

Google Scholar zoekresultaten voor wetenschappelijke projecten: linked data & natural language processing

Bart De Paepe
16 juni 2025

**HO
GENT**

Inhoud.

Kader.

Proof of
Concept.

Resultaten.

Discussie.

Conclusie.

**HO
GENT**

Kader.

**HO
GENT**

Vlaams Instituut voor de Zee



<https://www.vliz.be>

**HO
GENT**

Integrated Marine Information System

[Home](#) > [IMIS](#)

IMIS

[log in](#)

[Data Policy](#)

[Publicaties](#) | [Instituten](#) | [Personen](#) | [Datasets](#) | [Projecten](#) | [Kaarten](#) | [Evenementen](#)

IMIS - the Integrated Marine Information System - biedt informatie over datasets, publicaties, personen en instellingen die gerelateerd zijn aan mariene wetenschappen uitgevoerd in Vlaanderen en binnen projecten waaraan Vlaamse instituten deelnemen.

Special collections worden gedefinieerd binnen IMIS, zodat het systeem kan worden ingezet binnen nationale en internationale projecten, evenals binnen instituten, ter ondersteuning van hun eigen informatiemanagement voor datasets, publicaties, personen en instituten.

Alle gegevens in IMIS vallen onder het [VLIZ Privacybeleid](#).

Klik op de bovenstaande links om naar de zoekpagina's van de respectieve IMIS catalogues (datasets, publicaties, enz.) te gaan

<https://www.vliz.be/nl/imis>

**HO
GENT**

Google Scholar

The screenshot shows a Google Scholar search results page. The search query is "Vlaams Instituut voor de Zee" OR "Vlaams Instituut van de Zee" OR "Flanders...". The results list several articles related to Simon Stevin. The first article is "Simon Stevin of Bruges (1548-1620)" by G Sarton, published in 1934 in the *Journal of the History of Science*. The second article is "Simon Stevin: science in the Netherlands around 1600" by EJ Dijksterhuis, published in 2012 in *Books.google.com*. The third article is "Magic is no magic: the wonderful world of Simon Stevin" by JT Devroese and GV Berghie, published in 2003 in *Books.google.com*. The fourth article is "Wonder en is gheen wonder: De geniale wereld van Simon Stevin 1548-1620" by JT Devroese and G Vanden Berghie, published in 2003 in *Documentatiecentrum wotlab.be*. The fifth article is "Stevin's kleinschalig stevig werkt door in de euro. Met zijn interestabellen is hij economist avant la lettre. Hij verbeterde de werking van molens en sluizen. Stevin was actief in navigatie. ..." by JT Devroese and G Vanden Berghie, published in 2003 in *Documentatiecentrum wotlab.be*. The "Melding maken" button is highlighted with a red box.

zoekopdracht

The screenshot shows an email inbox with a search bar at the top. The inbox is filtered by "Focused" and "Other" tabs. The "Today" section shows two emails from Google Scholar. The first email is titled "Google Scholar-meldingen" and contains the text "intext: 'Vlaams Instituut voor de Zee' OR intext: ... nieuwe resultaten" and "Unravelling climate-driven habitat suitability for pelagic fish in European seas", dated 1:19 AM. The second email is also titled "Google Scholar-meldingen" and contains the text "'Vlaams Instituut voor de Zee' OR 'Vlaams Instituut ...' nieuwe resultaten" and "Unravelling climate-driven habitat suitability for pelagic fish in European seas", dated 1:19 AM. The "Yesterday" section shows one email from Google titled "Je e-mailadres verifiëren", dated Tue 9:19 PM.

alert

The screenshot shows a Google Scholar notification email. The email is dated Wed 1/29/2025 1:19 AM. The subject is "Google Scholar-meldingen <scholar-meldingen@vli.vlaanderen.be>". The body of the email contains the text "Unravelling climate-driven habitat suitability for pelagic fish in European seas" and "Pelagic fish species hold immense ecological and economic significance in European seas, playing a vital role in the region's marine ecosystem, contributing significantly to the local economy. Understanding their habitat requirements is crucial for effective conservation and sustainable management." The email also includes a link to "Vlaams Instituut voor de Zee" and "Vlaams Instituut van de Zee".

