

Eindrapport Persistente Identificatie & Open Cultuur Data

Auteur: Alina Saenko, Bert Lemmens, Rony Vissers

Datum: 31 maart 2016

Inhoudsopgave

Inhoudsopgave	2
Inleiding	3
Deel 1. Project 'Hoe word ik data-uitgever?'	4
Doelstelling.....	4
Methodologie.....	4
Conclusies en aanbevelingen	9
Deel 2. Project 'Event-based objectbeschrijvingen'	12
Doelstelling.....	12
Methodologie.....	12
Conclusies en aanbevelingen	26
Deel 3. Linked Open Data publicatie via Wikidata	28
Doelstelling.....	28
Methodologie.....	28
Conclusies.....	31
Bijlage 1: Events-data statistieken.....	33

Inleiding

In 2013 maakte PACKED vzw in opdracht van de Vlaamse Kunstcollectie (VKC) een doorlichting van de VKC Online Catalogus en de onderliggende systemen van de drie musea voor schone kunsten die deze catalogus voeden. Een gelijkaardige doorlichting werd gemaakt voor Contemporary Art Heritage Flanders (CAHF), maar hier ging de aandacht vooral naar het interne beheer van digitale collectiedata en beelden in de vier musea voor hedendaagse kunsten. Tot slot werd ook voor LUKAS een gelijkaardige oefening gemaakt rond de opslag en bewaring van digitale reproducties die aangeboden worden via lukasweb.be.

Omwille van de inhoudelijke verbanden tussen drie organisaties en de overtuiging dat relatief kleine organisaties uit de cultureel-erfgoedsector de technische uitdagingen die de digitalisering van hun werking meebrengt best collectief aanpakken, stelde PACKED vzw voor om met deze drie organisaties en de zeven aangesloten musea een gezamenlijke actieplan rond digitale duurzaamheid op te stellen. Vanaf eind 2013 werden de eerste stappen van dit actieplan uitgevoerd:

- het project 'Persistente Identificatie' (2014) identificeerde ruim 34.000 kunstwerken, hun vervaardigers, objectnamen, dateringen en bewaarinstellingen met een 'persistente URI';
- gelijktijdig met het project 'Persistente Identificatie' (2014) werden ook projecten opgestart door CAHF rond een verbeterd beschrijvingsmodel voor hedendaagse kunst, door VKC rond een gelaagd licentiemodel voor online publicatie van metadata en beelden;
- het vervolgtraject 'Persistente Identificatie en Open Cultuur Data' (2015) werkte verder aan (1) het operationeel maken van de persistente URI's, (2) de verdere verrijking van de metadata en het uittesten van het uitwisselingsformaat LIDO, (3) de publicatie van museumdata over kunstwerken als linked open data.

Dit rapport behandelt de eindresultaten van het traject 'Persistente Identificatie & Open Cultuur Data' (2015) en de drie deelprojecten: 'Hoe word ik data-uitgever?', 'Event-based objectbeschrijvingen' en 'Linked Open Data publicatie via Wikidata'.

Dit traject werd uitgevoerd in opdracht van het Departement Cultuur, Jeugd, Sport en Media van de Vlaamse overheid en ging van start op 17 november 2014.

Deel 1. Project ‘Hoe word ik data-uitgever?’

Doelstelling

Het doel van het deelproject ‘Hoe word ik data-uitgever?’ was:

- de opmaak van een plan van aanpak voor het activeren van de persistente URI's voor kunstwerken;
- de uitrol van een Resolver-service in de tien organisaties uit het voorgaande ‘Persistente identificatie’-project;
- de opmaak van een handboek voor het beheer van persistente URI's;
- de opleiding van een data-uitgever in elk van de tien organisaties.

Methodologie

Overlegmomenten en roadmaps

Dit deelproject werd uitgevoerd door PACKED-medewerkers Alina Saenko Joris Janssens en Bert Lemmens.

Het hele project werd uitgevoerd in de periode november 2014 - februari 2015.

In overleg met de collectieverantwoordelijken, de IT-verantwoordelijken en met aftoetsing bij het management, werd voor elke partnerinstelling een *roadmap* of stappenplan gemaakt voor het activeren van de persistente URI's en de implementatie van de Resolver-tool. Het vertrekpunt van deze *roadmaps* waren gesprekken met collectie- en IT-verantwoordelijken in elke partnerinstelling.

Tijdens deze overlegmomenten werden de volgende thema's behandeld:

- de rol van een ‘data-uitgever’ binnen elke instelling;
- documentatie van persistente URI's in het collectiebeheersysteem;
- installatie van de Resolver-tool;
- activatie van de persistente URI's.

De volgende overlegmomenten vonden plaats in het kader van het project ‘Hoe word ik data-uitgever’:

- 03.12.2014 – Groeningemuseum
- 08.12.2014 – M HKA
- 11.12.2014 – KMSKA
- 12.12.2014 – VKC en LUKAS
- 15.12.2014 – Middelheimmuseum (enkel met Jeroen De Meester)
- 16.12.2014 – S.M.A.K.
- 18.12.2014 – MuZEE
- 22.12.2014 – CVG
- 07.01.2015 – MSKGent
- 04.02.2015 – Middelheimmuseum (met de rest van de team)

Workshops

Daarnaast werden voor de data-uitgevers van elke partnerinstelling twee workshops georganiseerd.

Op de eerste workshop werden de volgende thema's behandeld¹:

- Wat is een data-uitgever en wat is zijn/haar plek in de organisatie? Waar is een Resolver voor nodig?
- Hoe kan je een CSV-importbestand voor de Resolver opstellen m.b.v. Open Refine?
- Op welke principes steunt de werking van de Resolver-tool en wat zijn de basisfunctionaliteiten? Hoe kan ik de persistente URI's activeren in de Resolver?
- Wat zijn praktische toepassingen van de Resolver en de actieve persistente URI's?

Tijdens de tweede workshop werd met behulp van een bordspel het traject van de publicatie van data over een nieuw kunstwerk op het web via de Resolver doorlopen². Het doel van dit spel was om stapsgewijs dit traject te ontdekken vanaf de registratie tot de verspreiding van persistente URI's. De startsituatie voor elke instelling werd gedefinieerd aan de hand van de reële stand van zaken in dit project, waardoor elke data-uitgever voor zichzelf kon definiëren wat de volgende stappen zijn in de eigen Roadmap.

Volgende thema's werden daarbij behandeld:

- Wat is het Data-uitgevershandboek? Wanneer moet je het gebruiken en actualiseren?
- Waarom bewaar je pURI's in je collectiebeheersysteem en wanneer doe je dat?
- Wanneer gebruik je de Resolver en hoe maak je persistente URI's aan?
- Hoe beslis je welke webdocumenten je aan persistente URI's koppelt? Hoe ga je om met de beperkingen van intellectuele eigendomsrechten?
- Hoe en aan wie verspreid je de kennis over persistente URI's binnen je eigen instelling?
- Wie zijn de externe gebruikers van persistente URI's? Hoe zorg je voor een zo groot mogelijke verspreiding en hergebruik van je persistente URI's?
- Welke invloed hebben de veranderingen in je interne IT-infrastructuur op de toegang tot data en beelden via persistente URI's?

Handboek

Voor elke partnerinstelling werd er een Handboek gemaakt³. Een Handboek is een hulpinstrument voor data-uitgevers bij het online publiceren van data via persistente URI's. In dit document worden de bijhorende procedures en het gebruik van de Resolver beschreven, m.n.:

- een definitie van het begrip 'persistente URI';
- een functieomschrijving van de data-uitgever;
- de werking van de Resolver;
- een actieplan voor de implementatie van persistente URI's in de instelling.

Daarnaast werd ook een Functieprofiel van een data-uitgever opgesteld waarin de competenties en verantwoordelijkheden worden vastgelegd⁴.

¹ Slides workshop 1: https://www.projectcest.be/images/5/58/20150130_Workshop1_DataUitgever_V1_0.pdf

² Slides workshop 2: https://www.projectcest.be/images/d/da/20150227_Workshop2_DataUitgever_V1_0.pdf

³ Modelhandboek: http://projectcest.be/images/3/39/20150422_Modelhandboek_Datauitgever_V1_1.pdf

⁴ Functieprofiel data-uitgever: https://www.projectcest.be/images/7/77/20160330_Functieprofiel_Datauitgever_v0_8.pdf

Op basis van deze Handboeken werden de volgende museummedewerkers aangeduid als verantwoordelijk voor of belast met de publicatie van persistente URI's voor kunstwerken in hun collectie⁵:

- Groeningemuseum: Guenevere Souffreau;
- MSKGent: Carine Van Bruwaene, Els Schepens;
- KMSKA: geen data-uitgever, toewijzing rol wordt intern besproken;
- VKC: Matthias Vandermaesen;
- LUKAS: geen data-uitgever, geen concrete stappen ondernomen in dit project;
- CVG: Pascale Van Kerckvoorde;
- Middelheimmuseum: geen data-uitgever, aanwerving data-uitgever word intern besproken;
- MuZEE: Barbara De Jong;
- S.M.A.K.: vroeger: Veronique Despodt, nu: toewijzing rol wordt intern besproken;
- M HKA: Evi Bert;
- M Museum: Alexandra Pauwels (pas in september 2015 in het project gestapt, een grote inhaalactie uitgevoerd).

Documentatie van persistente URI's

In de loop van dit traject werden de volgende acties ondernomen om pURI's te documenteren in de collectiebeheersystemen:

ADLIB

Tussen januari en december 2015 werden verschillende acties ondernomen om het voorstel voor de implementatie van persistente URI's in Adlib Museum onder de aandacht te brengen van het softwarebedrijf Axiell en de gebruikers van Axiell. Daarbij werden ook internationale contacten gelegd en werd actief gebruik gemaakt van het Nederlandse initiatief rond persistente URI's in het Netwerk Digitaal Erfgoed:

- 2014: een voorstel voor Axiell met toelichting bij de nieuwe velden uitgeschreven;
- 16.01.2015: bijeenkomst met Erfgoedinzicht om de aanpassingen in Adlib te bespreken;
- 21.01.2015: werkgroepbijeenkomst van Adlib-deskundigen rond aanpassingen in Adlib. Aanwezig: Jeroen De Meester (Applicatiebeheerder Adlib Stedelijke Musea Antwerpen), Christoph Coevoet (Erfgoedinzicht, PWV), Marc Cornelis (Erfgoedinzicht, POV), Mario Commeyne (Kunsten en Erfgoed, Collectie Vlaamse Gemeenschap), Bert Lemmens (PACKED vzw), Alina Saenko (PACKED vzw);
- werkende testversies met geïmporteerde persistente URI's bestaan in Adlib van het Middelheimmuseum en Collectie Vlaamse Gemeenschap;
- samenwerking met het gelijkaardige project van DEN rond persistente identificatie in Nederland (in het kader van het NDE-traject, projectleider: Gijsbert Kruithof, Nationaal Archief) om de grote Nederlandse musea achter ons voorstel voor nieuwe velden in Adlib te krijgen;

⁵ Status bij publicatie van dit eindrapport i.e. 25 februari 2016.

- via DEN Franse, Britse en Deense musea (gebruikers van Adlib) gecontacteerd om te polsen of ze bij ons voorstel willen aansluiten en samenwerken om Axiell van deze aanpassingen te overtuigen;
- via VKC het NEMO-netwerk van musea gecontacteerd te polsen of Adlib-gebruikers in dit netwerk bij ons voorstel willen aansluiten en samenwerken om Axiell van deze aanpassingen te overtuigen;
- 30.03.2015: bijeenkomst met IT-verantwoordelijken van Erfgoedinzicht om het voorstel nog een keer te bekijken;
- 30.04.2015: tweede presentatie van het voorstel voor PID-velden tijdens Adlib-gebruikersgroep in Breukelen;
- 17.09.2015: vergadering met Rene van den Heuvel (Axiell) en Livia Snauwaert (Erfgoedinzicht) tijdens 'Informatie aan zee' i.v.m. de ontwikkeling van PID-velden in Adlib;
- 27.10.2015; NDE meeting 'Beslissingsboom persistente identificatie';
- 05.11.2015: gesprek via Skype met Rene van den Heuvel (Axiell) om ons voorstel voor nieuwe velden gedetailleerd te bespreken: dit resulteerde in een consensus over de nood van het opnemen van die velden in Adlib, en Rene beloofde om een demoversie van toepassing van het voorstel uit te werken en aan PACKED voor te stellen;
- 25.11.2015: NDE meeting 'Waarom persistente identificatie?';
- 26.11.2015: PACKED vzw analyseert de demoversie: het gaat om de toevoeging van twee velden aan de interface en een elementair script om inventarisnummers te genereren op basis van het inventarisnummer;
- 07.12.2015: bespreking van de demoversie met Rene van den Heuvel (Axiell) tijdens DISH: afspraak om begin 2016 een nieuwe demoversie samen te bekijken.
- 16.12.2015 - NDE meeting 'Leveranciersmiddag'.

