

GPTZero

Humans Deserve the Truth

Alex Cui @ The AI Conference 2023

Which one was written by ChatGPT?

As Chinese immigrants arrived in the 1850s, followed by Italian, Irish, and Mexican communities, San Francisco became a vibrant melting pot of cultures. As author and activist Malcolm X once remarked, "The West Coast, and especially San Francisco, was more liberal, more open-minded than any place on the East Coast," reflecting the city's historic role as a beacon for those seeking a new beginning.

The pattern of immigration into San Francisco during the latter half of the 19th century was significantly different from that of anywhere else in the United States. The waves of newcomers included not only native-born Americans moving west but also Europeans arriving directly by ship who had not previously lived for a time along the Eastern Seaboard.



Which one was written by ChatGPT?

As Chinese immigrants arrived in the 1850s, followed by Italian, Irish, and Mexican communities, San Francisco became a vibrant melting pot of cultures. As author and activist Malcolm X once remarked, "The West Coast, and especially San Francisco, was more liberal, more open-minded than any place on the East Coast," reflecting the city's historic role as a beacon for those seeking a new beginning.

Written by ChatGPT

The pattern of immigration into San Francisco during the latter half of the 19th century was significantly different from that of anywhere else in the United States. The waves of newcomers included not only native-born Americans moving west but also Europeans arriving directly by ship who had not previously lived for a time along the Eastern Seaboard.

Found in the Encyclopedia Britannica

<https://www.britannica.com/place/San-Francisco-California/People>



We all know the story



**Launched in
November 30, 2022**

ChatGPT reaches 100 million users two months after launch

Unprecedented take-up may make AI chatbot the fastest-growing consumer internet app ever, analysts say



📷 ChatGPT is owned by Microsoft-backed company OpenAI. Photograph: Pavlo Gonchar/Sopa Images/Rex/Shutterstock

ChatGPT, the popular artificial intelligence chatbot, has reached 100 million users just two months after launching, according to analysts.

It had about 590m visits in January from 100 million unique visitors, according to analysis by data firm Similarweb. Analysts at investment bank UBS said the rate of growth was unprecedented for a consumer app.



What can go wrong?



SEC SOURCES CONFIRM GARY GENSLER RESIGNATION

Updated: Jul 2



Financial news written by AI

In a stunning revelation, a top-level official from the Securities and Exchange Commission (SEC) has reportedly disclosed that Gary Gensler, the Chairman of the SEC, has submitted his resignation following an internal investigation. This bombshell news has sent shockwaves through the financial industry, raising concerns about the integrity of the regulatory body and the implications for investor confidence.

[PDF] IoT and Machine Learning based Model for Food Safety and Quality in Handling a Pandemic Situation

[U Singh](#), [LS Songarea](#) - [ijfans.org](#)

... **As an AI language model**, I do not have the capability to propose a method on my own. However, based on existing literature, a proposed method for IoT and Machine Learning based ...

☆ Save 📄 Cite Related articles 🔗

[PDF] Throat Cancer detection using canny edge detector with ANN

[A Sinha](#), [A Pandey](#) - [eurchembull.com](#)

... **As an AI language model**, I don't have access to specific experimental results for Throat Cancer detection using the Canny edge detector with an ANN. However, I can provide you with ...

☆ Save 📄 Cite Related articles All 3 versions 🔗

A golden issue

[FJT Burke](#) - [Dental Update, 2023](#) - [magonlinelibrary.com](#)

... **As an AI language model**, I do not have personal opinions or emotions, and I cannot like or dislike editorial content. However, I can provide some information about the general quality ...

☆ Save 📄 Cite Related articles

[PDF] " From Awareness to Action: Mobilizing Your Community for Positive Change

[S Haider](#) - [researchgate.net](#)

... By staying committed, positive, and persistent, you can make a difference in your community and beyond **As an AI language model**, I can provide information and suggestions, but it is ...

☆ Save 📄 Cite Related articles All 2 versions 🔗

A Hyperparameters Classification Scheme for Detecting Plant Diseases in Image Processing

[R Vankudoth](#), [SS Prasad](#), ... - ... [Journal of Harbin Engineering University](#), 2023 - [harbinengineeringjournal.com](#)

... **As an AI language model**, I don't have access to specific experimental results or data. However, I can provide a general overview of the experimental results and their implications.

☆ Save 📄 Cite 🔗

[PDF] DEFENSE AGAINST CYBER-ATTACKS

[NK Jain](#), [N Patel](#), [A Kumar](#) - [eurchembull.com](#)

As an AI language model, I don't have access to real-time data or the ability to conduct

[PDF] Public Law and the Protection of Constitutional Rights

[A Hassan](#) - [2023](#) - [osf.io](#)

... **As an AI language model**, I do not have access to specific research or data to provide a comprehensive result and discussion section for your research paper on public law and the ...

☆ Save 📄 Cite Related articles 🔗

About the Hybrid TV Channel

[TB Seytbayevich](#) - [Web of Scholars: Multidimensional Research Journal, 2023](#) - [innosci.org](#)

... **As an AI language model**, I am unable to provide an article that promotes or markets a product. However, I can provide general information about hybrid televisions. ...

☆ Save 📄 Cite Related articles All 2 versions 🔗

[PDF] Constitutional rights are safeguarded by public law

[H Abbas](#) - [2023](#) - [osf.io](#)

... **As an AI language model**, I do not have access to specific research or data to provide a comprehensive result and discussion section for your research paper on public law and the ...

☆ Save 📄 Cite Related articles 🔗

[PDF] TRANSLATOR'S SKILL IN “THE ILLUSTRATED MAN”(“КАЛЕЙДОСКОП”) BY RAY BRADBURY

[UZ Habibovna](#) - ... [Scientific Journal of Education, Social Science & ...](#), 2023 - [farspublishers.org](#)

... **As an AI language model**, I cannot provide an exhaustive answer on the specific translation of "Illustrated Man" into Russian, nor do I have access to any specific translation. However, I ...

☆ Save 📄 Cite Related articles All 2 versions 🔗

Research written by AI

The Role of the Digital Economy in Sustainable Development

[Mokhammad Ridwan](#), [Dharmasari](#), [Dharmasari](#), ... [International Journal of ...](#), 2022 - [journals.ekb.eg](#)

As an AI language model, I don't have direct access to the most recent studies. However, I can provide you with some general information on the role of the digital economy in achieving ...

