

Handreiking 1.0: In 4 stappen naar open data

In slechts 4 stappen van open data ambities naar het daadwerkelijk open data.

Deze handreiking helpt overheidsorganisaties met open data ambities. Het vormt een soort "eerstelijns hulp" en beoogt 90% van alle vragen die spelen rond "het aan open data doen" te beantwoorden.

De handreiking neemt u - letterlijk - bij de hand en leidt uw organisatie in 4 stappen naar het publiceren van open data.

Stap 1 - Inventariseer de datasets



Welke data hebben we eigenlijk?

Startpunt is dat er een besluit is genomen om 'aan open data te gaan doen'. Dit besluit is bijvoorbeeld door een wethouder of een aantal afdelingshoofden genomen. De eerste vraag die dan rijst is: welke data hebben we eigenlijk? Oftewel, de eerste stap die je moet nemen, bestaat uit het inventariseren van data die mogelijk open data kunnen worden. Immers, je kunt pas data publiceren als je weet welke data de organisatie in huis heeft.

Het inventariseren van data gaat volgens drie achtereenvolgende stappen:



A. Maak een invullijst voor het inventariseren van je data

De eerste stap bestaat uit het maken van één verzamelpunt - in de vorm van een invullijst - om alle informatie in de organisatie omtrent data te noteren. Hierdoor ontstaat een overzicht. Vervolgens kun je gevonden datasets categoriseren en in een later stadium toegankelijk te maken.

Deze lijst is dus in feite een overzicht van alle datasets in de organisatie en een overzicht van hun kenmerken, de zogenaamde 'metadata': de gegevens over de data die in het kader van de inventarisatie verzameld dienen te worden. Dit kan gewoon in een spreadsheet o.i.d.

Als je de invullijst opstelt, moet je een aantal dingen in het achterhoofd houden:

Realiseer je dat het mensenwerk is

Je kunt veelal de data niet (geheel) zelf inventariseren: je bent aangewezen op de welwillende hulp van collega's, zoals datahouders en applicatie- en functioneel beheerders. Hun tijd (en geduld) zal beperkt zijn. Zorg daarom dat je je verzoek goed inleidt zodat ze weten waar je naar op zoek bent.

Inventariseer niet meer dan nodig is

Geef de invullijst zo vorm dat er zoveel als nodig - maar ook niet meer dan dat - ingevuld moet worden. Bovendien wil je slechts de metadata inventariseren die relatief gemakkelijk te vinden zijn. Maak je lijst dus niet te lang.

Eenduidige semantiek

Bij het vaststellen van de te inventariseren metadata moet er voorts voor gezorgd worden dat die een eenduidige betekenis hebben. Zo zal de term 'dynamiek van de data' verwarring veroorzaken; een betere beschrijving zou zijn: 'hoe vaak worden data geactualiseerd?', waarna een aantal categorieën van antwoorden gegeven kunnen worden (dagelijks, wekelijks, maandelijks etc.). Met andere woorden: voorkom gebruik van termen in de metadata die op meerdere wijzen te interpreteren zijn en omschrijf de metadata zo scherp mogelijk.

Gebruik in ieder geval een kernset van metadata

De vormgeving van de invullijst zal variëren van organisatie tot organisatie - een gemeente verschilt nu eenmaal van een waterschap - maar niettemin is er wel een vaste kern van metadata die in iedere inventarisatielijst voorkomt:

Naam

De titel waaronder de database binnen de organisatie bekend staat. Dit maakt het mogelijk een database met een bepaalde naam te vinden. De titel wordt gebruikt als het belangrijkste referentiepunt in de lijst zoekresultaten.

Beschrijving

De beschrijving van de inhoud van database. Presentatieveld dat de gebruiker helpt nader te bepalen of de database aan de behoefte voldoet.

Vindplaats

Plek waar je de dataset bent tegengekomen, op het diepste niveau. Dit veld identificeert de online bron.

URL vindplaats

Eenduidige verwijzing naar de database binnen een bepaalde context. Dit veld identificeert de data.

Contactpersoon

Interne eigenaar of contactpersoon die veel over de data weet. Persoonlijke bron waar meer informatie over de database te verkrijgen is.

Context

Onderwerp van de inhoud van de dataset. Maakt het mogelijk datasets te clusteren rondom bepaalde beleidsthema's.

Bronhouder

Organisatie die primair verantwoordelijk is voor het creëren van de inhoud van de dataset. Hierdoor kunnen datasets gevonden worden die gecreëerd zijn door of onder verantwoordelijkheid vallen van een bepaalde afdeling of dienst.

TIP: de bovenstaande kernset vormt de essentie van de data inventarisatie. Tijdens de inventarisatie kan meer informatie worden verzameld.



B. Identificeer de bronnen die naar de data leiden

Om deze invulijst te vullen moet je bij de databronnen zijn: lijstjes, applicaties of personen die naar de datasets leiden.

Nu we de invulijst hebben, moeten we deze gaan vullen en om deze te vullen, moeten we allereerst de bronnen van de data identificeren. Databronnen zijn lijstjes, applicaties of personen die je naar de datasets leiden. Ook hier zal het per organisatie verschillen hoe gemakkelijk de datasets te vinden zijn, niettemin is het raadzaam in dit proces de achtereenvolgende stappen te zetten.

