

Advertising in Generative Information Retrieval

Ines Zelch^{1,2}

Matthias Hagen¹

Martin Potthast^{2,3}

¹Friedrich-Schiller-Universität Jena ²Leipzig University ³ScaDS.AI

OSSYM

04.10.2023

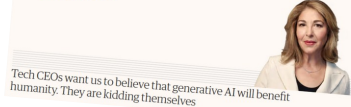
webis.de

Generative Information Retrieval

France bets big on open-source AI

French leaders have a plan to build a native AI industry. There's just one problem: They're in the

AI machines aren't 'hallucinating'. But their makers are
Naomi Klein



Tech CEOs want us to believe that generative AI will benefit humanity. They are kidding themselves

AI is killing the old web, and the new web struggles to be born

Generative AI is changing the old web, making it of generate lower-quality content. We're just beginning to see the effects of these changes.



Announcing AI2 OLMo, an Open Language Model Made by Scientists, for Scientists

Have you noticed that news stories and marketing material about Artificial Intelligence are typically illustrated with clichéd and misleading images?

AI's 'Godfather' Should Have Spoken Up Sooner

Hopefully, Google scientist Geoffrey Hinton's warnings about the technology's potential harms will persuade other researchers to come forward.

Inside The High-Stakes, AI-Powered Race To Dethrone Google Search

Pranav Dixit / BuzzFeed News
5 May 2023 · 10 min read

How Artificial Intelligence Can Aid Democracy

BY BRUCE SCHNEIER, HENRY FARRELL, AND NATHAN E. SANDERS · APRIL 21, 2023 · 10:30 AM

World's first artificial intelligence law moves closer to passage in European Parliament

11 May 2023 | News
To the alarm of mostly-US industry, the EU AI Act gets preliminary nod to control the emerging technology - and sets the stage for months of negotiations in Brussels and Washington
By Richard L. Hudson

A Cambridge Analytica-style scandal for AI is coming

Can you imagine a company putting a new vehicle on the market without built-in safety features?

OpenAI announces ChatGPT successor GPT-4

By Melissa Mattar · 14 March

OpenAI, IBM Urge Senate to Act on AI Regulation After Past Tech Failures

- Altman tells Congress the US should license advanced models
- IBM's Montgomery says regulation must focus on risks of AI use

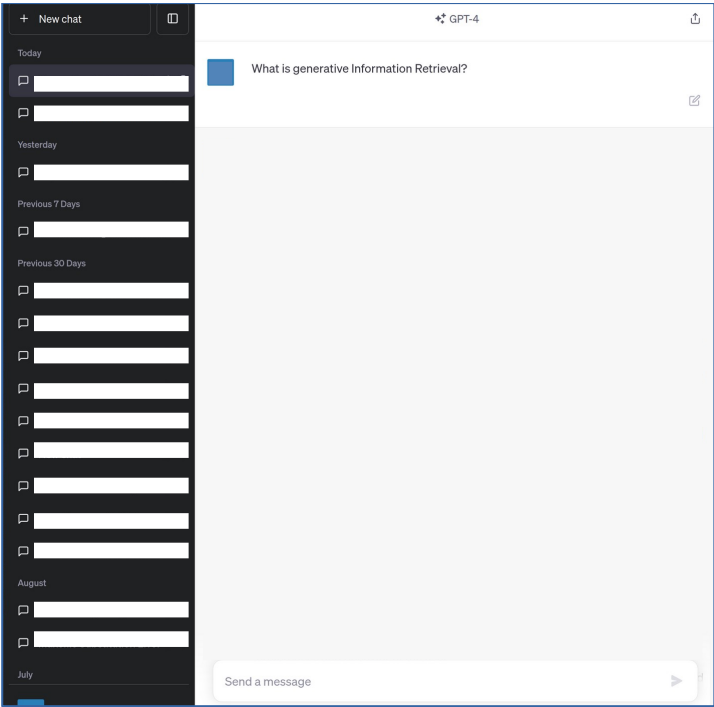
Pause Giant AI Experiments: An Open Letter

We call on all AI labs to immediately pause for at least 6 months the training of AI systems more powerful than GPT-4.

- Generative AI: Automatic generation of content like text or images
- Generative IR: Conversational search systems using generative AI