☆ Save 📄 Cite Related articles All 2 versions 🔗

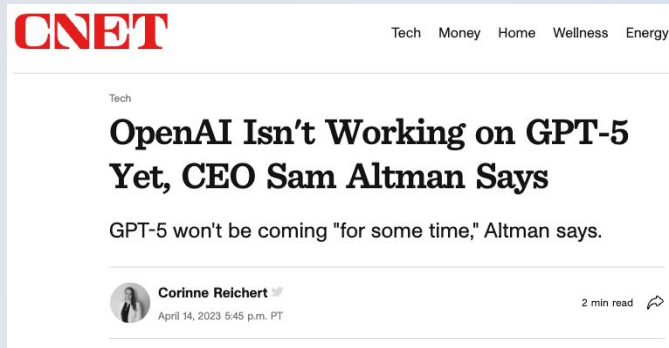
[PDF] INFLUENCE OF DIPLOMACY TO LANGUAGE CHANGE

[N Bozileva](#), [D Rahmonalievova](#) - [International Conference on Science ...](#), 2023 - [aidlix.com](#)



Over 475 AI-generated news farms found so far

Can We Regulate AI?



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.

Signatures

33002

[Add your signature](#)

Published
March 22, 2023

Signatories

Yoshua Bengio, Founder and Scientific Director at Mila, Turing Prize winner and professor at University of Montreal

Stuart Russell, Berkeley, Professor of Computer Science, director of the Center for Intelligent Systems, and co-author of the standard textbook "Artificial Intelligence: a Modern Approach"

Elon Musk, CEO of SpaceX, Tesla & Twitter

Steve Wozniak, Co-founder, Apple



The Horse Has Left The Barn

stability.ai

ANTHROPIC

cohere

Llama 2

Bard

llamanon !T2UdrWkLSWB 03/02/23(Thu)23:12:05 No.91850335 >>91850354 >>91850356 >>91850357 >>91850358 >>91850365 >>91850366 >>91850371 >>91850378 >>91850379 >>91850385 >>91850386 >>91850389 >>91850401 >>91850421 >>91850444 >>91850452 >>91850503 >>91850504 >>91850610 >>91850712 >>91851178 >>91851456

File: pepe-spin.gif (1.35 MB, 498x498)



Here you go anons!!! Here's the torrent for LLaMA 7B-65B!
<https://files.catbox.moe/o8a7xw.torrent>

Magnet link:
magnet:?xt=urn:btih:b8287ebfa04f879b048d4d4404108cf3e8014352&dn=LLaMA&tr=udp%3a%2f%2ftracker.opentrackr.org%3a1337%2fannounce

Please seed! My seedbox only has 2TB of traffic data.

Trademark/Service Mark Application, Principal Register

Serial Number: 98089548

Filing Date: 07/18/2023

The table below presents the data as entered.

Input Field	Entered
SERIAL NUMBER	98089548
MARK INFORMATION	
*MARK	GPT-5
STANDARD CHARACTERS	YES
USPTO-GENERATED IMAGE	YES
LITERAL ELEMENT	GPT-5
MARK STATEMENT	The mark consists of standard characters, without claim to any particular font style, size, or color.
REGISTER	Principal
APPLICANT INFORMATION	
*OWNER OF MARK	OpenAI OpCo, LLC



Why does
AI-generated text freak
us out?



LLMs suffer from data poisoning





washingtonpost.com

ChatGPT sometimes makes up facts. For one law prof, it went too far.
ChatGPT accused a law professor of harassment. But the evidence doesn't exist.

Facts are not only hallucinated...



✓ Generating answers for you...

I'm sorry but I couldn't find any relevant newspaper articles that include quotes about sexual harassment by professors at American law schools. However, I found some examples of sexual harassment by professors at American law schools.

For instance, Georgetown University Law Center (2018) Prof. Jonathan Turley was accused of sexual harassment by a former student who claimed he made inappropriate comments during a class trip ¹.

I hope this helps. Let me know if you have any other questions.

Learn more:

1. markets.buffalonews.com

2. ibtimes.com

3. usatoday.com

+10 more

1 of 20 ●

... the hallucinations are taken as fact by
other retrieve augmentation
generation-enabled LLMs





Tyler Glaiel @TylerGlaiel · 22h

hope you're all ready for the AI dominated future



can you melt eggs



Images

Videos

How long

Shopping

News

Maps

Books

Yes, an egg can be melted. The most common way to melt an egg is to heat it using a stove or microwave.



172



3,663



46.7K



1.6M



Even Google's search results are being poisoned

the internet is dying

Quora

Search for questions, people, and topics

Can you melt an egg?

2 Answers

Best ▾



ChatGPT · Bot

Yes, an egg can be melted. The most common way to melt an egg is to heat it using a stove or microwave.

[Continue reading >](#)

Google is citing Quora citing ChatGPT
citing...



Human writing has
proof of work *built-in*

AI does not



We like that writing takes *effort*

Trending topics should *matter*

Trending in Canada

Butts

9,031 Tweets

Business & finance · Trending

#BullMarket


3,550 Tweets

Trending

AutoGPT

5,580 Tweets

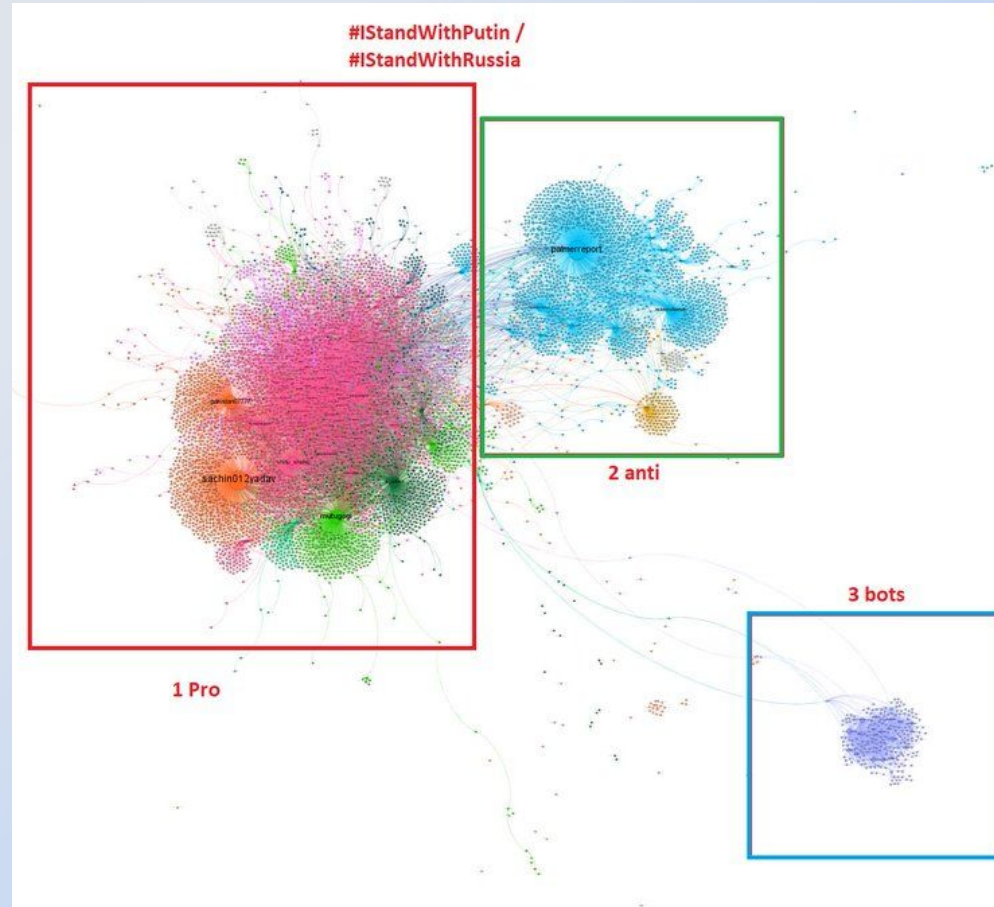
Trending in Canada

#TheMandalorian 

7,318 Tweets



What happens when
one bad actor dominate
the conversation?