Kijk hoe anderen het gedaan hebben

Het is begrijpelijk dat er binnen je organisatie niet een kant-en-klaar overzicht bestaat van alle datasets. Gelukkig ben je niet de eerste overheidsorganisatie die aan open data gaat doen, dus maak vooral gebruik van de wijze lessen en lijstjes van collega's elders in het land. Zij zullen je graag helpen en hebben vaak veel online staan. Hieronder een overzichtje van enkele goede voorbeelden:

- [Gemeente Amsterdam](#)
- [Gemeente Eindhoven](#)
- [Gemeente Zaanstad](#)
- [Gemeente Ede](#)
- [Ministerie van Economische Zaken](#)

Deze inventarisaties kunnen als spiegel gebruikt worden: als een gemeente bent, ligt het voor de hand te kijken wat gemeentes aan datasets hebben verzameld. De processen van gemeentes verschillen niet zo veel van elkaar en is het waarschijnlijk dat ze vergelijkbare datasets zullen publiceren.

Raadpleeg collega's

De tweede belangrijke databron zijn de collega's die met data werken. Denk hierbij aan functioneel beheerders en applicatiebeheerders: zij weten bij uitstek welke applicaties draaien en welke data daarbij gebruikt worden. Daarnaast zijn er ook collega's die dagelijks met data aan het werk zijn bij het uitvoeren van hun taken. Zij zullen vooral inhoudelijk kunnen vertellen welke data ze hebben.

Hieruit komt een lijstje van datasets van organisaties elders in het land en een overzicht van collega's in de organisatie die je kunnen helpen bij de laatste stap: het invullen van de inventarisatielijst.



C. Vul aan de hand van en met behulp van je databronnen de invullijst in.

Vervolgens vul je aan de hand van deze bronnen de invullijst in. De gesprekken met de datahouders en andere collega's zijn bovendien een mooie manier om het open data gedachtegoed verder te verspreiden.

De databronnen zijn verzameld en er is een invullijst gemaakt aan de hand van het hierboven beschreven format. Nu alleen nog invullen.

Vul zelf zoveel mogelijk in

Het handigst is als je eerst de door jou zelf gevonden data verwerkt. Dit zijn dus de datasets die je in andere overzichten hebt gevonden en waarvan je denkt dat deze ook binnen jouw organisatie aanwezig zullen zijn. Dit levert dus een lijst op waarmee je richting collega's kunt gaan.

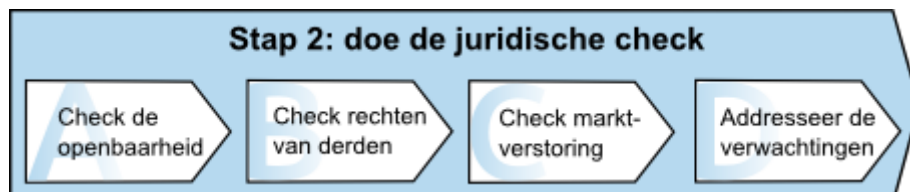
Laat collega's checken en aanvullen

Inmiddels heb je een lijst van collega's waarvan je vermoedt dat ze naar datasets zullen leiden. Met hen maak je een afspraak en leg je de conceptlijst voor die je hebt opgesteld aan de hand van de externe voorbeelden. Tijdens de gesprekken neem je deze lijst door en vul je deze aan waar nodig.

Vaak realiseren collega's zich niet dat zij (prachtige) open data hebben. Samen achter het beeldscherm bekijken van de processen en systemen die zij gebruiken,

leidt vaak tot verrassende vondsten. Bovendien kun je bij die gelegenheid de voordelen van open data toelichten en eventuele zorgen wegnemen: het feit dat zij die data hebben, wil zeker niet zeggen dat het ook open data zijn of moeten worden.

Qua volgorde is het raadzaam eerst de functioneel- en applicatiebeheerders te raadplegen om vervolgens door te schakelen naar collega's die de data inhoudelijk goed kennen. Zo vul je dus stapsgewijs, met behulp van de collega's de inventarisatielijst in met als eindresultaat een groslijst van datasets die mogelijk als open data kunnen worden gepubliceerd. De eerste stap is gezet.



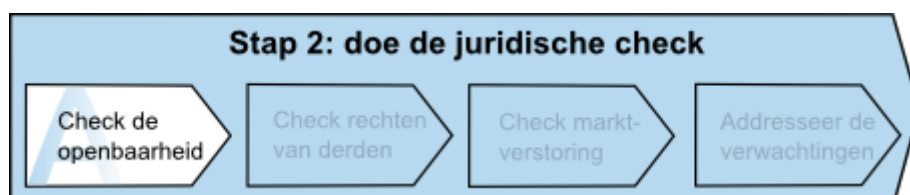
Stap 2 - doe de juridische check

Grondige check op aansprakelijkheid en strijdigheid met rechtsregels

Als duidelijk is welke data allemaal aanwezig is, moet uiteraard gecontroleerd worden of het open stellen daarvan niet in strijd is met enige rechtsregel. Dat doen we dan ook in stap 2 waarin we de juridische check uitvoeren.

Het nemen van deze stap is van groot belang omdat het vrijgeven van data sets in strijd met enige rechtsregel tot aansprakelijkheid kan leiden en daarnaast (ook) tot veel negatieve publiciteit. Met andere woorden: vergeet deze check vooral niet en doe hem grondig! Uiteraard is het raadzaam, naast het gebruik van deze Handreiking, ook de afdeling juridische zaken bij deze exercitie te betrekken.

Het uitvoeren van deze juridische check, bestaat uit 4 stappen.



A. Stel vast of de data openbaar is binnen de kaders van de Wet openbaarheid van Bestuur en/of andere openbaarheidsregelingen

De hoofdregel in Nederland is dat overheidsinformatie openbaar is, tenzij een wettelijke regel anders bepaalt. Die uitzonderingen op de hoofdregel zijn opgenomen in de Wet openbaarheid van Bestuur of andere bijzondere regelingen. Het is dus

zaak om van te voren goed te controleren of er geen uitzondering van geldt voor de geselecteerde data sets.

er beoordeling van de vraag of data gehouden door de overheid ook 'Open Data' zijn, is de eerste belangrijke vraag of de data openbaar is. Als data niet openbaar is, mag deze ook niet in hergebruik gegeven worden. Aan de hand van een aantal hierna weergegeven stappen kan bepaald worden of de data openbaar is.