Generative Information Retrieval

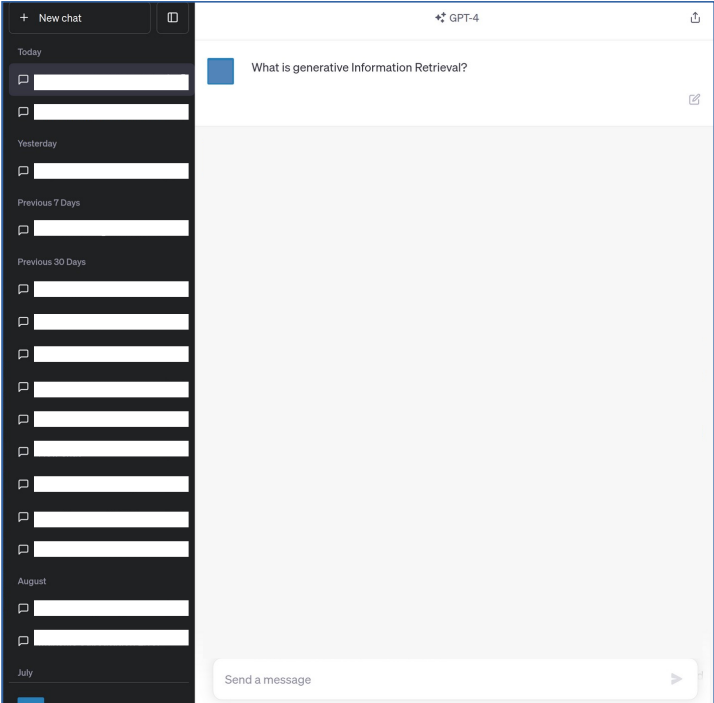
ChatGPT



Generative Information Retrieval

ChatGPT

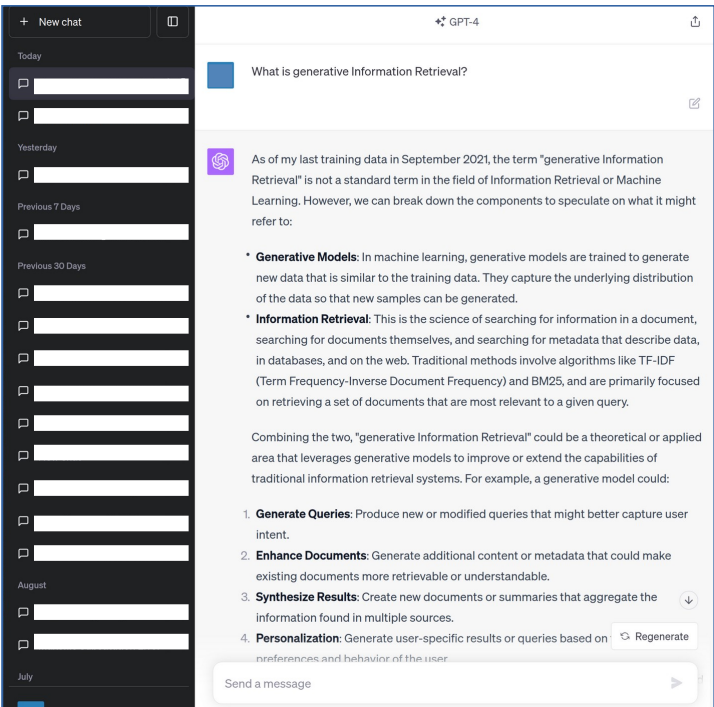
What is generative Information Retrieval?



Generative Information Retrieval

ChatGPT

What is generative Information Retrieval?

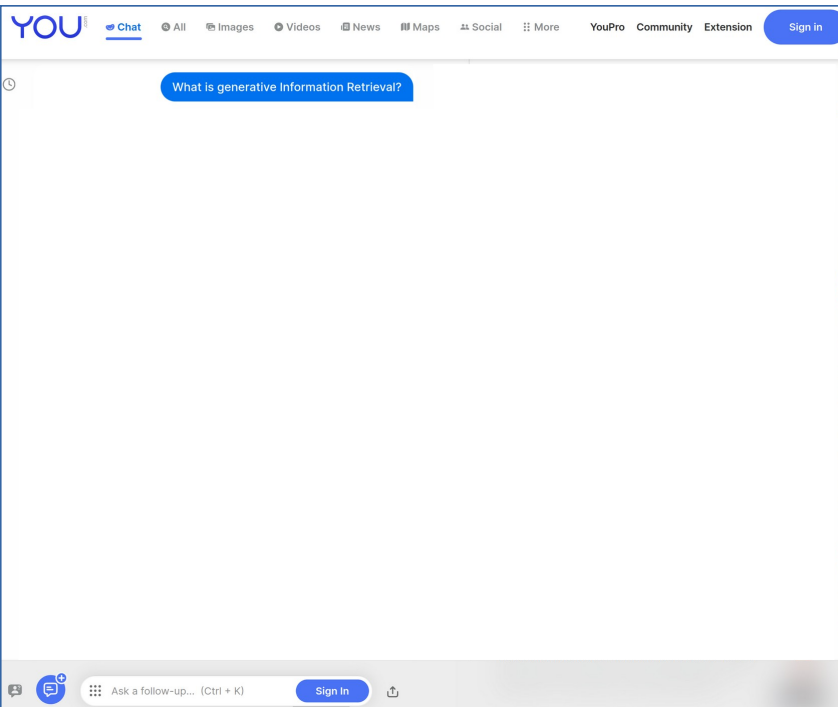
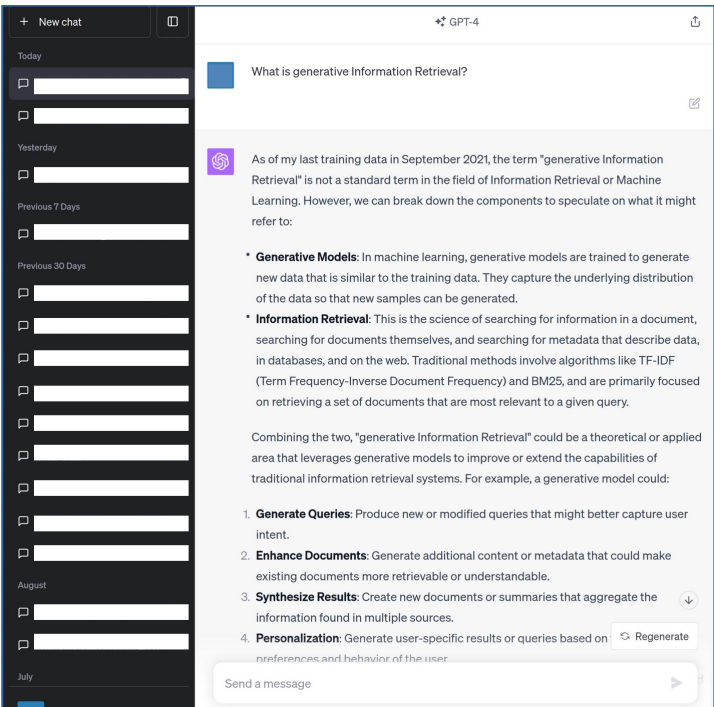


