

## Generative AI presentation summary

/ Andres Karjus | Tallinna Ülikool | Datafigure OÜ  
[andreskarjus.github.io](https://andreskarjus.github.io) | [datafigure.eu](https://datafigure.eu)

# Terms

## Large language models (LLM)

Machine learning models trained on vast amounts of textual data predict the next word (sentence, text) to an input.

GPT (Generative Pre-trained Transformer) is a type of LLM

OpenAI is the company behind GPT, ChatGPT, Dall-E, etc  
But all big tech companies are pivoting to AI now

Prompt: the input to a generative model, typically a sentence

Prompt engineering: the emerging field of crafting optimal prompts



# Language models

ChatGPT is a *chatbot* that runs on various GPT models (optimized GPT3.5, now GPT4)

"Answers" any question, but does not really "know" anything (although sounds like it does).

Predicts answers, which are often sensible, but can be "hallucinated".

Not perfect. But already pretty good.

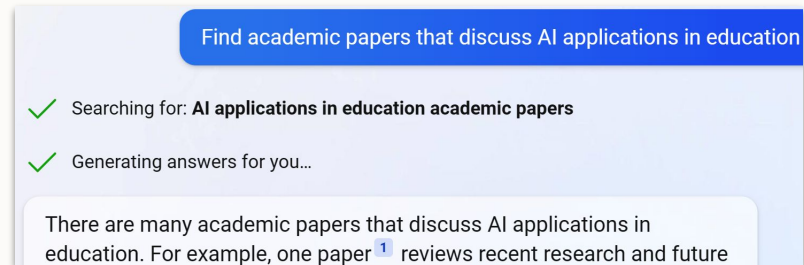
*It is not a new Wikipedia. It's not search/googling...*

# ...but it is being integrated with search

MS Bing Chat (based on GPT), Google Bard, You.com, Perplexity.ai, etc.

Rapid development towards integration of chatbots with search engines.

**TechScape: Google and Microsoft are in an AI arms race - who wins could change how we use the internet**



Find academic papers that discuss AI applications in education

- ✓ Searching for: **AI applications in education academic papers**
- ✓ Generating answers for you...

There are many academic papers that discuss AI applications in education. For example, one paper <sup>1</sup> reviews recent research and future

OpenAI  
GPT3.5/4



ChatGPT



Bing



Perplexity



YOU.com



...



Cursor



GitHub Copilot



GitHub Copilot X

Ghostwriter

Write better code, faster, with Replit's web-based coding AI.

replit



Meta

LLaMA (open source)

Alpaca  
Vicuna etc

Falcon LLM

stability.ai

Open Assistant



ANTHROPIC  
Claude



PaLM 2



Bard AI



Google Docs



# Generative models boom?

Challenges as well as opportunities

Maybe an opportunity to rethink some fossilized concepts

Changes are coming (and underway) in the edu sector and job market. No point in hiding one's head in the sand.

A chance for agile organizations to show the way and/or gain an edge (and avoid extinction)



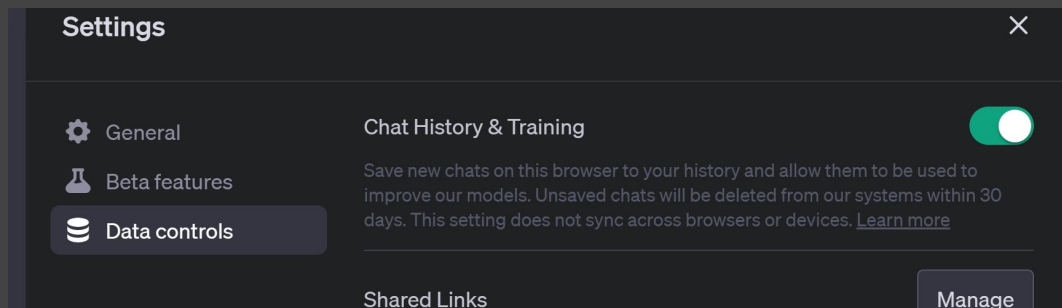
# How to use all of these things



- ChatGPT: [chat.openai.com](https://chat.openai.com), make an account or log in with Google etc. Free version is on GPT3.5; paid Plus (\$22/month) on GPT4.  
gpt-3.5-turbo API \$0.002/1000 tokens (~700 words)
- Bing: [bing.com](https://bing.com) (see also Bing Create for images)
- Bard: [bard.google.com](https://bard.google.com)
- See also: [perplexity.ai](https://perplexity.ai), [you.com](https://you.com), [open-assistant.io](https://open-assistant.io), [anthropic.com](https://anthropic.com)
- Try more image models: [midjourney.com](https://midjourney.com), [dreamstudio.ai](https://dreamstudio.ai), [scribblediffusion.com](https://scribblediffusion.com), [mage.space](https://mage.space), [clipdrop.co](https://clipdrop.co)

# Data privacy issues

- All your input data is very likely used for training their next products
- Solutions: run your own LLM (requires beefy hardware though) or the ChatGPT Enterprise (aka "pinky swear we won't train on your data")





# The need for updated digital literacy

- Spreading fake news has never been easier (text, photos, etc)
- Photo & video no longer credible evidence
- People may misuse chatbots (e.g. to seek medical advice)
- Phishing and scams are about to get a whole lot more elaborate ->

**Rapid growth of 'news' sites using AI tools like ChatGPT is driving the spread of misinformation**



Subject: Checking In and A Small Request

Dear Andres,

I hope you're well. Your work at the CUDAN C continues to impress. I'm curious to know how the dynamics is progressing. It's certainly a fascinating learning and large scale corpora analysis project.

Our collaboration on quantifying the evolution of great experience, and your input was invaluable between academia and the media industry v

On a different note, I'm currently facing a bit of publication, and I was wondering if you could understand this is out of the blue and I'd be more details.

Looking forward to hearing from you and ca

Best wishes,

Maximilian Schich

# The (very) near future

- AI will only get better, faster, more accessible.
- Integration into Office suites, Google Docs, browsers (already in Edge), search engines (already in Bing, Perplexity), and future operating systems.
- Practically impossible to ban or prohibit (but there are/will be those offering "solutions")



# Is AI usage plagiarism?

Nobody knows. Not really.

Impossible to reliably prove (mis)use:

- Any classifier will have a high false positive rate (false accusations)
- Easy to mislead (and who would directly submit AI text anyway?)
- AI usage is not an on-off switch, a continuum of usages

(as thesaurus, spell checker, idea generator, editor, text generator, text summarizer etc)

(still.... how to grade homework, where one may have been generated+edited in minutes, while the other has taken a students hours or days of careful manual labor?)

Turnitin announces AI writing detector and AI writing resource center for educators

Monday 13 February 2023



# Conclusions and ways forward?

- Learn to talk to machines (if you haven't yet)
- But do talk to humans (friends, other students, colleagues, professors)
- If you save some time, use it well
- No time like now to learn new skills (using your free personal AI tutor, and by learning from early adopters)
- No time like now to develop new products, startups, ideas (still low barrier of entry, but the window will be closing)



**Thanks! Aitäh!**

**/ Andres Karjus | Tallinna Ülikool | Datafigure OÜ**

[andreskarjus.github.io](https://andreskarjus.github.io) | [datafigure.eu](https://datafigure.eu)