

CDL Solution Architectuur

versie 2.0.0

Inhoudsopgave

Inhoudsopgave	1
1. Inleiding	2
2. Beschrijving van de oplossing voor CDL	2
2.1 Kenmerken oplossing CDL	2
3. Componenten van het CDL	3
4. CDL Connector	4
5. CDL Centraal	5
5.1 Aanleveren	5
5.2 Valideren	5
5.3 Samenvoegen	6
5.4 Beschikbaar stellen	6
IMWA	6
INSPIRE	7
GWSW	7
Uitleveren	7
5.5 Beheren	8
6. Bijlagen	8
6.1 Afkortingen en acroniemen	8

1. Inleiding

Het CDL draagt bij aan de realisatie van de doelstellingen van de waterschappen op het gebied van het leveren van ruimtelijke informatie. Door informatie van individuele overheden sectoraal te uniformeren, te bundelen en vindbaar te maken wordt het voor klanten (informatie-afnemers) eenvoudiger om deze informatie te gebruiken. De scope van het project CDL beperkt zich tot het realiseren van een gezamenlijke voorziening die data bij de waterschappen verzamelt en als open data beschikbaar stelt aan alle afnemers. Dit is inclusief de wettelijke verplichting voor het leveren van INSPIRE gegevens.

In de Project Start Architectuur van het CDL (versie 1.2) is de verbinding gelegd tussen de wensen en eisen van het project en de architectuur. Het is een vertaling van de beleidslijnen en architectuurprincipes naar projectkaders. In de PSA is de context van de oplossing beschreven, niet de oplossing zelf. In dit document, de solution architectuur, wordt de daadwerkelijke oplossing beschreven van het CDL.

2. Beschrijving van de oplossing voor CDL

De oplossing voor het CDL is gebaseerd op herbruikbare componenten van meerdere vergelijkbare data en services oplossingen, waar de gestandaardiseerde aanlevering en ontsluiting via een API en de integratie van datasets van meerdere bronhouders de kern vormt. Dit hoofdstuk beschrijft de verschillende componenten die samen de solution architecture vormen in meer detail.

2.1 Kenmerken oplossing CDL

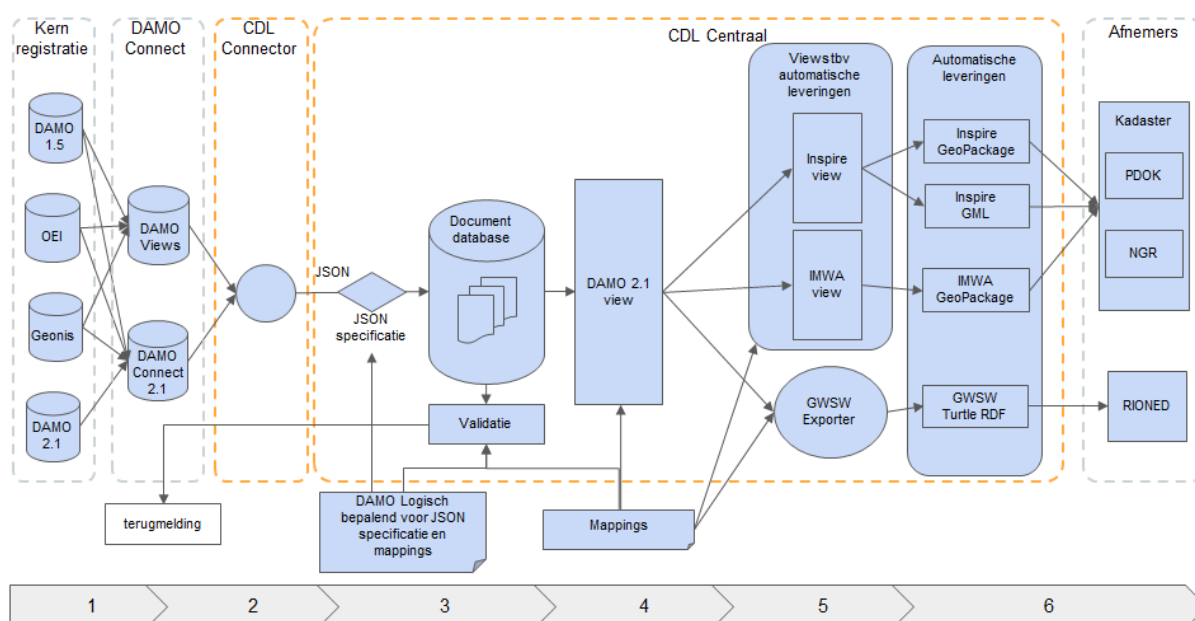
De algemene kenmerken van de oplossing voor CDL zijn:

