

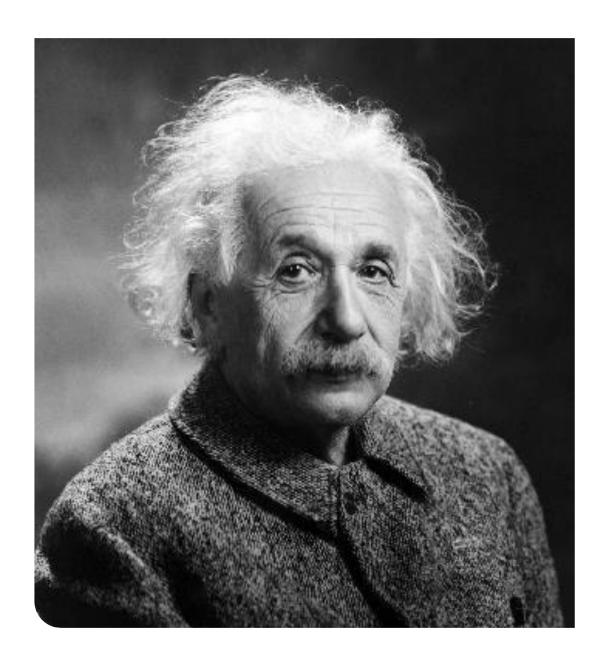
How (not) to use chatGPT as a researcher

Prof. David Martens

Faculty of Business & Economics Co-director Antwerp Center for Responsible Al Doctoral Day - October 2023

<u>Warning</u>: Presentation of October 2023 - Likely outdated, given the advances in LLM.







Einstein's Warning

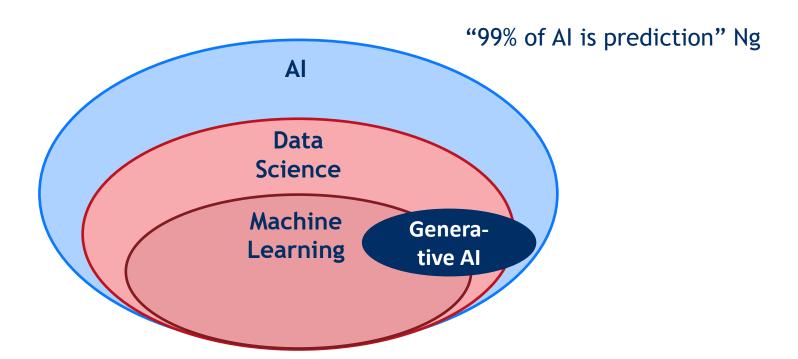
"Technological progress is like an axe in the hand of a pathological criminal."

Outline

- What is AI, generative AI and chatGPT
- Uses of chatGPT as a researcher
- Dangers of using chatGPT

Machine Learning, Data Science and Al

- Machine Learning: automatic extraction of patterns from data
- Data science: a set of fundamental principles that guide the extraction of knowledge from data
- Al: methods for improving the knowledge or performance of an intelligent agent over time, in response to the agent's experience in the world



Machine Learning, Data Science and Al



Mat Velloso @matvelloso

Gillatvettooo

Difference between machine learning and AI:

If it is written in Python, it's probably machine learning

If it is written in PowerPoint, it's probably AI

Post vertalen

2:25 a.m. · 23 nov. 2018

How does chatGPT work?

- Deep neural network: Generative Pretrained Transformer
 - Generative: Next word prediction
 - Pretrained: using 500 billion words to estimate 175 billion parameters

Input	Predicted next word
What are 10 applications of AI?	1.
What are 10 applications of AI? 1.	Credit
What are 10 applications of AI? 1. Credit	scoring
What are 10 applications of AI? 1. Credit scoring	has
What are 10 applications of AI? 1. Credit scoring has	been

Outline

- What is AI, generative AI and chatGPT
- Uses of chatGPT as a researcher
- Dangers of using chatGPT

- Content creation
- 2. Language translation
- 3. Email and communication
- 4. Idea generation
- 5. Coding
- 6. Data analysis and visualisation

- Content creation
- 2. Language translation
- 3. Email and communication
- 4. Idea generation
- 5. Coding
- 6. Data analysis and visualisation



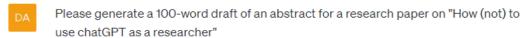
I'm giving a presentation to PhD researchers on how t use chaGPT. I have 6 general uses:

- 1. Content creation
- 2. Language translation
- 3. Email and communication
- 4. Idea generation
- 5. Coding
- 6. Data analysis and visualisation

Make a chatPGT prompt for each of the 6 uses that I can then execute, as to show interesting examples on how to use chatGPT. Try to make it already relevant for a phd researcher.

