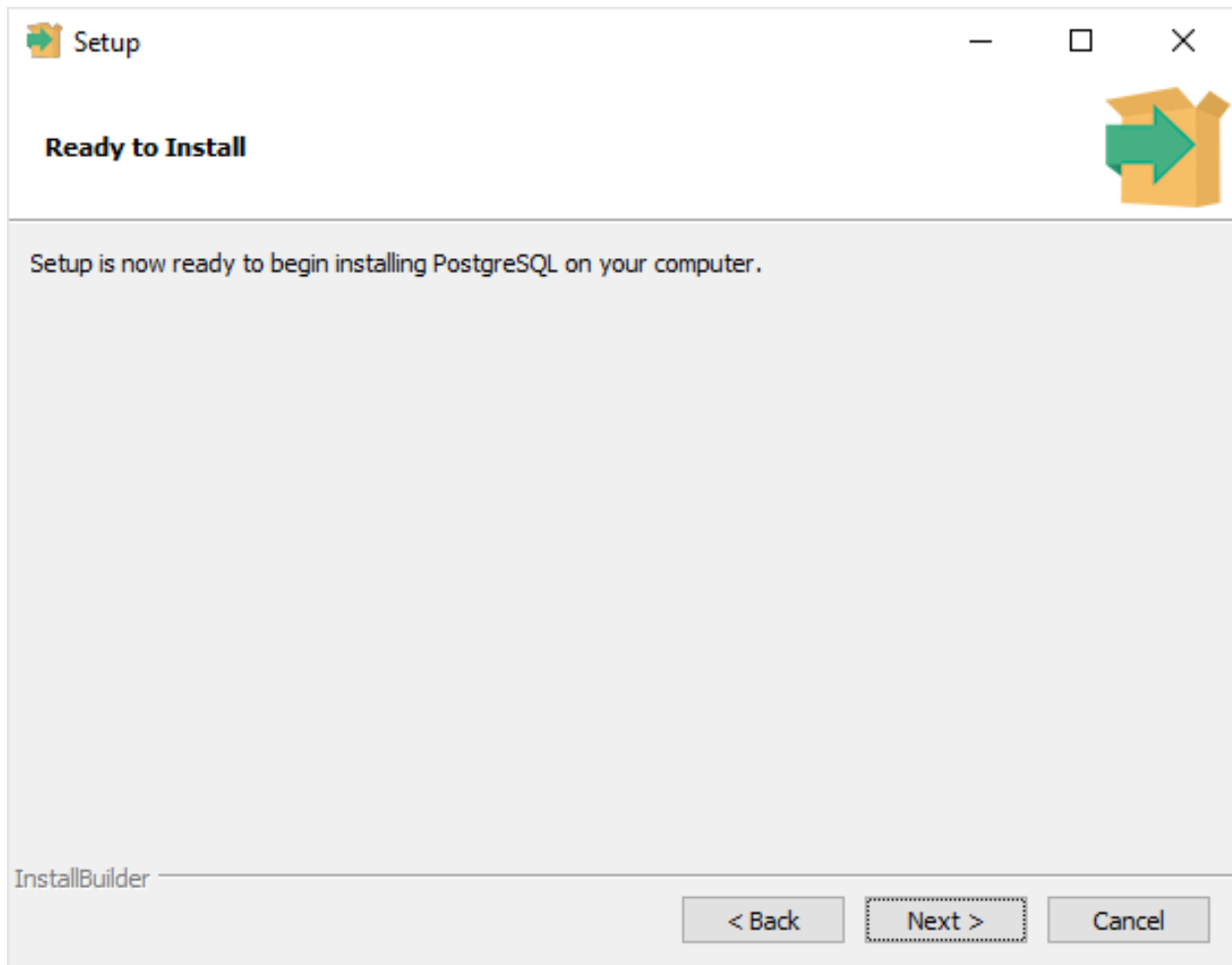
Behoud het standaard poortnummer.



De 'Default locale' is prima.

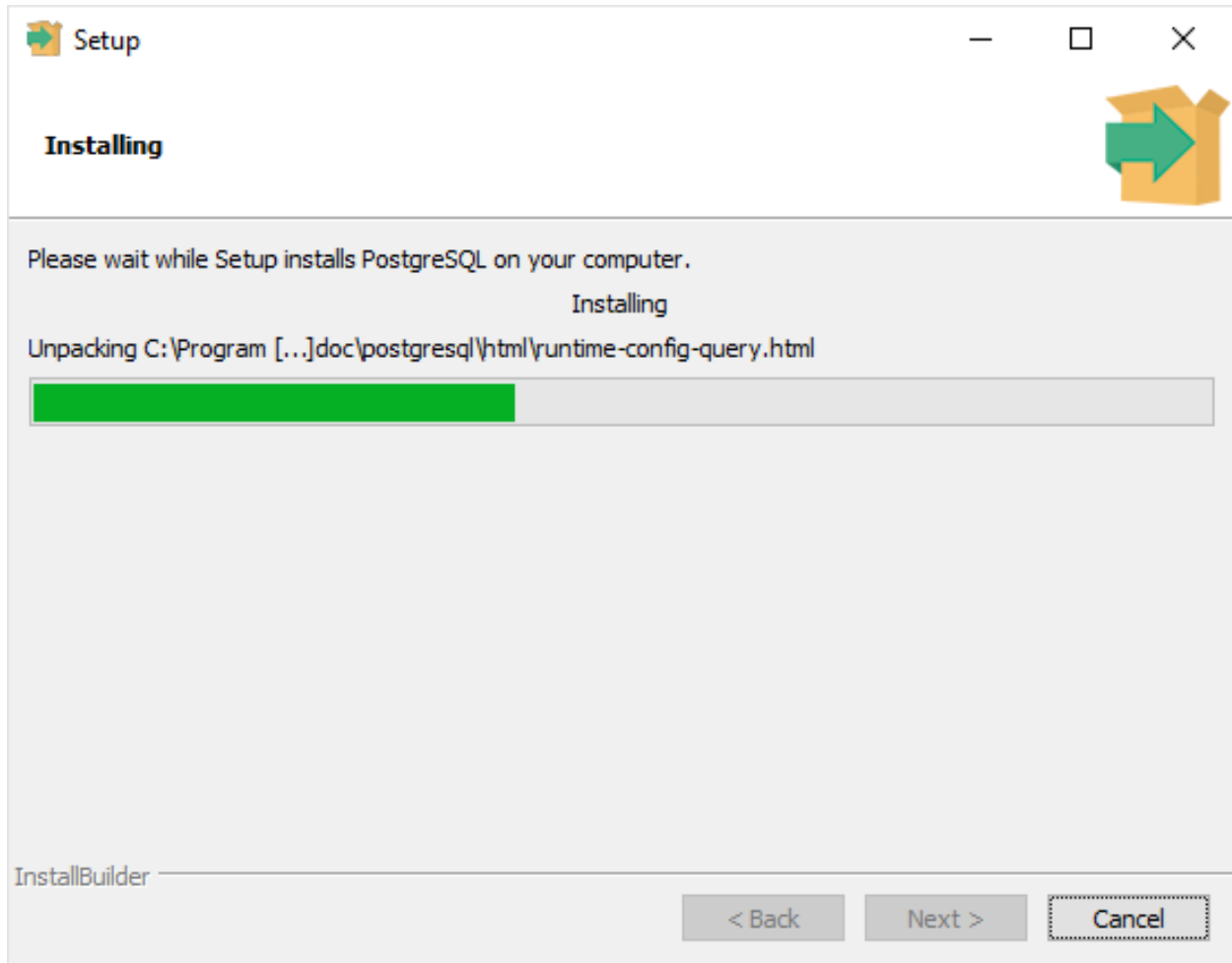


Alles is gereed om de installatie te beginnen.

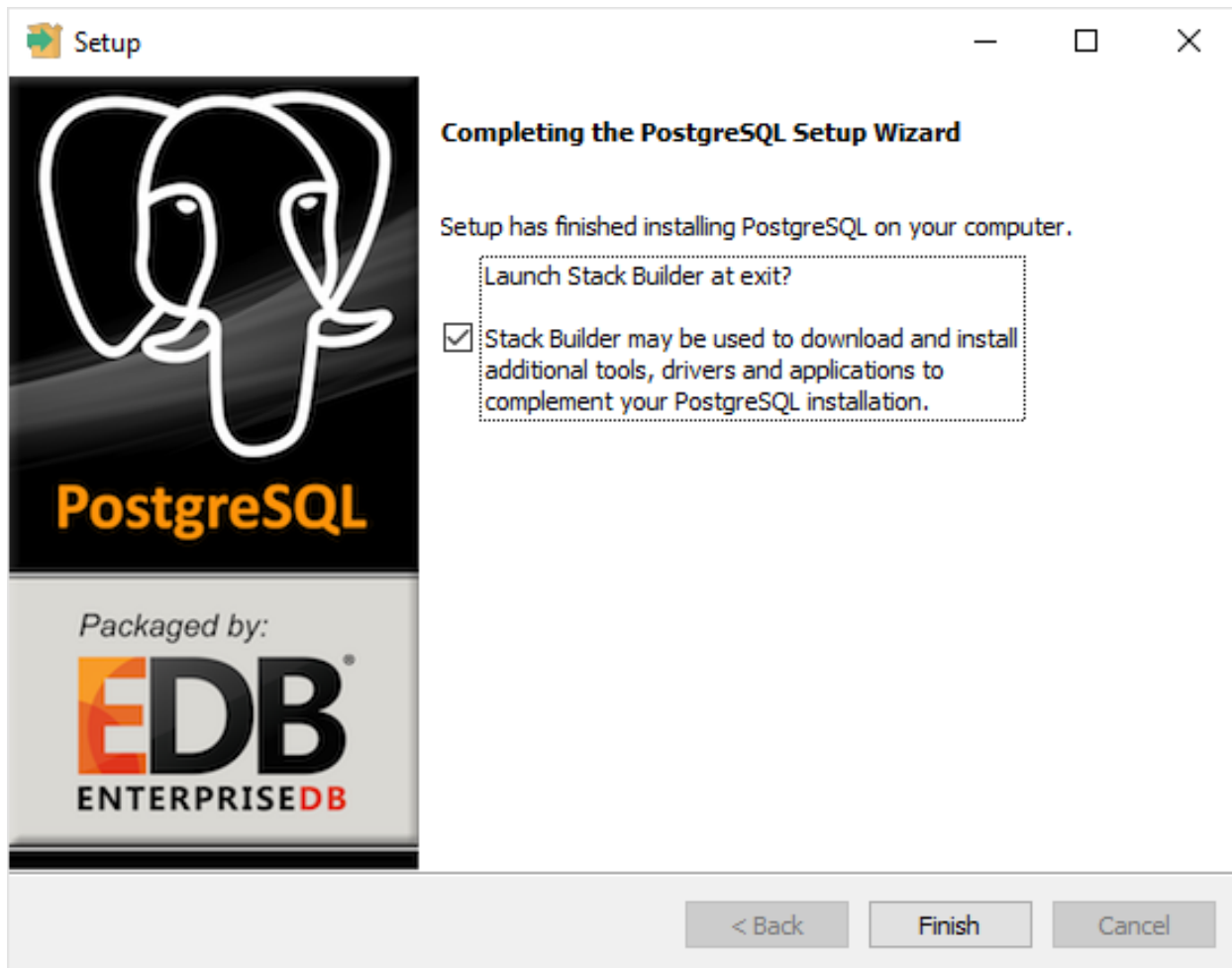


Als het goed is, zal de installatie vlot verlopen.





Nu is PostgreSQL geïnstalleerd. Het laatste scherm biedt de mogelijkheid om uitbreidingen toe te voegen. We laten het vinkje voor de Stack Builder aan staan, zodat we PostGIS kunnen installeren. Dit wordt hierna beschreven.