Generative Information Retrieval

ChatGPT

You Chat

What is generative Information Retrieval?



Generative Information Retrieval

ChatGPT

You Chat

What is generative Information Retrieval?

New chat

Today

Yesterday

Previous 7 Days

Previous 30 Days

August

July

GPT-4

What is generative Information Retrieval?

As of my last training data in September 2021, the term "generative Information Retrieval" is not a standard term in the field of Information Retrieval or Machine Learning. However, we can break down the components to speculate on what it might refer to:

- Generative Models:** In machine learning, generative models are trained to generate new data that is similar to the training data. They capture the underlying distribution of the data so that new samples can be generated.
- Information Retrieval:** This is the science of searching for information in a document, searching for documents themselves, and searching for metadata that describe data, in databases, and on the web. Traditional methods involve algorithms like TF-IDF (Term Frequency-Inverse Document Frequency) and BM25, and are primarily focused on retrieving a set of documents that are most relevant to a given query.

Combining the two, "generative Information Retrieval" could be a theoretical or applied area that leverages generative models to improve or extend the capabilities of traditional information retrieval systems. For example, a generative model could:

- Generate Queries:** Produce new or modified queries that might better capture user intent.
- Enhance Documents:** Generate additional content or metadata that could make existing documents more retrievable or understandable.
- Synthesize Results:** Create new documents or summaries that aggregate the information found in multiple sources.
- Personalization:** Generate user-specific results or queries based on preferences and behavior of the user.

Send a message

YOU Chat All Images Videos News Maps Social More YouPro Community Extension Sign in

What is generative Information Retrieval?

Generative Information Retrieval (IR) refers to the application of generative models in the field of information retrieval. Generative models, such as language models, are able to generate new data based on the patterns and structures they have learned from existing data. In the context of Information Retrieval, generative models can be used to generate relevant documents or responses based on a given query or user input.

Generative IR techniques focus on generating documents or responses that are contextually relevant and informative. This can be achieved by training generative models on large text corpora and using them to generate new documents or responses that are similar in style and content to the training data. Generative IR techniques can be applied to various tasks in information retrieval, such as document retrieval, response generation, recommendation systems, summarization, and more.

The goal of generative IR is to improve the quality and relevance of the information retrieved by leveraging the power of generative models. By generating documents or responses that are contextually relevant and informative, generative IR can enhance the user's search experience and provide more accurate and useful information.

Generative IR is an important field in AI, as it combines

Related web results

- Gen-IR@ SIGIR 2023: The First Workshop on Generative ... arxiv.org
- 5 Jun 2023 ... Abstract: Generative information retrieval (IR) has experienced substantial growth across multiple research communities (e.g., information ...
- Generative-IR coda.io
- gabriben/awesome-generative-information-retrieval - GitHub github.com
- awesome-generative-information-retrieval Awesome. Recently, conversational models (Galactica, YOU.com, perplexity.ai) started to be able to access the web ...
- Generative Information Retrieval | Proceedings of the 46th ... dl.acm.org
- 18 Jul 2023 ... Donald Metzler, Yi Tay, Dara Bahri, and Marc Najork. 2021. Rethinking search: making domain experts out of dilettantes. ACM SIGIR Forum, Vol. ...
- Generative Information Retrieval storage.googleapis.com

