

DEEL 2

Dataloket: instrument voor de datagedreven organisatie

Al uw data eenvoudig vindbaar via één gebruiksvriendelijk zoekportaal

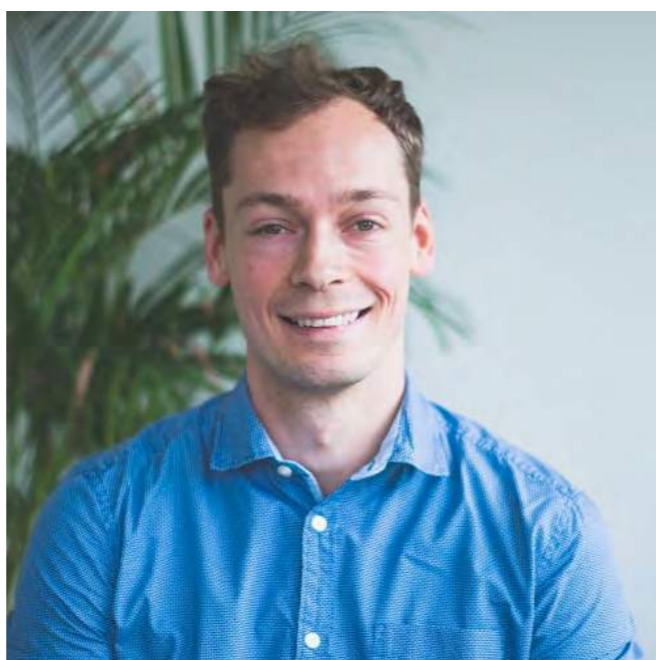
Als je het whitepaper *Dataloket: instrument voor de datagedreven organisatie deel 1* nog niet hebt gelezen, dan adviseren wij je om dat eerst te doen. Hierin geven wij uitleg over de toegevoegde waarde van het Dataloket voor een organisatie, en over de functionaliteit van de oplossing.

Deel 1 lezen >

Als je nu verder leest, dan ben je zelf (en/of een van jouw collega's geïnteresseerd in de mogelijkheden die het Dataloket jouw organisatie kan bieden voor het vinden, delen en hergebruiken van data (services) en informatieproducten. Je vervult mogelijk een rol als (technisch) projectleider, architect of beheerder van een dataplatform. Je wilt graag meer weten over hoe de Dataloket oplossing werkt, om te beoordelen of de oplossing voldoet aan de eisen en architectuurprincipes van jouw organisatie. Het doel van dit whitepaper is om een eerste verdieping te bieden, waarin een aantal veel gestelde vragen over de werking (en de onderliggende inrichtingsprincipes) van het Dataloket worden behandeld.

Als je na het lezen van het whitepaper nog openstaande vragen hebt, behandelen wij die vragen uiteraard graag in een gesprek met een van onze consultants.