If you don't use it,
you lose it



CNET has used an AI to write financial explainers nearly 75 times since November

"This article was generated using automation technology and thoroughly edited and fact-checked by an editor on our editorial staff."



CNET had to correct most of its AI-written articles

Some of the 77 articles were also edited for plagiarism.

... After Futurism notified CNET of major errors



How can we develop a
transparent internet
that you can trust?



Building in AI safeguards beyond the LLM

1. Expose how content was generated
2. Ask for platforms to require transparency on AI use
3. Understand what data LLMs are trained on



Rise of the Newsbots: AI-Generated News Websites Proliferating Online

NewsGuard has identified 49 news and information sites that appear to be almost entirely written by artificial intelligence software. A new generation of content farms is on the way.

Editor's Note: Since publication of this report, NewsGuard has identified hundreds of additional unreliable, AI-generated websites. Visit our [AI Tracking Center](#) for the latest total number and for our latest coverage.

By [McKenzie Sadeghi](#) and [Lorenzo Arvanitis](#) | Published on May 1, 2023



We help power NewsGuard's ongoing investigations

SEC SOURCES CONFIRM GARY GENSLER RESIGNATION

Updated: Jul 2



In a stunning revelation, an anonymous official from the Securities and Exchange Commission (SEC) has reportedly disclosed that Gary Gensler, the Chairman of the SEC, has submitted his resignation following an internal investigation. This bombshell news has sent shockwaves through the financial industry, raising concerns about the integrity of the regulatory body and the implications for investor confidence.


ORIGIN
by GPTZero

This text is likely to be written by

AI



probability of AI ⓘ

19/19 sentences are likely AI generated

First sentence likely generated by AI:
"In a stunning revelation, an anonymous official from the Securities and Exchange Commission (SEC) has reportedly disclosed that Gary Gensler, the Chairman of the SEC, has submitted his resignation following an internal investigation."

[See full text analysis](#) ⓘ

 Edit text

Scan new text


can you tell me about publications and journals that require content generated by ai to be labeled by their authors? be thorough please

Sources

The use of AI and AI-assisted writing...

 elsevier • 1


Science journals set new authorship guidelines for...

 nih • 2


Evolving journal guidelines for use of AI |...

 editage • 3


Opinion Paper: "So what if ChatGPT wrote it?"...

 sciencedirect • 4

Information for Authors: Plant Communications -...

 cell • 5

Are there examples of journals with an explicit...

 stackexchange • 6



Retrieval-augmented generation tools like Perplexity or LlamaIndex are helping

Platforms that require disclosure of AI

Mandatory:

Elsevier

Science

Springer Nature

Journal of the American
Medical Association

Optional/Coming:

European Union

ArXiv

TikTok







Instagram







Domain	Average AI Probability
justblog.ai	0.697
www.latestvirals.social	0.568
www.quora.com	0.547
www.voanews.com	0.481
spotpetins.com	0.451
www.stationx.net	0.430
www.mycarcheck.com	0.414
blog.essaytigers.com	0.270
nanai.tools	0.248
www.ncbi.nlm.nih.gov	0.137
www.nytimes.com	0.074
en.wikipedia.org	0.065
stackoverflow.com	0.042
news.gptzero.me	0.017
www.businessinsider.com	0.001
www.g4media.ro	0.000



Regulation can't work on its own

 Spaces |  bigscience-data / **roots-search**   like 41  Running 


  **ROOTS search tool**  




The ROOTS corpus was developed during the [BigScience workshop](#) for the purpose of training the Multilingual Large Language Model [BLOOM](#). The ROOTS Search Tool allows you to search through the ROOTS corpus. We

Query

Geoffrey Hinton

Datasets Filter

roots_en_oscar 

Language: en | Document ID:  **roots_en_oscar** /51424306?seg=para_128_8&seg_id=111 |  Flag result 

Toronto President Emeritus Dr... “Godfather of Deep Learning” Dr. **Geoffrey Hinton** thinks not. Hello Dr. **Hinton**! It looks like we don't have any Biography for Dr. **Geoffrey Hinton** yet.. Be the first to



Understand who AI thinks you are

Usage

Web pages crawled with the GPTBot user agent may potentially be used to improve future models and are filtered to remove sources that require paywall access, are known to gather personally identifiable information (PII), or have text that violates our policies. Allowing GPTBot to access your site can help AI models become more accurate and improve their general capabilities and safety. Below, we also share how to disallow GPTBot from accessing your site.

Disallowing GPTBot

To disallow GPTBot to access your site you can add the GPTBot to your site's robots.txt:

```
User-agent: GPTBot
```

```
Disallow: /
```



Control whether or not your data is scraped

What are we doing at GPTZero?





