

IIF-beeldinfrastructuur gekoppeld aan VKC-ecosysteem, fase 2: Hybride collecties, annotaties, multilayerviewer en IIF-koppeling met meemoo

WP2. IIF binnen meemoo en uitwisseling van relevante metadata tussen het VKC-ecosysteem en het meemoo-ecosysteem

Taak 2: IIF binnen meemoo

Binnen deze taak wordt ernaar gestreefd om de beelden van meemoo via de IIF-Image-API uitwisselbaar te maken. Dit veronderstelt een analyse van de huidige infrastructuur, het zoeken naar de juiste componenten en de feitelijke implementatie en testing.

Deliverable 5. Een IIF-API-endpoint op de meemoo-MAM

Medewerkers werkpakket 2

Matthias Priem, Bart Debunne, Miel Vander Sande, Rony Vissers, Nastasia Vanderperren, Timothy Naessens, Bart De Sitter (meemoo)
Pascal Ennaert, Rob Wyse, Karin Van der Poorten (VKC),
Alec Van den broeck (VKC-UGent)
Michiel Hebben (Hero Solutions)



Belangrijk disclaimer	1
Een IIF-endpoint op de meemoo-MAM in proefopstelling	2
Functionele processing workflow	5
Dynamisch versus vast herschalen	6
Openstaande issues	8
Openstaande issue 1: Metadata toegevoegd aan de digitale beelden	8
Openstaande issue 2: Wat in geval van vigerende auteursrechten?	8

Belangrijk disclaimer

In het IIF2-project engageerde de VKC zich om tal van applicaties en tools 'in een proefversie' op te leveren. Gezien de applicaties en tools 'proef- of testversies' zijn wil dit zeggen dat ze na oplevering verder worden getest en eventueel aangepast voor ze in productieversie worden uitgerold.

Dit geldt ook voor de IIF-Endpoint op de Meemoo-MAM. Als deze Endpoint niet actief is, dan wordt er mogelijk aan verbeterd, of is er ondertussen een productieversie actief. Stuur een mailtje naar info@vlaamsekunstcollectie.be als je vragen hebt over de stand van zaken. We nemen dan zo snel mogelijk contact op om een demonstratie te geven of om toegang te bieden.

Een IIIF-endpoint op de meemoo-MAM in proefopstelling

Binnen deze taak werd ernaar gestreefd om de beelden van meemoo via de IIIF Image API uitwisselbaar te maken. Dit veronderstelde een analyse van de huidige infrastructuur, het zoeken naar de juiste componenten en de feitelijke implementatie en testing.

Uiteindelijk werd een IIIF-endpoint op de meemoo-MAM in proefopstelling opgeleverd. De proefopstelling bevat een tweehonderdtal gedigitaliseerde foto's uit de collecties van de VKC-partnermusea.

Bedoeling is dat de proefopstelling nu verder wordt getest door de VKC en de VKC-partnermusea. Wanneer de proefopstelling goed wordt bevonden kunnen de volledige gedigitaliseerde fotocollecties van de VKC-partnermusea die aan meemoo in beheer werden gegeven worden vrijgegeven via de IIIF-endpoint op de meemoo-MAM. Dit is voorzien voor na de projectperiode.

De architectuur van de IIIF-Endpoint op de meemoo-MAM ontstond in onderling overleg tussen de collega's van meemoo en VKC. De door meemoo geschreven architectuurnota vind je onder meer via volgende links:

https://www.projectcest.be/wiki/Publicatie:IIIF_Image_API_endpoint_bij_meemoo en <https://github.com/viaacode/iiif-deliverables/blob/main/pdf/d4-workflow.pdf>.

De broncode van de proefopstelling met alle relevante documenten en issues kan je terugvinden in GitHub via de volgende link: <https://github.com/viaacode/iiif-image-processing>

Meemoo leverde t.b.v. de slotmeeting van dit werkpakket ook een presentatie aan met de uitleg over de werking van de IIIF-Endpoint op de meemoo-MAM en de mogelijkheid tot het herschalen van de foto's. Deze presentatie vind je hier:

https://github.com/viaacode/iiif-deliverables/blob/main/presentaties/IIIF_meemoo-VKC_slot_W2.pdf

De nota over de Image Processing Workflow en het herschalen van de foto's vind je hier:

<https://github.com/viaacode/iiif-deliverables/blob/main/pdf/d4-workflow.pdf>

De testomgeving voor de IIIF-image-endpoint op de meemoo-MAM (tst) vind je hier:

<https://images-tst.meemoo.be/iiif>

Een csv met de in de proefopstelling beschikbaar gemaakte digitale beelden en hun verschillende meemoo-id's en andere inventarisnummers kan op eenvoudig verzoek door meemoo of de VKC worden aangeleverd. Verder in deze deliverable vind je een aantal voorbeelden van dergelijke meemoo-id's.