Kevin Otjes,
Managing Partner BI



06 - 42 53 31 33
kevin.otjes@kbenp.nl

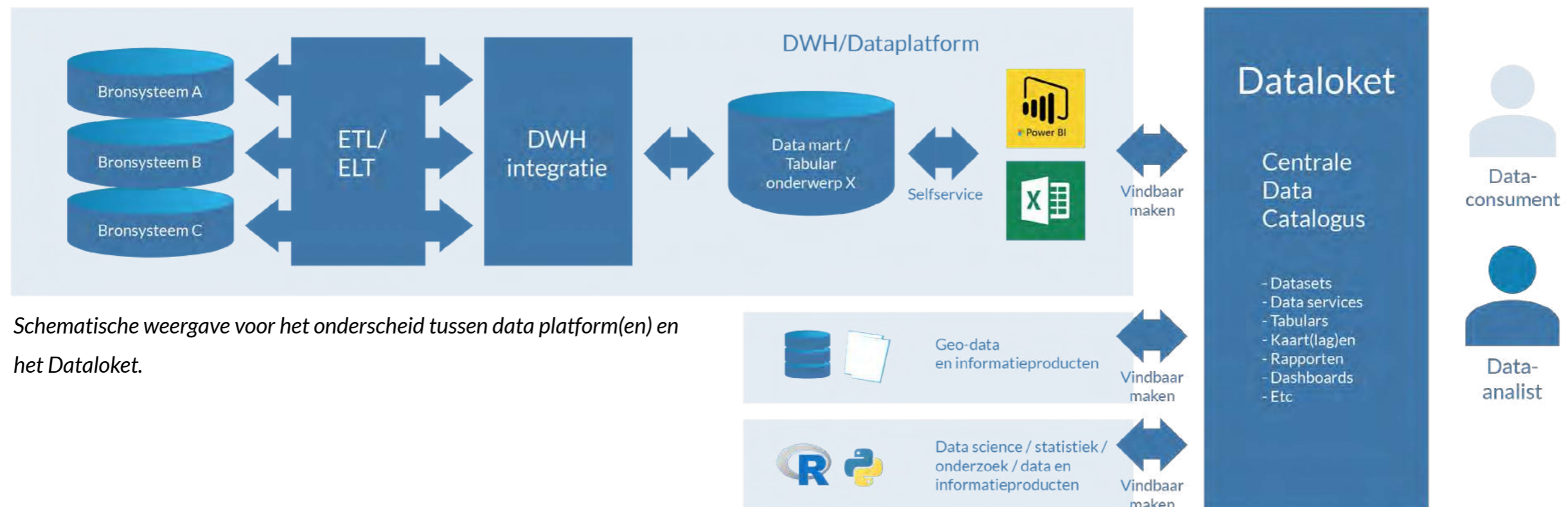
Inhoudsopgave

- 1 Wat is het verschil tussen het Dataloket en een Data Platform? 4
- 2 Hoe kom je van data aan de bron tot het Dataloket (en andersom)? 6
- 3 Hoe is de beveiliging van data geregeld in het Dataloket? 10
- 4 Hoe speelt het Dataloket een rol bij Open Data en standaardisatie? 12
- 5 Hoe kan het Dataloket mij helpen te sturen op het gebruik van data assets? 14

Wat is het verschil tussen het Dataloket en een Data Platform?

Het Dataloket is naar zijn aard een data catalogus. Een data catalogus is zeer geschikt voor het bundelen van dataservices, data sets en informatieproducten die op verschillende platformen zijn gecreëerd in een index. Het bundelen in een index wil zeggen: het voor een gebruiker samenbrengen, presenteren en standaardiseren van de beschrijvende metadata van de verschillende items, met als doel het vereenvoudigen van het vinden van relevante items en het vervolgens opvragen/raadplegen van de gevonden relevante items.

Wat is het Dataloket daarmee niet? Een Data platform, met als primair doel om de data vanuit verschillende bronsystemen inhoudelijk te verbinden en te combineren (via een ELT/ETL proces) tot een nieuw informatieproduct. Het Dataloket is zeer behulpzaam voor een data analist om relevante data services of data sets te vinden, maar in zichzelf is het niet de omgeving om een nieuw informatieproduct mee te maken.



Schematische weergave voor het onderscheid tussen data platform(en) en het Dataloket.

Waarom niet? Waarom hebben wij vanuit ons bedrijf gekozen voor het richten op een data catalogus functie? Hieronder onze kernredenen:

Weinig organisaties hebben het vermogen om alle informatie-producten, dataservices en datadiensten vanuit deze data platformen voor de eindgebruiker samen te brengen als ware het één geheel.

Bijna iedere grote organisatie heeft al min. 3 tot 4 dataplatformen in gebruik, geschikt voor het maken van informatieproducten voor geografische data (zoals: ArcGis, Q-Gis, GeoServer) of voor administratieve data (zoals: Power BI, Cognos, QlikSense, Tableau als presentatielaag boven een vorm van een Data Platform-oplossing).

Ons Dataloket past uitstekend in een situatie waarin er tenminste voor de middellange termijn meerdere data platformen in gebruik zullen zijn. En zelfs al is er maar één dataplatform om uit te lezen, het Dataloket biedt een goed ontkoppelpunt om data voor verschillende doelgroepen op een gebruiksvriendelijke manier vindbaar te maken.