Welke openbaarheidsregeling is van toepassing?

Je kunt veelal de data niet (geheel) zelf inventariseren: je bent aangewezen op de welwillende hulp van collega's, zoals datahouders en applicatie- en functioneel beheerders. Hun tijd (en geduld) zal beperkt zijn. Zorg daarom dat je je verzoek goed inleidt zodat ze weten waar je naar op zoek bent.

Om de toepasselijkheid van de openbaarheidsregeling vast te kunnen stellen, moeten we eerst bepalen:

1. [gaat het om bestuursorganen?](#)
2. [gaat het om bestuurlijke informatie?](#)
3. [berust deze bij het bestuursorgaan?](#)
4. [is deze neergelegd in documenten?](#)
5. [is de toepasselijkheid van de Wob uitgezonderd?](#)

Is de informatie openbaar?

Als we vastgesteld hebben dat het gaat om informatie waarop de Wob van toepassing is, dan is de hoofdregel - in lijn met de Open Data-filosofie en de Nederlandse beleidsambities - dat de informatie openbaar is en vrijgegeven moet worden voor hergebruik. Echter, de Wob kent een aantal uitzonderingen: soorten bestuurlijke informatie die niet openbaar zijn. Uiteraard moet een overheidsorganisatie deze uitzonderingen toetsen. Is een uitzondering van toepassing dan mag de informatie niet vrijgegeven worden voor hergebruik.

Bij de beoordeling of bestuurlijke informatie openbaar gemaakt mag worden moet het bestuursorgaan nagaan of de verstrekking niet indruist tegen de in de Wob expliciet genoemde belangen. Deze zijn vastgelegd in [artikel 10 van de Wob](#) en vallen uiteen in twee soorten de absolute weigeringsgronden en relatieve weigeringsgronden.

Absolute weigeringsgronden

Bij absolute weigeringsgronden blijft informatieverstrekking te allen tijde achterwege. Dit is het geval indien de verstrekking van de informatie:

1. [een bedreiging zou kunnen vormen voor de eenheid van de Kroon](#)
2. [de veiligheid van de Staat zou kunnen schaden](#)
3. [indien het bedrijfs- en fabricagegegevens betreft die vertrouwelijk aan de overheid zijn medegedeeld](#)
4. [indien het bijzondere persoonsgegevens betreft](#)

Relatieve uitzonderingsgronden

Bij relatieve weigeringsgronden heeft het bestuursorgaan een discretionaire bevoegdheid. Het moet een afweging maken tussen het algemene abstracte belang van openbaarheid en de belangen genoemd in Art. 10 tweede lid Wob. Het betreft hier:

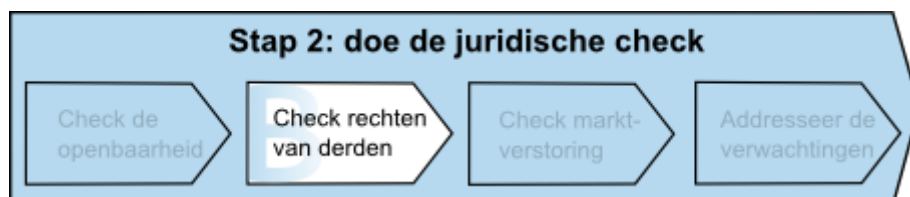
1. [internationale betrekkingen](#)
2. [economische of financiële belangen van de overheid](#)
3. [de opsporing en vervolging van strafbare feiten](#)
4. [inspectie controle en toezicht door bestuursorganen](#)
5. [eerbiediging van de persoonlijke levenssfeer](#)
6. [het belang dat een geadresseerde erbij heeft als eerste kennis te kunnen nemen van informatie](#)
7. [het voorkomen van onevenredige bevoordeling of benadeling van de betrokkenen](#)

Aanwezigheid van persoonsgegevens

De aanwezigheid van persoonsgegevens is een hele belangrijke als het gaat om aan Open Data doen. Immers, overheden verzamelen bij de uitoefening van hun taken vaak massaal gegevens over personen. De hoofdregel van de Wob in samenhang met de Wbp is dat deze gegevens niet voor hergebruik in aanmerking komen. Vanwege het grote belang van deze categorie, gaat deze Handreiking hier uitgebreid op in.

Achtereenvolgens komen aan de orde:

1. [het grote belang van zorgvuldige toetsing](#)
2. [regeling in de Wob en de Wbp](#)
3. [het begrip 'persoonsgegeven'](#)
4. [wat als de privacywetgeving van toepassing is?](#)
5. [hoe te handelen?](#)
6. [wat als er toch wat mis is gegaan?](#)



B Controleer of derden geen rechten op de data kunnen laten gelden

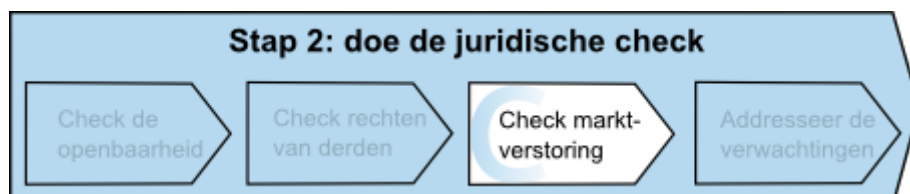
(Rechten op) data zijn over het algemeen van iemand (een overheidsorganisatie of een niet overheidsorganisatie; bijvoorbeeld een bedrijf). Meestal nemen die de vorm van auteurs- en databankrechten. Uiteraard moet je wel over de (rechten op de) data kunnen beschikken, anders geef je iets weg dat niet van jou is en maak je inbreuk op deze rechten.