Het oproepen van de beelden via de IIIF-endpoint op de meemoo-MAM kan via een URL. Deze URL wordt gevormd door een string met volgende IIIF-syntax:

<https://images-tst.meemoo.be/iiif/<meemoo-id>.jpg2/full/max/0/default.jpg>

Om beelden uit de proefopstelling op te roepen volstaat het bijgevolg dat je de external_id van meemoo (fluo-deel van de hierboven afgebeelde URL) uit de hierboven vermelde csv toevoegt. Bijvoorbeeld:

<https://images-tst.meemoo.be/iiif/416sx7vq3v.jp2/full/max/0/default.jpg>

<416sx7vq3v> is de unieke meemoo-ID voor de digitale afbeelding van het werk *De portretten van Jan zonder Vrees en Margaretha van Beieren* van een anoniem meester uit de Zuidelijke Nederlanden uit de collectie van het MSK Gent.

De identifier voor het bijhorende manifest kan je oproepen via volgende string:

<https://images-tst.meemoo.be/iiif/<meemoo-id>.jp2>

Het manifest bij het in vorige alinea genoemde beeld is dan

<https://images-tst.meemoo.be/iiif/416sx7vq3v.jp2>

De base URL's resollen automatisch naar: <iiif-url>/info.json

<https://images-tst.meemoo.be/iiif/416sx7vq3v.jp2/info.json>

De proefopstelling laat niet enkel IIIF toe. Ook volgende protocollen zijn mogelijk:

- IIP: Internet Imaging Protocol
- <https://images-tst.meemoo.be/fcgi-bin/iipsrv.fcgi?FIF=/416sx7vq3v.jp2&CVT=png>
- Zoomify
- Deepzoom

Voor meer info over de verschillende protocollen, zie:

<https://iipimage.sourceforge.io/documentation/protocol/>

Probeer gerust de volgende voorbeeldwerken uit via de toegevoegde unieke meemoo-ID's:

Uit de collectie van de Musea Brugge:

De Moreel-triptiek van Hans Memling, Foto Dominique Provost

meemoo-ID: 1z41r8cz66

Madonna met kanunnik Joris van der Paele van Jan van Eyck, Foto Dominique Provost

meemoo-ID: xk84j23793

Kaart van Brugge en de feodale gronden van Joannes Lobberecht, Foto Hugo Maertens

meemoo-ID: t43hx2xz13

Uit de collectie van het MSK Gent:

De baden in Oostende van James Ensor, Foto Hugo Maertens

meemoo-ID: tm71v7392h

De nimf Salmacis en Hermaphroditus van François-Joseph Navez, Foto Hugo Maertens

meemoo-ID: z02z33df0n

Vechtende mannen van George Minne, Foto Dominique Provost

meemoo-ID's: wp9t16kg44, dr2p57529r en p843r2fm4m

De heilige Hiëronymus van Jheronimus Bosch, Foto Dominique Provost

meemoo-ID: r785h93n1g

Portret van Laurent Van de Velde van Henry Van de Velde, Foto Hugo Maertens

meemoo-ID: 9g5gb3ph3z

Uit de collectie van het KMSKA:

Triptiek van de Zeven Sacramenten van Rogier Van der Weyden, Foto Hugo Maertens

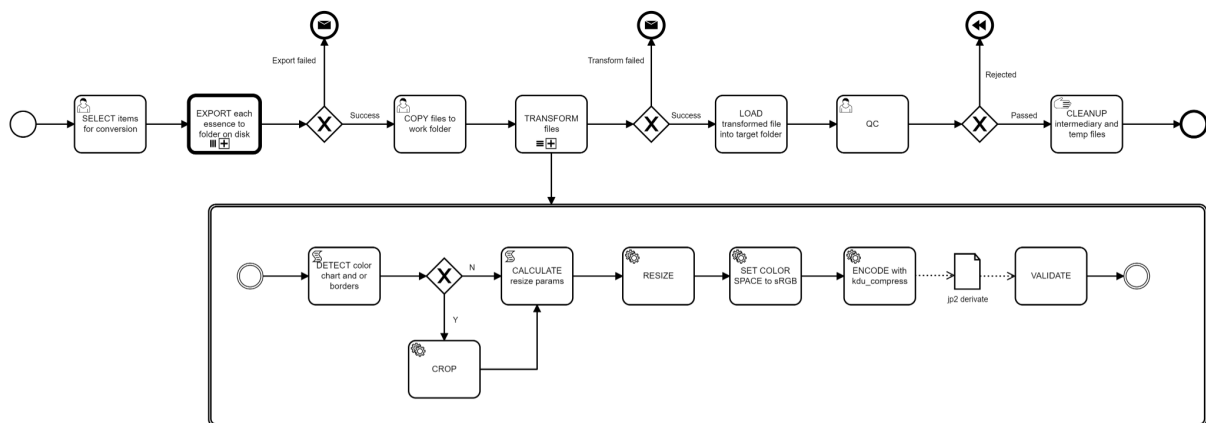
meemoo-ID's: 3775t56m1w, br8md0h46t, zg6g17k910 en gh9b584x12

