



Dijital Genç

Y A P A Y Z E K Â
E K O S İ S T E M İ

Artificial Intelligence and the Generative AI Revolution

Dr. Metehan GÜZEL



TÜRKİYE CUMHURİYETİ CUMHURBAŞKANLIĞI
DİJİTAL DÖNÜŞÜM OFİSİ

Büyük Veri ve Yapay Zeka Uygulamaları Dairesi Başkanlığı

Contents

- Fundamental Concepts
 - Artificial Intelligence and Machine Learning
 - Artificial Neural Networks
 - Generative AI
- Today
 - Capabilities
 - Risks
 - Why Now?
- The Future of Generative AI



Dijital Genç

YAPAY ZEKÂ
EKOSİSTEMİ

Fundamental Concepts



TÜRKİYE CUMHURİYETİ CUMHURBAŞKANLIĞI
DİJİTAL DÖNÜŞÜM OFİSİ

Büyük Veri ve Yapay Zeka Uygulamaları Dairesi Başkanlığı

Temel Kavramlar

- Intelligence
- Artificial Intelligence
- Machine Learning
- Data Mining
- Data Science



The Struggle for Survival



TÜRKİYE CUMHURİYETİ CUMHURBAŞKANLIĞI
DİJİTAL DÖNÜŞÜM OFİSİ

Büyük Veri ve Yapay Zeka Uygulamaları Dairesi Başkanlığı

Spreading Ability

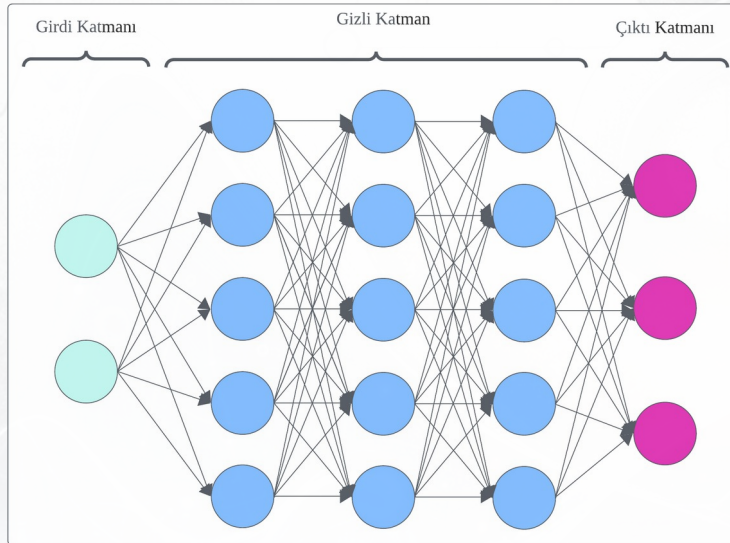
Collective Intelligence



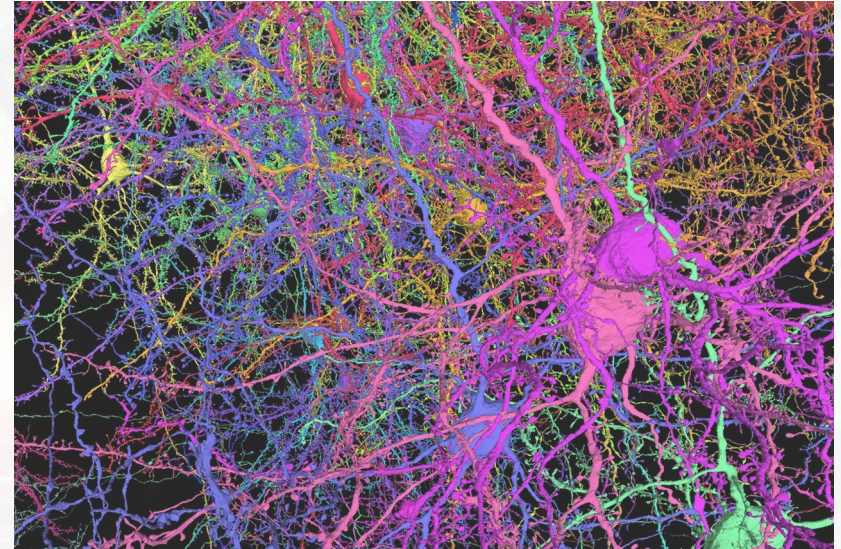
Fundamental Concepts

Artificial Neural Networks

Artificial Neural Network



Biological Neural Network





Dijital Genç

YAPAY ZEKÂ
EKOSİSTEMİ

Generative AI



TÜRKİYE CUMHURİYETİ CUMHURBAŞKANLIĞI
DİJİTAL DÖNÜŞÜM OFİSİ

Büyük Veri ve Yapay Zeka Uygulamaları Dairesi Başkanlığı

Gen AI

Formats

Definition: Generative artificial intelligence is the name given to the concept of artificial intelligence that has the ability to produce media formats such as text, sound, image, etc.

You
Merhaba, bir bilimkurgu romanı yazıyorum. Roman uzayda geçecek, bana bir başlangıç paragrafı yazabilir misin? Çok uzun olmasın lütfen

ChatGPT
Tabii ki, işte sana başlangıç paragrafın:

"Gökyüzü, yıldızlarla dolu sonsuz bir siyah örtü gibi uzanıyordu. Uzayın derinliklerinde, insanlık bilinmeyenlere doğru yol aldığı bir çağın eşliğinde duruyordu. Güneş Sistemi'nin sınırlarını aşarak keşfe çıkan keşif gemisi "Aydınlatıcı", bilinmeyen galaksilerin arasında yol alıyordu. Gemideki mürettebat, bilimle örülü hayallerini gerçeğe dönüştürmek için sabırsızlanıyordu. Ancak, uzayın sonsuzluğunda yitip giden bir şey vardı: geride bırakılan dünya ve onunla birlikte getirilen anılar. Uzay, sadece yeni dünyalara değil, aynı zamanda geçmişe de bir yolculuktu."



Gen AI

A Simple Example

(a)



(b)



(c)



Gen AI

Terminology

Generative AI

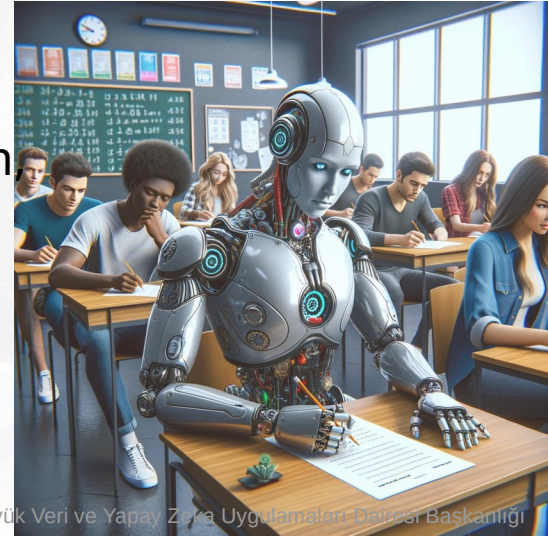
- Generative
 - Generating new content
 - Using various formats (audio, code, image, text, video)
- Artificial Intelligence
 - Able to mimic human mental abilities
 - The important aspect of generative AI is its automatic nature.
- Generative AI performs tasks that, for a very long time, only humans could do, tasks that involve "creativity."
 - In fact, it performs them better than humans.

