

**Forum Standaardisatie**

Wilhelmina van Pruisenweg 52
2595 AN Den Haag
Postbus 84011
2508 AA Den Haag
www.forumstandaardisatie.nl

notitie

COLLEGE STANDAARDISATIE

Agendapunt:	5. Open standaarden voor Open data		
Bijlagen:	Eindrapport Open Standaarden voor Open Data		
Aan:	College Standaardisatie		
Van:	Forum Standaardisatie		
Datum:	29 mei 2012	Versie	1.0
Betreft:	Open Standaarden voor open data		

Gevraagd besluit

1. Het College herbevestigt het belang van de beleidslijn "open data tenzij", conform de kamerbrief van de minister van Binnenlandse Zaken¹ en onderschrijft het belang van standaardisatie om de overheidsdienstverlening, marktontwikkeling en innovatie met open data te bevorderen.
2. Het College stelt de "Het onderdeel van de handreiking over standaarden voor Open Data" vast op basis van het model van Tim Berners-Lee (5 niveaus voor standaardisatie). Dit onderdeel van de handreiking biedt aanbieders van open data een handvat voor publicatie en hergebruik van open data. Op dit moment is stimuleren van publicatie van open data het belangrijkste. Het uitgangpunt "Ruwe data nu!"² is vanwege de prille fase van het beleidsveld nu de eerste stap.
3. Het College geeft het Bureau Forum Standaardisatie opdracht om in samenwerking met Binnenlandse Zaken, afdeling KOOP, een voorbeeldijst voor bruikbare open standaarden voor open data voor eind juni op te stellen.
4. Het College vraagt het ministerie van Binnenlandse Zaken de richtlijn met de voorbeeldijst met open standaarden te publiceren op data.overheid.nl.
5. Het College vraagt het ministerie van Binnenlandse Zaken, afdeling KOOP, het invulformulier voor aanlevering van open datasets voor publicatie op data.overheid.nl aan te vullen met een invulveld voor betekenis/context van de data. Hierdoor ontstaat een de facto open (metadata)standaard voor betere vindbaarheid en gebruik van open data.

¹ <http://www.rijksoverheid.nl/documenten-en-publicaties/kamerstukken/2011/05/30/kamerbrief-hergebruik-overheidsinformatie-en-open-data.html>

² Deze uitspraak van Tim Berners-Lee: "Raw data nu!" (2009) is een kernachtige samenvatting van de conclusie uit de internationale benchmark in het kader van dit onderzoek uitgevoerd door Jeff Rothenberg, zie <http://www.forumstandaardisatie.nl/themas/open-data/>

Toelichting op bovenstaande beslispunten

Datum
29 mei 2012

Ad 2:

Uitwerking het onderdeel van de handreiking over formaten voor publicatie van Open Data voor de Nederlandse overheid

Volgens de classificatie van Tim Berners-Lee zijn er vijf niveau's waarop open data aangeboden kan worden voor hergebruik. Meer informatie over het 5-sterren model van Tim Berners-Lee (grondlegger WWW) is te vinden op: <http://www.w3.org/DesignIssues/LinkedData.html>. Veel overheden bieden nu nog ruwe data aan en zullen naar verwachting in de tijd doorgroeien naar het aanbieden van meer gestructureerde open data. Het opgestelde onderdeel van de handreiking over standaarden voor open data gaat uit van de indeling van Tim Berners-Lee en het 'pas toe of leg uit'-beleid voor open standaarden van de Nederlandse overheid.

Hierna volgt een toelichting op de vijf niveau's van Tim Berners-Lee, vertaald naar de Nederlandse situatie.