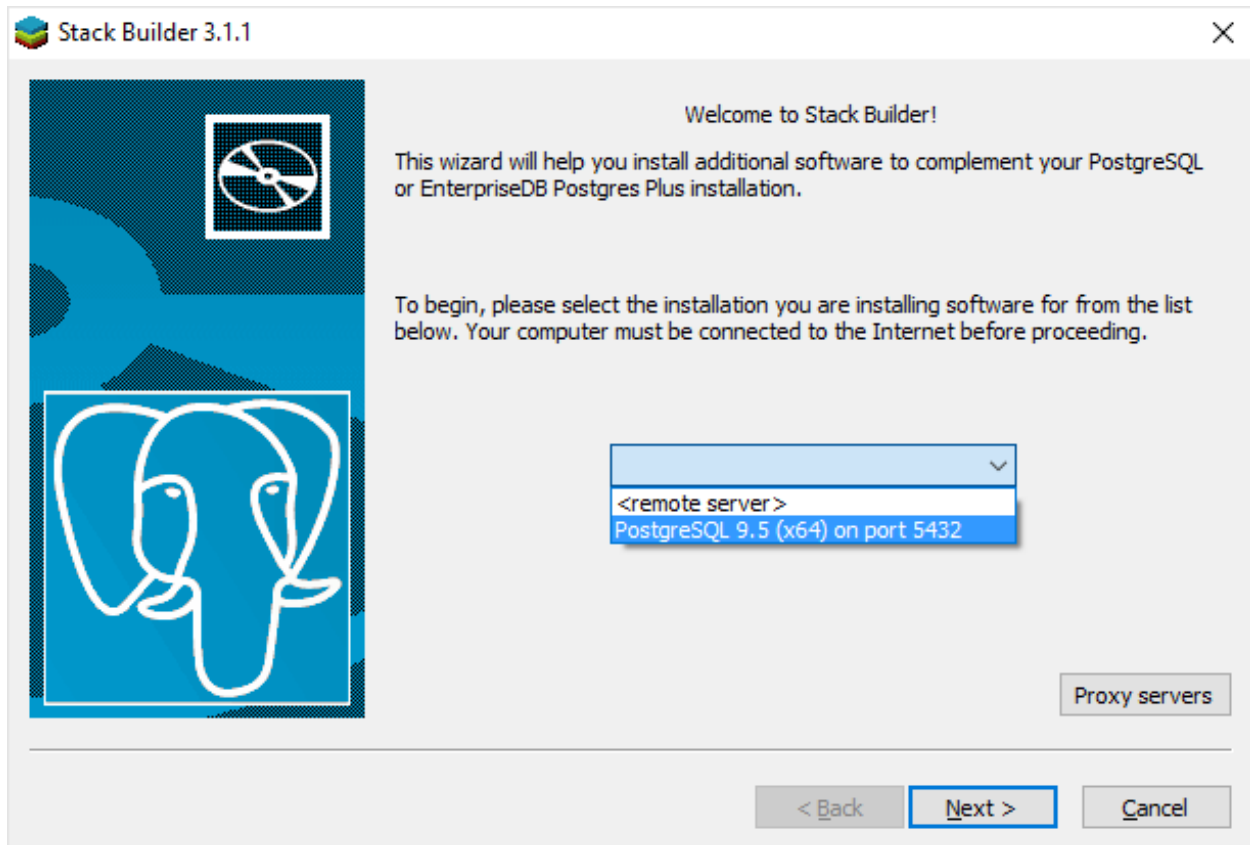


### 8.1.3 PostGIS Installeren

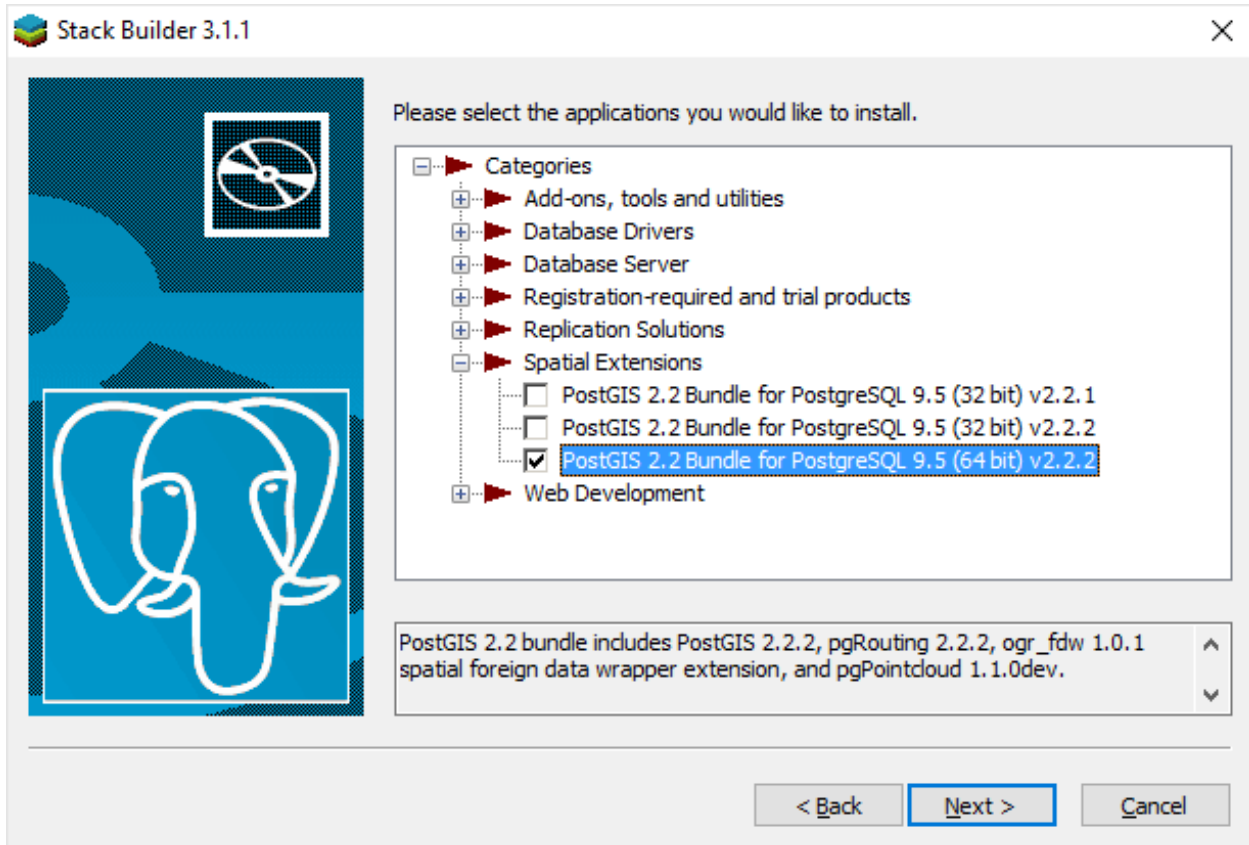
PostGIS is de uitbreiding op PostgreSQL die het mogelijk maakt om data met geometrische / geografische gegevens op te slaan en te verwerken.

We zien het beginscherm van de Application Stack Builder, dat in navolging op de PostgreSQL wordt uitgevoerd. Eventueel kan deze ook handmatig worden gestart.

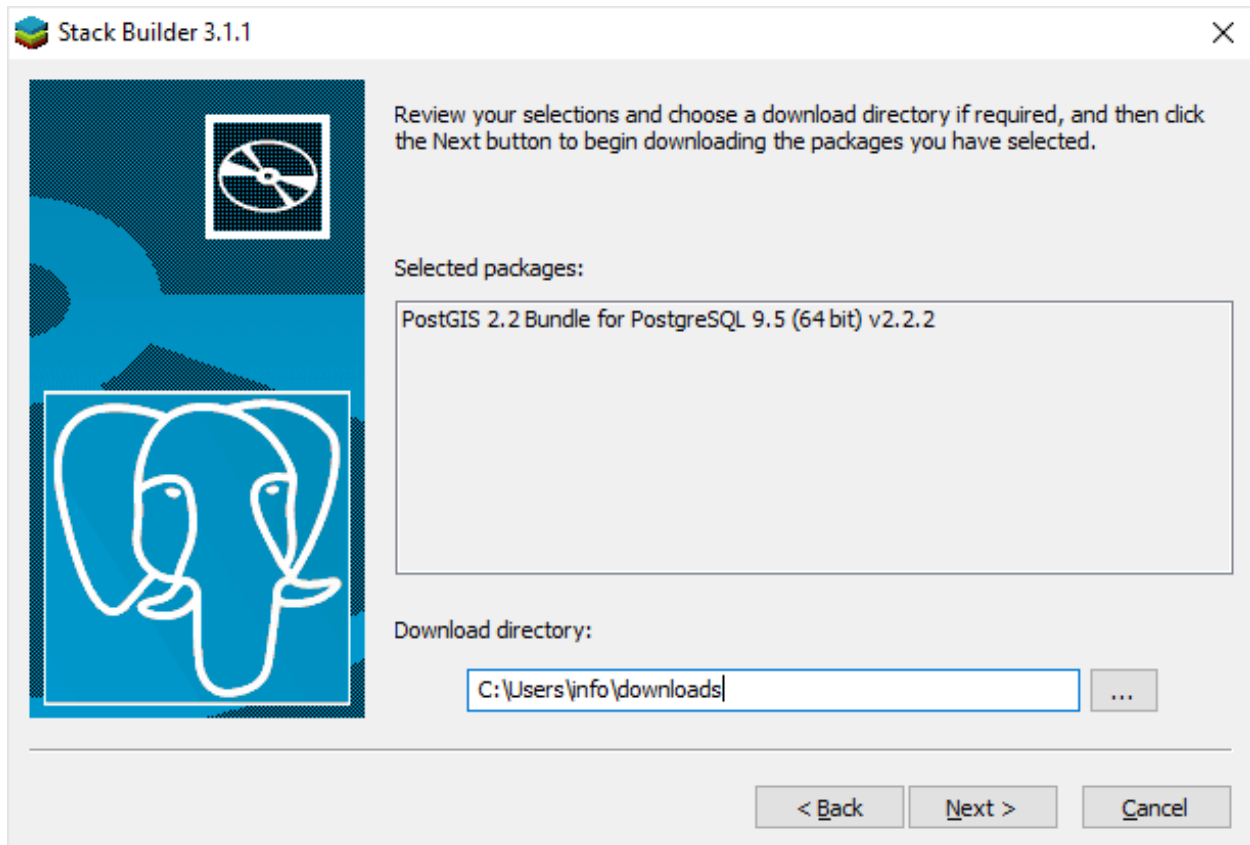
Kies in het keuzemenu de PostgreSQL server.



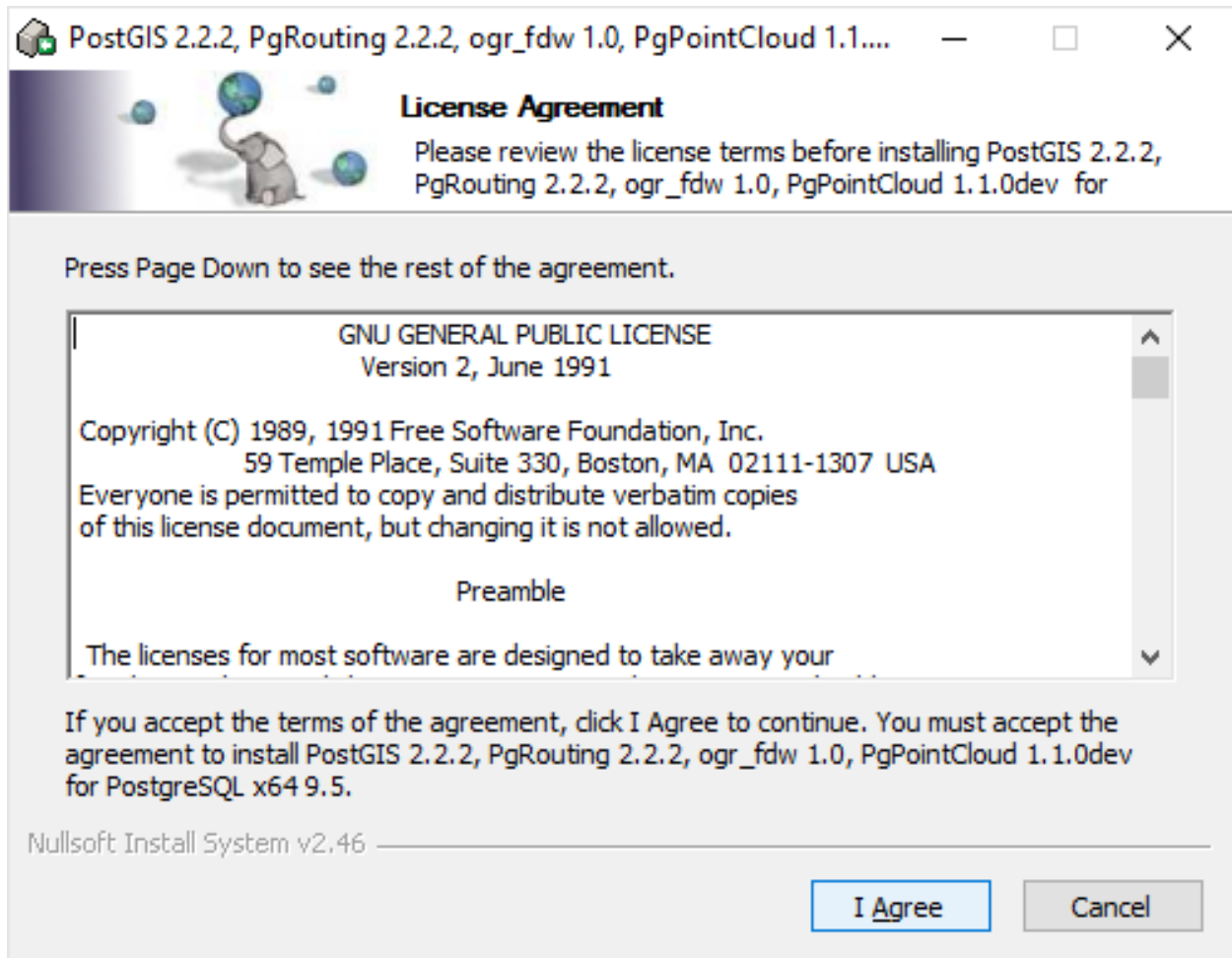
Kies in het volgende scherm de benodigde PostGIS bundle door een vinkje te plaatsen. Kies de bundle die bij de PostgreSQL installatie past. In deze instructie is dat 64 bit, PostgreSQL versie 9.5.



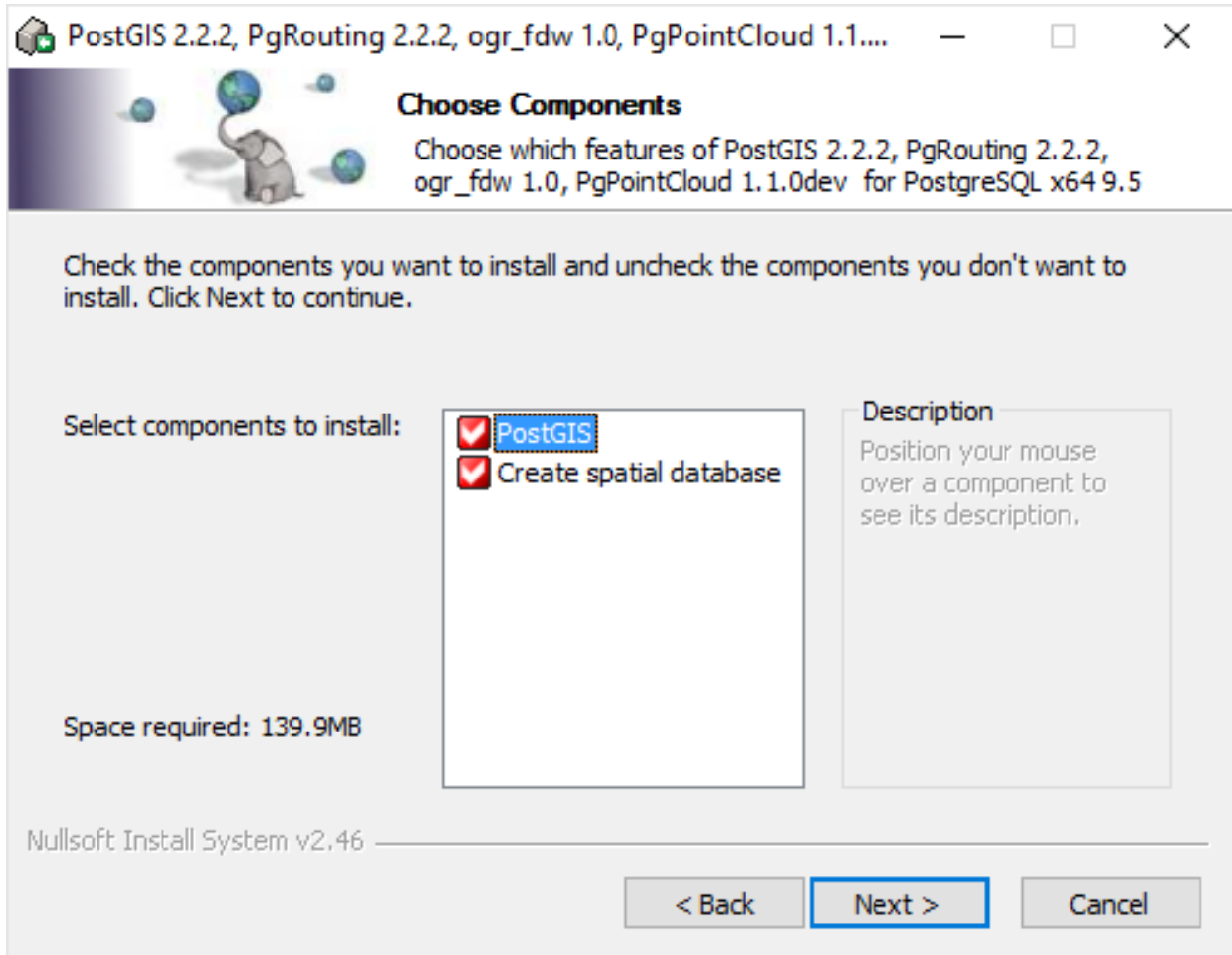
Stack builder vraagt waar de gedownloadte uitbreidingen geplaatst mogen worden. In deze instructie kiezen we voor de standaard map met de naam 'downloads' (waar alle downloads gebruikelijk toch al in komen). Een andere map zou ook prima moeten werken; het is bedoeld om tijdelijk te gebruiken.



