Deliverable

Project: IIIF-Beeldinfrastructuur gekoppeld aan de VKC-Datahub. Fase 1: proefopstelling
Werkpakket: WP3 Analyse koppeling met de onderbouw.

Subsidie: Projectsubsidie voor een internationaal cultureel-erfgoedproject
Periode: 2 juli 2018 - 31 augustus 2019

D3: Analyse van de koppeling met de onderbouw

Auteurs: Matthias Vandermaesen (VKC), Jolan Wuyts (VKC), Sarah Eycken (VKC)

Inleiding

Deze deliverable heeft als doel de creatie, borging, overdracht, ontsluiting en archivering van beelden in het huidige ecosysteem te beschrijven. Deze omgang met beelden en beeldbeheer in de breedste zin is wat in deze deliverable gedefinieerd wordt als 'onderbouw'. Het gaat hier om beelden die ontstaan door het maken van digitale reproducties van kunstwerken, zij het om materiaaltechnische redenen, ter reproductie van het werk zelf, voor onderzoeksdoelen, etc. Deze huidige situatie wordt in deze deliverable geanalyseerd, op basis waarvan een situatie wordt geschetst waar naartoe kan gewerkt worden. Via een SWOT-analyse identificeren we de sterktes, zwaktes, opportuniteiten en risico's in de huidige situatie. In de beschrijving van een mogelijke toekomstige situatie wordt gekeken hoe op de opportuniteiten en sterktes van de huidige situatie kan worden ingezet en hoe de risico's en zwaktes kunnen worden aangepakt.

Het VKC-ecosysteem

Wat is het VKC-ecosysteem?

De Vlaamse Kunstcollectie (VKC) is een dienstverlenende organisatie. Ze staat in voor het begeleiden van museale actoren in hun digitale trajecten, werkend op het kruispunt van musea en digitaal cultureel erfgoed. Momenteel zijn het KMSKA, het Groeningemuseum, M Leuven, Mu.ZEE, en het MSK Gent aangesloten bij de Vlaamse Kunstcollectie. Een andere actor waar de Vlaamse Kunstcollectie nauw mee samenwerkt is de fusieorganisatie VIAA/PACKED/LUKAS (VIAA+).

Het “ecosysteem” is een verzameling van softwaretoepassingen en -systemen die toelaten om gegevens over de collectie via het web te publiceren. We kunnen ze groeperen in 3 lagen die onderling gegevens uitwisselen:

- Laag 1: **Lokale systemen**: de collectiebeheersystemen en beeldbanken die onder het beheer van de musea vallen.
- Laag 2: **VKC-platform**: de databanken die de data uit de collectiebeheersystemen integreren en ter beschikking stellen van toepassingen. Het beheer van die platformen is de verantwoordelijkheid van VKC en wordt deels uitbesteed aan technische leveranciers. Het VKC-platform bestaat momenteel uit twee componenten: de Datahub en de Imagehub. De server voor de Datahub wordt ingehuurd bij Inuits, en deze voor de imagehub bij VIAA.
- Laag 3: **Toepassingen**: gebruikerstoepassingen zoals websites en mobiele apps die data en beelden in het VKC-ecosysteem beschikbaar maken op het web. De Vlaamse kunstcollectie ziet de Arthub als zijn grootste zelfbeheerde gebruikerstoepassing. De server voor de Arthub wordt ingehuurd bij Inuits.

Huidige situatie van het beeldbeheer

Het uitwisselen van beelden van de VKC-musea verloopt moeizaam en is zeer arbeidsintensief. Dat heeft te maken met het feit dat de uitwisseling nog steeds gebaseerd is op manuele processen waarbij gebruik gemaakt wordt van o.a. e-mail, WeTransfer, USB-sticks en harde schijven. Dat stelde PACKED (thans VIAA+) vast tijdens hun doorlichting van de technische onderbouw van de Vlaamse Kunstcollectie, waarvan de conclusies te vinden zijn op CEST. In dit subhoofdstuk wordt verder ingegaan op die conclusies om de huidige situatie te schetsen per partner van de VKC.