** Ruwe data, nu*

Publiceer open data in het ruwe, interne formaat via data.overheid.nl. Overheidsorganisaties verzamelen data in het kader van de uitvoering van hun publieke taak. Publiceer de data in het formaat, waarin de data door de overheidsorganisatie voor de uitvoering van haar publieke taak is verzameld en/of is bewerkt. Ook als dat (nog) geen open standaard en/of gestructureerd formaat is. Indien noodzakelijk kan dit ook om PDF/A en PDF1.7, die beide op de 'pas toe of leg uit'-lijst staan maar geen gestructureerde formaten zijn. Deze standaarden zijn vanwege het ontbreken van structuur niet aan te bevelen³.

***/** Open standaarden:*

Formaten conform open standaarden zijn de norm voor open data. "Pas toe of leg uit" geldt dus ook voor open data. Dat betekent dat bij nieuwe aanschaf van ICT-systemen rondom open data de relevante standaarden op de 'pas toe of leg uit'-lijst geëist moeten worden. Bij afwijkingen moet een overheid zich verantwoorden via het jaarverslag en de open standaard aanmelden.

**** Gestructureerde formaten van 'pas toe of leg uit'-lijst:*

Hanteer in geval van nieuwe aanschaf van ICT-systemen rondom open data bij voorkeur gestructureerde formaten van de 'pas toe of leg uit'-lijst. Dit vergroot de mogelijkheden tot hergebruik van de data. Deze standaarden hebben 3 sterren conform het model van Tim Berners-Lee (bijv. CSV of StUF). Hoewel overheden worden opgeroepen om ook beschikbare data te publiceren als zij dit niveau nog niet halen, is dit het minimale niveau in geval van nieuwe ICT-investeringen rondom open data.

*****/***** Verwijsbare en verwijzende formaten:*

Neem kennis van standaarden waarbij anderen naar jouw data kunnen verwijzen ("linken") en jij naar andermans data verwijst ("linkt"). Dit zijn de hoogste niveaus (4 en 5) in het model van Tim Berners-Lee. Deze open standaarden zijn op dit moment niet opgenomen op de 'pas toe of leg uit'-lijst. Overheden zijn vrij om deze linked data standaarden in te zetten en

³ De ervaring leert dat databestanden meestal in een ander formaat worden bijgehouden binnen een de organisatie, maar voor publicatie via internet worden omgezet naar een PDF formaat, dat is voor het beleidsveld open data zonde van de moeite en niet handig voor hergebruik

aan te melden voor de lijst open standaarden. In de praktijk zetten private partijen ook al beschikbare open data om in deze standaarden; er lijkt niet altijd noodzaak te zijn om als overheid deze standaarden in te zetten. Als een organisatie hiervoor kiest, blijft verantwoording in het jaarverslag noodzakelijk.

Datum
29 mei 2012

Ad 3:

Toelichting voorbeeldlijst

In het veld bestaat behoefte aan een overzicht van concrete open standaarden, zoals eerder goedgekeurd door het college, die een voorkeur krijgen voor de *publicatie* van Open Data op grond van de classificatie van het TBL model. De voorbeeldlijst zal uitgaan van 3 sterren als referentiepunt voor de meest wenselijke manier van publiceren van open data. Iedere concrete standaard op de voorbeeldlijst krijgt daardoor een aanbeveling of negatief advies. Zo is de publicatie van een dataset op basis van PDF als standaard af te raden. Door het gebrek aan structuur is hergebruik erg moeilijk.

De standaarden in de lijst kunnen een algemeen karakter hebben (1^e categorie) of een domein specifiek karakter (2^e categorie). Wanneer een partij een dataset publiceert die te relateren is aan een specifiek domein heeft een domeinstandaard de voorkeur. Wanneer een dataset beschikbaar is in meerdere gelijkwaardige standaarden volgens het TBL model, dan kan de aanbiedende partij een keuze maken. De voorbeeldlijst zal tevens de "Best Practice" standaarden bevatten als resultaat van het onderzoek. Een aanbiedende partij kan gebruik maken van de Best practice open standaarden. De aanleiding kan zijn een mogelijk hogere classificatie in het TBL model en/of vanwege beschikbaarheid in verband met het principe "raw data now". Alle partijen die hiervoor kiezen, worden uitgenodigd de gekozen standaard uit deze lijst aan te melden voor de 'pas toe leg uit'-lijst en een korte motivatie te geven voor de afwijking op de 'pas toe leg uit'-lijst in het jaarverslag.

