



Open data: delen, verbinden en verrijken

Auteurs:

Maarten Zeinstra
Nikki Timmermans



Creative Commons Nederland is een samenwerkingsverband tussen
Stichting Nederland Kennisland, Waag Society en het Instituut voor Informatierecht

postbus 2960, 1000 CZ Amsterdam - tel: +31(0)20 5756720
www.creativecommons.nl - info@creativecommons.nl

Aanleiding

Open data is een relatief nieuwe vorm van publiceren voor culturele instellingen. Bij deze vorm van publicatie ontstaan vragen waar de culturele sector nog maar weinig ervaring mee heeft. In een serie van drie white papers behandelt Creative Commons Nederland vraagstukken waar instellingen in de praktijk mee te maken krijgen als ze aan de slag gaan met open data. In de eerste paper uit het drieluik (Auteursrecht en open data in de culturele sector) werd beschreven welke auteursrechtelijke vraagstukken voorkomen bij het open beschikbaar stellen van culturele data. In de tweede white paper (Creative Commons-licenties en methoden voor het openstellen van culturele data) werd ingegaan op de technische vraagstukken die voorkomen bij het openstellen van data in de culturele sector.

In deze derde en laatste white paper gaan we in op de manieren waarop open data in de culturele sector hergebruikt kan worden. De drie hoofdredenen voor het publiceren van open cultuur data zijn: stimuleren van hergebruik /nieuwe toepassingen, breed ontsluiten en het verrijken van culturele data. Wat kun je doen met open data of wat kunnen anderen doen met open data? En wat kun je daar vervolgens als culturele instelling weer uithalen? Voor deze paper wordt geput uit de opgebouwde kennis van culturele instellingen uit het Open Cultuur Data-netwerk. Voor deze paper hebben een aantal deelnemers uit het Open Cultuur Data-netwerk, die in 2011 meegedaan hebben aan de Apps voor Nederland-competitie¹, ervaringen vastgelegd in een case study beschrijving. We bedanken de mensen die hieraan bijgedragen hebben. In deze paper citeren we uit deze case studies.

De white papers zijn geschreven in het kader van de activiteiten van Open Cultuur Data². Open Cultuur Data is een netwerk van cultuurprofessionals, ontwikkelaars, designers, auteursrechtsspecialisten en open data-deskundigen, dat culturele data open maakt en de ontwikkeling van waardevolle culturele toepassingen stimuleert. Eind 2011 zijn Hack de Overheid³ en het Innovators Netwerk Erfgoedsector⁴ Open Cultuur Data gestart. Open Cultuur Data organiseert in 2012 met ondersteuning van Creative Commons Nederland⁵ en Beelden voor de Toekomst⁶ een aantal activiteiten, waaronder een masterclass open data voor culturele instellingen en de documentatie van kennis over open data in de culturele sector in deze serie white papers.

Wij bedanken het Nationaal Archief, het Amsterdam Museum, ab-c Media, EYE Film Instituut Nederland en het Rijksmuseum voor de bijdrages die ze geleverd hebben aan het onderzoek dat ten grondslag ligt aan deze papers en de R&D afdeling van het Nederlands Instituut voor Beeld en Geluid voor hun toegevoegde inzichten.

¹ <http://www.appsvoornederland.nl>.

² <http://www.opencultuurdata.nl>.

³ <http://www.hackdeoverheid.nl>.

⁴ <http://www.inerfgoed.nl>.

⁵ <http://www.creativecommons.nl>.

⁶ <http://www.beeldenvoordetoeekomst.nl>.

Open data in de organisatie

Op de vraag hoe open data op dit moment in de organisatie ingebed is, wordt een divers beeld geschetst door de culturele instellingen die zich als eerste aansloten bij het Open Cultuur Data-netwerk. Bij het Nederlands Instituut voor Beeld en Geluid speelt de Research & Development afdeling volgens Maarten Brinkerink een voortrekkersrol als het om open data gaat. Er wordt gewerkt aan beleid waarin open projecten in balans komen met de projecten waarin de ambitie is om geld te verdienen. Voor de meeste instellingen is open data echter een nieuw fenomeen. Uit de case study opgesteld door Irene Haan blijkt dat open data voor het EYE Film Instituut Nederland een nieuw onderwerp was dat voor deelname aan de Apps voor Nederland competitie ad hoc intern goedgekeurd werd. Voor het Nationaal Archief was het ook een nieuw onderwerp. Door deelname aan de wedstrijd is alles echter in een stroomversnelling geraakt. Er is bij het Nationaal Archief volgens Tim de Haan inmiddels een interne werkgroep opgericht die de open data agenda als speerpunt verder gaat uitwerken.

Maarten Brinkerink (Nederlands Instituut voor Beeld en Geluid): *“Op dit moment vervult de Research & Development afdeling waar wij werken een voortrekkersrol. We werken samen met de unit Marketing & Communicatie en informeren ze over de stappen die we zetten”.*

Irene Haan (EYE Film Instituut Nederland): *“Dit staat nog in de kinderschoenen. Ontsluiting van de collectie wordt door iedereen erg belangrijk gevonden, maar het open beschikbaar stellen is nog niet ‘top of mind’. Het ging nu in overleg tussen de afdeling Collectie, Digitalisering en Digitale Presentatie. Het Management Team heeft het opstellen van de data wel geaccordeerd; maar nog niet echt een duidelijk beeld van de mogelijkheden en eventuele voordelen ervan.”*

Tim de Haan (Nationaal Archief): *“In eerste instantie leek het moeizamer te gaan dan ik had voorzien. Men snapte niet waar ik het over had, waar ik naar toe wilde, wat ik nodig had. Maar sneller dan ik had verwacht is het begrip open data binnen de organisatie geland. Het staat intern prominent op de agenda en we willen er dit jaar al meer mee doen [...] Intern is er nu een werkgroep samengesteld waarmee we open data verder willen brengen en waaruit concrete acties moeten komen om aan het einde van de rit goed te kunnen zeggen wat bij ons open data is en wat niet, onder welke voorwaarden enzovoort.”*

Er blijken drie argumenten voor culturele instellingen te zijn om data open beschikbaar te stellen: om de collectie beschikbaar te stellen, voor het hergebruik van data en om samen te kunnen werken met andere culturele instellingen. Het eerste argument heeft betrekking op de publieke missie van culturele erfgoedinstellingen om bronnen zo goed mogelijk en zo breed mogelijk *beschikbaar te stellen* voor iedereen. Open data draagt bij aan het vergroten van de openbare toegankelijkheid van collecties. Algemeen rijksarchivaris Martin Berendse publiceerde in dat kader eind maart 2012 een artikel op Villamedia.nl⁷ waarin hij een lans breekt voor open data. Hij ziet het als speerpunt voor het Nationaal Archief om “data uit het publieke domein open te stellen en voor hergebruik beschikbaar te maken”. Martin

⁷ <http://www.villamedia.nl/opinie/bericht/open-data-speerpunt-nationaal-archief/>.

