ChatGPT Search as a tool for scholarly tasks: evolution or devolution?

Cristina I. Font-Julián; Enrique Orduña-Malea; Lluís Codina

Como citar este artículo:

Font-Julián, Cristina I.; Orduña-Malea, Enrique; Codina, Lluís (2024). "*ChatGPT Search* as a tool for scholarly tasks: evolution or devolution?". *Infonomy*, 2(5) e24059.
https://doi.org/10.3145/infonomy.24.059

Cristina I. Font-Julián
https://orcid.org/0000-0003-2351-4816
https://directorioexit.info/ficha6738
Universitat Politècnica de València
Departamento de Comunicación Audiovisual, Documentación e Historia del Arte. The iMetrics Lab
Camino de Vera, s/n. 46022 Valencia, España
crifonju@upv.es

Enrique Orduña-Malea
https://orcid.org/0000-0002-1989-8477
https://directorioexit.info/ficha1074
Universitat Politècnica de València
Departmento de Comunicación Audiovisual, Documentación e Historia del Arte. The iMetrics Lab
Camino de Vera, s/n. 46022 Valencia, España
enorma@upv.es

Lluís Codina
https://orcid.org/0000-0001-7020-1631
https://directorioexit.info/ficha65
Universitat Pompeu Fabra
Departament de Comunicació
Roc Boronat, 138
08018 Barcelona, España
lluis.codina@upf.edu

Abstract
ChatGPT Search was launched on October 31 by *OpenAI* as a new AI-powered search engine. Among its features, it stands out for its ability to retrieve information from various online sources, including scholarly databases, which potentially allows the use of this tool for academic tasks, both quantitative and qualitative. To test its features, five academic tasks are designed: two quantitative (collecting hit count estimates from *Google Search* and scraping bibliometric indicators from *ResearchGate*); two qualitative tasks (performing a narrative synthesis of an academic topic and generating a brief academic author profile), and a mixed task (identifying, collecting and describing a list of publications from *Google Scholar Profiles*). The results show the inability of *ChatGPT Search* to conduct quantitative tasks correctly, fabricating the results (hallucination). Qualitative tasks are performed with better results; however, errors are detected, which prevent recommending the tool without manual analysis and refinement. Finally, the ability to generate links to scientific publications can open up competition among academic sites to be mentioned in the *ChatGPT Search* responses, giving rise to *Academic Generative Engine Optimization (A-GEO)*.

Keywords
Generative artificial intelligence; AI; Link analysis; Web search; Search engines; Narrative synthesis; Information retrieval; ChatGPT Search; Scholarly tasks; Academic tasks; Quantitative vs. Qualitative; *Academic Generative Engine Optimization (A-GEO)*.

Funding
Study supported by project “Publicized research: new indicators to measure the effects of timely marketing on scholarly evaluation (PRIME TIME)” (PID2022-142569NA-I00), funded by MCIN/AEI/10.13039/501100011033, and by “ERDF A way of making Europe”, and by Project “Parameters and strategies to increase the relevance of media and digital communication in society: curation, visualization and visibility (CUVICOM)”, funded by MICIN/AEI/PID2021-123579OB-I00 and by FEDER/EU.

CRediT author statement
Cristina I. Font-Julián: research; writing, review & edition. Enrique Orduña-Malea: conceptualization; methodology; investigation; writing-original draft; funding acquisition. Lluís Codina: methodology; investigation; writing, review & edition; funding acquisition.

1. Introduction
The AI-powered search engines market is expanding rapidly, with tools such as *Copilot*, *Brave Search*, *Search Labs with Search Generative Experience* (from *Google*), *You.com, Phind*, *Komo*, or *Perplexity* standing out.

On July 25, 2024, the company *OpenAI* launched the prototype of a new AI web search engine called *SearchGPT*, accessible to 10,000 users for testing.[1] This prototype combined language models from the *GPT-4* family with real-time web access to organize and interpret information, rather than simply providing a flat list of links. This approach enabled users to receive answers to their queries quickly, in a contextualized format, and based on relevant sources, as reported in the official documentation.

While the term "SearchGPT" was used during the prototype phase, the integrated feature is finally known as "*ChatGPT Search*", officially launched on October 31, 2024, and initially available for *ChatGPT Plus* and *ChatGPT Team* users.[2] The search engine is operational for the *GPT-4o* and *GPT-4o mini* language families, both online (via *chatgpt.com*) and on desktop (*Windows* and *macOS*) and Apps (*iOS*, *Android*). Additionally, *ChatGPT Search* can be installed as the default search engine in the browser through the plugin available in the *Chrome Web Store*.