Begin 2016 kan enkel besloten worden dat de Adlib Museum gebruikers die deelnamen aan de voorbije projecten rond persistente identificatie niet beschikken over de noodzakelijke velden in hun collectiebeheersystemen om persistente URI's voor kunstwerken en aanverwante contextuele informatie te documenteren in hun collectiedata. Dit betekent concreet dat de persistente URI's voor deze musea voorlopig enkel bewaard worden in losse CSV-bestanden, waardoor ze moeilijk te actualiseren zijn en het risico groot is dat ze verloren zullen gaan. Het aanpassen van Adlib Museum met het oog op de documentatie van persistente URI's blijft een belangrijk aandachtspunt voor vervolprojecten omtrent persistente URI's en de datahub.

TMS

In 2015 werd aan KMSKA een concreet voorstel (in het Engels) overgemaakt voor de implementatie van velden voor persistente URI's in TMS. Volgens onze informatie ondernamen noch KMSKA, noch TMS verdere stappen om dit voorstel te implementeren.

Collective Access

Implementatie van velden voor persistente URI's werd reeds voltooid in 2014.

Ensembles van M HKA

In 2015 werd aan M HKA een voorstel overgemaakt voor de implementatie van velden voor persistente URI's in de Ensemblebank. Volgens onze informatie werd met dit voorstel geen verder actie ondernomen door M HKA.

Resolver

Versie 1.0 van de Resolver-tool werd ontwikkeld in het project Persistente identificatie (2014). Tijdens het eerste deelproject werd de Resolver uitgetest door de data-uitgevers en verder ontwikkeld. Alle aanpassingen werden gepubliceerd op de Github-pagina⁶ van deze tool. De volgende functionaliteiten werden in de loop van dit project aan de Resolver toegevoegd:

- een RESTful API, waarmee data-uitgevers in theorie vanuit hun eigen applicatie (CMS) updates aan de PIDs kunnen uitvoeren;
- documentatie, zowel voor de interne werking van de Resolver (documentatie van de code) als voor het gebruik van de Resolver (documentatie van installatie en configuratie);
- de oplossing voor diverse bugs die zich voordeden na de eerste uitrol;
- om de Resolver sneller te kunnen uitrollen ("deployment") naar partners zijn er ook aanpassingen gebeurd aan dat deel van de code (packages) en is er extra documentatie geschreven.

De huidige applicatie (versie 1.5.1) is volledig functioneel en kan gebruikt worden door alle organisaties.

De volgende functionaliteiten werden tijdens dit proces voorgesteld door het testpanel, maar werden op de lijst gezet voor een volgende ontwikkelfase:

- ondersteuning van meertaligheid;
- vereenvoudiging van deployments door aanpassingen in de code.

Op het einde van dit traject werden Resolvers geactiveerd bij de volgende partnerinstellingen:

- Groeningemuseum: geactiveerd;
- MSKGent: geactiveerd;
- KMSKA: testfase (koppeling met domeinnaam vormt nog een probleem);
- VKC: geactiveerd;
- CVG: geactiveerd;
- MuZEE: test-fase (update naar v1.5.1 nodig);
- S.M.A.K.: test-fase (update naar v1.5.1 nodig);
- M HKA: test-fase (update naar v1.5.1 nodig);
- M Museum: nog niet geactiveerd;
- LUKAS: nog niet geactiveerd;
- Middelheimmuseum: nog niet geactiveerd.

Op de stuurgroep van 28 oktober 2015 werd aan PACKED vzw de vraag gesteld om de instellingen te blijven ondersteunen bij de installatie van de Resolver en het activeren van persistente URI's. Deze ondersteunende acties zijn gedeeltelijk uitgevoerd in de loop van het derde deelproject en gedeeltelijk in januari-februari 2016.

⁶ <https://github.com/PACKED-vzw/resolver>

Conclusies en aanbevelingen

In de loop van dit deelproject hebben de data-uitgevers van alle partnerinstellingen (1) de persistente URI's voor kunstwerken via de Resolver leren activeren, (2) het volledige traject van publicatie van data en beelden via persistente URI's van begin tot einde doorlopen en (3) de principes en het potentieel van het gebruik van persistente URI's gezien. Als resultaat van dit project hebben de data-uitgevers een referentiedocument gekregen, het Handboek voor het online publiceren van data, dat zal bijdragen aan de nodige continuïteit indien de verantwoordelijkheid voor het uitgeven van data wordt overgedragen naar andere personen binnen de instelling.

In de loop van dit project werden de volgende vaststellingen gedaan.

1. Documentatie van URI's in collectiebeheersystemen

De implementatie van velden voor het documenteren van persistente URI's in de collectiebeheersystemen, m.n. Adlib Museum en TMS, verloopt uiterst moeizaam. Ondanks de intensieve communicatie met leveranciers en applicatiebeheerders van deze systemen, blijkt het moeilijk om op een open en constructieve manier de nodige aanpassingen aan de systemen en werkprocessen door te voeren.

Wat betreft de samenwerking met Axiell leeft bij PACKED vzw de overtuiging dat er bij de leverancier intern geen draagvlak bestaat om op een open en constructieve manier met PACKED vzw en partnerinstellingen samen te werken rond de documentatie van persistente URI's in Adlib Museum. Axiell communiceert naar klanten dat ze rekening houden met de voorstellen van PACKED vzw. Maar in de praktijk verloopt de dialoog over aanpassingen aan Adlib Museum enkel via informele kanalen. De dialoog verloopt ook steeds in kringetjes: eerst wordt aangekondigd dat het probleem is opgelost, dan blijft verdere berichtgeving uit en tot slot blijken er nog weinig tot geen concrete stappen genomen te zijn, waarna de vraagstukken die reeds bij aanvang van de dialoog beantwoord werden opnieuw op tafel worden gelegd.

Wat betreft de samenwerking met collectiemedewerkers in de musea leeft bij PACKED vzw de overtuiging dat er een fundamenteel gebrek is aan technische expertise in de instellingen om zelf verbeteringsprocessen met de leverancier doeltreffend te begeleiden. PACKED vzw en de VKC data manager adviseren voortdurend de collectiemedewerkers, maar dit vertaalt zich zeer moeizaam naar resultaatgerichte acties bij de leveranciers van de instellingen. In de communicatieketen ontstaan er misverstanden over de doelstelling en de denkwijze achter gevraagde aanpassingen, en de risico's en de impact van gevraagde aanpassingen worden verschillend ingeschat. De oorzaak hiervan is het gebrek aan technische expertise in de instellingen zelf. PACKED vzw en de VKC data manager zijn niet gemachtigd om rechtstreeks contact op te nemen met leveranciers omwille van bestaande afspraken. Dit bemoeilijkt een resultaatgerichte aanpak.

Aanbevelingen:

- Om de persistente URI's die gemaakt werden in het project 'Persistente identificatie' (2014) duurzaam te verankeren in de collectiebeheersystemen, blijft het noodzakelijk om in vervolprojecten tijd te investeren in de begeleiding van instellingen bij het aanpassen van hun collectiebeheersysteem. Bovendien moet ook de druk op leveranciers verder verhoogd worden door het voorstel

rond de documentatie van persistente URI's in collectiedata op de agenda te plaatsen bij zowel concurrerende leveranciers als bij instellingen en museumnetwerken.

- Om meer technische expertise in de instellingen onder te brengen moet de noodzaak van een data-uitgever in de instelling verder bepleit te worden. Een belangrijk instrument in dit pleidooi is het functieprofiel voor de data-uitgever. Dit profiel dient door PACKED veel concreter uitgewerkt te worden in nauwe samenwerking met het Departement CJSM en de betrokken musea. Opleiding, vaardigheden en de plaats van de data-uitgever in de organisatiestructuur dienen verder geconcretiseerd worden. Een tweede belangrijke strategie om data-uitgevers te introduceren in de museumwerking is om hun taken vast te leggen in informatiebeleidsplannen die het algemene informatiebeleid van musea beschrijven.
- Om het draagvlak voor de functie van data-uitgever te vergroten is het wenselijk om op korte termijn een voorbeeld te realiseren waarmee aangetoond wordt dat de data-uitgever de kwaliteit en efficiëntie van museumprocessen die gebruik maken van collectiedata beduidend verbetert.

2. Uitrol Resolver-netwerk

De uitrol van het Resolver-netwerk verloopt moeizaam. Hoewel de Resolver ontwikkeld werd met technologie die vrij beschikbaar is en de broncode en installatieprocedure uitvoerig gedocumenteerd werden, waren doorgaans meerdere contacten nodig tussen de leverancier en de collectiebeheerder vooraleer een Resolver operationeel werd. De ambitie was echter om een tool te ontwikkelen waar installatie en beheer een minimale werklast met zich meebrengen. De evaluatie van de uitrol van de Resolvers tot nog toe wijst de volgende knelpunten aan.

Bij de keuze voor de technologie ging PACKED vzw van het principe uit dat het gebruik van open en goed gedocumenteerde technologie waarborgt dat men niet afhankelijk is van één bepaalde ontwikkelaar en een vrije keuze heeft uit verschillende ontwikkelaars. In de praktijk waarborgt de keuze voor Python dat de technologie voor de Resolver vrij beschikbaar en goed gedocumenteerd is, maar Python-ontwikkelaars blijken minder talrijk en duurder zijn. Een keuze door PACKED vzw voor PHP zou, achteraf bekeken, een betere keuze geweest zijn met het oog op een makkelijke toegang tot hosting en ontwikkeldiensten.

Collectieverantwoordelijken leggen zich doorgaans te snel neer bij de technische keuzes die hun IT-diensten maken. Ze beschikken niet over de technische kennis om het belang van het gebruik van open technologieën te bepleiten bij hun IT-diensten. Dat betekent dat ze vaak ook beperkt worden tot het gebruik van technologieën die weinig duurzaam zijn. De overtuiging van PACKED vzw is dat collectieverantwoordelijken een kritischer houding moeten aannemen ten opzichte van de IT-diensten die hen ter beschikking worden gesteld. Zo blijken veel problemen bij de installatie van de Resolver voort te vloeien uit het niet of onvoldoende grondig lezen van de installatie-instructies. De documentatie werd doorheen het project voortdurend door PACKED vzw geactualiseerd en uitgebreid. Er werden door PACKED vzw ook kant-en-klare installatiepakketten voorzien, maar dit leidde zelden of nooit tot een vlottere installatie van de Resolver.

Ook hier speelt het hierboven reeds aangehaalde fundamenteel gebrek aan technische expertise in de instellingen zelf om de uitrol van de Resolver doeltreffend te begeleiden. De communicatieketen PACKED/VKC data manager – instelling – leverancier bemoeilijkt een resultaatgerichte aanpak. Voor een aantal instellingen heeft de VKC data manager deze taak op

zich genomen, maar dit vergt ook een aanzienlijke tijdsinvestering van de VKC, waardoor andere acties in het gedrang komen.

Aanbevelingen:

- Om te vermijden dat de toegang tot duurzame en actuele technologie beperkt wordt door gesloten of verouderde serverconfiguraties is het noodzakelijk om instellingen te sensibiliseren rond de keuze van hostingdiensten. Instellingen moeten de kostprijs van hun basis IT-infrastructuren kunnen afstemmen op de noden van hun collecties.
- Om meer technische expertise in de instellingen onder te brengen dient het belang van een data-uitgever in de instelling verder bepleit te worden. Een belangrijk instrument om in dit pleidooi is het functieprofiel voor de data-uitgever. Dit profiel dient door PACKED veel concreter uitgewerkt te worden in nauwe samenwerking met het Departement CJSM en de betrokken musea. Opleiding, vaardigheden en de plaats van de data-uitgever in de organisatiestructuur dienen verder geconcretiseerd worden. Een andere belangrijke strategie om data-uitgevers te introduceren in de museumwerking is om hun taken vast te leggen in informatiebeleidsplannen die het algemene informatiebeleid van musea beschrijven.