KMSKA

Sinds 2010 zet het KMSKA in op het systematisch fotograferen van haar volledige collectie. Op twee- tot driehonderd langdurige bruiklenen na, is de volledige collectie intussen gefotografeerd. Alle moeder- en raadplegingsbestanden worden centraal bewaard op een server die de IT-dienst van het KMSKA beheert. Er werden geen specifieke maatregelen genomen voor het beheer (ontdubbeling, selectie, creatie van verschillende raadplegingsformaten) en de preservering (integriteitscontrole, transcoderen naar archiveringsformaten) van die beelden. De wildgroei aan kopieën op de verschillende computers binnen het KMSKA wordt als een probleem ervaren. Het KMSKA startte reeds een project voor de ingebruikname van het DAM-systeem (Digital Asset Management) ResourceSpace. Hiertoe werkte het KMSKA reeds een metadatamodel uit voor het beschrijven van afbeeldingen in die software. De bedoeling is om alle beelden die momenteel verspreid over verschillende dragers en locaties bewaard worden, samen te brengen en beschikbaar te maken via één platform voor alle medewerkers.

MSK Gent

Net als bij het KMSKA worden de moederbestanden van de digitale beelden verspreid bewaard. De beelden die het MSK zelf maakt, staan op een server van de stad Gent. De beelden die door Lukas (thans VIAA+) gemaakt worden, bewaart VIAA+ (Lukas) zelf. Toch bezit het MSK in principe ook een kopie van de moederbestanden.
De raadplegingsbestanden die beschikbaar worden gesteld aan Erfgoedinzicht en VKC, worden in Adlib bewaard. Er worden geen specifieke maatregelen genomen voor het beheer (ontdubbeling, selectie, verschillende raadplegingsformaten) en de preservering (integriteitscontrole, transcoderen naar archiveringsformaten) van de digitale beelden. Beelden worden uitgewisseld via DVD's en USB-sticks. Soms laat de Stad Gent het MSK toe beelden uit te wisselen via een FTP-verbinding.

Musea Brugge

De digitale beelden van het Groeningemuseum zijn afkomstig van drie bronnen:

- kunstwerken gefotografeerd door de stadsfotograaf;

- schilderijen uit de depots en de tekeningencollectie gefotografeerd door een fotograaf in opdracht van LUKAS;
- de prentencollectie en analoge foto's gescand door het Groeningemuseum.

Van al die beelden bewaart het Groeningemuseum een kopie. Aangezien de kwaliteit van de oudere gedigitaliseerde beelden inferieur is, wordt doorgaans een nieuw beeld gemaakt bij elke vraag om een digitaal afbeelding van een werk uit de collectie. Alle versies blijven bewaard, deels omdat men dit historisch interessant vindt, deels omdat men geen duidelijk overzicht heeft van kwantiteit en kwaliteit van de digitale beelden. De stad Brugge bewaart de moederbestanden op een server. Ze staan op een externe harde schijf als back-up. Er worden geen specifieke maatregelen genomen voor het beheer (ontdubbeling, selectie, verschillende raadplegingsformaten) en de preservering (integriteitscontrole, transcoderen naar archiveringsformaten) van de digitale beelden. De uitwisseling van de beelden met de VKC verloopt via DVD's of via een FTP-server.
Op korte termijn zijn er plannen om het beheer van de moederbestanden onder te brengen in een 'mediabank'. Deze beeldbank zou zowel de preservering van de moederbestanden als de online terbeschikkingstelling van raadplegingsbestanden voor zijn rekening nemen. Het is nog niet duidelijk of daarvoor een oplossing binnen of buiten de stad zou worden uitgedokterd.

Gezien M Leuven en Mu.ZEE binnen het bestek van dit project geen contentpartner zijn, laten we het beeldbeheer van deze musea voorlopig buiten beschouwing.