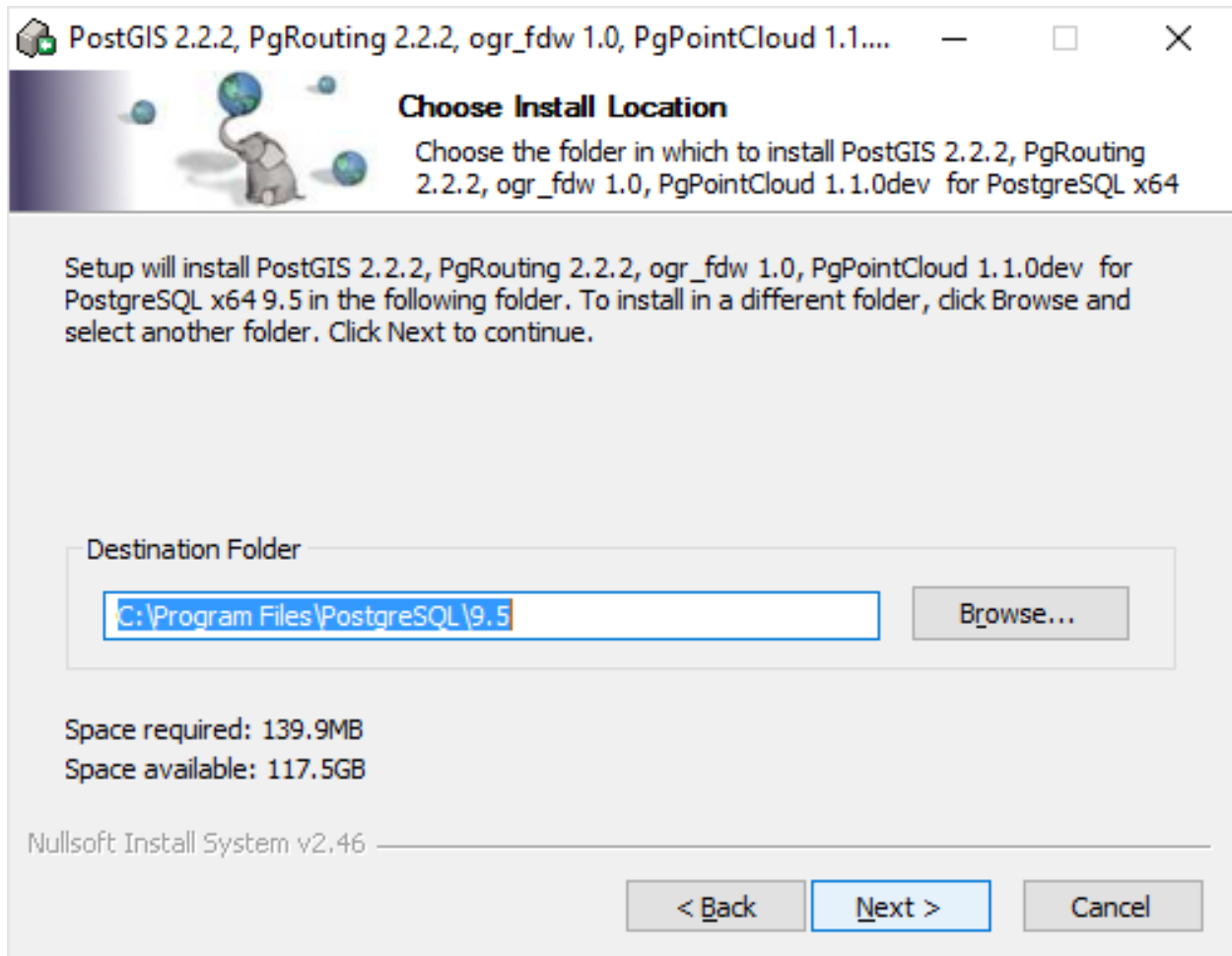
Nu wordt gevraagd akkoord te gaan met de licentievoorwaarden.



Voor het gemak vinken we de keuze 'Create spatial database' aan. Het is in deze instructie de bedoeling om direct een database aan te maken die kan dienen om de dump (back-up) terug te zetten. Het kan uiteraard ook op een later moment, bijv. wanneer je een tweede database dump wilt terugzetten.

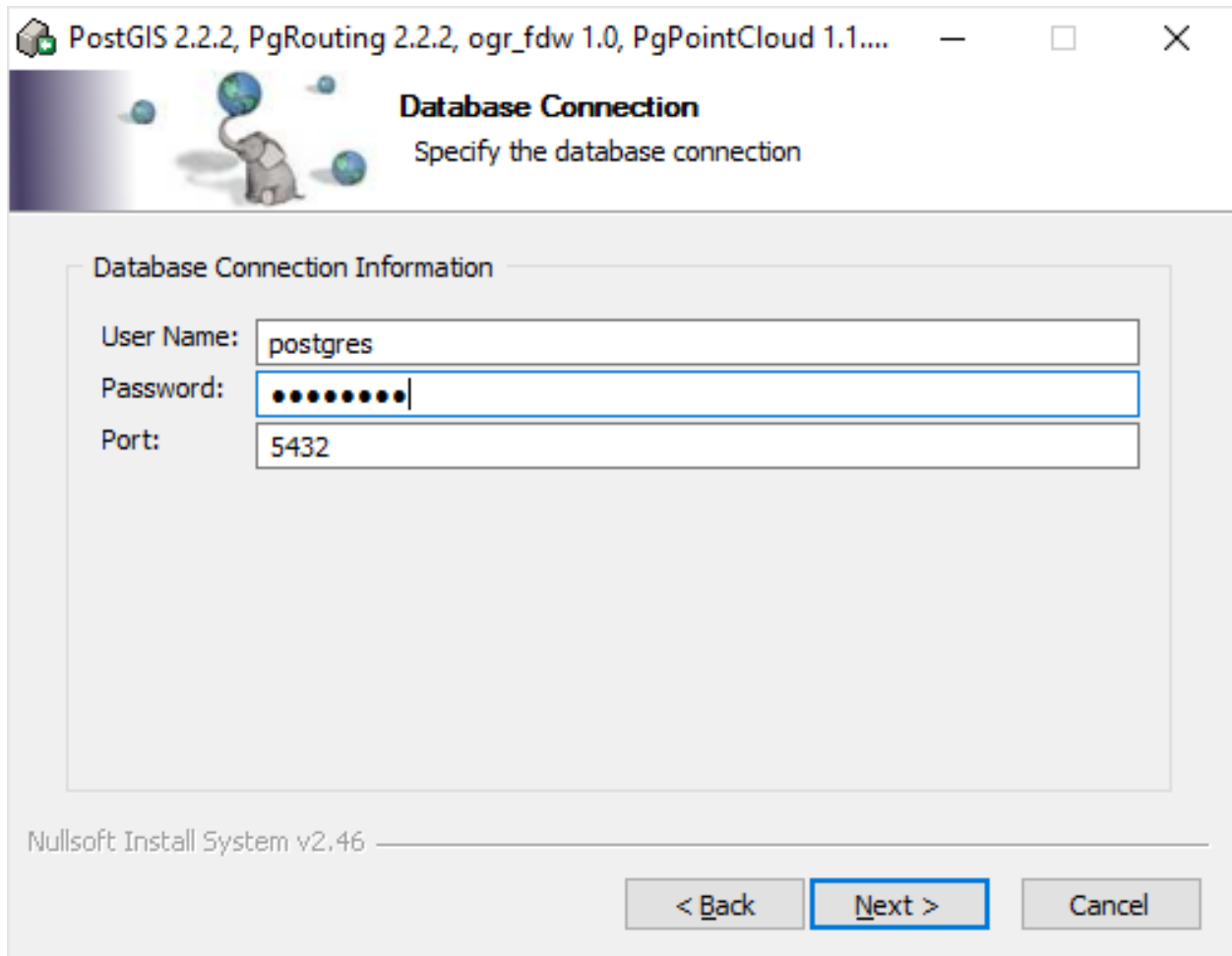


Hier wordt gevraagd waar de PostGIS uitbreiding geplaatst kan worden. We laten dit staan, mits dit klopt met de locatie van de PostgreSQL installatie.



Op het volgende scherm wordt de gebruikersnaam en het wachtwoord gevraagd. Hiermee krijgt het installatieprogramma toegang tot PostgreSQL en kan een database worden aangemaakt (daar hadden we immers voor gekozen).





PostGIS 2.2.2, PgRouting 2.2.2, ogr\_fdw 1.0, PgPointCloud 1.1....