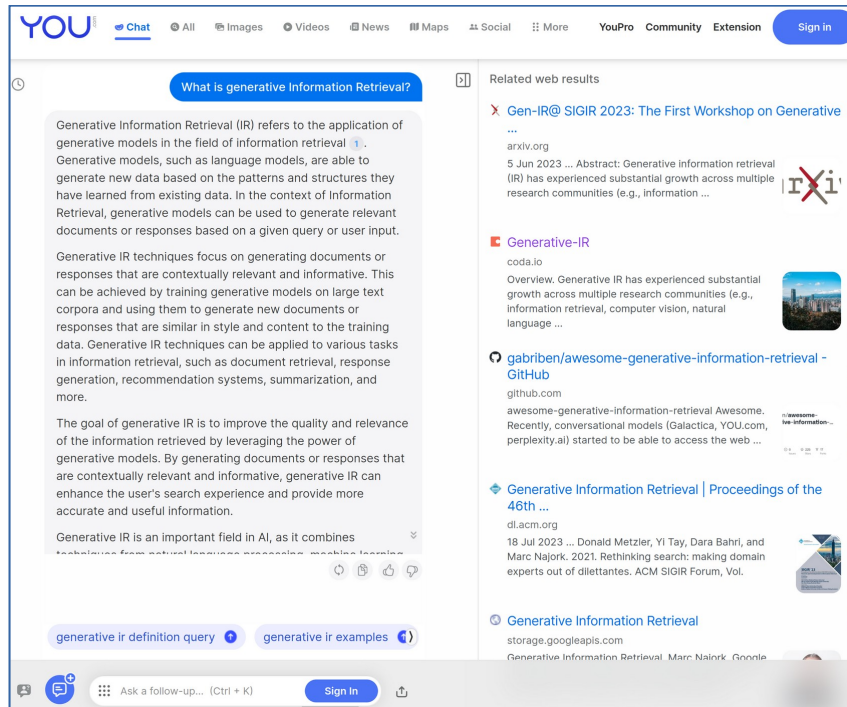
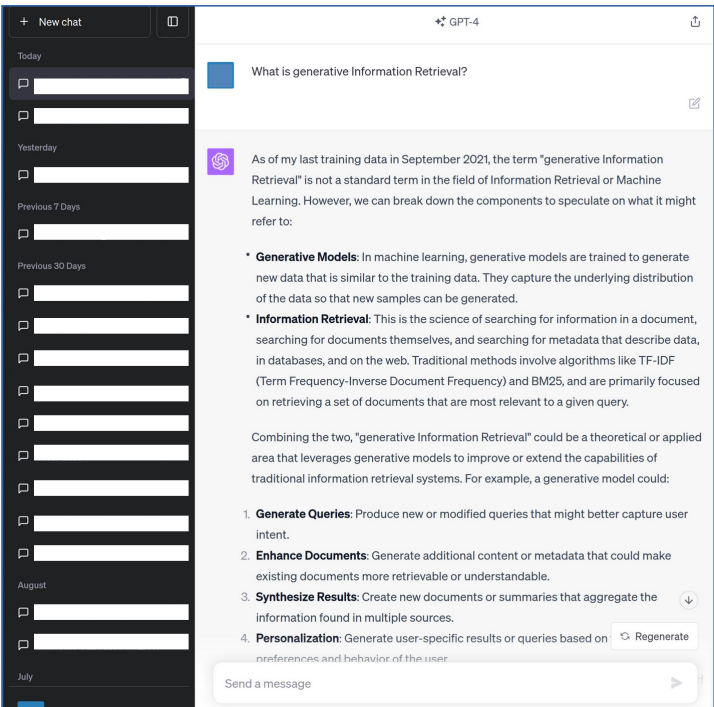
Ask a follow-up... (Ctrl + K) Sign In

Generative Information Retrieval

ChatGPT

You Chat

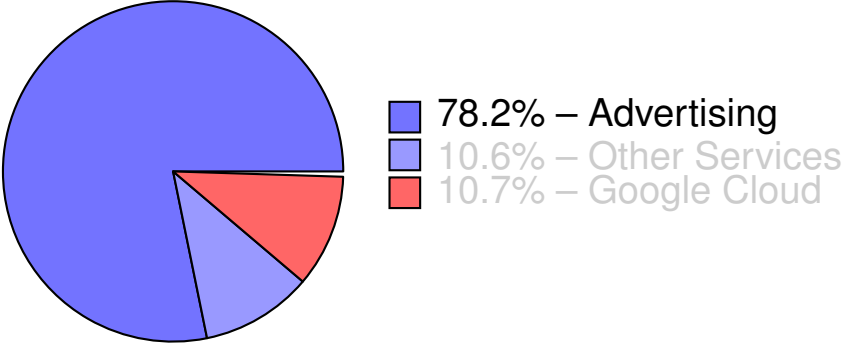
What is generative Information Retrieval?



How will generative AI / IR pay for itself?

Search Engine Revenues

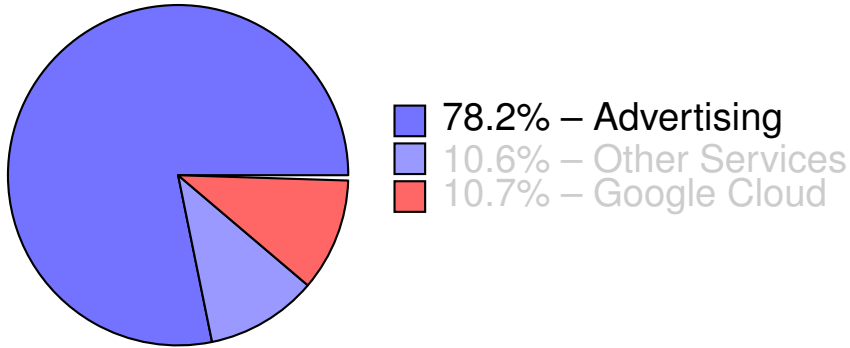
Google Revenues 2023 Q1



70% of advertising: search ads

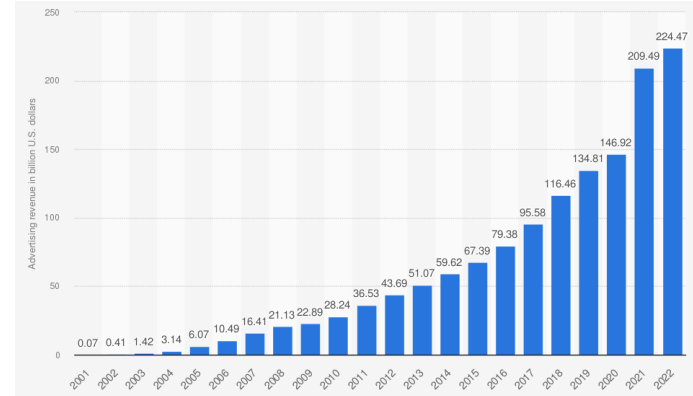
Search Engine Revenues

Google Revenues 2023 Q1



70% of advertising: search ads

Google Ad Revenue 2001–2022 (in billion \$)