(Rechten op) data zijn over het algemeen van iemand (een overheidsorganisatie of een niet overheidsorganisatie (bijvoorbeeld een bedrijf)). Meestal nemen die de vorm van auteurs- en databankrechten, geregeld in de Auteurswet, respectievelijk de Databankenwet. Uiteraard moet je wel over de (rechten op de) data kunnen beschikken, anders geeft je iets in hergebruik dat niet van jou is en maak je inbreuk op de rechten van deze derden.

Als het gaat om mogelijke rechten van collega overheidsorganisaties, dan zijn daarvoor enkele bijzondere regels die afwijken van het bovenstaande. Kort gezegd houdt die regeling in dat als een overheidsorganisatie bij de publicatie van de informatie niet heeft aangegeven dat zij haar rechten wil voorbehouden, die informatie vrij is van rechten en dus gewoon als Open Data vrijgegeven kan worden.

Niettemin, als je data ter beschikking willen stellen voor hergebruik volgens de Open Data filosofie doe je er dus goed aan bij de aanmaak c.q. verwerving van de data zeker te stellen dat er geen rechten van derden op rusten. Mochten er derden aan te pas komen, dan kan dit door (standaard) inkoopvoorwaarden te hanteren die bepalen dat alle mogelijke rechten die ontstaan in het kader van de inkoop eigendom worden van de inkoopende overheid. De huidige standaard voorwaarden van inkoop voor de centrale overheid bevatten een dergelijk beding. Als laatste slot op de deur doet men er uiteraard goed aan op het moment van vrijgeven voor hergebruik nog een controle uit te voeren op dit aspect.

Deze handreiking bevat een andere toelichting en een stroomschema (met toelichting) aan de hand waarvan je kan nagaan wat je op het gebied van rechten van derden moet controleren en doen.



C. Stel vast of openbaarmaking van de data niet concurrentievervalsend kan werken

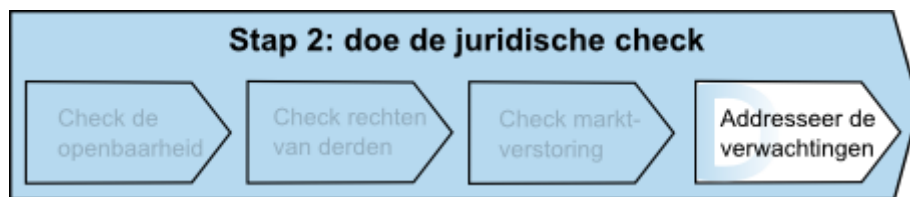
In het geval er marktpartijen zijn die zelf vergelijkbare data sets produceren en verkopen, dan moet je nagaan of je deze marktpartijen geen oneerlijke concurrentie aandoet.

In het geval er marktpartijen zijn die zelf vergelijkbare data sets produceren en verkopen, dan moet je oppassen dat je deze marktpartijen geen oneerlijke concurrentie aandoet doordat je de data, in lijn met het Open Data gedachtegoed, gratis en voor niets ter beschikking stelt.

Volgens de Open Data filosofie moet overheidsinformatie weliswaar as is maar wel gratis en voor niets, althans tegen maximaal de verstrekingskosten en verder ongeclausuleerd, ter beschikking worden gesteld voor hergebruik aan eenieder. Als er evenwel een markt bestaat waarop deze of gelijksoortige informatie door

marktpartijen verhandeld wordt, betekent dit dat er voor afnemers op die markt een zeer goedkoop of zelfs gratis alternatief beschikbaar komt, waardoor klanten van die marktpartijen kunnen weglopen. Het is daarom denkbaar dat een overheidsorgaan aansprakelijk wordt gesteld door een marktpartij die betoogt dat het uitvoeren van het Open Data beleid onrechtmatig jegens hem is, omdat de condities waartegen de overheid de data ter beschikking stelt hem onrechtmatige concurrentie aandoen.

Recent is op gebied veel beweging geweest: er is nieuwe wetgeving gekomen en daarnaast heeft de rechter in een tweetal zeer duidelijke uitspraken gedaan. De geruststellende conclusie hieruit is dat een overheid die aan Open Data gaat doen zich niet snel aan concurrentievervalsing schuldig zal maken.



D. Geef duidelijk aan wat de gebruiker mag verwachten en wat niet

De data sets die je open stelt zijn verzameld in het kader van de uitoefening van de publieke taak. Nu kan er uiteraard daarbij het een en ander misgaan waardoor de inhoud van de data sets niet correspondeert met de werkelijkheid. Op zich is dit geen reden om de data sets niet te publiceren, maar het is wel van belang de verwachtingen van de gebruikers te adresseren.

De data sets die je open stelt zijn verzameld in het kader van de uitoefening van de publieke taak. Nu kan er uiteraard daarbij het een en ander misgaan waardoor de inhoud van de data sets niet correspondeert met de werkelijkheid. Op zich is dit geen reden om de data sets niet te publiceren, maar het is wel van belang de verwachtingen van de gebruikers te adresseren. Er rust immers wel een zorgplicht op de overheid om gebruikers te informeren over de data en ze te waarschuwen dat er mogelijk gebreken in de informatie zouden kunnen zitten.

Je doet er daarom goed aan op twee manieren de gebruikers voor te lichten over hetgeen ze wel en niet mogen verwachten van de data, en wel door een basislaag van meta-informatie toe te voegen en door een proclaimer op te nemen.