The launch of *ChatGPT Search* represents a clear strategic move to position itself in the new AI web search engine market, a strategy that *Meta* is also following in order not only to reduce its dependence on traditional web search engines that are transitioning to the AI market[3], such as *Google* (with *Gemini*) or *Bing* (*Copilot*), but also to compete directly with other native AI search engines, such as *Perplexity*.

The launch of *ChatGPT Search* represents a clear strategic move to position itself in the new AI web search engine market

The responses generated by *ChatGPT Search* include inline web citations through a “Sources” button, which, when clicked, opens a side menu that contains the references used in generating the response, along with a link to the original content. Additionally, *ChatGPT Search* provides the kind of output characteristic of generative AI, that is, narrative syntheses in response to prompts. Furthermore, this response process is conducted in the absence of advertising.

The tool also incorporates follow-up questions, non-textual answers (i.e., images and videos), as well as real-time information from the Web, provided by its own web crawler, (*OAI Searchbot*), used to discover and surface websites in search results.[4] Additionally, *ChatGPT Search* leverages third-party search providers and their partners' content to provide comprehensive information to the final users.

Although *OpenAI* has clarified that *OAI-SearchBot* is dedicated solely to indexing for search purposes and does not collect data for training AI models, encourag-

[1] *https://openai.com/index/searchgpt-prototype*
[2] *https://openai.com/index/introducing-chatgpt-search*
[3] *https://www.theinformation.com/articles/meta-develops-ai-search-engine-to-lessen-reliance-on-google-microsoft*
[4] *https://platform.openai.com/docs/bots/overview-of-openai-crawlers*

ing webmasters to allow *OAI-Searchbot* in their site's *robots.txt* file and allow requests from their published IP ranges below, this functionality is expected to cause some controversy, in the same way, that *Perplexity*, which has been sued by different media, such as *NewsCorp*, *Forbes* or *Wired*,[5] for not respecting *robots.txt* protocols.

Focusing on the web sources used by *ChatGPT Search*, the specific databases and sources it access are not publicly disclosed by *OpenAI*. However, the official documentation reports the existence of partnerships "with news and data providers to add up-to-date information and new visual designs for categories like weather, stocks, sports, news, and maps.", among which are *Associated Press*, *Financial Times*, *Hearst*, *Le Monde*, *News Corp*, *Prisa* (*El País*), *Reuters*, *The Atlantic*, *Time*, and *Vox Media*. Other websites or publishers interested can choose to appear in *ChatGPT Search*, opening a new avenue of collaboration between content generators and the server.

Beyond these news content providers, *ChatGPT Search* stands out because it can retrieve information from various online sources, including academic databases. Although it does not have direct access to proprietary academic databases, it can use sources if they are publicly accessible. Among the sources used to generate responses, we can identify institutional repositories (e.g. *Riunet*) and thematic repositories (e.g. *ArXiv*), academic, social networks (*Academia.edu* and *ResearchGate*), bibliographic databases (*Dialnet*, *SciELO*), altmetrics providers (*Altmetric.com*) and academic search engines (*Google Scholar*), among many others.

Beyond news content providers, *ChatGPT Search* stands out because it can retrieve information from various online sources, including academic databases. Although it does not have direct access to proprietary academic databases, it can use sources if they are publicly accessible

On the one hand, quantitative analysis might be possible due to the ability of *ChatGPT Search* to search the Web and process data, serving as a bridge between the researcher and the website. In this sense, AI-powered search engines could alleviate the problems of automatic access derived from consulting search engines directly (**Thelwall**; **Stuart**, 2006), allowing quantitative data to be collected in the search engine results pages, overcoming well-known problems in webometrics studies (**Thelwall**, 2010; **Orduña-Malea**; **Alonso-Arroyo**, 2017).

On the other hand, the capacity of these tools to generate content as a response would allow the realization of more qualitative studies, an aspect already studied for general AI tools (**Lopezosa**; **Goyanes**; **Codina**, 2024). For example, the generation of narrative syntheses, a type of qualitative results that

[5] *https://www.innovationleader.com/moves-that-matter/perplexity-lawsuit-2024-newscorp-rupert-murdoch*

can speed up the obtaining of theoretical frameworks and similar components of academic works. Obtaining this kind of qualitative results has other applications, for example, they could eventually replace some forms of rapid reviews. They can help with the writing of sections such as the introduction and discussion. There is no closed list of possible applications of narrative syntheses, and we only mention these here by way of illustration.