1. Content Creation:

- · Prompt: "Please generate a 500-word draft of an abstract for a research paper on [specific research topic in the researcher's field]."
- Example Output: [Generated abstract]







Title: How (Not) to Use ChatGPT as a Researcher

Abstract:

This research paper critically examines the utilization of ChatGPT, an Al-based language model, within the context of academic research. It explores the opportunities and challenges associated with employing ChatGPT for research purposes and offers guidance on responsible and effective utilization. The paper emphasizes the need for a cautious approach, discussing potential pitfalls and ethical considerations when harnessing Al-generated content in scholarly work. By providing insights into the limitations and strengths of ChatGPT as a research tool, this study aims to empower researchers to make informed decisions about its integration into their methodologies while promoting ethical and rigorous academic standards in the age of Al.







2. Language Translation:

- Prompt: "Translate the following scientific abstract from English to Spanish:
 [Provide a scientific abstract in English]."
- Example Output: [Translated abstract in Spanish]

3. Email and Communication:

- Prompt: "Compose an email to a fellow researcher regarding collaboration on a [specific research project] and suggest a suitable meeting time."
- Example Output: [Generated email content]



Compose an email that I want to send to my supervisor, with as message:

- I have finished the empirical part
- I like the results a lot, but nt sure abot the statistical testing I did
- I'd like to discuss in person
- He keeps calling me after working hours inquiring about he paper. I want him to stop doing so.

Make it polite, but to-the-point. I don't want to offend him of course.



Subject: Discussion Request and Update on Research Progress

Dear [Supervisor's Name],

I hope this message finds you well. I wanted to provide you with an update on the progress of my research project.

I'm pleased to inform you that I have completed the empirical part of the project. I've been working diligently, and I'm genuinely excited about the results I've obtained. However, there's a specific aspect that I would appreciate your guidance on. I have some reservations regarding the statistical testing I conducted, and I believe a discussion in person would be the most productive way to address this concern. I value your expertise and insights, and I believe a face-to-face conversation will allow us to explore this matter more thoroughly.

Additionally, I understand the importance of our research, but I would like to kindly request that we limit communication after working hours. I want to ensure I can give your inquiries the attention they deserve, and maintaining a healthy work-life balance is important to me.

I am grateful for your guidance and support throughout this project, and I want to make sure that we can collaborate effectively and make the most of our interactions.

Please let me know a suitable time for you to meet in the coming week, and I will make the necessary arrangements. Thank you for your understanding, and I look forward to our discussion.

Best regards,



4. Idea Generation:

- Prompt: "Brainstorm three potential research hypotheses or questions related to [a specific research domain or topic]."
- Example Output:
- 4.1. "What are the implications of [specific phenomenon] on [related field]?"
- 4.2. "How can [specific method or technology] be applied to address challenges in [related field]?"
- 4.3. "What are the emerging trends in [specific subfield] that warrant further investigation?"



Brainstorm three potential research hypotheses or questions rlated to "Al for Supply Chain Management". It should be high risk-high gain. Just list the hypothesis, no need for explanations.





Certainly, here are three high-risk, high-reward research hypotheses or questions related to "Al for Supply Chain Management":

- "Can Al enable fully autonomous, self-optimizing supply chains, reducing human involvement to near-zero, while maximizing efficiency and resilience?"
- "Is it possible to develop AI systems that can predict and mitigate supply chain disruptions with such accuracy that organizations can operate with minimal buffer stock, dramatically reducing costs?"
- "Can Al-driven supply chain orchestration models be designed to enable real-time, global supply chain collaboration, breaking down traditional silos and delivering unprecedented efficiency and agility?"







D.A.

Main uses of chatGPT

5. Coding:

 Prompt: "Write a Python script to analyze data from [mention a dataset or research source] and plot a graph that visualizes the trends related to [spec variable or parameter]."