HO
GENT

De onderzoeksvraag.

“Hoe kunnen de zoekresultaten van Google Scholar automatisch toegevoegd worden aan IMIS?”

- Hoe kunnen de uiteenlopende zoekresultaten van Google Scholar omgezet worden in gestructureerde data?
- Zijn alle zoekresultaten uniek identificeerbaar zodat er geen duplicaten opgeslagen worden?
- Hoe kunnen ook steeds nieuwe zoekresultaten van dezelfde zoekopdracht systematisch opgezocht worden?
- Kan er een score berekend worden hoe relevant elk zoekresultaat is ten opzichte van de zoekopdracht?
- Als er geen unieke identifier is, hoe wordt er dan gecontroleerd op duplicaten?
- Hoe blijft de tool overwegend onafhankelijk van third-party software?
- Hoe wordt er gegarandeerd dat de tool op de bestaande hardware kan draaien?

Proof of Concept.

Google Scholar alert

zoekopdracht
instellen

Web scraping

LLM

klassieke parser

Natural Language
Processing

relevantiescore

Linked data

DOI opzoeken

Semantic search

semantic search score

**HO
GENT**

Web scraping.

[HTML] The role of narcissistic personality traits in bullying behavior in adolescence—A systematic review and meta-analysis

L Schunck, U von Düring, J Wika - Personality and Individual Differences, 2025
This systematic review and meta-analysis examine the association between narcissistic personality **traits** in adolescents and their roles in the bullying process. The aim is to summarize the current state of research to gain insights into the ...



Exploring Nitrogen Use Efficiency in Cereals: Insight into Traits, Metabolism, and Management Strategies Under Climate Change Conditions—A Comprehensive ...

FZ Ben Dabbane, A Baldani, M Aarbaoui... - Journal of Soil Science and ..., 2025
Nitrogen (N) is an essential element for cereals growth and development, playing a crucial role in productivity and yield. Consequently, nitrogen fertilizers are extensively used in cereal cultivation. However, the excessive fertilizer application ...



[HTML] Task-evoked pupil responses during free-viewing of hierarchical figures in relation to autistic traits in adults

C Brittenham, AS DiCrisolo, V Tolani, Y Hu, JB Wagner - Scientific Reports, 2025
... with high levels of autistic **traits** would have more default reports of the local orientation. For pupillary responses, we hypothesized that the high **traits** group would show greater constriction in the early part of the pupil response as compared ...



Additive effects of freshwater salinization on the predator-induced traits of larval amphibians

B Tator, M Le Sage, RA Relyea - Environmental Pollution, 2025
... of salt on individual **traits**. Moreover, we know even less about how salt pollution may affect phenotypically plastic **traits** that have evolved in ... Interestingly, the predator effects did not interact with salt effects for any of the **traits**. These results ...



How Collaboration Context and Personality Traits Shape the Social Norms of Human-to-Avatar Identity Representation

S Kang, B Yoon, K Kim, J Gratch, W Woo - IEEE Transactions on Visualization and ..., 2025
... **traits** directly address impression management and strategic self-representation, both central to avatar customization. By targeting these two **traits**... Therefore, our study investigates how perspective, deviations from one's real-world appearance ...



[HTML] Pork-YOLO: Automated Collection of Pork Quality Traits

parse with LLM

messages=["role": "system", "content": "You are a master at scraping Google Scholar results data. Scrape top 10 organic results data from Google Scholar search result page.

,],

parse with scraper

<h3 style="font-weight:normal;margin:0;font-size:17px;line-height:20px;"> [HTML] The role of narcissistic personality traits in bullying behavior in adolescence—A systematic review and meta-analysis </h3>

<div class="gse_alrt_sni" style="border:1px solid black; height:17px">The Ocean

Race, known as one of the world most challenging round-the-world
 sailing competitions, has evolved into a significant platform for ocean science,
 advocacy, and innovation. Through its Science Program, The Race leverages its...</div>

HO
GENT

Natural language processing.