- **Degelijkheid en hergebruik:** de oplossing is gebaseerd op kennis en componenten van bewezen vergelijkbare oplossingen zoals het Register Externe Veiligheid. Het CDL wordt gerealiseerd met dezelfde onderliggende standaard applicatiecomponenten en software-bibliotheken.
- **Professionele build-infrastructuur:** voor al onze applicaties hanteert Geodan een build infrastructuur voor het ontwikkelen en onderhouden van data en software. Hierdoor zijn wij in staat wijzigingen in zeer korte tijd goed gecontroleerd en goed getest door te voeren naar productie.
- **Schaalbaarheid, robuustheid en hoge beschikbaarheid:** voor het CDL wordt gebruik gemaakt van een schaalbare cloudinfrastructuur. Toekomstige uitbreidingen van datalagen of gebruikers (waterschappen) is hierdoor mogelijk.
- **Open Source componenten:** de oplossing bestaat uit uiteenlopende open source componenten.
- **Datamodel onafhankelijk:** een groot deel van de programmatuur is datamodel onafhankelijk, zoals het CDL Centraal. Hierdoor kunnen later relatief eenvoudig aanpassingen of uitbreidingen van DAMO worden doorgevoerd.
- **Hoge beschikbaarheid:** het CDL draait mee in de monitoring van voorzieningen met beschikbaarheidseisen waar de service levels voor het CDL ruim binnen vallen.

- Eenvoudige doelgerichte user-interface: een eenvoudige en sterk doelgerichte user-interface is de sleutel tot succes. Hoe eenvoudiger de user-interface, hoe lager de instapdrempel voor waterschappen.

3. Componenten van het CDL

De essentie van de Centrale Distributie Laag (CDL) bestaat uit gegevenslevering door de waterschappen op basis van DAMO, het ophalen en samenvoegen van deze gegevens door het CDL en het ter publicatie aanbieden van de landsdekkende gegevens aan de verschillende afnemers. Figuur 1 geeft een overzicht van het CDL in de CDL Keten. De oranje omcirkelde componenten zijn onderdeel van het CDL en worden in de volgende paragrafen in meer detail beschreven. De grijs omcirkelde componenten zijn geen onderdeel van het CDL, maar maken wel deel uit van de CDL Keten en hebben een relatie met het CDL.



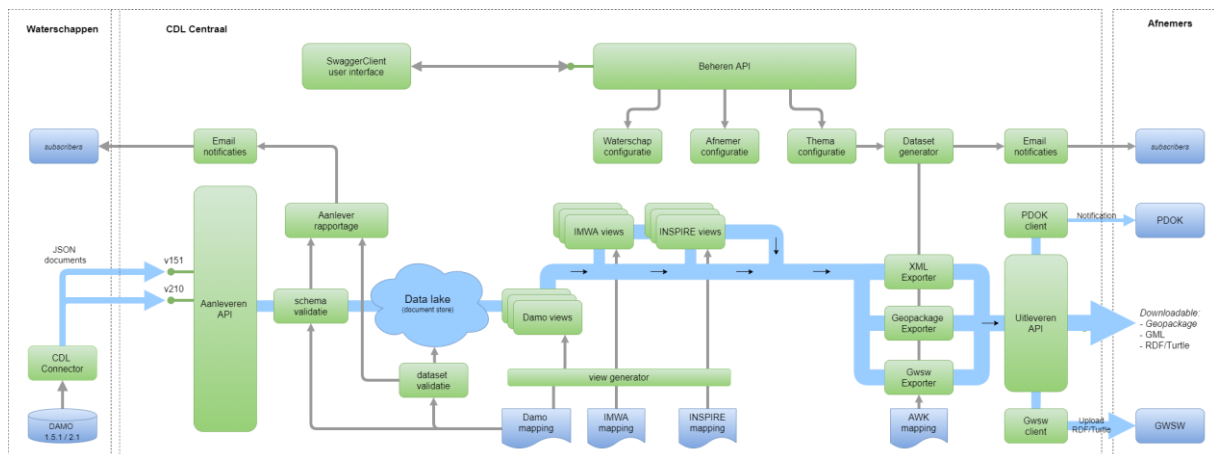
Figuur 1: overzicht CDL in de CDL Keten

De data stroomt als volgt door de CDL Keten:

