

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

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

Taak 3: Uitwisseling van metadata tussen de VKC-datahub en de meemoo-mediahaven-DAM

Binnen deze taak wordt ernaar gestreefd om de beelden van meemoo via de IIIF 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 7. Component die de metadata uit de datahub synchroniseert met records in het meemoo-MAM

Medewerkers werkpakket 2

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



In deliverable 6 kon je lezen dat het niet alleen de bedoeling is dat de digitale afbeeldingen uit de collecties van de VKC-musea via een IIIF-Endpoint op de meemoo-MAM (Mediahaven) geautomatiseerd kunnen doorstromen van het meemoo-ecosysteem naar het VKC-ecosysteem, maar tevens dat de museale metadata uit de museale collectiebeheersystemen van de VKC-musea via de datahub moeten kunnen doorstromen naar de meemoo-mediahaven-MAM. Op die manier kunnen de digitale afbeeldingen van de museumobjecten in het meemoo-systeem geautomatiseerd worden gemetadateerd. Dit zou op termijn een behoorlijke efficiëntiewinst betekenen gezien de metadata voor nieuwe beeldbestanden niet meer manueel zouden moeten worden aangemaakt en toegevoegd in het meemoo-systeem, maar automatisch via de datahub zouden kunnen worden opgehaald.

Om dit te realiseren werd er door meemoo een permanente pipeline opgezet die de VKC-datahub dagelijks bevroegt en de records uit de VKC-datahub exporteert naar de beeldenserver van meemoo, en dat in het metadatamodel van meemoo.

Deze pipeline werd in het kader van het IIIF2-project in een proefopstelling ontwikkeld:

- Een mapping werd gemaakt tussen LIDO/XML en het meemoo-metadatamodel in Mediahaven: Voorbeeld:
<https://docs.google.com/spreadsheets/d/1QClaxEujfMqWdrwDn16Qv1a9rPL890TmCXWBRwQY2Ko/edit#gid=0>;
- Records worden opgevraagd via de publiek toegankelijke OAI-PMH endpoint op de datahub (waarmee nog eens wordt aangetoond dat iedereen de metadata uit de datahub kan hergebruiken);
- Via een ETL-pipeline worden de gegevens omgezet en ingeladen in Mediahaven:
 - o Deze ETL-pipeline maakt gebruik van Apache Airflow (<https://airflow.apache.org/>) die een aantal Python scripts in een workflow plaatst die op geregelde tijdstippen kan worden herhaald. De transformatie gebeurt op basis van XSLT met de Saxon C library (<https://www.saxonica.com/saxon-c/index.xml>)

Alle informatie over de mapping en de pipeline is te vinden op GitHub:

<https://github.com/viaacode/datahub-metadata-sync>

Deze proefopstelling wordt in de komende maanden verder getest door meemoo. Als er geen problemen opduiken kan een productie-opstelling worden klaargezet via welke dan niet alleen de metadata in de meemoo-mediahaven-MAM via de datahub rechtstreeks worden aangepast cf. de beschrijvingen in de collectiebeheersystemen van de musea, maar deze metadata ook gepubliceerd worden op het meemoo-portaal, de Art in Flanders-website <https://artinflanders.be/nl> of andere meemoo-portalen.