3. Projectmatige benadering persistente identificatie

De ervaring van PACKED vzw in dit project is dat de rol van data-uitgever door de musea beschouwd wordt als een projectresultaat, m.a.w. de aanwezigheid van een data-uitgever is enkel relevant voor de duur van de PACKED-projecten rond persistente identificatie. Dat werd duidelijk in de wijze waarop de toewijzing van de rol van data-uitgever verliep. In twee gevallen werd de collectiemedewerker toegestaan data-uitgever taken uit te voeren in de 'eigen tijd'. In andere gevallen wordt gearzeld om de rol formeel aan iemand toe te wijzen.

Daarnaast wordt er in recent opgestarte ontsluitingsprojecten geen gebruik gemaakt van persistente URI's om collectie-informatie op een duurzame manier ter beschikking te stellen. Data-uitgevers die deelnamen aan het project en de workshops zijn geen ambassadeurs voor het gebruik van persistente URI's in de instellingen. Een aantal van hen geven aan dat dit komt omdat persistente identificatie vooralsnog een abstracte materie blijft, die ze moeilijk uitgelegd krijgen aan collega's binnen hun eigen instelling.

Aanbevelingen:

- Om data-uitgevers verder te sensibiliseren en te ondersteunen bij het implementeren van persistente URI's in hun museumwerking, moet PACKED vzw de training voor data-uitgever die bij aanvang van dit project plaatsvond op regelmatige basis georganiseren. Door dit verhaal op regelmatige tijdstippen opnieuw te brengen wordt de duurzaamheidsstrategie achter het gebruik van 'persistente URI's verder ingeslepen..
- Deze trainingen moeten ook een veel breder publiek van museumwerkers bij persistente URI's betrekken. Om effectief draagvlak te creëren binnen musea voor het gebruik van persistente URI's, moeten niet alleen voor data-uitgevers, maar ook voor communicatiemedewerkers en directies aangepaste trainingen georganiseerd worden.

Deel 2. Project ‘Event-based objectbeschrijvingen’

Doelstelling

Dit deelproject concentreert zich op het concept van *events*: clusters van informatie over wie, waar, wanneer een bepaalde handeling heeft gesteld met betrekking tot een kunstwerk. Het doel van dit project was te onderzoeken hoe de brede waaier van contextuele informatie over kunstwerken geclusterd kan worden in een generieke datastructuur, gecodeerd op een machineleesbare manier en herbruikbaar gemaakt worden voor webapplicaties. Het project onderzoekt de nieuwe mogelijkheden ontstaan wanneer je met die rijke informatie de levensloop van een kunstwerk te visualiseert?

Hiermee verkent dit project een deel van de collectiedata uit de zeven musea dat vooralsnog weinig aandacht kreeg, zoals gegevens over de wijze waarop het kunstwerk in een museum terecht is gekomen, welke weg het werk voordien heeft afgelegd, wie over het werk geschreven heeft, waar het werk ooit getoond is geweest, de conditie en waarde van het werk en wat het werk precies afbeeld, enz. Deze informatie werd door de musea tot nog toe beschouwd als te onvolledig, te vervuild en te complex om in aanmerking te komen voor hergebruik in webapplicaties. Dit project maakt een betrouwbare inschatting van de omvang en kwaliteit van dit reservoir aan contextuele data en test een strategie uit om ze toegankelijk te maken voor ontwikkelaars. Daarbij wordt gebruik gemaakt van een eigenschap van het LIDO-uitwisselformaat, namelijk *events*. Een *event* is een generieke container (wrapper) waarin je uiteenlopende informatie op een, voor computers, gelijkaardige manier kan verpakken door ze te ontleden aan de hand van vijf aspecten:

1. Wie is betrokken in het event?
2. Waar vond het event plaats?
3. Wanneer vond het event plaats?
4. Om welk type event gaat het?
5. Hoe wordt het event omschreven?

Door met behulp van die vijf aspecten uiteenlopende contextuele informatie op een eenvormige manier te coderen, kan de omvang en kwaliteit beter gemeten worden, en kan deze informatie ook door webapplicaties gevisualiseerd worden in tijdlijnen, kaarten en diagrammen. Dit project onderzoekt ook of op deze manier vernieuwende perspectieven op de collectiedata geopend kunnen worden.

Methodologie

Dit deelproject werd uitgevoerd door PACKED-medewerkers Alina Saenko, Pieter Depraetere, Bert Lemmens en jobstudent Nils Van Geele.

Het hele project werd uitgevoerd in de periode maart 2015 - oktober 2015.

Bestaande collectiedata: exports en analyse

Events worden niet als dusdanig gedocumenteerd in een collectiebeheersysteem, maar zitten vaak verscholen in de informatie over bv. de productie, verwerving en bruikleen van het object. Daarom werd bij aanvang van dit project onderzocht welke contextuele data geregistreerd werd in het collectiebeheersysteem.

In de interviews met de collectieverantwoordelijken en de analyse van de data-exports werden de volgende vaststellingen gemaakt:

- Het is tijdrovend om velden te identificeren in het collectiebeheersysteem die relevante data bevatten over *events*. De registratie verschilt naargelang het museum en het systeem dat ze gebruiken. De beschrijvingsregels voor contextuele data zijn ook minder strikt dan voor de identificatiedata, waardoor verschillende beschrijvingspraktijken circuleren. Bijgevolg moest voor elke instelling een volledige data-export gemaakt worden om daarin de relevante velden met wie-waar-wanneer data te identificeren.
- De digitale registratie van contextuele informatie in het collectiebeheersysteem gebeurt slechts fragmentair. Veel contextuele informatie wordt beheerd buiten de collectiebeheersystemen en/of in een niet-gestructureerde vorm. Deze data was voor het project niet bruikbaar om *events* te maken.
- In interviews met de collectieverantwoordelijken werd voor de volgende types events aangegeven dat er gegevens terug te vinden zijn in de data-exports:
 - vervaardiging (o.a. creator en production);
 - opschriften;
 - verwerving;
 - eigendomsgeschiedenis;
 - tentoonstellingen;
 - bruiklenen;
 - bibliografie;
 - onderzoek;
 - waardebepaling;
 - verzekering;
 - rechtenoverdracht;
 - iconografie (content, association);
 - locatiecontrole;
 - verplaatsingen (location.history, movement, despatch);
 - verpakking (package);
 - reproductie;
 - conditiecontrole;
 - volledigheidscntrole;
 - conservatie aanvraag;
 - restauratie (treatment);
 - verlies/beschadiging;
 - afstoting (disposal).

De verdere analyse van de data-exports moet uitwijzen hoeveel reële data voor deze *event* types aanwezig is om betekenisvolle visualisaties te creëren.

Structureren van event-clusters in LIDO xml

Mapping naar LIDO xml

De analyse van de omvang en kwaliteit van de data (m.a.w. of de data-exports voldoende 'feiten' bevatten zoals persoon, datering en plaats om er *events* van te maken) gebeurde door de contextuele data uit de data-exports te *mappen* naar LIDO⁷. LIDO is een uitwisselformaat voor het online toegankelijk maken van collectiedata uit musea. Omdat LIDO hiërarchisch en flexibel is opgebouwd, is het zeer bruikbaar voor beschrijvingen van verschillende types objecten, bv. kunst, architectuur, culturele geschiedenis en technologische geschiedenis. Op het hoogste niveau vormen alle gegevens over een object een *LIDO record*. In elke LIDO record is er ruimte voorzien voor administratieve en beschrijvende gegevens in de vorm van zeven *wrappers*, waaronder ook een *event-wrapper*. Die wrapper maakt het mogelijk om verschillende soorten feiten uit de levensloop van een object generiek te structureren en dus eenvoudig herbruikbaar te maken.

Om de mapping van de oorspronkelijke data-exports naar LIDO uit te voeren werd er in dit project gewerkt met de XML-editor *Oxygen XML*⁸. Per instelling werd een XSL-bestand⁹ ontwikkeld voor de vertaling van het oorspronkelijke dataformaat van het collectiebeheersysteem naar LIDO. In dit project werd vooral met mappings van Adlib-exportbestanden naar LIDO gewerkt. De resultaten van die mappings werden gepubliceerd op Github: <https://github.com/PACKED-vzw/adlib-to-lido/>.

De mapping van de verschillende CSV-bestanden uit de TMS-databank van KMSKA naar LIDO werd niet voltooid in dit project. De export van data uit TMS bleek een tijdrovende klus te zijn. En de CSV-bestanden die uiteindelijk beschikbaar werden gesteld bleken te complex om binnen de resterende tijd om te zetten naar LIDO. In dit project werd duidelijk dat data-exports rechtstreeks uit TMS in de praktijk niet werkbaar zijn als methode om collectiegegevens uitwisselbaar te maken. In voorgaande projecten werd gebruik gemaakt van een alternatieve strategie, namelijk exports uit de CollectionConnection aggregator. KMSKA gebruikt CollectionConnection voor het beschikbaar stellen van data aan o.a. de VKC-online catalogus en interne systemen. De koppeling tussen CollectionConnection en TMS wordt beheerd door de leverancier CiT. CollectionConnection vormde echter geen oplossing voor het verzamelen van *event-data*, omdat er slechts een beperkte hoeveelheid data in de aggregator aanwezig is.

Voor elke partnermuseum werd er een dataset gecreëerd met de oorspronkelijke data in vorm van LIDO-records en deze werden geanalyseerd m.b.v. twee criteria:

- Voor welk aandeel (%) van de objecten zijn er bepaalde type events aanwezig?
- Welke gegevens zijn er voor data event aanwezig?

Bijlage 1 bevat een overzicht van het aantal gecreëerde events per collectie en het aantal 'feiten' die per event gekoppeld werden. Uit die getallen kan men afleiden welke gebeurtenissen er per collectie goed in kaart worden gebracht en welke meer aandacht en meer contextuele data nodig hebben.

⁷ <http://lido-schema.org/>

⁸ <https://www.oxygenxml.com/>

⁹ http://www.w3schools.com/xml/schema_intro.asp

De analyse wees uit dat voor de volgende types events voldoende data aanwezig zijn om een betekenisvolle visualisatie te maken:

event type	aantal kunstwerken / aantal events (collectie)
vervaardiging – production	90-100% (alle instellingen)
verwerving – acquisition	90-100% (Groeningemuseum, MSKGent, S.M.A.K., Mu.ZEE, M HKA, CVG)
eigendoms geschiedenis – provenance	70-115% (KMSKA, MSKGent, S.M.A.K., Mu.ZEE, CVG)
tentoonstelling – exhibition	+200% (S.M.A.K., M HKA)
publicatie – publication	100-170% (KMSKA, MSKGent, S.M.A.K.)
voorstelling – iconography	40% (MSKGent ¹⁰)

Persistente identificatie van events en contextuele data

De analyse van de data-exports wees ook uit dat voor geen enkel type event persistente URI's gedocumenteerd worden. Dit is meestal het geval omdat die voor het betrokken type event niet bestaan (cf. vervaardiging, eigendoms geschiedenis).

Het project ontwikkelde een visie op het identificeren van events. De persistente identificatie van events is een technische noodzaak bij het beschikbaar stellen van collectiedata als linked open data. Maar het kan ook het potentieel voor hergebruik van collectiedata via de datahub gevoelig vergroten.

Er zijn drie mogelijke strategieën om events persistent te identificeren:

- niet identificeren in de data die ter beschikking wordt gesteld, maar op het niveau van de applicatie;
- zelf PIDs voor events maken die beheerd worden door de collectie (via de resolver);
- PIDs ontlenen aan externe bronnen.

De keuze voor een strategie is afhankelijk van het type event en de omvang van de data.