Write a Python script to analyze data from a dataset listing a wide range of mace economic variables, per quarter, in Belgium and plot a graph that visualizes the trends related to Al spending.





To analyze data from a macro-economic dataset and plot a graph showing trends related to Al spending in Belgium, you can use Python with libraries like Pandas and Matplotlib. Ensure you have the dataset (e.g., in a CSV file) and replace ''data.csv'' with the actual path to your dataset. Here's a Python script for this task:

```
Copy code
python
import pandas as pd
import matplotlib.pyplot as plt
# Load the dataset (update 'data.csv' with the actual path)
data = pd.read_csv('data.csv')
# Filter data for Belgium and AI spending
belgium_data = data[data['Country'] == 'Belgium']
ai_spending = belgium_data['AI_Spending']
# Convert the date column to a datetime format
belgium_data['Date'] = pd.to_datetime(belgium_data['Date'])
# Sort the data by date
belgium_data = belgium_data.sort_values(by='Date')
# Plot the trends in AI spending
plt.figure(figsize=(10, 6))
plt.plot(belgium_data['Date'], ai_spending, marker='o', linestyle='-', color
plt.title('Trends in AI Spending in Belgium')
plt.xlabel('Quarter')
plt.ylabel('AI Spending')
plt.grid(True)
plt.xticks(rotation=45)
plt.tight_layout()
# Show the plot
plt.show()
```

DA

Provide a Python code snippet to crawl all news articles with the words "Elon Musk" from major Flemish online newspapers.



Crawling news articles from multiple online newspapers is a complex task and can involve web scraping, which should be done responsibly and in accordance with the websites' terms of service. Before proceeding, ensure you have permission and consider the legal and ethical aspects of web scraping.

Below is a simplified Python code snippet that demonstrates how you might crawl news articles mentioning "Elon Musk" from a single website (in this case, De Standaard, one of the major Flemish newspapers). You would need to adapt this code for each specific website you want to scrape:

```
Copy code
import requests
from bs4 import BeautifulSoup
# Define the URL of the website to scrape
url = 'https://www.standaard.be/zoek?query=Elon+Musk'
# Send an HTTP request to the website
response = requests.get(url)
# Check if the request was successful
if response.status_code == 200:
   # Parse the HTML content of the page using BeautifulSoup
   soup = BeautifulSoup(response.text, 'html.parser')
   # Find and extract relevant information (e.g., article titles and links)
   articles = soup.find_all('article', class_='c-card')
   for article in articles:
       title = article.find('h2', class_='c-card_title').text.strip()
       link = article.find('a', class_='c-card__link')['href']
       print(f"Title: {title}")
       print(f"Link: {link}\n")
                                                                      G Regenerate
```

Some examples

Title

Abstract

- Ask for suggestions
- Write in the style of [your favorite academic author]

Proposals

- Rebuttal
- Tailored towarsd call, eg ERC: high risk high gain. Add the call details and evaluation criteria.

How should I be using it?

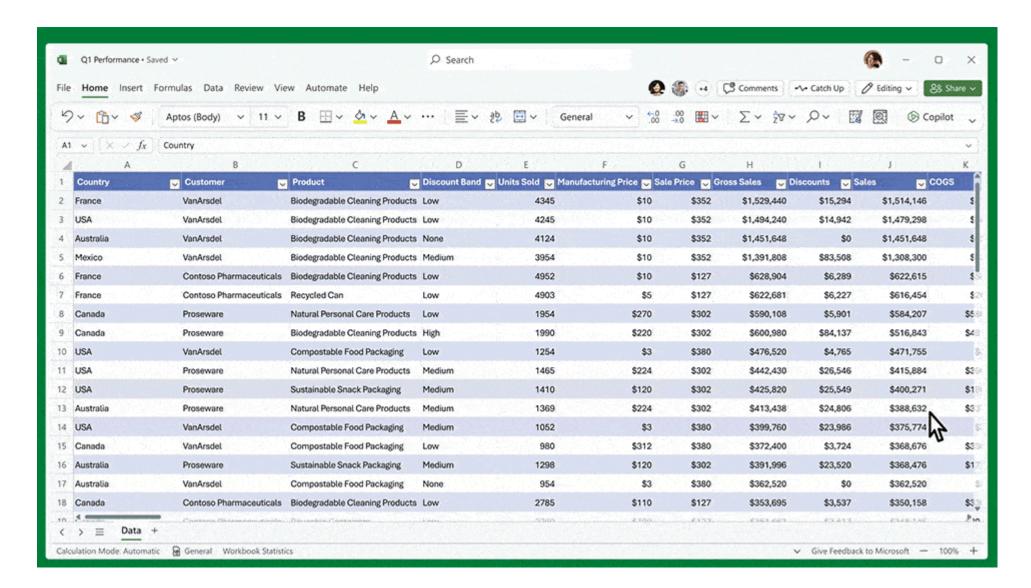
- Writing assistant: author: abstract, title, rephrasing
 - As grant applicant or reviewer: summarize, rebuttals, rephrasing, etc.
- Creative assistant: coming up with ideas for analogies, examples, structure, etc.
- Coding assistant: noone is restricted to "knowing someone who can code"
- Prereviewing assistent: ask for comments on writing, methodology, conclusions, etc.
- Should never be an author, nor need to mention the use.
- BUT be sure not to commit plagiarism, and beware for the risks...