Berendse, Lizzy Jongma en Maarten Brinkerink zeggen het volgende over open data en hoe dat bijdraagt aan brede beschikbaarheid:

Martin Berendse (Algemeen rijksarchivaris): *“Met open data krijgen de controleurs van onze democratie, onderzoekers en journalisten en private hergebruikers van informatie een enorme rijkdom aan bronnen tot hun beschikking. ‘Pas als het niet meer zoeken is naar een speld in een hooiberg wordt openbaarheid toegankelijkheid. Dat is waarom het NA zich zal blijven inspannen voor zowel openbaarheid als open data”* (Uit: Villamedia, 29 maart 2012)

Lizzy Jongma (Rijksmuseum): *“Het Rijksmuseum droomde al geruime tijd over implementatie van Open Data. De collectie is niet alleen van het Rijksmuseum, maar van iedereen. De kennis en gegevens die wij hebben over de objecten uit de collectie wilden we graag delen met anderen. Wij wilden vooral graag Open Data om onze collectie te kunnen delen met portals als Europeana. Maar, wij wilden geen eenmalige exports meer leveren zoals bij Het Geheugen van Nederland, omdat wij geleerd hebben dat data snel ‘veroudert’.”*

Maarten Brinkerink (Nederlands Instituut voor Beeld en Geluid): *“Vanuit het Innovators Netwerk Erfgoedsector was Beeld en Geluid één van de trekkers van het project, omdat wij het belangrijk vinden dat erfgoed zo open en breed mogelijk beschikbaar wordt gesteld. Het is immers de publieke missie van Beeld en Geluid om collecties onder een breed publiek kenbaar te maken.”*

Het tweede argument dat door culturele instellingen aangedragen wordt om data openbaar beschikbaar te stellen, is *hergebruik van de data*. Hergebruik van de data in nieuwe toepassingen of applicaties waar de instellingen zelf nog niet aan gedacht hadden. En hergebruik van data in bestaande of lopende samenwerkingsprojecten, zoals bijvoorbeeld een project als CATCHPlus⁸ waarin meerdere dataleveranciers data aanleveren voor een breed scala aan ontwikkelde diensten voor het erfgoedveld. Of denk bijvoorbeeld aan andere data harvesting-projecten zoals Europeana⁹, waarin meerdere bronnen geaggregeerd en ontsloten worden op een Europese portal voor erfgoed. Maarten Brinkerink en Lizzy Jongma zeggen het volgende over open data en nieuwe toepassingen:

Maarten Brinkerink (Nederlands Instituut voor Beeld en Geluid): *“Gezien de resultaten van Apps for Amsterdam, waarbij de data van het Amsterdam Museum mateloos populair bleek, waren we vooral heel erg benieuwd naar wat ontwikkelaars met onze data zouden doen. We gingen er vanuit dat de culturele datasets ook voor Apps voor Nederland veel gebruikt zouden worden en verwachtten vooral dat we nieuwe inzichten zouden krijgen in de mogelijkheden van het hergebruik van onze collecties.”*

Lizzy Jongma (Rijksmuseum): *“We hadden gezien dat het Amsterdam Museum bij de wedstrijd Apps voor Amsterdam enkele hele mooie apps had gekregen. Wij hoopten dat ten minste één app ontwikkeld zou worden met onze Open Data. [...] Het project*

⁸ <http://www.catchplus.nl>.

⁹ <http://www.europeana.eu>.

heeft ons veel meer opgeleverd dan we hadden verwacht. We hadden gehoopt op een enkele toepassing. Inmiddels zijn er 11 apps ontwikkeld¹⁰."

Het derde argument dat door culturele instellingen aangedragen wordt om data open beschikbaar te stellen, is dat het makkelijker wordt om je data verder te verrijken met andere datasets en bronnen, en je daardoor ook makkelijker met andere (culturele) instellingen kunt gaan *samenwerken*. Je hoeft niet eerst een project definiëren, je kunt zelf gaan kijken of er door andere instellingen data beheerd wordt die relevant is voor je eigen collectie, missie en doelstellingen. Het open aanbieden van de collectiedata van bijvoorbeeld het Rijksmuseum was een eerste stap voor een verdere samenwerking met Europeana en de onderwijscommunities van Kennisnet, die op dit moment verder uitgebouwd worden. Het Open Beelden-platform¹¹ integreerde de open datasets van het Rijksmuseum en het Amsterdam Museum voor betere ontsluiting van het historische videomateriaal.

Community

Hoe kun je organiseren dat je data hergebruikt wordt en breder beschikbaar komt? Binnen je eigen organisatie zullen veel ideeën en misschien ook al projecten bestaan rond specifieke toepassingen. De meerwaarde van het aansluiten bij een community van open data-experts, programmeurs en ontwikkelaars die het leuk vinden om met open data aan de slag te gaan, is dat het een *nieuw perspectief* of nieuwe perspectieven op jouw dataset kan geven waar je nog nooit over nagedacht had en (*technische*) *expertise* geleverd wordt die je niet in huis hebt. In zekere zin zijn de ideeën vanuit een organisatie zelf beperkt door het 'professionele vocabulaire' dat je hebt. Een collectiemedewerker heeft een specifiek idee over waarom een collectie relevant is, een ontwikkelaar kijkt weer met heel andere ogen naar deze collectie. Hij ziet bijvoorbeeld de potentie voor een technologie als gezichtsherkenning voor de ontsluiting van de collectie. Door het aantal perspectieven op je data te vergroten, vergroot je de reikwijdte. Open Cultuur Data vindt om die reden aansluiting bij de community van Hack de Overheid, een collectief van experts op het gebied van open data en open data-toepassingen dat zich sterk maakt voor de ontsluiting van meer open data door overheden. Tweemaal per jaar organiseert het collectief een hackathon waar de betrokkenen van verschillende gespecialiseerde werkgroepen zoals Open Onderwijs Data, Open Zorg Data en Open Cultuur Data hun data presenteren en samenwerken aan het maken van nieuwe toepassingen. Bijkomend voordeel van deze aanpak is dat hierdoor meerdere type datasets gecombineerd kunnen worden waarbij interessante kruisverbanden gelegd kunnen worden.

¹⁰ Een actueel overzicht van gemaakte toepassingen met de Rijksmuseum API kan gevonden worden op www.rijksmuseum.nl/api.

¹¹ <http://www.opencultuurdata.nl/2011/12/demo-open-beelden-video%E2%80%99s-verrijkt-met-open-data/>.



Afbeelding 1. Hack de Overheid hackathon in de Shell toren. Foto: Breyten Ernsting / CC-BY.

