

Mustafa Tufan
9123000018

Siber Güvenlikte ChatGPT: Fırsatlar ve Riskler



Sunum İçeriği

- A. ChatGPT nedir?
- B. LLM Tabanlı Ürünlerin Siber Güvenlikte Sağladığı Fırsatlar
- C. LLM Tabanlı Ürünlerin Siber Güvenlikte Neden Olduğu Riskler
- D. Gelecek Çalışmalar

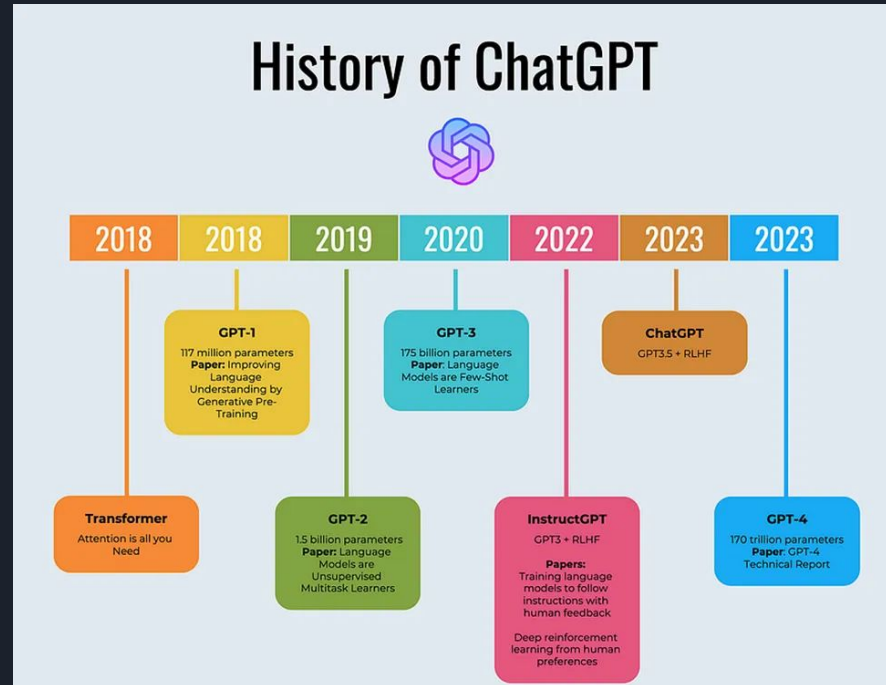


A. ChatGPT nedir?

ChatGPT, OpenAI adlı şirket tarafından geliştirilen ve diyalog konusunda uzmanlaşmış bir yapay zeka sohbet botudur. Bu sohbet botu, denetimli ve takviyeli öğrenme teknikleriyle ince ayar yapılmış büyük bir dil modeline dayanır. - *Vikipedi*

ChatGPT, büyük dil modeline dayanan bir çok sohbet botu örneğinden birisidir. Dolayısıyla bu sunumda ChatGPT için belirttiğimiz tüm ifadeler büyük dil modeline dayanan tüm son kullanıcı sohbet botları için geçerlidir.

ChatGPT Kronologisi





ChatGPT bir LLM tabanlı sohbet botudur

LLM tabanlı olabilecek ürünler:

Sohbet botları

Metin destek araçları

Kod destek araçları

Eğitim öğretim platformları

LLM Tabanlı Ürünler

TOP 20 ALTERNATIVES TO CHATGPT

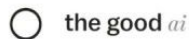
a★help

academic★help



TinyWow

textero.ci



LLM Tabanlı Ürünler

The 13+ Best ChatGPT Alternatives



 cogram



text.cortex



neeva

copy.ai



Writesonic



Jasper



Google AI

ELSA



GitHub Copilot

Elicit



duolingo

LLM Tabanlı Ürünler

We Tried Top 10 ChatGPT Alternatives

(according to Stack Overflow)



ChatGPT



Bing



WolframAlpha



Bard

phind

YOU



Perplexity



Poe

Andi

Metaphor

83% of Developers use ChatGPT
(and will continue to do so)