Casuscitaat gemeente: "Waar Amersfoort momenteel nog geen enkele aandacht besteedt aan standaardisatie om het proces niet te bemoeilijken, heeft Rotterdam ervaring met de Pas toe Leg uit lijst, die waar mogelijk wordt toepast. Een risico ligt in een overvloed aan standaarden, waardoor het alsnog complex is om zaken te koppelen en aan de financiële middelen die nodig zijn om te voldoen aan standaarden."

Ad 5:

Toelichting ontwikkeling open standaard voor metadata.

- a. Uit het onderzoek blijkt de behoefte aan toevoeging van het veld betekenis/context op het formulier voor aanlevering van datasets door overheidsdienstverleners voor publicatie data.overheid.nl. Dit bevordert gebruik van datasets.
- b. Het zoeken en vinden van een dataset is afhankelijk van de invoervelden op het formulier bij aanlevering van een dataset. Deze invulvelden moeten juist gedefinieerd zijn, maar liefst niet uniek zijn voor data.overheid.nl. Aansluiting op bestaande standaarden voor indelingen op de 'pas toe leg uit'-lijst, zoals OWMS, en op de EU ontwikkelingen verhogen de kans op vindbaarheid. Data.overheid.nl wordt uitgenodigd hier een trekkersrol in te vervullen als beheerder van het portaal voor open data.

Casuscitaat KOOP: "Voor de vindbaarheid van data moet er een *paradigmashift* plaatsvinden. Er wordt nog veel zelf georganiseerd binnen de eigen muren, maar dat zou eigenlijk moeten zonder muren eromheen. Eigenlijk moet men er van meet af aan van uit gaan dat overheidsdata Open data is. Er wordt nu nog veel gedacht in 'mijn data' en aarzelingen zoals als de data 'open' wordt is het dan nog mijn data. Maar collecties vullen elkaar aan, moeten overlap hebben. Dit maakt het beeld completer. Het belang, de noodzaak om gevonden te worden ontbreekt nog. Er wordt vaak nog voor een traditionele manier van publiceren gekozen – via communicatieafdelingen etc. – in plaats van het online zetten van de data."

Datum

Waarom is open data belangrijk?

Uit verschillende onderzoeken is gebleken dat het openbaar, vindbaar en herbruikbaar aanbieden van open data positieve maatschappelijke en economische effecten kan hebben (bron digitale agenda EL&I 2011):

1. Overheidsinformatie voorziet in een behoefte: Er is een toenemende belangstelling voor hergebruik van overheidsinformatie.
2. Economische waarde: De groeiende aandacht voor hergebruik van overheidsinformatie komt ook voort uit de potentiële economische waarde. In studies wordt de waarde van open data hoog ingeschat. Voor de Nederlandse economie wordt geschat dat commercieel hergebruik van open overheidsdata 200 miljoen euro kan opleveren (naar analogie van buitenlandse studies, zoals in Denemarken). De opbrengsten van commercieel hergebruik van open overheidsdata voor de hele Europese economie wordt geschat op 28 miljard euro. Bron: <https://zoek.officielebekendmakingen.nl/kst-22112-1338.html>
3. Het economisch potentieel van open data zal pas benut worden als er nieuwe innovatieve toepassingen met open data worden gemaakt. Hergebruik van data kan voor verbetering van efficiency en kwaliteit zorgen, waardoor besparingen op de rijksbegroting mogelijk zijn.
4. Participatie en controle door burgers: De toegankelijkheid van overheidsinformatie draagt er aan bij dat burgers beter in de gelegenheid worden gesteld tot participatie in de maatschappij en controle van de overheid.

