

Het gebruik van Computer Vision API's voor de beschrijving van cultureel-erfgoedcollecties

Onderzoeksvoorstel Bachelorproef

Nastasia Vanderperren¹

Samenvatting

Om de registratieachterstand bij collectiebeherende instellingen in Vlaanderen weg te werken, dient een inhaalbeweging gemaakt te worden. We stellen in dit onderzoek voor om hiervoor Computer Vision API's in te schakelen. Het onderzoek wil zich focussen op het analyseren van de metadata-velden die geschikt zijn voor Computer Vision, de voorwaarden die nodig zijn om tot een succesvol resultaat te komen en een vergelijking te maken tussen de bestaande diensten. We willen weten of deze software in de toekomst gebruikt kan worden om het werk van de registratoren te verlichten en achterstanden weg te werken.

Sleutelwoorden

Onderzoeksdomein. Machine learning en kunstmatige intelligentie — Computer Vision — Digital humanities

Contact: ¹ nastasia.vanderperren.v8632@student.hogent.be

Inhoudsopgave

1	Introductie	1
2	Stand van zaken	2
2.1	Computer Vision en (cultureel) erfgoed	2
2.2	Theoretisch onderzoek	3
2.3	Computer Vision API's	3
3	Methodologie	3
4	Verwachte resultaten	4
5	Verwachte conclusies	5
	Referenties	5

1. Introductie

De Vlaamse Overheid investeert in het Vlaamse cultureel erfgoed door collectiebeherende instellingen te ondersteunen en kwaliteitslabels uit te reiken. In ruil voor die steun verwacht de Vlaamse Overheid dat die instellingen een aantal taken uitvoeren, waaronder het registreren, inventariseren en metadata-reken van cultureel-erfgoedobjecten op een gestandaardiseerde manier, het onderzoeken (en faciliteren van onderzoek) en het presenteren van de collectie. (Agentschap Kunsten en Erfgoed, 2014; Gatz, 2016)

De collectiebeherende instellingen lijden aan een historische achterstand m.b.t. de registratie van de eigen collectie (Gatz, 2016). In 2018 werd daarom een nieuwe subsidielijng opgestart om de digitale collectieregistratie weg te werken. Deze subsidielijng werd opgestart vanuit de vaststelling dat de competenties en strategieën ontbreken om een inhaalbeweging te realiseren (Departement Cultuur, Jeugd en Media, 2018a).

In de bachelorproef willen we onderzoeken in welke mate Computer Vision API's (vanaf nu afgekort als CVA), zoals Google Cloud Vision¹ of Microsoft Computer Vision API², ingezet kunnen worden om dit registratieproces te versnellen en als strategie gebruikt kunnen worden om een inhaalbeweging te realiseren. Momenteel moet ieder erfgoedobject geregistreerd worden door een domeinexpert. Dit is tijdrovend werk. Met behulp van artificiële intelligentie (AI) kan dit proces ten dele geautomatiseerd worden. Dit geeft de collectieregistrator de mogelijkheid om zich met minder basaal werk bezig te houden, zoals het verrijken van de data of het klaarmaken van de data om ze als open (linked) data vrij te geven, nog een beleidslijn waar de Vlaamse Overheid op wil inzetten (Departement Cultuur, Jeugd en Media, 2018b).

Figuren 1 t.e.m. 4 geven een voorbeeld van de registratieachterstand in musea. De cijfers zijn afkomstig van de VKC Datahub Dashboard.³ Dit dashboard analyseert de collectieregistratie van de Vlaamse musea voor Schone (en in de toekomst ook Hedendaagse) Kunsten die aangesloten zijn bij de VlaamseKunstCollectie (VKC)⁴. Het is bedoeld om de museale medewerker inzicht te geven in de staat van de collectiemetadata van zijn museum in functie van digitale bruikbaarheid. Ze meten onder meer hoeveel records ingevuld zijn volgens de minimale en basisregistratie.⁵ Dit zijn cijfers

¹<https://cloud.google.com/vision/>

²<https://azure.microsoft.com/nl-nl/services/cognitive-services/computer-vision/>