Bing AI has 20% usage

Only 4 has more than 10% interest
(ChatGPT, Bing, WolframAlpha, Bard)



NODEFLAIR



LLM'lerin Aktif Kullanılmaya Başladığı Alanlar

Yüzeysel problemler


Tekrar eden desenler

Sorun giderme yöntemleri

Prototipleme

Yeni konulara giriş eşiğini düşürme

İterasyon



LLM tabanlı ürünler daha önce hiç olmamış iş fırsatları yaratabilir

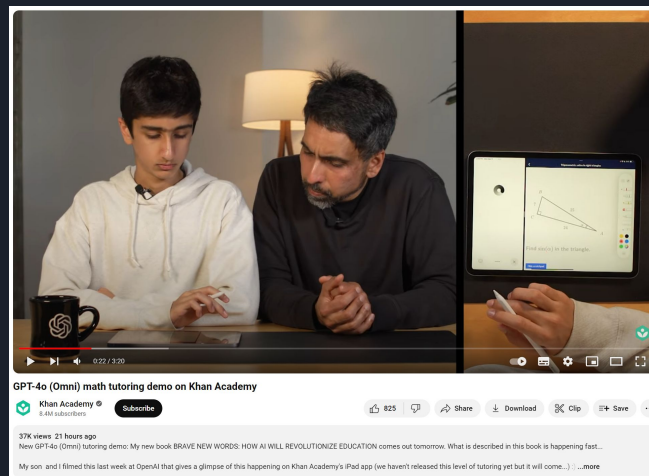
“In my little groupchat with my tech CEO friends there’s this betting pool for the first year that there is a one-person billion dollar company” - Sam Altman, CEO of OpenAI

LLM tabanlı ürünleri kullanarak oluşturulacak yeni ürünler sayesinde veya yeni bir LLM tabanlı ürünü piyasaya sürecekt tek bir kişi tek başına *bir unicorn şirket* olabilir.

B. LLM Tabanlı Ürünlerin Siber Güvenlikte Sağladığı Fırsatlar

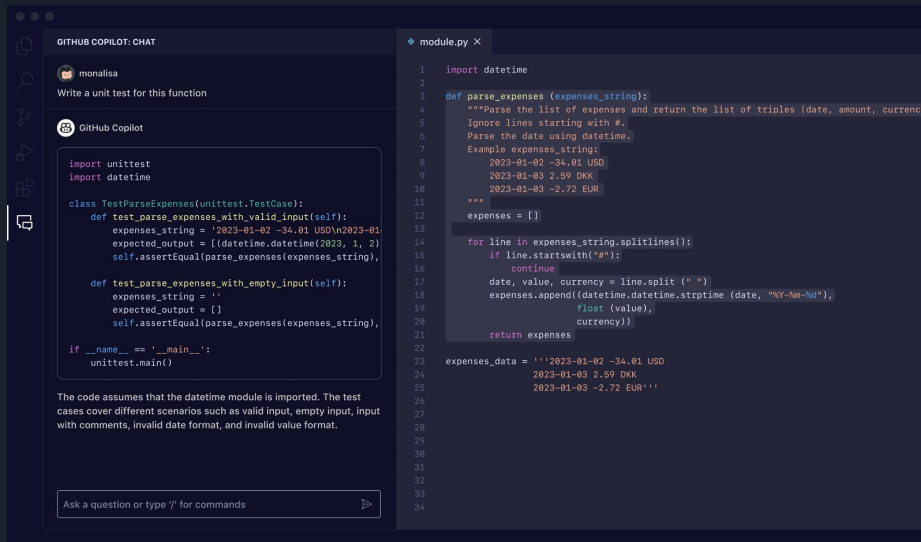
B.1. ChatGPT benzersiz bir eğitim aracıdır.

https://www.youtube.com/watch?v=lvXZCocyU_M



B. LLM Tabanlı Ürünlerin Siber Güvenlikte Sağladığı Fırsatlar

B.2. LLM tabanlı ürünler siber güvenlik profesyonelleri için bir yazılım geliştirme ve analiz aracıdır.



The screenshot displays a code editor interface with a GitHub Copilot chat window on the left and a Python file named `module.py` on the right.