Vlaamse Kunstcollectie

Niet enkel de drie musea, maar ook de VKC verzamelt beelden van de verschillende VKC-musea. Deze beelden werden op een externe server geplaatst voor de VKC-onlinecatalogus. Uitwisseling van de beelden met de musea verliep via manuele processen, onder meer door middel van DVD's, USB-sticks of FTP. Aanvankelijk werden hiervoor de raadplegingsbestanden gebruikt die in de collectiebeheersystemen zaten. Later werden beelden met een te lage kwaliteit ad hoc vervangen door nieuwe digitale beelden die gemaakt werden door Lukas (thans VIAA+). In de loop van het proces zijn er technische vereisten voor digitale representaties afgesproken, maar die zijn niet formeel vastgelegd.

Momenteel is het grootste deel van de beelden die de Vlaamse Kunstcollectie bezit en ontsluit afkomstig van een server, beheerd door Cit. Die beelden zijn van lage resolutie, dikwijls gedateerd, en kunnen niet direct beheerd worden door de VKC.

Lukas, Art in Flanders (VIAA+)

De beelden van de musea die aan VIAA+ (Lukas, art in Flanders) werden toevertrouwd, worden bij VIAA+ (Lukas) beheerd in Mediahaven. Zowel de metadata als een hogeresolutiebestand en een lageresolutiebestand worden in Mediahaven bewaard. Tot deze DAM is er in principe toegang via de Mediahaven REST API.

VAA heeft plannen om in 2019-2020 een pilot op te zetten om de beelden minstens via de IIIF Image API te ontsluiten. Er zijn ook plannen om de IIIF Presentation API te implementeren.[1]

SWOT-analyse van de huidige situatie

Een analyse van de huidige situatie moet de sterktes, zwaktes, opportuniteiten en risico's van het huidige veld blootstellen. We focussen ons op de uitwisseling van beelden bij de musea en haalden inspiratie uit de doorlichting van de technische infrastructuur van de VKC-musea door PACKED.

Sterktes

- Het zelf beheren van beelden door musea betekent dat zij de maximale controle behouden over die beelden en kunnen bepalen op welke manier en om welke reden ze worden ontsloten.
- Het gebruik van ad hoc-dragers voor beelden stelt museale medewerkers in staat op een simpele manier beelden uit te wisselen zonder dat er nood is aan technische bijstand (*nvdr. de keerzijde van deze medaille is dat alles ook trager, inefficiënter en manueel gebeurt*).
- De musea blijven verzekerd van de interne controle op de beelden. Beelden moeten immers niet bewaard worden op externe servers door externe organisaties.

Zwaktes

- Er is geen centraal overzicht van afbeeldingen die aangeven waar de 'moederbestanden' in verschillende formaten en op de hoogste resolutie gevonden kunnen worden.
- Er is weinig tot geen uitwisseling van beelden door de musea naar de buitenwereld. Beelden zijn enkel beschikbaar voor eindgebruikers via de website van Lukas, de koepelwebsite en thematische websites van de VKC en via Wikimedia Commons voor een deel van de beelden.

---

[1] VIAA+ (Packed) leverde tevens de volgende info aan:
Hoe worden ze ter beschikking gesteld via API?
- Syncrator > API laag op de verschillende beheersystemen (inclusief Mediahaven)
- Datastorage/Index > metadata reservoir dat metadata ophaalt via de syncrator en via een top API laag metadatasets aanlevert op maat van eindgebruiker applicaties

Hoe kan er toegang via IIIF worden gerealiseerd?
- Image API op niveau van het MAM > voor pilot: Cantaloup?
- Presentation API > in de API laag boven op de data storage

Hoe kan er toegang tot de API worden opgezet? Is er authenticatie noodzakelijke?
- Pilot: afgeschermde api.
- Lange termijn: open IIIF api met authenticatie

Welke beeldensets kunnen er voor het project ter beschikking worden gesteld?
- Pilot: subset van VKC beelden maken > sets van Groeninge, MSKGent, KMSKA, Muzee en MLeuven

Hoe kan voor het project een koppeling met de Lukas-infrastructuur worden opgezet?
- Cantaloup imageserver opzetten met http koppeling naar VKC beelden in object store.