Op basis van de eerste ervaringen met het inbrengen van open cultuur data op hackathons kunnen de volgende conclusies getrokken worden over wat vanuit het perspectief van een hacker en ontwikkelaar interessant bevonden wordt. Ten eerste, *(omvang)rijke content* is gewild. Zorg ervoor dat je zo veel mogelijk en zo interessant mogelijke content beschikbaar stelt, en niet enkel metadata. Uiteindelijk spreekt het meer tot de verbeelding om iets te doen met de afbeeldingen uit collecties, zoals de Nachtwacht van het Rijksmuseum of de zonnebloemen van het Van Gogh museum, dan met alleen de metadata van deze stukken, zo blijkt. Het EYE Film Instituut Nederland bracht voor Apps voor Amsterdam bijvoorbeeld metadata in onder CC0 beschikbaar van hun collectie Nederlandse films. Deze data werd door niemand gebruikt. Het Rijksmuseum bracht naast metadata van haar gehele collectie uiteindelijk ook een groot deel van hun content uit de collectie in. Aanvankelijk was het Rijksmuseum van plan om contentselecties te maken rond specifieke onderwerpen, maar uiteindelijk is ervoor gekozen om de selectie zo breed mogelijk te houden. Deze strategie leverde vlak na de afsluiting van de wedstrijd (binnen een maand) al elf applicaties op. Irene Haan en Lizzy Jongma zeggen het volgende over de inbreng van content:

Irene Haan (EYE Film Instituut Nederland): *“Omdat we helaas nog geen beeldmateriaal met onze dataset konden meegeven, hadden we geen hoge verwachtingen. Wel hoopten we dat er interessante verbindingen zouden ontstaan met andere ontsloten collecties [...] Het beschikbaar stellen van de data was een goede eerste stap, maar zoals gezegd bevatte de data geen beeldmateriaal, waardoor het voor de ontwikkelaars waarschijnlijk geen aantrekkelijke dataset is (dit was al onze verwachting). De data is dan ook niet gebruikt. De volgende keer zullen we alleen data aanleveren als we ook beeldmateriaal beschikbaar kunnen stellen.”*

Lizzy Jongma (Rijksmuseum): *“We begonnen aan de wedstrijd met selecties van datasets waarmee ontwikkelaars (in onze ogen) leuke apps zouden kunnen bouwen. We verzandden echter in de vraag wat appbouwers leuk zouden gaan vinden. Het hele idee van sets naar onderwerpen hebben we toen losgelaten. We zijn kunstwerken gaan selecteren die auteursrechtelijk vrij zijn en waar we beeldmateriaal van hebben, vanuit de veronderstelling dat appbouwers altijd beeldmateriaal willen gebruiken. Deze veronderstelling blijkt juist. Zodoende zijn wij tot een selectie van ruim 100.000 kunstwerken van voor 1840 gekomen [...] We*

merken inderdaad dat de meeste appbouwers op zoek zijn naar de Rembrandts en werken uit de Gouden Eeuw. Dit hebben wij uit kunnen leveren. We vinden het wel spijtig dat wij onze modernere parels niet als Open Data beschikbaar hebben kunnen stellen en zodoende ook niet een ander beeld van het Rijksmuseum hebben kunnen bieden.”

Ten tweede geldt hoe *opener* de data beschikbaar is, hoe beter dat is voor het stimuleren van hergebruik. In deze context betekent open dat (commercieel) hergebruik is toegestaan, dat de kwaliteit van de aangeboden content zo hoog mogelijk is (geen thumbnails) en dat het via een open standaard aangeboden wordt op zo’n manier dat je er ook iets mee kan. Arjan den Boer, ontwikkelaar bij webbureau ab-c media, zegt hier het volgende over:

Arjan den Boer (ab-c media): “In ieder geval graag met zo min mogelijk beperkingen qua licentie en met name ook het gebruik van beeldmateriaal. Alleen metadata en geen beeld of slechts thumbnails hebben we weinig aan.”

Toepassingen

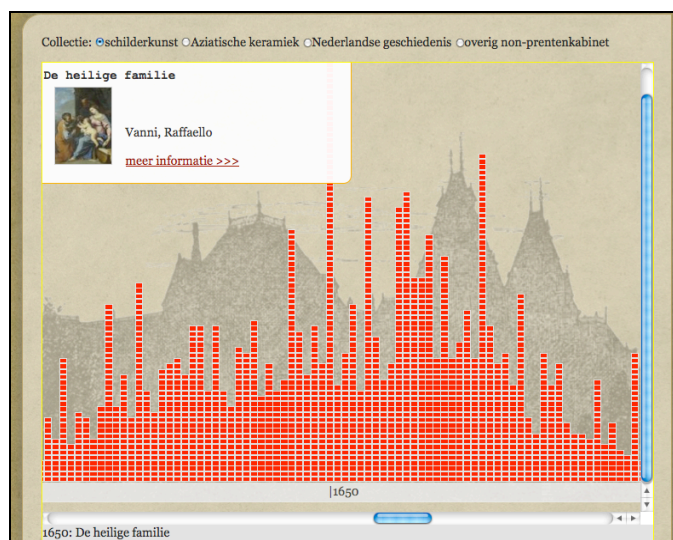
Wat voor toepassingen kunnen gemaakt worden met open cultuur data? Op basis van de ingezonden applicaties van Apps voor Nederland waar open cultuur data is gebruikt, geven we een indruk van de mogelijkheden die er zijn. In totaal werden 13 applicaties van de ruim veertig inzendingen met open cultuur datasets gemaakt. Op OpenCultuurData.nl wordt een overzicht van toepassingen gegeven¹². Hier kan een actueel overzicht gevonden worden van open cultuur data-toepassingen en de links naar alle onderstaande applicaties. Tevens worden de meeste van de ingezonden applicaties toegelicht in een zogenaamde ‘Demoblog’¹³, deze wordt geschreven door de developers zelf. Hieronder wordt uit ervaringen beschreven in deze Demoblogs geput.

Collecties doorzoekbaar maken

Verreweg de meeste applicaties zijn gemaakt om de data beter of op een nieuwe manier doorzoekbaar te maken. Door de data bijvoorbeeld te koppelen aan een classificatiesysteem, aan data van andere culturele instellingen of door nieuwe technologieën als spraakherkenning te koppelen aan de data. Ronald Klip, ontwikkelaar van contente content en de Rijksmuseum Tijdlijn applicatie: *“[Ik wilde iets maken] waardoor je op andere manieren de collectie van het Rijksmuseum kunt ontsluiten dan de opties die het museum zelf al op de website biedt. Standaard manieren zoals zoeken op titel of op naam van de kunstenaar: die zijn er al. Kenmerkend van die manieren is dat je vindt wat je zoekt – op zich wenselijk, maar het kan ook interessant zijn om iets te vinden wat je juist niet zoekt. Het leek me daarom leuk om een tijdlijn te ontwikkelen, omdat dit weer een andere manier is om de collectie te benaderen”*. Hieronder een alfabetisch overzicht van applicaties gemaakt om de doorzoekbaarheid van collecties te vergroten.

¹² <http://www.opencultuurdata.nl/apps/>

¹³ Lees alle Demoblogs hier: <http://www.opencultuurdata.nl/category/blog/demo/>.