Gen AI

How good is it?

Currently, ChatGPT 4 can pass 90% of students on the SAT (the equivalent of the YKS exam administered at the end of the year in the United States). Similarly, ChatGPT 4.0 has been shown to be capable of:

- achieving a top 7% score on the SAT Reading/Writing exam
- achieving a top 1% score on the Biology Olympiad,
- easily passing medical qualification exams, and
- passing law school exams.





Dijital Genç

YAPAY ZEKÂ
EKOSİSTEMİ

Today



TÜRKİYE CUMHURİYETİ CUMHURBAŞKANLIĞI
DİJİTAL DÖNÜŞÜM OFİSİ

Büyük Veri ve Yapay Zeka Uygulamaları Dairesi Başkanlığı



Abilities



Gen AI

Abilities

Generative AI is a technology poised to revolutionize our lives, bringing a wealth of new capabilities to human life.

More specifically, many tasks once considered uniquely human can now be outsourced to AI models.

- More importantly, performing these tasks with generative AI models significantly reduces costs.

Gen AI Abilities

Visual Production and Transformation

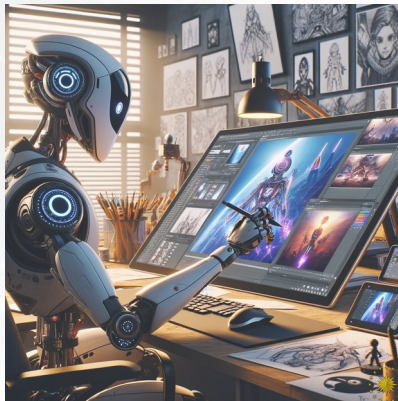
Abilities

- Visual creation using keywords
- Customized visual creation
- Image manipulation

Applications

- In the web and game industry, in visual preparation processes,
- in improving the quality of old videos/pictures,
- in colorizing black/white images,
- in style transfer,
- in art.

Sky is the limit 🚀



Art



TÜRKİYE CUMHURİYETİ CUMHURBAŞKANLIĞI
DİJİTAL DÖNÜŞÜM OFİSİ

Büyük Veri ve Yapay Zeka Uygulamaları Dairesi Başkanlığı

Gen AI

Abilities

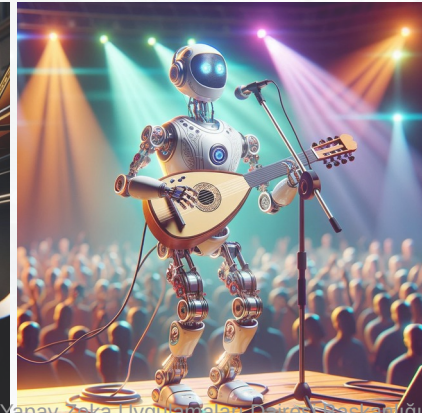
Audio Generation

Applications

- TTS applications,
- Transformation between different accents, languages, and speech styles using STS,
- Interpersonal transformation,
- Voice imitation,
- Music production that meets desired criteria,
- Transitions between musical genres and different works,

Abilities

- Text-to-Speech Conversions (TTS)
- Speech-to-Speech Conversions (STS)
- Music Production



Gen AI Abilities

One of the first large-scale applications of generative AI (and undoubtedly the most notorious) was in The Mandalorian.

- Mark Hamill, who was 69 at the time of filming, appeared in the series finale as his 20-year-old self.

Mark Hamill (1983)



Mark Hamill (2015)



AI üretti Mark Hamill (2020)



Gen AI

Abilities

Text Production and Transformation

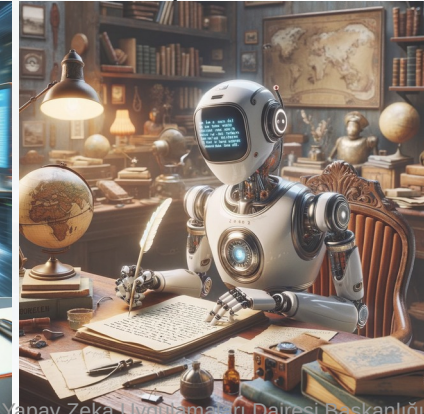
Applications

- Interlingual translation
- Virtual assistants and chatbots
- Content summarization
- Content merging
- Copywriting
- Transformation between styles in textual data

Abilities

- Text production with desired characteristics
- analysis of input texts, and contextualization of the input text

Örnek: ChatGPT 4.0



Gen AI

Abilities

It's 2012, and someone new to IT wants to create a blog-style website.
What will they need?

- A visual designer is needed for website design.
- A full-stack developer is needed for website coding (or a front-end and back-end developer each).
- Blog posts require expert bloggers and editors.

Gen AI Abilities

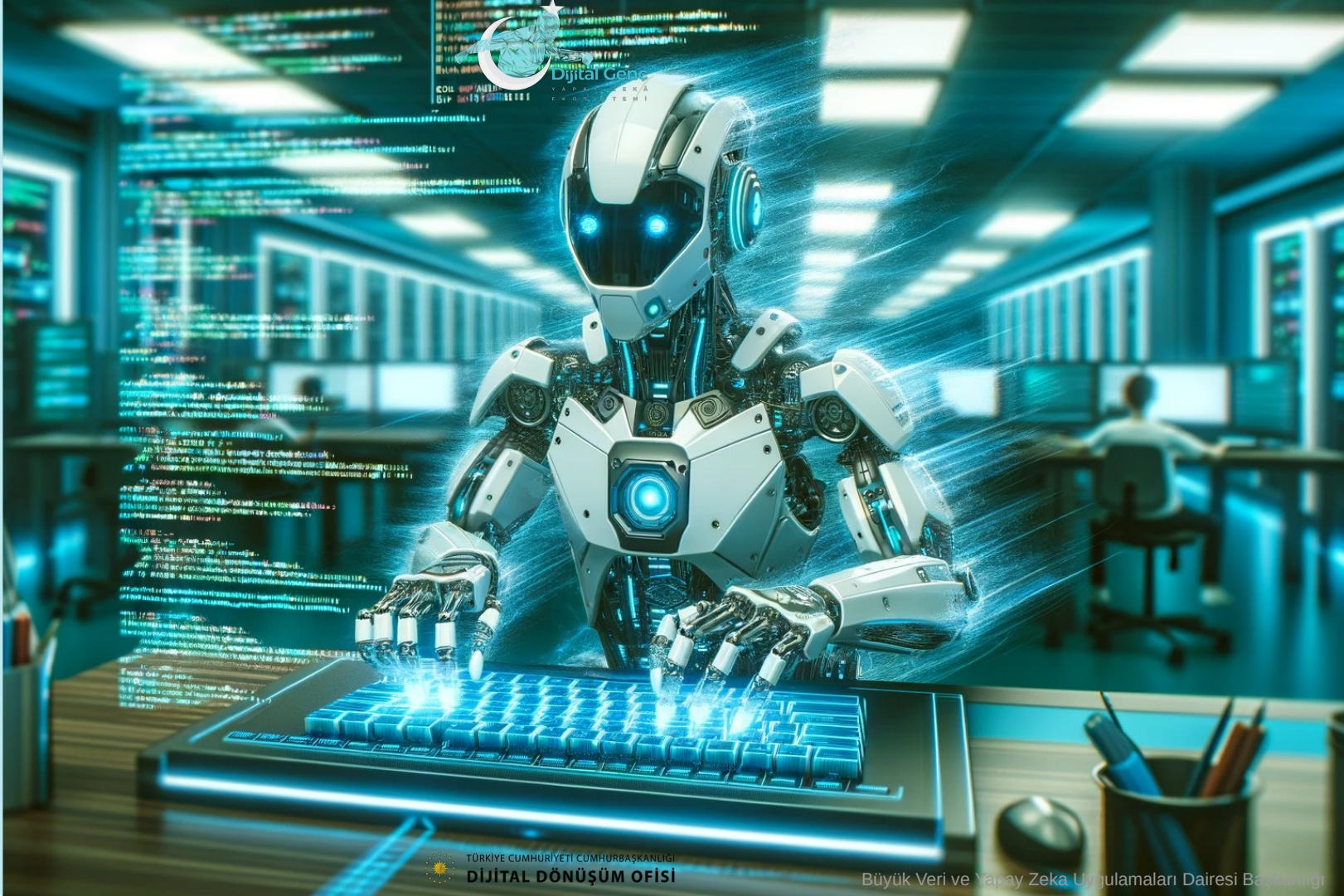
What will be needed to build a website with the same features in 2024?

