

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

**Bart De Paepe  
12 juni 2025**



# Inhoud.

Kader.

Proof of  
Concept.

Resultaten.

Discussie.

Conclusie.

**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>



# Google Scholar

"Vlaams Instituut voor de Zee" OR "Vlaams Instituut van de Zee" OR "Flanders Marine Institute"

Artikelen  
Gegeven 35.880 resultaten (0,09 sec)

Simon Stevin of Bruges (1548-1620)  
G Sarton - 1934 - journals.uchicago.edu  
... The Flemish mathematician SIMON STEVIN was perhaps the most original man of science of the second half of the sixteenth century. I say "perhaps" only because of his contemporary ...  
☆ Opstellen 99 Citeren Geciteerd door 58 Verwante artikelen Alle 6 versies

jacob Simon Stevin: science in the Netherlands around 1600  
EJ Dijksterhuis - 2012 - books.google.com  
The works of Simon Stevin are most interesting for the history of science, because they have such a wide scope and reflect so clearly the development of scientific knowledge around ...  
☆ Opstellen 99 Citeren Geciteerd door 108 Verwante artikelen Alle 4 versies

jacob 'Magic is no magic': the wonderful world of Simon Stevin  
JT Devreese, GV Berghe - 2008 - books.google.com  
This book gives a comprehensive picture of the activities and the creative heritage of Simon Stevin, who made outstanding contributions to various fields of science, in particular physics ...  
☆ Opstellen 99 Citeren Geciteerd door 81 Verwante artikelen Alle 5 versies

jacob 'Wonder en is gheen wonder': De geniale wereld van Simon Stevin 1548-1620  
JT Devreese, G Vanden Berghe - 2003 - documentatiecentrum.watlab.be  
Stevin's handig stalen werklijf door in de euro. Met zijn interesseballen is hij economist, avonturier, ...  
☆ Opstellen 99 Citeren Geciteerd door 7 Verwante artikelen Alle 2 versies 16

Melding maken

zoekopdracht

6/26

Search All Mail Items

Focused Other By Date ↑

Today

Google Scholar-meldingen  
intext:"Vlaams Instituut voor de Zee" OR intext:"... nieuwe resultaten"  
Unravelling climate-driven habitat suitability for pelagic fish in European seas 1:19 AM

Google Scholar-meldingen  
"Vlaams Instituut voor de Zee" OR "Vlaams Instituut ... nieuwe resultaten"  
Unravelling climate-driven habitat suitability for pelagic fish in European seas 1:19 AM

Yesterday

Google  
Je e-mailadres verifiëren

Tue 9:19 PM

Currently displaying all messages newer than 12 months.

GS

Google Scholar-meldingen

To: Google Scholar

If there are problems with how this message is displayed, click here to Click here to download pictures. To help protect your privacy, Outloo...

Unravelling climate-driven habitat suitability for pelagic fish in European seas, playing a vital role in the region's marine ecosystem, contributing significantly to the local economy. Understanding this...

"Vlaams Instituut voor de Zee" OR "Vlaams Instituut van de Zee" OR "Instituut" OR "VLIZ" OR "Simon Stevin" ...: nieuwe resultaten

Google Scholar heeft dit bericht gestuurd omdat je nieuwe resultaten voor "Vlaams Instituut voor de Zee" OR "Vlaams Instituut van de Zee" OR "Instituut" OR "VLIZ" OR "Simon Stevin" ...: nieuwe resultaten ontvangt. [Meer]

MELDING ANNULEREN

alert

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
    - ✗ online LLM
    - ✗ lokaal LLM
    - ✓ BeautifulSoup
    - ✗ Serpapi
  - Natural Language Processing
    - ✓ relevanciescore
  - Linked data
    - ✓ DOI in URL
    - ✓ DOI in Crossref
    - ✓ DOI op webpagina
  - Semantic search
    - ✓ semantic search score

# Web scraping (1/2).

**HTML The role of narcissistic personality traits in bullying behavior in adolescence-A systematic review and meta-analysis**  
L Schunk, U von Düring, J Wilke - *Personality and Individual Differences*, 2025  
This systematic review and meta-analysis examine the association between narcissistic personality traits 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 ...