### Database Connection

Specify the database connection

Database Connection Information

User Name: postgres

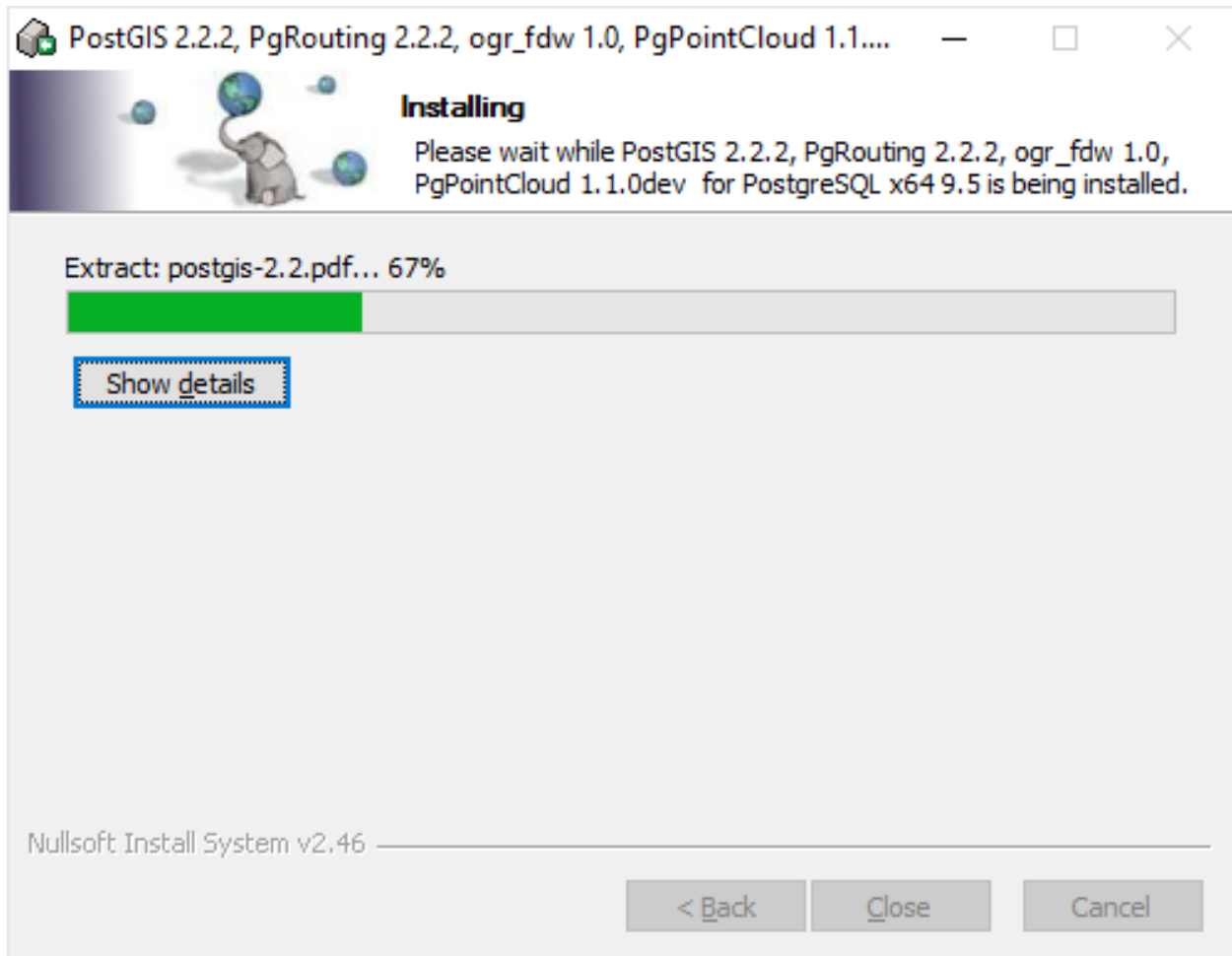
Password: ●●●●●●●●

Port: 5432

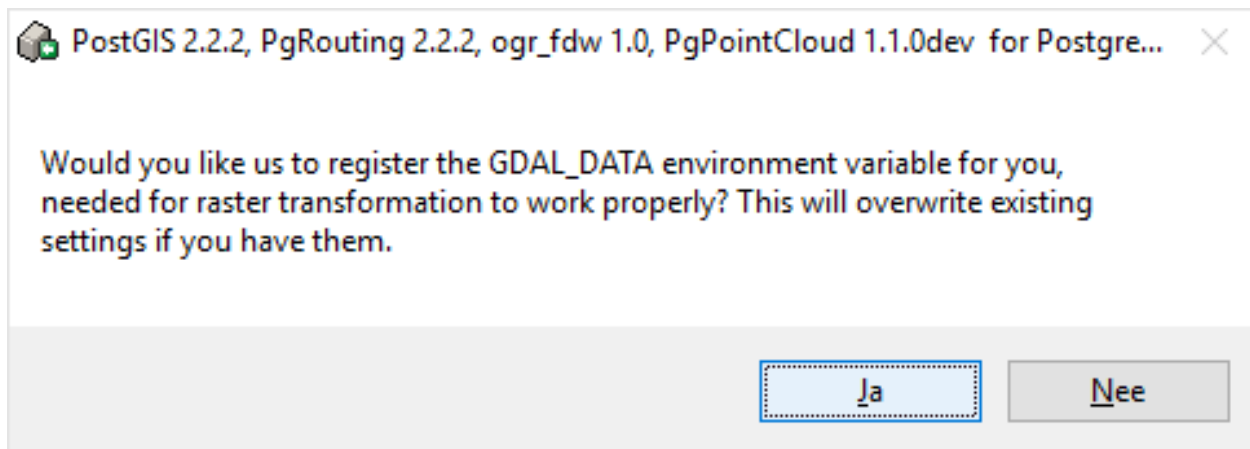
Nullsoft Install System v2.46

< Back   Next >   Cancel

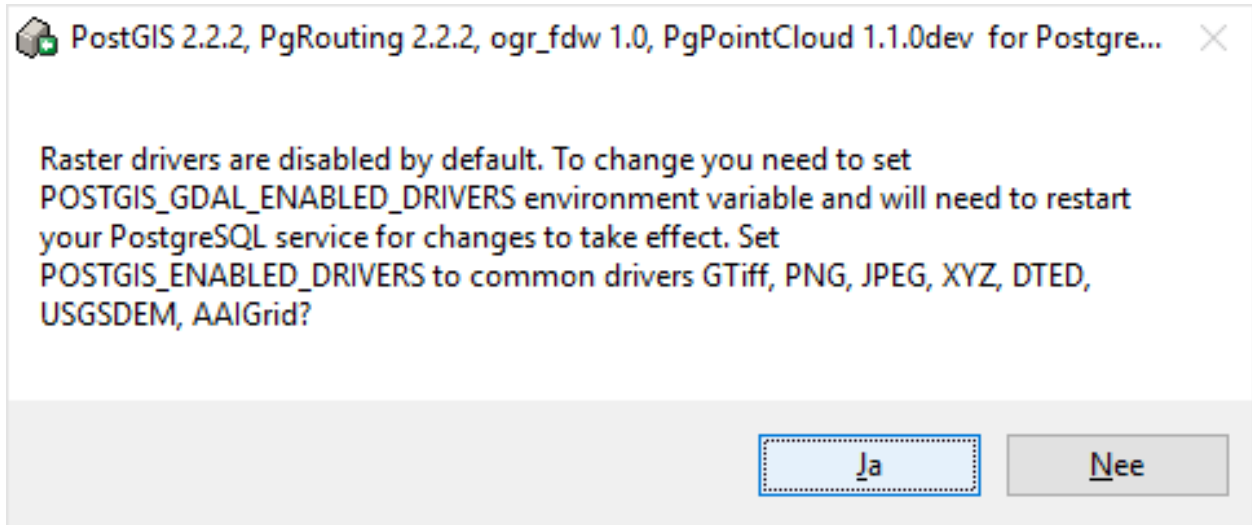
Nu wordt PostGIS geïnstalleerd. Dit verloopt redelijk vlot.



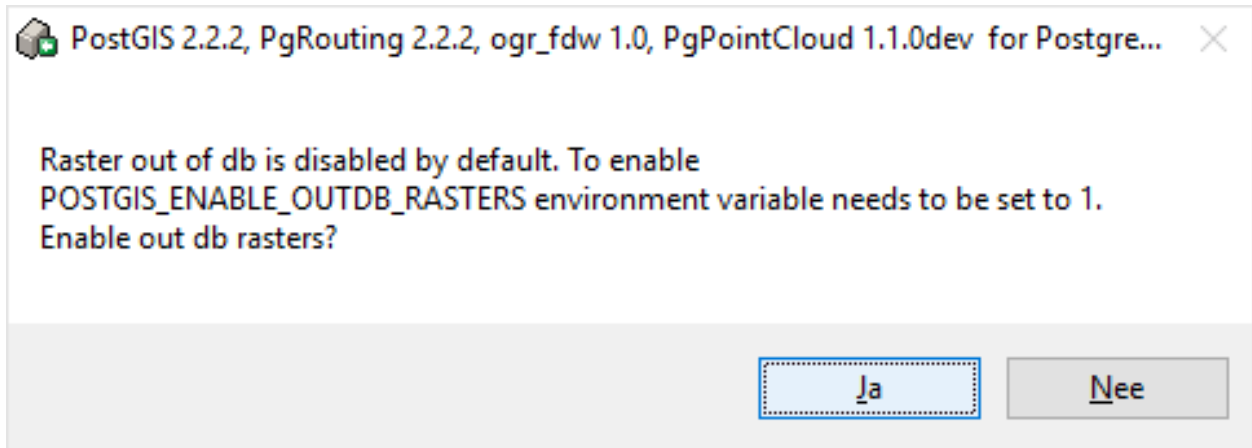
Het is afhankelijk van de toepassing van de database, voor het gemak wordt de vraag bevestigd met 'Ja'.



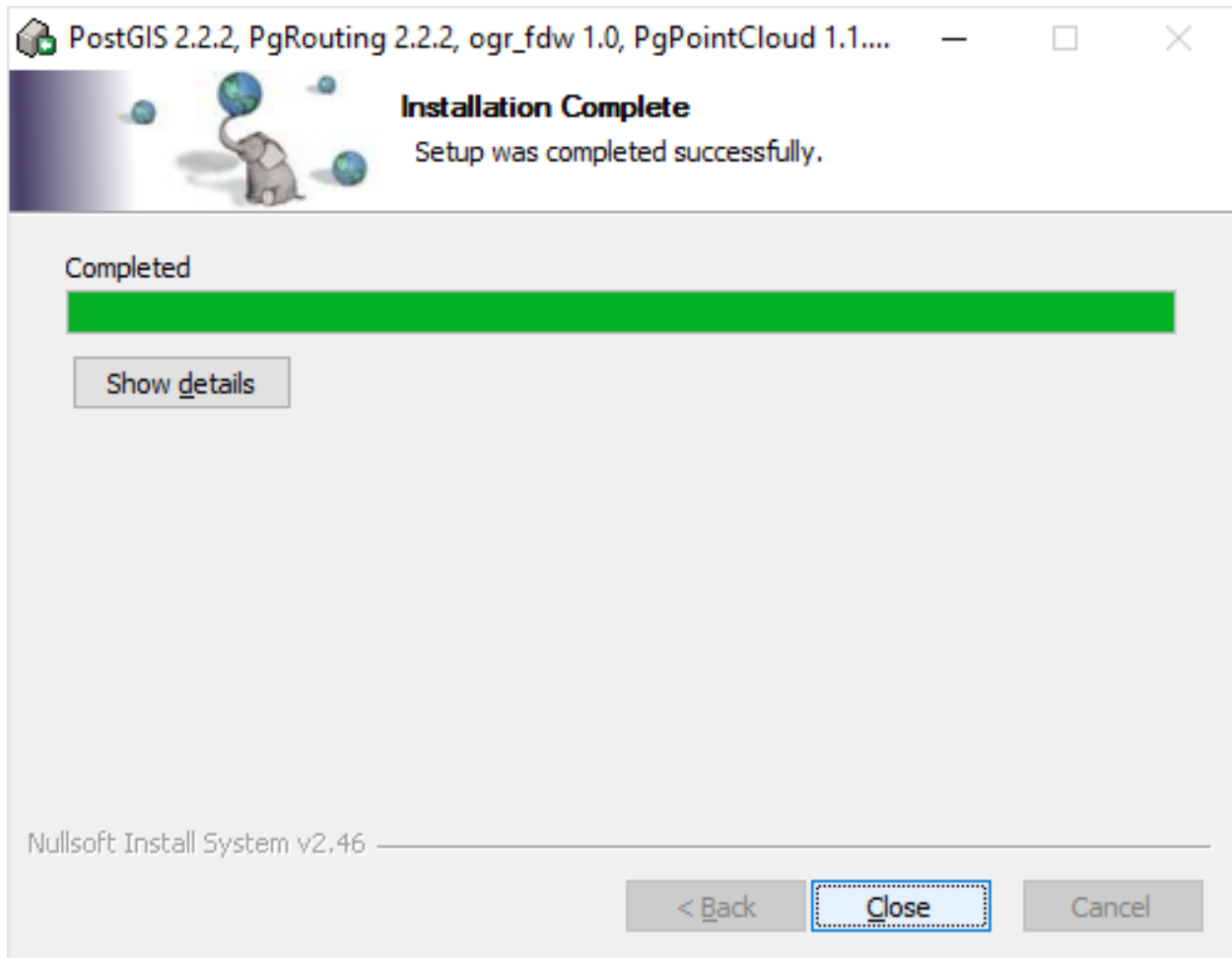
Het is afhankelijk van de toepassing van de database, voor het gemak wordt de vraag bevestigd met 'Ja'.



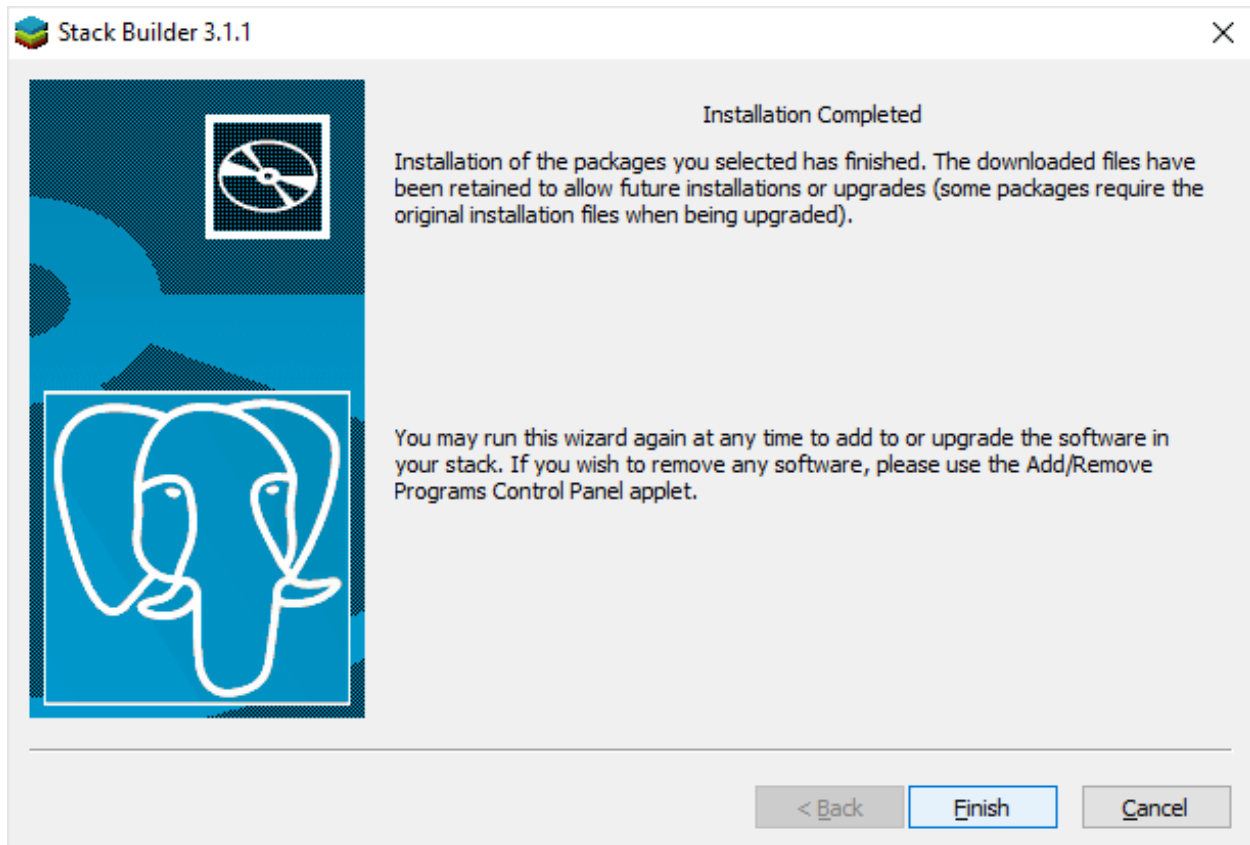
Het is afhankelijk van de toepassing van de database, voor het gemak wordt de vraag bevestigd met 'Ja'.



De installatie van PostGIS is geslaagd en de database is aangemaakt.

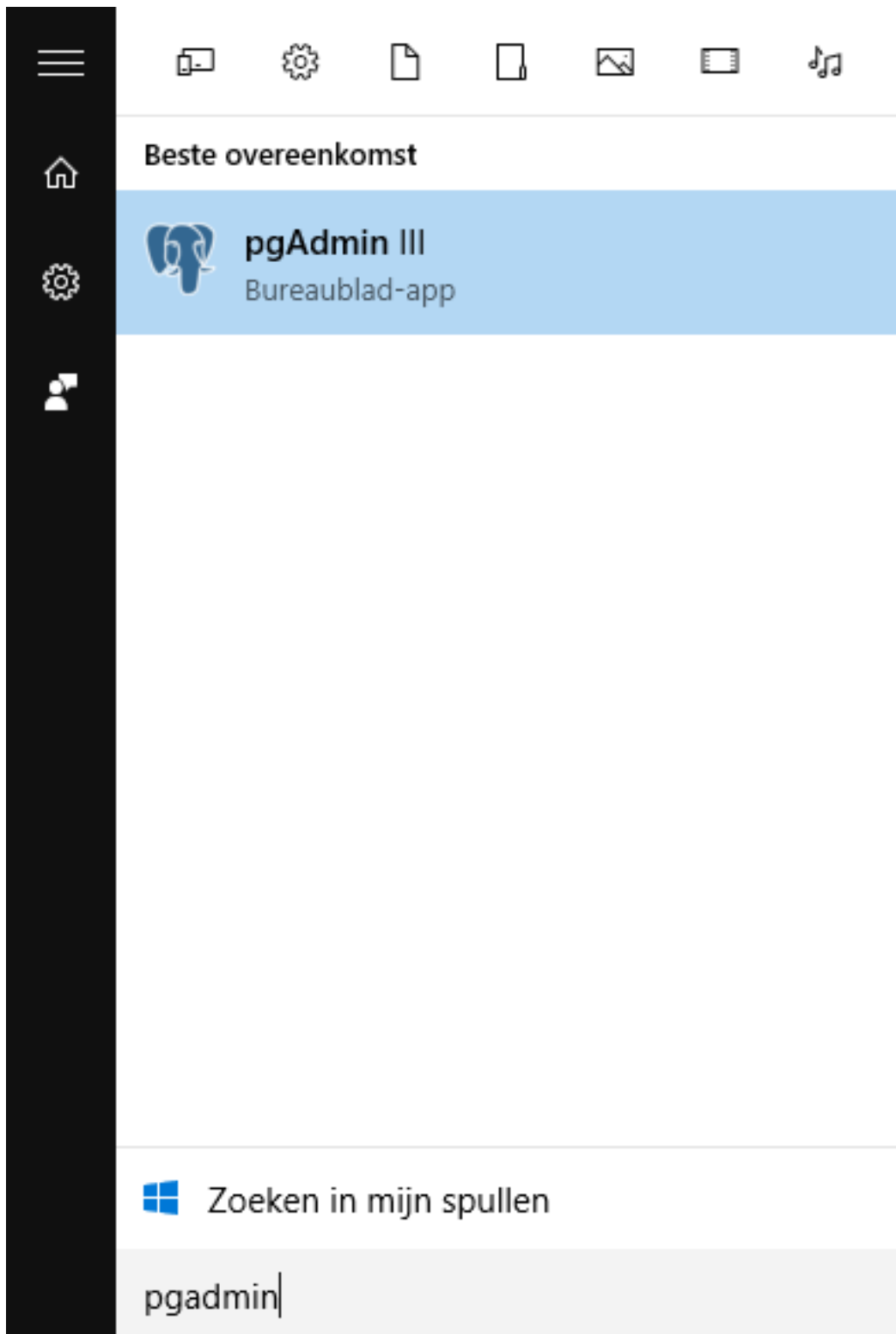


De Stack Builder geeft aan dat alle aangevinkte uitbreidingen zijn geïnstalleerd.