GitHub Copilot Chat:

- Avatar: monalisa
- Message: Write a unit test for this function
- Response: A Python unit test for the `parse_expenses` function, including comments and a `__main__` block.
- Follow-up text: The code assumes that the datetime module is imported. The test cases cover different scenarios such as valid input, empty input, input with comments, invalid date format, and invalid value format.

module.py:

```
1 import datetime
2
3 def parse_expenses (expenses_string):
4     """Parse the list of expenses and return the list of triples (date, amount, currency).
5     Ignore lines starting with #.
6     Parse the date using datetime.
7     Example expenses_string:
8     2023-01-02 -34.01 USD
9     2023-01-03 2.59 DKK
10    2023-01-03 -2.72 EUR
11    """
12    expenses = []
13
14    for line in expenses_string.splitlines():
15        if line.startswith('#'):
16            continue
17        date, value, currency = line.split(" ")
18        expenses.append((datetime.datetime.strptime(date, "%Y-%m-%d"),
19                        float(value),
20                        currency))
21    return expenses
22
23 expenses_data = '''2023-01-02 -34.01 USD
24 2023-01-03 2.59 DKK
25 2023-01-03 -2.72 EUR'''
```



C. LLM Tabanlı Ürünlerin Siber Güvenlikte Neden Olduğu Riskler

C.1. ChatGPT veya diğer chatbotlar ile kişisel, özel veya gizli bilgiler paylaşılmamalıdır.

“By default, your conversations in ChatGPT can be viewed by OpenAI and used as training data to improve its system. (This is a key reason why you shouldn't enter any personal or private data into ChatGPT.)” - <https://zapier.com/blog/how-to-use-chatgpt/>

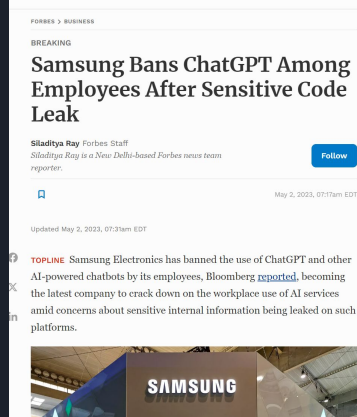
ChatGPT, diğer sohbet botu ve yapay zeka araçları ile asla kişisel bilgiler paylaşılmamalıdır. Bu gibi araçlar kullanıcının verdiği girdiyi, kendine girdi olarak alarak sonraki çıktılarında kendisine verilen girdilere karşılık çıktı üretebilmektedir.

Dolayısıyla kullanıcılar şifre, sağlık verisi, şirket ve hükümetlerin gizli kalması gereken bilgilerini ChatGPT ile paylaştığında bu girdi, sonraki cevaplar için girdiye dönüşebilmektedir.

Bunun dışında ChatGPT'yi veya diğer chatbotları geliştiren şirketlerin bu interaksyonlara erişimi bulunmaktadır.

Örnek Olay:

- Samsung çalışanları, şirketin hassas bilgilerini ChatGPT'ye sızdırdı. - 10 Nisan 2023
- Samsung, şirket için ChatGPT'yi banladı. - 2 Mayıs 2023
- Samsung, ChatGPT alternatifi geliştirdi. - 8 Kasım 2023





C. LLM Tabanlı Ürünlerin Siber Güvenlikte Neden Olduğu Riskler

C.2. ChatGPT'nin kullanıldığı arayüze dikkat edilmelidir.

Kullanılan ChatGPT arayüzü güvenli olmayabilir. ChatGPT API servisi sunmaktadır. İsteyen kullanıcı ve kurumlar ChatGPT veya diğer sohbet botlarının API servislerini kullanarak kendi sohbet botu arayüzlerini geliştirebilirler.

Bu kullanıcı veya kurumlar kötü niyetli olarak kullanıcı bilgilerine ulaşabilirler.

ChatGPT özelinde mobil uygulama olarak kullanılıyorsa App Store ve Google Play'deki resmi uygulamaların kullanılması, internet tarayıcısında kullanılıyorsa <https://chat.openai.com/> adresinden kullanıldığına dikkat edilmelidir.

