

Automatische beeldherkenning als instrument voor museumcollecties

Proces en resultaten - *lessons learned*

Annika Hendriksen

Symposium **Museumcollecties & AI**, 16 mei 2022

Automatische beeldherkenning

als instrument voor museumcollecties

- processtappen trainen model
- welke collecties en modellen: de pilots op een rij
- data issues - name resolver
- dataverrijking: aanvullende soortinformatie online
- pilot: 'lessons learned'
- interface: modellen gebruiken in de praktijk



Projectfasen beeldherkenning collecties

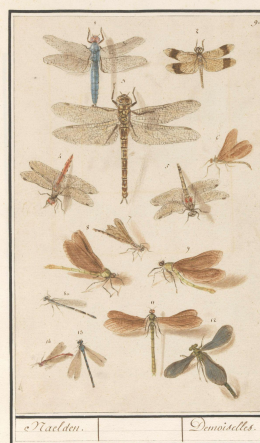
Model trainen in 7 stappen

1. Identificeer een soortgroep met potentiële kansen, bijv. Conidae of Papilionoidea
2. Afbeeldingen verzamelen van geïdentificeerde exemplaren
3. Data verzamelen en uniformeren
4. Modellen trainen
5. Modellen toepassen = identificeren van niet-geïdentificeerde exemplaren
6. Modellen evalueren
7. Modellen en/of workflow verbeteren



Collecties selecteren

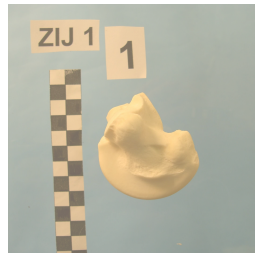
- beschikbaar gedetermineerd en gefotografeerd materiaal
- type object geschikt voor beeldherkenning
- vraaggericht: een toepassing in collectieregistratie of voor publiek
- Maskers Museon en NMVW
- Tekeningen Anselmus de Boodt Rijksmuseum
- Papillotten-collectie Naturalis
- Conidae (mollusca) NMR en Naturalis
- 'Nederlandse' Vogeleieren Museon en Naturalis
- Eikapsels roggen en haaien Naturalis
- Mineralen o.v.b.



Rijks Museum van Natuurlijke Historie-Leiden
4. *Limantopus limantopus*
Animal: *Limantopus*
Dat: Juni 1935
Loc: Wijkse Wormen
Coll: J. Oringer
Ingeheimd: 29-7-1935 Register no 6830

... en wat niet:

- niet voldoende *gedetermineerd* en *gefotografeerd* materiaal
 - bv. te veel variatie binnen het type object en te weinig beelden
 - andere pilots of projecten
 - andere oplossingen, zoals OCR en HTR
-
- Botten, schelpen, mineralen etc. - publieksapp wenselijk
 - Laden met opgeprikte insecten ARISE
 - Herbariumvellen: *Half-Earth Challenge* 2.5 m images/ 65.000 species
 - Fossielen: *LegaSea* NWO-project met AI-toepassing & *citizen-science* voor fossielen uit het Noordzeebekken



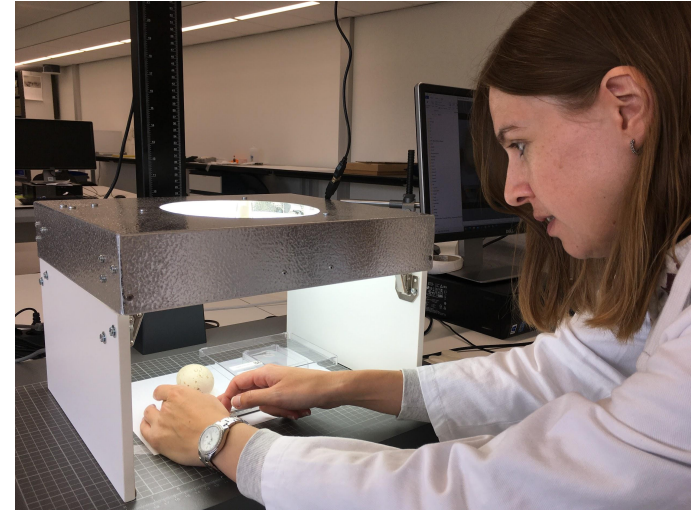
Projectfase

- gedetermineerde specimens selecteren



Projectfase

- collectie digitaliseren
- medewerkers fotograferen vogeleierencollectie.