- De infrastructuur is niet schaalbaar. De huidige oplossingen, zoals de koepelwebsite van de VKC en de thematische websites, zijn niet geschikt voor het ontsluiten van een groot aantal beelden op hoge resolutie uit verschillende instellingen.
- Voor de huidige oplossingen hangen musea nog bijna compleet af van externe actoren die het beheer in handen nemen en hun diensten aanbieden zonder centrale sturing van de musea.

Kansen

- Een intern DAM-systeem kan de basis vormen van een open en efficiënt beeldbeheer. Als musea hun afbeeldingen centraal en eenduidig beheren in een intern DAM-systeem dat open API's bezit, kan hierop voortgebouwd worden om beelden te laten doorstromen naar (externe) hergebruiksapplicaties.
- Een overkoepelend beleid voor beeldbeheer opent opportuniteiten om gedigitaliseerde werken makkelijk over te dragen tussen musea, om interinstitutioneel onderzoek te ondersteunen, en om tentoonstellingen en andere publiekswerking op te starten tussen de musea onderling.

Bedreigingen

- Door de afwezigheid van een centraal intern systeem is er een grote dreiging tot het verlies van data. Huidige ad hoc-dragers zoals optische schijven, USB-sticks en harde schijven zijn geen duurzame dragers voor digitaal cultureel erfgoed. Deze fysieke dragers zijn ook verbonden aan de medewerkers die ze gebruiken. Die dragers dreigen dan ook te verdwijnen of in onbruik te geraken wanneer medewerkers de organisatie verlaten of van functie veranderen.
- Een gedeelde interne schijf waar iedereen toegang toe heeft, is een stap in de goede richting, maar resources zijn nodig voor een actief onderhoud en archivering van deze digitale representaties. Zonder actief onderhoud worden mappenstructuren te complex, worden versies en formaten niet bijgehouden, met verlies van data tot gevolg. Ook is het moeilijk om een gelaagde toegang tot beelden mogelijk te maken met een gedeelde harde schijf. Gebruikersbeheer en beveiliging zijn minder evident bij een interne gedeelde harde schijf, waardoor de controle over afbeeldingen nogmaals verloren gaat.
- Het verspreiden van beelden over verschillende dragers zonder duidelijk beheer creëert een enorme overhead. Het vinden van de juiste beelden duurt lang en moet via verschillende kanalen gaan (mondelinge aanvraag, e-mail, WeTransfer, FTP-servers, USB-sticks…). Op deze manier werken is inefficiënt en zorgt voor onnodig dubbel werk.

Situatie “To Be”

Waar wil een toekomstige infrastructuur op inzetten?