Functionele processing workflow

Binnen de opgeleverde opstelling doorloopt het digitale beeld volgende functionele processing workflow:

- Het beeld wordt geëxporteerd uit het archief a.d.h.v. de meemoo external_id (andere id's zijn mogelijk).
- De metadata wordt opgehaald uit de MAM (REST API).
- Indien er een kleurenkaart aanwezig is wordt deze automatisch verwijderd.
- Er volgt een dynamische herschaling (voor meer uitleg, zie hieronder).
- De colorspace wordt naar sRGB omgezet (indien er een ander icc-profiel is).
- Er volgt een 'compress' en 'encode' als jp2 met behulp van Kakadu.
- De metadata-tags van het origineel worden toegevoegd (xmp, iptc).
- Het gegenereerde beeld wordt opgeslagen in de media folder t.b.v. de image server.

De workflow ziet er schematisch als volgt uit¹:



Meer info over de image processing workflow vind je hier:

<https://github.com/viaacode/iiif-deliverables/blob/main/pdf/d4-workflow.pdf>

¹ Een beter beeld vind je in de [powerpoint](#) waarvan sprake in dit document. Het beeld is beter leesbaar als je op de tekst inzoomt.

Dynamisch versus vast herschalen

Een belangrijk element in dit werkpakket was de denkoefening rond het herschalen van foto's. Het leek ons om verschillende redenen weinig zinvol om de allergrootste beelden via de endpoint ter beschikking te stellen. Dit zou immers de gebruikerservaring in belangrijke mate beperken (vertragen) zonder duidelijke meerwaarde.

In eerste instantie werd afgesproken dat er een vaste herschaling zou worden gehanteerd. Bij vaste herschaling wordt de langste of kortste zijde herleid tot een vooraf bepaalde waarde, waarbij het resulterend bestand voldoende detail bevat en tegelijk de performantie waarborgt.

Initieel werd dit voor het IIF2-project bepaald als (maximaal) 5.000 pixels voor de kortste zijde. Dit betekende sowieso reeds een aanzienlijke verbetering ten aanzien van de vorige situatie. Immers, in het door de VKC en haar partnermusea gehanteerde gelaagd licentiemodel uit 2016, werd op basis van de richtlijnen van Lukas, Art in Flanders, maximaal 1.920 pixels in de breedte afgesproken voor werken in het publiek domein.

Niettemin betekent 5.000 pixels voor de kortste zijde voor bestanden die merkkelijk groter zijn dan gemiddeld mogelijk nog steeds een ingrijpende herschaling met kwaliteitsverlies.

Op basis van een representatieve testset van 77 afbeeldingen bleek dat er in elk geval geen waarneembare correlatie kon worden vastgesteld tussen de fysieke afmeting van een werk en de resolutie in pixels. In de testset clusterde de resolutie rond een mediaan van 6.050 pixels, met het merendeel tussen 4.000 en 8.000 pixels, afnemend in aantal boven 10.000 pixels tot enkele zeldzame pieken boven de 20.000 pixels (max. 25.000 pixels).

Op basis van deze analyse werd besloten dat een betere balans tussen opslagcapaciteit, performantie en kwaliteit gewaarborgd moest worden. Een dynamische herschaling leek beter dan een vaste herschaling. Daarom werden volgende principes afgesproken.

- Afbeeldingen tot 5.000 pixels breedte worden ongemoeid gelaten
- Afbeeldingen van 5.001 tot 10.000 pixels worden op 50% herschaald.
- De breedte mag maximaal 10.000 pixels bedragen.

Hierop werd de referentieset als representatieve controleset gehanteerd om de impact van bovenvermelde principes op het herschalen na te gaan.

Hieruit bleek dat maximaal 15 procent van de originele archiefbestanden moest worden herschaald. Het verschil tussen herschalen of niet-herschalen leverde een besparing op van ongeveer 20% aan opslagcapaciteit, in de praktijk ongeveer 1 GB per 200 beelden.