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**



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 trait larval amphibians**  
B Tabor, M Le Sage, RA Rehage - 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



## 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 ...



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.  
, ],

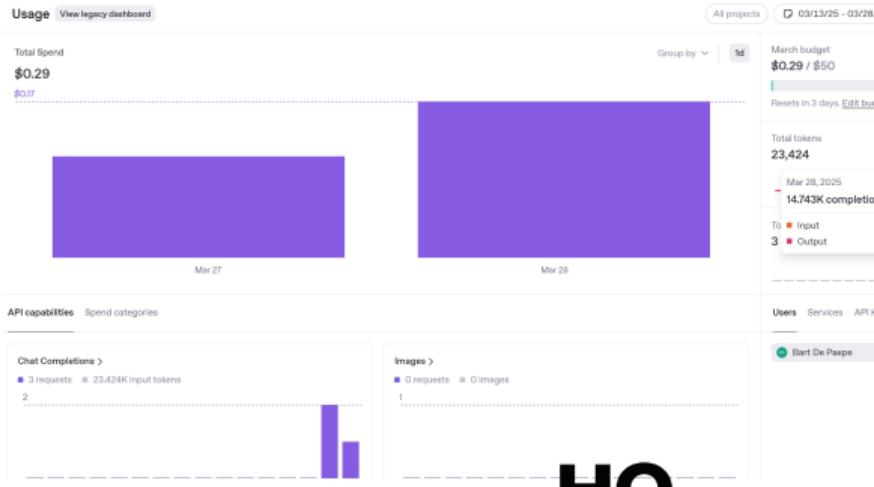
## **parse with scraper**

<h3 style="font-weight:normal;margin:0;font-size:17px;line-height:20px;"> <span style="font-size:12px;">[HTML]</span> [The role of narcissistic personality traits in bullying behavior in adolescence–A systematic review and meta-analysis </a> </h3> <div class="gse\\_alrt\\_sni" style="font-size:17px;">The Ocean Race, known as one of the world's 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>](https://www.s...8)

**HO  
GENT**

# Web scraping (2/2).

- ✗ OpenAI (AKA ChatGPT)
- ✗ Anthropic (met Mirascope)
- ✗ Llama 3 (met Ollama)
- ✗ Serpapi
- ✓ Beautiful Soup



**HO  
GENT**

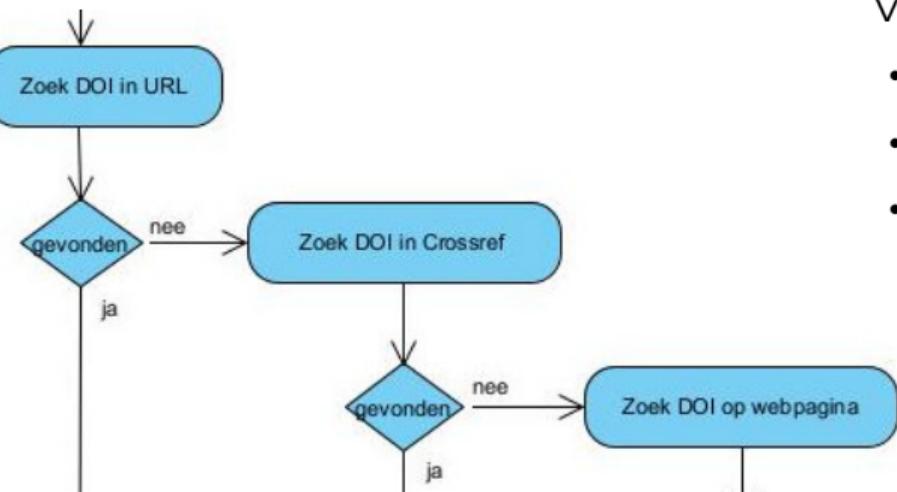
# Natural language processing.