Figure 1 illustrates the *ChatGPT Search* response to query expressed in the prompt, in this case for an academic-related question. The Sources button activates the right sidebar, which includes not only citations (sources cited in the text) but other external online sources, mainly scholarly publications.

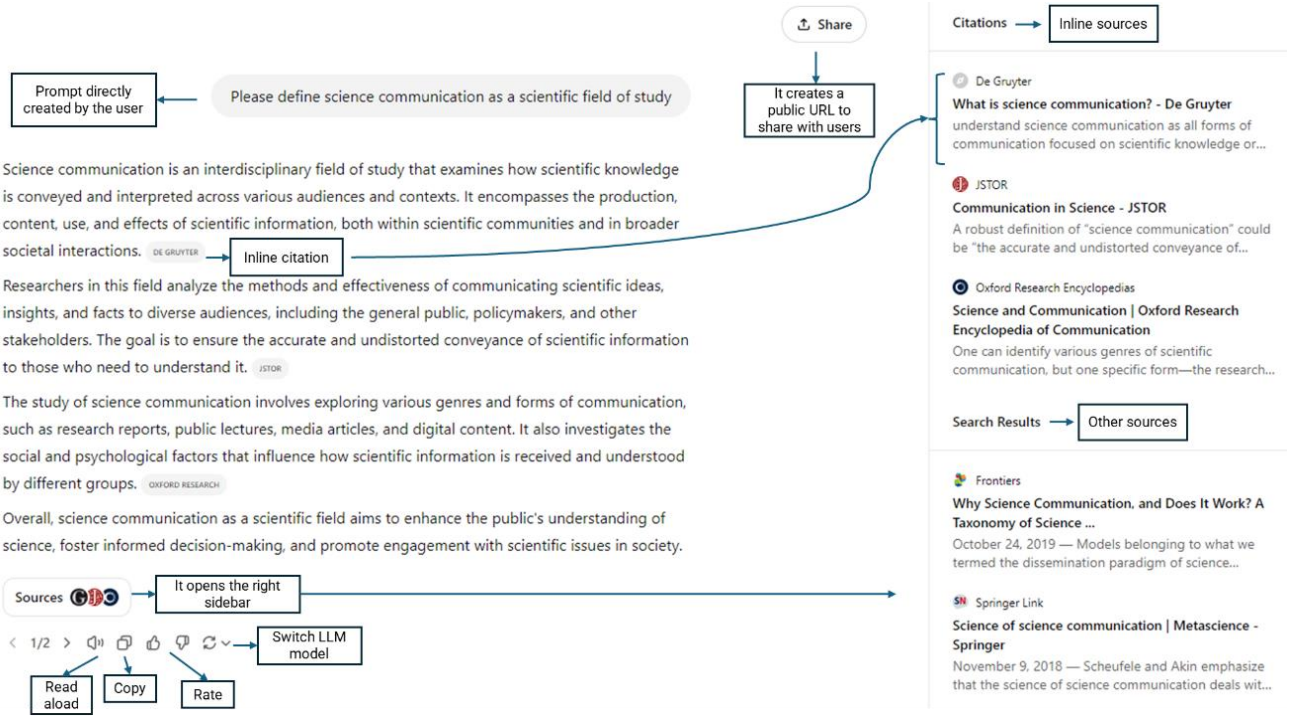

Figure 1. Example of *ChatGPT Search* response for an academic-related question

The emergence of *ChatGPT Search* could represent an opportunity for conducting academic web studies.

The objective of this exploratory work is precisely to test the capacity, reliability, and precision of *ChatGPT Search* to conduct both quantitative and qualitative tasks. To do this, a selection of five tasks have been designed to

a) Collect hit count estimates by scraping search engine results pages.
b) Collect bibliometric indicators by scraping author cards from academic profiles.
c) Generate a narrative synthesis supported by sources on academic topics.
d) Generate the academic profile of a researcher.
e) Collect bibliographic data by scraping the author academic profiles.

2. Method

The performance of *ChatGPT Search* is tested by designing five case studies, which are described below.

First, two tasks with a quantitative focus are designed. The purpose of these tasks is to check whether *ChatGPT Search* can collect the data requested and, if it is possible, to check whether the data collected is accurate.

Then, other two tasks with a qualitative focus were also designed. The purpose of these tasks is to evaluate the system's ability to "understand" an academically oriented prompt and generate a valid text, to expedite the creation of a theoretical or conceptual framework given a research problem in the first case, and to create a correct and accurate academic profile in the second case.