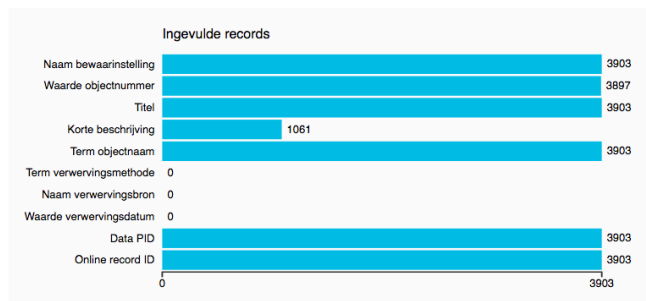
³<https://dashboard.vlaamsekunstcollectie.be>

⁴<http://vlaamsekunstcollectie.be/>

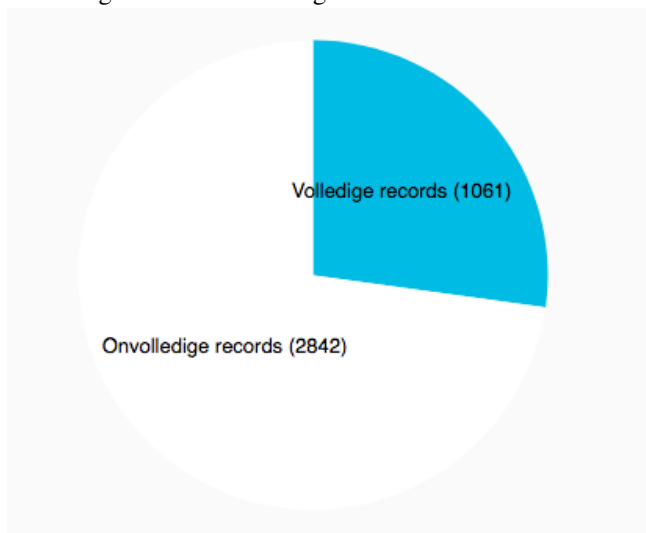
⁵De CIDOC-richtlijnen, ook wel de minimale registratie genoemd, gelden als de absolute minimum voor het beschrijven van objecten in museale collecties. Ook de basisregistratie is een minimumrichtlijn die afkomstig is van CIDOC. Voor meer info, zie CIDOC Fact Sheet No. 1: Registration step by step(1993) en CEST-Richtlijn: Objecten registreren

van de grootste kunstmusea van Vlaanderen.

Figuur 1. Het overzicht van de velden van de minimale registratie en het aantal keer dat ze aanwezig zijn in de collectiedata van MSK Gent



Figuur 2. Het aantal volledig ingevulde records van MSK Gent volgens de minimale registratie



Uit deze cijfers blijkt dat vooral formele en administratieve gegevens geregistreerd worden; inhoudelijke informatie, zoals afgebeelde persoon of afgebeeld concept, die interessant is voor ontsluiting en onderzoek, ontbreekt hoofdzakelijk.

In het verleden werd reeds (theoretisch) onderzoek verricht naar het gebruik van Computer Vision voor cultureel erfgoed en kunst.⁶ De resultaten van dit onderzoek waren veelbelovend. Nieuwe ontwikkelingen zorgen ervoor dat Computer Vision steeds meer accuraat wordt en steeds eenvoudiger in gebruik, waardoor het meer en meer een instrument wordt waarmee ook developers aan de slag kunnen (Hindle, 2017).