- Als de omvang van de event data beperkt is, wordt deze best eerst aangevuld in het collectiebeheersysteem vooraleer een identificatieschema wordt opgezet. Dit sluit niet uit dat al gemapt wordt naar event-wrappers. Dit vormt meteen ook een stimulans om de beschrijvingsregels voor event data te verfijnen.
- Production, provenance, condition... : Dit zijn events die uniek zijn voor het object in de collectie en dus niet voorkomen bij objecten in andere collecties: de vervaardigingen van

¹⁰ Hoewel minder dan de helft van de voorstellingen werd geïdentificeerd, werd de dataset van MSKGent toch weerhouden omdat het om classificatie met ICONCLASS-codes gaat. Aan deze codes is heel wat metadata verbonden en het leek ons interessant om te onderzoeken of we hieruit interessante visualisaties konden destilleren.

een object is uniek voor een kunstwerk in jou collectie. In andere collecties bevinden zich geen kunstwerken die deel uitmaken van hetzelfde event. Voor deze events kan de collectie dus optreden als autoriteit en valt de strategie om zelf persistente URI's te creëren te verantwoorden.

- Exhibition en publication: Publicaties en tentoonstellingen zijn events waarin meerdere kunstwerken in verschillende musea betrokken zijn. Voor deze events wordt voorgesteld om persistente URI's te gebruiken die beheerd worden door een externe autoriteit.
 - exhibition: op dit moment bestaat er nog geen centrale standaardterminologie die alle tentoonstellingen in kaart brengt, maar de databanken die beheerd worden door Cultuurnet of ODIS hebben het potentieel om uit te groeien tot een autoriteit. Hergebruik van persistente URI's uit deze databanken wordt aanbevolen, eerder dan zelf tentoonstellingen persistent te identificeren.
 - publication: Voor publicaties zijn zowel op regionaal als internationaal niveau verschillende externe autoriteiten actief. Er is ook een groeiende trend om deze autoriteiten onderling te linken. Voor deze events wordt voorgesteld om aan te sluiten bij de autoriteit die de grootste dekking van de publicaties in je collectiedata garandeert. Van daaruit kan op termijn gelinkt worden naar andere autoriteiten. Enkele voorbeelden: Worldcat en ISBN.

Voor contextuele informatie zoals personen en plaatsen gelinkt aan een gebeurtenis werd er in dit project gebruik gemaakt van externe standaardterminologieën.

- Voor personen hebben we gebruik gemaakt van de normalisering die in het voorgaande traject plaats heeft gevonden en waar we persistente URI's van VIAF, RKDartists&, Wikidata en ODIS hebben gebruikt
- Voor plaatsnamen hebben we in dit project de API van Geonames uitgetest¹¹, waarbij we Geonames ID's en coördinaten van plaatsen hebben opgenomen in onze dataset. Die data hebben we kunnen gebruiken om kaarten aan te maken bij de visualisaties (zie deel 'Visualisaties').

Visualisaties

Het doel van de visualisaties is demonstreren hoe je contextuele data kan gebruiken om nieuwe aspecten van je collectie zichtbaar te maken op het web. Events bieden nieuwe mogelijkheden om, vaak visueler en intuïtiever dan de klassieke resultaatlijsten, je collectie doorzoekbaar te maken. Deze demonstratie moet uitwijzen of het gebruik van LIDO *events* de ontwikkeling van zulke visualisaties eenvoudiger maakt.

De visualisatiesoftware die gebruikt werd om deze visualisaties te creëren bestaat uit twee componenten. Enerzijds is er de component die verantwoordelijk is voor het verwerken en masseren van de oorspronkelijke data-export. Dit onderdeel is geschreven in Python en maakt gebruik van voornamelijk de XML en CSV-bibliotheken. De andere component zet de aangepaste data om in grafische visualisaties. De code hiervoor is grotendeels in Javascript (sommige onderdelen ook in Python), met gebruik van de nvd3 en d3-bibliotheken voor visualisatie. De broncode voor beide componenten is te vinden op <https://github.com/PACKED-vzw/lido-event-visualisaties>.

¹¹ voor de gebruikte GRELfuncties in Open Refine zie http://www.projectceest.be/images/3/38/20150430_Handout_Open_Refine_workshop.pdf

Voor het maken van de visualisaties op basis van de vijf types events, werd de relevante collectiedata gemapt naar LIDO en verrijkt met informatie uit externe standaardterminologieën. De collectiedata werd door de acht instellingen enkel ter beschikking gesteld voor de duur van het project. Bijgevolg is deze data dus niet opgenomen in dit eindrapport. Een voorbeeldrecord van een LIDO record zoals gebruikt in de demo is te vinden op Github¹².

De visualisaties werden ontwikkeld op basis van vijf use cases:

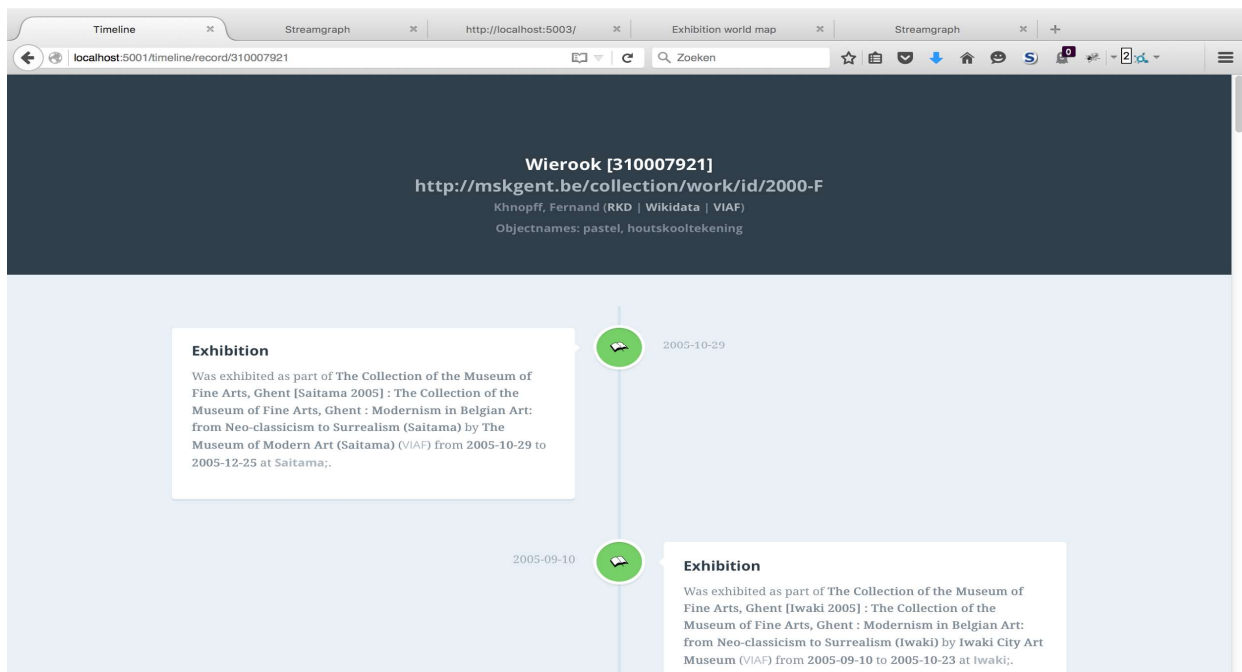
1. Levensloop

Vraag: Wat is er doorheen de tijd met het kunstwerk gebeurd en waar is het kunstwerk te zien geweest?

Uit de collecties van vijf instellingen (Groeningemuseum, S.M.A.K., MSKGent, KMSKA en MuZEE) werden records gekozen met een heel lange levensloop en veel geregistreerde events om uitgebreide tijdslijnen te kunnen bouwen. De visualisatie heeft een vorm gekregen van een 'facebook'-tijdslijn waar alle gebeurtenissen elkaar chronologisch opvolgen van oud naar nieuw, en waar er ook een kaart aanwezig is die de gebeurtenissen geografisch plaatst. De gebruiker 'scrolt' doorheen de levensloop van een kunstwerk en ziet het verband tussen verschillende gebeurtenissen in het leven van het kunstwerk.

Deze tijdlijn wordt mogelijk gemaakt door verschillende types gebeurtenissen op een zelfde manier te coderen in een LIDO-event, waardoor de applicatie, op basis van de 'wanneer'-informatie in elk event, verschillende types events chronologisch kan rangschikken.

Niet alle gebeurtenissen hadden echter een datering of een plaatsnaam, waardoor ze niet op een tijdlijn of kaart geplaatst konden worden. Dit komt deels door onvolledige registratie, maar in een aantal gevallen ook omdat deze gegevens traditioneel niet gedocumenteerd worden. Bijvoorbeeld de plaats waar een kunstwerk is gemaakt, wordt vrijwel nooit gedocumenteerd, hoewel dit zeer bruikbaar kan zijn om de herkomst van kunstwerken te visualiseren.



¹² https://github.com/PACKED-vzw/adlib-to-lido/blob/master/MSKG_LIDO1testrecords.xml

Timeline Streamgraph http://localhost:5003/ Exhibition world map Streamgraph

localhost:5001/timeline/record/310007921 Zoeken

Wales en de Eerste Wereldoorlog (Heino-Wijhe) by Hannema-de Stuers Fundatie (VIAF) from 2002-03-30 to 2002-06-02 at Heino/Wijhe;

2002-01-12

Exhibition

Was exhibited as part of Kunst in ballingschap [Gent 2002] Kunst in ballingschap: Vlaanderen, Wales en de Eerste Wereldoorlog by Museum voor Schone Kunsten Gent (Wikidata | VIAF) from 2002-01-12 to 2002-03-17 at SMAK (Stedelijk Museum voor Actuele Kunst) Gent;Gent.

2000-10-26

Acquisition

Acquired by Museum voor Schone Kunsten Gent from Derom, Patrick (VIAF)

2000-10-26

Provenance

Changed hands from Derom, Patrick (VIAF) to Museum voor Schone Kunsten Gent

Timeline Streamgraph http://localhost:5003/ Exhibition world map Streamgraph

localhost:5001/timeline/record/310007921 Zoeken

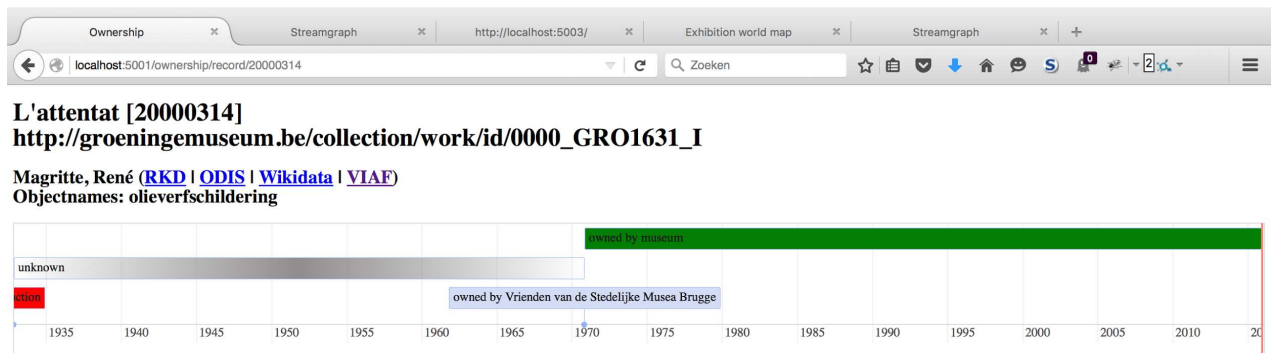
Publication

Appeared in the publication Museum voor Schone Kunsten Gent : Catalogus schilderkunst ; Deel I: 14de - 18de eeuw ; Deel II: 19de - 20ste eeuw by Hoozee, Robert (Wikidata | VIAF) .

Locations

2. Gespecialiseerde tijdlijn (eigendoms geschiedenis)

Vraag: Hoe is het kunstwerk doorheen de tijd van eigenaar verwisseld?



