



kadaster



Locatieserver

Frank Steggink
Gerwin Hulsteijn

Agenda

- Introductie
- Deel 1:
 - Nieuwe datasets
 - Reverse geocoderen
 - Tips en tricks
- Deel 2: van Docker containers naar Kubernetes
- Vragen

Frank Steggink

- 18 jaar software-ontwikkelaar
- 18 jaar ervaring in de geo-sector
- 3 jaar ZZP-er
- Ingehuurd o.a. door Kadaster, Politie, IFV en Sweco
- Open source: NLExtract, ondersteuning OpenTopo



Gerwin Hulsteijn

- 17 jaar softwareontwikkelaar
- 4 jaar hosting architect
- 0,5 jaar ervaring in de geo-sector
- Ervaring in de telecom, notariaat en de politie
- Sinds begin dit jaar in dienst bij het Kadaster in het team PDOK

Historie Locatieserver

- Ontwikkeld bij Politie ter vervanging van een eerdere geocodeerservice
- Code overgedragen aan PDOK, samenwerkingsovereenkomst
- PDOK:
 - Locatieserver 2.0 (api v2): alleen BAG, opvolger van de geocoder, release december 2016
 - Locatieserver 3.0 (api v3): incl. Kadastrale percelen + koppeling met adressen, release juni 2017
 - Locatieserver 4.0 (api v3): huidige versie, release mei-juli 2018

Belangrijkste wijzigingen 2018

- Data:
 - NWB wegen en hectopunten
 - Appartementsrechten
 - CBS wijken en buurten
 - Bestuurlijke grenzen
 - Waterschapsgrenzen
- Functionaliteit: reverse geocoderen, momenteel testversie beschikbaar
- Diverse bugs opgelost

Nationaal Wegenbestand (NWB)

- Wegen: gekoppeld aan BAG openbare ruimtes indien match op naam, anders toegevoegd (bijv. rijkswegen)
- Hectopunten: ontsloten als nieuw objecttype (*hectometerpaal*)

Highlights:

- Goede samenwerking met Rijkswaterstaat
- Kwaliteitsslag voor NWB zelf (BAG compliant, extra controles)
- Veruit meeste openbare ruimtes die een locatie hebben (98,3%) kunnen gekoppeld worden aan het NWB

Appartementsrechten

- Kadastrale percelen waren in 2017 al toegevoegd
- Appartementsrechten als nieuw objecttype (*appartementsrecht*) toegevoegd en gelinked vanuit adressen (attribuut *gekoppeld_appartement*)
- Opbouw: string met gemeentecode, sectie, perceelnummer, “A” en volgnummer
- Herkomst data: Kadaster (Basisregistratie Kadaster, BRK)
- Nog geen plannen om adres/perceel- en adres/appartement-koppelingen als aparte download beschikbaar te stellen

Adres zoeken bij perceel/appartement

- Van perceel naar adres:
`http://geodata.nationaalgeoregister.nl/locatieserver/suggest?fq=*&q=gekoppeld_perceel:<gem>-<sectie>-<perceel>&fl=*`
- Van appartement naar adres:
`http://geodata.nationaalgeoregister.nl/locatieserver/suggest?fq=*&q=gekoppeld_appartement:<gem>-<sectie>-<perceel>-A-<nr>&fl=*`

CBS wijken en buurten

- Twee nieuwe objecttypen: *wijk* en *buurt*
- Voorbeeld: alle wijken van Apeldoorn:
http://geodata.nationaalgeoregister.nl/locatieserver/suggest?q=type:wijk%20AND%20gemeentecode:0200&fq=*&fl=identificatie,wijknaam,weergavenaam&rows=20

Bestuurlijke grenzen en waterschapsgrenzen

- Nieuwe objecttypen: *provincie* en *waterschap*
- Objecttype *gemeente* nu vanuit bestuurlijke grenzen i.p.v. geaggregeerd vanuit de BAG
- Voorbeeld: alle waterschappen:
http://geodata.nationaalgeoregister.nl/locatieserver/suggest?q=type:waterschap&fq=* &fl=identificatie,weergavenaam&rows=25