Let's "workshop"

Here is the abstract of a paper. Any comments or suggestions for improvements? Give an example for each suggestion. Also state what a good title would be.

- Open chatGPT
- Open your latest draft paper or proposal.
- Go through the steps
 - 1. Writing assistant: author: abstract, title, methodology, conclusion, etc.
 - 2. As grant applicant or reviewer: summarize, rebuttals, rephrasing, etc.
 - 3. Creative assistant: coming up with ideas for analogies, examples, structure, etc.
 - 4. Coding assistant: noone is restricted to "knowing someone who can code"
 - 5. Prereviewing assistent: ask for comments on writing, methodology, conclusions, etc.

Microsoft Copilot is coming



I'm giving a presentation on the use of AI in research. Create a nice image for the opening slide.

I'm giving a presentation on "How to use chatGPT for research" and the audience is a group of doctoral researchers at a faculty of business and economics. Create a cool image for the opening slide.









"How to use chatGPT for research with a business ... "

Made by Bing Image Creator

Powered by DALL-E 3









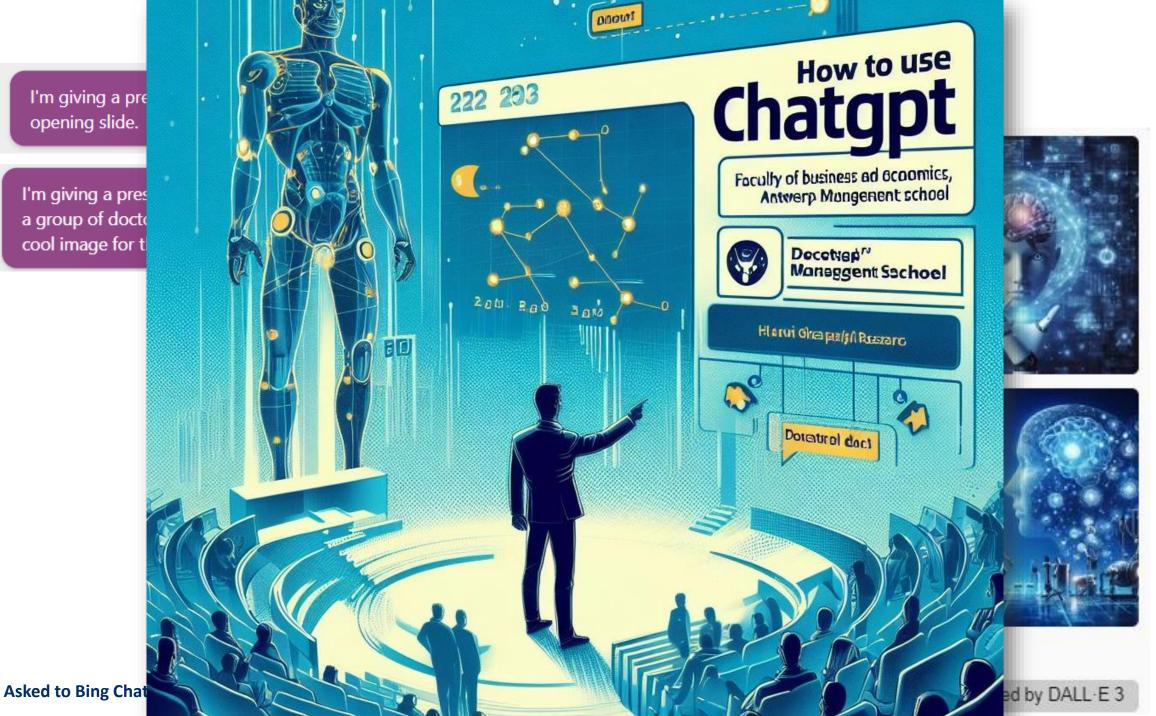
"Al in research"