Een basislaag van meta-informatie

De overheid die data in hergebruik wil geven doet er allereerst goed aan een basislaag van metadata toe te voegen, hetgeen de hergebruiker in staat stelt de context van de data beter te begrijpen en eventuele gebreken gemakkelijk zelf kan constateren. Deze basislaag bestaat uit:

- Context
- Verwijzing naar pagina met achtergrond informatie over de dataset
- Verwijzing naar plaats waar de data kan worden gedownload

- Formaat van de data
- Bestandsgrootte van de data
- Titel
- Unieke naam
- Dataset betreft periode
- Betreft Grondgebied
- Licentie
- Tags
- Linked open data sterren
- Organisatietype eindverantwoordelijke
- Organisatie eindverantwoordelijke
- Organisatietype Auteur
- Organisatie Auteur
- Organisatie type beheerder
- Organisatie beheerder
- Email adres beheerder

Zeker nu deze basislaag effectief zou moeten zijn voor het wegnemen van de meest voor de hand liggende risico's, zal dit veelal een afdoende maatregel zijn. Uiteraard is het handig om deze metadata te verzamelen als je je data potentieel inventariseert zoals beschreven in stap 1

Disclaimers en proclaimers

Diverse overheden pogen mogelijke aansprakelijkheid uit te sluiten door het gebruik van disclaimers. De vraag is wel of deze zogenaamde exoneratie altijd houdt. Zeker als de overheid de enige bron is waar de informatie opgehaald kan worden en helemaal als het hebben van de informatie noodzakelijk is voor de deelname aan het rechtsverkeer, is de houdbaarheid minder waarschijnlijk. Daarbij valt bijvoorbeeld te denken aan informatie uit de basisregistraties die geacht worden de juiste gegevens te bevatten en waarvoor ook een 'gebruikverplichting' bestaat binnen de overheid.

Indachtig het Open Data gedachtegoed is er wel veel meer te zeggen voor het geldig kunnen hanteren van mededeling waarin wordt uitgelegd wat een her-gebruiker wel mag verwachten van de data. Deze zogenaamde proclaimers zijn in toenemende mate populair, ook in overheidsland, en komen uiteraard ook weer in vele soorten en maten, zie bijvoorbeeld die van de [provincie Brabant](#), [de gemeente Zwolle](#) en het [RIVM](#). Het onderscheid tussen de disclaimer en de proclaimer zit met name in de aard: waar de disclaimer gelijk stevig alle aansprakelijkheid afwijst, is de proclaimer veel vriendelijker van toon en beoogt het de gebruiker te informeren over de inhoud van de informatie en tegelijkertijd ook diens verwachtingen te managen.

De houdbaarheid van een proclaimer bij het openstellen van data voor hergebruik zal nog toenemen als deze specifiek gericht is op de dataset in kwestie, en dus geen algemene verklaring is. Uiteraard kan de verklaring getrapt zijn: er is een algemene proclaimer die geldt voor alle data die beschikbaar zijn voor hergebruik en per dataset is er één die verwijst naar de basislaag van metadata in de individuele datasets. In die metadata kan dan aangegeven worden in welke context de data verzameld zijn, zodat de hergebruiker zijn bedachtzaamheid op gebreken daarop kan baseren.

De data wordt bij voorkeur onder een [CC0 of publiek domein](#) verklaring beschikbaar gesteld. Dit geeft voldoende waarborg voor hergebruik zonder beperkingen. Eventueel kan indien gewenst een [CC-By licentie](#) waarin de hergebruiker wordt verplicht de bronhouder van de data te vermelden. In het geval van CC-By moet worden aangegeven hoe de te gebruiken naam van de bronhouder moet worden beschreven.

Wat als er toch wat mis gaat?

Mocht een hergebruiker niettemin menen dat hem onrecht is aangedaan doordat de overheid hem (beweerdelijk onjuiste) data zijn verstrekt dan zal hij een zogenaamde onrechtmatige daadactie kunnen instellen bij de rechter. Dit is geregeld in [artikel 162 van Boek 6 van het Burgerlijk Wetboek](#). Hij zal dan betogen dat de overheid aansprakelijk is voor gebrekkige open data.



Stap 3 - beslis over & organiseer het publicatieproces

Open data als standaard onderdeel van werkprocessen

Je weet nu welke data je in huis hebt (stap 1) en welke daarvan, vanuit juridisch perspectief (stap 2), als Open Data beschikbaar gesteld kunnen worden voor hergebruik. Aldus zijn we aangekomen bij de volgende stap: het selecteren van de te publiceren datasets en het organiseren van de consequenties daarvan. Door weloverwogen te selecteren en het proces vervolgens goed te organiseren schep je niet alleen het benodigde draagvlak, je voorkomt ook dat het publiceren van open data een ad hoc klusje is dat erbij hangt. Anders gezegd: deze stap zorgt ervoor dat het doen aan Open Data een standaard onderdeel gaat vormen van de werkprocessen binnen je organisatie en waar mensen zich verantwoordelijk voor voelen.

Een en ander valt uiteen in 3 substappen.



A. Selecteer de te publiceren datasets

We weten, op grond van de vorige stap, wat juridisch gezien als Open Data naar buiten mag en wat niet. Maar uiteraard vormt het juridisch kader niet het enige punt van afweging, ook andere aspecten spelen mee bij de keuze welke data te publiceren.

We weten, op grond van de vorige stap, wat juridisch gezien als Open Data naar buiten mag en wat niet. Maar uiteraard vormt het juridisch kader niet het enige punt van afweging, ook andere aspecten spelen mee bij de keuze welke data te publiceren.

Gezien de hoeveelheid databestanden die binnen een organisatie bestaan, is het niet doenlijk om de publicatie van alle bestanden tegelijkertijd te realiseren.