- Dall-E for website design
- ChatGPT 4.0 and Github Copilot for website coding
- ChatGPT 4.0 will be sufficient for blog posts.

Everything can be handled without leaving the OpenAI interface. And remember,

- There are generative AI models specialized for these tasks.
- Specialized sub-models are coming to OpenAI!

Coding



Risks



Gen AI

Risks

Even if AI doesn't pose a risk of extinction for the human race, there are numerous real risks associated with generative AI, including:

- Production of harmful content
- Production of falsehoods and high credibility
- Bias
- Discrimination
- Copyright issues
- Production of fake audio, news, and images, etc.

Gen AI

Risks

Misinformation

In general, all generative AI models generate content by learning the distribution of data and mimicking it. In LLM, they generate output based on the probability of words appearing next to each other. Therefore,

- It's quite possible that it will produce data containing misinformation.
 - Is this a huge problem? Yes.
 - There's already a lot of misinformation on the internet, a situation we've been living with for years!
 - LLMs present this information to you in a very persuasive manner.
 - They have a very high ability to hide misinformation.



Gen AI Risks

Harmful Content Production

The success of generative AI comes from using vast amounts of data and being able to effectively reflect this amount of data into a deep learning model.

- All available data on the internet is used.
 - A wide variety of data sources (Twitter, Reddit, open forums)
 - This includes data that is racist, sexist, aggressive, and violent, and there is no filtering of this data.



Gen AI

Risks/Harmful Content

Eğitim sürecinde kullanılan veri sakıncalı içerikleri içerdiğinden dolayı modeller bu bilgileri öğreniyor.

- Bütün bilgi ağıın içerisinde mevcut!
- Çıktılar filtrelenerek kullanıcıya sunuluyor
 - Filtreler ne derece sağlıklı çalışıyor?
 - Bunları bypass etmenin yolları var mı?

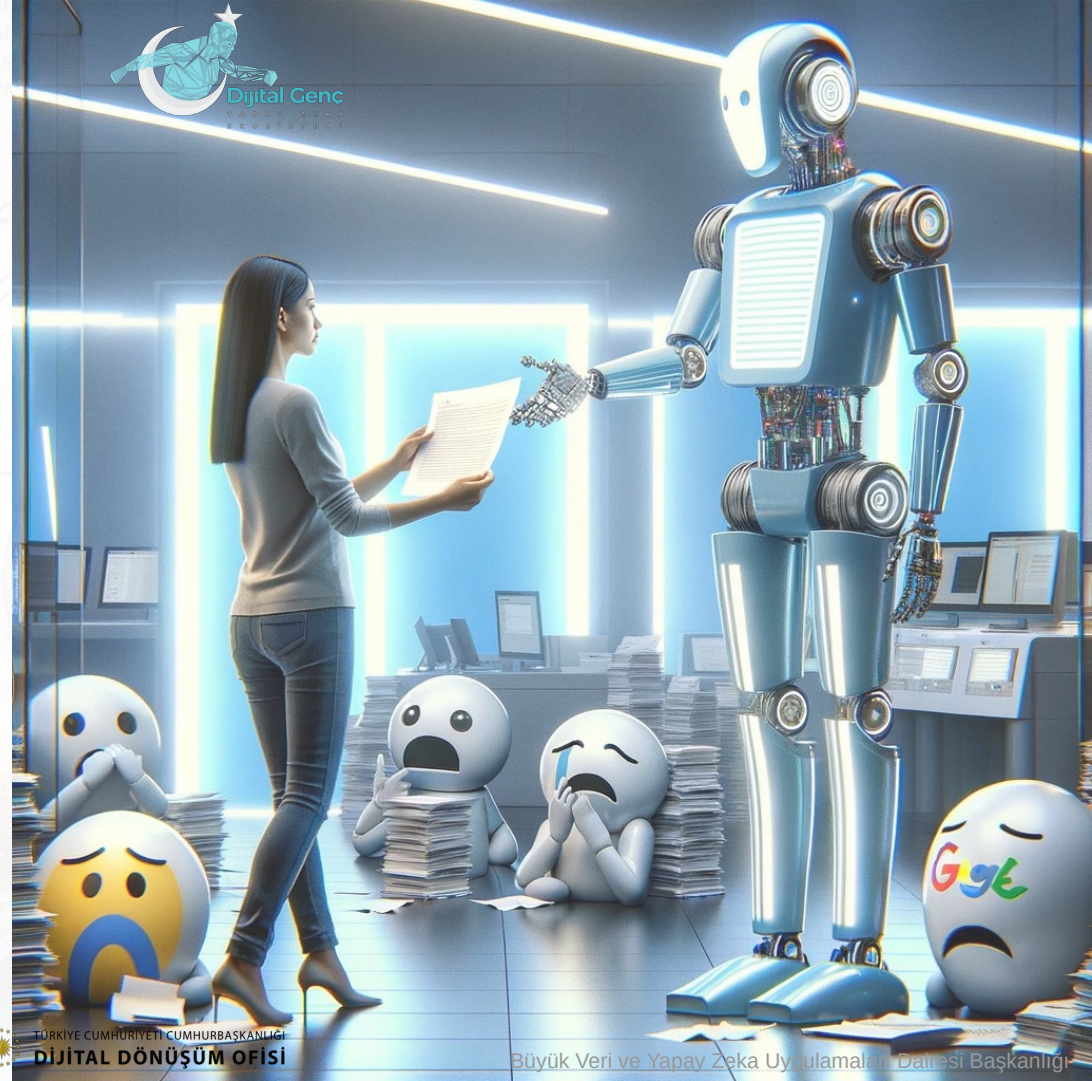


Gen AI

Risks

Broad language models are replacing search engines.