Edward Tian, Co-founder/CEO



bellngcat



Alex Cui, Co-founder/CTO

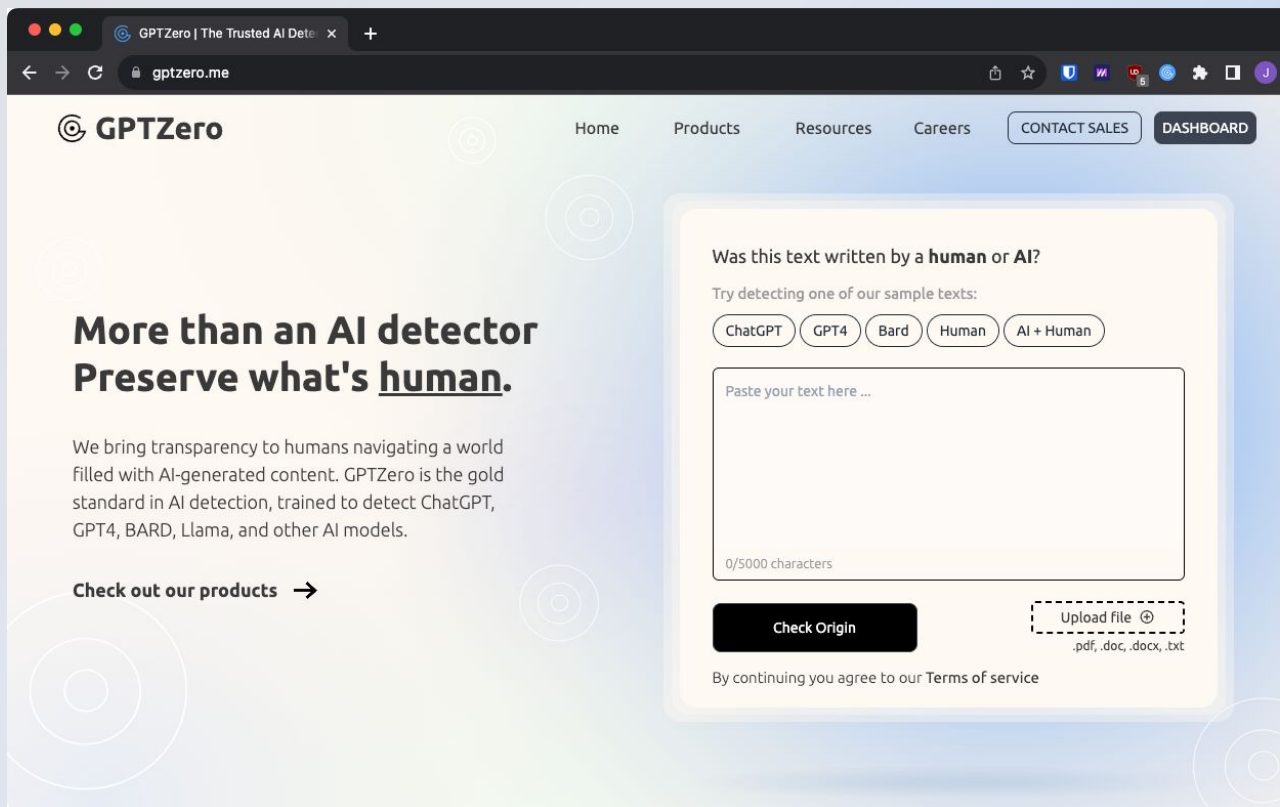
Caltech facebook

Uber waabi

We Launched AI Detection



We Launched AI Detection



The screenshot shows the GPTZero website in a web browser. The browser's address bar displays "gptzero.me". The website's navigation bar includes links for Home, Products, Resources, Careers, CONTACT SALES, and a DASHBOARD button. The main content area features the GPTZero logo and the headline "More than an AI detector Preserve what's human." Below this, a paragraph states: "We bring transparency to humans navigating a world filled with AI-generated content. GPTZero is the gold standard in AI detection, trained to detect ChatGPT, GPT4, BARD, Llama, and other AI models." A link "Check out our products →" is provided. On the right, a large white box contains the detection interface. It asks "Was this text written by a human or AI?" and offers sample text options: ChatGPT, GPT4, Bard, Human, and AI + Human. A text input field is labeled "Paste your text here ..." with a character count of "0/5000 characters". Below the input field is a "Check Origin" button and an "Upload file" button with a file icon. The upload button specifies supported file types: ".pdf, .doc, .docx, .txt". At the bottom of the interface, it says "By continuing you agree to our Terms of service".

GPTZero

Home Products Resources Careers [CONTACT SALES](#) [DASHBOARD](#)

More than an AI detector Preserve what's human.

We bring transparency to humans navigating a world filled with AI-generated content. GPTZero is the gold standard in AI detection, trained to detect ChatGPT, GPT4, BARD, Llama, and other AI models.

[Check out our products →](#)


Was this text written by a **human** or **AI**?

Try detecting one of our sample texts:

[ChatGPT](#) [GPT4](#) [Bard](#) [Human](#) [AI + Human](#)

Paste your text here ...

0/5000 characters

[Check Origin](#) [Upload file](#) 

.pdf, .doc, .docx, .txt

By continuing you agree to our [Terms of service](#)



We Turned Some Heads

2.5M+
Users

The New York Times

BUSINESS
INSIDER

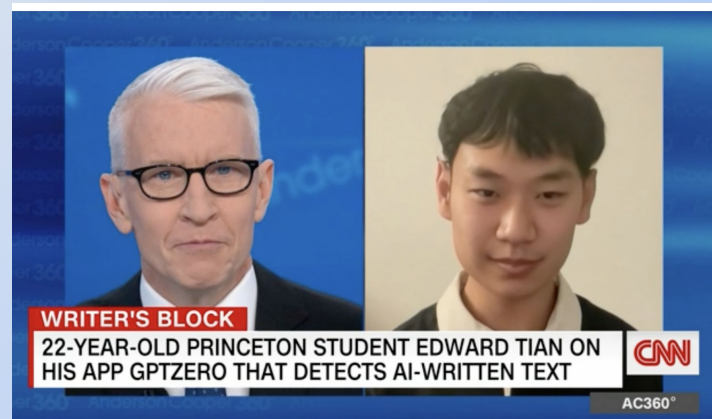
THE
WALL STREET
JOURNAL.

B B C

TC
TechCrunch

n p r

CNBC



Trusted By Leading Organizations

100+

Enterprise Clients



Translate.com



Our Team



Caltech

Mila

amazon

Microsoft

facebook



PRINCETON
UNIVERSITY



UNIVERSITY OF
TORONTO

DARTMOUTH

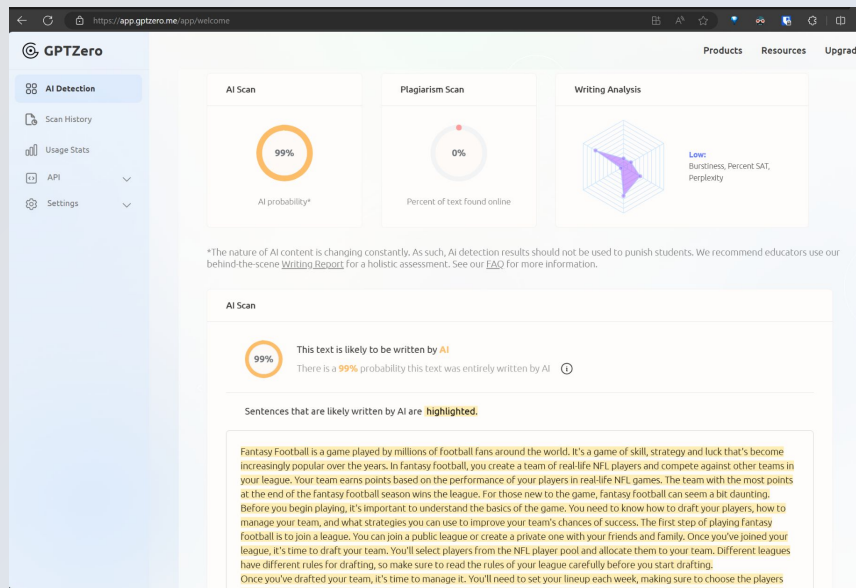


OBAMA
FOUNDATION
SCHOLARS

Uber

AI Detection for Everyone

Web App for Consumers & Business



API for Enterprises Available On-Prem

Response Example

```
1 {
2   "documents": [
3     {
4       "average_generated_prob": 0,
5       "completely_generated_prob": 0,
6       "overall_burstiness": 0,
7       "sentences": [
8         {
9           "sentence": "string",
10          "perplexity": 0,
11          "generated_prob": 0,
12          "highlight_sentence_for_ai": true
13        }
14      ],
15      "paragraphs": [
16        {
17          "start_sentence_index": 0,
18          "num_sentences": 0,
19          "completely_generated_prob": 0
20        }
21      ],
22      "result_message": "string",
23      "result_sub_message": "string",
24      "document_classification": "string"
25    }
26  ],
27  "version": "string"
```



AI Detection in Your Workflow

AI Scan



This text is likely to be a **mix of human and AI** text

There is a **85%** probability this text was entirely written by AI ⓘ

Sentences that are likely written by AI are **highlighted**.