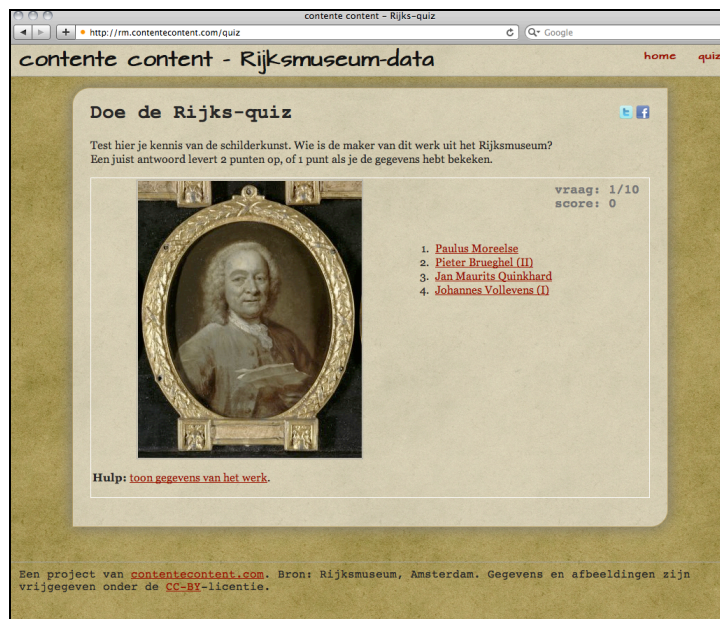
Afbeelding 2. Screenshot Rijksmuseum Tijdlijn applicatie.

- Arkyves ICONCLASS Browser: de collectie van het Rijksmuseum is doorzoekbaar gemaakt met gebruik van het ICONCLASS classificatiesysteem. Developer: Etienne Posthumus, Arkyves. Gebruikte dataset: Rijksmuseum.
- Connected Collection: applicatie die relaties legt tussen collecties van verschillende musea met gebruik van open data. Developer: Wout van Wezel, Cit. Gebruikte datasets: o.a. Amsterdam Museum, Nationaal Archief, Open Beelden (Beeld en Geluid), Rijksmuseum.
- Gezichten van het Rijksmuseum: gezichtsherkenningstechniek maakt het mogelijk gezichten op schilderijen, tekeningen en prenten van het Rijksmuseum op een nieuwe manier te bekijken en daarmee de museumcollectie op een 'menselijke' manier te verkennen. Developers: Hans van den Berg en Arjan den Boer, ab-c media. Gebruikte dataset: Rijksmuseum.
- Rijksmuseum Tijdlijn: doorzoek de collectie van het Rijksmuseum middels een tijdlijn in diverse categorieën. Developer: Ronald Klip, Contente Content. Gebruikte dataset: Rijksmuseum.
- Het Virtuele Rijksmuseum: virtueel Rijksmuseum gemaakt met alle schilderijen, met de juiste verhoudingen in grootte. Developer: Hay Kranen. Gebruikte dataset: Rijksmuseum.

Games

Een ander type toepassing die kan worden ontwikkeld met open cultuur data is de game. Op basis van de onderwerpen van de individuele objecten van het Rijksmuseum zijn twee quizzes ontwikkeld waarbij ofwel de juiste naam van het werk geraden moet worden, ofwel de maker van het werk. In de praktijk is dit voor de eindgebruiker moeilijker dan het op het eerste gezicht lijkt, zo geeft Ronald Klip, die naast de Rijksmuseum Tijdlijn ook de applicatie Rijks-Quiz maakte, aan: *"Het is een pittige quiz geworden. Zelf heb ik nog niet meer dan 13 punten gehaald, van de 20 die maximaal te behalen zijn. Het helpt wel als je aan het schilderij kunt zien of de maker een Italiaan of een Nederlander is. Of dat het werk niet*

Middeleeuws is: dan kun je namen als 'Dirck' of 'Meester van ...' wegstrepen". Je zou je natuurlijk kunnen voorstellen dat door meer (meta) data toe te voegen aan die van het Rijksmuseum, je de vragen veelzijdiger zou kunnen maken waardoor het aantrekkelijker wordt om de quiz te spelen. Door bijvoorbeeld informatie toe te voegen over de nationale identiteit van de maker, welke andere werken er in die tijd gemaakt werden, gerelateerde kunstenaars of andere metadata waar het Rijksmuseum zelf misschien niet direct over beschikt, kunnen er gerichte vragen gesteld worden en wordt de eindgebruiker geholpen met het beantwoorden van de vragen. Hieronder de twee games die gemaakt zijn op basis van open cultuur data.



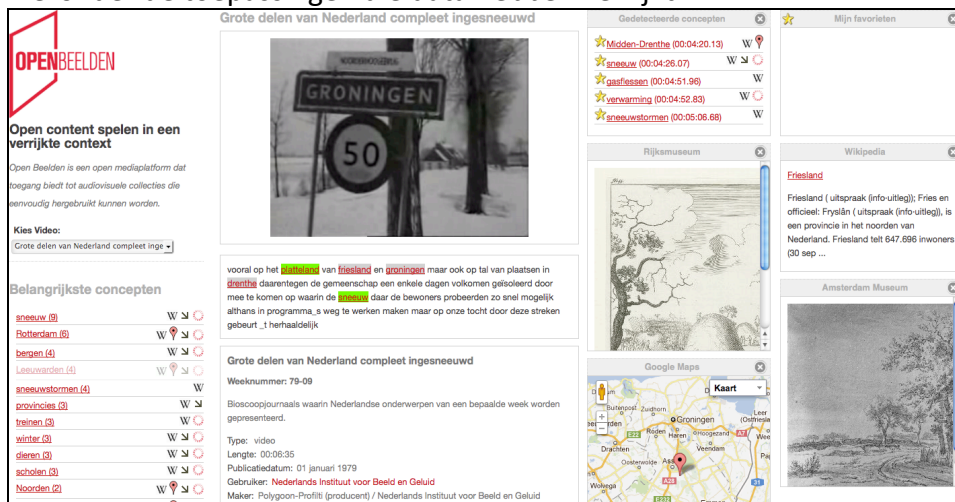
Afbeelding 3. Screenshot Rijks-Quiz applicatie.

- Rijks-Quiz van Contente Content: daagt iedereen uit zijn of haar kennis van de schilderkunst te testen. Developer: Ronald Klip, Contente Content. Gebruikte dataset: Rijksmuseum.
- Rijksmuseum Quiz: quiz over de collectie van het Rijksmuseum: kies uit 4 opties de juiste titel van een gegeven afbeelding. Developer: Nico Witteman. Gebruikte dataset: Rijksmuseum.

Verrijken met nieuwe bronnen

Andere toepassingen zijn gemaakt door de aangeboden data te verrijken met nieuwe datasets of het bijdragen van de datasets aan andere platforms waar al meer data te vinden is, bijvoorbeeld door de open data te verrijken met Wikipedia-artikelen of deze op Google Maps beschikbaar te maken. Jaap Blom, ontwikkelaar bij Beeld en Geluid, verrijkte historisch videomateriaal van het Open Beelden-platform met nieuwe datasets, zodat het mogelijk wordt dat als er iets over Drenthe te zien is in een filmpje, hierbij ook een Wikipedia-artikel over Drenthe getoond wordt. Jaap Blom zegt hier het volgende over: *“Aan de basis van de werking van de demo liggen de spraaktranscripten van deze video’s, die zijn geleverd door de spraakherkenningssoftware van X-MI. Het idee voor de demonstratie ligt in het contextualiseren van bekeken video’s om de kijker zodoende te voorzien van nuttige,*

leuke en onverwachte achtergrondinformatie over hetgeen er gesproken wordt in het filmpje”. Hieronder de toepassingen die data hebben verrijkt.



Afbeelding 4. Screenshot Open Beelden video's verrijkt met open data.