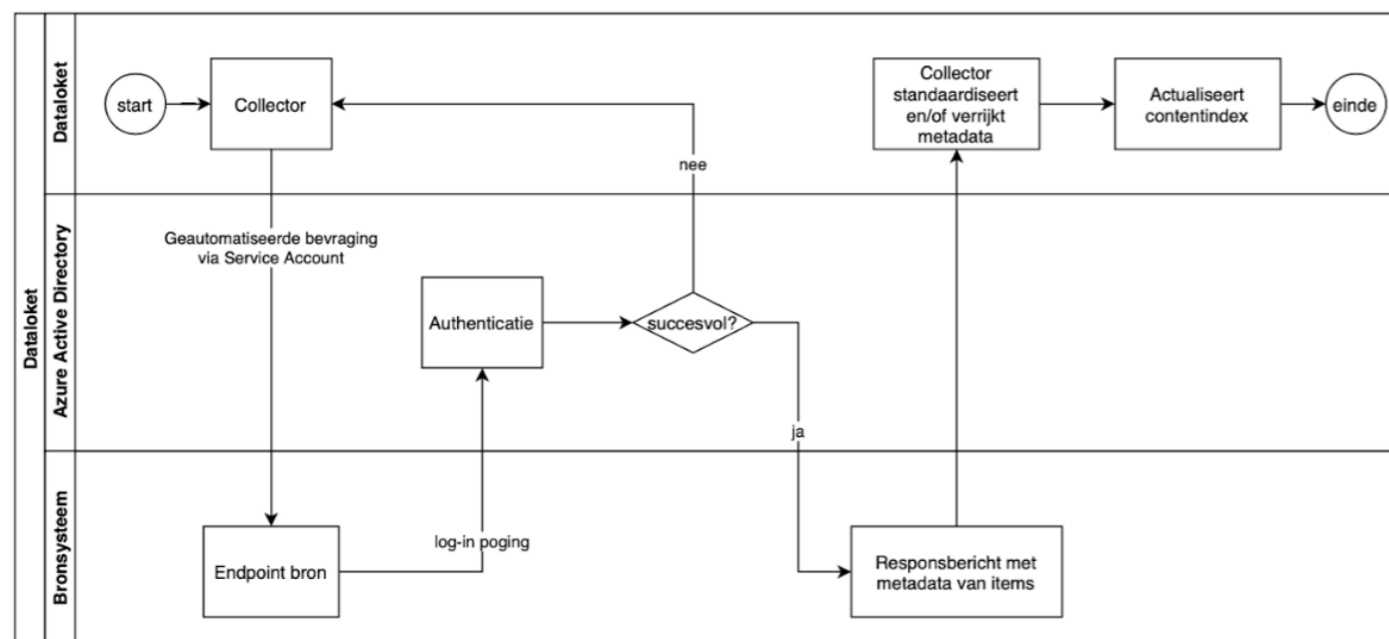
Hoe kom je van data aan de bron tot het Dataloket (en andersom)?

Een belangrijk onderdeel van hoe het Dataloket werkt, is door de bronsystemen (de data platformen) geautomatiseerd uit te lezen voor wat betreft de metadata van de aanwezige datasets, dataservices en informatieproducten. Voor verder leesgemak noemen wij deze opsomming van items uit een data platform in de rest van dit whitepaper: de “data assets”. Voor de ontsluiting van data assets via het Dataloket onderscheiden wij drie lagen:

1. **Het bronsysteem** – het data platform met daarin de data assets;

2. **De Active Directory (het Identity & Access Management systeem)** - Hierin is vastgelegd wie waarvoor is geautoriseerd, welke gebruikers toegang hebben tot specifieke data en informatieproducten;