1. Vanuit het CDL gekeken is de data van de waterschappen beschikbaar in DAMO 1.5.1 formaat via DAMO Views (_EVW) of in DAMO 2.1 formaat via DAMO Connect.
2. De connector ondersteunt beide versies van DAMO en stuurt de data in JSON formaat via de REST API naar CDL Centraal. Als DAMO Connect is geïmplementeerd bij de waterschappen hoeft de connector niet meer aangepast te worden bij nieuwe versies van DAMO.
3. De data wordt gevalideerd tegen de JSON specificatie (schema validatie) en opgeslagen in een document database. Daar vindt vervolgens nog een validatie plaats op basis van de DAMO mapping. De validatie resultaten van de schema validatie en de DAMO mapping worden samengevoegd in het upload report en teruggemeld aan het Waterschap en desgewenst ook aan Het Waterschapshuis.
4. CDL Centraal maakt op basis van de DAMO mapping een landelijk geconsolideerd DAMO 2.1 beeld van alle data in de database (de DAMO 2.1 view).

5. CDL Centraal maakt van de DAMO 2.1 view een INSPIRE en IMWA view die de basis vormen voor de bijbehorende automatische levering. Voor de GWSW data wordt geen aparte view gegenereerd. De mappings zijn bepalend voor de programmatuur.
6. Op basis van de INSPIRE view wordt een INSPIRE GeoPackage en INSPIRE GML gemaakt. De IMWA view vormt de basis voor de IMWA GeoPackage. De automatische leveringen voor INSPIRE en IMWA worden uitgeleverd aan PDOK. Na uitlevering aan PDOK zorgt Kadaster voor de publieke ontsluiting van deze data via PDOK en NGR.

De GWSW Exporter genereert op basis van de DAMO 2.1 view een Turtle RDF bestand conform de GWSW standaard welke wordt uitgeleverd aan RioNed. De GWSW mapping is bepalend voor de GWSW Exporter.



Figuur 2: overzicht solution architectuur

Figuur 2 geeft een overzicht van de solution architectuur met de verschillende applicatie componenten van het CDL.

4. CDL Connector

- De CDL Connector is een Java applicatie die lokaal wordt geïnstalleerd bij het deelnemende Waterschap en verbinding maakt met de Oracle database van het Waterschap.
- De CDL Connector zorgt voor de conversie van de DAMO objecten uit de databases bij de Waterschappen naar een JSON bestand conform DAMO 1.5.1 of DAMO 2.1 structuur.
- Het JSON bestand wordt afhankelijk van de grootte van de dataset als één of meerdere zip-bestand(en) naar CDL Centraal gestuurd.
- De CDL Connector stuurt altijd een volledige gegevensset op naar CDL Centraal. Bij een nieuwe aanlevering wordt de oude aanlevering van het waterschap overschreven. Het is daarom niet mogelijk om een volledige gegevensset in delen naar CDL Centraal te sturen.
- Ondersteuning 2D en 2.5D geometrieën: In DAMO kunnen objecten geometrieën hebben met x,y en/of x,y,z coördinaten. Beide formaten worden automatisch ondersteund in het JSON bestand dat naar CDL Centraal wordt verzonden.
- De CDL Connector zorgt voor het overhevelen van het JSON bestand naar CDL Centraal. De Connector communiceert met CDL Centraal via het REST API endpoint van CDL Centraal. Meer informatie over de REST API; zie de technische beschrijving van de CDL API.
- De CDL Connector verstuurt een GET request met versleutelde token in de header naar CDL Centraal. Op deze manier weet het CDL Centraal welk Waterschap het request verstuurt. De CDL Connector krijgt ook een versleutelde token terug. Vervolgens verstuurt de CDL Connector de

dataset van het Waterschap via een POST request met daarin de 2 tokens; de eerste token geeft aan welk waterschap het betreft, het tweede token geeft aan welke bestand(en) horen bij het request. Een levering aan CDL Centraal kan dus bestaan uit meerdere zip-bestanden. De CDL Connector bepaalt aan de hand van de grootte van de dataset of het versturen van meerdere zip-bestanden nodig is.

- De CDL Connector ondersteunt 2 versies van DAMO, de vigerende versie en de versie daarvoor.
- De CDL Connector kent een eenvoudige configuratie. In het configuratiebestand kan de volgende informatie worden ingesteld door het waterschap:
 - Database-instellingen DAMO (hostnaam, poortnummer, gebruikersnaam en wachtwoord)
 - DAMO versie (1.5.1 of 2.1)
 - Waterschapscode
 - Waterschapstoken (afkomstig van CDL Centraal)
- Tijdens het uitvoeren van de CDL Connector wordt de logging van de applicatie bijgehouden. In het daaruit volgende logbestand zijn ook de eventuele foutmeldingen beschreven. Het logbestand wordt lokaal opgeslagen.