- Open Beelden video's verrijkt met Open Data: de demo laat zien hoe je een video in een verrijkte context kunt afspelen door open databronnen te koppelen aan termen die voorkomen in de spraaktranscripten van de video. Developer: Jaap Blom, Beeld en Geluid, afdeling R&D. Gebruikte datasets: Amsterdam Museum, Open Beelden (Beeld en Geluid), Rijksmuseum.
- Rijksmonumenten: website waarop de 61.000 rijksmonumenten in Nederland te doorzoeken zijn, verrijkt met informatie van Wikipedia en een koppeling met Google Maps. Developer: Terry van Elk, Elk ICT Services. Gebruikte dataset: o.a. Rijksmonumenten.

Mobiele applicatie en tablets

Als je aan applicaties of apps denkt, dan denk je in eerste instantie toch wel aan toepassingen voor je mobiele telefoon of een tablet. Ook deze zijn gemaakt met open cultuur data. Een (mobiele) applicatie waarbij je op basis van locatie stukken uit de collectie kunt bekijken. Bij de applicatie Vistory is er ook de mogelijkheid om nieuwe (user generated) content toe te voegen, zodat de data verder verrijkt kan worden door de eindgebruiker.

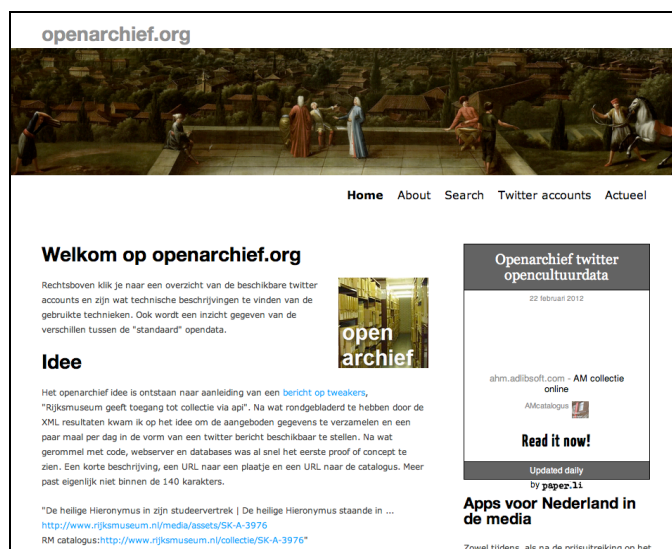


Afbeelding 5. Screenshot Vistory applicatie.

- Historische kaarten op de iPad: iPad app gemaakt waarmee je historische kaarten uit de 16e tot 19e uit de collectie van het Nationaal Archief kan bekijken. Developers: Kars Alfrink en Chris Eidhof. Gebruikte dataset: Nationaal Archief.
- Rijksmonumenten.info: Mobiele app met alle 61.000 rijksmonumenten van Nederland mobiel beschikbaar, met hierbij ook informatie van de Rijksdienst voor het Cultureel Erfgoed en Wikipedia, afbeeldingen van Wikimedia Commons, RCE, Flickr en van historische archieven en musea. Developers: Hans van den Berg en Arjan den Boer, ab-c media. Gebruikte dataset: o.a. Rijksmonumenten.
- Vistory – The Interactive Historical Video App: gebruikt geolocatie technologie op smartphones om te bepalen waar je bent en laat je op basis hiervan historische films zien die in de buurt zijn opgenomen. Developers: Jonathan Carter, Paul Manwaring, Jeroen van der Linde, Martin Elshout, Deniz Tezcan (gezamenlijk: Glimworm IT). Gebruikte dataset: Open Beelden (Beeld en Geluid).

Sociale mediaplatforms

Het koppelen van open data aan sociale platforms en daar een sociale dimensie aan geven (technisch gezien) is een vorm van hergebruik die we ook terug zien in de gemaakte applicaties. Koppelingen met Twitter, Facebook of bijvoorbeeld Flickr worden al gebruikt bij veel culturele instellingen, ook is Wikipedia als sociaal platform volwassen geworden. Bij de hierboven genoemde Rijks-Quiz wordt bijvoorbeeld ook een koppeling gemaakt met Twitter en Facebook, waardoor de quiz ook een sociale dimensie kan krijgen. Als je punten behaald hebt in de quiz, kun je deze prestatie delen met je vrienden op Facebook en je volgers op Twitter, waardoor zij ook geïnteresseerd zouden kunnen raken in de collectie van het Rijksmuseum. Hans Nouwens ontwikkelde verder de OpenArchief Twitterstreams, waarbij gebruikers elke dag één of meerdere tweets krijgen waarin ze geattendeerd worden op voor hen interessant cultuuraanbod. Hij zegt hier het volgende over: *“Gebruikers kunnen zich abonneren op een van de vele onderwerpen die worden aangeboden. De onderwerpen variëren van voetbal tot historische penningen en trapgevels. Per onderwerp worden per dag een aantal tweets verstuurd”*.



Afbeelding 6. Screenshot OpenArchief Twitterstreams.

- Openarchief Twitterstreams: Tweets over het Nederlandse culturaanbod. Developer: Hans Nouwens. Gebruikte datasets: o.a. Amsterdam Museum, Rijks.

Kortom, in een korte tijd zijn met beperkte middelen veel nieuwe soorten toepassingen gemaakt waar je als culturele instelling misschien niet zo snel aan had gedacht. Hoewel de toepassingen nog relatief eenvoudig lijken, zijn dit eerste (werkende) prototypes binnen een genre waarbij nieuwe mogelijkheden in de toekomst nog verder verkend kunnen worden en bij kunnen dragen aan het faciliteren van brede toegang tot je data en het stimuleren van hergebruik in nieuwe contexten voor nieuwe publieken.

Meerwaarde open data

Wat kun je vervolgens als culturele instelling doen met de toepassingen die gemaakt zijn met jouw open datasets? Wat kun je eruit halen? De hierboven genoemde argumenten die vaak worden gebruikt om tot het besluit te komen data open te stellen komen overeen met de ervaringen: door de collectie beschikbaar te stellen neemt de toegang en het hergebruik toe en ontstaat er grotere samenwerking tussen instellingen. Hierdoor wordt de waarde van je collectie vergroot, hoewel hierbij met name moet worden gedacht aan 'soft benefits', zoals hogere (online) bezoekersaantallen, contextualisering van collecties en objecten op externe platforms als Wikipedia en verrijking van metadata dan verhoging van financiële baten. Deze resultaten zullen hieronder worden toegelicht.

Toegang vergroten en verbeteren

Ten eerste, het zelf publiceren van open data kan, zoals in bovenstaande voorbeelden naar voren is gekomen, leiden tot het verder verrijken van je datasets, waardoor uiteindelijk de *toegang vergroot en verbeterd* kan worden. Musea en archieven zijn al lange tijd bezig met het verrijken van hun metadata door delen van hun collectie online te publiceren. Het Nationaal Archief publiceerde bijvoorbeeld een grote set foto's en data op de fotowebsite Flickr: The Commons¹⁴, waardoor veel extra informatie toegevoegd kon worden aan hun collectie door commentaar op foto's te lezen op Flickr (user generated content). De Deutsche Fotothek stelde veel foto's ter beschikking aan Wikimedia Commons, de

¹⁴ <http://www.flickr.com/photos/nationaalarchief/>.

beeldbank van Wikipedia¹⁵. Wikimedianen gaven commentaar op de duizenden bestanden, waardoor de Deutsche Fotothek haar metadata kon verrijken. Verder werden de foto's ook in de Wikipedia-artikelen gebruikt, zoals getoond in Afbeelding 6 hieronder.