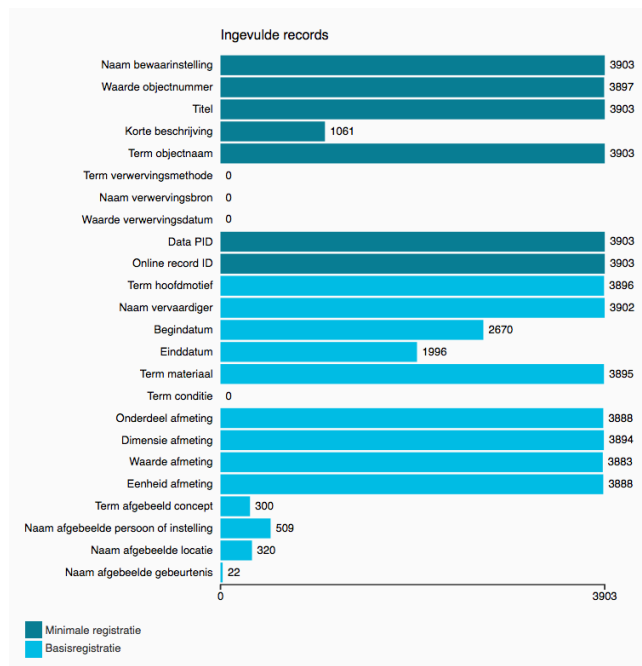
2. Stand van zaken

2.1 Computer Vision en (cultureel) erfgoed

Een aantal instellingen maken al gebruik van Computer Vision voor de ontsluiting van hun collectie:

⁶Zie infra, in het deel *Stand van zaken*.

Figuur 3. Het overzicht van de velden van de basisregistratie en het aantal keer dat ze aanwezig zijn in de collectiedata van MSK Gent



Figuur 4. Het aantal volledig ingevulde records van MSK Gent volgens de basisregistratie

Er zijn geen volledig ingevulde records.

- **The Museum of Modern Art (MoMA)**⁷ gebruikte diensten van Google via *Google Arts & Culture Lab* om aan de historische foto's van afgelopen tentoonstellingen in MOMA foto's uit de kunstcollectie te koppelen. Het algoritme analyseerde hiervoor alle foto's van tentoonstellingszichten. Als het een kunstwerk op de foto's kon herkennen, dan werd er een koppeling gelegd met de afbeelding van dit kunstwerk in de collectie van MOMA.⁸ MOMA stelde hierbij vast dat het logaritme vooral goed scoort op het vlak van tweedimensionale, statische afbeeldingen (bv. een afbeelding van een kunstwerk), maar dat het minder goede resultaten geeft met afbeeldingen van 3D-objecten (bv. een standbeeld) of bewegend beeld (MOMA, 2018?).
- Het **Smithsonian**⁹ gebruikt AI-technologie om stalen van planten te classificeren. Het Smithsonian is gestart met het systematisch digitaliseren van de collectie voor wetenschappers en in functie van online ontsluiting. De AI-technologie slaagde erin om via deze afbeeldingen twee gelijkaardige planten te onderscheiden met een

⁷<https://www.moma.org/>

⁸Bekijk een voorbeeld over een tentoonstelling over Cézanne, Gauguin, Seurat en Van Gogh uit 1929

⁹<https://www.smithsonianmag.com/>

succesgraad van meer dan 90% (Smith, 2017).

- Het **Nasjonalmuseet**¹⁰ was het onderwerp van het *Principal Components* onderzoek. Hierbij werd met het deep learning framework Caffe¹¹ gezocht naar compositionele gelijkenissen tussen kunstwerken en werden ze geïconiseerd op basis van de Iconclass-termen¹². Dit resulteerde in een vernieuwende publiekstoegang tot de collectie waarbij de kunstwerken op basis van gelijkenissen gevisualiseerd werden. Hoe meer gelijkenissen, hoe dichter de kunstwerken bij elkaar staan.¹³ (Nasjonalmuseet, 2017?; Westvang, 2017?)

2.2 Theoretisch onderzoek

Daarnaast is er eerder theoretisch onderzoek gedaan naar het gebruik van Computer Vision binnen een erfgoedcontext. In onderstaande lijst worden de onderzoeken vermeldt die zich focussen op museumcollecties:

- **Using Machine Learning for Identification of Art Paintings (2013)**: In dit onderzoek werd machine learning gebruikt voor het classificeren van kunstwerken van zeven kunstenaars: Cézanne, Dali, Dürer, Monet, Picasso, Rembrandt en Van Gogh. Per kunstenaar werden er tweehonderd kunstwerken gezocht. In 87,13% van de gevallen was de computer correct. De onderzoekers vermoedden dat dit resultaat nog beter kan zijn als de training set groter is. (Blessings & Wen, 2013)
- **The Rijksmuseum Challenge: Museum-Centered Visual Recognition (2014)**: Beelden die publiek beschikbaar zijn via de API van Het Rijksmuseum¹⁴ werd gebruikt voor het oplossen van vier challenges:
 - voorspel de kunstenaar van het afgebeelde object;
 - voorspel het materiaal dat gebruikt werd voor het afgebeelde object;
 - voorspel het jaar waarin het object gemaakt werd;
 - voorspel het soort object (schilderij, tekening, standbeeld...) dat afgebeeld wordt.