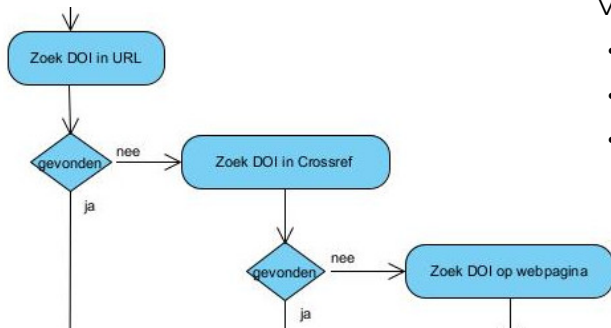
- “Het VLIZ verwelkomt nieuwsgierige particulieren en kleine groepen voor een bezoek aan haar Campus en haar Mariene Station Oostende.”
- 1. Coreference resolution
 2. Stopwoorden
 3. Lemmatisering
- “VLIZ verwelkom nieuwsgierig particulier klein groep bezoek VLIZ Campus VLIZ Marien Station Oostende.”

$$\text{relevantiescore} = \text{frequentie} * \log \left(\frac{\text{totaal aantal zoekresultaten}}{\text{aantal zoekresultaten waar term voorkomt}} \right)$$

**HO
GENT**

Linked data.

<https://pubs.acs.org/doi/full/10.1021/acsami.4c21991>

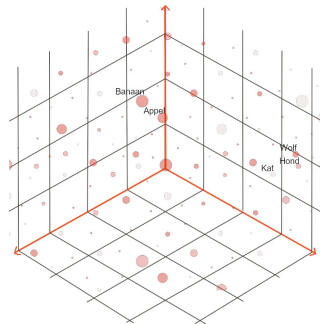
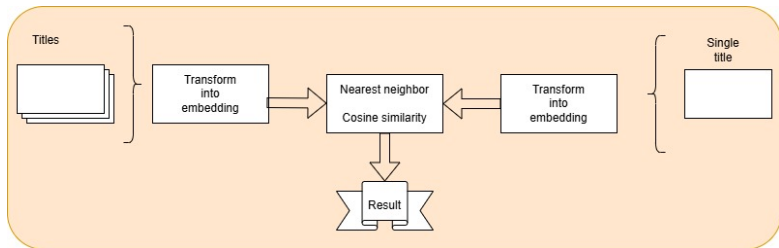


verschillende types webpagina's

- een gewone HTML pagina
- een PDF document
- een webpagina met een embedded PDF document

**HO
GENT**

Semantic search.

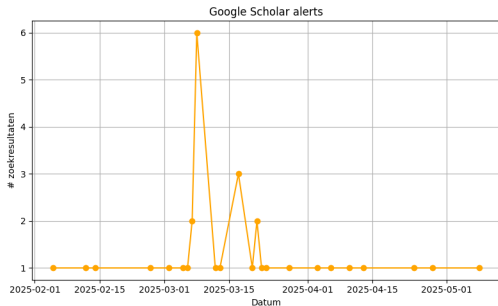


**HO
GENT**

Resultaten.

**HO
GENT**

Resultaten (1/4).



Zoekresultaten voor de
zoekopdracht “VLIZ”

- 23 alerts
- 1-6 zoekresultaten per alert
- 18% van de zoekresultaten zijn volledig

BULLETIN OF THE BELGIAN MATHEMATICAL SOCIETY SIMON

TDM JOSUAT-VERGUÉS, FCMUE MARTÍNEZ...

T. DOUVROPOULOS–M. JOSUAT-VERGUÉS. Cluster parking functions. 1 F.
CRESPO–M. URIBE–E. MARTÍNEZ. Melnikov method for perturbed completely
integrable systems. 45 S. GARTI–S. SHELAKH. Poésie Aristotélienne. 67 WF XUAN–D .



**HO
GENT**

Resultaten (2/4).