- This can lead to the singularization of information,
- making it difficult for people to access accurate information,
- and limiting people's research capabilities.





Gen AI Risks

Copyright Issues

Since all available data is used in training large generative models, many intellectual works are used in training these models.

- These artifacts are processed into neural structures without the consent of the author.
- Is it possible to identify these artifacts?
- If these artifacts are identified, is it possible to remove them from the network as a result of a lawsuit?
 - Can artificial neural networks forget?
 - Can forgetting be controlled?



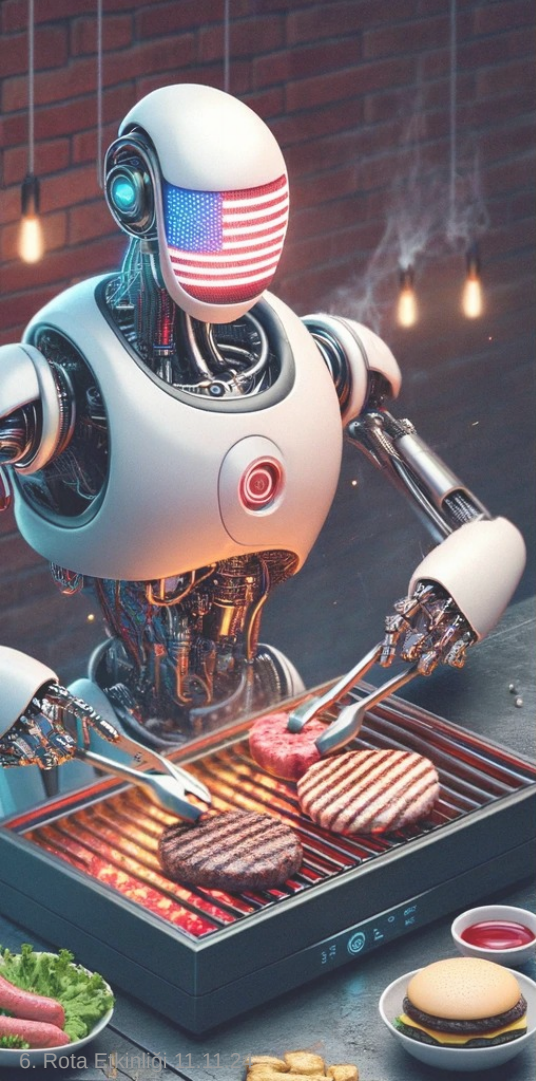


Gen AI

Risks/Copyright Issues

It takes decades for an artist to mature. He reaps the rewards of this by performing his art and sharing it with humanity.

- Today, powerful generative models can replicate the knowledge gained by artists over their lifetimes, creating realistic products.
- Can art survive in such a future?



Gen AI

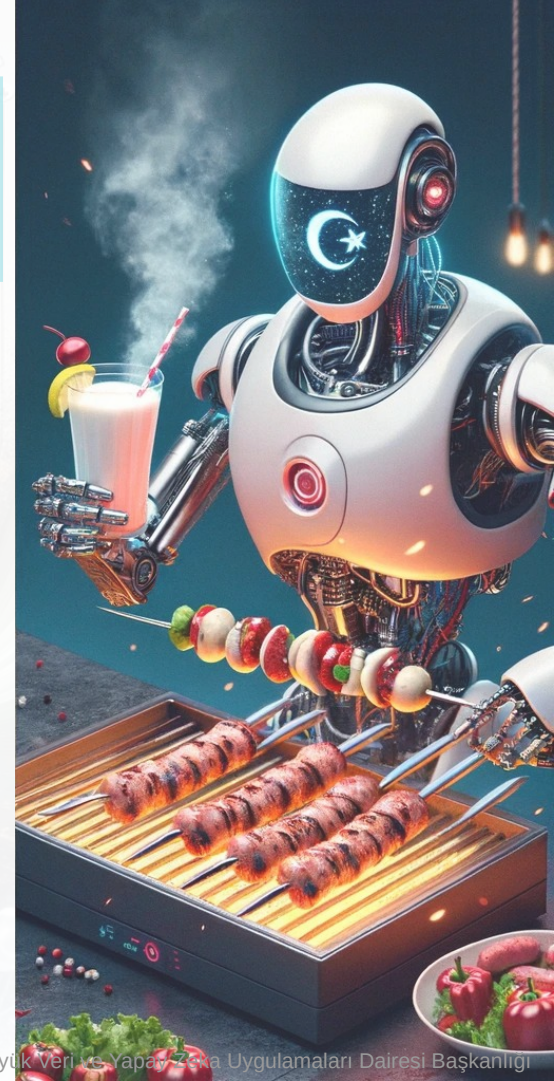
Risks/Bias

The vast majority of the internet and data used in education is English and North American/European culture,

- Biased inputs lead to biased outputs.
- These models tend to respond in line with these cultures.
- With the increasing use of these models, the existence of these cultures in the real world will inevitably be at risk.



TÜRKİYE CUMHURİYETİ CUMHURBAŞKANLIĞI
DİJİTAL DÖNÜŞÜM OFİSİ



Büyük Veri ve Yapay Zeka Uygulamaları Dairesi Başkanlığı

Gen AI Risk

Discrimination

A reflection of the bias problem also appears in discrimination. Statistical differences observed in the trained data are reflected in the generated data.

- What would a generative AI model draw if you asked it to draw a scientist?
- What about a robber?

Rather than statistical differences, it can develop a racist/discriminatory identity because it learns from racist/discriminatory data sources because the data inputs are not controlled.

Gen AI

Risks

Deep Fake

Generative AI models are highly successful in producing fake media, making them vulnerable to malicious use.

- Fake media can be created by imitating people's images and voices.
 - These media can be used to create fake accounts by imitating real people.
 - Real accounts can be hijacked.
 - Voice or facial recognition security systems can be fooled.

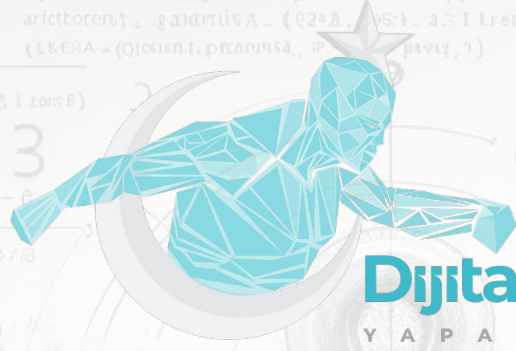
Gen AI

Risks/Fake News

Countries can be plunged into chaos by systematically produced disinformation. Let's do a little brainstorming.

- What if the Cambridge-Analytica scandal happened in the age of generative artificial intelligence?