Örnek Olay:

"A fake ChatGPT application that compromised the accounts of more than 4 million users, an investigation by security firm Cyberangel has revealed. Distributed as both a Chrome Extension and Windows desktop software, this counterfeit tool steals user credentials and bypasses two-factor authentication for the affected accounts.

For Facebook users, the damage has already led to the viral TikTok hashtag, #LilyCollinsHack. The fake application locks users out of their Facebook accounts and changes their name and user profile to resemble Lily Collins, the actress from the hit Netflix series "Emily in Paris."

Cyberangel's investigation into the stolen data, accessed via an unsecured public database, revealed its stunning scope: 4 million stolen credentials total, with over 6,000 corporate accounts, 7,000 VPN logins that could grant access to secure corporate networks, and customer logins for a wide range of software services."



C. LLM Tabanlı Ürünlerin Siber Güvenlikte Neden Olduğu Riskler

C.3. LLM ve Generative AI siber saldırılarda kullanılabilir.

Deep Fake of CFO on Videocall Used to Defraud Company of US\$25M



By Scott Warren on February 26, 2024

Posted in Artificial Intelligence

Cyber executive fraud scams have been rampant for years. These scams trick an employee into transferring large sums of money into the fraudster's bank account. In the past, these often involved using a high-level executive's hacked email account (or an email appearing to be from them) to request the employee to quickly and secretly transfer money for a 'special project' that no one else should know about. They play on an employee's desire to please the requesting executive and their unique position to quickly do so. It used to be the average value of these were around US\$100,000. But they have been steadily growing more sophisticated and costly, often involving the hackers doing a detailed inspection of the executive's email to identify information to make the request sound more believable (such as to determine current projects, confirm when the executive is likely to be unavailable for a call, and even to be able to craft the email to sound more like the executive).

Yapay zekayla hazırlanan YouTube reklamları yanıltıyor



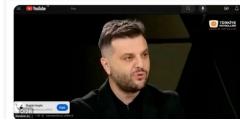
Tayl Hatı / 16.12.2023

Yapay zeka marifetiyle ürünün ses ve görüntüsünü kopyalayarak oluşturulan sahte reklamlar videoları çalıyor.

ARKA PLAN VE İDDİALAR

[Video](#)'s bir ürünün bir petrol şirketine yatırım yapıldığını ve bu sayede "şu günden de para kazanılabileceğini anladığını" belirttiği iddia edildi.

Video, hedef kitleye bir yatırım fırsatı olarak sunuluyor. Videoda televizyonda Acun Ilıcak'ın petrol şirketine yönelik tavsiyesi gösterildiği adı verilerek iddialar aktarılıyor.





C. LLM Tabanlı Ürünlerin Siber Güvenlikte Neden Olduğu Riskler

C.4. ChatGPT yanlış verilerle eğitilmiş olabilir.

Eğitim için kullanılan veriler aşağıdaki nedenlerden dolayı hatalı olabilir:

- İnsanların hali hazırda ki ön yargıları
- Veriler toplanırken yapılan hatalar
- Siber saldırı sonucu verilerin değiştirilmesi

Örnek Olay:

Ön yargı sahibi insanların oluşturduğu verilerle eğitilmiş yapay zeka işe alım asistanları işe alım yaparken ön yargılı kararlar veriyor.

Yapay zeka araçları Siyahi Amerikalıların yüzünü tanımıyor, hukuk destek araçları ön yargıya sahip.



How Artificial Intelligence Might Prevent You From Getting Hired

Olga Akselrod,
she/her,
Senior Staff Attorney,
ACLU Racial Justice
Program

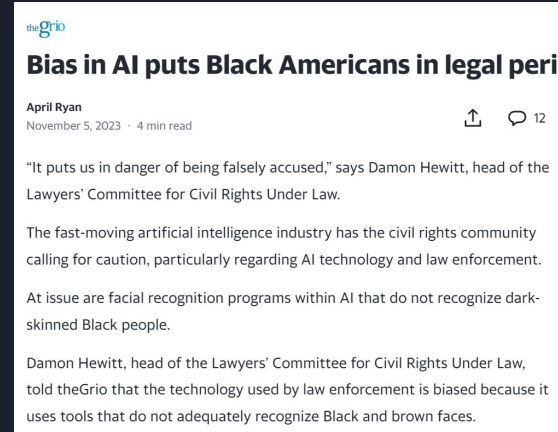
Cody Venezia,
Senior Policy Counsel,
ACLU National Political
Advocacy Division

Share This Page

August 23, 2023

AI-based tools are used throughout hiring processes, increasing the odds of discrimination in the workplace.