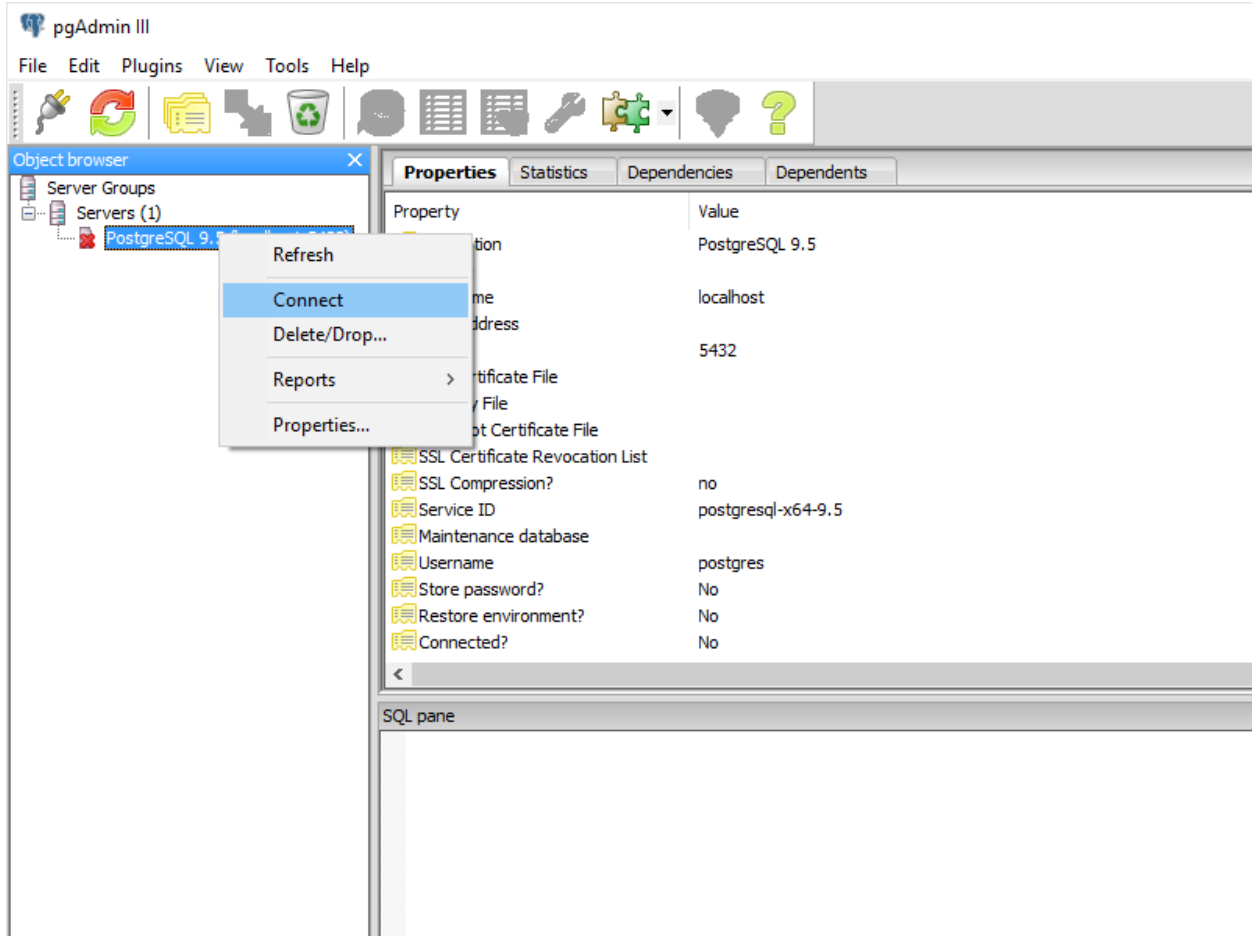


#### 8.1.4 PgAdmin III - BAG login toevoegen

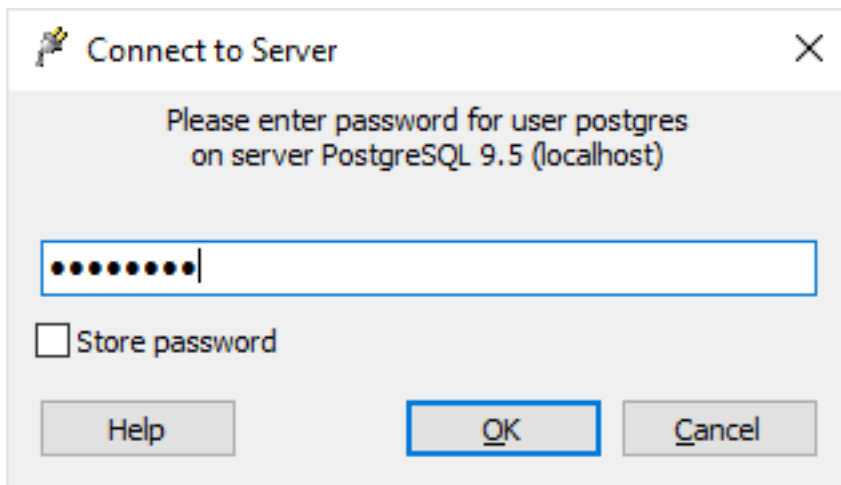
De tool pgAdmin III is geïnstalleerd tijdens de installatie van PostgreSQL. Zoek deze op in het menu (of druk op de Windows toets en type 'pgadmin') en start het programma.



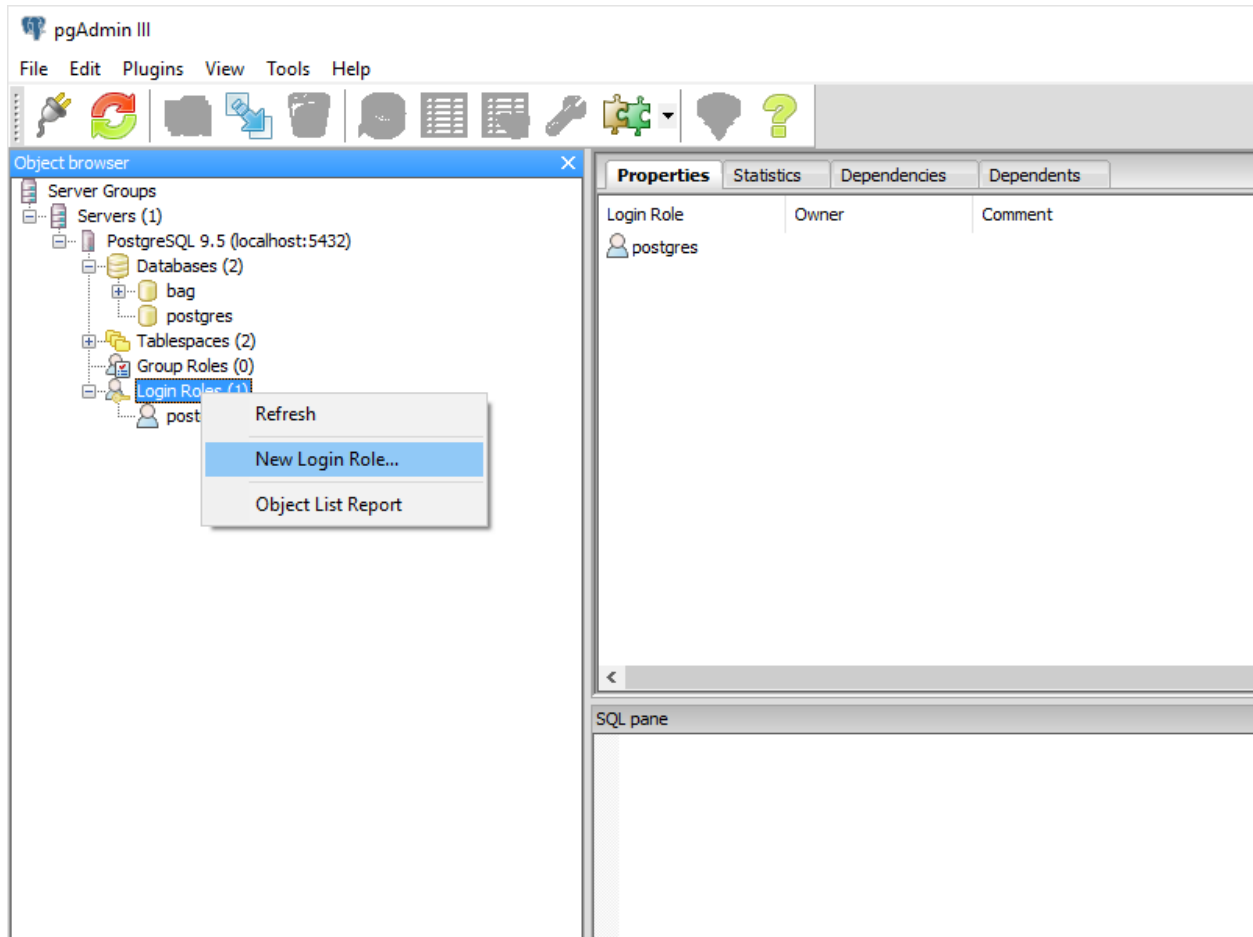
Nadat pgAdmin III is geopend, maak een verbinding (Connect) met de PostgreSQL server.



Voer het wachtwoord in dat behoort bij de gebruiker 'postgres'.



Maak een nieuwe 'login' aan. Dit is een gebruiker van de database. Het is de bedoeling om een login te hebben die past bij de terug te zetten database dump, zoals in deze instructie staat beschreven.



De nieuwe login krijgt de naam 'basuser'. Natuurlijk kun je ook een andere naam kiezen.



New Login Role...