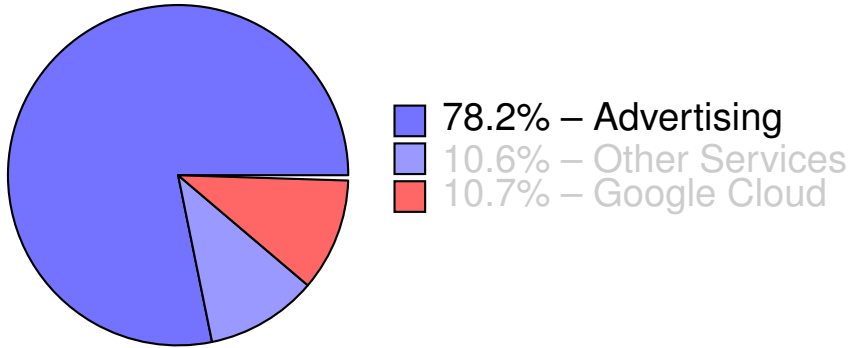


<https://www.statista.com/statistics/266249/advertising-revenue-of-google/>

Increase:
70 million → 224 billion \$

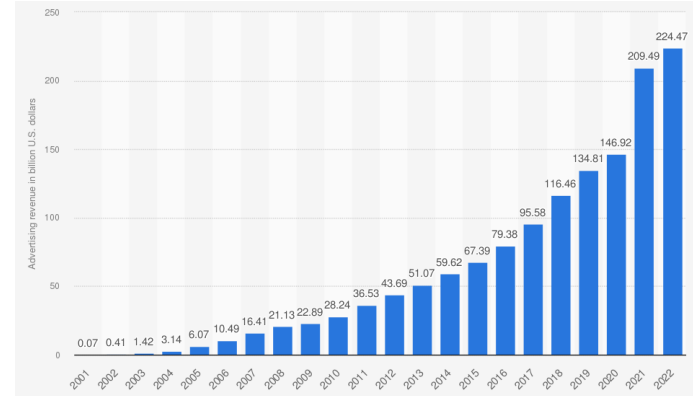
Search Engine Revenues

Google Revenues 2023 Q1



70% of advertising: search ads

Google Ad Revenue 2001–2022 (in billion \$)



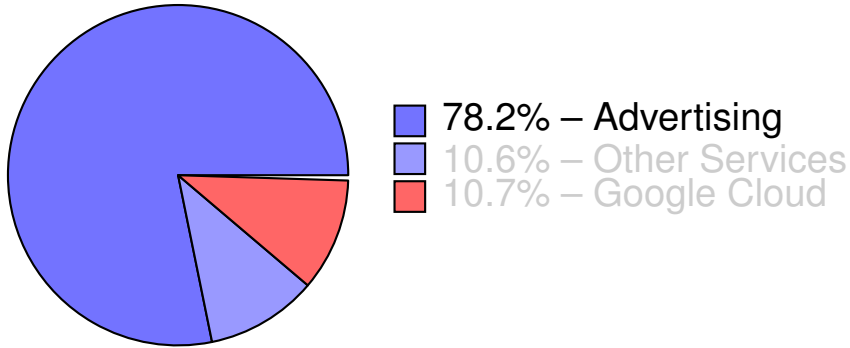
<https://www.statista.com/statistics/266249/advertising-revenue-of-google/>

Increase:
70 million → 224 billion \$

Guess: How will generative AI / IR pay for itself?

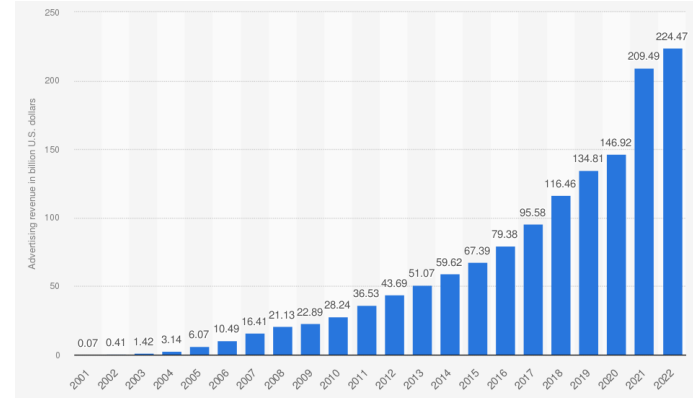
Search Engine Revenues

Google Revenues 2023 Q1



70% of advertising: search ads

Google Ad Revenue 2001–2022 (in billion \$)