Reverse geocoderen

- Doel: zoeken van objecten in Locatieserver o.b.v. afstand tot een opgegeven punt (o.b.v. *X/Y* of *lat/lon*)
- Extra filtering o.b.v. *type* en/of een maximale afstand (*distance*)
- Alle objecttypen zijn beschikbaar
- Applicatie is in Kotlin geschreven
- Vergelijkbare parameters als Solr (rows, start, wt, fl, fq)
- Momenteel beschikbaar voor beta test
- Meer informatie:
<https://forum.pdok.nl/t/nieuw-reverse-geocoder-api-beschikbaar-om-te-testen-pdok-locatieserver/1544>

Inhoud data Locatieserver

Type	ID-prefix	Geometrie	Standaard in suggest/free	Bron	Aantal objecten
provincie	pro	vlak	nee	Bestuurlijke grenzen	12
gemeente	gem	vlak	ja	Bestuurlijke grenzen	380
woonplaats	wpl	vlak	ja	BAG	2.501
weg	weg	lijn of punt	ja	BAG + NWB	266.761
postcode	pcd	punt	ja	BAG (afgeleid)	466.253
adres	adr	punt	ja	BAG	9.179.190
perceel	pcl	vlak	nee	BRK	7.879.539
appartementsrecht	app	multipunt	nee	BRK	1.354.334
hectometerpaal	hmp	punt	nee	NWB	148.530
waterschap	wsa	vlak	nee	HWH	21
wijk	wij	vlak	nee	CBS	3.143
buurt	buu	vlak	nee	CBS	13.386

Query aantallen: http://geodata.nationaalgeoregister.nl/locatieserver/suggest?fq=*&facet=true&facet.field=type&rows=0

Solr faceting queries

- Handig om aantallen per attribuut te bepalen
- Combinatie van q, fq en attribuut (facet.field)
- Voorbeelden:
 - [Aantal wegen per bron](#)
 - [Aantal adressen per woonplaats \(code\)](#)
 - [Aantal percelen per kadastrale gemeente](#)
 - [Aantal hectometerpalen per wegnummer](#)
- Attributen: facet=true, facet.field (1 per veld), facet.limit=-1, rows=0, facet.sort=count/index (o.b.v. naam)
- Meer informatie: <https://wiki.apache.org/solr/SimpleFacetParameters>

Scoreberekening Apache Solr

- Voorkeur voor korte namen
- Bij meerdere zoektermen hogere score bij meer matches
- Boosting parameters
- Berekening zichtbaar bij *debug=true*
- Meer informatie:
https://lucene.apache.org/solr/guide/6_6/the-extended-dismax-query-parser.html

(Ik denk dat ik deze sheet ga schrappen)

Locatieserver stack

- Solr
- Zookeeper
- Conductor
- Hadoop

Wat is Docker?

Technisch: Virtualisatie op kernel niveau. Virtuele omgeving bestaande uit een proces en een netwerkruimte.

Praktisch: Een manier om dependencies en applicaties te combineren. Het mogelijk maken microservices bestaande uit een stukje logica en bijbehorende dependencies te creëren.

Pets vs cattle vraagstuk



Voordelen cattle

- Schaalbaar
- Uptime
- Minder resources!?
- Verminderde beheerlast
- Kosten

Zijn docker containers cattle?

Niet zonder meer. Om van pets naar cattle te gaan moet aan de volgende voorwaarden worden voldaan:

- Containers zijn vervangbaar
 - Geen communicatie rechtstreeks naar containers
 - Omschrijf de kaders, niet de concrete situatie (wat, niet hoe)

Docker met beperkte orkestratie



Kubernetes

Bij PDOK gaan we over **van** een omgeving waarin de docker servers **silo's** zijn **naar** een omgeving waarin de docker servers zijn weg geabstraheerd in een **cluster**.

Hoe helpt Kubernetes?

- Containers worden niet meer rechtstreeks aangeroepen.
 - Kubernetes maakt gebruik van een **Deployment**, **ReplicaSet**, **DaemonSet** of **StatefullSet**
- Yaml files worden gebruikt om de gewenste situatie aan te geven. Kubernetes doet de rest.
 - **Rolling-upgrade**
 - **Autoscaling**
 - **Health checks**

ETL

- Dagelijks complete verwerking van de datasets
- Naar een delta verwerking

Vragen





kadaster