Properties Definition Role privileges Role membership Variables Security

Role name basuser

OID

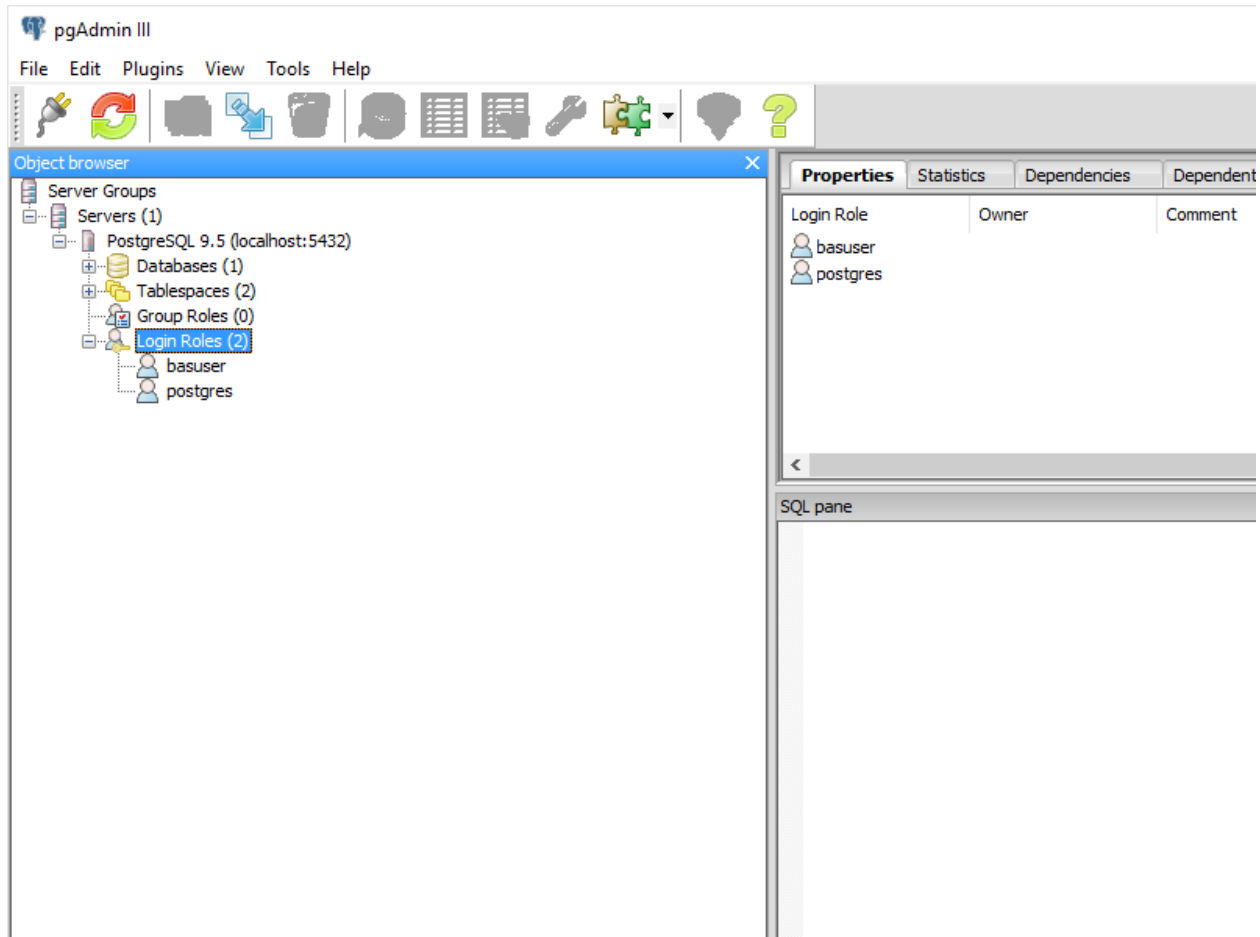
Comment

Use Slony

Help OK Cancel

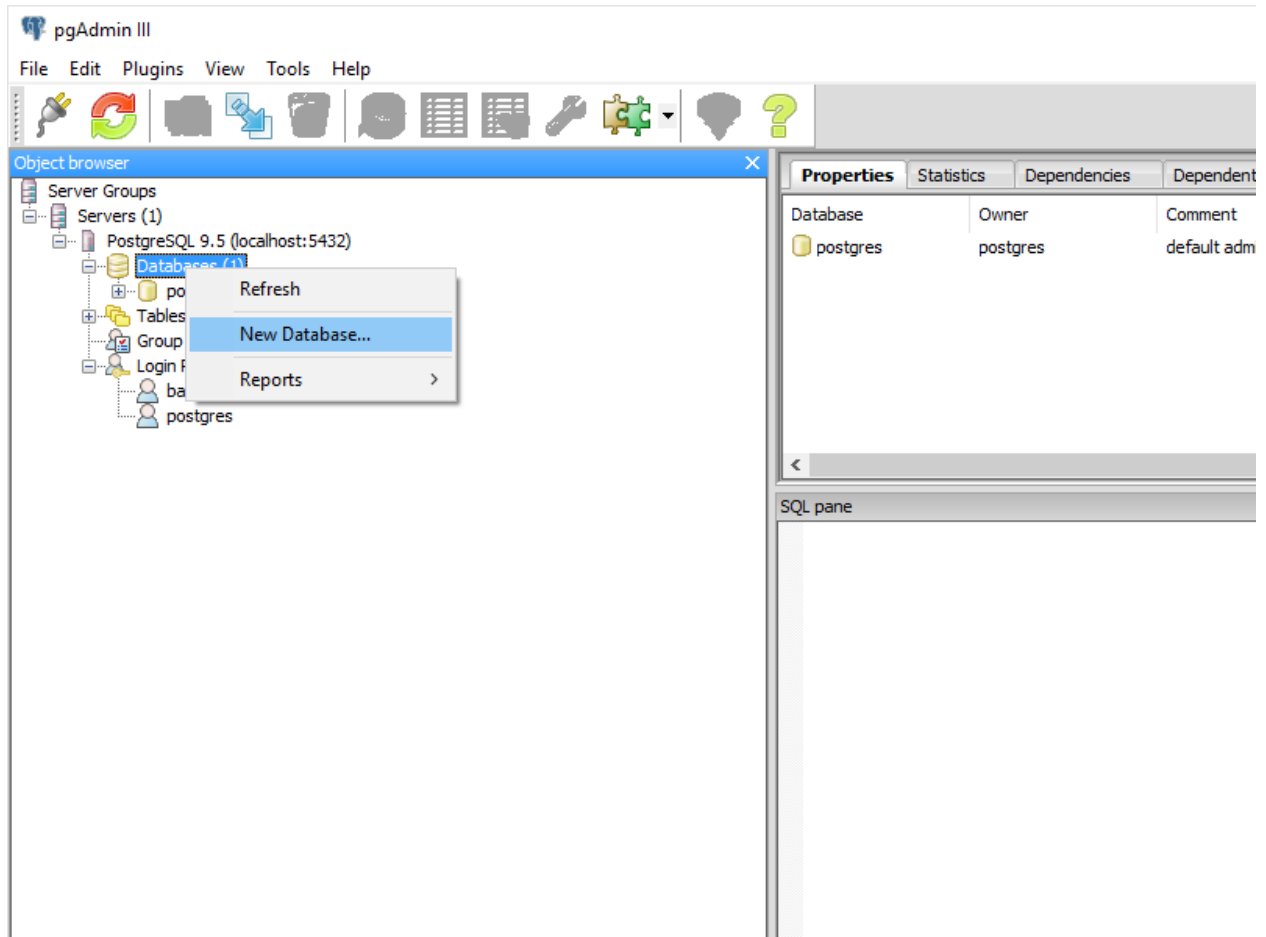
Kies een wachtwoord voor deze login, vul deze twee keer in en onthoud deze. Sluit af met 'OK'.

De PostgreSQL database server heeft nu twee logins.

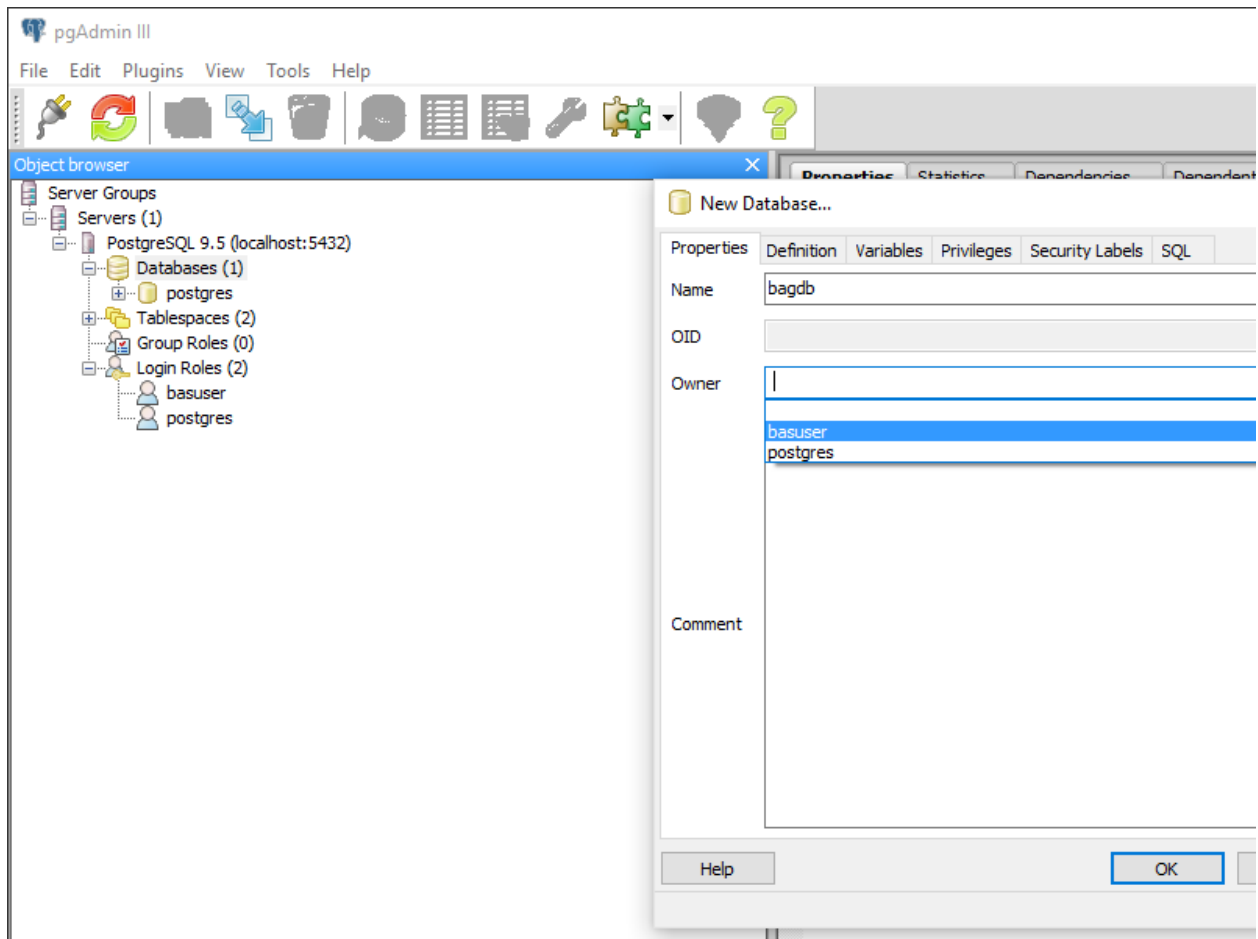


### 8.1.5 PgAdmin III - BAG database aanmaken

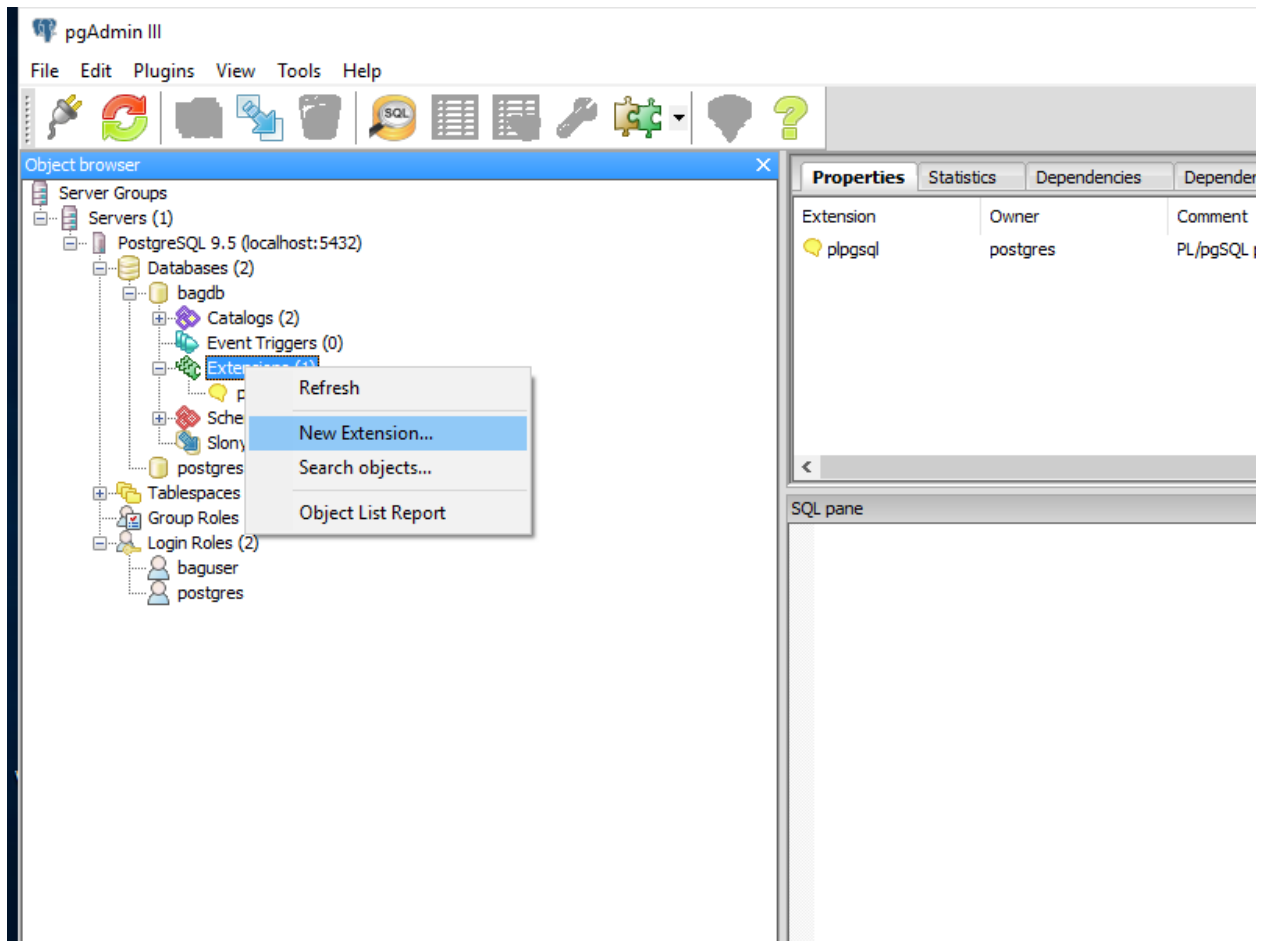
De BAG database aanmaken omvat twee stappen, namelijk een PostgreSQL database aanmaken en een PostGIS uitbreiding toevoegen aan de database. Open voor het aanmaken van een nieuwe database het contextmenu (rechter muisknop) op 'Databases' en kies 'New Database'. NB belangrijk is dat de database met de character-set UTF-8 wordt aangemaakt!



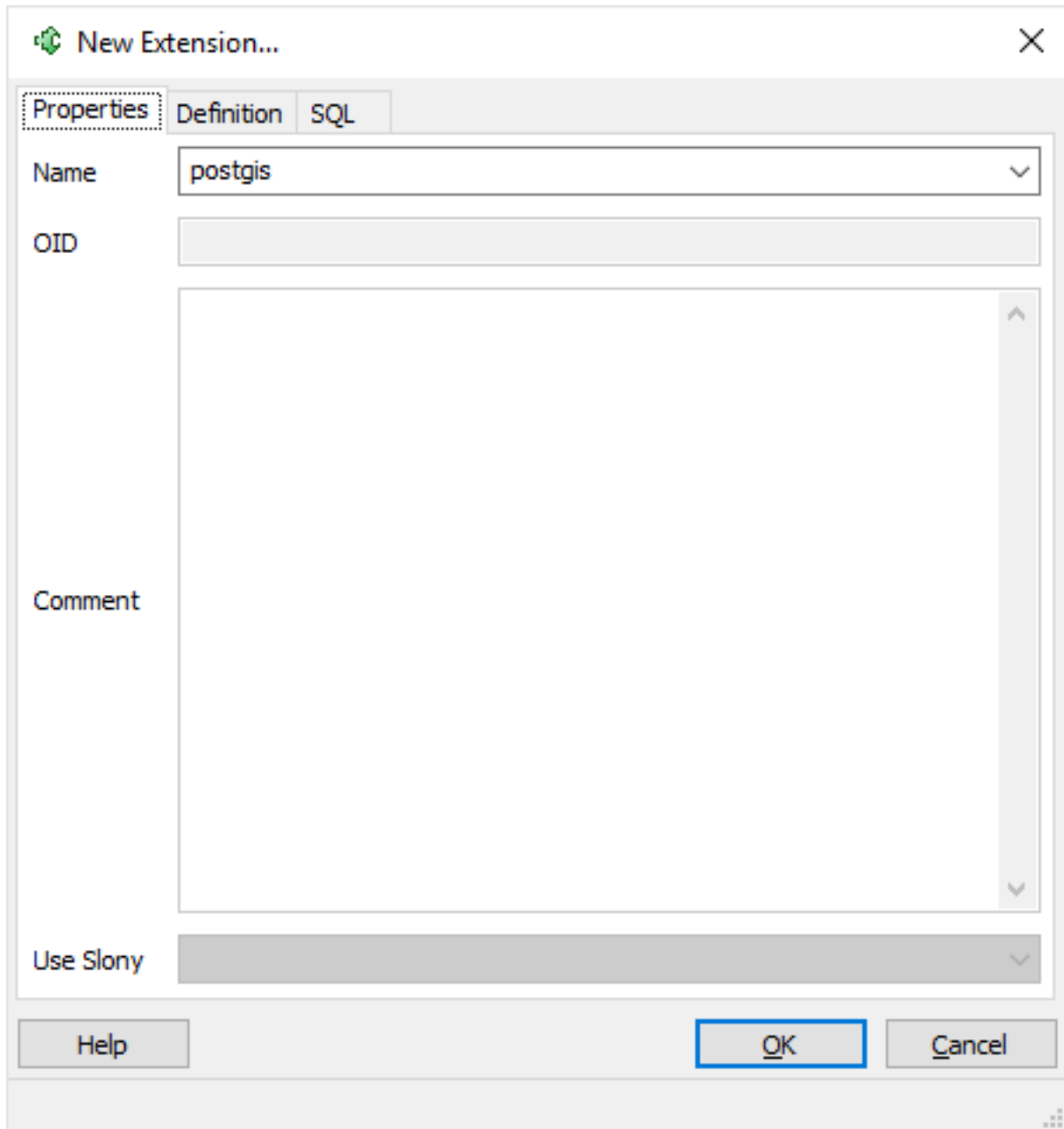
We noemen de database 'bagdb'. Kies de zojuist gemaakte inlog 'baguser' als eigenaar van de database.