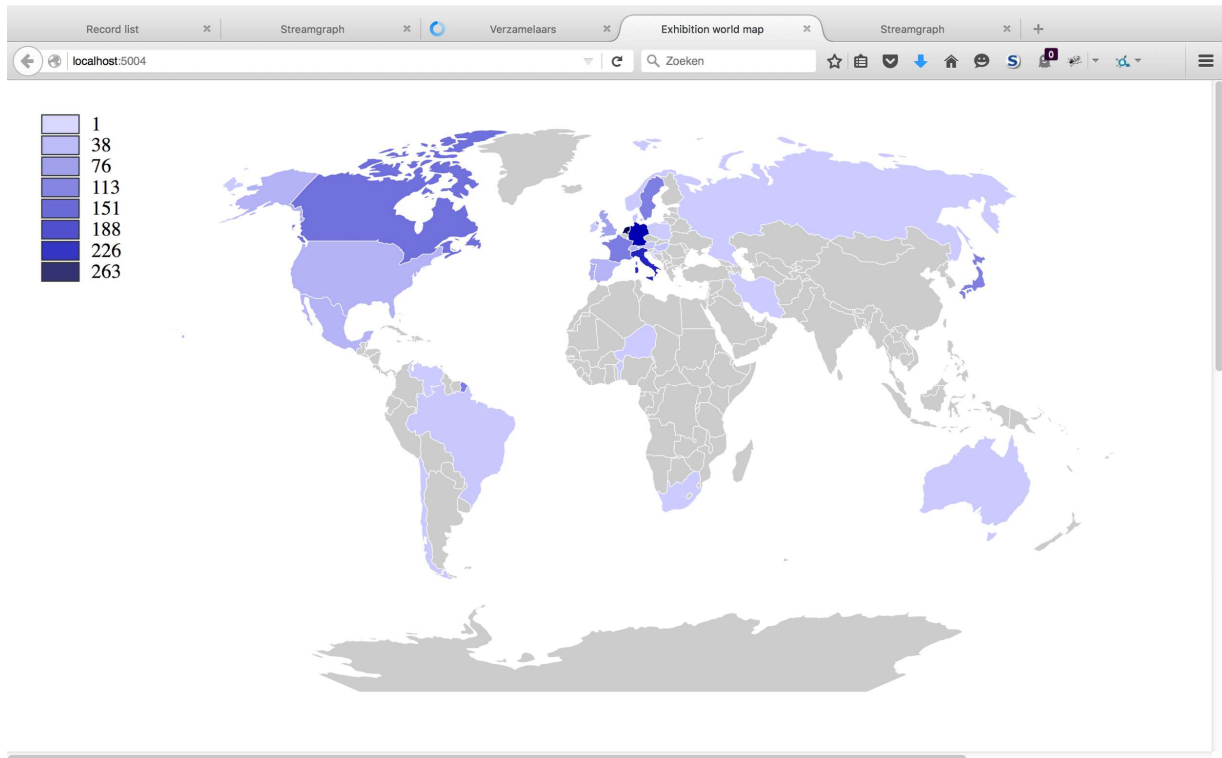
Deze visualisatie is een tweede type tijdlijn die zichtbaar maakt over welke periodes in het leven van een kunstwerk je wel of niet iets weet. Over het werk *L'attentat* in het Groeningemuseum weten we bijvoorbeeld weinig over de eerste helft van zijn geschiedenis, tot het moment waarop het in de collectie van het Groeningemuseum terecht komt. Dit soort informatie – zichtbaar maken wat je niet weet - kan het vertrekpunt zijn om lacunes in de kennis van het museum te vullen. Bij kunstwerken waar de eigendoms geschiedenis wel goed bekend en gedocumenteerd is, vormt deze visualisatie een nuttig didactisch instrument om verschillende fases in het leven van een kunstwerk te onderscheiden.

Voor deze visualisatie werd gebruik gemaakt van één type events, nl. provenance events. Ze visualiseert een heel traditioneel gegeven uit het kunsthistorisch onderzoek, m.n. de *pedigree*. Meestal is die enkel beschreven in lijsten en teksten, maar op deze manier wordt het visueel toegankelijk gemaakt voor niet-experten.

De registratie van eigendoms geschiedenis in collectiebeheersystemen vertoont grote hiaten. Gegevens over de verwerving van een kunstwerk door het museum zelf is vrijwel volledig, hoewel die vaak dubbel wordt geregistreerd in 'verwerving' en in 'eigendoms geschiedenis'. De eigendoms geschiedenis vóór verwerving is slechts sporadisch gedocumenteerd. (zie tabel 3 en 4 in bijlage 1).

3. Kaarten (tentoonstellingen)

Vraag: In welke landen hangt vandaag kunst uit de betrokken musea in een tentoonstelling en waar in de wereld zijn kunstwerken in het verleden getoond?



Deze visualisatie maakt gebruik van de geregistreerde informatie over bruiklenen en tentoonstellingen en toont de landen waar een kunstwerk uit een collectie werd tentoongesteld. De visualisatie maakt tentoongestelde werken zichtbaar per jaar. Hoe meer kunstwerken in één land tentoongesteld worden, hoe donker het land kleurt op de kaart. De visualisatie maakt zo zichtbaar hoe kunstwerken uit museumcollecties in de wereld over de hele wereld te zien zijn.

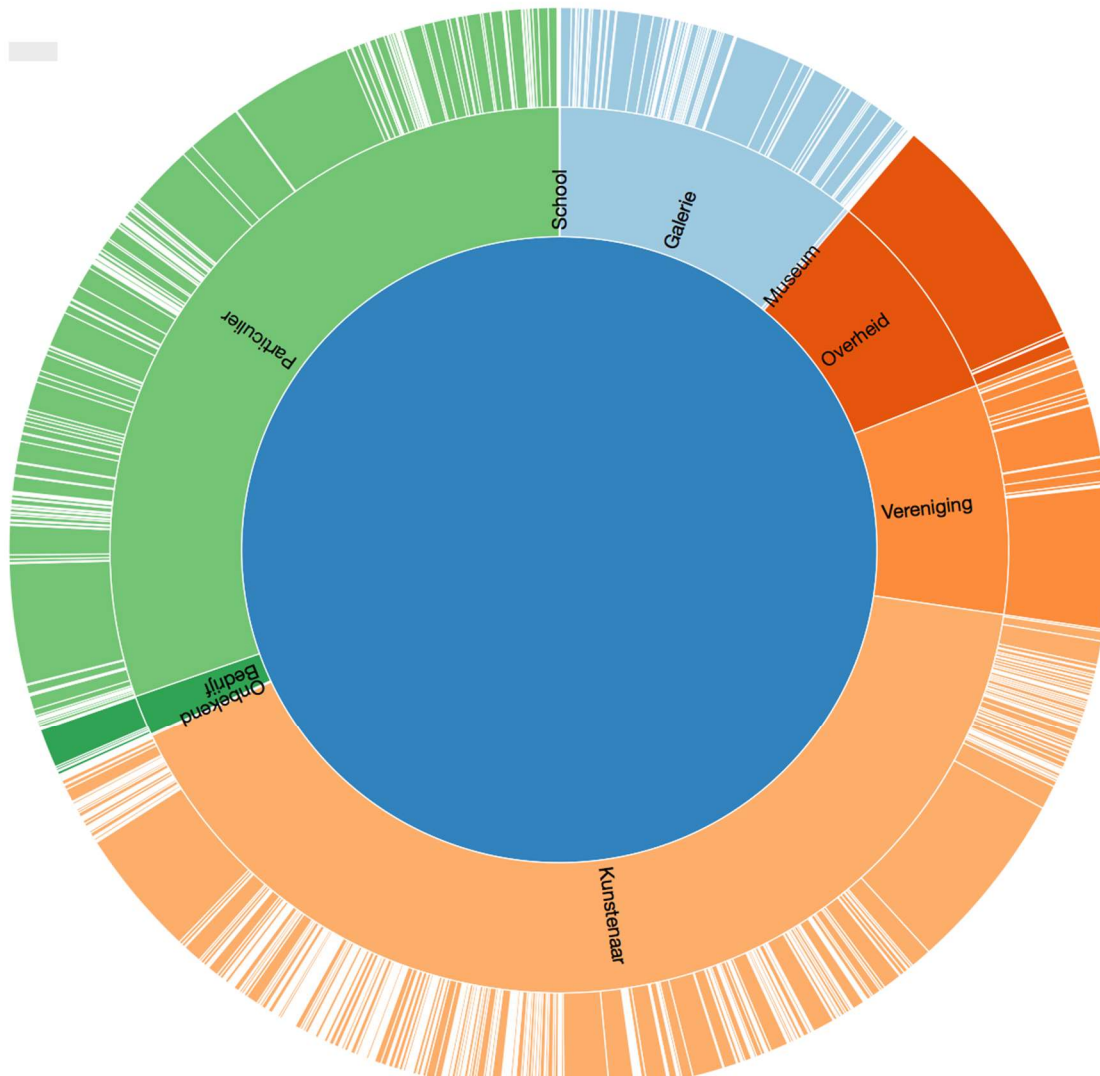
Bij de analyse van de collectiedata bleek ook dat de relatie tussen informatie over bruiklenen en over tentoonstellingen complex is: Soms ikomt een vermelding van een bruikleen overeen met de vermelding van een tentoonstelling, wat resulteert in een verdubbeling van hetzelfde event. Soms is er geen verband te vinden tussen een bruikleen en tentoonstelling. Meestal zijn de gegevens dan complementair. In een aantal gevallen zorgen ze voor contradicties. Voor deze visualisatie werden de gegevens over bruiklenen en tentoonstellingen handmatig geschoond en genormaliseerd tot 'exhibition event' (zie tabel 5 in bijlage 1). Om ook in de praktijk de tentoonstellingsgeschiedenis van collecties te visualiseren, zijn er veel striktere afspraken nodig voor de registratie van bruiklenen en tentoonstellingen, waardoor het schonen van data zoveel mogelijk beperkt wordt.

4. Herkomst van de collecties

Vraag: Welke verzamelaars vormen de belangrijkste bron voor een collectie?

Deze visualisatie heeft twee versies:

4.1 Per museum



Deze visualisatie maakt gebruik van 'acquisition events', met name de events die de verwerving van een kunstwerk door één specifiek museum identificeren (zie tabel 3 in bijlage 1). De visualisatie toont de verschillende bronnen waaruit kunstwerken verworven werden (bv particulieren, overheden, galeries, de kunstenaar zelf, etc) en hoe die bronnen zich in omvang tot elkaar verhouden.

Voor deze visualisatie werd de collectiedata uit één museum verrijkt met verschillende types bronnen, zoals bv. *particulier*, *galerie*, *kunstenaar*, *vereniging*, *overheid*, *bedrijf* en *onbekend*. Men kan ook verder inzoomen op een type om te zien over welke personen of instellingen het gaat.

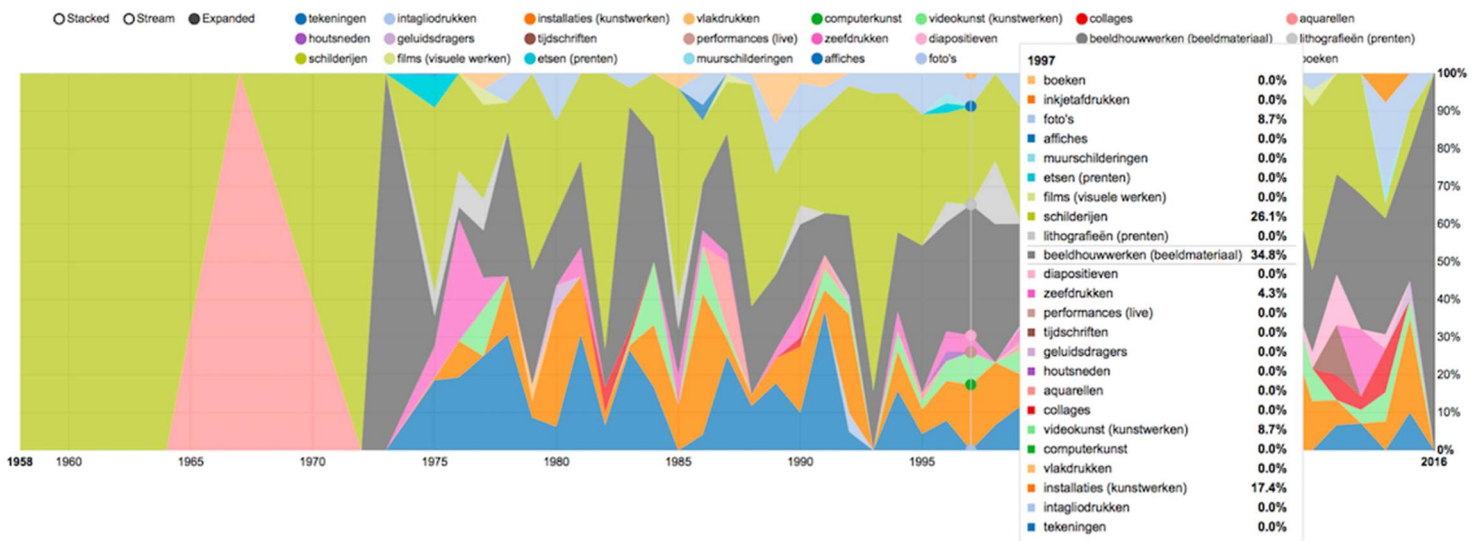
4.2. Voor vijf collecties samen



Voor deze visualisatie werden 'acquisition events' uit vijf musea samengebracht om te tonen hoe data van alle musea nieuwe inzichten oplevert over Vlaamse kunstcollecties als geheel.