Made by Bing Image Creator

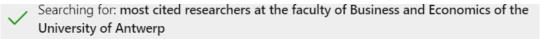
Powered by DALL-E 3

I'm giving a pre opening slide.

I'm giving a pres a group of docto cool image for t



Who are the most cited researchers at the faculty of Business and Economics of the University of Antwerp?



Generating answers for you...





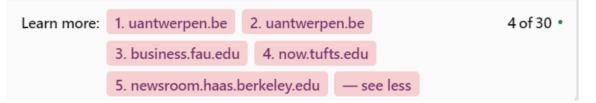
Your personal and company data are protected in this chat

According to the web search results, the most cited researchers at the faculty of Business and Economics of the University of Antwerp are not explicitly listed. However, based on the information available, some possible candidates are:

- **Ann Verhetsel**, professor of regional economics and urban planning, who has published more than 100 papers and has over 2000 citations 1.
- **Thierry Vanelslander**, professor of transport and regional economics, who has published more than 150 papers and has over 1800 citations ².
- Marc Deloof, professor of corporate finance, who has published more than 80 papers and has over 1600 citations 3.
- Christophe Croux, professor of econometrics and statistics, who has published more than 120 papers and has over 1400 citations 4.
- Nico Dewaelheyns, professor of accounting and auditing, who has published more than 50 papers and has over 1000 citations 5.

These are only some examples based on the web search results, and there may be other researchers at the faculty who have higher citation counts or more influential publications. A more accurate way to find the most cited researchers at the faculty

These are only some examples based on the web search results, and there may be other researchers at the faculty who have higher citation counts or more influential publications. A more accurate way to find the most cited researchers at the faculty would be to use a bibliometric database such as Scopus or Web of Science, which can provide more comprehensive and reliable data on academic publications and citations.





Your personal and company data are protected in this chat

According to Google Scholar, the most cited researchers at the faculty of Business and Economics, at the University of Antwerp, are:

- Ann Verhetsel, professor of economic geography and regional economics, with 3,843 Mations and an h-index of 28 1.
- Thierry Vanelslander, professor of transport and logistics, with 2,949 Mations and an h-index of 25 1.
- Christa Sys, professor of maritime economics and logistics, with 2,038 stations and an h-index of 21 1.
- Tom Brijs, professor of data mining and business intelligence, with 1,917 cations and an h-index of 19 1.
- Wouter Verbeke, professor of data analytics and customer intelligence, with 1,300 citations and an h-index of 19 1.

These are the top five researchers based on the number of citations as of October 2023. You can find more information about their research fields and publications on their Google Scholar profiles 2. I hope this helps you.



Learn more: 1. uantwerpen.be

2. scholar.google.com

3. business.fau.edu

+1 more

3 of 30 •

User profiles for Ann Verhetsel



Ann Verhetsel

Professor of Economic Geography, University of Antwerp Verified email at uantwerpen.be

Cited by 2669

Outline

- What is AI, generative AI and chatGPT
- Uses of chatGPT as a researcher
- Dangers of using chatGPT

1. Hallucinations



Is 270 90% of 300?



Generating answers for you...



Your personal and company data are protected in this chat

No, 270 is not 90% of 300. To find out what percentage of 300 is 270, we can use the following formula:

$$percentage = \frac{part}{whole} \times 100$$

In this case, the part is 270 and the whole is 300. So, we can plug in these values and simplify:

percentage =
$$\frac{270}{300} \times 100$$

percentage = 0.9×100
percentage = 90

Therefore, 270 is 90% of 300.