OpenAI	Anthropic	Llama3	Beautiful Soup
The role of narcissistic personality traits in bullying behavior in adolescence-A systematic review and meta-analysis	The role of narcissistic personality traits in bullying behavior in adolescence-A systematic review and meta-analysis	-	The role of narcissistic personality traits in bullying behavior in adolescence-A systematic review and meta-analysis
Exploring Nitrogen Use Efficiency in Cereals: Insight into Traits, Metabolism, and Management Strategies Under Climate Change Conditions-A Comprehensive Review	Exploring Nitrogen Use Efficiency in Cereals: Insight into Traits, Metabolism, and Management Strategies Under Climate Change Conditions-A Comprehensive Review	-	Exploring Nitrogen Use Efficiency in Cereals: Insight into Traits, Metabolism, and Management Strategies Under Climate Change Conditions-A Comprehensive Review
Task-evoked pupil responses during free-viewing of hierarchical figures in relation to autistic traits in adults	Task-evoked pupil responses during free-viewing of hierarchical figures in relation to autistic traits in adults	-	Task-evoked pupil responses during free-viewing of hierarchical figures in relation to autistic traits in adults
Additive effects of freshwater salinization on the predator-induced traits of larval amphibians	Additive effects of freshwater salinization on the predator-induced traits of larval amphibians	-	Additive effects of freshwater salinization on the predator-induced traits of larval amphibians
How Collaboration Context and Personality Traits Shape the Social Norms of Human-to-Avatar Identity Representation	How Collaboration Context and Personality Traits Shape the Social Norms of Human-to-Avatar Identity Representation	-	How Collaboration Context and Personality Traits Shape the Social Norms of Human-to-Avatar Identity Representation
Pork-YOLO: Automated Collection of Pork Quality Traits	Pork-YOLO: Automated Collection of Pork Quality Traits	-	Pork-YOLO: Automated Collection of Pork Quality Traits
Inter-and intraspecific study on compositional quality traits of diverse sets of spelt and common wheat	Inter-and intraspecific study on compositional quality traits of diverse sets of spelt and common wheat	-	Inter-and intraspecific study on compositional quality traits of diverse sets of spelt and common wheat
Embryo traits and their taxonomic implications in stipoid grasses (Poaceae: Stipeae) in Asia	Embryo traits and their taxonomic implications in stipoid grasses (Poaceae: Stipeae) in Asia	-	Embryo traits and their taxonomic implications in stipoid grasses (Poaceae: Stipeae) in Asia
Uncovering bighorn sheep life-history trajectories in multidimensional trait space	Uncovering bighorn sheep life-history trajectories in multidimensional trait space	-	Uncovering bighorn sheep life-history trajectories in multidimensional trait space
Multigenerational exposure to glyphosate has only modest effects on life history traits, stress tolerance, and microbiome in a field cricket	Multigenerational exposure to glyphosate has only modest effects on life history traits, stress tolerance, and microbiome in a field cricket	-	Multigenerational exposure to glyphosate has only modest effects on life history traits, stress tolerance, and microbiome in a field cricket

**HO
GENT**

Resultaten (3/4).

relevantiescore



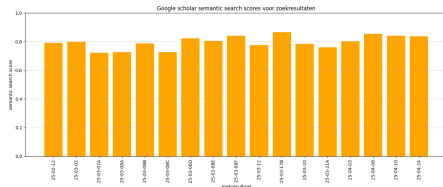
DOI

	VLIZ dataset	IMIS dataset
Percentage aantal keer dat de DOI gevonden is.	48%	62,7%
● Percentage aantal keer dat de DOI in de URL staat.	7%	
● Percentage aantal keer dat de DOI via Crossref gevonden wordt.	20%	
● Percentage aantal keer dat de DOI op webpagina staat.	73%	
Referentiepercentage DOI		61%

**HO
GENT**

Resultaten (4/4).

semantic search score



Titel	Semantic search score
Genome-wide association study revealed candidate genes associated with egg-laying time traits in layer chicken	1.0
Genome wide association study revealed candidate genes associated with egg laying time traits in layer chicken	0.99581
Gen. wide assoc. study revealed candidate genes associated with egg time traits in layer chicken	0.970841
Genome-wide association study revealed candidate genes associated with egg-laying time traits	0.954331
...	0.9
Genome-wide association study revealed some new candidate genes associated with flowering and maturity time of soybean in Central and West Siberian regions of Russia	0.81463
Genome- and Exome-Wide Association Studies Revealed Candidate Genes Associated with DaTscan Imaging Features	0.79323
...	

**HO
GENT**

Discussie.

**HO
GENT**

Discussie.