Het ging om objecten die afkomstig waren van de oudheid tot de late 19e eeuw en om een veelheid aan objecten: schilderijen, foto's, keramiek, meubels, etc. Ook in dit onderzoek waren de resultaten veelbelovend. (Mensing & van Gemert, 2014)

- **INSIGHT (2017-2020)**: Hier wil men onderzoeken hoe AI kan gebruikt worden om collecties uit de cultureel-erfgoedsector van beschrijvende metadata te voorzien. De collecties van de Koninklijke Musea voor Schone Kunsten van België en de Koninklijke Musea voor Kunst & Geschiedenis worden als testcase gebruikt. De focus ligt op het vrijgeven van die data als open datasets. (Universiteit Antwerpen, 2017?) Uit de paper *Deep Transfer Learning for Art Classification Problems*

van Sabatelli, Kestemont, Daelemans en Geurts (2018) kwamen veelbelovende resultaten naar boven voor het voorspellen van materiaal, objecttype en kunstenaar.

- **Automated Image Analysis with IIIF (2017)**: Dit onderzoek is eerder praktisch gericht en werd uitgevoerd buiten de academische wereld. Het werd uitgevoerd door CogApp, een bedrijf dat software ontwikkelt voor online archieven en musea.¹⁵ Zij hebben verschillende testen gedaan met machine learning. Het meest interessant voor dit voorstel is het onderzoek waarbij drie Computer Vision API's (Google Cloud Vision, Microsoft Computer Vision en Clarifai¹⁶) getest werden om beelden te voorzien van extra tags om de doorzoekbaarheid te verbeteren. Naast de grappige resultaten (Roddis, 2018), leidde dit tot goede resultaten waarmee de beeldencollectie op een interessante manier doorzocht kan worden: Je kan er een selectie van beelden verkrijgen door te filteren op verschillende tags, bv. *ik wil een kunstwerk uit de Renaissance van iemand met een snor en een cape*.¹⁷ De onderzoekers concludeerden dat de CVA accurate beschrijvingen geven en dat Computer Vision steeds eenvoudiger in gebruik wordt. De foutieve beschrijvingen, zoals in Figuur 6, komen vermoedelijk doordat CVA getraind worden met hedendaagse beelden gemaakt met een smartphone of computer. (Hindle, 2017)

2.3 Computer Vision API's

Tot slot lichten we kort CVA toe. CVA, ook wel image/visual recognition API's genoemd, zijn API's die afbeeldingen automatisch kunnen taggen, organiseren en zoeken via machine learning. Het biedt de mogelijkheid om machine learning toe te passen zonder dat je hier zelf een expert in moet zijn. Je kan er zelf modellen mee creëren om de API's te trainen naar je eigen behoefte. Zo leerde Matt Fraser Google Cloud Vision verschillende spinnen te herkennen door de dienst te trainen met honderd afbeeldingen per spin (Fraser, 2018). De bekendste diensten zijn Clarifai, IBM Visual Recognition¹⁸, Microsoft Computer Vision en Google Cloud Vision. De diensten zijn zo toegankelijk dat iedere ontwikkelaar zonder machine learning kennis modellen kan bouwen, maar dat ook, volgens TechCrunch, iedereen zonder programmeerervaring de diensten kan gebruiken om afbeeldingen te taggen en categoriseren (Lardinois, 2018). Ook uit het onderzoek van CogApp bleken de diensten eenvoudig in gebruik en precies te trainen (Hindle, 2017).

3. Methodologie

Dit voorstel wil onderzoeken wat de kwaliteit is wanneer beelden van erfgoedcollecties geanalyseerd worden door CVA.