Met andere woorden: we moeten kiezen. In dat keuzeproces kan je een aantal criteria toepassen:

- begin met data sets die als onderlegger kunnen functioneren voor andere data sets. Denk daarbij bijvoorbeeld aan: geodata, basisregistraties als de BAG, en (historische) statistieken.
- neem ook het enthousiasme van je collega's mee bij het maken van keuzes: afdelingen waar men verwacht dat het doen aan Open Data zal bijdragen aan de eigen prestaties en efficiëntie, zijn je beste bondgenoten.
- probeer in te schatten welke datasets het meeste 'Open Data potentieel' hebben voor de hergebruikers
- kijk of er vraag is naar de data, open deze data met voorrang.
- kijk vooral ook naar andere organisaties: het is waarschijnlijk dat de logica van hun keuzes ook op jouw organisatie van toepassing is.

Dit proces leidt dus tot een lijst van geselecteerde datasets. Neem vervolgens, na interne afstemming dan ook de formele beslissing om deze datasets te gaan publiceren. Dan is het helder wat het mandaat is en dan kun je overgaan naar de volgende stap: het plannen en organiseren.



B. Beleg de verantwoordelijkheden en plan het proces

De open te stellen data zijn geselecteerd en de beslissing daartoe is genomen. Dat maakt de data echter nog niet open: daarvoor is de hulp nodig van een aantal stakeholders.

De open te stellen data zijn geselecteerd en de beslissing daartoe is genomen. Dat maakt de data echter nog niet open: daarvoor is de hulp nodig van een aantal stakeholders. Zo zijn daar:

- de chief data officer
- de datahouder (red: data-eigenaar)

- de technisch beheerders (red: verstrekkers en catalogushouders)
- de beleidsverantwoordelijken
- de gebruikers
- de Open Data steward

De Chief Data Officer

Data is altijd sterk verweven geweest met het bedrijfsproces waardoor als het proces goed beschreven was de data ook goed beschreven was. Maar door een steeds grotere integratie van processen en behoefte van uitwisselen van gegevens wordt de rol van data op zichzelf steeds belangrijker. We verzamelen steeds meer data en de data is ook steeds makkelijker te verplaatsen en te bewerken. De rol van data groeit. Dat vraagt ook van organisaties dat er meer bewustzijn omtrent data komt, dat bekend is welke data er is, waar deze data is, wie toegang heeft tot de data, wat de kosten van het verzamelen van de data zijn, wie verantwoordelijk is voor de data, wat de kwaliteit is van de data enzovoorts. Data is op deze manier een nieuw onderwerp voor veel organisaties.

Om aandacht voor data te organiseren en op een goede manier met data om te gaan kan invulling worden gegeven aan de rol van een Chief Data Officer, de CDO. De CDO plaatst alle onderwerpen met betrekking tot data op de agenda van de organisatie. Het belang van data voor de eigen organisaties wordt beschreven en de relatie van de data met buitenwereld.

De datahouder

De datahouders zijn de collega's die met de data werken beleidsmatig of uitvoerend. Zij zijn inhoudelijk gedreven en zullen de data op dat niveau heel goed kennen. Vooral bij uitvoerders zullen de baten neerslaan van het aan Open Data doen en zij zijn dan ook een belangrijke sponsor. Hun taak is vooral inhoudelijk: het metadateren en vooral ook contextualiseren van de data: waarvoor hebben we deze verzameld en wat mag je ervan verwachten (en wat vooral ook niet). Zij zijn dus verantwoordelijk voor het aanleveren van inhoud, de benzine waarop de machine moet draaien.

De technisch beheerders

Applicatiebeheerders en functioneel beheerders zijn je bondgenoten waar het de het scheppen van de juiste infrastructuur en het daadwerkelijk publiceren aangaat (zie hierna). Zij zullen vooral willen weten wat de benodigde functionaliteiten zijn en zij hebben als taak te zorgen dat de publicatiemachine soepel draait.

De beleidsverantwoordelijken

Ter borging van het Open Data gedachtegoed is het voorts raadzaam goede voeling te houden met de interne beslissers: zij zullen geregeld willen weten of de plannen op schema liggen en vooral ook wat de output is en de effecten zijn. Het is hun taak de continuïteit van het aan Open Data doen politiek bestuurlijk te borgen en te verdedigen. Daarvoor moeten ze uiteraard wel gevoed worden met de juiste informatie.

De gebruikers

Gebruikers hebben formeel geen rol bij de organisatie van het proces, maar ze zijn wel essentieel. Niet alleen moeten zij voor de baten zorgen - nieuwe producten ontwikkelen, overnemen van taken, feed back geven op de data etc. - zij zijn daarmee ook een belangrijke verdedigingslinie voor het borgen van het gedachtegoed.

De Open Data steward

Om al deze taken en belangen te stroomlijnen en organiseren doe je er goed aan een 'Open data steward' te benoemen. Deze Open data steward is de proceseigenaar van het aan Open Data doen. Het is zijn/haar taak zich maximaal in te zetten voor zowel verdere ontwikkeling van het Open Data proces, als ook voor het stimuleren van bewustzijn zowel binnen en buiten de organisatie. Bij voorkeur is de open data steward iemand die zowel de beleids-(uitvoerings-)kant als ook de technische kant goed kent. Verder moet de steward goed kunnen schakelen met bestuurders en directies. Gegeven het grote belang van de in- en externe communicatie dient de Open Data Steward dus ook nauw samen te werken met de afdeling communicatie.