5. Streamgraph (objectnaam/iconografie):

Vraag: Hoe belangrijk is een objecttype of voorstellingstype voor een collectie doorheen de tijd?



Voor deze visualisatie werd gebruik gemaakt van andere contextuele data per kunstwerk die niet rond een 'event' werd geclusterd, maar ook in een LIDO-record werd opgenomen: objectnaam, iconografie en vervaardigingsdatum. Op die manier kon deze visualisatie tonen dat het structureren van eender welke data over een object in vorm van LIDO nieuwe mogelijkheden biedt tot hergebruik. Deze streamgraph combineert namelijk 'feiten' over een kunstwerk die normaal gezien zelden naast elkaar worden geanalyseerd: trefwoord en vervaardigingsdatum. De visualisatie geeft een beeld van de verhouding tussen types van kunstwerken van een collectie doorheen de tijd. Voor onderzoekers en het breder publiek is dit een interessante inzicht in een collectie.

De combinatie van vervaardigingsdatum en iconografie werd ook uitgetest in het kader van dit project om de vaak voorkomende thema's in kunstwerken te kunnen visualiseren, maar er was onvoldoende data in de oorspronkelijke data-exports om te kunnen structureren in LIDO en uitspraken te kunnen maken over de volledige collectie.

Kosten-baten analyse

Om de haalbaarheid van de ontwikkeling van visualisaties op basis van LIDO-data in te schatten, werd een kosten-baten analyse gemaakt van het ontwikkelproces dat in het project doorlopen werd.

Tijdsinvestering

Eén van de doelstellingen van dit project was om te onderzoeken of het coderen van contextuele gegevens in events de ontwikkeling van vernieuwende visualisaties van collecties eenvoudiger maakt. Een belangrijk aspect hiervan is de omvang van de tijd en middelen die nodig zijn om collectiedata om te zetten naar een webapplicatie.

Voor dit project werd het aantal mandagen geregistreerd om:

- Een **analyse** te maken van de vijf visualisaties die het project heeft gemaakt. Dit omvat het uitschrijven van use cases en het oplijsten van de functionele en technische vereisten voor elke visualisatie.
- De **bewerking** van de data. Dit omvat het selecteren van de nodige metadata-elementen en het maken van XSL-bestanden om de collectiedata om te zetten naar LIDO.
- De **ontwikkeling** van de webapplicaties. Dit omvat het schrijven van een *parser*¹³ die LIDO omzet naar een formaat dat door de webapplicatie geanalyseerd kan worden, de nodige open source softwarecomponenten bij elkaar zoeken om de visualisatie te maken en te testen.

De visualisaties werden gemaakt door twee projectmedewerkers: één ontwikkelaar en één data-beheerder.

Hieronder volgt een overzicht van het aantal mandagen die voor PACKED vzw nodig waren om voor de vijf use cases visualisaties te maken.

- Analyse: 2 personen/12 dagen = 24 mandagen
 - brainstorm/overleg met collectiebeheerders
 - ideeën uitschrijven voor de ontwikkelaar
- Data bewerking: 2 personen/32 dagen = 64 mandagen
 - datasets selecteren
 - data naar LIDO mappen
 - LIDO-data verder verrijken
- Ontwikkeling: 2 personen/25 dagen = 50 mandagen
 - LIDO-parser schrijven op maat van de visualisatie
 - visualisatiesoftware zoeken en testen
 - de visualisatie maken

De broncode die in dit project geschreven werd is vrij beschikbaar voor hergebruik, waardoor een gelijkaardig project gebruik kan maken van de resultaten van dit onderzoek.

¹³ Een parser is een computerprogramma dat de grammaticale structuur van een importbestand ontleedt. Daarbij controleert de parser of het importbestand correct is opgebouwd en vervolgens zet het de data om in een vorm waardoor het door een applicatie geanalyseerd en verwerkt kan worden.

Sterkte-zwakte analyse

Aan het eind van dit project werd een zelf-evaluatie gemaakt waarbij de projectmedewerkers een analyse van de sterktes en zwaktes van het project vanuit het standpunt van de musea.

Voor deze zelfevaluatie werd vertrokken vanuit het volgende standpunt: Als een museum in een nieuw project zijn collectiedata (zowel de contextuele als de identificatiedata) zou omzetten naar LIDO en op basis van deze dataset zijn collectie visueel zichtbaar en doorzoekbaar zou maken:

- Welke nuttige effecten heeft dat voor de collectie?
- Welke schadelijke effecten heeft dat voor de collectie?
- Welke kansen biedt dit aan de collectie?
- Welke bedreigingen vormt dit voor de collectie?

	nuttig	schadelijk
interne factoren	<p>sterkte:</p> <ul style="list-style-type: none"> • Het project maakt het mogelijk om één visualisatie te maken voor verschillende datasets uit verschillende musea. • Het project maakt het mogelijk om de collectie op een visuelere en intuïtievare manier doorzoekbaar te maken. • Het project maakt de geschiedenis van de collectie beter zichtbaar. • Het project maakt de context en de keuzes die gemaakt werden bij de vorming van de collectie beter zichtbaar. 	<p>zwakte:</p> <ul style="list-style-type: none"> • Het project maakt de hiaten in de collectieregistratie intern zichtbaar. Als er weinig gegevens over bijvoorbeeld tentoonstellingen beschikbaar zijn, zal dit ook duidelijk worden in de beperkte omvang van de visualisatie. • Het project vereist een investering in tijd en middelen om de technische expertise in huis te halen die nodig is om de collectiedata te mappen naar LIDO en ze beschikbaar te stellen aan ontwikkelaars.
externe factoren	<p>kans:</p> <ul style="list-style-type: none"> • Het project biedt de mogelijkheid om inhaaloperaties in de registratie beter te plannen, zowel in tijd, focus als omvang. • Het project biedt de mogelijkheid om de beschrijvingsregels voor contextuele data fijner te stellen, in overleg met andere musea. 	<p>bedreiging:</p> <ul style="list-style-type: none"> • Het project maakt de hiaten in de collectieregistratie publiek zichtbaar. • Het project creëert de verwachting dat het museum de contextuele data op lange termijn beschikbaar houdt en verder verbetert. • Ontwikkelaars willen LIDO niet gebruiken omdat het formaat te complex is en te tijdrovend om het te parsen.

Conclusies en aanbevelingen

In dit deelproject hebben we de contextuele data uit zeven collecties gemapt naar LIDO-events, verrijkt en vervolgens gebruikt om vijf use cases te vertalen in visualisaties van de collectiedata. Het doel van dit project was te testen of het LIDO-formaat, en in het bijzonder de event-wrapper, een bruikbaar instrument is om uiteenlopende contextuele data in een generiek dataformaat te coderen dat eenvoudig ingezet kan worden om nieuwe inhoudelijke perspectieven op de collecties te openen.

Op basis van dit project werden de volgende algemene vaststellingen gedaan met betrekking tot het gebruik van events voor het coderen van contextuele data:

1. Volledigheid van contextuele data

Voor een vijftal type events (vervaardiging, verwerving, eigendomsgeschiedenis, tentoonstelling en publicatie) is voldoende data aanwezig in de collectiebeheersystemen om op een betekenisvolle en vernieuwende manier collecties te visualiseren. Een groot deel van deze contextuele data bevindt zich spijtig genoeg nog buiten de systemen en vaak in een niet-gestructureerde vorm. Een groot deel van de informatie over de eigendomsgeschiedenis bevindt zich in analoge collectiedossiers. Een groot deel van de informatie over bruiklenen en tentoonstellingen bevindt zich in mailboxen en rekenbladen. De registratie van contextuele gegevens staat laag op de prioriteitenlijst van registratoren (hun aandacht gaat vooralsnog voornamelijk naar minimale en basisregistratie), waardoor een groot deel van deze informatie niet toegankelijk is voor hergebruik in webapplicaties. Een groot deel van de kennis die aanwezig is in musea blijft daardoor onzichtbaar voor de buitenwereld.

Aanbeveling: ‘

- De business case voor het mappen van contextuele data naar LIDO kan nog aanzienlijk versterkt worden als ook deze ‘analoge’ data digitaal te registreren. Hoe vollediger de data in de collectiebeheersystemen, hoe groter de impact van de visualisatie op de eindgebruiker. Op basis van de gegevens uit dit project kunnen instellingen deelprojecten opzetten om voor specifieke event types de lacunes in collectiedata weg te werken. De visualisaties van deze event types vormen een nuttig instrument om het resultaat van zo een deelproject meteen zichtbaar te maken.

2. Beschrijvingsregels voor contextuele data

De mapping van Adlib xml naar LIDO xml is gelukt, maar was tijdrovend door de subtiele verschillen in de structuur van de oorspronkelijke Adlib-exportbestanden en door de verschillende registratiepraktijk die in de musea bestaat voor contextuele data. Wanneer musea bepaalde gegevens systematisch en volledig documenteren, blijkt de mapping van deze data nog steeds complex, omdat dezelfde gegevens in verschillende velden zitten of omdat eenzelfde veld op verschillende manieren wordt ingevuld.

Aanbeveling:

- Door de beschrijvingsregels voor contextuele data in het MovE-invalboek aan te scherpen kan de kwaliteit van de data gevoelig verbeteren en wordt de mapping ook eenvoudiger. Op basis van de gegevens uit dit project kan PACKED vzw voor elk afzonderlijk event type een voorstel gemaakt worden om de beschrijvingsregels in musea aan te scherpen waardoor wie-wat-waar-wanneer gegevens eenvoudig gemapt kunnen worden naar LIDO

of andere formaten. Deze oefening zal gemaakt worden binnen het lopende traject in samenwerking met Erfgoedinzicht om het MovE-invulboek over te brengen naar CEST.

3. Mapping van contextuele data

Voor de mapping van data exports werd gebruik gemaakt van XSL en XSLT. Het gebruik van XSL werd als tijdrovend en complex ervaren. Het omzetten van XML-bestanden met behulp van XSL-bestanden vraagt veel tijd en rekencapaciteit.

Aanbeveling:

- Voor de uitwerking van een mappingprocedure in de praktijk, wordt aanbevolen om alternatieve mappingmethodes te onderzoeken die minder tijdrovend zijn en minder belastend zijn voor de IT-infrastructuur. Met name voor het uitwerken van een aggregatie-strategie voor de datahub moeten alternatieve technologieën getest worden. Deze oefening zal gemaakt worden binnen het lopen Datahub project dat VKC uitvoert in samenwerking met PACKED.

4. Ontwikkeling van visualisaties

De tools die gebruikt werden voor het visualiseren van de data (kaarten, tijdlijnen en diagrammen) blijken in de praktijk allemaal hun eigen dataformaat te gebruiken. Bijgevolg was in alle gevallen, naast de mapping van collectiedata naar LIDO, nog een bijkomende mapping nodig van LIDO naar het datamodel van de specifieke visualisatie. (bv. van Adlib/xml naar LIDO/xml naar een specifiek CSV-formaat).

De mapping van collectiedata naar LIDO op zich maakt de ontwikkeling van visualisaties niet eenvoudiger, maar zorgt wel dat data uit verschillende systemen samengebracht wordt in één dataset. Van daaruit hoeft slechts één mapping gemaakt te worden naar een applicatie-specifiek formaat.