¹⁰<http://www.nasjonalmuseet.no/en/>

¹¹Voor meer info, zie: <http://caffe.berkeleyvision.org/>

¹²Iconclass is een gespecialiseerd kunsthistorisch classificatiesysteem, <https://nl.wikipedia.org/wiki/Iconclass>

¹³Bekijk bijvoorbeeld schilderijen op basis van hun motief.

¹⁴<https://www.rijksmuseum.nl/en/api>

¹⁵<https://www.cogapp.com/about/>

¹⁶<https://www.clarifai.com/>

¹⁷Probeer het zelf op: <http://labs.cogapp.com/iiif-ml/>

¹⁸<https://www.ibm.com/watson/services/visual-recognition/>

Figuur 5. Een vrouw wordt door Microsoft Computer Vision geïdentificeerd als een teddybeer (Roddis, 2018).

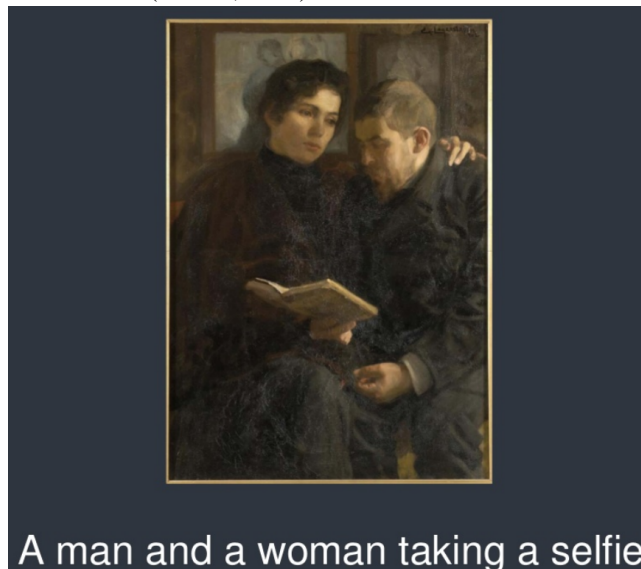


Kunnen de ingebouwde modellen van de API's gebruikt worden, of is er nood aan (doorgedreven) training? Welke metadata velden zijn het meest geschikt om door een CVA te laten taggen? Welke kwaliteit van afbeelding is nodig om een goed resultaat te verkrijgen? Kunnen enkel beelden op hoge resolutie gebruikt worden, of kan een museummedewerker gewoon foto's nemen van de erfgoedcollectie met zijn/haar smartphone en dit laten analyseren door de CVA? Dit zal onderzocht worden op basis van een proefopstelling.

Het onderzoek zal ook bestaan uit een vergelijkende studie van de verschillende CVA. Potentiële items waarop de CVA beoordeeld kunnen worden, zijn:

- **Kwaliteit:** Welke dienst behaalt de beste resultaten?
- **Training:** In welke mate kan de dienst getraind worden?
- **Automatisering:** Kan de API ingebouwd worden, of moeten afbeeldingen één per één geanalyseerd worden? Kan dit geautomatiseerd worden via scripting?
- **Prijs:** Cultureel-erfgoedinstellingen hebben doorgaans

Figuur 6. Vrouw en man met boek worden door Microsoft Computer Vision geïdentificeerd als man en vrouw die een selfie nemen (Roddis, 2018).



een beperkt budget¹⁹ en worden gesubsidieerd met gemeenschapsgeld.

- **Export en import van gegevens:** In welk dataformaat wordt de verkregen informatie ontvangen en kan ze geïmporteerd worden in de registratiesystemen die de instellingen gebruiken?

Oberoi (2016) heeft software geschreven om de vijf belangrijkste Computer Vision API's te vergelijken.²⁰ Er zal bekeken worden of deze software gebruikt en/of aangepast kan worden om de verschillende diensten te vergelijken.