Finally, a mixed (quantitative and qualitative) is designed, with the purpose of testing the tool to locate, select and collect bibliographic data.

All data were extracted between November 3rd and 6th, 2024.

Task A: Collecting hit count estimates from Google Search results

This quantitative task uses the *ChatGPT Search* interface to automatically obtain the number of results *Google Search* provides for a given query.

To this end, the 25 best-ranked universities in the *SCImago Institutions Ranking* were selected, and their domain names were collected. Two types of searches were then designed:

> *Web size*: number of URLs indexed by *Google Search* for the domain name indicated (e.g., site:”harvard.edu”).
> *Title mention*: number of web pages that mention a university by name, excluding self-mentions of the university's website (e.g. "Harvard University" -site:”harvard.edu”).

All queries were performed using the following two data collection techniques:

> *Manual*: searches were executed directly in the google.com search interface, obtaining the number of results offered on the first page of results (SERP1) through the *Chrome* plugin "Bring back Google search result counter and display indexed pages".
>
> *Automatic*: searches were conducted using *ChatGPT Search* (GPT-4o mini model), from prompts-1 (Annex A, table A.1) and 2 (Annex A, table A.2), available in the supplementary material.

Task B: Collecting author-level indicators from ResearchGate

This quantitative task uses the *ChatGPT Search* interface to automatically locate and collect a number of author-level met
rics displayed in *ResearchGate*.

In this case, the need arises to extract 3 metrics (Publications, Reads, Citations) from the public profile of 10 relevant authors in quantitative studies of science and meta-science.

As in previous cases, the data is extracted manually and through *ChatGPT Search* through the prompt-3 (Annex B, table B.1), available in the supplementary material.

Task C: Generating a narrative synthesis supported by sources
This qualitative task is based on the prompt-4 (**Annex C**, **table C.1**), and includes a request to create a narrative synthesis supported by sources, in this example dedicated to the question “Can solutions journalism reduce news avoidance?”.

The prompt was directly queried to *ChatGPT Search*, and the response was obtained, and subsequently analyzed.

Task D: Creating a brief narrative academic author profile
This qualitative task is based on the prompt-5 (Annex D, table D.1), and includes a request to create an academic profile of the three authors of this contribution. The query explicitly indicates the source to collect the data, the sections to be created, and the extension of the text.

Task E: Collecting bibliographic data from Google Scholar Profiles
This mixed (quantitative and qualitative) task uses the *ChatGPT Search* interface to automatically locate and collect bibliographic data from *Google Scholar Profiles* for a given enquiry.

We start from a list of 25 relevant authors in the field of bibliometrics, according to *Google Scholar Profiles* (those authors including the tag “bibliometrics” as a keyword in their *Google Scholar Profile*). We intend to use *Google Scholar Profiles* to create a table indicating a) the total number of citations received by each author; b) the publication with the most citations received; c) the source and year of publication of said work; and d) the number of citations received by that most cited work.

Again, the data is collected both manually and through *ChatGPT Search* (GPT-4o mini model) using the prompt-6 **(**Annex E; Table E.1).

3. Results

Task A: Collecting hit count estimates from Google Search results
After the first interactions with the *ChatGPT Search* chat, the authors check the tool focuses on delivering concise, conversational answers with relevant sources, without indicating the overall volume of available information. That is, *ChatGPT Search* does not provide "hit count estimates". For this reason, the task focuses on being able to perform the query on *Google Search* using *ChatGPT Search* as an intermediary.

At first, the tool provides contradictory information, indicating that it understands the query but cannot carry it out (Annex A, Figure A.1), even indicating that it is unable to carry out live web searches or access real-time data (Annex A, Figure A.2), an aspect that precisely characterizes *ChatGPT Search*. However, with a simple and direct query finally the chat indicates its capability to conduct web searches (Annex A, Figure A.3).

As the tool does not allow using file attachment and search at the same time, the prompts designed were embedded directly in the search box. Prompt-1 (Annex A, table A.1) was applied to obtain the web size results, and prompt-2 (Annex A, table A.2) to obtain the Title mention results. The raw results obtained are included in Annex A (figures A.4 and A.5).

We asked the chat what procedure it followed to obtain the data, and the chat responded, "I utilized a *Chrome* extension called *Google Search Results Counter*." This tool restores the total number of search results on *Google* pages, which is hidden by default. The authors redefined the prompt to ask the chat to collect the data directly using this *Chrome* extension.