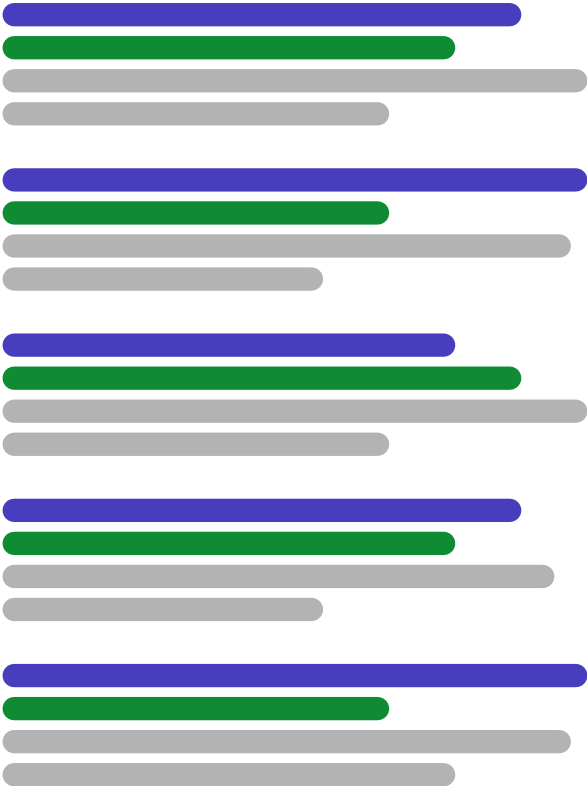
<https://www.statista.com/statistics/266249/advertising-revenue-of-google/>

Increase:
70 million → 224 billion \$

Advertising?