Dit kan onderzocht worden op beelden van het KMSKA, MSK en/of Groeningemuseum. De reeds bestaande metadata van de kunstwerken bevinden zich als open dataset in de Arthub²¹/Datahub²² en Wikidata²³; hun kunstwerken werden gedigitaliseerd door LUKAS²⁴ en zijn, met uitzondering van die van MSK, ook als open dataset beschikbaar op Wikimedia Commons²⁵

4. Verwachte resultaten

De uitkomst van dit onderzoek wordt een methodiek om de verschillende API's te vergelijken, getrainde modellen en testresultaten. Er is zicht op de velden uit de basisregistratie waarvoor CVA gebruikt kan worden en een idee hoeveel de modellen getraind moeten worden om goede resultaten te

¹⁹Zie VRT NWS: "Vlaamse musea krijgen er 10 miljoen euro per jaar bij".

²⁰Cloudy Vision: <https://github.com/goberoi/cloudy-vision>

²¹<https://arthub.vlaamsekunstcollectie.be/nl>

²²<http://datahub.vlaamsekunstcollectie.be/>

²³https://www.wikidata.org/wiki/Wikidata:Main_Page

²⁴<http://www.lukasweb.be/>

²⁵Schilderijen in het Koninklijk Museum voor Schone Kunsten Antwerpen op Wikimedia Commons; Schilderijen in het Groeningemuseum op Wikimedia Commons

geven. Daarnaast is geweten welke CVA de beste kwaliteit aanbiedt (weliswaar voor de uitgewerkte use case). We hebben een overzicht van de verschillende CVA en hen beoordeeld op verschillende punten.²⁶

Afgaand op de resultaten uit reeds eerder gevoerd onderzoek, vermoeden we dat de CVA goed scoren op het herkennen van materiaal, het type object en de voorgestelde objecten en figuren op het object. De kwaliteit van de resultaten zal afhankelijk zijn van de specificiteit die verwacht wordt. Hoe specifieker de resultaten van de CVA moeten zijn, hoe meer de modellen getraind moeten worden.

5. Verwachte conclusies

CVA kan de metadata van een erfgoedobject verrijken met inhoudelijke informatie die nu niet voorkomt in de basisregistratie, zoals de sfeer en het gevoel dat op een kunstwerk weergegeven wordt en de gebruikte kleuren. Mogelijk is het zelfs in staat om bekende personen op de afbeelding te herkennen, zoals bij Hindle (2017). Deze informatie wordt nu niet opgenomen in collectiebeheersystemen, maar is wel interessant voor onderzoekers en het publiek. We verwachten dat de meerwaarde van de CVA vooral op dit vlak ligt: het toevoegen van inhoudelijke informatie waar een registrator nu geen tijd voor heeft.

De CVA zullen te beperkt zijn om gebruikt te worden als strategie om een inhaalbeweging te realiseren. Niettemin kunnen ze een goede aanvulling zijn voor de data en informatie bezorgen die zowel voor het museum, de onderzoeker als de erfgoedgeïnteresseerde interessant is. De CVA zal de registrator niet vervangen, maar kan een meerwaarde bieden aan diens werk. Registratoren zijn nu domeinexperten. We vermoeden dat de museumwereld zelf kritisch zal staan tegenover de technologie. "Kan zo'n computer wel zo goed registreren als ik?"

Ook als er goede resultaten komen uit dit onderzoek, dan zal de opname van de informatie in de collectiebeheersystemen een knelpunt zijn, zeker voor informatie waar nu geen velden voor voorzien zijn. In het Vlaamse erfgoedlandschap wordt hiervoor hoofdzakelijk Adlib²⁷ en The Museum System (TMS)²⁸ gebruikt. Dit zijn gesloten systemen waarvan de aanbieders niet gekend zijn om hun bereidwilligheid om het systeem te veranderen, noch beschikken de musea over het budget om de aanpassingen te kunnen bekostigen.

Deze technologie zal wel meer en meer ingebouwd worden in DAM-systemen²⁹. DAM-software met Computer Vision is al op de markt: Gezichtsherkenning is ingebouwd in verschillende DAM-systemen.³⁰

²⁶Zie *Methodologie*.