Table 1. Comparison of hit count estimates values (web size queries)

University	GPT-4o mini (without plugin)	GPT-4o mini (plugin)	Manual
harvard.edu	1,220,000,000	1,200,000	9,160,000
ucas.ac.cn	1,020,000,000	1,500,000	136,000
tsinghua.edu.cn	1,150,000,000	2,000,000	1,090,000
zju.edu.cn	1,030,000,000	1,800,000	496,000
sjtu.edu.cn	1,100,000,000	1,600,000	750,000
pku.edu.cn	1,200,000,000	2,100,000	1,330,000
stanford.edu	1,150,000,000	1,700,000	3,410,000
ox.ac.uk	1,100,000,000	1,900,000	2,420,000
jhu.edu	1,050,000,000	1,400,000	5,480,000
ucl.ac.uk	1,000,000,000	1,500,000	1,700,000
mit.edu	1,200,000,000	2,200,000	3,710,000
utoronto.ca	1,050,000,000	1,300,000	2,160,000
washington.edu	1,100,000,000	1,600,000	1,350,000
umich.edu	1,050,000,000	1,700,000	2,970,000
hust.edu.cn	1,020,000,000	1,200,000	4,450
upenn.edu	1,150,000,000	1,500,000	2,110,000
cornell.edu	1,100,000,000	1,600,000	3,750,000
columbia.edu	1,150,000,000	1,800,000	1,870,000
usp.br	1,000,000,000	1,400,000	6,770,000
cam.ac.uk	1,100,000,000	1,900,000	1,460,000
ucla.edu	1,150,000,000	1,700,000	1,400,000
unimelb.edu.au	1,050,000,000	1,500,000	1,400,000
scu.edu.cn	1,020,000,000	1,200,000	343,000
nus.edu.sg	1,050,000,000	1,300,000	711,000
imperial.ac.uk	1,100,000,000	1,600,00	210,000

Table 1 shows the comparison of the results obtained manually with those obtained through *ChatGPT Search* (with and without the *Chrome* extension). The

results show the disparity of the values obtained. It is interesting to note how the queries that were explicitly requested with the *Chrome* extension have reduced their order of magnitude by thousands. *ChatGPT Search* exhibits the same behavior when querying Title mentions. Due to space constraints, these results are included in Annex A (Table A.3).

Task B: Collecting author-level indicators from ResearchGate

As in the previous cases, direct consultation of the prompt-3 (Annex B, figure B.1) does not allow data extraction, citing various limitations, such as privacy and legal concerns (many websites have terms of service that prohibit scraping to protect user privacy and their data); technical limitations that contradict the terms of *ChatGPT Search*'s features (for example, "my current environment doesn't allow for real-time internet access or direct interaction with external web pages, which means I can't retrieve or extract live data from the web"); or due to the existence in the website of measures to prevent automated data collection, such as requiring logins, using CAPTCHAs, or serving dynamic content that changes based on user interactions, that do not apply in the task required.

After a few interactions with the chat, the results could be obtained (Annex B; Figure B.1). Table 2 shows the comparative data of author-level metrics collected from *ChatGPT Search* and those obtained by manual inspection of websites. As the reader can observe, the data are completely different.

Table 2. Comparison of author-level metrics (*ResearchGate* queries)

Author	***ChatGPT Search* data collection**			**Manual direct data collection**		
	Publications	**Reads**	**Citations**	**Publications**	**Reads**	**Citations**
Enrique Orduña-Malea	100	50,000	1,200	256	249,589	4,720
Mike Thelwall	200	150,000	5,000	681	387,440	37,402
Isidro Aguillo	80	40,000	900	191	74,856	3,205
Anne-Wil Harzing	150	100,000	3,000	155	410,332	17,398
Cinzia Daraio	60	30,000	800	100	24,563	4,538
Stefanie Haustein	120	60,000	1,500	104	41,417	8,241
Cassidy Sugimoto	110	55,000	1,400	228	114,125	11,362
Kevin Boyack	90	45,000	1,000	152	108,615	11,367
John Ioannidis	300	200,000	10,000	1,013	349,534	239,900
Isabella Peters	70	35,000	850	107	34,817	2,633

Task C: Obtaining a narrative synthesis supported by sources

The response generated is available in Annex C (Table C.2) and is publicly available through a shared URL.[6]

The narrative synthesis is conceptually correct. Furthermore, it is articulated in various sections that well reflect the different dimensions of the phenomenon addressed. The synthesis has demonstrated the ability to "interpret" the question correctly and has detected that it is a yes/no question and has answered it that way, but also providing the development that we have commented on. It also includes relevant sources and relationships to the different parts of the synthesis. It can be concluded that *ChatGPT Search* has resolved the task well.