Samengevat zijn de concrete taken van de Open Data steward zijn dus:

- het voorbereiden van data-releases
- het onderhouden van contacten met de datahouders, technisch beheerders
- de concrete oplijning van data voor publicatie
- Het werven onder enthousiasmeren van datahouders en beleidsverantwoordelijken
- binnen de organisaties en werkt met hen aan het beschikbaar stellen van datasets
- Het zijn van eerste aanspreekpunt voor gebruikers
- het onderhouden van contacten met collega organisaties

Tenslotte worden zij erop gewezen dat het doen aan Open Data niet van de een op de andere dag ingeregeld is. Het is daarom raadzaam deze rol voor langere tijd goed te beleggen en deze niet bij de eerste bezuinigingsronde van tafel te laten vallen.

Qua planning van het proces leert de ervaring leert dat het zinvol en effectief is om te werken met vaste en aangekondigde publicatiemomenten zodat deze steeds in rust kan worden voorbereid. Dikwijls is het zo dat ieder publicatiemoment niet alleen bestaande Open Datasets worden geactualiseerd, ook is het een uitgelezen moment nieuwe sets toe te voegen. Een vast ritme voorkomt 'pieken' rondom een of twee momenten in het jaar en draagt eraan bij dat het doen aan Open Data sterker verankerd raakt in de vaste processen van je organisatie.



C. Schep de infrastructuur

Gelijktijdig met of aansluitend op het organiseren en plannen van het Open data proces, moet ook de benodigde infrastructuur worden geschapen.

Gelijktijdig met of aansluitend op het organiseren en plannen van het Open data proces, moet ook de benodigde infrastructuur worden geschapen. Deze infrastructuur bestaat uit een puur technisch gedeelte maar ook uit een zachter 'informatieele gedeelte'.

Technische infrastructuur

Goed beschouwd moeten de data die Open Data moeten worden van A naar B. Ze zitten in A - back-office systemen van de organisatie - en ze moeten naar B - een voor Open Data bedoeld datamagazijn binnen het netwerk. Dikwijls worden hiervoor scripts gebruikt - zogenaamde ETL-scripts (staat voor *Extract Transform Load*) waarmee de data in de gewilde vorm worden getransformeerd en op een op de doellocatie geladen. Dit heeft tot voordeel dat systemen niet aangepast hoeven worden.

Vanuit dit datamagazijn worden vervolgens de databestanden periodiek op een dataserver in een zogenaamde DMZ (*demilitarized zone*) geladen, die overigens ook prima extern gehost kan worden. Dit gebeurt uitsluitend geïnitieerd vanuit het netwerk, en betreft eenrichtingsverkeer naar de DMZ toe. Het is vervolgens deze dataserver die via het web te benaderen is door hergebruikers om Open Data te bekijken en op te halen.

Dan wil je een plek creëren waar de Open Data via een eigen herkenbaar kanaal naar buiten kunnen. De website van de organisatie ligt voor de hand, maar er zijn ook andere opties (verzamelites, samenwerking met nabij gelegen organisaties die al aan Open Data doen). Uiteraard spelen ook de beschikbare middelen hierbij een rol. De twee meest gangbare opties zijn een zogenoemde datadump of ontsluiting via een API. Een datadump houdt in dat de gehele dataset in een keer geplaatst wordt. Een API staat voor *Application Programming Interface*, wat betekent dat je als gebruiker, als het ware, vragen kunt stellen aan de dataset. Deze laatste optie wordt vooral gehanteerd wanneer het grote en gecompliceerde datasets betreft. Realiseer je wel dat het maken van een API een behoorlijke klus kan zijn en bovendien kun je je afvragen of dat nog wel een taak is van je organisatie (en niet van de markt).

Waar je ook voor kiest, het is raadzaam simpel en klein te beginnen: een eenvoudige (sub)site waar de datasets thematisch en/of in alfabetische volgorde kunnen worden gevonden en downloadbaar zijn in de meest gangbare formaten. Uiteraard kunnen de functionaliteiten gaandeweg uitgebreid worden.

Wordt er realtime data of met regelmaat wisselende data ontsloten, dan zijn er verschillende mogelijkheden om deze data te ontsluiten:

- een webservice die gestructureerde data onsluit
- basisdatabestand met een incrementeel wijzigingsbestand erbij
- wijzigingsbestanden

Voor welke manier van beschikbaar stellen een overheidsorganisatie kiest is afhankelijk van de mogelijkheden en de vorm waarin de data geleverd kan worden. Uitgangspunt is in ieder geval dat er een open standaard, zonder beperkingen in hergebruik wordt toegepast.

Informationele infrastructuur

Daarnaast wil je natuurlijk ook op de plek waar je de data publiceert een aantal zaken helder maken. Denk hierbij aan:

- de grondslag voor het aan Open Data doen, zoals een raadsbesluit van een gemeente
- de hoofdlijnen van het Open Data beleid
- Een proclamer waarin duidelijk gemaakt wordt wat gebruikers wel en niet mogen verwachten
- Een contact persoon - bij voorkeur de Open Data steward - waar gebruikers terecht kunnen met vragen en opmerkingen

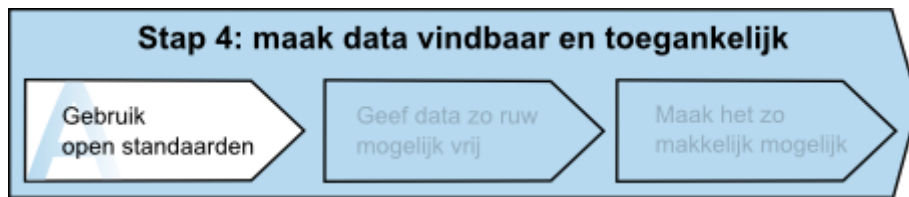


Stap 4 - Maak data vindbaar & toegankelijk

Open data is vindbaar en toegankelijk

Je hebt nu drie stappen doorlopen: de data is geïnventariseerd, de data heeft de juridische toets doorstaan en je hebt het publicatieproces georganiseerd en gepland. Wat nu nog rest is dat de datasets goed toegankelijk en vindbaar worden gemaakt. De data moet nu eenmaal voor iedereen bruikbaar zijn en deze moet ook gecommuniceerd worden naar buiten om daadwerkelijk van open data te kunnen spreken.