²⁷<http://www.adlibsoft.com/>

²⁸<https://www.gallerysystems.com/products-and-services/tms-suite/tms/>

²⁹Digital Asset Management Systemen, een systeem voor het opslaan, beheren en distribueren van digitale assets, zoals afbeeldingen.

³⁰Zoals bij ResourceSpace: <https://www.resourcespace.com/knowledge-base/user/facial-recognition>

Referenties

- Agentschap Kunsten en Erfgoed (Red.). (2014). *Handleiding bij het Cultureel-Erfgoeddecreet: Het Kwaliteitslabel*.
- Blessings, A. & Wen, K. (2013). *Using Machine Learning for Identification of Art Paintings*. Stanford University.
- Departement Cultuur, Jeugd en Media. (2018a, juli 12). Inhaalbeweging voor digitale collectieregistratie. Verkregen 8 december 2018, van <http://www.kunstenenerfgoed.be/nl/nieuws/inhaalbeweging-voor-digitale-collectieregistratie>
- Departement Cultuur, Jeugd en Media. (2018b, juni). Visie-nota: Een Vlaams cultuurbeleid in het digitale tijdperk. Vlaamse Regering.
- Fraser, M. (2018, maart 14). Using Google Cloud AutoML to classify poisonous Australian spiders. Verkregen 9 december 2018, van <https://shinesolutions.com/2018/03/14/using-google-cloud-automl-vision-to-classify-poisonous-australian-spiders/>
- Gatz, S. (2016, maart). Conceptnota aan de Vlaamse Regering. Naar een duurzame cultureel-erfgoedwerking in Vlaanderen. Een langetermijnvisie voor cultureel erfgoed en cultureel-erfgoedwerking in Vlaanderen. Vlaamse Regering.
- Hindle, A. (2017, juni 20). Automated image analysis with IIIF: Using Artificial Intelligence for bulk image analysis. Verkregen 8 december 2018, van <https://blog.cogapp.com/automated-image-analysis-with-iiif-6594ff5b2b32>
- Lardinois, F. (2018). Google's AutoML lets you train custom machine learning models without having to code. Verkregen 9 december 2018, van <https://techcrunch.com/2018/01/17/googles-automl-lets-you-train-custom-machine-learning-models-without-having-to-code>
- Mensink, T. & van Gemert, J. (2014). *The Rijksmuseum Challenge: Museum-Centered Visual Recognition*. ISLA Lab - Informatics Institute University of Amsterdam.
- MOMA. (2018?). Identifying art through machine learning: A project with Google Arts & Culture Lab. Verkregen 8 december 2018, van <https://www.moma.org/calendar/exhibitions/history/identifying-art>
- Nasjonalmuseet. (2017?). Project: "Principal Components". Verkregen 8 december 2018, van http://www.nasjonalmuseet.no/en/collections_and_research/collection_management/digital_collection_management/Project%3A+%C2%ABPrincipal+Components%C2%BB.b7C_wJjU4L.ips
- Oberoi, G. (2016, juli 11). Comparing the Top Five Computer Vision APIs. Verkregen 9 december 2018, van <https://goberoi.com/comparing-the-top-five-computer-vision-apis-98e3e3d7c647>
- Roddis, T. (2018, mei 16). When automated analysis goes wrong. Verkregen 8 december 2018, van <https://www.slideshare.net/Europeana/when-automated-analysis-goes-wrong-by-tristan-roddis-europeantech-conference-2018>

- Sabatelli, M., Kestemont, M., Daelemans, W. & Geurts, P. (2018). *Deep Transfer Learning for Art Classification Problems*. Universiteit Antwerpen en Université de Liège.
- Smith, R. P. (2017, november 3). How Artificial Intelligence Could Revolutionize Archival Museum Research. Verkregen 8 december 2018, van <https://www.smithsonianmag.com/smithsonian-institution/how-artificial-intelligence-could-revolutionize-museum-research-180967065/>
- INSIGHT. (2017?). Verkregen 8 december 2018, van <http://uahost.uantwerpen.be/insight/index.php/about/>
- Westvang, E. (2017?). Principal Components: Machine learning in search of the uncanny. Verkregen 8 december 2018, van <http://bengler.no/principalcomponents>