Het mappen van LIDO naar een productspecifiek formaat vereist een parser voor het LIDO-formaat. Een dergelijke parser is momenteel niet beschikbaar voor het LIDO-formaat. Voor dit project werd een specifieke parser ontwikkelt op maat van de data die nodig was voor de visualisaties. Deze ontwikkeling heeft een aanzienlijk deel van de ontwikkeltijd gevraagd en vormt een ernstige hindernis voor de duurzaamheid van LIDO als uitwisselformaat tussen verschillende applicaties.

Aanbeveling:

- Om de mapping van LIDO naar een applicatie-specifiek formaat zo eenvoudig mogelijk te maken is de ontwikkeling van een standaard LIDO-parser noodzakelijk die het volledige gamma aan data dat gedocumenteerd kan worden in LIDO toegankelijk maakt. Door de gelaagdheid van het LIDO-formaat is dit echter een grote technische uitdaging. Maar de beschikbaarheid van zo een parser is een doorslaggevend argument voor de adoptie van LIDO als uitwisselingsformaat voor museumdata.

Deel 3. Linked Open Data publicatie via Wikidata

Doelstelling

Dit project bouwt verder op de resultaten van het project 'Hoe wordt ik data-uitgever?' (2014-2015) en maakt de identificatiegegevens van de kunstwerken uit het PID-voortraject (en die van enkele andere partners)¹⁴ beschikbaar als Linked Open Data (LOD) met behulp van het Wikidata-platform. Q-nummers uit Wikidata, i.e. persistente URI's voor items die beschreven worden door de Wikipedia-gemeenschap, werden in het eerste project, 'Persistente Identificatie' (2013-2014), gebruikt om kunstenaars die een pagina hebben in Wikipedia persistent te identificeren. In dit deelproject stroomt een deel van de collectiedata terug naar Wikidata, met name de persistente URI's van de kunstwerken van deze kunstenaars die in de zeven bovenvermelde musea hangen. Op die manier wordt de informatie in Wikidata verder verrijkt, zijn alle kunstwerken in de zeven musea vindbaar in Wikidata en hebben de musea een krachtig instrument om gebruikers van Wikipedia te betrekken bij de verder ontwikkeling van hun collectiedata..

Dit project onderzoekt of de infrastructuur die de Wikidata gratis ter beschikking stelt, door musea gebruikt kan worden om:

- hun opendatabeleid in de praktijk te brengen;
- hun collectiedata als Linked Open Data te publiceren zonder daarvoor zware IT-investeringen te doen.

Methodologie

Dit deelproject werd uitgevoerd door PACKED-medewerkers Bert Lemmens, Alina Saenko en Barbara Dierickx. Wegens de specificiteit van de Wiki-gemeenschap (community, platformen, technische know-how) werkten we in onderaanneming samen met [Sandra Fauconnier](#)¹⁵, en op vrijwillige basis met [Maarten Dammers](#)¹⁶. Beiden brachten de nodige ervaring en info over Wiki-processen in het project binnen.

Het hele project werd uitgevoerd in de periode oktober 2015 - januari 2016.

Actie 1. White paper open data beheer in Wikidata

In samenwerking met projectmedewerker Sandra Fauconnier werd een white paper opgemaakt en tijdens de eerste stuurgroep voorgesteld aan de projectpartners. Deze paper werd ook

¹⁴ In de marge van het project werden een aantal musea gecontacteerd om aan te sluiten bij de zeven musea uit het PID-voortraject. Een belangrijke voorwaarde was wel dat ze in korte termijn het PID-voortraject konden doorlopen. Museum M ging uiteindelijk in op de uitnodiging en leverde een importbestand met identificatiedata aan voor upload naar Wikidata.

¹⁵ <https://nl.wikipedia.org/wiki/Gebruiker:Spinster>

¹⁶ https://commons.wikimedia.org/wiki/User:Multichill#About_me

vertaald naar het Engels ter verspreiding in de bredere Wiki-gemeenschap ([beschikbaar online¹⁷](#)). De paper bestaat uit drie hoofdstukken:

- introductie tot de Wikimedia-beweging en de projecten Wikipedia en Wikidata;
- omschrijving van de business case om kunstwerken te documenteren in Wikidata voor respectievelijk de musea, de Wikimedia-gemeenschap en de samenleving;
- een crosswalk die weergeeft hoe collectiedata in het datamodel van Wikidata wordt gemapt.

De conclusie van deze white paper bevat een sterkte-zwakte analyse van het gebruik van Wikidata voor het beschikbaar stellen van collectiedata op het web. Het document diende als introductiemateriaal voor de projectpartners in dit deeltraject.

Op basis van deze paper werd ook een inleidende screencast gemaakt ([beschikbaar online¹⁸](#)), die kort uitlegde wat er in het project zou gebeuren en hoe dit alles in zijn werk zou gaan.

Actie 2. Open data publicatie van de CSV dataset

De datasets die als resultaat van het project Persistente Identificatie gecreëerd werden, konden gebruikt worden als importset voor dit deeltraject. Het ging om informatie m.b.t. de eigenlijke objecten uit de museumcollectie, de persistente links die daaromtrent gecreeërd werden, de vervaardigers van de objecten en de objecttypes.

Het doel was deze (zeer feitelijke) data op Wikidata te publiceren als open data. Dit betekende dat de deelnemers hun akkoord moesten geven over het vrijgeven van de informatie onder een CC0-licentie. Om dit te bewerkstelligen werd een samenwerkings- en datagebruiksovereenkomst afgesloten met elk van de partners. Deze werden uiteindelijk ondertekend door de volgende instellingen: CVG, VKC, KMSKA, S.M.A.K., Groeningemuseum, MSKGent, MuZEE en M Leuven. Ondanks de korte duurtijd van het deelproject wisten de vermelde partners de overeenkomsten te ondertekenen.

MHKA, Middelheimmuseum en LUKAS participeerden uiteindelijk niet in dit deelproject en tekenden geen samenwerkings- en datagebruiksovereenkomst. Elk van deze drie instellingen heeft uitgebreide toelichting ontvangen bij beide overeenkomst, maar ondanks herhaalde vragen tot feedback en uitnodigingen voor workshops en stuurgroepen heeft geen van deze drie instellingen gereageerd op de voorstellen tot overeenkomst.

Actie 3. Workshop en stuurgroep I

De eerste Stuurgroep van dit deelproject werd gehouden op 28 oktober 2015. Naast de projectpartners en PACKED-medewerkers was ook Sandra Fauconnier aanwezig. De volgende zaken werden besproken:

- voorstelling projectplanning en stappen;
- voorbeeld van de benodigde elementen in de aan te leveren dataset(s);
- toelichting bij gebruik van CC0-licentie en wat de publicatie als 'open data' zou impliceren

¹⁷

https://www.wikidata.org/wiki/Wikidata:Flemish_art_collections,_Wikidata_and_Linked_Open_Data/Whitepaper

¹⁸ <https://www.youtube.com/playlist?list=PL2XsjAg21n-jTNQI6X3UbZO2f8II2xn5r>

- voorstelling van de whitepaper, en mogelijkheid tot feedback;
- illustratie van de mogelijkheden voor hergebruik op basis van data die tot het Wikidata-platform behoort.

De workshop markeerde ook meteen de start van de uploadprocedure van de aangeleverde data in het Wikidata-platform.

Actie 4. Upload CSV-dataset in Wikidata

PACKED vzw contacteerde daartoe Maarten Dammers (Wikidata-vrijwilliger) om afspraken te maken rond de praktische uploadprocedure. Sandra Fauconnier coördineerde het uploadwerk en voerde manueel enkele correcties/verrijkingen door.

De aangeleverde data werd voorafgaand aan de upload nog op een paar manieren bewerkt. Zo werd er onder andere gekeken

- of alle werken in de lijst wel voldoende 'notable' waren: items zoals (alledaagse) gebruiksvoorwerpen horen niet thuis op Wikidata; religieuze gebruiksvoorwerpen werden om die reden bv. weerhouden uit de up te loaden set van M Leuven;
- of de aanwezige 'creators' in de lijst nog beter verrijkt en gelinkt konden worden, en of zij al een eigen pagina hadden binnen Wikidata;
- of de aanwezige 'objecttypes' konden gemapt worden naar de types die binnen Wikidata gangbaar zijn.

De eigenlijke upload gebeurde in grote mate door het programmeren van een bot; een programma dat automatische nieuwe Wikidata-records aanmaakt op basis van de aangeleverde data, of reeds bestaande records aanvult met deze data. Op die manier kon meteen snel en in grote getale informatie worden toegevoegd aan het platform.

Actie 5. RDF-export uit Wikidata

De gepubliceerde data kan op drie verschillende manieren uit het platform geëxporteerd worden. Deze manier staan beschreven in het Handboek (zie volgend actiepunt). De projectpartners hebben met deze mogelijkheid een belangrijke troef in handen om nu semantische applicaties te gaan maken. Het eigenlijke maken van de export wordt overgelaten aan de projectpartners aangezien het aan hen is om echt iets met de data te gaan 'doen' en dus ook best kunnen bepalen welke vorm de export aanneemt.

Actie 6. Handboek open data beheer in Wikidata

PACKED vzw realiseerde i.s.m. Sandra Fauconnier een handboek¹⁹ dat de projectpartners in staat stelt de opgeladen data in Wikidata zelf te beheren. Onder beheer wordt zowel het aanpassen van bestaande records, als het toevoegen van nieuwe records verstaan. Het document is in eerste instantie relevant voor collectiebeheerders, maar is ook interessant voor vrijwilligers die informatie over kunst toevoegen en aanpassen op Wikidata.

Het handboek bestaat uit vier hoofdstukken. Eerst wordt aangegeven hoe je bestaande data op Wikidata kan beheren (data opzoeken, aanpassen, referenties meegeven). In een volgend hoofdstuk leer je hoe je nieuwe items op Wikidata kan plaatsen; zowel manueel als met de hulp

¹⁹ https://drive.google.com/open?id=1e7g_MSgC9tGr8BYHZddh9PtC87AV9qcTNvDQGbdBu8

van een bot (vrijwilliger), indien het bijvoorbeeld om grote hoeveelheden informatie gaat. Nadien wordt bekeken hoe de data kan gebruikt worden en hoe je wijzigingen en andere gebruikers van Wikidata in de gaten kan houden. Tot slot lees je meer over de beschikbare export(s) van Wikidata-gegevens (zie vorig actiepunt), en tonen we voorbeelden van data-toepassingen.

Het handboek bevat ook een kleine appendix met daarin enkele handige tools die je binnen Wikidata kan gebruiken, een overzicht van de mapping van door musea gebruikte objectnamen met Wikidata-termen, informatie over 'notability' en enkele voorbeelden van SPARQL-queries.

Actie 8²⁰. Workshop en stuurgroep II

De afsluitende stuurgroepvergadering van het project had eigenlijk moeten plaatsvinden voor eind december 2015, maar gezien vertragingen in de voorgaande deelprojecten besloten we het project met één maand te verlengen. Om die reden vond de afsluitende stuurgroep pas plaats op 29 januari 2015. Volgende zaken werden onder andere besproken:

- resultaten project en upload;
- blik op toekomstige acties en afronding data-upload;
- voorbeelden van hergebruik/visualisaties.

Na de middag werd een workshopgedeelte voorzien, waarin de projectpartners op basis van het Handboek zelf (hun) data leerden beheeren in Wikidata, enkele nieuwe records aanmaakten, bepaalde creators opzochten, afbeeldingen linkten, etc. Deze workshop (en stuurgroepvergadering) vond plaats onder begeleiding van Sandra en Maarten, zodat er ook mogelijkheid was tot vragen stellen over het hele projecttraject.

Conclusies

1. Data publicatie op Wikidata

Op 29 januari 2016 stonden er 26.680 kunstwerken uit de partnermusea op Wikidata. In vergelijking: in oktober 2015, aan het begin van het project, waren dat er 65. Die 26.680 kunstwerken zijn gemaakt door 3.615 kunstenaars en daarvan zijn enkele honderden kunstenaars door dit project nieuw op Wikidata ingevoerd. Er werden 399 verschillende objectnamen uit de partnercollecties gemapt met begrippen op Wikidata: van acrylschilderij tot zwart-witfoto. Dit totale aantal opgeladen werken representeert nog niet alle aangeleverde informatie. Zo zijn er bv. nog werken van S.M.A.K. opgeladen na de stuurgroep (29/1/2016).

