

Memo

onderwerp Template impactanalyse HVDL

datum 9 februari 2023

status definitief

blad 1 van 6

van Wideke Boersma, Ton Zijlstra, Ine de Visser

aan Dataproviders

1 Inleiding

Eind december 2022 is de uitvoeringsverordening voor Europese hoogwaardige datasets (High Value Datalijst; HVDL) van kracht geworden¹. Deze verordening legt verplichtingen op aan overheden t.a.v. het via APIs en via bulkdownload publiceren van openbare gegevens in zes thema's. Voor een deel van deze thema's is in detail gespecificeerd om welke data het gaat. Voor andere thema's, met name binnen het thema Aardobservatie en milieu, lijkt er meer ruimte voor keuze en prioritering.

De implementatieperiode duurt 16 maanden en moet medio 2024 zijn afgerond. Om keuzes te kunnen maken in de implementatie voert het Ministerie van Binnenlandse Zaken en Koninkrijksrelaties (BZK) een verkenning uit onder betrokken dataproviders naar de benodigde inspanningen, en mogelijke opbrengsten en scenario's van de implementatie.

Om de uitkomsten van die verkenning zoveel mogelijk vergelijkbaar te laten zijn, en om met alle betrokken dataproviders te kunnen beoordelen welke elementen bijvoorbeeld gezamenlijk kunnen worden uitgevoerd, voorziet dit document in een sjabloon voor de beantwoording van de vragen in de verkenning.

De lijst van HVDL wordt door de Europese Commissie periodiek uitgebreid. Dit betekent dat ook voor latere toevoegingen aan de HVDL zulke vergelijkbaarheid van nut kan zijn.

2 Uitgangspunten verkenning: opbrengsten meewegen en scenario's vergelijken

De Europese hoogwaardige datasets zijn gekozen vanwege de reeds zichtbare en te verwachten maatschappelijke waarde van het hergebruik van die gegevens door derden.

Niet voor ieder van de zes thema's in de HVDL is gelijksoortig omschreven welke data precies binnen het thema valt.

Voor alle data zijn er afwegingen te maken t.a.v. de keuze voor API's of andere technische voorzieningen.

Waar de te publiceren data voldoende helder omschreven is in de uitvoeringsverordening (met name t.a.v. geodata, bedrijven en bedrijfseigendom, statistiek en meteorologie) blijven de te maken keuzes daar toe beperkt. Waar zulke helder omschreven data uit de HVDL al gepubliceerd wordt voor hergebruik betreffen die keuzes wellicht meer het al dan niet aanpassen van die publicatievoorzieningen.

¹ [https://eur-lex.europa.eu/legal-content/NL/TXT/HTML/?uri=PI_COM:C\(2022\)9562&from=EN](https://eur-lex.europa.eu/legal-content/NL/TXT/HTML/?uri=PI_COM:C(2022)9562&from=EN)

Waar de te publiceren data minder helder omschreven is en ruimte laat voor keuzes (aardobservatie en milieu, en mobiliteit) liggen er additionele keuzes voor t.a.v. welke data zal worden gepubliceerd, naast hoe die publicatie zal worden gerealiseerd. Dit omdat er interpretatieruimte is in welke gegevens uit de genoemde INSPIRE thema's of op basis van de genoemde milieurichtlijnen wel of niet in scope is.

Dit leidt tot meerdere mogelijke scenario's (combinaties van organisatorische keuzes t.a.v. welke data wordt gepubliceerd, en de meer technische aspecten t.a.v. hoe die data wordt gepubliceerd). Zo kun je in de aangewezen INSPIRE thema's mogelijk onderscheid maken tussen varianten waarbij 'alle' data binnen het thema in scope wordt geacht, waarbij data wordt geprioriteerd op basis van gebruiksnut, of waarbij alleen de meest geschikte dataset in scope wordt geacht. Ook zijn er verschillende manieren om invulling te geven aan het begrip 'historische data'. De verschillende scenario's zijn niet alleen om de benodigde inspanningen (en mits bekend daarmee gepaard gaande kosten) in kaart te brengen, maar ook de daarmee samenhangende (verwachte) maatschappelijke baten van hergebruik te bepalen.

De intentie van dit document is om te helpen bij het in kaart brengen van zowel de inspanningen als de opbrengsten kant, en de verschillende scenario's naast elkaar te zetten. De gezamenlijke vergelijking van de uitkomsten door BZK is dan de basis om met elkaar te kiezen hoe de implementatie zal worden vormgegeven.

2.1 Karakterisering van opbrengsten

Waardevol hergebruik heeft heel verschillende verschijningsvormen. Naast de verschillende verschijningsvormen is hergebruik bovendien vaak minder zichtbaar voor de dataprovider en in eerste instantie veelal klein van aard. Bij elkaar opgeteld (over alle vormen en toepassingen heen) is de hergebruikswaarde echter vaak vele malen groter dan de inspanning van beschikbaarstelling. De EU High Value Datalijst (HVDL) is tot stand gekomen op basis van een studie² (2020) van de Europese Commissie waarin meerdere dimensies van waarde creatie door hergebruik zijn beschouwd.