Waarom zijn open standaarden voor open data belangrijk?

Een open standaard vergemakkelijkt het gebruik van open data. Het is een belangrijke randvoorwaarde om de bovenstaande beleidsuitgangspunten waar te maken. Het bevordert de interoperabiliteit.

Overheidsinstanties zoeken naar houvast op dit nieuwe beleidsterrein. Daarnaast biedt een breed gedragen metadatering standaard voor de vindbaarheid en voor de beschrijving van de dataset verbeterde toegankelijkheid van de dataset. Deze open standaard kan bijdragen aan het voorkomen van verkeerde interpretatie.

Impact en voordelen voor betrokken stakeholders zijn:

- Overheidspartijen kunnen bij publicatie van open data de het onderdeel van de het onderdeel van de handreiking volgen, waardoor een praktische invulling van de beleidslijn ontstaat.
- Applicatiebouwers kunnen de datasets eenvoudiger in de eigen infrastructuur omzetten.
- Overheidspartijen vermijden misbruik en verkeerde interpretatie van de data door een het meegeven van de context van de dataset onder andere in de vorm van metadatering.

Waarom is de rol van College Standaardisatie hierin belangrijk?

Het college bestrijkt een breed terrein en kan daarmee het gebruik van open standaarden voor open data bevorderen.

Waarom kunt u met een 'gerust hart' ja zeggen tegen dit advies?

Het advies is in lijn met de internationale ontwikkelingen. De toegevoegde waarde van het advies is breed gedragen door betrokken partijen. Voor overheidsorganisaties die twijfelen over publicatie van datasets kan het een steuntje in de rug zijn. De voorgestelde beslispunten ondersteunen (de uitvoering van) het beleid voor open data van BZK en EL&I.

Verantwoording onderzoeksopzet

Datum
29 mei 2012

Aanleiding en verloop

Het Forum Standaardisatie heeft in het werkplan 2012 acties opgenomen ter ondersteuning van de beleidslijn open data van de ministeries van BZK en EL&I. door het wegnemen van belemmeringen voor deze beleidslijn. Een van de belemmeringen bleek de onduidelijkheid te zijn over standaarden voor open data.

Opzet onderzoek

1. Het onderzoek van Logica Business Consulting moest leiden tot voorstellen om het aanbod- en distributieproces voor open data te verbeteren en tot een onderdeel van de het onderdeel van de handreiking (vanuit een technische invalshoek). Vanuit de aanbodkant zijn de volgende partijen betrokken: VU/KOOP, CBS, I&M, NDW, IPO, Geonovum, Gemeente Amersfoort en Rotterdam (1 casusbeschrijving) en Planbureau voor de Leefomgeving.
2. Het gebruikersperspectief, de bedrijven voor hergebruik, (1 casusbeschrijving) werd gerepresenteerd door: Antwoord voor bedrijven, Reed Elsevier, Open Sahara's.

In de internationale benchmark zijn de Verenigde Staten, het Verenigd Koninkrijk en Canada betrokken (recent gestarte Open data pilot (data.gc.ca)), alsmede Nieuw Zeeland. De resultaten zijn na synthese aangeboden aan de Belgische overheid ter toetsing.

Op basis van een analyse zijn door de onderzoekers aanbevelingen geformuleerd. De uitkomsten en aanbevelingen zijn in de vorm van stellingen door de onderzoekers via een G(roup)D(ecision)R(oom)-sessie voorgelegd aan de medewerkers van de betrokken overheidsorganisaties en bedrijven. Deze medewerkers hebben hierop i.c. de best-practices en de aanbevelingen van de onderzoekers, kunnen reageren. De resultaten van deze GDR sessie zijn uitgewerkt in een zogenaamd Syntheserapport. Op basis van het synthese rapport is de eindrapportage opgesteld waarin door Logica Business Consulting de centrale onderzoeksvraag is beantwoord en zijn aanbevelingen geformuleerd.