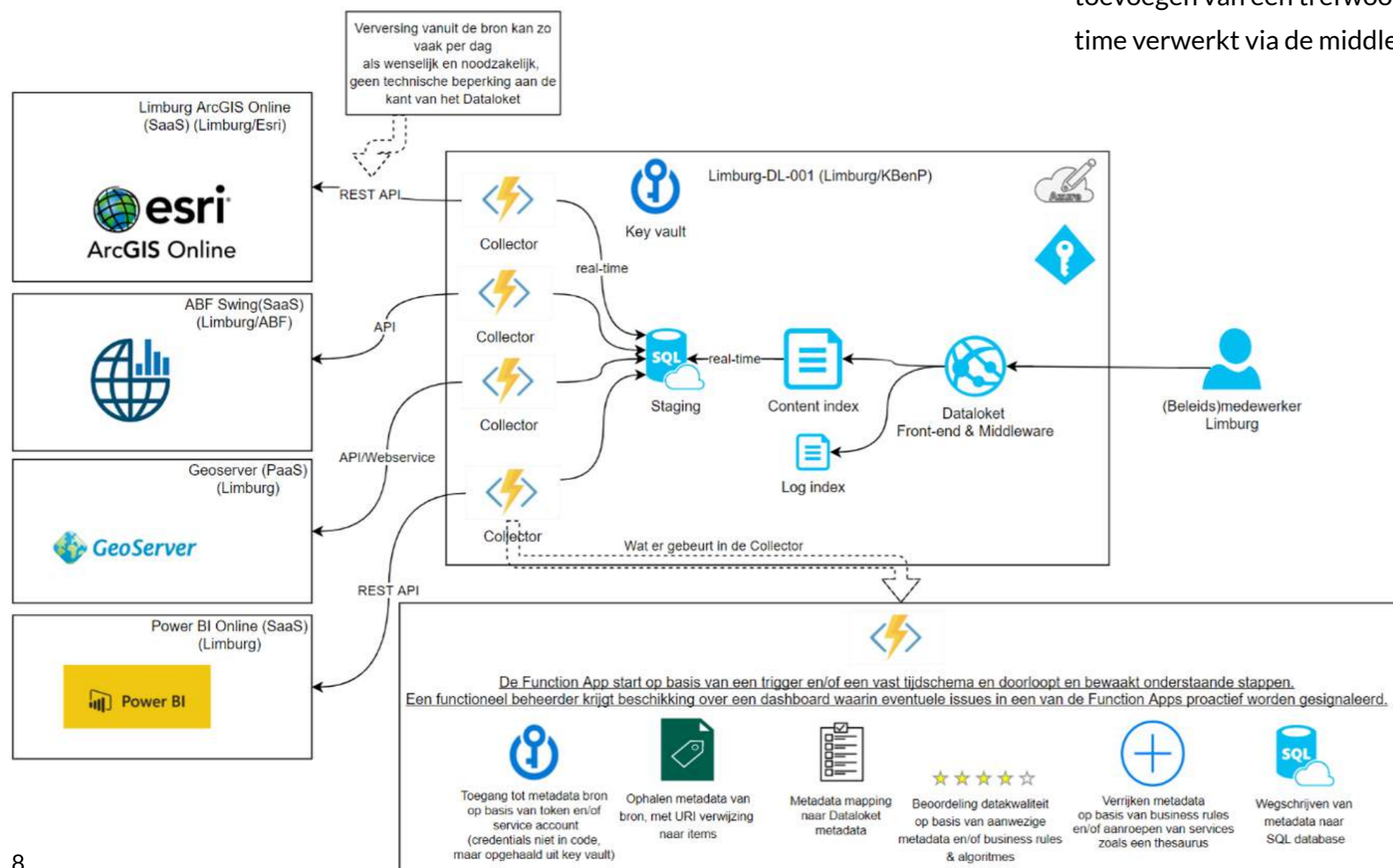
3. **Het Dataloket** - De gebruikersinterface (front-end & middleware) waarbinnen data assets beschikbaar worden gemaakt voor gebruikers met voldoende permissierechten aan de bron. In onderstaand schema hebben wij de functionele procesflow voor het ophalen van de metadata naar de index inzichtelijk gemaakt.



Schematische weergave functionele procesflow voor het ophalen van de metadata naar de index.



Hieronder een voorbeeld van het schematische verloop van de laadprocessen van het Dataloket bij de Provincie Limburg voor het ophalen van de metadata. Het "Collector" component vervult hierin zowel de rol van orchestrator (bewaker van de juiste procesuitvoering) als van het inhoudelijk verrijken en standaardiseren van de metadata in een modulair "wasstraat-principe". De genoemde connectoren in onderstaand diagram zijn een aantal (niet-uitputtende) voorbeelden van bronsystemen waarmee wij ervaring hebben in de koppeling. Wij kunnen in de basis ieder dataplatform dat beschikt over een API of webservice uitlezen en wij kunnen daarnaast ook diverse vormen van fileshares, websites (HTML) en database-oplossingen ontsluiten als daaraan behoefte is.