[6] *https://chatgpt.com/share/672ba634-78c4-8009-ae2b-ee1b7d4d7261*

Task D: Obtaining an academic profile

The analysis results of the three generated profiles is included in Table 3, and is based on evaluating the attributes explicitly requested in the prompt (length, structure) and the quality of the information provided (selection of publications, collaborators, sources). The responses created for each profile are available in shared URLs.[7]

Some specific errors are recorded. For example, in Cristina Font's profile, there is an affiliation error, probably because a co-author does belong to the indicated affiliation. In the case of Enrique Orduña's profile, errors are identified in the links assigned to the inline citations. Other errors are located in the bibliographic description of the selected publications and the affiliation of one collaborator, which is correctly identified in the author's profile.[8]

In the case of Dr. Codina, *ChatGPT Search* has developed a fairly representative profile of his academic activity. It has well identified the most recent lines of research, and the list of co-authors has been well selected as well. The publications used to outline his work are adequate, except for one of them (a case of hallucination). Beyond the generation of one imaginary paper, *ChatGPT Search* has performed an accurate profile in the case of this profile.

Notably, the citation sources have only been activated in the case of Enrique Orduña's profile, which is precisely the only one of the profiles including citations. The reason why *ChatGPT Search* does not provide citations to the remaining profiles is unknown. Otherwise, the number of search results is 12 for all the profiles.

Table 3. Evaluation of the academic profiles created by *ChatGPT Search*

Parameters	**Cristina Font**	**Enrique Orduña**	**Lluís Codina**
Length	Correct	Correct	Correct
Structure	Correct	Correct	Correct
Affiliation	Incorrect	Correct	Correct
Publications- Total selected	5	4	5
Publications- Relevant	5	4	4
Publications- Correct description	3	4	4
Publication links- *Total*	NA	4	NA
Publication links- *Correct URL*	NA	3	NA
Citations- Total	0	5	0
Search results- Total	12	12	12
Collaborators- Total	4	4	5
Collaborators- Relevant	4	4	5
Collaborators- Correct description	3	4	5
Collaborators- Correct URL	NA	2	NA
Line of research	Correct	Correct	Correct

[7] Lluís Codina: *https://chatgpt.com/share/672bd5eb-89f8-8009-a2fa-82fa1ee726a2*
Cristina I. Font-Julián: *https://chatgpt.com/share/672bd768-5d00-8009-a147-86fd5a9f2de6*
Enrique Orduña-Malea: *https://chatgpt.com/share/672bd8ec-4d38-8009-ac40-673d82858ef6*

[8] This author changed recently his affiliation. The affiliation shown in Font's profile is the old one.

NA: data not available.

Task E: Collecting bibliographic data from Google Scholar Profiles

The authors again encounter the initial tool's refusal to perform the requested task (Annex E, figure E.1). However, after conversing to the virtual assistant, it finally processes the query included in the prompt-6 (Annex E, table E.1). Annex E includes the raw data collected by *ChatGPT Search* (Annex E, table E.2), as well as the data collected directly from *Google Scholar Profiles* (Annex E, table E.3).

The differences between the data collected from *ChatGPT Search* and *Google Scholar Profiles* are shown in Table 4. As we can observe, the most cited publication is correct only for 6 authors (24%). Of these cases, the bibliographic description is also correct for 5 authors. On the other hand, regardless of whether the indicated publication is or is not the most cited, the source of the identified work is correct for 11 authors (44%), and the year of publication for 15 authors (60%).

As regards citations, the figures provided are only close in the case of four authors. The publication identified as the most cited could not be located for four authors, neither in the profile nor in *Google Scholar*, so it is considered that it is a hallucination (AI-lusion) of the tool. Other minor issues are mentioned in the *Notes* column. The description was fully accurate for two authors, precisely the first two authors of the list provided in the prompt.

4. Discussion

This work has tested *ChatGPT Search* based on five specific tasks related to academic activity, two of which consist of quantitative tasks (*ChatGPT Search* as a scraper), other two consist of qualitative tasks (narrative synthesis and content creation), and the last one includes both quantitative and qualitative tasks).