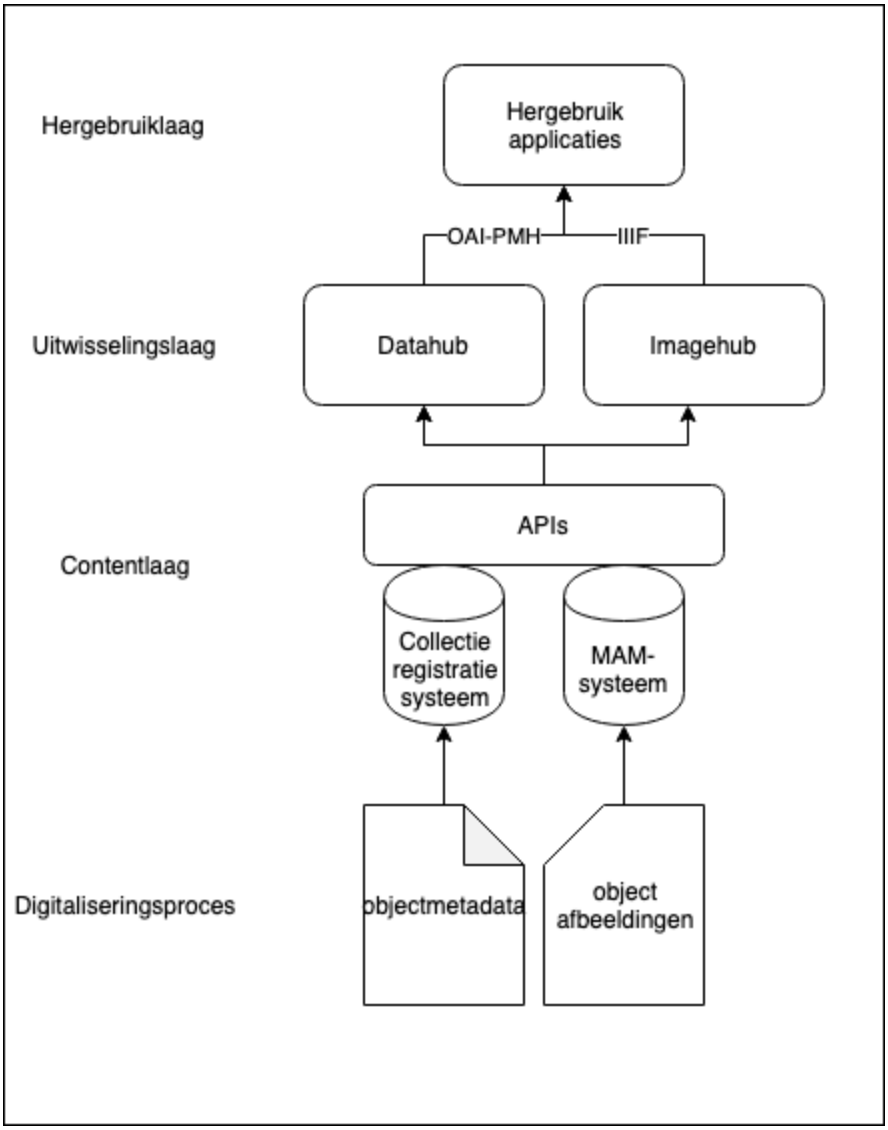
De twee grootste bedreigingen in het huidige beheer van de digitale reproducties zijn:

- De afwezigheid van een beleid rond beeldbeheer. Dit leidt tot een versplintering van de locatie van en toegang tot interne beeldbestanden. Dat leidt op zijn beurt dan weer tot dubbel werk, een grote overhead, en het potentieel verlies van waardevolle beelden. Hoe een beleid rond beeldbeheer kan worden gevoerd staat uitgebreid op CEST beschreven in de Blauwdruk Gedistribueerd Beeldbeheer.
- De afwezigheid van een duurzame infrastructuur die instaat voor de ontsluiting en het hergebruik van digitale reproducties door externe gebruikers. Actieve inzet, samenwerking en een duidelijk traject om de auteursrechtelijke status van de beelden uit te klaren, zijn elementen die door deze infrastructuur in de weg gestaan wordt.

Deze deliverable wil een aanpak voorstellen om die twee bedreigingen weg te werken. De infrastructuur die we hier voorstellen,houdt er rekening mee dat:

- Musea interne digitale platformen nodig hebben om afbeeldingen te beheren. Een beleid moet dicteren op welke manier deze DAM-systemen gebruikt worden en moet ervoor zorgen dat het beheer van beelden ingebed wordt in de dagelijkse werking van de musea.
- Musea beelden op een geautomatiseerde manier moeten kunnen aanleveren voor hergebruiksapplicaties.

De voorgestelde infrastructuur kan op onderstaande manier schematisch worden voorgesteld. Het IIIF-beeldinfrastructuurproject is een eerste stap om dit diagram te implementeren.
Het deel van dit schema over de objectmetadata werd al grotendeels geoperationaliseerd in de Datahub Fase 1 en Fase 2-projecten. Objectmetadata worden via de collectieregistratiesystemen



van de VKC-musea beschikbaar gesteld via de Datahub, die via OAI-PMH protocol doorstroomt naar hergebruiksapplicaties zoals de Arthub.
Het deel van dit schema voor de beelden werd getest in het proefproject IIIF-beeldinfrastructuur Fase 1. Met eventuele volgende fasen van dit project willen we de volgende operationele doelstellingen realiseren:

- Het aanbieden van beelden op hoge resolutie binnen de Arthub en de thematische clusters, voor alle records van werken in de Datahub. Dit veronderstelt een verdere doorontwikkeling van de proefopstelling.
- Metadatering van beeldmateriaal binnen het ecosysteem, met bijzondere aandacht voor de interdependenties en de auteursrechtenstatus, zowel van het originele werk als van het gedigitaliseerde beeld.