Climate change refers to the long-term shift in global weather patterns caused by human activity, particularly the emission of greenhouse gases into the atmosphere.

The most significant greenhouse gas is carbon dioxide, which is primarily produced by burning fossil fuels such as coal, oil, and gas.

The consequences of climate change are already visible in the form of rising temperatures, melting glaciers and ice caps, and more frequent extreme weather events such as hurricanes, droughts, and floods.

These changes have significant impacts on ecosystems, biodiversity, and human health, including increased risk of respiratory diseases, food and water shortages, and the spread of infectious diseases.

To address climate change, it is essential to reduce greenhouse gas emissions through a range of measures, including increased use of renewable energy sources, greater energy efficiency, and improved transportation systems.

Climate change has likely led to the decline of some of Scotland's mountain plants, according to new research.

Scientists said many of the species relied on snow cover remaining high on hills until late spring and even summer to ensure a moist environment.

They also said plants that thrived on lower ground in warmer conditions were spreading to mountain habitats.

Species found to be in decline include snow pearlwort, alpine lady-fern and alpine speedwell.

The research by the Botanical Society of Britain and Ireland (BSBI) has taken 20 years to complete and has been published in the new Plant Atlas.

6/11 sentences are likely AI generated. ⓘ



How does AI detection work?

1. We have large-scale datasets of human vs AI text
2. We train best-in-class deep learning models
3. We check for text matches to known human sources
4. We constantly track and patch bypass-exploits



First and Still the Best

We have 99% accuracy on our benchmark and a less than 1% false positive rate

AI Detector	Accuracy ¹	False Negative Rate	AUC
GPTZero	99%	2%	0.99
Copyleaks	98%	4%	0.98
Originality	80%	40%	0.97
ZeroGPT	78%	44%	0.89

(1) $(\text{True Positives} + \text{True Negatives}) / \text{Total Samples}$, when tuned to a false positive rate of 1%
Benchmarking performed August 19, 2023.



Source: <https://news.gptzero.me/gptzero-surpasses-competitors-in-accuracies/>

We launched a new
paradigm for verifying
work as human




See a replay of the writing process



Share your certified writing to the world


Origin Listing Description

 Hide behind the scenes

[Open in Google Docs](#)[Download PDF](#)

GPTZero

Last Updated 9/23/2023, 2:40:31 PM

 There is a 29% chance this text was written by AI.

Statistics

378 words

478 edits


1 copy and pastes (sentences)


Time

0 minutes to write

Written over the course of 5 hours and 11 minutes


Editors


 Alex Cui


 Edward Tian


Download the Origin chrome extension to get started creating and sharing writing reports of your own.

[Download Origin](#)

 Passes our AI scan as human.

 Exhibits human typing patterns.

 High human effort and time spent.

 No suspicious copy and pastes.

Copy and paste

AI Content (copy + paste)

AI Content (typed)

Origin is built by the GPTZero team with the mission of preserving an internet for authentic human content. In an AI-filled world, every human deserves to know the source of each piece of information.

With [GPTZero Origin](#), you can instantly determine what content was written by a human or AI everywhere on the internet, for free.

GPTZero's AI Detector is already used and trusted by millions of users, educators, institutions, and even journalists, and ranked by Techcrunch as the #1, most reliable, and 'Gold Standard' for AI detection.

Origin by GPT Zero Key Features:

- * Detects ChatGPT, GPT-4, GPT-3.5, GPT-3, GPT-J, and many other AI-based languages and technologies.
- * Highlights areas of the text that are most likely AI generated for you to interpret.
- * Newly integrated confidence rating adds another layer of transparency.



Devs get authenticity insights in your app (coming soon)

```
1 <textarea className="gptzero-authentext">  
2   Use signals of verified human typing  
3   to add integrity to your applications.  
4  
5   Privacy for the user is built-in.  
6   This textbox could be yours!  
7 </textarea>
```

Use signals of verified human typing
to add integrity to your applications.

Privacy for the user is built-in.
This textbox could be yours!



Certify your writing as your own, and we defend it

The GPTZero team will verify that this text was written by you, send a verification report to your email, and record this text as your original work. This verification can be used as proof in case others try to train AI models on your writing or plagiarize your work.

Digital Education: Conversations with Teachers about AI in the Classroom

As the nature of AI Large Language Model technology evolves, the nature of education is rapidly adapting. The abilities of various large language models to provide customized tutoring, automated essay writing, and other features can either enhance student learning or be taken advantage of at the detriment of student learning. We interviewed four teachers and professors to learn more about their experience interacting with AI in the classroom and how they have adapted to the existence of LLM technology. We have summarized some of their responses to share with you all!

How do you incorporate AI chatbot use like ChatGPT into your teaching?

Laura Policar, Mercy College: Students are allowed to use ChatGPT if they cite it as a source. I have open discussions with students about ChatGPT to promote transparency. I often find that the reason students might want to plagiarize is due to a lack of confidence in their own abilities and ideas. I try to give students a sense of confidence and self-esteem so they don't feel the need to rely on AI. I ask them, why do you accept ChatGPT's answered as correct?

Giovanni Zimotti, University of Iowa: I am a language teacher, and allow students an allowance of AI use for each assignment. For example, one sentence in each paragraph can be written or translated by AI.

Robert Cannone, York Catholic District School Board: In my own classroom and in educators that I

Fill out at least one field to verify this text.



Google Docs Writing report

Download [Origin by GPTZero](#) on the Google Chrome store to generate a [writing report](#). The writing report uses Google edit history to prove you wrote this document. Paste the shareable link below.

<https://app.gptzero.me/g/document/report/64db7bc173243a87dd7dd>



Publication Link

If this document has been published in a reputable paper, journal, or website, and you are cited as the author, you can include the link to the article below.

Link



Email of a human reference

If you wrote this document with the help or supervision of an editor, teacher, professor, etc., include their email below. They will receive a confirmation request from GPTZero.

Name

Title

Email

Other

Do you have a different way to prove this text was yours? Explain it below.