Elektronische computers [bewerken]

Door de Tweede Wereldoorlog kreeg de ontwikkeling van computers een snelle vlucht. In het Verenigd Koninkrijk werd van de Colossus gebruikgemaakt om Duitse geheime codes te kraken, onder andere die van de Enigma-codeermachine. De Colossus was de eerste elektronische computer, gebruikmakend van elektronenbuizen. De eerste computer in de VS was de ENIAC, die enkele klaslokalen in beslag nam. De eerste computer in Nederland was de ARRA bij het Mathematisch Centrum. De eerste computer in een commerciële omgeving was de Miracle, een Ferranti Mark I bij het Shell-laboratorium in Amsterdam. In de periode dat het permanente geheugen (de harde schijf) nog niet algemeen bestond, was het invoeren van gegevens of programma's in een computer vrij moeizaam. Dit gebeurde oorspronkelijk met schakelaartjes en ponsband, nog iets later met ponskaarten, en in een nog later stadium met magneetbanden.

De computers in de jaren 1950-1980 waren vooral mainframes: zeer grote computers, waar honderden tot duizenden gebruikers gelijktijdig op konden werken. Vooral banken en verzekeringsmaatschappijen gebruikten zulke mainframes op grote schaal. De mainframe was verbonden met de gebruikers via een simpele applicatie op een bureaucomputer (vroeger ook wel via een zogenaamde domme terminal). De mainframe is met de komst van de kleine computers nog niet volledig uitgestorven en wordt nog steeds gebruikt door professionele instellingen. De bekendste bouwer van mainframes is IBM.

Miniaturisatie [bewerken]

Met de enorme ontwikkeling van de elektronica en de halfgeleiders, toegepast in transistoren, kon de computer veel kleiner en sneller worden. Later werden de transistors geïntegreerd in een geïntegreerde schakeling. De microprocessor is zo'n geïntegreerde schakeling. Hoewel microprocessor-gebaseerde computers zoals de Commodore PET (Personal Electronic Transactor) en de Apple II al vanaf het midden van de jaren zeventig opgang deden, was de IBM PC uit 1981 het eerste systeem dat expliciet met de naam personal computer op de markt werd gebracht. De pc werd steeds goedkoper en gemakkelijker te gebruiken waardoor steeds meer bedrijven en huisgezinnen er een kochten. De ontwikkelingen gaan voort, zakenmensen gebruiken veelal een laptop om met hun computer op stap te gaan. De steeds verdere miniaturisering leidt er toe dat de kleine Personal Digital Assistant (pda) met steeds meer mogelijkheden in beeld komt. Ook veel apparaten zoals wasmachines, videorecorders, digitale camera's en dergelijke bevatten tegenwoordig een computer om allerlei zaken te regelen, deze worden dan meestal een ingebed systeem of - in het Engels - embedded system genoemd.



Vrouw gebruikt computer voor haar werk (1983) [bewerken]

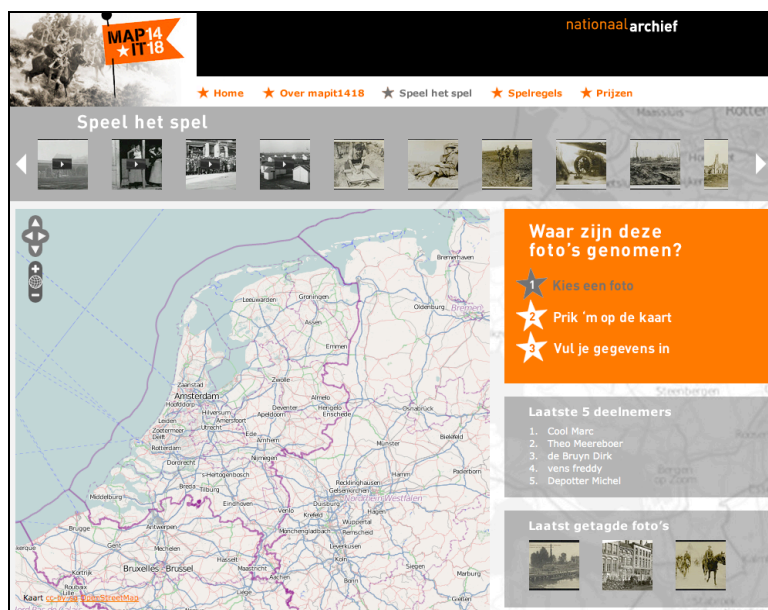
Afbeelding 7. Wikipedia-artikel met foto uit Fotothek collectie.

In het kader van Beelden voor de Toekomst hebben het Nationaal Archief en het Nederlands Instituut voor Beeld en Geluid enkele crowdsourcingprojecten uitgevoerd om user generated metadata aan hun collectie toe te voegen. De aangeleverde content voor deze projecten is overigens niet onder een open licentie beschikbaar. Waisda?¹⁶ is het video tagging spel van Beeld en Geluid waarbij mensen spelenderwijs video's beschrijven met trefwoorden. Mapit1418¹⁷ is een website waarbij mensen gevraagd wordt de exacte locatie van een foto aan te geven op de kaart. Beide zorgen voor enorme verrijking van het materiaal. Door deze metadata toevoegingen is het materiaal sneller te vinden, te gebruiken en in context te plaatsen.

¹⁵ http://commons.wikimedia.org/wiki/Category:Images_from_the_Deutsche_Fotothek?uselang=nl.

¹⁶ Zie het project met het televisieprogramma Man Bijt Hond: <http://woordentikkertje.manbijthond.nl/>.

¹⁷ <http://www.mapit1418.nl> (niet meer geactualiseerd).



Afbeelding 8. Screenshot Mapit1418.nl.

Open data wordt natuurlijk niet alleen hergebruikt door particuliere ontwikkelaars. Collectiehoudende instellingen kunnen zelf ook gebruik maken van open data. Doordat de producent van de data bekend is, kunnen datasets als autoriteit beschouwd worden. Het Amsterdam Museum en het Rijksmuseum kunnen bijvoorbeeld gebruik maken van elkaars datasets omdat ze beide informatie hebben over Nederlandse kunstenaars uit de Gouden Eeuw. Door gebruik te maken van elkaars datasets wordt de verzamelde informatie accurater, betrouwbaarder en waardevoller. Europeana's verzamelde dataset zal een grote rol spelen in het verrijken van metadata bij culturele instellingen. Europeana heeft van meer dan 20 miljoen culturele objecten in Europea metadata verzameld. Deze datasets worden door Europeana gecombineerd waardoor een rijke gemetadateerde virtuele collectie ontstaat. Culturele instellingen kunnen zelf op zoek gaan naar meer informatie over de objecten die ze zelf bezitten. De gecombineerde metadata van meer dan tweeduizend erfgoedinstellingen in Europa zijn toegankelijk voor alle partners van Europeana en is later dit jaar via een API onder CC0 beschikbaar.