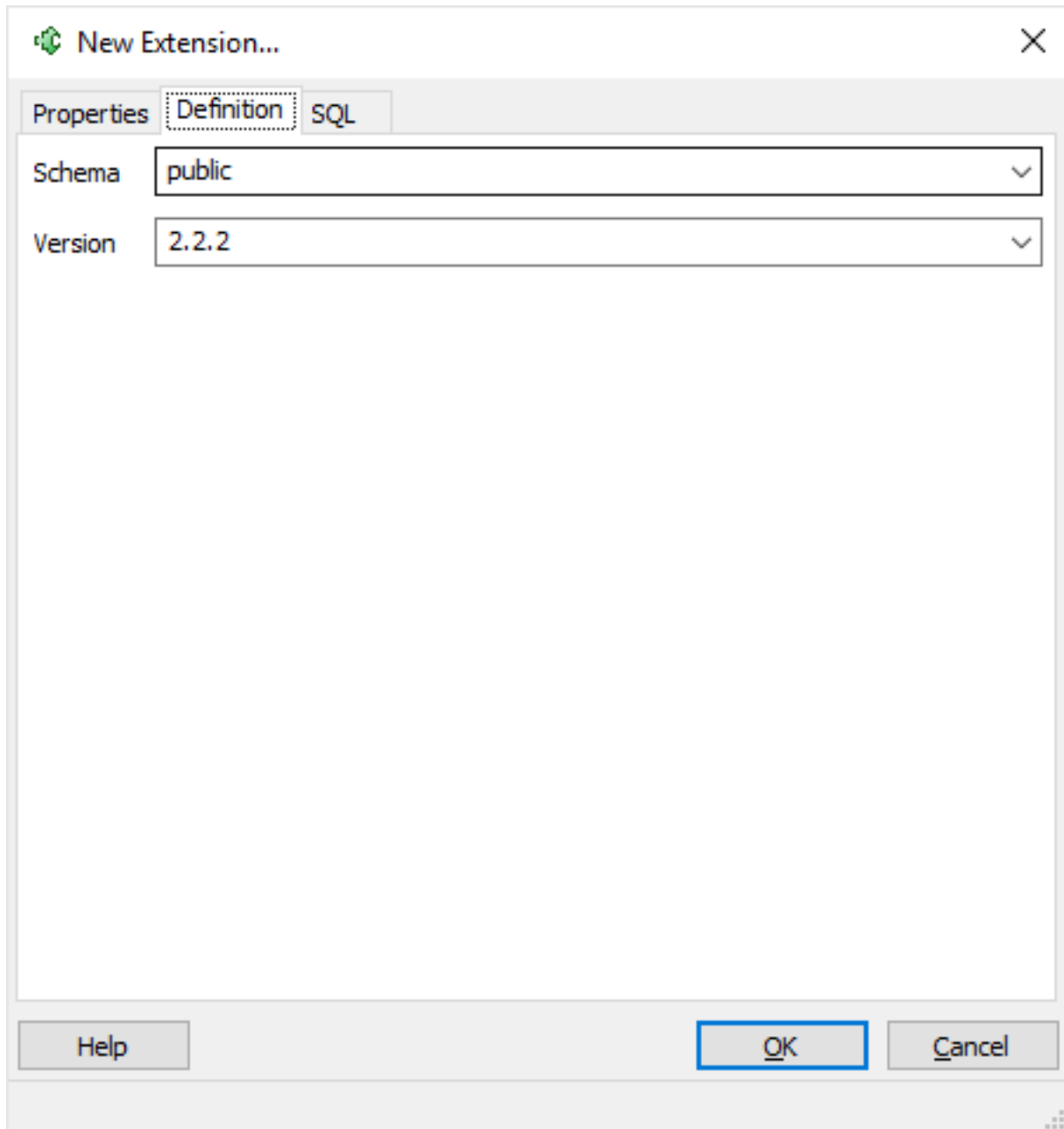


Er is nu een nieuwe database aangemaakt, echter bevat deze nog geen PostGIS mogelijkheden. Hiervoor gaan we een uitbreiding toevoegen aan de database. Kies 'New Extension' in het contextmenu (rechter muisknop) op 'Extensions', binnen de database 'bagdb'.



Kies in het scherm de uitbreiding 'postgis'.



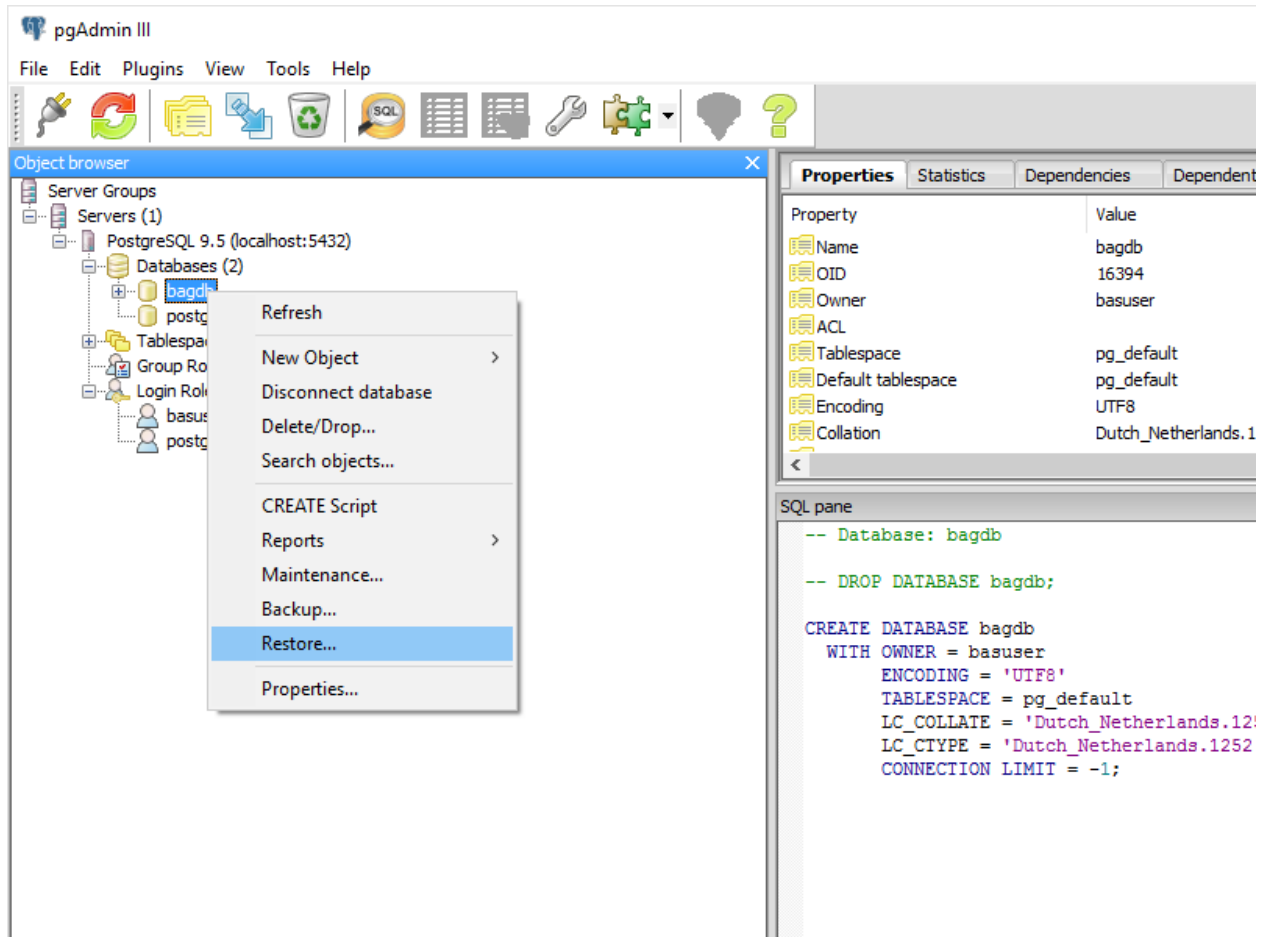


Klik op 'OK' en de uitbreiding 'postgis' zal zichtbaar worden.

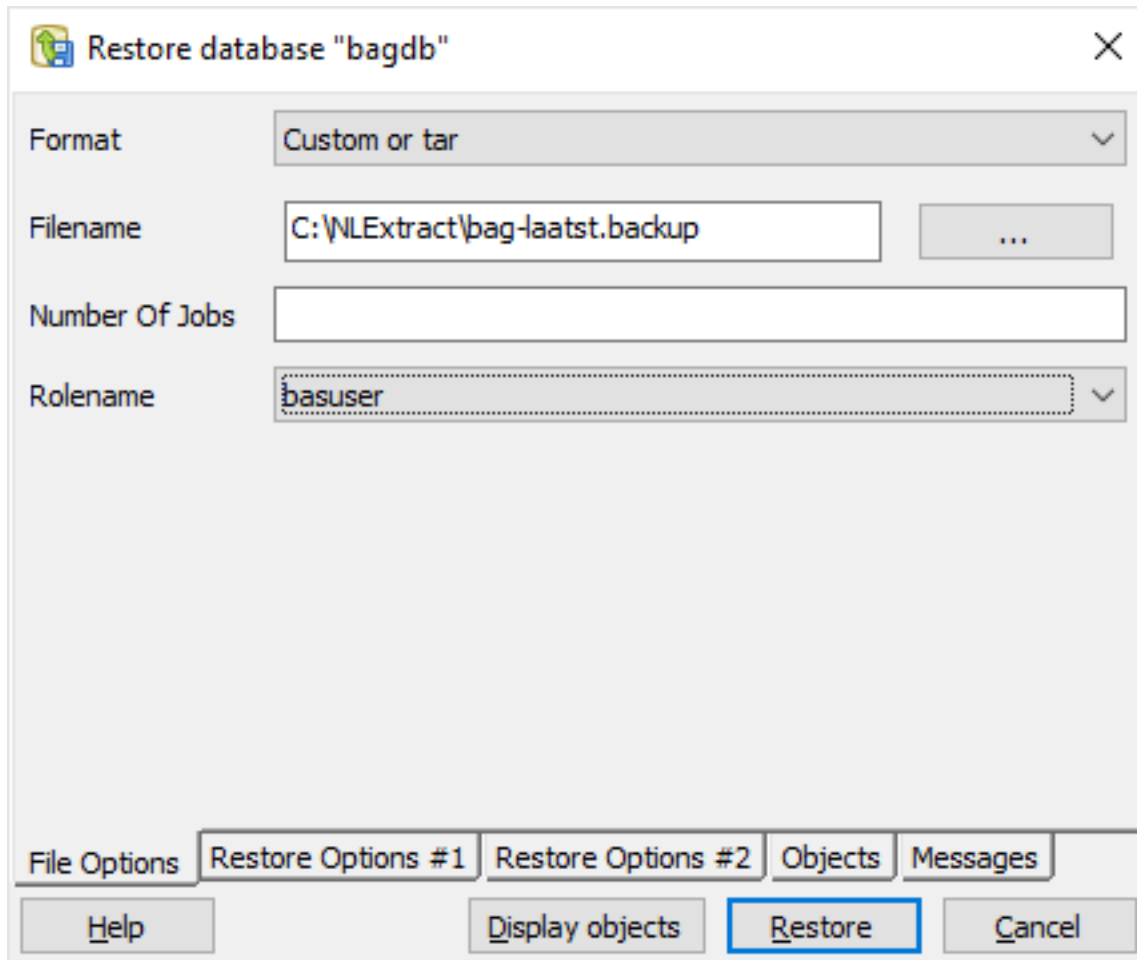
### 8.1.6 PgAdmin III - BAG dump terugzetten

Kies optie 'Restore' middels het contextmenu (rechter muisknop) op de juiste gemaakte database 'bagdb'.

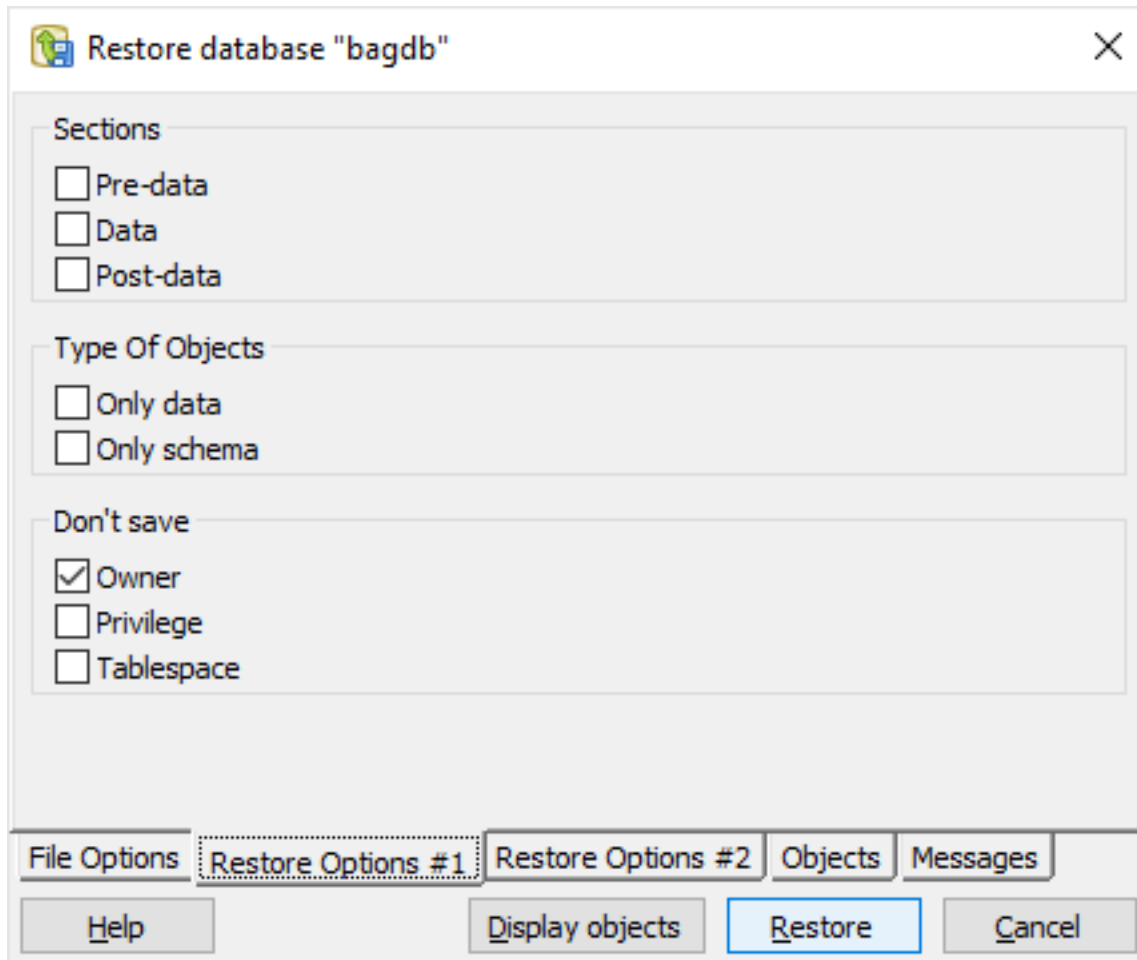




Kies als 'Filename' het eerder gedownloade bestand 'bag-laastst.backup' en kies als 'Rolename' de eerder aangemaakte login 'basuser'.