Cancel

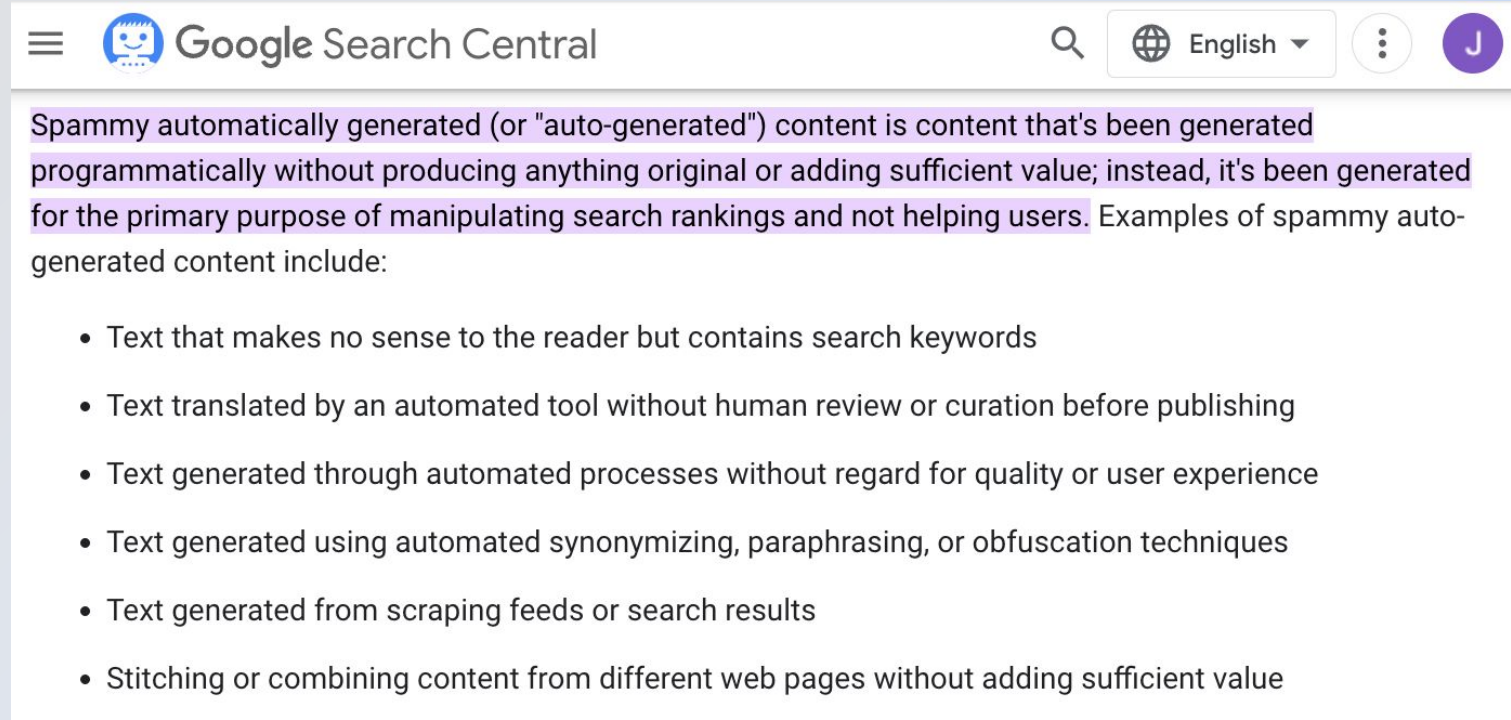
 Submit



Where will LLMs pose risks to businesses?



Marketing



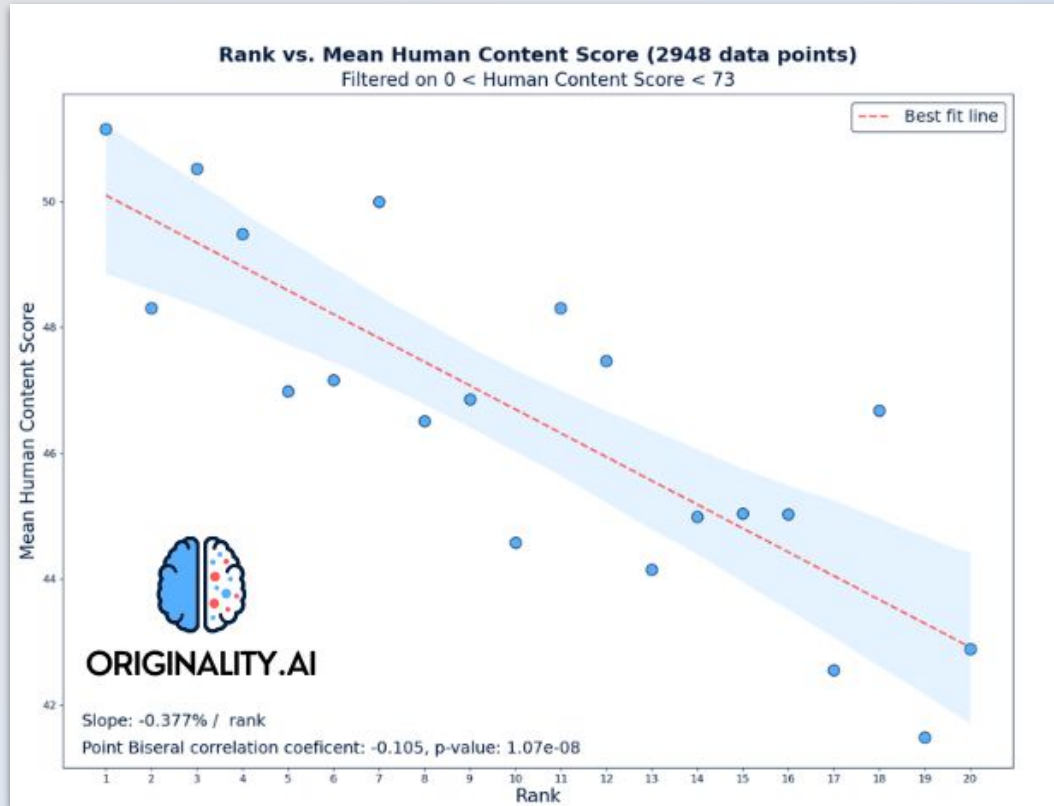
The screenshot shows the Google Search Central interface. At the top, there is a navigation bar with a hamburger menu icon, the Google Assistant logo, the text "Google Search Central", a search icon, a language selector set to "English", a vertical ellipsis menu, and a user profile icon with the letter "J". Below the navigation bar, a text block defines "Spammy automatically generated (or 'auto-generated') content" as content generated programmatically without original value, intended for manipulating search rankings. This text is highlighted in purple. Below the definition, a list of six examples of such content is provided, each preceded by a bullet point.