Het herschalen gaat vanzelfsprekend gepaard met een visueel kwaliteitsverlies bij het inzoomen. Bij de grootste beelden die verhoudingsgewijs meer herschaald zijn, is een verzachting van de contouren merkbaar. Het kwaliteitsverlies is progressief in die zin dat een origineel beeld van 15.000 pixels breed relatief meer effect zal ondervinden van herschaling dan een beeld van 7.000 pixels breed. Hoewel er geen detail of kleur verloren lijkt te gaan,

zijn de edges en details minder scherp waardoor de afbeelding minder 'crisp' overkomt op de hogere zoomniveau's.

Voor materiaal gelijkaardig qua resolutie aan de referentieset is de resulterende kwaliteit aanvaardbaar. Indien in de toekomst nog grotere beelden worden aangeboden dienen de drempelwaardes voor het herschalen echter herbekeken te worden. In elk geval betekende het lineair of vast herschalen tot een grens van 5.000 pixels voor sommige werken een te grote degradatie van de kwaliteit van de beelden en was dergelijke vaste of lineaire herschaling daarom niet wenselijk.

Meer info over het herschalen vind je via volgende links:

https://www.projectcest.be/w/images/D4_IIIF_Image_endpoint_bij_meemoo_workflow.pdf of <https://github.com/viaacode/iiif-deliverables/blob/main/pdf/d4-workflow.pdf>

Openstaande issues

Hieronder worden twee issues belicht die in de toekomst kunnen worden opgenomen.

Openstaande issue 1: Metadata toegevoegd aan de digitale beelden

Het overlegd beeldbeleid van de VKC en haar partners veronderstelt dat de belangrijkste aan het beeld gerelateerde metadata ook in het beeld worden meegegeven. De in de proefopstelling gebruikte IIPImage kan metadata uit de bronbestanden meenemen naar de output. Maar dit is voorlopig beperkt tot TIFF- en XMP-outputbestanden. Vooralsnog laat IIPImage niet toe dat de XMP metadata uit jp2-files worden toegevoegd aan de output tiles. Meemoo bracht dit probleem aan bij de beheerders van IIPImage in de hoop dat deze belangrijke tekortkoming zo snel mogelijk wordt verholpen.

Openstaande issue 2: Wat in geval van vigerende auteursrechten?

De VKC heeft momenteel een kaderovereenkomst met SABAM/Unisono lopen voor de publicatie op verschillende websites van werken van kunstenaars die nog niet onder het publiek domein vallen en in de portfolio van SABAM zitten. In de overeenkomst met SABAM/Unisono staat volgende bepaling over dergelijke werken: *'De resolutie van het gereproduceerde beeld mag niet meer dan 640 X 480 pixels zijn en een definitie van maximaal 72 DPI hebben.'* Het spreekt voor zich dat de kwaliteit van deze beelden mijlenver afstaat van de eerder genoemde pixelbreedte en resolutie. Bovendien laat SABAM/Unisono door haar overeenkomsten niet toe dat dergelijke auteursrechtplichtige werken kunnen worden gekopieerd of gedownload.

In de proefset in de IIF-Endpoint op de meemoo-DAM zitten een beperkt aantal auteursrechtplichtige werken (van Jean Brusselmans, Dan Van Severen, Roger Raveel, Georges Vantongerloo, Frans Masereel, Max Ernst, Henri Van de Velde, Edgard Tytgat, Oskar Kokoschka, Marcel Broodthaers, Constant Permeke, Jan Cox, Henri Victor Wolvens, René Magritte, Kees Van Dongen).

Het lijkt ons niet haalbaar om dergelijke auteursrechtplichtige werken in dezelfde formaten te publiceren en ter beschikking te stellen als de Public Domain-werken. Er zal moeten worden afgewogen of dergelijke werken al initieel, in de interne workflow van de IIF-Endpoint op de meemoo-DAM, moeten worden herschaald, of als dat dient te gebeuren op het moment van publicering in de IIF-vensters. Deze knoop moet worden doorgehakt alvorens de proefopstelling van de IIF-Endpoint op de Meemoo-DAM wordt uitgerold voor de volledige collecties van de VKC-partnermusea.

Gezien ook meemoo een interne workflow hanteert voor het ter beschikking stellen van specifieke foto's en de copyrightstatus daar een belangrijke rol in heeft, lijkt het voor de VKC althans belangrijk dat dergelijke kwaliteitsbeperking door de copyrightstatus reeds via het ter beschikking stellen via de interne workflow van de Endpoint wordt opgenomen.