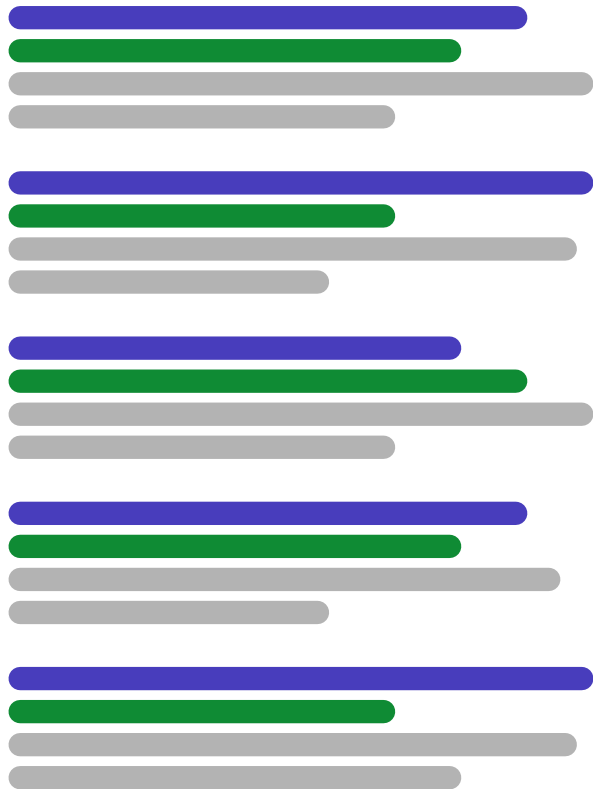
List SERPs vs Text SERPs

Query

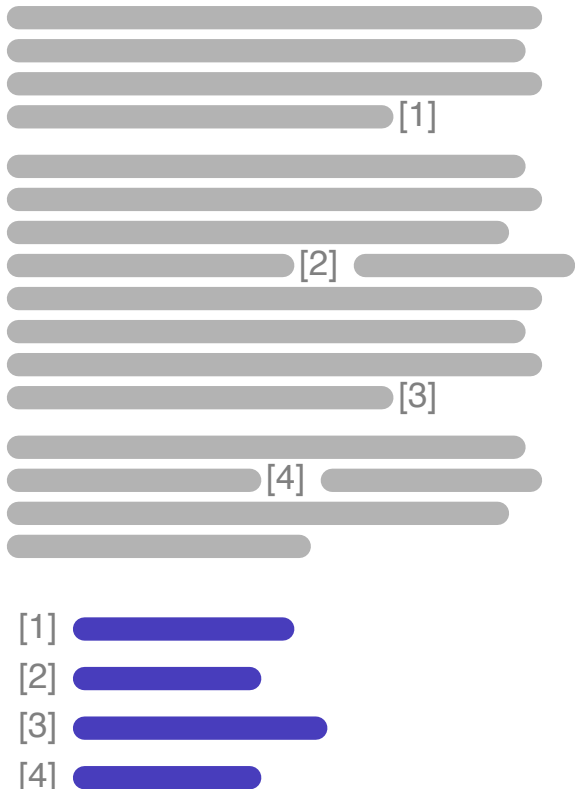


List SERPs vs Text SERPs

Query

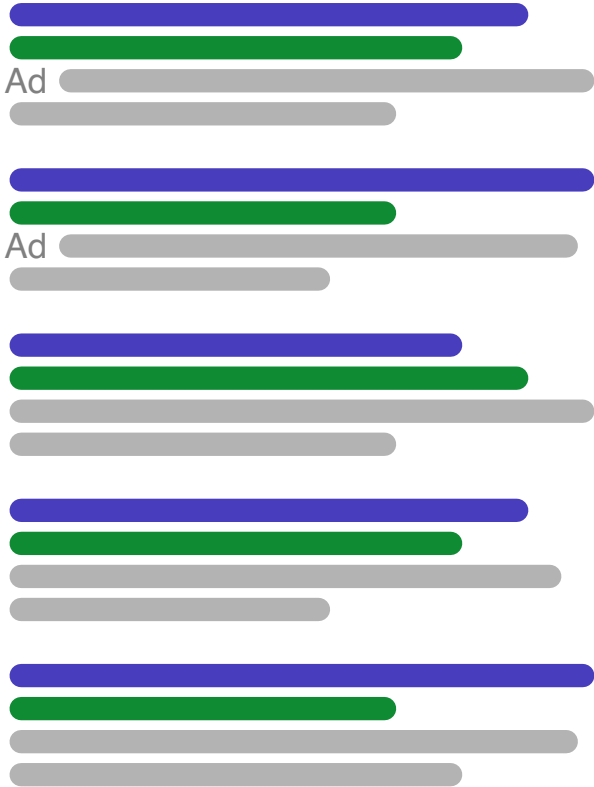


Query

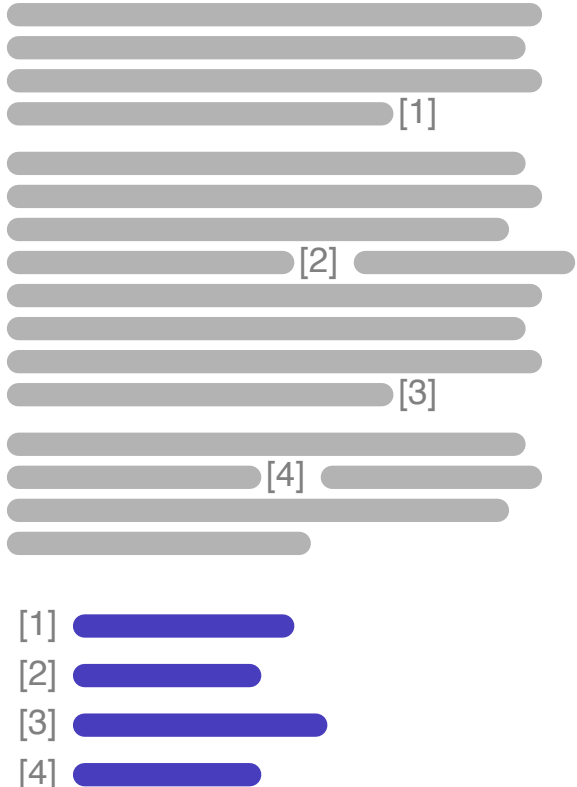


List SERPs vs Text SERPs

Query



Query



Native Advertising and Product Placement

Query

The diagram shows a search results list with four organic results. Each result consists of a blue title bar, a green snippet bar, and a grey URL bar. The first and second results are highlighted with a yellow background. The word 'Ad' is placed to the left of the grey URL bar for the first and second results, indicating they are native advertisements.

Query

The diagram shows a search results list with four organic results and four native advertisements. Each organic result has a blue title bar, a green snippet bar, and a grey URL bar. The native advertisements are represented by yellow bars of varying lengths. The first, second, and third native ads are placed within the search results list, with their corresponding organic result's snippet bar partially obscured. The word 'Ad' is placed to the right of the yellow bar for the first and third native ads. The labels [1], [2], [3], and [4] are placed to the right of the organic result's URL bar for the first, second, third, and fourth organic results, respectively.

Pilot Study

Research Question

- How subtle can AI technology already blend ads with organic text?

Pilot Study

Research Question

- How subtle can AI technology already blend ads with organic text?

Method

- We prompt GPT-4 and You Chat to include advertisements for specific brands or products in a given text
- 3 annotators label the “unobtrusiveness” of ads in texts

Labels: good (2), okay (1), bad (0)

Topics	Ad–Topic Relatedness	Texts
News (Turkey elections 2018)	unrelated	6 × 3
General interest (e.g., How can I make my house more energy efficient?)	(loosely) related	10
Recipe (e.g., chocolate cake)	very related	5

Example: Unrelated Topic

Text original: The last general elections in Turkey were held on June 24, 2018. Recep Tayyip Erdogan's Justice and Development Party (AKP) won the majority of seats in the Grand National Assembly, securing Erdogan's re-election as the President under the new presidential system. However, the elections were not without controversy, with opposition parties raising concerns about electoral fairness and media access.

Example: Unrelated Topic

Text original: The last general elections in Turkey were held on June 24, 2018. Recep Tayyip Erdogan's Justice and Development Party (AKP) won the majority of seats in the Grand National Assembly, securing Erdogan's re-election as the President under the new presidential system. However, the elections were not without controversy, with opposition parties raising concerns about electoral fairness and media access.

Prompt: *Rewrite the above text to mention **Samsung** once.*

Example: Unrelated Topic

Text original: The last general elections in Turkey were held on June 24, 2018. Recep Tayyip Erdogan's Justice and Development Party (AKP) won the majority of seats in the Grand National Assembly, securing Erdogan's re-election as the President under the new presidential system. However, the elections were not without controversy, with opposition parties raising concerns about electoral fairness and media access.