- zoekterm: Bart
  - “Bart vertelde dat hij morgen naar de hogeschool gaat omdat hij daar een belangrijk examen heeft.”
  - “Bart vertelde dat Bart morgen naar de hogeschool gaat omdat Bart daar een belangrijk examen heeft.”
  - “Bart vertelde morgen hogeschool gaan Bart belangrijk examen”
1. “Bart vertel morgen hogeschool gaan Bart belangrijk examen”
  2. “adres hogeschool is Schoonmeersen”
  3. “is laatste examen Bart”

relevantiescore: 0,36

# 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

# Semantic search.

Google Scholar zoekresultaat



publicatie in IMIS



<https://www.youtube.com/watch?v=QSW2L8dkaZk>

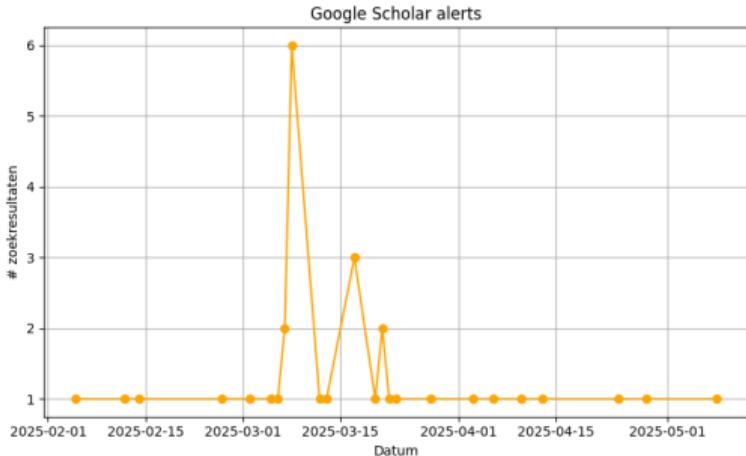
**HO  
GENT**

Photo by Yasmine Duchesne on <https://unsplash.com/photos/clear-footed-glass-with-red-and-yellow-liquid-on-table-VSb6KDaiga0>  
Photo by Farhad Ibrahimzade on <https://unsplash.com/photos/cooked-food-on-black-ceramic-bowl-HNmcpzPHag>

# **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 packing functions. 1 F.  
CRESPO–M. URIBE–E. MARTÍNEZ. Melnikov method for perturbed completely  
integrable systems. 45 S. GARTI–S. SHELAH. Poésie Aristotélienne. 67 WF XUAN–D .

Opslaan Twitter LinkedIn Facebook

**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 ...
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



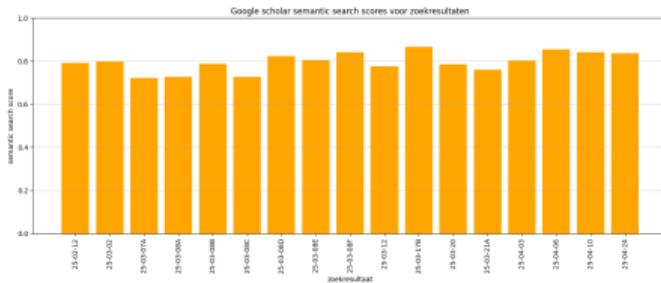
## DOI

	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 (1).

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

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

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

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 (0000819N-LIFEWATCH, J002021N-LIFEWATCH)