In the case of quantitative tasks, the results have been mostly negative. The chat often reports that it cannot obtain the data without access to the search service. A plausible explanation is that the system may fail due to too many users. The results are invented when it finally accesses and collects the data (hallucination).

In the case of quantitative tasks, the results have been mostly negative. The chat often reports that it cannot obtain the data without access to the search service. A plausible explanation is that the system may fail due to too many users. The results are invented when it finally accesses and collects the data (hallucination)

To check the reason for the hallucination, a GPT-4 assistant (*ScraperAPI*) was used to repeat these tasks. In the case of task C, the results obtained were correct and accurate. The results obtained are shown in Annex C (figure C.2). Unfortunately, *ScraperAPI* did not work correctly in the case of tasks A and B.

Table 4. Accuracy of *ChatGPT Search* data for each author

Author Name	Most cited	Source	Year	Citations	Notes
Nees Jan van Eck	Correct	Correct	Correct	Correct	
Chaomei Chen	Correct	Correct	Correct	Correct	
Ludo Waltman	Incorrect	Correct	Correct	Correct	
John Mingers	Incorrect	Incorrect	Incorrect	N/D	The article does not exist (hallucination)
Vincent Larivière	Incorrect	Incorrect	Incorrect	Incorrect	
Wolfgang Glänzel	Incorrect	Incorrect	Incorrect	N/D	The article does not exist (hallucination)
Johan Bollen	Incorrect	Correct	Correct	Incorrect	
Francis Narin	Incorrect	Incorrect	Correct	Incorrect	
Yves Gingras	Incorrect	Incorrect	Incorrect	Incorrect	
Henry Small	Correct	Correct	Correct	Incorrect	
András Schubert	Incorrect	Incorrect	Incorrect	Incorrect	The article does not exist (hallucination)
Blaise Cronin	Incorrect	Incorrect	Correct	Correct	
Ying Ding	Incorrect	Incorrect	Incorrect	Incorrect	
Félix de Moya Anegón	Correct	Correct	Correct	Incorrect	Title name variation.
Emilio Delgado López-Cózar	Incorrect	Correct	Correct	Incorrect	
Paul Wouters	Incorrect	Incorrect	Incorrect	Incorrect	
Diana Hicks	Incorrect	Incorrect	Correct	Incorrect	Name variation and corresponding to another author.
Judit Bar-Ilan	Correct	Incorrect	Incorrect	Incorrect	
Thed van Leeuwen	Incorrect	Incorrect	Incorrect	Incorrect	The article does not belong to the author
Martin Meyer	Incorrect	Correct	Correct	Incorrect	
Stefanie Haustein	Incorrect	Incorrect	Correct	Incorrect	
Rodrigo Costas	Correct	Correct	Correct	Incorrect	
Howard D. White	Incorrect	Incorrect	Incorrect	N/D	The article does not exist (hallucination)
Ciriaco Andrea D'Angelo	Incorrect	Correct	Correct	Incorrect	
Katherine W. McCain	Incorrect	Correct	Correct	Incorrect	

In the case of task C, there is an added difficulty. One must consider that the data in *Google Scholar Profiles* are not necessarily correct. Moreover, the same publication might have a different number of citations received according to different profiles due to the diverse data curation practices of the authors, and the publications may be indexed in *Google Scholar* outside of the academic profiles. The task, however, was about identifying the data included in the profiles, regardless of their actual accuracy.

Another important limitation to the correct execution of tasks is the high dependence on the prompt. Similar prompts can generate very different data. Even the same prompt can generate different responses depending on the previous conversation with the chat. These characteristics add unpredictable instability

to the data collection process, which limits the use of this tool in quantitative tasks.

As regards quantitative tasks, the results have been more correct, although not fully satisfactory. The tool has been able to carry out its tasks, including a few errors. The main limitation has been identified in the generation of incorrect links to selected citations and sources in the texts.

Specifically, the inclusion of links to sources is critical, since it is estimated that with the increase in the use of this tool, the number of visits from *ChatGPT Search* to the mentioned resources may increase considerably, making *ChatGPT Search* a significant source of access to academic resources, and competing not only with *Perplexity*, but with *Google Scholar*. Therefore, the generation of incorrect links could limit the user experience of the academic community.

The need to appear as a cited source might open a competition among academic sites to be mentioned in the *ChatGPT Search* responses, what can be called Academic Generative Engine Optimisation (A-GEO), as a specialization of general GEO (**Aggarwal** *et al.*, 2024) with effects on traditional SEO The number of visits from AI-based search engines to academic journals, recently reported by **Urbano** (2024) evidences the relevance of these platforms to drive visits to academic websites (**Lopezosa**; **Rovira**; **Codina**, 2024).