Voor een beter
herkenningsmodel worden de
eieren tegen een witte én een
zwarte achtergrond
gefotografeerd

Projectfase

- **collectie digitaliseren**
vrijwilligers fotograferen 'papillotten', een grote verzameling gevouwen vlinders.

er zijn al 130.000
dagvlinders
(Papilionoidea)
gedigitaliseerd

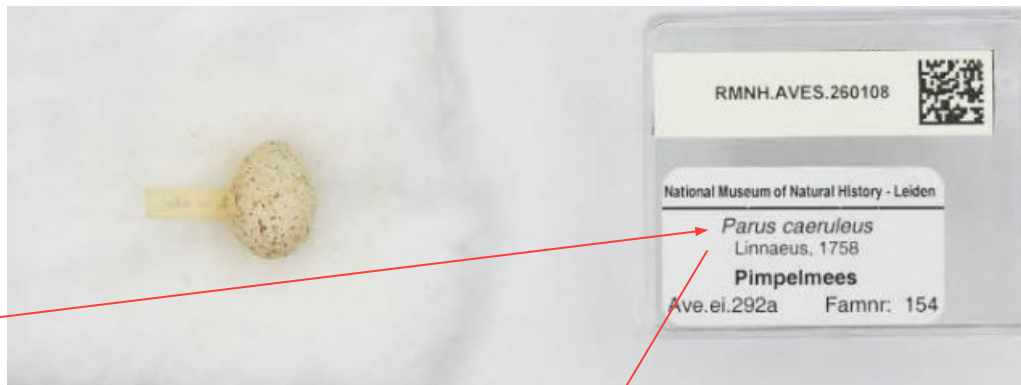


Projectfase

- soortnamen normaliseren

De 'referentiecollectie Nederlandse Vogeleieren' bestaat uit gedetermineerde eieren van vogels die in Nederland voorkomen.

Verouderde soortnamen uit het collectiebeheersysteem > matchen met GBIF en CoL voor het herkenningmodel.



Naamgeving

Nederlandse naam	Pimpelmees
Geldige naam	<i>Cyanistes caeruleus</i> (Linnaeus, 1758)
Synoniem	<i>Parus caeruleus</i> Linnaeus, 1758

Projectfase

- soortnamen normaliseren

⇒ spellingsvarianten, verouderde namen, variaties in taxonauteur en -jaar
via **Open Refine** > data analyseren, schrijfwijzen clusteren en opschonen