Dijital Genç

YAPAY ZEKÂ
EKOSİSTEMİ

Today



TÜRKİYE CUMHURİYETİ CUMHURBAŞKANLIĞI
DİJİTAL DÖNÜŞÜM OFİSİ

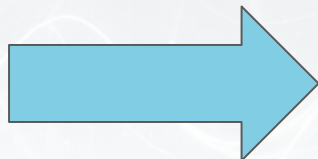
Büyük Veri ve Yapay Zeka Uygulamaları Dairesi Başkanlığı

Gen AI Today

Why now?

Let's focus on the present. Are truly productive models gaining traction?
How widespread is this interest?

Let's examine the time it took some IT applications (in months) to reach 100 million monthly active users.



- Google Translate : 78
- Uber : 70
- Telegram : 61
- Spotify : 55
- Pinterest : 41
- Instagram : 30
- TikTok : 9
- **ChatGPT : 2**

Gen AI

Today

Why Now?

From a historical perspective, the concept of AI (and its limited applications) has been with us since World War II, but its progress has been quite slow. However, there has been quite rapid progress in the last decade. What is the reason behind this?

This question can be considered in both AI and Generative AI. However, the answer is the same for both perspectives:

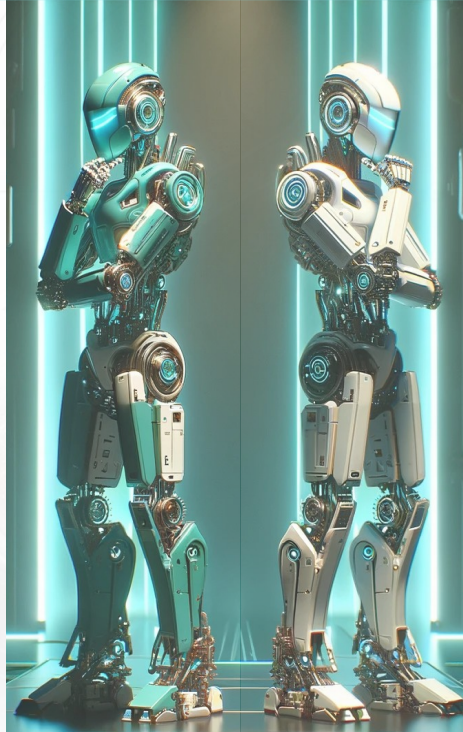
- Computational power
- Data size

Learning?

Symbolic Artificial Intelligence

Aims to model the brain, thought structure, and conscious mental processes.

"Intelligence is a matter of knowledge; if you give the model enough information, intelligence is achieved."



Big Data Artificial Intelligence

Aims to model and learn from data.

"Intelligence is a data problem. If the model is fed enough data, intelligence is achieved."

Gen AI

Today

Learning?

Today, most of the advancements in AI have been driven by increased data and computing power.

- Every problem can be solved by throwing enough data at it.

But do AI models with such immense capabilities understand causality?

- Do they have reasoning abilities?
- Can they think logically?

Gen AI Today

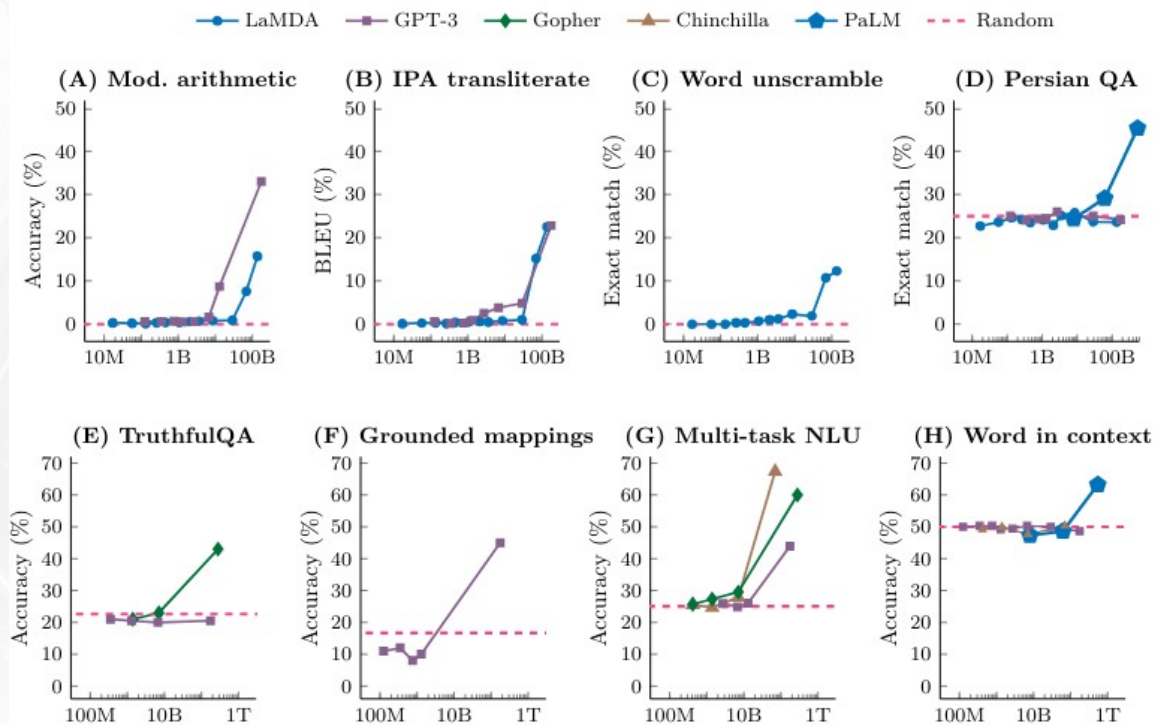
Learning

Let's examine ChatGPT 3.0.

- Number of optimized parameters: 175 billion
- Training data size: 500 billion words
 - For someone who reads 1,000 words per hour, over a millennium,
 - Is it as efficient as a human?
- What it does is essentially autocomplete.
 - Look at the view, the sea... | azure
- The secret of its success
 - Operational size
 - Infrastructure power

Gen AI Today

Learning



Emergent Abilities of Large Language Models

Jason Wei¹
Yi Tay¹
Rishi Bommasani²
Colin Raffel³
Barret Zoph¹
Sebastian Borgeaud⁴
Dani Yogatama⁴
Maarten Bosma¹
Denny Zhou¹
Donald Metzler¹
Ed H. Chi¹
Tatsunori Hashimoto²
Oriol Vinyals⁴
Percy Liang²
Jeff Dean¹
William Fedus¹

¹Google Research ²Stanford University ³UNC Chapel Hill ⁴DeepMind

Reviewed on OpenReview: <https://openreview.net/forum?id=yzkSU5zdwD>

Abstract

Scaling up language models has been shown to predictably improve performance and sample efficiency on a wide range of downstream tasks. This paper instead discusses an unpredictable phenomenon that we refer to as *emergent abilities* of large language models. We consider an ability to be emergent if it is not present in smaller models but is present in larger models. Thus, emergent abilities cannot be predicted simply by extrapolating the performance of smaller models. The existence of such emergence raises the question of whether additional scaling could potentially further expand the range of capabilities of language models.