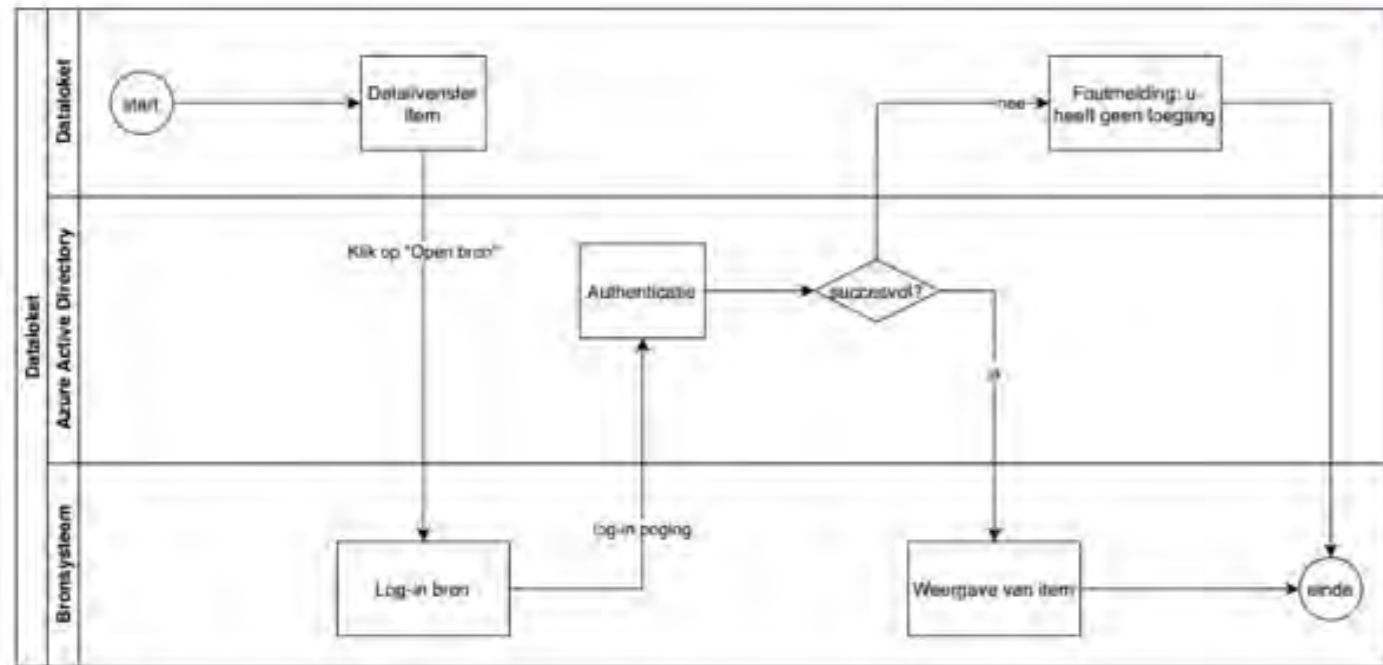
De bronsystemen kunnen op ieder gewenst moment en met iedere gewenste frequentie worden gesynchroniseerd, zodat eindgebruikers via het Dataloket altijd de beschikking hebben over de meest actuele data. Synchronisatie van de bronsystemen kan worden uitgevoerd op afgesproken momenten (ieder uur, dagelijks, wekelijks, of anders) en/of wanneer informatie aan een bronsysteem is toegevoegd of gewijzigd. De frequentie waarmee het Dataloket de data en informatieproducten uit deze bronsystemen ophaalt, is een kwestie van maatwerk op basis van de wensen vanuit jouw organisatie. Daarvoor gelden bij ons geen technische beperkingen. Het proces van indexeren vormt slechts een beperkte belasting voor de bronsystemen, aangezien er uitsluitend metadata wordt opgehaald. Acties van gebruikers, zoals het instellen van een favoriet, het toevoegen van een trefwoord of het uploaden van een document, worden real-time verwerkt via de middleware en gelogd.

Schematische weergave van de laadprocessen van het Dataloket bij de Provincie Limburg voor het ophalen van de metadata.

Hoe is de beveiliging van data geregeld in het Dataloket?

Regelmatig krijgen wij de vraag of het Dataloket de toegang tot data assets wijzigt. Het antwoord daarop is: nee. Het Dataloket blijft ten allen tijden de rechten vanuit het bronsysteem hanteren. Wel voegt het Dataloket als een extra metadata veld een data classificatie voor de vertrouwelijkheid (op basis van geautomatiseerde beslisregels of een beslissing door een data steward) toe, welke kan worden gebruikt voor het op voorhand wegfilteren van bepaalde resultaten voor gebruikersgroepen met onvoldoende permissies. Hierdoor is ook het onderscheiden van openbare data en bedrijfsvertrouwelijke data in de gepresenteerde zoekresultaten eenvoudig te doen.