choices/0/v	choices/1/v	choices/2/v	choices/3/v
§ Papilio (Menelaides) aegaeus cf. ormenus Guérin-Ménéville, 1829	1 Papilio (Menelaides) aegaeus ormenus (Guérin-Ménéville, 1829)	1 Papilio (Menelaides) aegaeus ormenus Guérin-Ménéville, 1829	1 Papilio (Menelaides) aegaeus ormenus Gu
§ Graphium (Graphium) agamemnon agamemnon (Linnaeus, 1758)	1 Graphium (Graphium) agamemnon agamemnon Linnaeus, 1758	1 Graphium (Graphium) agamemnon cf. agamemnon Linnaeus, 1758	1 Graphium agamemnon agamemnon (Lin
§ Graphium (Graphium) evemon evemon (Boisduval, 1836)	1 Graphium (Graphium) evemon evemon Boisduval, 1836	1 Graphium cf. evemon (Boisduval, 1836)	1 Graphium evemon evemon (Boisduval, 1
§ Zeuxidia luxeri Hübner, 1820	1 Zeuxidia luxeri Hübner, 1820/26	1 Zeuxidia luxeri luxeri Hübner, 1820/26	1
§ Delias hyparete (Linnaeus, 1758)	1 Delias hyparete Linnaeus, 1758	1 Delias hyparete hyparete Linnaeus, 1758	1
§ Graphium (Graphium) bathycles bathycles (Zinken, 1831)	1 Graphium (Graphium) bathycles bathycles Zinken, 1831	1 Graphium bathycles bathycles (Zinken, 1831)	1
§ Graphium (Graphium) evemon evemon (Boisduval, 1836)	1 Graphium (Graphium) evemon evemon Boisduval, 1836	1 Graphium evemon evemon (Boisduval, 1836)	1
§ Troides (Troides) cf. helena (Linnaeus, 1758)	1 Troides (Troides) helena helena (Linnaeus, 1758)	1 Troides helena helena (Linnaeus, 1758)	1
§ Papilio (Menelaides) aegaeus ormenus (Guérin-Ménéville, 1829)	1 Papilio (Menelaides) aegaeus ormenus Guérin-Ménéville, 1829	1 Papilio (Menelaides) aegaeus ormenus Guérin-Ménéville, 1831	1
§ Parantica aspasia (Fabricius, 1787)	1 Parantica aspasia Fabricius, 1787	1 Parantica aspasia aspasia (Fabricius, 1787)	1
§ Graphium (Graphi	Drupadia ravindra Horsfield, 1828	Drupadia ravindra Horsfield, 1829	
§ Graphium (Graphi	Vindula arsinoe (Cramer, 1777)	Vindula arsinoe (Cramer, 1779)	
§ Papilio (Menelaide	Papilio (Menelaides) polytes timorensis Felder, 1864	Papilio (Menelaides) polytes timorensis Felder, 1866	
§ Papilio (Menelaide	Graphium (Pathysa) antiphates javanicus (Eimer, 1888)	Graphium (Pathysa) antiphates javanicus (Eimer, 1889)	
§ Graphium (Pathysa	Papilio (Menelaides) polytes javanus Felder, 1862	Papilio (Menelaides) polytes javanus Felder, 1864	
§ Atrophaneura (Atr			
§ Atrophaneura (Atr			
§ Graphium (Graphi			

Projectfase

- Data normaliseren

Grote variëteit in schrijfwijzen in collectie management systeem

⇒ typefouten, spellingsvarianten, verouderde namen, variaties in auteur en jaartal (met en zonder haakjes)

} normaliseren

⇒ **name resolver** als tijdelijke oplossing



Projectfase

- Data normaliseren

Grote variëteit in schrijfwijzen in collectie management systeem

⇒ typefouten, spellingsvarianten, verouderde namen, variaties in auteur en jaartal (met en zonder haakjes)

} normaliseren

⇒ **name resolver** als tijdelijke oplossing

⇒ **duurzaam** standaardiseren

via **thesauri** en/of interne/externe **identifiers**



Projectfase

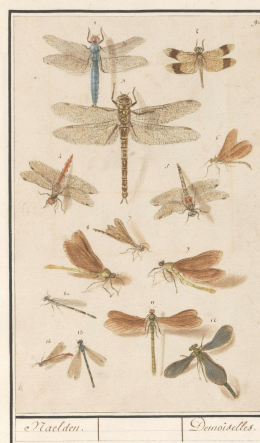
- Data normaliseren
- Model **Conidae**
 - getraind met 5 collecties, 4 museale en 1 particuliere collectie
 - heeft een support van 15.877 afbeeldingen om 797 soorten te herkennen
 - De naamgeving van de soorten uit de 5 collecties is geüniformeerd m.b.v. GBIF en WoRMS



5 Modellen

6 collecties

- Maskers *Presentatie Corine Blik* Museon
- Papillotten *Presentatie Mónica Guimarães Cruz*
- Conidae (mollusca) NMR en Naturalis
- 'Nederlandse' Vogeleieren Museon en Naturalis
- Roggen- en haaieneieren Naturalis
- Tekeningen Anselmus de Boodt Rijksmuseum



Rijks Museum van Natuurlijke Historie - Leiden
4. *Limantopus limantopus*
Antal: 1 - tegel.
Dat: Jun 1935
Loc: Wijk Wormer
Coll: J. Oringer
Ingeleverd: 29-7-1935 Register no 6830

ISA Campagne Anselmus de Boodt

Historia Naturalis - albums voor keizer Rudolf II



taggen van alle afgebeelde soorten planten, dieren, schelpen en paddestoelen op tekeningen, gemaakt en verzameld tussen 1596 en 1610 door Anselmus de Boodt.