De dimensies van hergebruikswaarde waar de kwalificatie high value in de HVDL op gebaseerd is, zijn:

- Klimaatverandering en milieu (open data hergebruik om klimaat- en milieufactoren beter te begrijpen, en te verbeteren)
- Economische waarde (productiviteit, werkgelegenheid, consumentenbaten, en impact of marktcompetitiviteit o.b.v. hergebruik)
- Innovatie, o.a. m.b.t. AI (ondernemendheid m.n. MKB, nieuwe diensten en applicaties, publieke innovatie)
- Publieke dienstverlening en overheid (verantwoording, inkoop, belastingopbrengsten van commercieel hergebruik, beheer/vernieuwing publieke diensten, open data publiceren als beleidsinstrument om stakeholders te activeren rondom opgaven)
- Maatschappelijk/sociaal (algemene opbrengsten voor de maatschappij, w.o. transparantie, participatie en inclusiviteit, gezondheidsvoorspellingen / preventie, mobiliteitssystemen, rechtssysteem)

Binnen deze dimensies werden voor alle deelfactoren indicatoren geformuleerd waarop de hergebruikswaarde van datasets kon worden beoordeeld. In de impactanalyse is het nuttig deze

² <https://www.access-info.org/wp-content/uploads/Deloitte-Study-2020.pdf>

dimensies als denkrichting te gebruiken, en uiteraard is het mogelijk om de studie² te gebruiken om de achterliggende indicatoren en voorbeelden van hergebruik ter reflectie te vinden.

Vragen om te stellen bij het beschouwen van het hergebruik van nu reeds gepubliceerde data zijn:

- Welke vormen van hergebruik zijn nu zichtbaar? Welke typen stakeholders zijn de zichtbare hergebruikers?
- Welke vragen m.b.t. data worden nu ontvangen waaraan nu geen gevolg kan worden gegeven (zichtbare vraag)?
- Welke groepen stakeholders, gebaseerd op huidig zichtbaar gebruik of zichtbare vraag, zijn potentiële hergebruikers (redelijkerwijs te vermoeden vraag)?
- Welke (groepen) stakeholders zou je zelf graag als hergebruiker zien, omdat dat bij kan dragen aan de uitvoering en effectiviteit van eigen beleidstaken (aan te moedigen vraag)?

Vragen om te stellen bij het beschouwen van het mogelijke hergebruik van nog niet gepubliceerde data zijn:

- Welke vragen m.b.t. data worden nu ontvangen waaraan nu geen gevolg kan worden gegeven (zichtbare vraag)?
- Welke vormen van hergebruik van deze data zijn elders zichtbaar of te verwachten? Welke typen stakeholders zijn de mogelijke hergebruikers? (te vermoeden vraag)
- Welke (groepen) stakeholders zou je zelf graag als hergebruiker zien, omdat dat bij kan dragen aan de uitvoering en effectiviteit van eigen beleidstaken (aan te moedigen vraag)?

2.2 Scenario's: Organisatorische versus inrichtingsaspecten

Voor de inschatting van de implementatie-inspanningen is het de bedoeling meerdere scenario's naast elkaar te plaatsen, alvorens (met elkaar) beslissingen te nemen. Daarbij maken we onderscheid in keuzes die te maken hebben met de scope van gegevens die gepubliceerd (zullen) worden, en keuzes die te maken hebben met hoe je die gegevens publiceert. In de tabel onderaan dit document zijn horizontaal de scenario's voor de meer organisatorische aspecten genoemd, en verticaal de meer technische inrichtingsaspecten. Hieronder worden beiden nader toegelicht.

2.2.1. Organisatorische keuzes / scope van data

T.a.v. de datasets die in scope zijn is voor de HVDL thema's geodata, statistiek, bedrijven en eigendom van bedrijven en meteorologie in de uitvoeringsverordening gedetailleerd aangegeven om welke data het gaat.

Voor het HVDL thema aardobservatie en milieu, en in mindere mate voor het HVDL thema mobiliteit, is minder gedetailleerd aangegeven welke data in scope is. Daar worden 19 INSPIRE thema's in scope geplaatst, en wordt gewezen op informatie behorend bij een dertigtal milieurichtlijnen, en is er in het algemeen enige interpretatieruimte.

Voor de slechts in de breedte aangeduide INSPIRE thema's onderscheiden we drie varianten in de impactanalyse (zie ook de factsheet Uitgangspunten voor de High Value Dataset implementatie):

- Alle binnen het INSPIRE thema vallende data wordt in scope geacht.
- Geprioriteerde datasets op basis van bestaand en potentieel hergebruik worden in scope geacht.
- De meest geschikte dataset zoals aangemerkt voor INSPIRE wordt in scope geacht.