Zie de CDL Connector Installatiehandleiding voor meer informatie.

5. CDL Centraal

CDL Centraal is het centrale onderdeel van het CDL en is verantwoordelijk voor het valideren, samenvoegen en beschikbaar stellen van informatie.

5.1 Aanleveren

- De CDL Aanleveren API verzorgt een gestructureerde aanlevering door de Waterschappen op basis van het Request-Reply communicatiepatroon.
- Deze REST API wordt gebruikt door de CDL Connector, maar zou ook gebruikt kunnen worden door externe systemen.
- De CDL Aanleveren API is beschikbaar in twee versies; één voor het aanleveren van data in DAMO 1.5.1 en één voor het aanleveren van data in DAMO 2.1. Beide API's functioneren technisch gezien hetzelfde, alleen de JSON specificaties zijn verschillend en afhankelijk van de betreffende DAMO versie.
- Vanuit de CDL Aanlever API worden ook de events gelogd die een beheerder kan bekijken. Denk hierbij aan het tijdstip van aanleveren van data door een Waterschap via de CDL Connector.

5.2 Valideren

Tijdens de eerste fase van de verwerking van een aanlevering vinden de volgende validaties plaats:

- CDL Centraal controleert of de aangeleverde bestanden voldoen aan de afgesproken structuur aan de hand van de JSON specificatie (schema validatie). Alleen bestanden die voldoen aan de JSON specificatie worden verwerkt door CDL Centraal. In de JSON specificatie is onder andere gedefinieerd welke geometrie types zijn toegestaan per featuretype en welke attributen al dan niet verplicht zijn.
- CDL Centraal controleert of de aangeleverde geometrieën technisch valide zijn (bijvoorbeeld: is een polygoon sluitend).
- Alleen objecten die voldoen aan de JSON specificatie en waarvan de geometrie (indien aanwezig) valide is worden opgeslagen in CDL Centraal

In de tweede fase worden de opgeslagen objecten gevalideerd:

- CDL Centraal controleert de referentiële integriteit van de opgeslagen objecten. Dat betekent dat alle gerelateerde objecten in een levering moeten voorkomen. Een afsluitmiddel dat naar een niet-bestaande duiker verwijst wordt als invalide beschouwd.
- Objecten waarvan de referentiële integriteit onjuist is worden in CDL Centraal als invalide gemarkeerd, waardoor deze niet in de DAMO 2.1 view cq. INSPIRE, IMWA en GWSW uitleveringen voorkomt.
- Alle validatie resultaten worden verzameld en vastgelegd in het upload report. Per objecttype worden alleen de eerste 10 validatiefouten in het upload report opgenomen, zodat het upload report niet te lang wordt en leesbaar blijft. Het upload report wordt met een e-mail notificatie aan het waterschap (en desgewenst Het Waterschapshuis) verzonden.

5.3 Samenvoegen

- CDL Centraal voegt de data per waterschap samen tot landelijke datasets.
- De data van de waterschappen wordt eerst samengevoegd tot één landelijke dataset in DAMO 2.1 formaat; dit is de DAMO 2.1 view in CDL Centraal. De DAMO 2.1 mapping is hiervoor bepalend.
- Vanuit de DAMO 2.1 view worden de views in IMWA formaat samengesteld in CDL Centraal. De IMWA mapping is hiervoor bepalend.
- Vanuit de DAMO 2.1 view worden de views in INSPIRE formaat samengesteld in CDL Centraal. De INSPIRE mapping is hiervoor bepalend.
- Er is geen aparte GWSW view in CDL Centraal. Het Turtle RDF bestand conform GWSW wordt door de GWSW Exporter direct vanuit de DAMO 2.1 view gegenereerd op basis van de GWSW mapping.

5.4 Beschikbaar stellen

IMWA

- De IMWA uitleveringen worden gemaakt op basis van de data in de IMWA views.
- De IMWA data wordt aan PDOK uitgeleverd in GeoPackages.
- Aan PDOK wordt per thema een dataset beschikbaar gesteld als zip-bestand met per objecttype een GeoPackage.
- Het CDL zet de volgende landelijke datasets klaar voor PDOK:
 - IMWA Oppervlaktewateren
 - IMWA Keringen
 - IMWA Kunstwerken
 - IMWA Waterbeheergebieden
- De functioneel beheerder kan via de CDL API zelf de thema definitie bepalen:
 - definiëren welke objecttypes per thema (dataset) uitgeleverd moeten worden, op basis van de in IMWA aanwezige objecttypes
 - wie bij een valide of invalide uitlevering een notificatie moet ontvangen
 - aan welke afnemer(s) de dataset uitgeleverd moet worden (op dit moment worden de datasets alleen uitgeleverd aan PDOK)