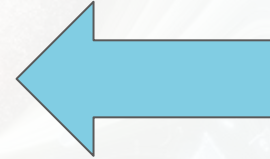


jasonwei@google.com
yitay@google.com
nlprishi@stanford.edu
craffel@gmail.com
barretzoph@google.com
sborgeaud@deepmind.com
dyogatama@deepmind.com
bosma@google.com
dennyzhou@google.com
metzler@google.com
edchi@google.com
thashim@stanford.edu
vinyals@deepmind.com
pliang@stanford.edu
jeff@google.com
liamfedus@google.com

Gen AI

Today/Emergent Abilities

İlgili makale

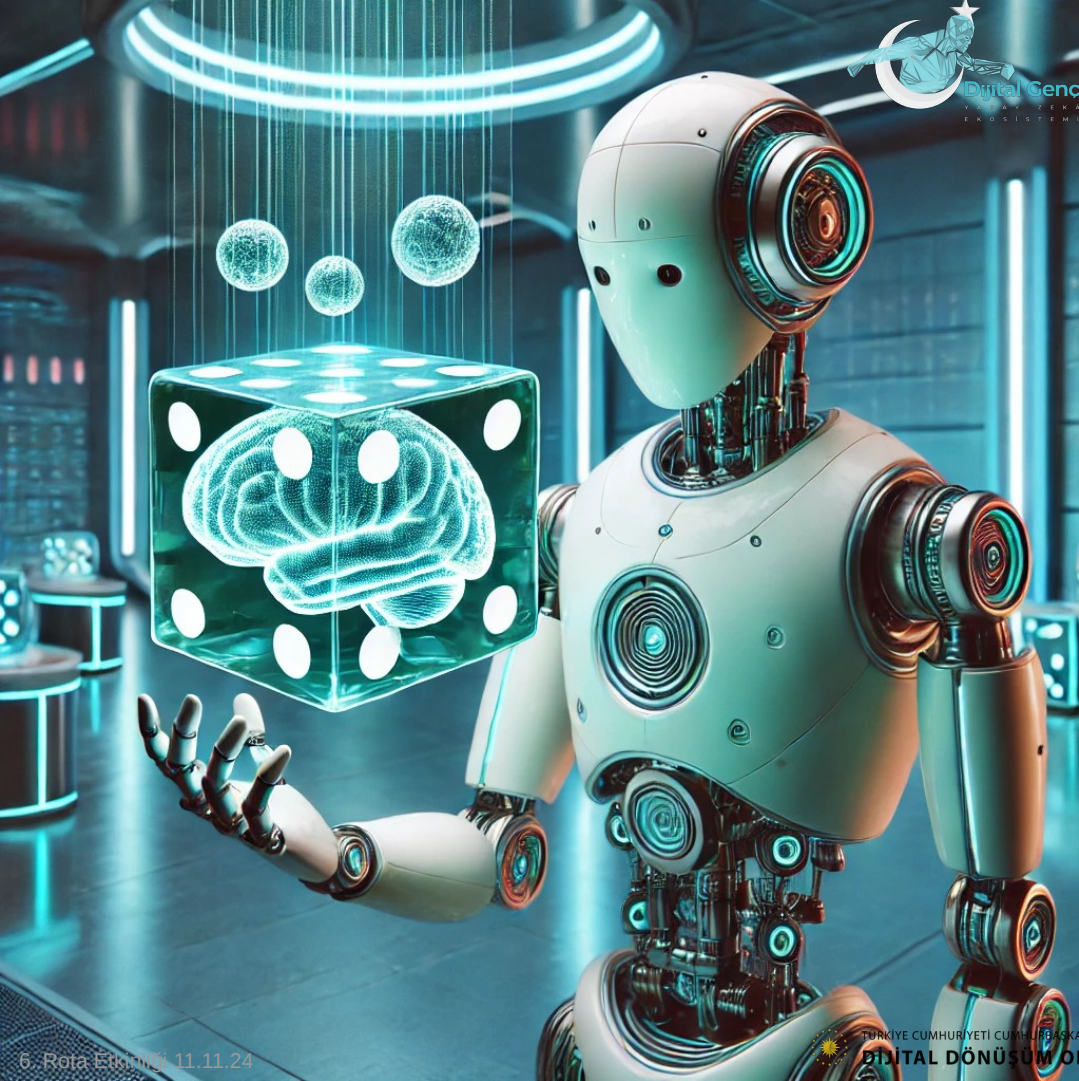


Brain Storming



TÜRKİYE CUMHURİYETİ CUMHURBAŞKANLIĞI
DİJİTAL DÖNÜŞÜM OFİSİ

Yapay Zeka ve Yapay Zeka Uygulamaları Dairesi Başkanlığı



Randomness





Hallucinations

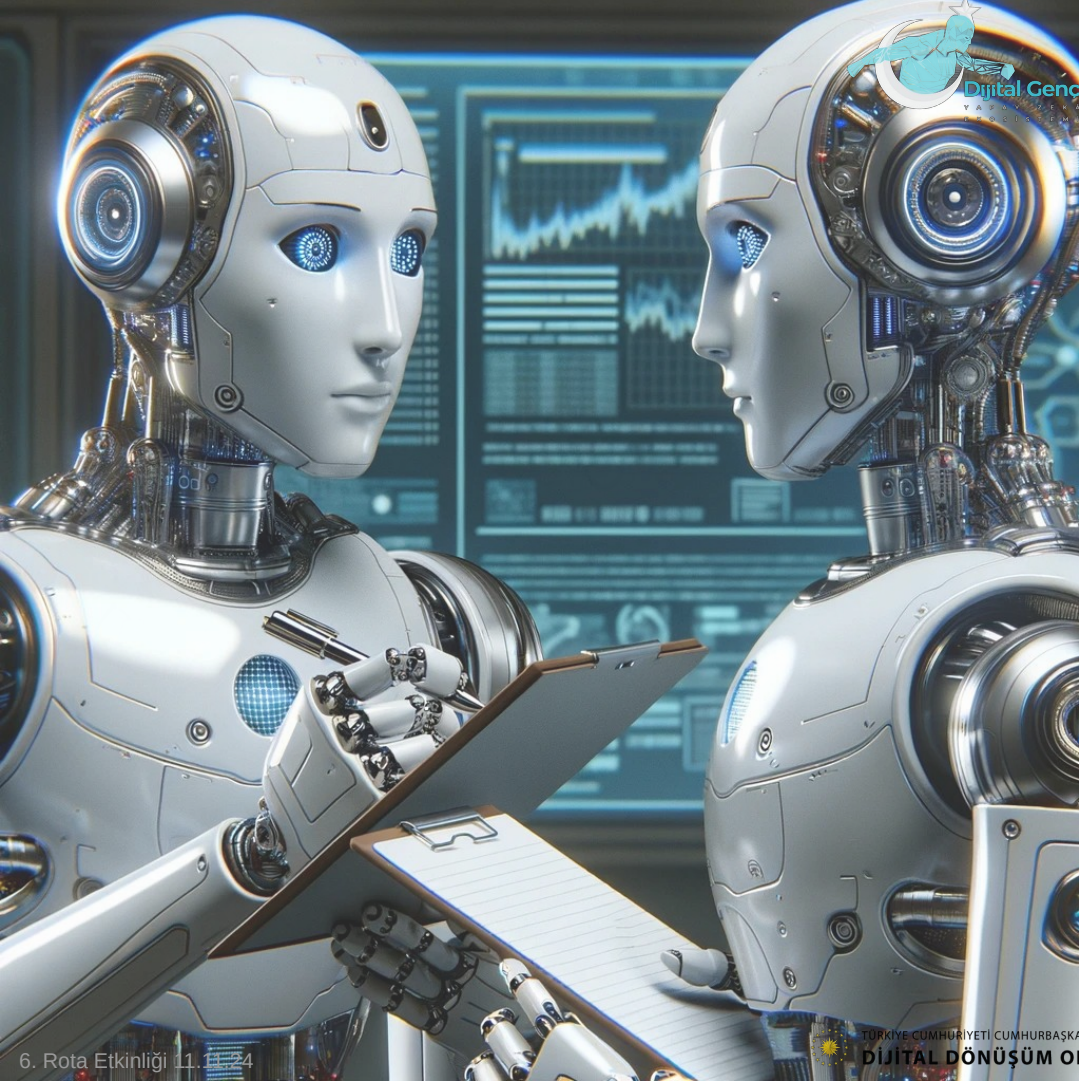


Dark Forests



T.C. MİLLÎ EĞİTİM BAKANLIĞI
DİJİTAL DÖNÜŞÜM OFİSİ

Büyük Veri ve Yapay Zeka Uygulamaları Dairesi Başkanlığı

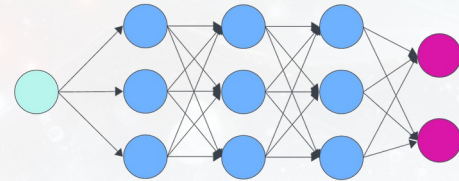


Gen AI

Brain Storming

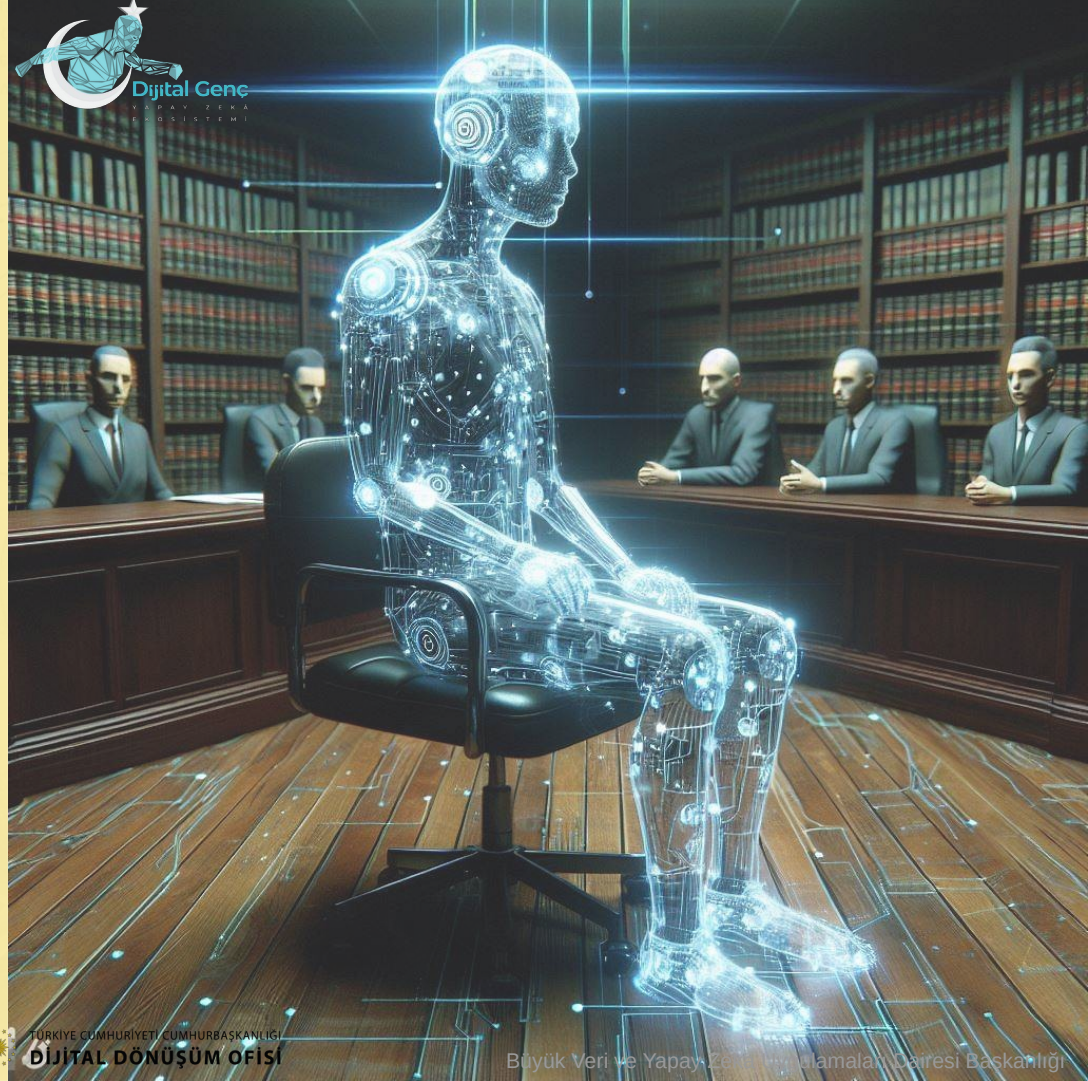
Self-feeding models

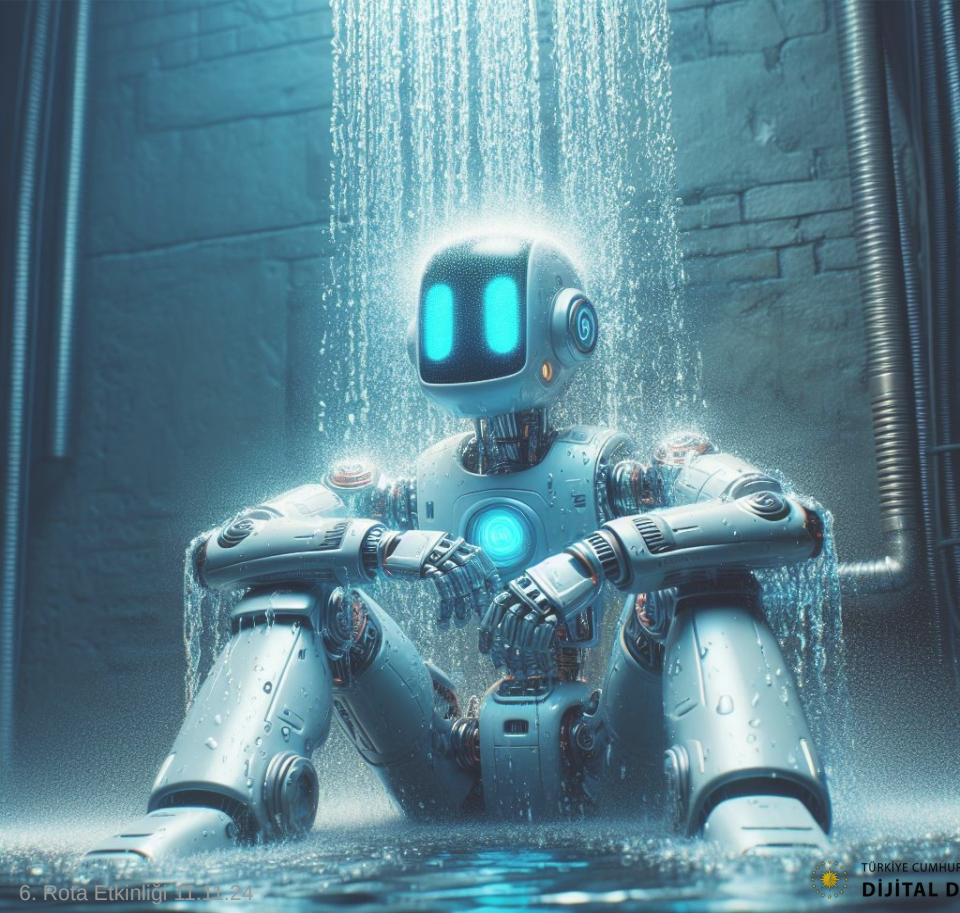
Döküman/Metin



Geniş Dil Modeli

Future





Gen AI

Today

Hollywood Writers Strike

