

Wereldwijd meest complete online archief

# Natuurhistorie Indonesië online

Onderzoekers in het project 'Making Sense of Illustrated Handwritten Archives' maken een complex negentiende-eeuws onderzoeksarchief toegankelijk met behulp van kunstmatige intelligentie en semantische technologie.

Maarten Heerlien

Tussen 1820 en 1850 bracht de Natuurkundige Commissie voor Nederlands-Indië de flora, fauna en natuurlijke hulpbronnen van de Indonesische archipel in kaart. De onderzoekers stuurden duizenden zoölogische en botanische objecten terug naar Nederland en produceerden een omvangrijk corpus van aantekeningen en publicaties. Naturalis Biodiversity Center in Leiden, beheerder van de collectie, liet het geschreven corpus tussen 2008 en 2011 digitaliseren. Toch bleef het vanwege de complexiteit van de handschriften grotendeels ontoegankelijk.

## Waardevolle inzicht

"Zonde," zegt Andreas Weber, universitair docent Science, Technology and Policy Studies (STePS) aan de Universiteit Twente, "de veld-



Pagina uit een bundel veldaantekeningen met een beschrijving en afbeelding van de vleermuisensoort *Rhinolophus larvatus*.

Credits: Naturalis Biodiversity Center, archief van de Natuurkundige Commissie voor Nederlands-Indië. Rechten: Public Domain Mark 1.0.

aantekeningen en dagboeken geven een waardevolle inzicht in de dagelijkse praktijk van historisch natuuronderzoek. Bovendien vormt de collectie van de Natuurkundige Commissie waarschijnlijk wereldwijd het meest complete archief van de negentiende-eeuwse biodiversiteit van Indonesië, een belangrijke biodiversiteitshotspot."

Weber is als PostDoc verbonden aan 'Making Sense of Illustrated Handwritten Archives'.

In dit NWO Creatieve Industrie-project, met cofinanciering van Brill Scholarly Publishing, maken Naturalis, de universiteiten van Leiden, Groningen en Twente en Brill de handschriften van de Natuurkundige Commissie doorzoekbaar en koppelbaar aan de verzamelde objecten. Dit doen ze met behulp van automatische handschriftherkenning en semantische annotatie. Een uitdaging volgens Weber: "Dergelijke natuurhistorische archieven zijn

complexe visuele landschappen waar gerelateerde informatie op verschillende plekken en in verschillende vormen en talen te vinden is. Bijvoorbeeld als tekeningetjes, uiteenlopende notatiewijzen, enzovoort." Als oplossing wordt MONK gebruikt, een geavanceerde ICT-architectuur voor automatische handschriftherkenning, ontwikkeld door het team van Lambert Schomaker, hoogleraar Kunstmatige Intelligentie aan de Rijksuniversiteit Groningen. Weber: "MONK zoekt naar visuele patronen in de vorm van woorden en andere beeldelementen. Door bovendien een betekenis aan een woordbeeld te koppelen, leert het systeem woordklassen te identificeren. Daar kan het systeem bij vergelijkbare nieuwe manuscripten weer gebruik van maken."

## Koppeling

De koppeling tussen collectieobjecten en gerelateerde informatie in de handschriften wordt verder gefaciliteerd door de Semantic Field Book Annotator, een tool waarvoor Lise Stork, PhD bij het Leiden Institute of Advanced Computer Science (LIACS) in oktober 2019 de Young eScientist Award ontving. Eind 2020 lanceert Brill het platform Natuurkundige Commissie Online, waarin het archief open doorzoekbaar zal zijn gemaakt. Een afsluitend congres vindt plaats van 22 tot en met 24 november bij Naturalis.

[makinenseproject.org](http://makinenseproject.org)

Machine Learning toegepast op Citizen Science-data

# GeoHealth-onderzoek zet data op de kaart

Hoe groot is het risico op een tekenbeet? En wat is de impact van extreme weergebeurtenissen? In dit artikel aandacht voor het GeoHealth-onderzoek van jong talent Irene Garcia-Martí. *Marion Wittenberg*

## JONG TALENT

### 'Burgerobservaties combineren met omgevingsvariabelen'

Elk jaar worden meer dan 1 miljoen Nederlanders gebeten door een teek, van dit aantal worden ongeveer 27.000 mensen besmet met de ziekte van Lyme. Bij 1.000 tot 2.500 mensen leidt dit tot ernstige en langdurig klachten. Data scientist Irene Garcia-Martí promoveerde 27 september jongstleden bij de Universiteit van Twente op het proefschrift 'Mapping tick dynamics and tick bite risk using data-driven approaches and volunteered observations'.

## Voorspellen

Garcia-Martí combineerde burgerobservaties met een breed scala aan omgevingsvariabelen (weergegevens, remote sensing, officiële geo-data) en ontwikkelde door middel van machine learning, modellen die tekenactiviteit en het risico op tekenbeten op nationaal niveau kunnen voorspellen. De burgerobservaties zijn afkomstig van de Natuurkalender ([naturetoday.com](http://naturetoday.com)), een fenologisch waarnemingsnetwerk, waaraan burgers al sinds 1868 waarnemingen in de natuur doorgeven. In 2006

startte de Natuurkalender met het registreren van tekenbeten. Door de grote hoeveelheid aanmeldingen is in 2012 voor deze registratie een aparte website gelanceerd: [tekenradar.nl](http://tekenradar.nl). Daarnaast hebben tussen 2006 en 2016 vrijwilligers teken gevangen op 12 locaties in Nederland, gecoördineerd door onderzoekers van de Universiteit van Wageningen. Momenteel zijn Garcia-Martí's modellen nog statisch, een vervolgstap is om ze dynamisch te maken, waarmee fluctuaties in het tekenrisico door het jaar heen te voorspellen zijn.

## Blogposts over onderzoek

Inmiddels werkt Garcia-Martí bij het Koninklijk Nederlands Meteorologisch Instituut (KNMI), waar ze onderzoek doet naar de impact van extreme weergebeurtenissen. Ook voor haar huidige werk gebruikt ze burgerobservaties verzameld via de weerstations van



Na haar promotie zijn de gegevens en code van Garcia-Martí gearchiveerd in EASY zodat ook anderen kunnen bestuderen hoe de modellen zijn ontwikkeld. Het was de 100.000ste dataset in EASY (DOI: [10.17026/dans-zre-tggd](https://doi.org/10.17026/dans-zre-tggd)). Om dit te vieren, stuurde DANS haar een taart. Op de foto van links naar rechts: Marga Koelen (research support coördinator), Tom Veldkamp (decaan ITC Faculteit), beide van de Universiteit Twente en Irene Garcia-Martí. Credits: John Horn

het Weather Observations netwerk WOW-NL ([wow.knmi.nl](http://wow.knmi.nl)). Garcia-Martí en haar collega's onderzoeken hoe de data van WOW-NL kunnen helpen bij het bestuderen van lokale fenomenen zoals stedelijke hitte-eilanden. Belangrijk, omdat het juist de steden zijn die bij de huidige wereldwijde temperatuurstij-

ging warmer worden, wat gezondheidsproblemen van de bevolking met zich meebrengt. Irene Garcia-Martí houdt een blog bij over haar onderzoek op [irenegarciamarti.com](http://irenegarciamarti.com). Haar proefschrift is beschikbaar via: [library.itc.utwente.nl/papers\\_2019/phd/garciamarti.pdf](http://library.itc.utwente.nl/papers_2019/phd/garciamarti.pdf)