INSPIRE

- De INSPIRE uitleveringen worden gemaakt op basis van de data in de INSPIRE views.
- De INSPIRE data wordt aan PDOK uitgeleverd in XML bestanden.
- Het XML bestand wordt aan PDOK beschikbaar gesteld in een zip-bestand.
- Het CDL zet de volgende landelijke datasets klaar voor PDOK:
 - INSPIRE Hydrography
- De functioneel beheerder kan via de CDL API zelf de thema definitie bepalen:
 - definiëren welke objecttypes per thema (dataset) uitgeleverd moeten worden, op basis van de in INSPIRE aanwezige objecttypes
 - wie bij een valide of invalide uitlevering een notificatie moet ontvangen
 - aan welke afnemer(s) de dataset uitgeleverd moet worden (op dit moment worden de datasets alleen uitgeleverd aan PDOK)

GWSW

- Het Turtle RDF bestand conform GWSW wordt door de GWSW Exporter gegenereerd vanuit de DAMO 2.1 view op basis van de GWSW mapping.
- Het Turtle RDF bestand wordt elke nacht naar RioNed verzonden middels een met apikey versleutelde POST request op de GWSW API. Met de POST request op de GWSW API wordt het Turtle RDF bestand geüpload naar de GWSW server van RioNed.
- De functioneel beheerder kan via de CDL API zelf een GWSW thema definiëren:
 - de inhoud van dataset is vastgelegd in de GWSW mapping
 - wie bij een valide of invalide uitlevering een notificatie moet ontvangen
 - aan welke afnemer(s) de dataset uitgeleverd moet worden (RioNed)

Uitleveren

- CDL Centraal stelt de IMWA en INSPIRE datasets beschikbaar aan PDOK. Het CDL zet dagelijks de bestanden klaar op het applicatie hosting platform waarbij een unieke URL wordt gegenereerd voor iedere dataset. Het CDL genereert ook een unieke token voor iedere dataset. Deze token is onderdeel van de URL van de dataset. Deze URL's worden verstuurd in een JSON bestand naar PDOK zodat zij weten onder welke URL's de datasets te downloaden zijn. De verbinding met het CDL is een https verbinding. Dit betekent dat wij een certificaat meesturen die PDOK dient te accepteren alvorens de bestanden gedownload kunnen worden.
- Een dag (24 uur) later wordt een volgende dataset klaargezet voor PDOK met andere, unieke URL's voor de te downloaden datasets.
- Elke nacht voordat een nieuwe dataset wordt aangemaakt, worden oude datasets opgeschoond waardoor alleen de twee meest recente valide datasets voor dat thema worden behouden. Daarna wordt de nieuwe dataset voor dat thema aangemaakt, waardoor - als die valide is - maximaal drie datasets per thema beschikbaar zijn; de huidige dataset en de twee voorgaande datasets. Als een nieuwe dataset invalide is, blijven alleen de twee meest recente valide datasets beschikbaar.
- Als een gebruiker na het opschonen van een dataset op de downloadlink van de betreffende dataset in de bevestigingsmail klikt, zal de gebruiker error 404 pagina zien, omdat het bestand niet meer bestaat.
- Een GeoPackage uitlevering wordt als invalide beschouwd als één van de objecttypes in de dataset geen records bevat, de dataset wordt dan niet uitgeleverd.

5.5 Beheren

- De CDL API kent een eenvoudige Swagger user-interface die is gericht op het samenstellen van datasets, gebruikersbeheer en het inzien van logbestanden. De user interface wordt gebruikt door de functioneel beheerder van het CDL.
- De functioneel beheerder kan:
 - zelf GeoPackage datasets samenstellen op basis van IMWA objecttypes
 - gebruikers (waterschappen) aanmaken, wijzigingen en verwijderen
 - logbestanden inzien (wie heeft wat en wanneer geupload?)

6. Bijlagen

6.1 Afkortingen en acroniemen

Afkorting / acroniem	Toelichting
AHP	Applicatie Hosting Platform van Geodan
API	Application Programming Interface
CDL	Centrale Distributielaag
DAMO	Data Afspraken Modelmatig Ondersteund
GWSW	Gegevenswoordenboek Stedelijk Water
IMWA	Informatiemodel Water
INSPIRE	Infrastructure for Spatial Information in Europe
JSON	JavaScript Object Notation
NGR	Nationaal Georegister
PDOK	Publieke Dienstverlening Op de Kaart
REST	Representational state transfer