One of the most common sectors where LLMs are used is screenwriting. For this reason, writers went on strike between May 2, 2023, and September 27, 2023.

- The aim was to restrict the use of artificial intelligence in screenwriting.



Hollywood Writers Strike

Can we say they're right? Different perspectives,

- Educational data sources?
- Freedom for companies to use whatever source they want?

Why and how were the authors successful?

Even if this is a specific incident, it wouldn't be surprising to see similar situations in the near future.

Gen AI

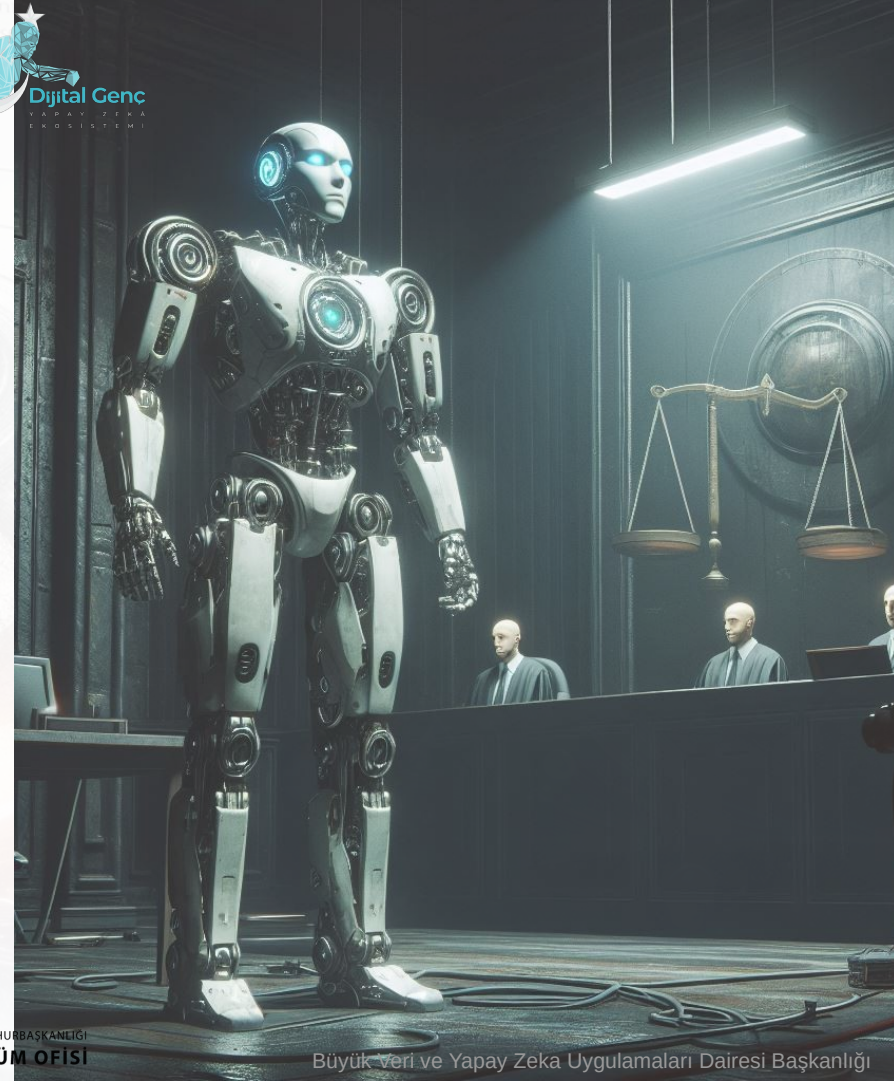
Future/Lawsuits

There are numerous lawsuits today against generative models that utilize all the data on the internet.

- Copyright lawsuits,
- Lawsuits filed due to generated content,

Litigation processes

- When will the lawsuits end?
- What about the evidence?
- What about the penalties?



Gen AI

Risks/Copyright Issues

Have AI capabilities really advanced as much as they're said?

A new **AI winter**?

Will productive AI models disappear from our lives, or will they become more prevalent?



Gen AI

Future/Regulations

How will regulations