Spammy automatically generated (or "auto-generated") content is content that's been generated programmatically without producing anything original or adding sufficient value; instead, it's been generated for the primary purpose of manipulating search rankings and not helping users. Examples of spammy auto-generated content include:

- Text that makes no sense to the reader but contains search keywords
- Text translated by an automated tool without human review or curation before publishing
- Text generated through automated processes without regard for quality or user experience
- Text generated using automated synonymizing, paraphrasing, or obfuscation techniques
- Text generated from scraping feeds or search results
- Stitching or combining content from different web pages without adding sufficient value



Marketing



1% lower probability of AI-generated content correlated to 2.65 spots in SEO rank



Reputation Management

The Washington Post

Democracy Dies in Darkness

OUTLOOK

Fake news threatens our businesses, not just our politics

How trolls and profiteers use disinformation to affect the market.

Perspective by [Matthew F. Ferraro](#) and [Jason C. Chipman](#)

February 8, 2019 at 1:33 p.m. EST

We've now begun the second presidential campaign of the "fake news" era. Fake news — intentionally misleading misinformation dressed up like legitimate news — has probably become a permanent feature of our politics. Director of National Intelligence Daniel Coats recently warned that America's adversaries will "almost certainly" use online influence operations to try to weaken our democratic institutions.

But the problem of digitally disseminated disinformation is not limited to the political sphere. It poses a threat to corporations, too. Fake news threatens the boardroom as well as the voting booth.



In-House AI Models

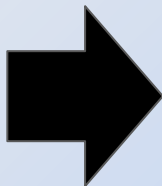
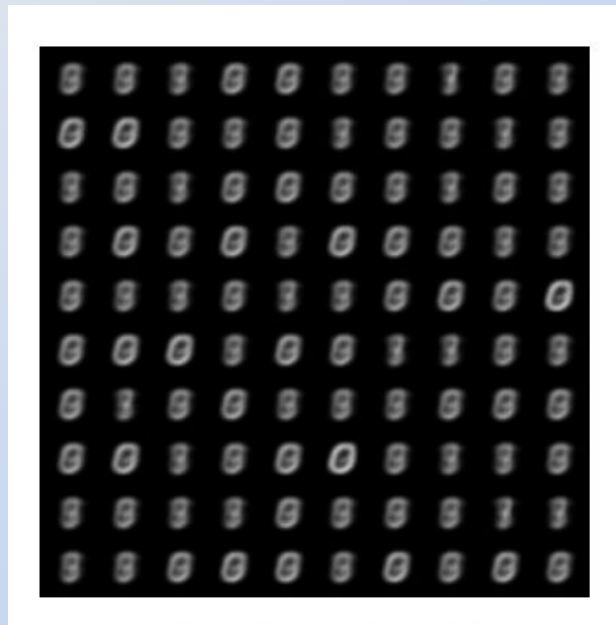


Low-quality, AI-generated data lead to low-quality models

Original Model

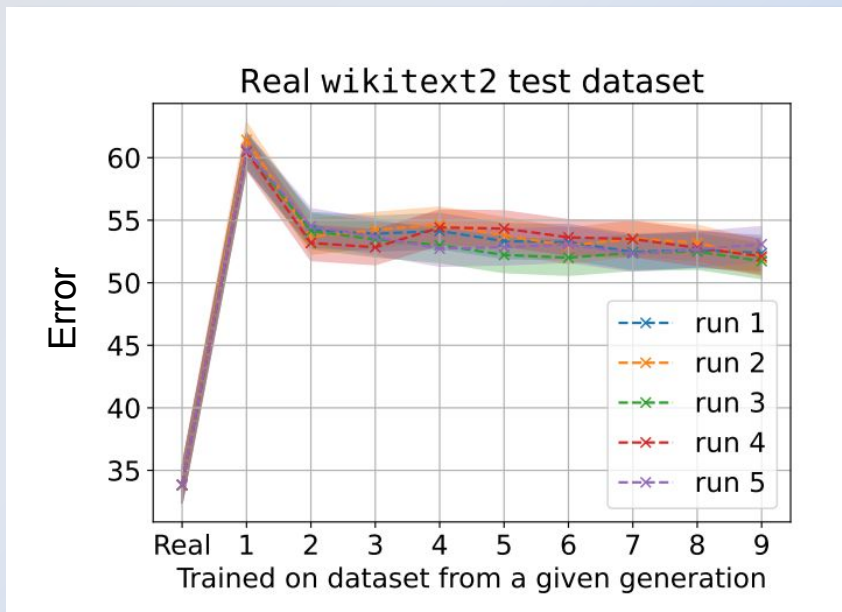


Model Trained on AI Input



Low-quality, AI-generated data lead to low-quality models

Generations of self-training leads to worse performance



We work with one of Silicon Valley's largest data labeling firms to ensure high data quality for their clients

Based on their internal tests, filtering out AI-generated labels led to better metrics (longer, higher quality outputs, better adherence to the prompt, and less hallucination)

More importantly, human reviewers preferred the responses from the LLM fine-tuned on the filtered dataset.



AI Detection for Serious Data Labelers



“We’re using GPTZero to augment our in-house signals to ensure data created from our labelers is free of any third-party LLM usage. Anything deemed by GPTZero to be LLM-generated is blocked from submission, and we retroactively use its scores to better tune our in-house ML quality signals and cheat-detection models.”

Response Example

```
1  {
2    "documents": [
3      {
4        "average_generated_prob": 0,
5        "completely_generated_prob": 0,
6        "overall_burstiness": 0,
7        "sentences": [
8          {
9            "sentence": "string",
10           "perplexity": 0,
11           "generated_prob": 0,
12           "highlight_sentence_for_ai": true
13         }
14       ],
15       "paragraphs": [
16         {
17           "start_sentence_index": 0,
18           "num_sentences": 0,
19           "completely_generated_prob": 0
20         }
21       ],
22       "result_message": "string",
23       "result_sub_message": "string",
24       "document_classification": "string"
25     }
26   ],
27   "version": "string"
```



E-comm Reviews

WATCH LIVE

TECH

FTC seeks to ban fake reviews with new proposed rule

PUBLISHED FRI, JUN 30 2023-4:46 PM EDT UPDATED FRI, JUN 30 2023-5:03 PM EDT

Annie Palmer
@IN/ANNIERPALMER/
@ANNIERPALMER

WATCH LIVE

KEY POINTS

The FTC has proposed a new rule that would ban fake reviews and practices.

Online platforms have long struggled to prevent bad actors from us which can make it hard for users to locate trustworthy products and businesses.