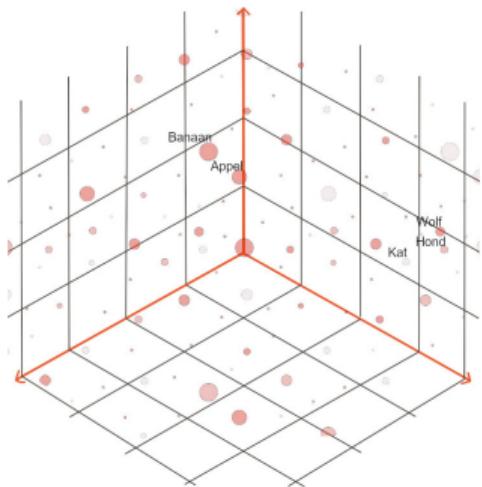
Authors: Anouk Ollevier | Jonas Mortelmans | Wieter Boone | Klaas Deneudt | Marleen De Troch | Roeland Develder | Cedric Goossens | Lorenz Meire | Klas Ove Möller | Leandro Ponsoni | Pascal I. Hablitzel

https://doi.org/10.1002/lom3.10674 | Actions

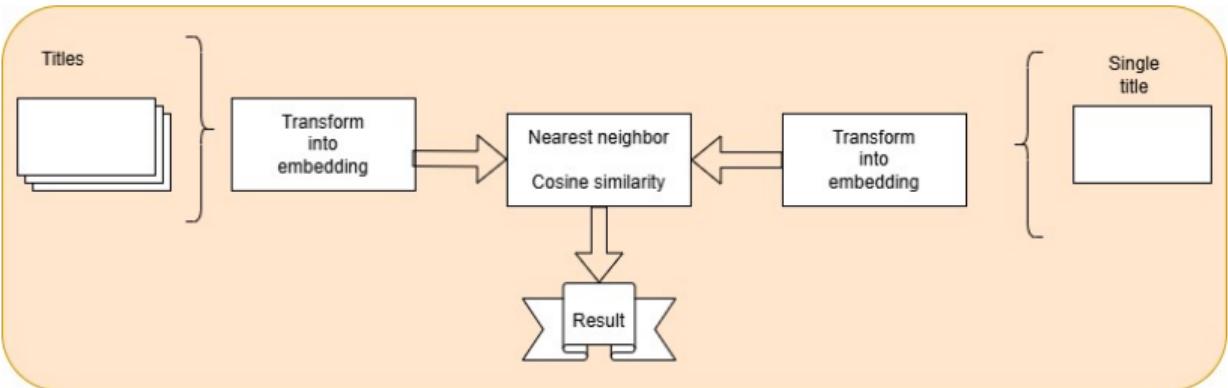
DOI in Crossref

**HO  
GENT**

# Semantic search (2).



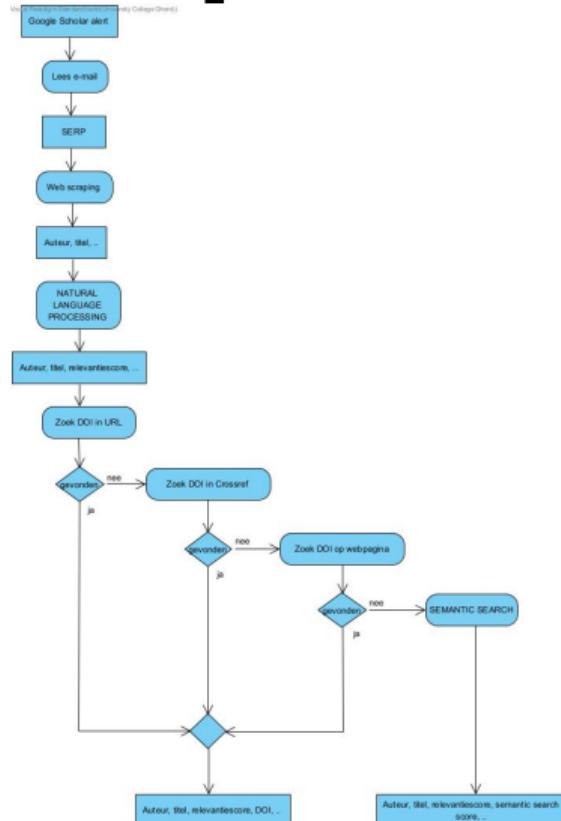
embeddings



schema

**HO  
GENT**

# Proof of Concept



**HO  
GENT**