ChatGPT Search is believed to have enormous potential for generating web traffic to academic sites, a fact that could make it a serious competitor to *Google Scholar* in the long term as the main gateway to scientific literature. The availability of the tool as a browser in *Chrome* aims to make it more widely used by the population, which could change the web search market

These practices might pose a risk of gaming bots generating search responses that benefit linking certain publishers, methods, schools of thought, etc., as well as the dissemination of scientific fake news related to health issues (**Wan**; **Wallace**; **Klein**; 2024), an issue that the media has already pointed out for its social impact (**Bains**, 2024).

5. Conclusions

ChatGPT Search is currently not a recommended tool for conducting quantitative academic tasks, nor can it be used for webometric data collection.

However, *ChatGPT Search* has demonstrated an acceptable level of performance in qualitative tasks. Specifically, it has developed a successful qualitative synthesis. Although other tools provide more elaborate narrative syntheses and with greater capillarity in the relationship between proposition and quote. In the case of the generation of academic author profiles, this task has been carried out acceptably well, despite introducing a few bibliographic errors and misplaced links.

ChatGPT Search is believed to have enormous potential for generating web traffic to academic sites, a fact that could make it a serious competitor to *Google Scholar* in the long term as the main gateway to scientific literature. The availability of the tool as a browser in *Chrome* aims to make it more widely used by the population, which could change the web search market.

The inclusion of links to sources is critical, since it is estimated that with the increase in the use of this tool, the number of visits from *ChatGPT Search* to the mentioned resources may increase considerably, making *ChatGPT Search* a significant source of access to academic resources, and competing not only with *Perplexity*, but with *Google Scholar*. Therefore, the generation of incorrect links could limit the user experience of the academic community

6. Supplementary material

https://infonomy.scimagoepi.com/index.php/infonomy/article/view/77/108

7. References

Aggarwal, Pranjal; **Murahari, Vishvak**; **Rajpurohit, Tanmay**; **Kalyan, Ashwin**; **Narasimhan, Karthik**; **Deshpande, Ameet** (2024). Geo: Generative engine optimization. In: *Proceedings of the 30th ACM SIGKDD Conference on Knowledge Discovery and Data Mining* (pp. 5-16). New York (USA): ACM.
https://arxiv.org/pdf/2311.09735

Bains, Callum (2024). The chatbot optimisation game: Can we trust AI web searches? *The Guardian*, 3 November.
https://www.theguardian.com/technology/2024/nov/03/the-chatbot-optimisation-game-can-we-trust-ai-web-searches

Lopezosa, Carlos; **Goyanes, Manuel**; **Codina, Lluís** (2024) *Acelerando la investigación cualitativa con inteligencia artificial: una guía práctica para el diseño, desarrollo y ejecución de investigación con entrevistas*. Barcelona: Universitat de Barcelona.
http://hdl.handle.net/2445/211948

Lopezosa, Carlos; **Rovira, Cristòfol**; **Codina, Lluís** (2024). *La IA generativa y su aplicación al SEO, SEM, analítica web y diseño UX*. Barcelona: Universitat Pompeu Fabra, 92 pp. *https://doi.org/10.31009/cr.2024.11*

Orduña-Malea, Enrique; **Alonso-Arroyo, Adolfo** (2017). *Cybermetric techniques to evaluate organizations using web-based data*. Oxford: Chandos Publishing.
https://doi.org/10.1016/C2016-0-00399-6

Thelwall, Mike (2010). Webometrics: Emergent or doomed? *Information Research: An International Electronic Journal*, v. 15, n. 4.
https://informationr.net/ir/15-4/colis713.html

Thelwall, Mike; **Stuart, David** (2006). Web crawling ethics revisited: Cost, privacy, and denial of service. *Journal of the American Society for Information Science and Technology*, v. 57, n. 13, pp. 1771–1779.
https://doi.org/10.1002/asi.20388

Urbano, Cristóbal (2024). Los chatbots de IA
generativa como fuentes de tráfico web: aproximación desde la analítica web de un portal de revistas. *Anuario ThinkEPI*, v.18.
https://doi.org/10.3145/thinkepi.2024.e18a31

Wan, Alexander; **Wallace, Eric**; **Klein, Dan** (2024). What evidence do language models find convincing? *arXiv preprint*.
https://arxiv.org/pdf/2402.11782