- Het aanbieden van een gebruiksvriendelijke viewer die alle functionaliteiten bevat die door gebruikers als belangrijk worden gezien.
- Een robuuste infrastructuur waarbij afbeeldingen altijd online beschikbaar en duidelijk identificeerbaar zijn.
- Het openstellen van beelden van werken in het publieke domein als open data onder Creative Commons-licenties.
- Een betere controle over het beeldmateriaal binnen het museale ecosysteem via een centrale ingest module, die gebouwd is op de MAM-systemen gebruikt door musea of daartoe aangestelde derden.
- Een service-oriented architectuur met losse herbruikbare componenten. Deze infrastructuur moet kunnen worden gekoppeld onafhankelijk van softwareoplossingen (D/MAM's) die door musea worden gekozen voor hun collectieregistratie of beeldbeheer.
- Een opensourceinfrastructuur die door derden inzetbaar is.
- Een platform waar andere musea makkelijk op kunnen aansluiten, los van hun technische infrastructuur.

Korte beschrijving van een toekomstige infrastructuur

Verschillende componenten dragen bij aan een duurzame beeldbeheerinfrastructuur voor musea. We beschrijven ze van de meest belangrijke en interne naar de meest gebruiksgerichte componenten.

Beeldbeheerbeleid

Een duurzaam beeldbeheer begint bij het opzetten van een duurzaam beeldbeleid zoals uitvoerig werd beschreven in de Blauwdruk Gedistribueerd Beeldbeheer. Bij de digitalisering van erfgoedobjecten moet een duidelijk beleid uitgewerkt zijn voor de opslag, metadatering, verrijking, bewaring, en ontsluiting van de gecreëerde data. De data moeten bij het aanmaken in een archiveringsformaat opgeslagen en gearchiveerd worden, zodat 'moederfiles' altijd teruggevonden en gebruikt kunnen worden. Bij het ontstaan van data moet het museale personeel een persistente identifier aanmaken die 1 file persistent identificeert met 1 onveranderlijk nummer dat verwijst naar 1 uniek data-object. Deze afbeelding-PID moet ingeschreven worden in de ingebedde metadata van de afbeelding, en moet vermeld worden bij de metadata over het object zelf als een verwijzing naar een *resource* van dat object.
Een beleid voor beeldbeheer moet ook in detail treden over waar afgeleide bestanden van de moederfiles worden bewaard, en wie er toegang toe heeft. Een personeelslid dat instaat voor het beeldbeheer moet controle hebben over de opslag, toegang en ontsluiting van de moederfiles en afgeleide bestanden. Een MAM-systeem helpt in het beheer van die afbeeldingen. De afbeeldingen moeten gekoppeld kunnen worden aan de inhoudelijke, administratieve en technische metadata over het object dat ze representeren. Er moet nagedacht worden over hoe er moet omgegaan worden met updates aan afbeeldingen, de verwijdering van afbeeldingen, en de toegang tot afbeeldingen van werken die nog niet in het publieke domein liggen.

Een traject voor het klaren van auteursrechten op beelden

Een beleid over beeldbeheer moet duidelijk kunnen stellen hoe omgegaan wordt met auteursrecht op digitale reproducties van een cultureel object. Dit beleid volgt idealiter het gezamenlijk standpunt over beeldbeheer dat de Vlaamse Kunstcollectie uitschreef, alle VKC-partnermusea ondertekenden en werd overgemaakt aan VIAA/PACKED/Lukas (of VIAA+). Volgens het standpunt kunnen er geen auteursrechten meer worden geclaimd op digitale reproducties van werken die in het publieke domein zitten of op een andere manier geklaard zijn van alle auteursrecht. Een traject voor het klaren van rechten op gedigitaliseerde werken moet een eerste stap zijn om alle digitale reproducties van werken van een museale collectie open te stellen.