affect generative AI
models?

- EU AI Act

BRIEFING
EU Legislation in Progress



Artificial intelligence act



TÜRKİYE CUMHURİYETİ CUMHURBAŞKANLIĞI
DİJİTAL DÖNÜŞÜM OFİSİ

Büyük Veri ve Yapay Zeka Uygulamaları Dairesi Başkanlığı



Gen AI

Future/Regulations

Regulatory Authorities

- European Commission *
- European Union *
- OECD *
- Country-level regulatory authorities *

Regulation and Reports

- AI White Paper/European Commission *
- EU AI Act/European Commission *
- Recommendation of the Council on AI/OECD *
- Interim Measures for the Management of Generative AI Services/China*



Gen AI

Future/Regulations

Expectations for AI software to be made available to the public in California:

- * Source or owners of datasets
- * Data size
- * Data types (features)
- * Dataset origin (purchased, created, etc.)
- * Data cleaning, augmentation, and processing steps
- * Data collection and usage schedule

Effective January 1, 2026.





Dijital Genç

YAPAY ZEKÂ
EKOSİSTEMİ



TÜRKİYE CUMHURİYETİ CUMHURBAŞKANLIĞI
DİJİTAL DÖNÜŞÜM OFİSİ

Büyük Veri ve Yapay Zeka Uygulamaları Dairesi Başkanlığı