Tag the species



Nederlandsche vogelen van Nozeman en Sepp

Compendium (ca. 1800) with more than 2000 drawings of Dutch ('Nederlandsche') birds. The latin names of the birds are often mentioned in file name or Commons category.

Typen objecten

variatie verschilt per soort:
eieren *versus* vlinders



Model vogeleieren

- veel variatie binnen legfels van één soort

zwarte kraai
Corvus corone Linnaeus, 1758



Model vogeleieren

- veel variatie binnen legsels van één soort
- weinig verschil tussen soorten

Model ziet geen verschil in afmetingen



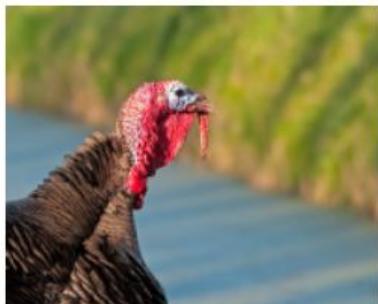
Columba oenas Linnaeus, 1758
RMNH.AVES.259806



Columba palumbus Linnaeus, 1758
RMNH.AVES.259819

Modellen evalueren

- Model Vogeleieren
 - ondersoorten niet als klasse onderscheiden
 - welke soorten worden goed herkend en welke niet en waarom niet?
 - aanvullende soortinformatie



Meleagris gallopavo
Linnaeus, 1758
Populaire naam: [Kalkoen](#)



Meleagris gallopavo
Linnaeus, 1758
ZMA.AVES.50144
Populaire naam: **Kalkoen**



Meleagris gallopavo
Linnaeus, 1758
ZMA.AVES.50144
Populaire naam: **Kalkoen**

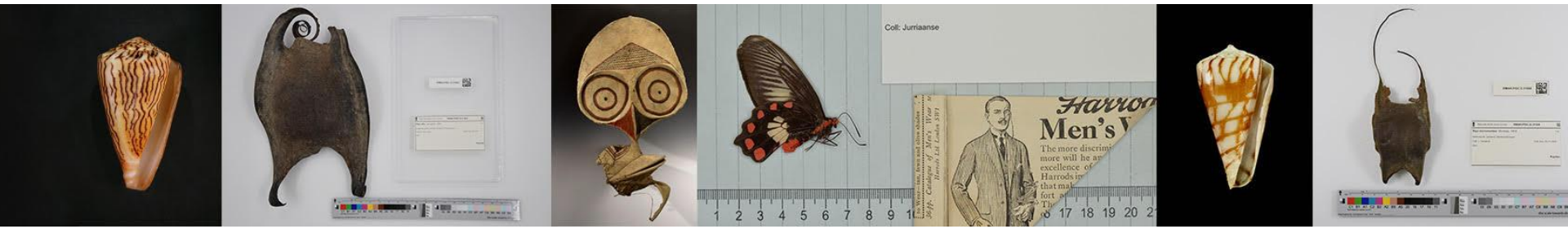


yes; *Meleagris gallopavo*
Linnaeus, 1758
ZMA.AVES.50108
Populaire naam: **Kalkoen**

User Interface: modellen

→ Modellen beschikbaar stellen aan musea en/of publiek via **interface**:

- geschikt voor losse uploads, voor het snel categoriseren van een specimen
- geschikt voor bulkinvoer door musea
- te gebruiken met smartphone door het publiek
- in te zetten in een werkproces
- door een *name resolver* in te zetten ⇒ gestandaardiseerde naamgeving
- link naar bijv. GBIF
- aanvullende soort-informatie bij de determinaties



Interface

<https://museum.identify.biodiversityanalysis.nl>



aanvullende soortinformatie bij de determinaties ter verificatie

 website **Malesian Butterflies**

Malesian Butterflies

MENU

Papilio gambrisius Cramer, 1777

Morphology

Pictures

Occurrence

Map of life

Biology

Collection Specimens

Taxonomy

Search taxa

Abisara abnormis

MORPHOLOGY

External characteristics

Male

Forewing length (mm): 76. Wingspan (mm): 134.