Europeana gaat haar data daarnaast ook aanbieden als linked open data. Bij linked (open) data gaat het niet alleen maar om de vrije toegang tot de data (open), maar ook om de interoperabiliteit tussen meerdere datasets. Linked open data maakt het mogelijk om zoekmachines te maken die informatie kunnen vinden in verschillende databases, zonder dat daarvoor veel extra handelingen nodig zijn. Zo zou informatie over geportretteerde bekende Nederlanders verbonden kunnen worden met rijksmonumenten zonder dat een ontwikkelaar hier heel veel tijd in hoeft te stoppen. Een culturele instelling kan daardoor haar collectie-informatie in theorie beperken tot het open beschikbaar stellen van de objecten van een collectie zelf, overige relevante informatie wordt automatisch geleverd door andere datasets. Tot op heden is dit echter nog geen standaard praktijk voor culturele instellingen. Er zijn namelijk nog maar weinig culturele datasets beschikbaar als linked open data.

Verhogen bereik

Een tweede antwoord op de vraag wat je als culturele instellingen kunt halen uit de open data die je zelf publiceert, is dat je het *bereik van je collectie significant kunt vergroten*. Uit onderzoek van Kennisland¹⁸ blijkt dat open cultuur data in zijn 'ruwe' vorm niet direct een middel kan zijn om de bezoekersaantallen op je eigen website te vergroten. Het bezoek aan en bereik van je collectie op andere platforms of websites buiten je eigen website of beeldbank kun je daarentegen wel noemenswaardig vergroten. Het Nationaal Archief en Spaarnestad Photo stelden bijvoorbeeld ruim 1.000 foto's beschikbaar aan Wikimedia Commons¹⁹. Kennisland onderzocht het effect hiervan op de bezoekersaantallen op de eigen beeldbank. In maart 2012 werden Wikipedia-artikelen met deze foto's ruim 500.000 keer bekeken. De Wikimedia-foto's werden op beeldbank van het Nationaal Archief en Spaarnestad Photo niet vaker bekeken dan andere foto's. De conversie van de bezoekers van Wikipedia naar de beeldbank van het Nationaal Archief en Spaarnestad Photo bleven dus uit. Dit zou kunnen liggen aan het feit dat wanneer een afbeelding bij een artikel op Wikipedia gepubliceerd wordt, de verwijzing naar de culturele instelling en de link naar het werk niet mee gepubliceerd wordt. De eindgebruiker moet daardoor een aantal keer doorklikken voordat hij bij de 'bron' uitkomt. In het onderzoek hebben slechts 114 bezoekers door geklikt naar de website van het Nationaal Archief en Spaarnestad Photo. Het onderzoek wees verder uit dat een enkel bezoek aan de foto's van de beeldbank gelijk stond aan maar liefst 17.000 bezichtigingen van hetzelfde materiaal op Wikipedia-artikelen. Een significante vergroting van het bereik voor het Nationaal Archief en Spaarnestad Photo.

Geld verdienen?

Veel culturele instellingen verkeren in de veronderstelling dat je ook wezenlijke economische baten kunt genereren met alle data die je in huis hebt. Het massadigitaliseringstraject Beelden voor de Toekomst dat in 2007 startte met het digitaliseren en beschikbaar stellen van een kerncollectie van de Nederlandse audiovisuele geschiedenis had bij aanvang van het project een terugverdienverplichting. Het geld dat verdiend zou worden met het gedigitaliseerde materiaal zou weer terugvloeien naar de overheid. Het project heeft inmiddels na vier jaar uitgewezen dat het project wel degelijk maatschappelijke waarde creëert. De elektronische verspreiding van content en metadata levert echter geen grote economische baten op. Onderzoek van TNO wees uit dat er bij het beschikbaar stellen van gedigitaliseerd cultureel erfgoed vaak geen rendabele online business modellen zijn. Ook staat een verdienmodel vaak de missie van een culturele instelling in de weg. De onderzoekers van TNO concluderen als volgt²⁰:

“De terugverdienverplichting is volgens de onderzoekers onrealistisch, gezien de auteursrechtelijke obstakels en de huidige marktontwikkelingen. Met de opkomst van platforms als YouTube is de norm voor digitale audiovisuele diensten ‘gratis’. Bovendien valt de bereidheid om te betalen binnen het onderwijs tegen, mede omdat de overheid haar beleid inmiddels heeft gericht op gratis beschikbaarheid van leermiddelen die publiek zijn gefinancierd. Ondanks de achterblijvende economische

¹⁸ Download het onderzoeksrapport via <http://www.kennisland.nl/filter/nieuws/effectmeting-persfotos-veel-vaker-bekeken-in-context>.

¹⁹ Zie <http://beeldenvoordetoekomst.nl/nl/activiteiten/nationaal-archief-joins-wikipedia>.

²⁰ Download het onderzoeksrapport via <http://beeldenvoordetoekomst.nl/nl/news/tussentijdse-evaluatie-beelden-voor-de-toekomst>.

opbrengsten, ziet TNO wel aantoonbare maatschappelijke baten, en beveelt aan deze meer inzichtelijk te maken.”

Op basis van de ervaringen van Beelden voor de Toekomst kan gesteld worden dat het onrealistisch is om te verwachten dat je als culturele instelling geld kan verdienen met diensten die culturele data ontsluiten. Toch ligt het streven naar maatschappelijke baten niet altijd voor de hand voor culturele instellingen. Maarten Brinkerink (Nederlands Instituut voor Beeld en Geluid): *“Er bestaat uiteraard een spanningsveld tussen enerzijds ‘open’ projecten en de commerciële exploitatieprojecten anderzijds, omdat in beide gevallen data en content wordt ingezet waar wij zelf de rechten over bezitten. Er wordt door alle betrokkenen gewerkt aan een beleid en principeafspraken teneinde dit te stroomlijnen.”*

Voorbij de apps..?!

Culturele instellingen en specifiek erfgoedinstellingen (met name archieven en musea) zijn bedoeld om voor toekomstige generaties erfgoed te beheren en deze toegankelijk te maken voor het publiek. Open data zal dus op lange termijn ook bij moeten kunnen dragen aan deze missie. Het is natuurlijk leuk om een heleboel apps te maken, maar hoe draagt het ook structureel bij aan een (verbeterde) toegang tot de data die in huis is? Met andere woorden, welke elementen zien we in de verdere verduurzaming van het open cultuur data veld? We vroegen de culturele instellingen die als eerste participeerden in het Open Cultuur Data-netwerk of en hoe ze de open data in de toekomst verder wilden aanbieden en wat ze daarvoor nodig hebben. De volgende drie punten kwamen naar voren.