Auteursrechten op een gedigitaliseerde reproductie én op het cultureel erfgoedobject zelf moeten opgenomen worden in de metadata en ingebed worden in het beeld. Die auteursrechten moeten verwijzen naar een Creative Commons of een rightsstatement.org-licentie. Auteursrechten en rechten op hergebruik moeten op een gebruiksvriendelijke en simpele manier worden meegedeeld aan gebruikers van hergebruiksapplicaties. Dit alles conform de Blauwdruk Gedistribueerd Beeldbeleid.

Dergelijke stappen zijn uitermate belangrijk om de auteursrechtenstatus technisch en geautomatiseerd kenbaar te kunnen maken via de uitwisselingsinfrastructuur en hergebruiksapplicaties zoals de Arthub.

DAM-systeem

Een Digital Asset Management Systeem of DAMS helpt bij het operationaliseren van het museaal beleid voor beelden. Een DAMS zorgt dat de afbeeldingen die museale medewerkers en gebruikers respectievelijk intern en extern raadplegen op één plaats te vinden, beheren en bewaren zijn. Medewerkers en instanties die instaan voor beeldbeheer, kunnen dit DAM-systeem beheren en ervoor zorgen dat de juiste mensen toegang hebben tot de juiste afbeeldingen. Een DAM-systeem zorgt ervoor dat beelden op eenvoudige wijze gebruikt kunnen worden voor tentoonstellingen, campagnes, en het dagelijkse werk van museale medewerkers. Een DAM-systeem beschikt over API's die de koppeling met de hergebruikslaag en uitwisselingslaag mogelijk makkelijk maken. Een DAM-systeem moet schaalbaar en inwisselbaar kunnen zijn, wat wil zeggen dat het zonder problemen duizenden beelden moet kunnen verwerken, en dat het makkelijk ingewisseld moet kunnen worden door een ander systeem.

IIIF-uitwisselingslaag

Eens afbeeldingen via de API's van een DAM-systeem uitgewisseld kunnen worden, moet de uitwisselingslaag ervoor zorgen dat deze afbeeldingen op een automatische, open en herbruikbare manier aangeleverd kunnen worden ten behoeve van hergebruiksapplicaties. IIIF zorgt voor de ontsluiting van beelden op een open en herbruikbare manier. De Imagehub is een oplossing van de Vlaamse Kunstcollectie voor de implementatie van de IIIF Image API en Presentation API in de infrastructuur van Vlaamse musea. Die oplossing kan gekoppeld worden

aan de infrastructuur van de musea en opent zo de weg naar (intermuseale samenwerking voor) het opzetten van hergebruiksapplicaties, die de content van (verschillende) musea als bron nemen.

Via een imagehub per museum kunnen we desgewenst een decentraal netwerk van IIIF-endpoints vormen, die gebruikt kan worden voor collectiewebsites, thematische websites, … of om een arthub of gelijkaardige koepelsite op te zetten. Een centrale imagehub kan zorgen voor een gezamenlijke ontsluiting van beelden van verschillende musea in een enkele hergebruiksapplicatie, zoals de VKC-Arthub. Een gedetailleerde analyse van de uitbouw van de Imagehub is te vinden in *D4: een aanpak voor Fase 2*.

Hergebruikslaag

Door het gebruik van IIIF kunnen hergebruiksapplicaties makkelijk voorzien worden van content uit de uitwisselingslaag. Een IIIF-viewer, zoals de Universal Viewer, kan gebruikt worden om de beelden en hun metadata te tonen. Collectieluiken, digitale tentoonstellingen, thematische websites kunnen éénvoudig gebouwd worden met content uit de uitwisselingslaag.