Als procesflow ziet het raadplegen van een data asset door een gebruiker van het Dataloket er als volgt uit (waarin er gewerkt wordt met een single-sign on principe, waarbij onze middleware namens de ingelogde gebruiker om toegang vraagt):



Schematische weergave procesflow raadplegen van een data asset door een gebruiker van het Dataloket.



Hoe gaat het Dataloket om met de opslag van de data van jouw organisatie?

Het Dataloket verzamelt alleen de metadata, al de inhoudelijke data vanuit jouw organisatie blijft staan aan de bron en verlaat daarmee nooit jouw organisatie. De metadata die wordt opgeslagen is beveiligd in onze Microsoft Azure cloud-omgeving, waarbij iedere Dataloket implementatie gesegmenteerd is van alle andere instanties van het Dataloket. Je hebt als klant zowel de beschikking tot een productie omgeving als een pre-productie-omgeving, die volledig exclusief voor jouw organisatie is ingericht. Bij het eventueel aflopen van onze dienstverlening worden jouw omgevingen, inclusief alle opgeslagen data, permanent verwijderd. Wij koppelen de Dataloket omgeving met het Identity & Access Management systeem van jouw organisatie, waardoor deze rollen en rechten bepalend zijn voor het stellen van toegangen en andere permissies. Wij beschikken over het ISO:27001 certificaat voor Informatiebeveiliging en besteden veel aandacht aan data security bij de ontwikkeling en levering van onze producten.

Hoe speelt het Dataloket een rol bij Open Data en standaardisatie?

Standaardisatie vinden wij een belangrijk thema. Wij hebben het Dataloket gemaakt om als een volledige SaaS-oplossing te functioneren die volledig toegankelijk is via de webbrowser en geen installatie van programmatuur vraagt vanuit jouw organisatie. De standaardfunctionaliteit is volledig configurabel, het is per organisatie alleen nog maar een kwestie van het toepassen van huisstijl, thema-indeling en het leggen van de koppeling naar bronsystemen, de Active Directory en eventueel een servicedesk-systeem van jouw organisatie (optionele koppeling voor het direct vanuit het Dataloket kunnen indienen van tickets voor het updaten, aanpassen of beschikbaar stellen van data assets).

Standaardisatie passen wij ook toe op metadata. Wij hanteren daarbij DCAT als de basis van ons metadatamodel. Het Dataloket kan worden ingezet om direct open data beschikbaar te maken voor een doelgroep, maar ook om open data aan te dragen naar een ander platform (zoals hieronder te zien data.overheid.nl). In beide scenario's blijkt de data aan de bron staan en dient het Dataloket voor het vindbaar maken, structureren van de metadata en het bijhouden van daadwerkelijk gebruik (via logging).