Nieuwe samenwerkingen met developers aangaan

Een belangrijk punt is de samenwerking met de developerscommunity buiten de muren van je eigen instelling. Op een hackathon kan dit beginnen met iemand die een applicatie met jouw open cultuur data maakt. Deze incidentele samenwerking zou kunnen uitgroeien tot een langdurige samenwerking of tot samenwerking met weer nieuwe partijen. Arjan de Boer en Lizzy Jongma zeggen hier het volgende over:

Arjan den Boer (ab-c media): *“We blijven de Rijksmonumenten API aanbieden”.*

Lizzy Jongma (Rijksmuseum): *“Onze collectie [is] door een aantal grote partijen als Kennisnet en Artstor geharvest en kunnen wij in samenwerkingsprojecten zoals nu met de VU (CATCHPlus) gebruik maken van de API om data te leveren”*

Meer open data ontsluiten

Een tweede punt dat naar voren kwam, was om in de toekomst (samen) verder te blijven werken aan het openstellen van meer data. Hoe meer open beschikbaar komt, hoe beter we (elkaar kunnen helpen om) collecties toegankelijk kunnen maken en verrijken. Een van de dingen die hiervoor moet gebeuren volgens het Rijksmuseum is auteursrechtenonderzoek. Je moet natuurlijk wel toestemming hebben van rechthebbenden om data en content te publiceren. Lizzy Jongma en Arjan de Boer zeggen hier het volgende over:

Lizzy Jongma (Rijksmuseum): *“Wij gaan door, wij willen in de toekomst meer sets en meer objecten beschikbaar stellen. Wij werken verder op het gebied van auteursrecht (o.a. door middel van het achterhalen van rechthebbenden, het opstellen van een*

separaat auteursrechtcontract en een verbeterde weergave en uitleg op de website) en hopen binnenkort tijd en geld te hebben om de techniek beter op orde te krijgen. Wij willen daarnaast graag onze kennis met derden delen en ervaring uitwisselen.”

Arjan den Boer (ab-c media): *“We blijven de Rijksmonumenten API aanbieden en zijn deze momenteel aan het verbeteren, zodat de data completer en actueler is. Graag zouden we de API uitbreiden met andere relevante bronnen, bijvoorbeeld op het gebied van architectuur of hergebruik.”*

Aansprekende voorbeelden

Open data is een onderwerp dat voor veel professionals bij culturele instellingen weinig tot de verbeelding spreekt. Het creëren van werkende voorbeelden en applicaties helpt om inzicht hierin te geven. We zouden samen meer moeten experimenteren en onderzoeken met het maken van nieuwe toepassingen om te laten zien hoe open data bij kan dragen aan het breder toegankelijk maken en verrijken van cultuur. Daarnaast wordt aangegeven dat er meer onderzoek gegenereerd zou moeten worden naar de effecten van open data en ook gewerkt zou moeten worden aan modellen om deze effecten in kaart te brengen. Wat zou je moeten meten? Het aantal bezoekers aan je data, op je eigen website of aan je fysieke gebouw? Lizzy Jongma en Maarten Brinkerink zeggen hier het volgende over:

Lizzy Jongma (Rijksmuseum): *“Wij hebben vanaf het begin getracht onze collegae uit te leggen wat Open Data is en wat er mee kan gebeuren. Maar, het is een technisch verhaal dat voor veel mensen pas gaat leven en begrijpelijk wordt als er resultaten zichtbaar zijn. Voor veel mensen is deze openheid van zaken een grote stap, maar de uitgangspunten worden door de directie gesteund (en zijn in beleidsplannen vastgelegd) en door voorlichting en presentatie van de behaalde resultaten ook intern gedragen.”*

Maarten Brinkerink (Nederlands Instituut voor Beeld en Geluid): *“Waar Beeld en Geluid nog wel behoefte aan heeft is modellen om de resultaten – en dus het succes – van dergelijke initiatieven systematisch te meten, om zo een gedegen onderbouwing van de effecten en uitkomsten van dergelijke initiatieven te kunnen rapporteren”.*

Voor de verdere verduurzaming van het open cultuur data veld is voor zowel het aangaan van nieuwe samenwerkingen, het ontsluiten van meer open data en het genereren van aansprekende voorbeelden, de rode draad dat culturele instellingen dit niet alleen kunnen doen. Hiervoor is participatie van collega-instellingen (zowel nationaal als internationaal) nodig en een community van betrokken developers die hierin mee willen experimenteren.

Afsluiting

Het stimuleren van hergebruik van data, het zo breed mogelijk ontsluiten en het verrijken van culturele data zijn voor culturele instellingen de drie hoofdargumenten om met open data aan de slag te gaan. Open data kan helpen om je publieke functie als instelling beter te vervullen. Het kan de toegankelijkheid vergroten en verbeteren, je bereik verhogen, maar open data kan niet als verdienmodel gezien worden. Hoewel je van tevoren niet helemaal kunt vaststellen welke toepassingen er gemaakt worden, kunnen we wel een aantal algemene adviezen geven over hoe je het maximale uit open data kunt halen.

OPEN CULTUUR DATA

DELEN. VERBINDEN. VERRIJKEN.

- **Dataset:** zorg ervoor dat je zoveel mogelijk en zo interessant mogelijke metadata én content beschikbaar stelt, metadata alleen is voor hergebruik minder interessant. Hoe meer open deze data beschikbaar gemaakt kan worden, hoe beter dat is voor het stimuleren van hergebruik en verrijking van je data met nieuwe relevante data. Voor de verdere verduurzaming van open cultuur data is het belangrijk om samen te werken met collega-instellingen bij het beschikbaar stellen van meer open cultuur data;
- **Community:** sluit aan op een community van open data experts en ontwikkelaars waar je samen nieuwe toepassingen mee kunt maken. Het stelt je in staat om nieuwe perspectieven op jouw dataset te genereren en (technische) expertise aan te wenden die je zelf niet direct in huis hebt. Voor de verdere verduurzaming van open cultuur data is het belangrijk om nieuwe samenwerkingsverbanden met ontwikkelaars te starten waarin op langere termijn samengewerkt kan worden;
- **Toepassingen:** probeer zoveel mogelijk experimenten te stimuleren met nieuwe en innovatieve toepassingen met jouw open data set(s). Toepassingen die je data breder toegankelijk maken, verrijken en bezoek kunnen vergroten. De mogelijkheden zijn legio. Toepassingen die collecties beter doorzoekbaar maken, op basis van de data games ontwikkelen, mobiele en tablet applicaties, toepassingen die je data verrijken met nieuwe bronnen en via sociale media platforms een sociale dimensie kunnen toevoegen aan je data. Voor de verdere verduurzaming van open cultuur data is het belangrijk om meer succesvolle voorbeelden te genereren en als instelling ook te gaan meten wat de impact van open cultuur data is.

Met andere woorden: deel, verbind en verrijk!

Licentie

Op dit werk is een Creative Commons Naamsvermelding-licentie van toepassing. Dit werk mag worden bewerkt, doorgestuurd, geüpload, gedownload, gekopieerd en anderszins verspreid, zolang de namen van de auteurs vermeld worden.



De gehele licentietekst is te vinden op: <https://creativecommons.org/licenses/by/3.0/nl/>

De bovenstaande informatie en aanbevelingen zijn gebaseerd op onderzoek en onze praktijkervaringen. Kennisland is er zich van bewust dat elke situatie zijn eigen mogelijkheden, beperkingen en randvoorwaarden kent. Daarom benadrukken wij het algemene karakter van bovenstaande, en accepteert Stichting Nederland Kennisland geen aansprakelijkheid voor de eventuele negatieve gevolgen van het toepassen van deze informatie.