Overzicht van aangeleverde informatie vs. wat gepubliceerd was op Wikidata op 29 januari 2016:

	records in csv	items op Wikidata
KMSKA	7.799	7.798
Groeningemuseum	2.858	2.859
MSK Gent	6.973	6.973
Mu.ZEE	3.979	3.978
S.M.A.K.	2.579	1.934
M - Museum Leuven	4.727	3.137
VKC	2.537	3
CVG	1.933	2

²⁰ Actie 7. refereerde naar dit Eindrapport.

Dat betekent dat er na het formele einde van het project nog een aantal acties moeten ondernomen worden.

2. Werkplan controle en vervollediging data op Wikidata

Samen met Sandra en Maarten plant PACKED vzw nog enkele uploads die mappings moeten actualiseren, werk aan het koppelen van delen met gehelen (bv. triptiek), aanvullen van informatie rond creators, Na afloop van het project werd hiervoor een werkplan gemaakt om de gepubliceerde data te controleren en te vervolledigen²¹. Dit werkplan omvat de volgende acties:

- correcties in de mapping van persistente URI's;
- upload data SMAK, M HKA, VKC en CVG;
- koppeling van items die één kunstwerk vormen (bv. diptieken, reeksen);
- controle materialen;
- controle dateringen;
- toevoegen van rollen voor vervaardigers;
- vervolledigen handboek;
- update projectpagina op Wikidata
- activeren van (dode) persistente URI's (cf. Resolvers).

Streefdatum om dit werkplan af te werken is eind juni 2016.

Partners vanaf dan zelf informatie over nieuwe aanwinsten in hun collectie opnemen in Wikidata.

3. Publicatie werkmaterialen

De materialen die in dit project zijn aangemaakt - vaak lijvige documenten zoals bv. de Whitepaper en het Handboek - worden in de komende weken integraal vertaald naar het Engels en online raadpleegbaar gemaakt via de [projectpagina op Wikidata](#)²² (de Whitepaper is op dit moment al in het Engels beschikbaar). Op die manier kan de ervaring die opgedaan werd in dit 'regionale' project een model van aanpak vormen voor andere musea die op een gelijkaardige wijze willen te werk gaan en wordt de impact ervan significant vergroot.

²¹ https://docs.google.com/spreadsheets/d/1_9Qur1ruBJq3BK7k6LONSWAakypvehjH6hbeQ7O6VE8

²² https://www.wikidata.org/wiki/Wikidata:Flemish_art_collections,_Wikidata_and_Linked_Open_Data

Bijlage 1: Events-data statistieken

De tabellen hieronder geven een overzicht van collectiedata die we naar 'events' hebben kunnen vertalen. Dit is niet per se een overzicht van alle geregistreerde data omdat niet alle data werd meegenomen tijdens de mapping naar LIDO, maar op die manier zien we wel waar er bruikbare gegevens zitten. Voor ons was dit een vertrekpunt om visualisaties te maken (licht groen)

Toelichting excel:

- KMSKA: CSV-bestand werd niet naar LIDO-xml overgezet, dus cijfers komen uit ruwe exports;
- M HKA: enkel Adlib export, geen data uit Ensembles;
- CVG: een volledige export gekregen;
- Percentages: gemiddeld aantal events per kunstwerk (dus geen uitspraken over volledigheid);
- Production: werd aangemaakt als er ten minste één stukje van de nodige data aanwezig was (datum/actor/place);
- Provenance: bevatten vaak dubbels omdat data vaak twee keer word geregistreerd in 'verwerving' en 'eigendomsgeschiedenis'

Tabel 1. Algemeen overzicht van beschikbare events-data

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
1		Kunstwerken	Events													
2			Production	Acquisition		Provenance		Exhibitions and loans		Publication and research		Condition		Iconografie		
3	Groeningemuseum	6584	6521	99%	4684	71%	3093	47%	889	14%	2149	33%	2329	35%	nvt	nvt
4	KMSKA	5370	5370	100%	nvt	nvt	5331	99%	2987	56%	5879	109%	nvt	nvt	2192	41%
5	MSKG	9059	8444	93%	6538	72%	10085	111%	3295	36%	9524	105%	1612	18%	3448	38%
6	SMAK	3917	3917	100%	3917	100%	4521	115%	8797	225%	6580	168%	1959	50%	nvt	nvt
7	MuZEE	5670	5621	99%	5434	96%	2463	43%	757	13%	293	5%	1686	30%	nvt	nvt
8	MHKA	2593	2590	100%	2514	97%	468	18%	5723	221%	1327	51%	497	19%	nvt	nvt
9	Middelheimmuseum	1285	1285	100%	0	0%	4	0%	243	19%	nvt	nvt	0	0%	nvt	nvt
10	CVG	9734	9202	95%	9407	97%	8643	89%	8842	91%	694	7%	5187	53%	nvt	nvt
11																
12	Totaal / Gemiddeld	44212	42950	98%	32494	76%	34608	65%	31533	84%	26446	68%	13270	29%	5640	39,44%
13																

Tabel 2. Production event

	A	B	C	D	E	F	G	H	I	J	K	L
1	Production	Aantal events	Identificatie		Actor		Place		Date		Title	
2	Groeningemuseum	6521	0	0%	6505	100%	0	0%	6028	92%	6316	97%
3	KMSKA	5370	0	0%	5360	100%	0	0%	nvt*	nvt	5368	100%
4	MSKG	8444	0	0%	8386	99%	0	0%	4214	50%	8405	100%
5	SMAK	3917	0	0%	3917	100%	0	0%	3065	78%	3917	100%
6	MuZEE	5621	0	0%	5584	99%	0	0%	3819	68%	5535	98%
7	MHKA	2590	0	0%	2585	100%	196	8%	2312	89%	2581	100%
8	Middelheimmuseum	1285	0	0%	1280	100%	57	4%	nvt**	nvt	1285	100%
9	CVG	9202	0	0%	8870	96%	18	0%	5309	58%	9158	100%
10												
11	Totaal / Gemiddeld	42950	0	0%	42487	99%	271	2%	24747	72,58%	42565	99%
12												
13	* opgenomen in de 'title' veld na een comma											
14	** production.period meegegeven ipv date.start en date.end											
15												

Tabel 3. Acquisition event

	A	B	C	D	E	F	G	H	I	J	K	L
1	Acquisition	Aantal events	Identificatie		Actor - repeller		Actor - acquirer		Place		Date	
2	Groeningemuseum	4684	0	0%	4238	90%	3585	77%	1811	39%	4110	88%
3	KMSKA	nvt	nvt	nvt	nvt	nvt	nvt	nvt	nvt	nvt	nvt	nvt
4	MSKG	6538	0	0%	5912	90%	6412	98%	3569	55%	5937	91%
5	SMAK	3917	0	0%	2488	64%	3916	100%	10	0%	1938	49%
6	MuZEE	5434	0	0%	4055	75%	4865	90%	4	0%	5036	93%
7	MHKA	2514	0	0%	1883	75%	1844	73%	500	20%	2148	85%
8	Middelheimmuseum	0	nvt	nvt	nvt	nvt	nvt	nvt	nvt	nvt	nvt	nvt
9	CVG	9407	0	0%	945	10%	55	1%	114	1%	9205	98%
10												
11	Totaal / Gemiddeld	32494	0	0%	19521	67%	20677	73%	6008	19%	28374	84%
12												

Tabel 4. Provenance event

	A	B	C	D	E	F	G	H	I	J	K	L
1	Provenance	Aantal events	Identificatie		Actor - repeller		Actor - acquirer		Place		Date	
2	Groeningemuseum	3093	0	0%	2865	93%	3042	98%	1532	50%	2831	92%
3	KMSKA	5331	nvt	nvt	nvt*	nvt	nvt	nvt	nvt	nvt	nvt	nvt
4	MSKG	10085	0	0%	6613	66%	9329	93%	4798	48%	6782	67%
5	SMAK	4521	0	0%	1436	32%	3560	79%	14	0%	3331	74%
6	MuZEE	2463	0	0%	1885	77%	2398	97%	4	0%	2049	83%
7	MHKA	468	0	0%	110	24%	450	96%	46	10%	427	91%
8	Middelheimmuseum	4	0	0%	1	25%	1	25%	1	25%	0	0%
9	CVG	8643	0	0%	0	0%	296	3%	19	0%	8112	94%
10												
11	Totaal / Gemiddeld	34608	0	0%	12910	45%	19076	70%	6414	19%	23532	72%
12												
13	* informatie is in een veld opgenomen mbv ;											
14	Provenance: ook verdubbeling van verwerking info aanwezig											
15												

Tabel 5. Exhibition event

	A	B	C	D	E	F	G	H	I	J	K	L
1	Exhibition	Aantal events	Identificatie*		Actor		Place		Date		Title	
2	Groeningemuseum	889	782	88%	743	84%	781	88%	889	100%	889	100%
3	KMSKA	2987	nvt	nvt	nvt**	nvt	nvt	nvt	nvt	nvt	nvt	nvt
4	MSKG	3295	3191	97%	3101	94%	3179	96%	3176	96%	3193	97%
5	SMAK	8797	5051	57%	4930	56%	5092	58%	5076	58%	5092	58%
6	MuZEE	757	74	10%	734	97%	634	84%	742	98%	682	90%
7	MHKA	5723	2283	40%	4686	82%	3086	54%	5704	100%	3437	60%
8	Middelheimmuseum	243	34	14%	52	21%	23	9%	23	9%	243	100%
9	CVG	8842	66	1%	311	4%	2070	23%	2890	33%	4311	49%
10												
11	Totaal / Gemiddeld	31533	11481	44%	14557	62%	14865	59%	18500	71%	17847	79%
12												
13	*databanknummer of dossiernummer											
14	** informatie is in een veld opgenomen mbv ;											
15												

Tabel 6. Publication event

	A	B	C	D	E	F	G	H	I	J	K	L
1	Publication	Aantal events	Identificatie		Actor		Place*		Date*		Title	
2	Groeningemuseum	2149	2146	100%	1649	77%	1772	82%	2088	97%	1820	85%
3	KMSKA	5879	5879	100%	nvt**	nvt	nvt	nvt	nvt	nvt	nvt	nvt
4	MSKG	9524	9512	100%	8453	89%	nvt	nvt	nvt	nvt	9509	100%
5	SMAK	6580	5960	91%	5655	86%	5932	90%	5960	91%	5960	91%
6	MuZEE	293	233	80%	169	58%	218	74%	231	79%	233	80%
7	MHKA	1327	0	0%	1169	88%	nvt	nvt	nvt	nvt	1327	100%
8	Middelheimmuseum	nvt	nvt	nvt	nvt	nvt	nvt	nvt	nvt	nvt	nvt	nvt
9	CVG	694	0	0%	4	1%	nvt	nvt	nvt	nvt	349	50%
10												
11	Totaal / Gemiddeld	26446	23730	67%	17099	66%	7922	82%	8279	89%	19198	84%
12												
13	* verrijkt waar mogelijk met data uit expot van de Adlib-Bibliotheek module											
14	** informatie is in een veld opgenomen mbv ;											
15												

Tabel 7. Condition event

	A	B	C	D	E	F	G	H	I	J
1	Condition assessment	Aantal events	Identificatie		Actor		Place		Date	
2	Groeningemuseum	2329	0	0%	1802	77%	0	0%	1730	74%
3	KMSKA	nvt	nvt	nvt	nvt	nvt	nvt	nvt	nvt	nvt
4	MSKG	1612	0	0%	1524	95%	0	0%	1521	94%
5	SMAK	1959	0	0%	1304	67%	0	0%	1571	80%
6	MuZEE	1686	0	0%	432	26%	0	0%	134	8%
7	MHKA	497	0	0%	34	7%	0	0%	41	8%
8	Middelheimmuseum	nvt	nvt	nvt	nvt	nvt	nvt	nvt	nvt	nvt
9	CVG	5187	0	0%	3920	76%	0	0%	4060	78%
10										
11	Totaal / Gemiddeld	13270	0	0%	9016	58%	0	0%	9057	57%
12										