Voor de milieugegevens o.b.v. zowel de Aarhusrichtlijn als de genoemde 31 specifieke milieurichtlijnen is het van belang ook aan te geven welke datasets worden meegenomen in de impactanalyse, en eventuele variaties daarop op basis van bestaand en potentieel hergebruik.

2.2.2 Inrichtingsaspecten

Voor elk van de meerdere van toepassing zijnde keuze-opties t.a.v. de scoping van de data, kan worden beschouwd wat de effecten van implementatiekeuzes ten aanzien van de inrichtingsaspecten zijn.

In die implementatiekeuzes valt te denken aan:

- Welke van de data wordt nu al als open data gepubliceerd, welke data is al wel digitaal ontsloten maar nog niet open?
- Voor al ontsloten data voldoet daar de bestaande API, moet er een API worden ontwikkeld, of moet de bestaande API worden verbeterd?
- Is er een bestaande en voldoende voorziening voor bulkdownload, of moet deze worden gerealiseerd?
- Zijn de nu beschikbare dataformats voldoende, of zijn er redenen om ook andere formats aan te bieden?
- Welke invulling van de definitie van 'historische data' wordt gehanteerd: een bredere invulling, of de minimale vereiste invulling?
- Welke generalisatieniveaus worden gepubliceerd? Enkel de vereiste of meerdere varianten, zowel onder als boven de vereiste?
- Welke documentatie moet worden opgesteld/gepubliceerd, zowel m.b.t. de data als de API?
- Ligt het voor de hand als datahouder deze inspanningen zelfstandig te doen, of is samenwerking waardevoller?

2.2.3. Inspanningen

De inspanningen die nodig zijn in elke combinatie van scenario en andere keuzes vallen vervolgens uiteen in:

- Veranderinspanningen die eenmalig zijn, zoals het realiseren van een publicatievoorziening of API, of documentatie waar die nog niet voorhanden zijn
- Periodieke inspanningen die terugkeren, zoals vaste kosten, beheer en onderhoud om publicatievoorzieningen in stand te houden, het actueel houden van documentatie etc.

Uiteindelijk zijn die inspanningen te beschouwen in de context van de te verwachten hergebruikswaarde.

3. Factsheets

Om de gezamenlijke afwegingen te vergemakkelijken stelt Geonovum een aantal factsheets op die voor een bepaald aspect (zoals Uitgangspunten voor de High Value Dataset implementatie, of wat als API geldt) de huidige stand van zaken, feiten en openstaande vragen weergeven. Deze factsheets worden telkens actueel gehouden (geversioneerd), zodat alle betrokkenen dezelfde begrippen kunnen hanteren.

De actuele factsheets zijn te vinden op <https://docs.geostandaarden.nl/eu/factsheets/>

4. Tabel

De hierboven genoemde dimensies en aspecten waarop de impact van de invoering van de HVDL kan worden beoordeeld zijn in de onderstaande tabel weergegeven (de tabel is echter niet bedoeld als limitatieve opsomming). Gebruik van de tabel vergroot de vergelijkbaarheid van uitkomsten van de impactanalyse. Horizontaal staan de verschillende scenario's t.a.v. de scope van data. Verticaal staan de meer technische inrichtingsaspecten t.a.v. ontsluiting. De tabel is niet zozeer als in te vullen sjabloon bedoeld, maar helpt om de samenhang tussen te beoordelen elementen te zien, en na te gaan of alle relevante elementen ook zijn beschouwd. Zowel de kolommen als de rijen eindigen met een veld m.b.t. effect op hergebruik, dit omdat bij elke keuze t.a.v. inspanning of scenario ook het mogelijk te verwachten effect op hergebruikswaarde meegewogen kan worden.

Scenario-> ----- -	Beschre- ven datasets	INSPIRE alles in thema	INSPIRE gepriori- teerd naar hergebruik	INSPIRE meest geëigende dataset	Milieu datasets	Effect op hergebruik van keuze
Inspanning en						
HVDL thema	Geo- informatie, Statistiek, Bedrijven,, Meteorolog ie	Aardobserv atie en milieu, Mobiliteit	Aardobserv atie en milieu, Mobiliteit	Aardobserv atie en milieu, Mobiliteit	Aardobserv atie en milieu	
Definitie- inspanning data in scope						
Data al open?						
Bestaande API wel/niet upgraden						
Nieuwe API maken						
Toevoegen bulkdownlo ad						
Dataformat transforma ties nodig?						
Historie meer dan vereist?						
Een, meer of alle generalisati eniveau's?						
Documenta tie						
Zelftstandi g of gezamenlij k						
Veranderin -spanning						
Periodieke inspanning en						
Effect op hergebruik van scenario						