Forewings: Upperside black with four pale yellowish subapical spots; underside streaks between the veins in the apical area. **Hindwings** crenulate, no tail; upper central pale dirty yellow area, stretching from inner margin to front margin, when base, outer edge of pale area toothed along veins; underside black, red spot ab

 website **Nederlands Soortenregister**



Kleine karekiet *Acrocephalus scirpaceus*

Foto: Hans van der Meulen

→ Samenvatting

→ Naamgeving

→ Herkenning

→ Foto's

→ Bedreiging en bescherming

→ Verspreiding

→ Biotopen

→ [Levenswijze](#)

→ Literatuur

→ Summary

Voedsel

Insecten, in najaar ook bessen.

Eieren

Aantal eieren in legsel meestal 4, zelden 3-6. Buikig. Glad en glanzend. Groenachtig-wit of zeer lichtgroen. Meestal sterk groen, olijfkleurig of dof grijs gespikkeld en gevlekt, vaak geconcentreerd aan of rond de stompe pool en vormt een zware kap; als de tekens fijner zijn schijnen ze dichter en gelijkmatig verspreid. Formaat 18,4 - 13,6 mm.

Publicatie

Harrison, C. & Taapken, J. 1977. Elseviers broedvogelgids: nesten, eieren en jongen van alle in Europa, Noord-Afrika en het Midden-Oosten broedende vogels. Elsevier.

Project beeldherkenning

→ *Lessons Learned*

- gestandaardiseerde data, via thesauri en/of identifiers
- niet alle beeldmateriaal is geschikt
- niet alle typen objecten zijn geschikt voor automatische beeldherkenning
- sommige soorten/typen behoeven meer data dan andere soorten/typen
- aanvullende informatie voor determinatie soms noodzakelijk
- het gekozen aanzicht is niet kenmerkend genoeg
- er gebeurt veel op dit gebied: kijk wat er elders gebeurt en wissel kennis uit!
- AI is een hulpmiddel; misschien is automatische beeldherkenning niet altijd de beste oplossing...



User interface: modellen gebruiken

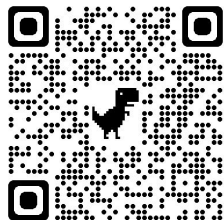
upload::dropzone

upload::button ☐ suppress_enrichments



Resultaten

1. *Raja brachyura* Lafont, 1871
0.6293196678161621
[GBIF](#)
2. *Leucoraja naevus* (Müller & Henle, 1841) ⓘ
0.1105806902050972
[meer informatie \[Naturalis\]](#) [GBIF](#)
3. *Raja clavata* Linnaeus, 1758 ⓘ
0.06981516629457474
[meer informatie \[Naturalis\]](#) [GBIF](#)



User interface: modellen gebruiken

Sleep je hier je afbeelding naartoe (max. 10 mb)

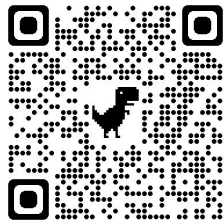
Upload afbeelding

☐ Geen aanvullende soort-/objectinformatie bij resultaten tonen



Resultaten

1. *Conus leopardus* (Röding, 1798)
0.9750891327857971
GBIF
2. *Conus betulinus* Linnaeus, 1758
0.01832430809736252
GBIF
3. *Conasprella lizarum* (Raybaudi Massilia & da Motta, 1992)
0.0014080095570534468
GBIF



Bedankt voor de aandacht!

contact: annika.hendriksen@naturalis.nl

Project Automatische beeldherkenning als instrument voor museumcollecties



colofon

Afbeeldingen collectie-specimens: CC-0

Afbeeldingen medewerkers /collectie Naturalis: CC-0

Afbeeldingen De Boodt (coll. Rijksmuseum): Publiek Domein (PD)

Afbeelding masker (coll. Museon): <https://cc.museon.nl?queryid=5d9f63b3-b1ef-4120-ac81-d1e0a7006b16>

Afbeelding Conus betulinis, Iconographia zoologica (coll. UvA): Publiek Domein (PD)

Automatische beeldherkenning als instrument voor museumcollecties