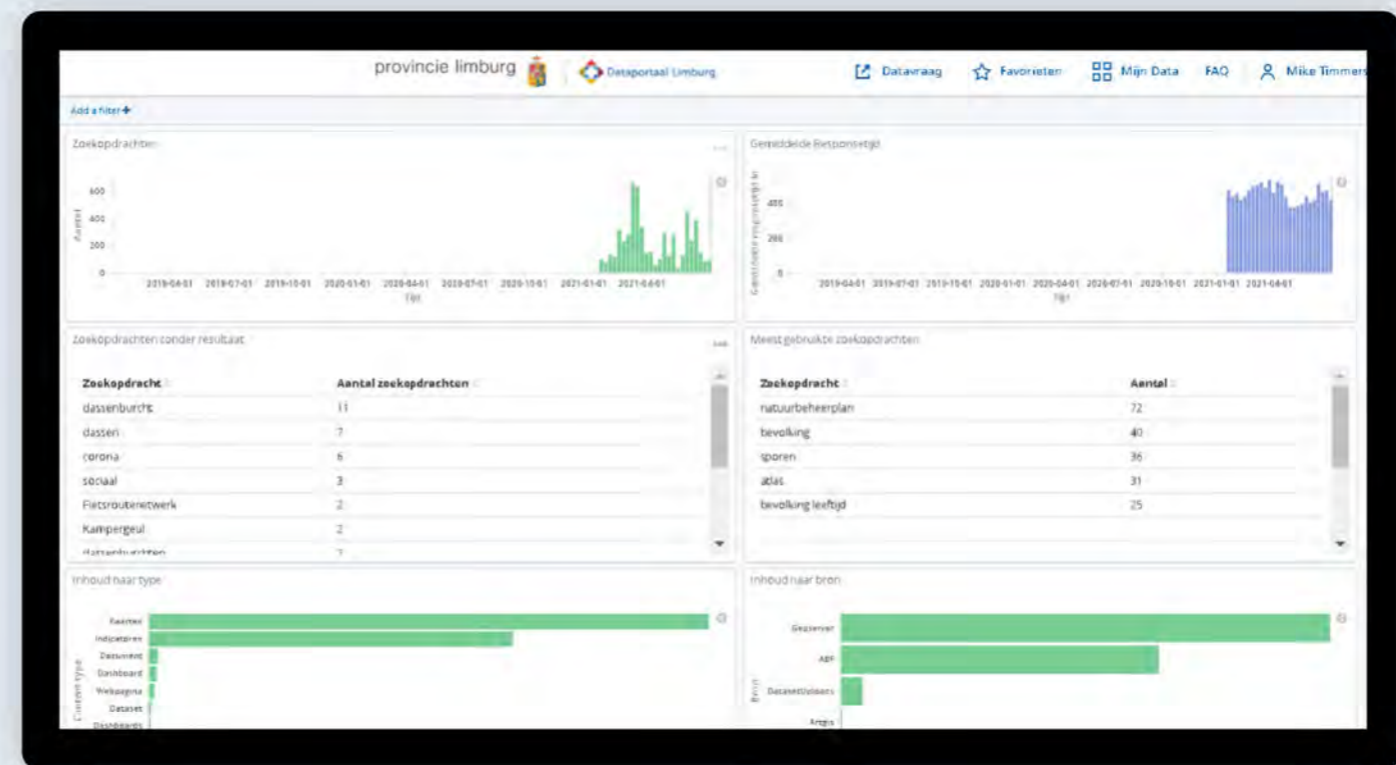


Hoe kan het Dataloket mij helpen te sturen op het gebruik van data assets?

Een waardevol aspect van het Dataloket is het bijhouden van logdata over de zoekopdrachten en andere acties die zijn uitgevoerd. Hiermee krijg je veel inzicht in de vraag naar informatieproducten en data binnen jouw organisatie. Vanuit een levenscyclusbeheer perspectief zijn ook veel beheerders van data platformen geïnteresseerd in welke van hun data assets wel en niet gebruikt worden binnen de organisatie.

Meerdere van onze klanten (zoals bijvoorbeeld de provincies Zuid-Holland en Limburg) gebruiken de informatie over daadwerkelijk gebruik van data assets vanuit het Dataloket voor hun eigen levenscyclusbeheer van zowel individuele data assets, als van gehele (legacy) data platformen. Het Dataloket past als zodanig uitstekend in een groter beheerplan om een doelarchitectuur te realiseren en legacy (verouderde, niet-doelarchitectuur) systemen geleidelijk uit te faseren door relevante content opnieuw op te bouwen in het doelplatform. Meerdere van onze klanten gebruiken deze inzichten ook voor het sturen op de adoptie van data-gedreven werken binnen hun organisatie, doordat zij goed zicht hebben op welke afdelingen het meest en minst actief gebruik maken van de data assets. Het inzicht in welke zoekopdrachten geen resultaat geven is ook interessant, om nieuwe behoeften naar data assets te ontdekken bij de gebruikers van het Dataloket. Wij helpen onze klanten bij het benutten van de sturingsinzichten van het dashboard, waardoor het Dataloket een belangrijk inzicht kan bieden voor de Digitale Transformatie van de organisatie. Wij nemen dit standaard mee als onderdeel van ons implementatieplan.

In onderstaande schermafbeelding wordt ons standaard dashboard voor zoekgedrag getoond. Deze dashboardfunctionaliteit is voor beheerders in te zien binnen de Dataloket webapplicatie.



OVER ANALYZE

Onder de naam Analyze ontwikkelt en implementeert KBenP datagedreven applicaties, zodat jouw klanten en medewerkers altijd en overal beschikken over relevante informatie. Zo hebben wij gestandaardiseerde producten ontwikkeld voor Smart City projecten en het dataplatform Dataloket.

Heb je nog vragen die niet zijn beantwoord? Of ben je enthousiast geworden over een vervolgstap? Neem dan contact met ons op voor een vrijblijvend intakegesprek.



info@kbenp.nl



070 – 300 06 84



<https://www.linkedin.com/company/kbenp>

[Demo plannen >](#)



ANALYZE