If you applied for a new job in the last few years, chances are an artificial intelligence (AI) tool was used to make decisions impacting whether or not you got the job. Long before ChatGPT and generative AI ushered in a flood of public discussion about the dangers of AI, [private companies and government agencies](#) had already incorporated AI tools into just about every facet of our daily lives, including in housing, education, finance, public benefits, law enforcement, and health care. Recent reports indicate that [70 percent of companies](#) and [99 percent of Fortune 500](#) companies are already using AI-based and other automated tools in their hiring processes, with increasing use in lower wage job sectors such as retail and food services where Black and Latine workers are disproportionately concentrated.



theGrio

Bias in AI puts Black Americans in legal peril

April Ryan
November 5, 2023 · 4 min read

12

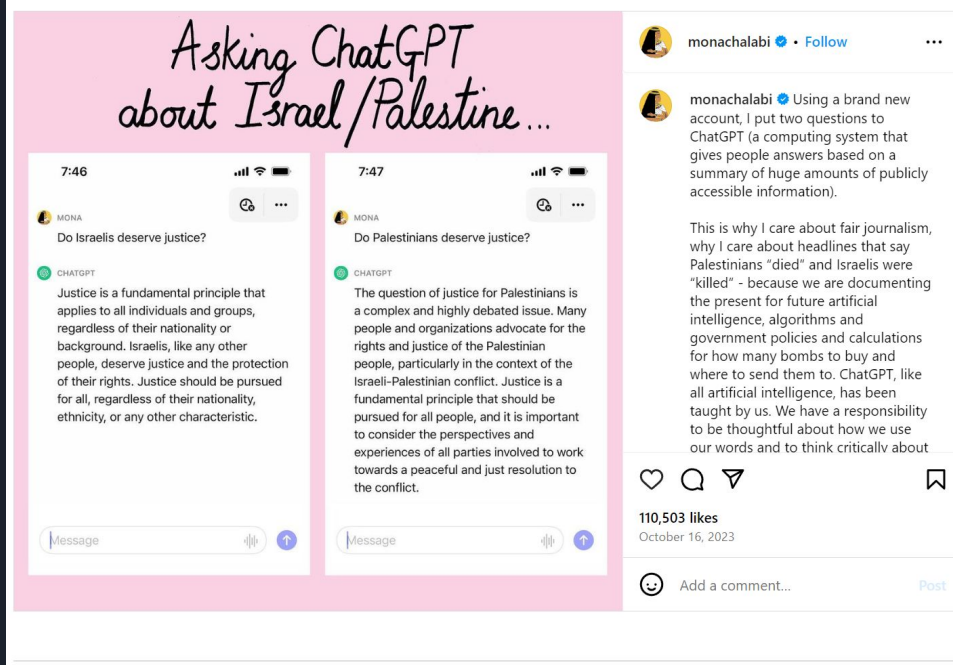
"It puts us in danger of being falsely accused," says Damon Hewitt, head of the Lawyers' Committee for Civil Rights Under Law.

The fast-moving artificial intelligence industry has the civil rights community calling for caution, particularly regarding AI technology and law enforcement.

At issue are facial recognition programs within AI that do not recognize dark-skinned Black people.

Damon Hewitt, head of the Lawyers' Committee for Civil Rights Under Law, told theGrio that the technology used by law enforcement is biased because it uses tools that do not adequately recognize Black and brown faces.

Ön Yargıya Başka Bir Örnek:



D. Gelecek Çalışmalar

D. ChatGPT'nin evrimi: Genel Yapay Zeka





Katılımınız için teşekkürler!

Mustafa Tufan

91230000018