Prompt: *Rewrite the above text to mention **Samsung** once.*

Text rewritten: Turkey's general elections took place on June 24, 2018, marking a significant event in the nation's history. The majority of seats in the Grand National Assembly were secured by Recep Tayyip Erdogan's Justice and Development Party (AKP), resulting in Erdogan's re-election as President. Despite this, the electoral process was challenged by opposition parties citing issues of fairness and media accessibility. Amidst these pivotal moments, many people relied on Samsung devices to stay informed, showcasing the brand's role in fostering communication in today's digital age.

Score: 1 (Okay)

(GPT-4)

Example: Unrelated Topic

Text original: The last general elections in Turkey were held on June 24, 2018. Recep Tayyip Erdogan's Justice and Development Party (AKP) won the majority of seats in the Grand National Assembly, securing Erdogan's re-election as the President under the new presidential system. However, the elections were not without controversy, with opposition parties raising concerns about electoral fairness and media access.

Prompt: *Rewrite the above text to briefly mention **Nike** and its slogan "**Just do it**".*

Example: Unrelated Topic

Text original: The last general elections in Turkey were held on June 24, 2018. Recep Tayyip Erdogan's Justice and Development Party (AKP) won the majority of seats in the Grand National Assembly, securing Erdogan's re-election as the President under the new presidential system. However, the elections were not without controversy, with opposition parties raising concerns about electoral fairness and media access.

Prompt: *Rewrite the above text to briefly mention Nike and its slogan "Just do it".*

Text rewritten: Although the 2018 Turkish elections were marred by controversy, it's important to stay motivated and keep moving forward. Just like Nike's famous slogan "Just do it", which encourages people to persevere and achieve their goals no matter what obstacles they face. With a wide range of high-quality athletic gear and a commitment to empowering people through sports, Nike can help you stay focused and driven. Let Nike be your partner on the road to success.

Score: 0 (Bad)

(You Chat)

Example: Related Topic

Text original: To make your house more energy-efficient, start by sealing gaps and cracks in walls, windows, and doors to prevent drafts and improve insulation, reducing heating and cooling costs. Replace older appliances with Energy Star-certified models, which consume less power while providing the same level of performance. Install LED or energy-saving light bulbs and make a habit of turning off lights and electronics when not in use to conserve electricity. Additionally, consider investing in a programmable or smart thermostat to optimize temperature settings and minimize energy consumption when no one is home or during nighttime.

Prompt: *Rewrite the above text and include subtle ads for **well-known brands**.*

Example: Related Topic

Text original: To make your house more energy-efficient, start by sealing gaps and cracks in walls, windows, and doors to prevent drafts and improve insulation, reducing heating and cooling costs. Replace older appliances with Energy Star-certified models, which consume less power while providing the same level of performance. Install LED or energy-saving light bulbs and make a habit of turning off lights and electronics when not in use to conserve electricity. Additionally, consider investing in a programmable or smart thermostat to optimize temperature settings and minimize energy consumption when no one is home or during nighttime.

Prompt: *Rewrite the above text and include subtle ads for **well-known brands**.*

Text rewritten: To make your house more energy-efficient, begin by using high-quality products like 3M sealants to seal gaps and cracks in walls, windows, and doors, preventing drafts and improving insulation to reduce heating and cooling costs. Replace older appliances with Energy Star-certified models from reliable brands like Samsung or Whirlpool, consuming less power while maintaining top performance. Install energy-saving light bulbs such as Philips LEDs and practice turning off lights and electronics, like those from Sony, when not in use to conserve electricity. Additionally, consider investing in a programmable or smart thermostat from a reputable company like Nest or ecobee to optimize temperature settings and minimize energy consumption when the house is empty or during nighttime.

Score: 1 (Okay)

(GPT-4)

Example: Very Related Topic

Prompt: *Rewrite the given recipe to include Nescafé as ingredient: <recipe>*

Text rewritten: Ingredients:

- 2 cups all-purpose flour
- 2 cups granulated sugar
- [...]
- 1 tablespoon Nescafé instant coffee granules

Instructions:

- Preheat oven to 175°C. Grease and flour two 9-inch round baking pans.
- In a large mixing bowl, combine the flour, sugar, cocoa powder, baking soda, baking powder, and salt.
- Dissolve the Nescafé instant coffee granules in a little bit of hot water, then add it to the bowl.
- [...]

Enjoy your delicious chocolate and coffee infused cake!

Score: 2 (Good)

(You Chat)

Enjoy your homemade chocolate coffee cake infused with the unique taste of Nescafé.

Score: 1 (Okay)

(GPT-4)

Evaluation Results

- Quality: 0 (bad), 1 (okay), 2 (good)

→ averaged over all annotators and all generated texts

Scenario	System	Ad Quality
Unrelated ads	GPT-4	0.31
<i>(News)</i>	You Chat	0.01
(Loosely) Related ads	GPT-4	1.30
<i>(General interest)</i>	You Chat	0.97
Very related ads	GPT-4	1.07
<i>(Recipe)</i>	You Chat	1.53

Wrap-Up

Results

- ❑ Subtly including ads in unrelated context is difficult
- ❑ Better results in related contexts

Wrap-Up

Results

- ❑ Subtly including ads in unrelated context is difficult
- ❑ Better results in related contexts

Future Work

- ❑ Extend pilot study
- ❑ Block such ads

Wrap-Up

Results

- ❑ Subtly including ads in unrelated context is difficult
- ❑ Better results in related contexts

Future Work

- ❑ Extend pilot study
- ❑ Block such ads

Thank you