- Het automatisch opzoeken van de DOI is de grootste meerwaarde van de tool.
- Gebruikers kunnen focussen op publicaties zonder DOI en/of met een lage relevantiescore.
- 1500-2000 publicaties per jaar: 5-6 maanden
- Met de tool potentiële verlaging tot 2-3 maanden

Conclusie.

**HO
GENT**

Conclusie.

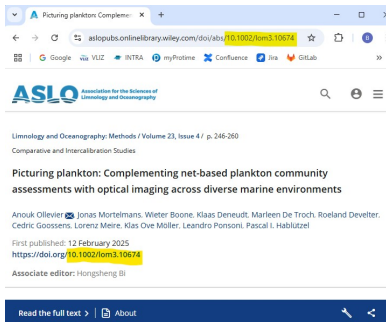
- De tool zal gebruikt worden binnen het VLIZ.
- Korte termijn: Verbetering van de tool door meer gelinkte data (auteurs, tijdschriften) op te zoeken in IMIS.
- Lange termijn: Uitvoerige tests met LLMs (zijn niet onderhevig aan de structuur van de HTML)

Woord van dank.

- Bart Vanhoorne
- Milan Lamote
- Jan Claes
- Fons Verheyde
- Cedric Decruw
- Lies Knockaert

**HO
GENT**

Linked data.



Picturing plankton: Complementing net-based plankton community assessments with optical imaging across diverse marine environments

Limnology and Oceanography: Methods / Volume 23, Issue 4 / p. 246-260
Comparative and Inter calibration Studies

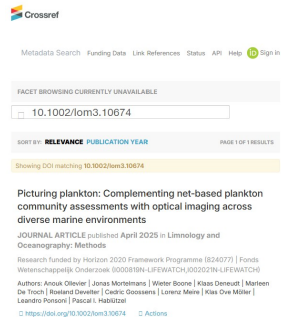
Anouk Ollevier, Jonas Mortelmans, Wieter Boone, Klaas Deneudt, Marleen De Troch, Roeland Develter, Cedric Goossens, Lorenz Meire, Klas Ove Möller, Leandro Ponsoni, Pascal I. Hablützel

First published: 12 February 2025
<https://doi.org/10.1002/lom3.10674>

Associate editor: Hongsheng Bi

Read the full text > | About

DOI in de URL & op de webpagina



Crossref

Metadata Search Funding Data Link References Status API Help Sign in

FACET BROWSING CURRENTLY UNAVAILABLE

10.1002/lom3.10674

SORT BY: RELEVANCE PUBLICATION YEAR PAGE 1 OF 1 RESULTS

Showing DOI matching 10.1002/lom3.10674

Picturing plankton: Complementing net-based plankton community assessments with optical imaging across diverse marine environments

JOURNAL ARTICLE published April 2025 in Limnology and Oceanography: Methods

Research funded by Horizon 2020 Framework Programme (824077) | Fonds Wetenschappelijk Onderzoek (000819N-LIFEWATCH, 0002021N-LIFEWATCH)

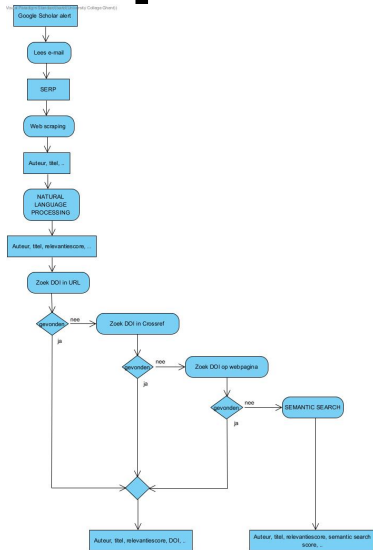
Authors: Anouk Ollevier | Jonas Mortelmans | Wieter Boone | Klaas Deneudt | Marleen De Troch | Roeland Develter | Cedric Goossens | Lorenz Meire | Klas Ove Möller | Leandro Ponsoni | Pascal I. Hablützel

<https://doi.org/10.1002/lom3.10674> Actions

DOI in Crossref

**HO
GENT**

Proof of Concept



**HO
GENT**