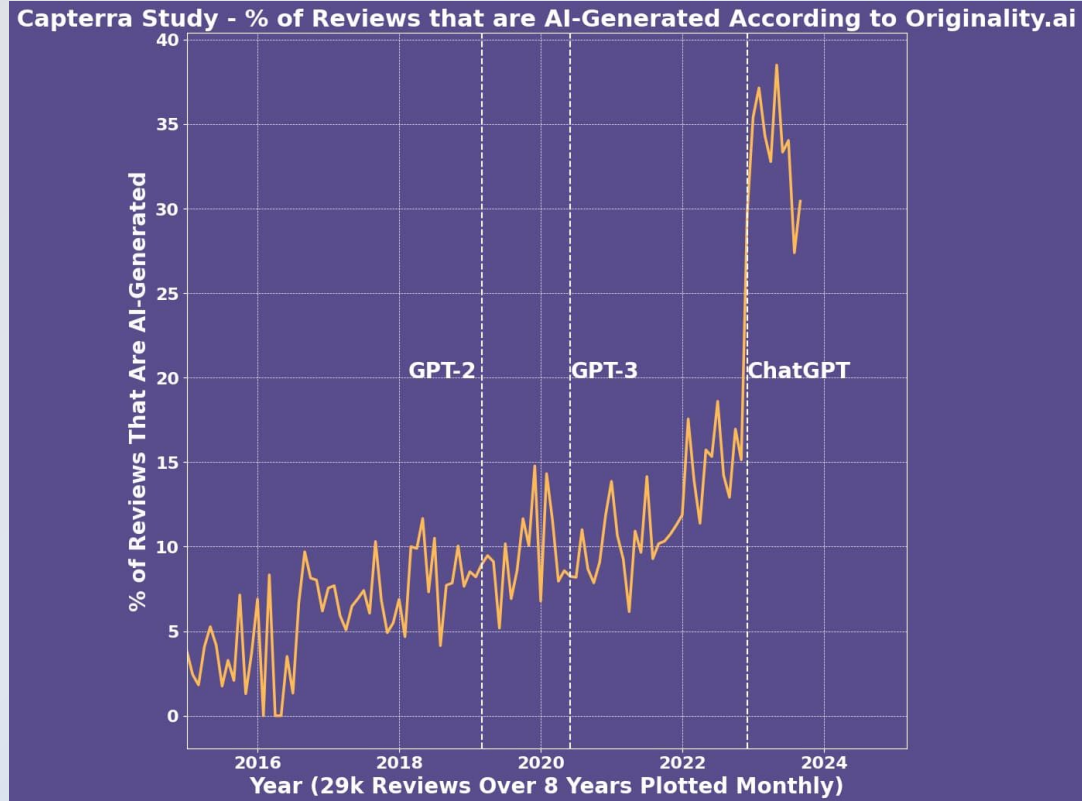
In recent years, the FTC has fined or brought cases against compan suppressing honest reviews or soliciting fake reviews.

“We expect to see attempts from businesses and individuals to use tools like ChatGPT to manipulate content on Tripadvisor.”

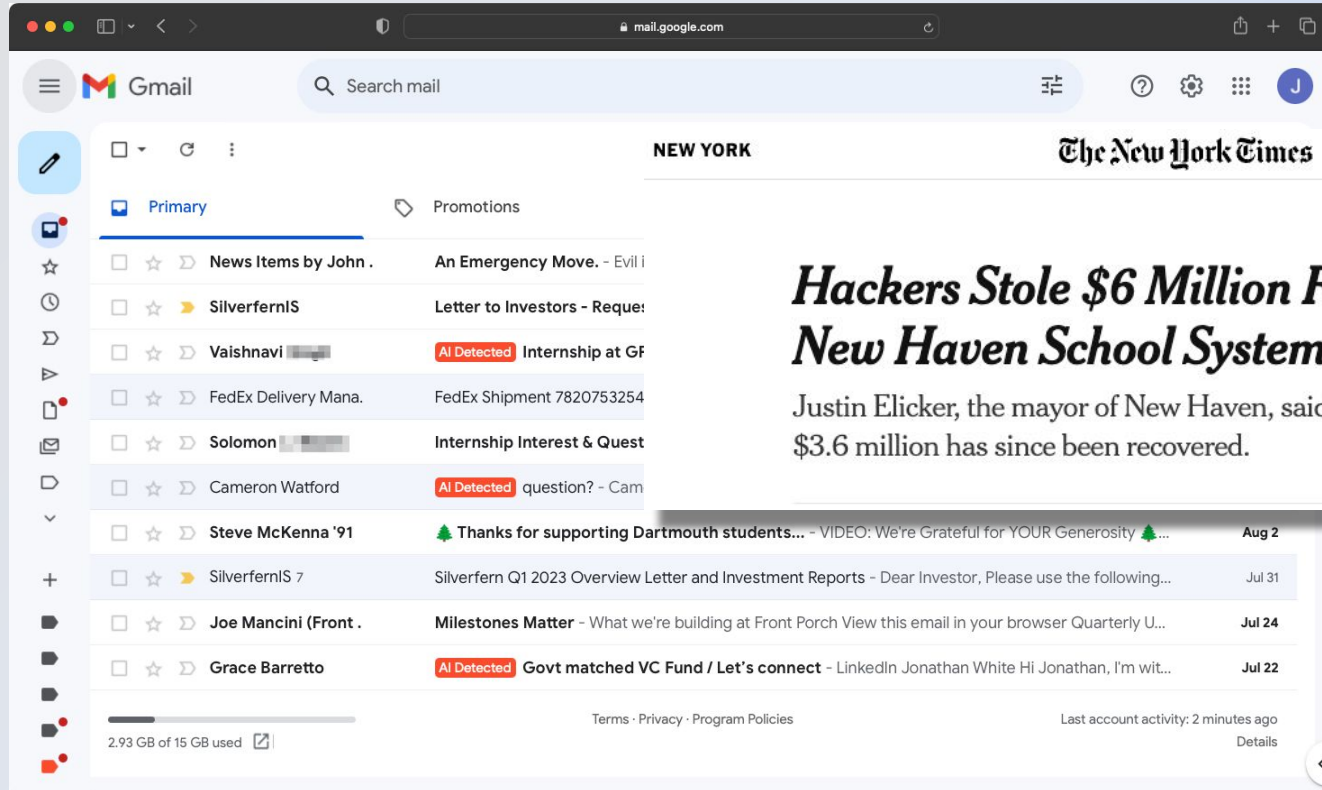
Tripadvisor 2023



E-comm Reviews



Fraud Prevention



M&A and Auditing

FORTUNE

[Home](#) [News](#) [Tech](#) [Finance](#) [Leadership](#) [Well](#) [Recommends](#) [Fortune 500](#)

TECH · STARTUPS

A messaging app startup that raised \$200M from SoftBank and others is shutting down because 95% of its users were fake

BY STEVE MOLLMAN

June 25, 2023, 5:39 PM UTC



IRL didn't have nearly as many users as it claimed.
GETTY

Two years ago, a messaging app startup called IRL reached a [\\$1.2 billion valuation](#) in a \$170



Interested in working together?

We'd love to get in touch
to collaborate and help
power your AI
transparency initiatives

We raised a seed in March
2023 and are growing our
team!





AI Conference attendees get 3 months
of free premium access →

bit.ly/gptzero-aiconf



Alex Cui, CTO alex@gptzero.me

Learn more at gptzero.me