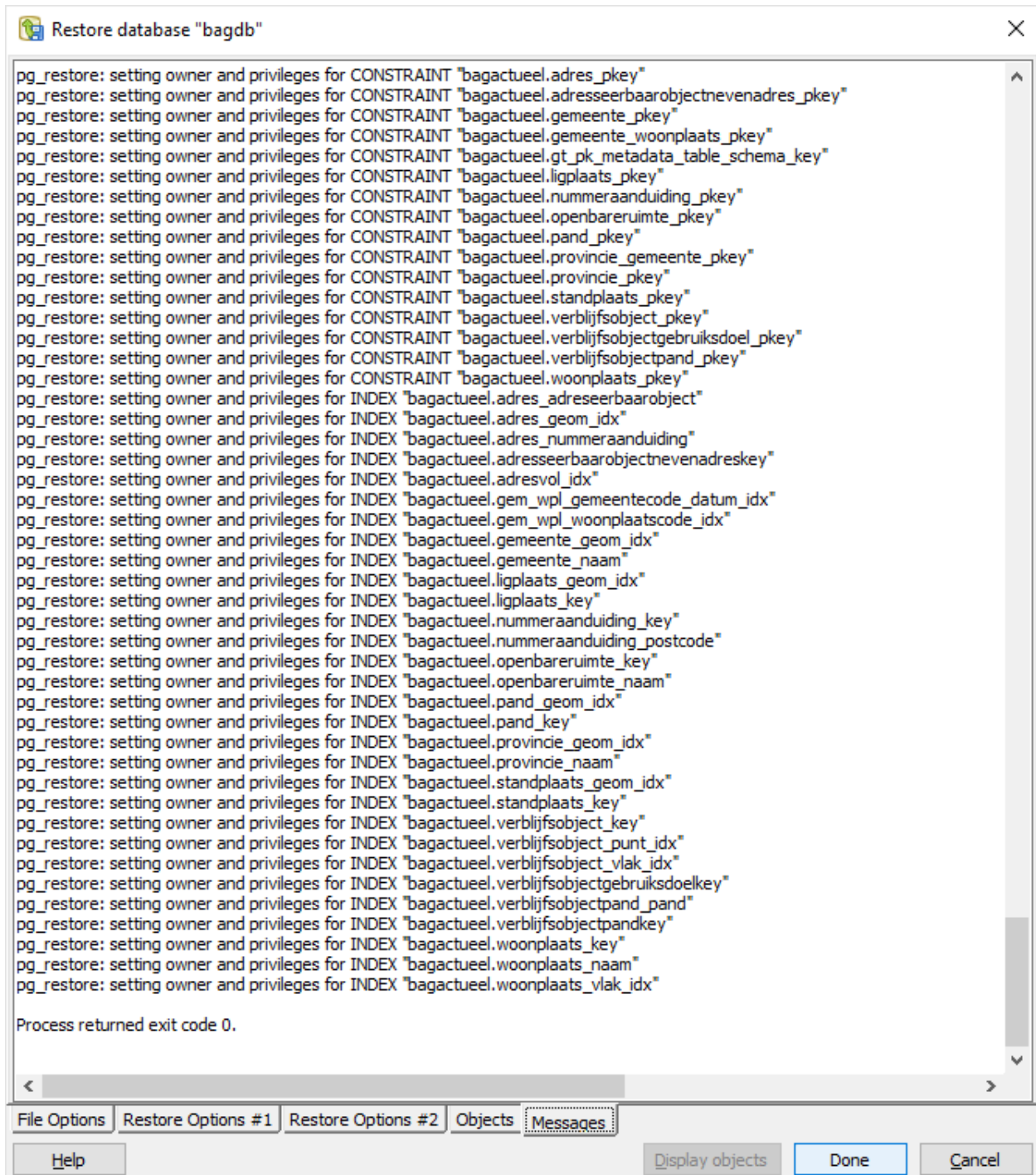


De restore opties:

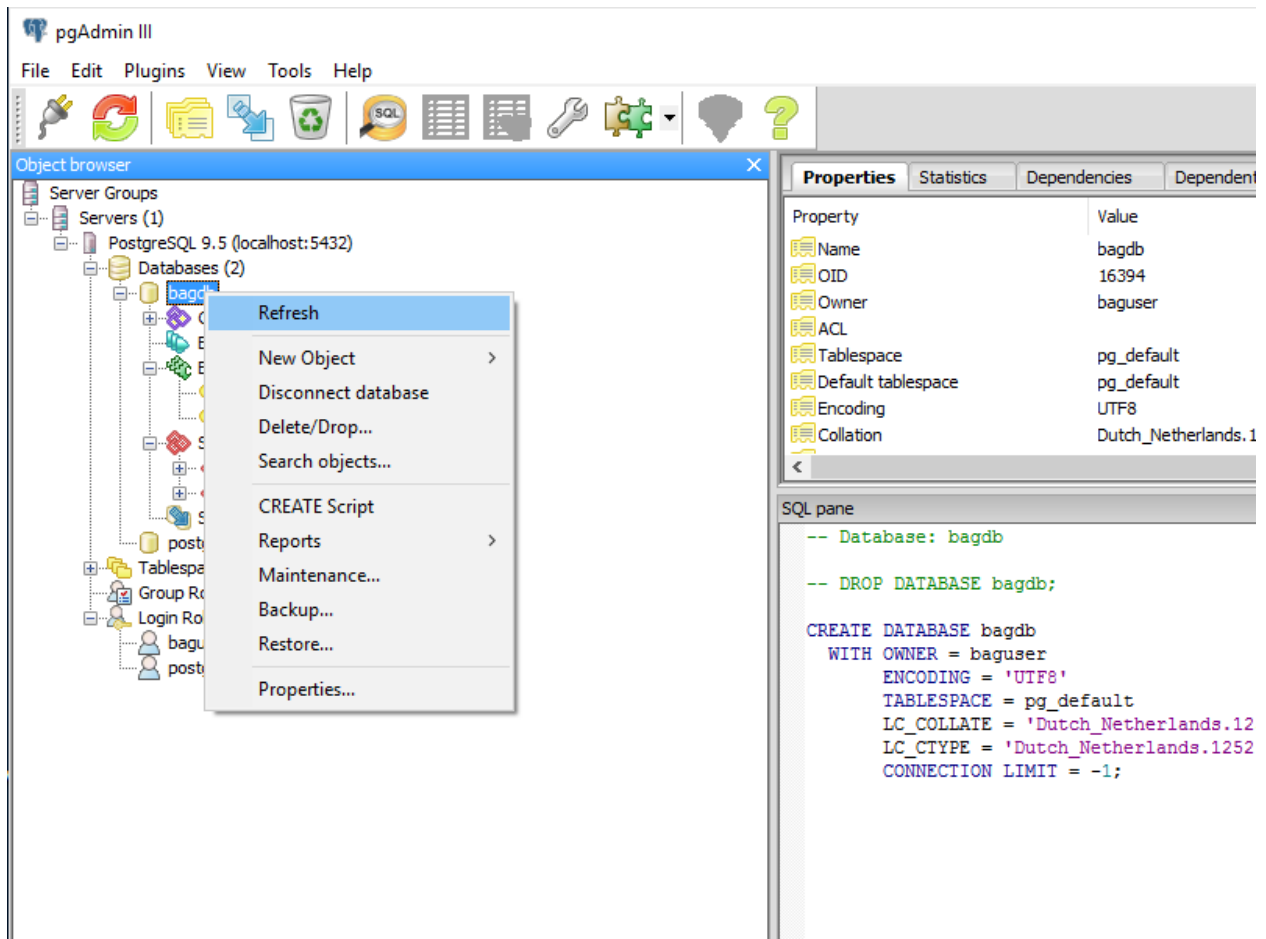


Kies 'Restore' en het terugzetten van de dump zal beginnen.

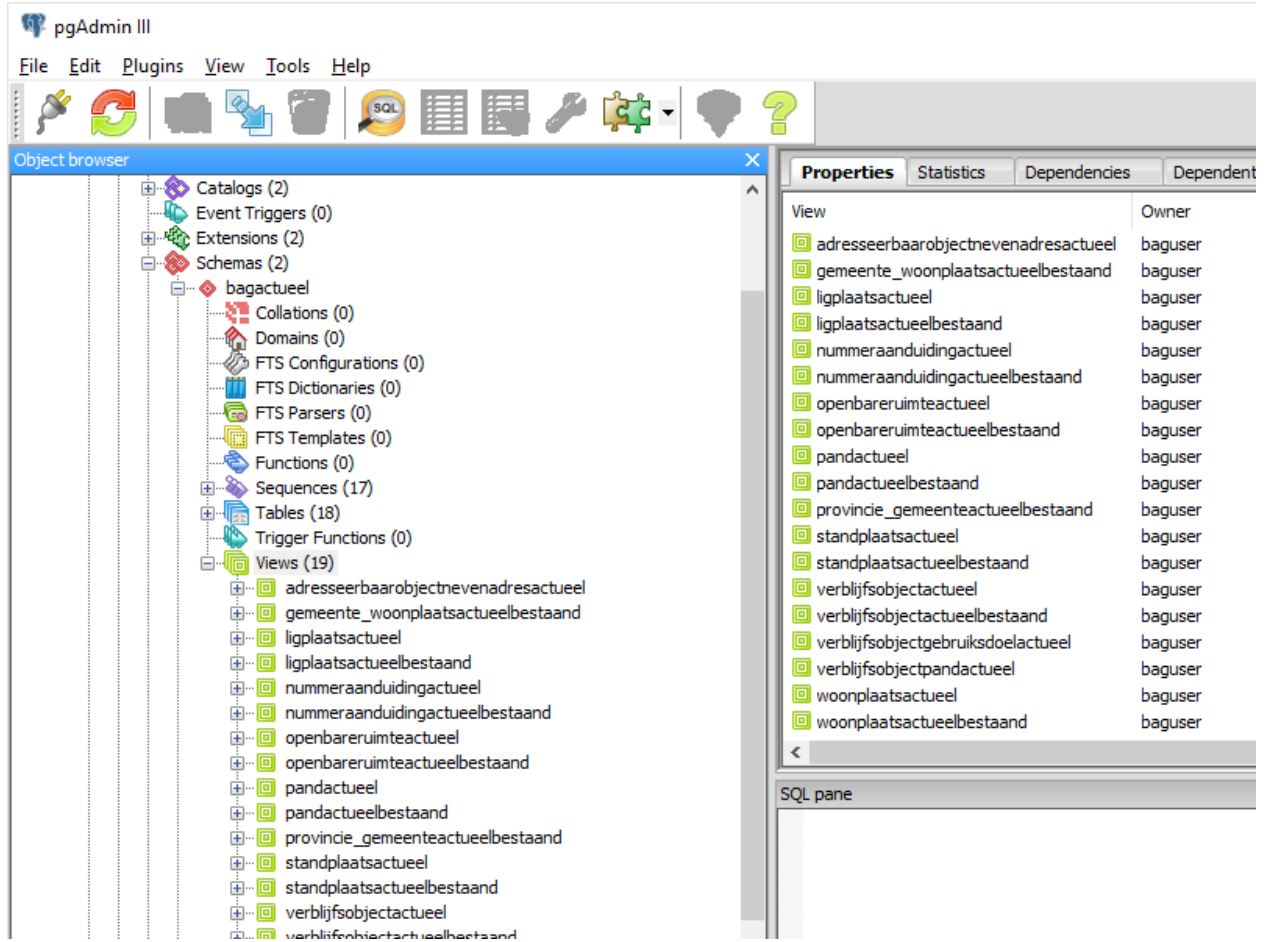
Nu is het herstelproces van de BAG database begonnen. Dit kan, afhankelijk van de capaciteit van de computer, enige tijd in beslag nemen (kwartier/half uur). Aan het eind moet 'Proces returned exit code 0' zichtbaar zijn, dat aangeeft dat het terugzetten gelukt is.



Kies eventueel een 'Refresh' om de navigatiestructuur te verversen, zodat alle herstelde data zichtbaar wordt.



De database is klaar voor gebruik:



De data bevindt zich in het schema 'bagactueel'. Houd hiermee rekening als je de BAG-data gebruikt in bijv. QGIS of een andere applicatie.

Hieronder links naar relevante zaken rond het project.

- NLEextract Downloads via Geotoko.nl - <https://geotoko.nl>
- NLEextract Downloads Metadata via Geocatalogus.nl - <https://geocatalogus.nl>
- GDAL/OGR, <https://gdal.org>
- lxml, <https://lxml.de>
- INSPIRE, <https://inspire.ec.europa.eu/>
- PostGIS/PostgreSQL, <http://postgis.net/>
- PDOK <https://pdok.nl>





## CHAPTER 10

---

### Contact

---

Onze website [www.nlextract.nl](http://www.nlextract.nl) biedt de meeste informatie over alle zaken rond het project.

Daar staan ook de [mensen](#) vermeld die werken aan NLEExtract.

Er is ook een mailing lijst voor vragen/opmerkingen en ontwikkelingen: <https://groups.google.com/group/nlextract>

Voor commerciële ondersteuning zoals gespecialiseerde extracties, downloads en andere services zie: <http://www.nlextract.nl/issues> of neem contact op met de volgende aanbieders:

- Just van den Broecke - via <https://justobjects.nl>
- Frank Steggink - via <https://steggink.it>
- OpenGeoGroep - via <https://opengeogroep.nl>



# CHAPTER 11

---

## Indices and tables

---

- `genindex`
- `search`