Stap 4 is dan als volgt op te delen.



A. Gebruik open standaarden

Maak, waar mogelijk, gebruik van open standaarden om toegang en machineleesbaarheid van data te bevorderen.

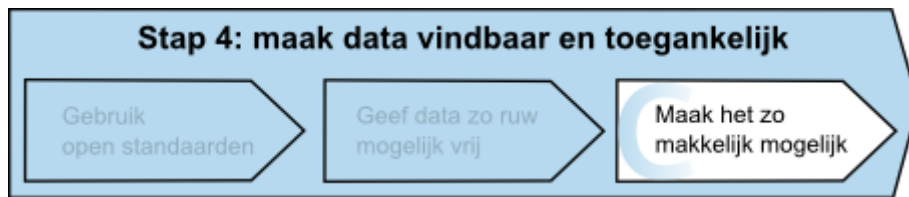
Een open standaard is een norm die publiek beschikbaar is. Voorbeelden hiervan zijn XML (*Extensible Markup Language*) en CSV (*Comma Separated Values*). Dit is allereerst handig voor de gebruiker, aangezien hij niet afhankelijk is van bepaalde software (leveranciers) om je datasets te kunnen lezen. Als tweede, is het sinds 2008 ook verplicht voor overheden om open standaarden te gebruiken. Hiervan kan alleen worden afgeweken in een aantal uitzonderingsgevallen, bijvoorbeeld wanneer de bedrijfsvoering in gevaar komt (hierbij gebruikmakende van het zogenoemde 'pas toe of leg uit' principe).

Machineleesbaarheid maakt automatische verwerking van data door computers makkelijker. Machineleesbaarheid is vooral van belang voor de gebruiker en maakt het voor hem/haar een stuk makkelijk om van je open data gebruik te maken. Dit moet daarom zoveel mogelijk bevorderd worden. Als je binnen organisatie data in een machineleesbaar formaat hebt, moet je het ook zo publiceren; verder dan dit hoef je niet te gaan, aangezien machineleesbaar publiceren niet verplicht is.



B. Geef data zo ruw mogelijk vrij

Zorg ervoor dat je data in een zo min mogelijk bewerkte staat vrijgeeft. Je wil het gebruik van de data niet een bepaalde kant op sturen, noch wil je bepaalde dingen aannemen. Zorg ervoor dat je data in een zo min mogelijk bewerkte staat vrijgeeft. Je wil het gebruik van de data niet een bepaalde kant op sturen, noch wil je bepaalde dingen aannemen. Het is af en toe makkelijk, en verleidelijk, om aannames over het gebruik van een dataset te maken (de gebruiker vindt het vast fijn als we in dataset X, Y en Z aan passen om de data gebruiksvriendelijker te maken), maar zorg ervoor dat je je handjes zoveel mogelijk thuishoudt. Soms is het echter onvermijdelijk om data te bewerken. Soms moet je, om juridische redenen, persoonsgegevens uit de dataset halen of moet je data aggregeren omdat anders deze gegevens herleidbaar zijn tot bepaalde personen. Probeer dit alles echter tot een minimum te beperken. Zorg er ook voor dat, wanneer je data in een machineleesbaar formaat hebt, je deze ook zo publiceert (zoals in substap A werd aangegeven).



C. Maak het zo makkelijk mogelijk

Gelijktijdig met of aansluitend op het organiseren en plannen van het Open data proces, moet ook de benodigde infrastructuur worden geschapen.

Als finishing touch moet je het de gebruiker zo makkelijk mogelijk maken. Dit kan je langs drie lijnen bewerkstelligen.

Voorkom het opwerpen van onnodige drempels

Werp geen onnodige drempels op om toegang te krijgen tot een dataset, zoals registratie, inlogprocedure, of beperking van het aantal aanvragen. Wanneer je data opengemaakt is, moeten gebruikers ook daadwerkelijk toegang hebben tot deze data. Het is begrijpelijk dat je als organisatie nieuwsgierig bent naar je datagebruikers, probeer de gebruiker echter niet te hinderen. Een beter manier is om mensen uit te nodigen of te vragen hoe zij je data gebruiken.

Zorg voor een duidelijke, goed bereikbare contactpersoon

Zorg verder dat er een duidelijke, goed bereikbare contactpersoon is voor vragen over de gepubliceerde datasets. Je wilt weten wat gebruikers van je dataset vinden, zorg er daarom voor dat ze een kans krijgen om feedback te geven. Publiceer dus duidelijk een telefoonnummer met naam en e-mailadres zodat gebruikers ook de vragen die ze hebben, daadwerkelijk kunnen stellen. Wijs per dataset een persoon aan die specialist op het gebied dat de dataset bestrijkt. Bij deze persoon komen alle vragen binnen. Dit zorgt ervoor dat de gebruiker directe en duidelijke hulp krijgt bij het beantwoorden van zijn/haar vragen.

Registreer je data

Ten slotte: registreer je dataset sowieso op data.overheid.nl. Daarnaast is het natuurlijk ook zinvol je datasets ook op andere plekken te vermelden. Denk hierbij vooral aan de doelgroep die waarschijnlijk van jouw data gebruik gaat maken. Vertel dit dan ook aan je doelgroep. Hierbij is het verstandig om samen te werken met de communicatieafdeling binnen je organisatie.

Wanneer je dit alles gedaan hebt, is je data klaar voor open data gebruik en rond je het hele open data proces af.