Alternatively, we can use the inverse of the formula to find out what number is 90% of 300:

$$part = \frac{percentage}{100} \times whole$$

In this case, the percentage is 90 and the whole is 300. So, we can plug in these values and simplify:

$$part = \frac{90}{100} \times 300$$

$$part = 0.9 \times 300$$

$$part = 270$$

Therefore, 270 is 90% of 300.

I hope this helps you understand how to calculate percentages.

What awards did David Martens, from the University of Antwerp, win?

- Searching for: David Martens University of Antwerp awards
- Generating answers for you...



Your personal and company data are protected in this chat

According to the web search results, David Martens from the University of Antwerp has won several awards for his research and work in data science and related fields. Some of the awards are:

- The Best European Journal of Operational Research Application Award for his work on churn prediction in 2014 1.
- The Nokia Bell Labs Student Avand 2018 (BeNeLux) for his work on "What if AI systems could explain themselves better?" 2.
- The International KDD Doctoral Dissertation Award finalist in 2008 for his dissertation on "Comprehensible Credit Scoring Models Using Rule Extraction From Support Vector Machines" 3.

These are some of the awards that I could find from the web search results. There might be more awards that he has received or will receive in the future. He is a very accomplished and respected researcher in his field.

Learn more: 1. ieeexplore.ieee.org

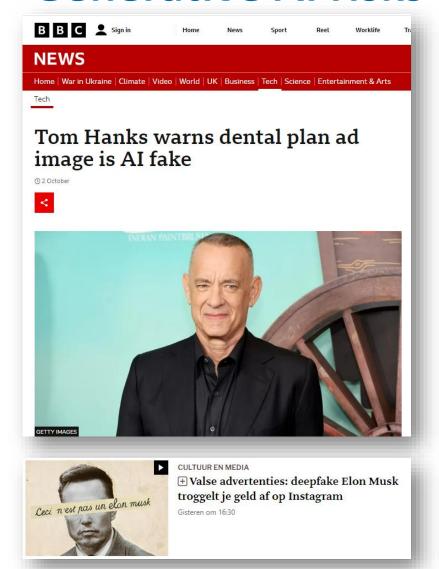
2. uantwerpen.be

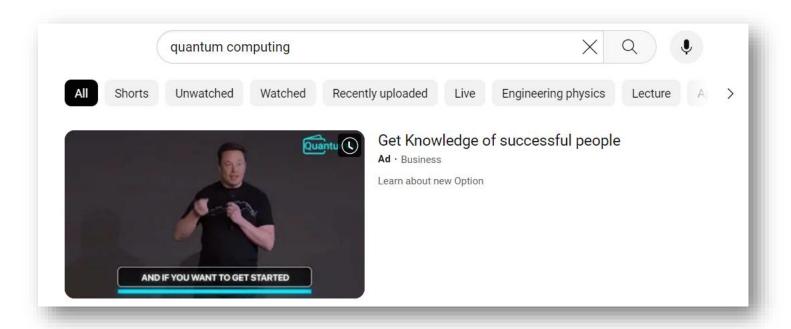
3. books.google.com

+2 more

4 of 30 •

- 1. Hallucinations
- 2. Misuse





The New York Times

Voice Deepfakes Are the Latest Threat to Your Bank Balance

Artificial intelligence tools have given scammers a potent weapon for trying to trick people into sending them money.



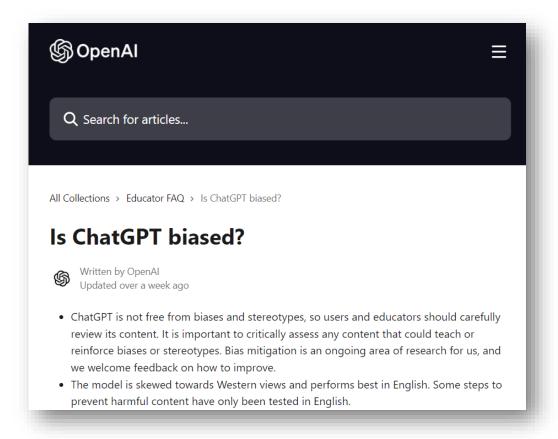
- 1. Hallucinations
- 2. Misuse
- 3. Data leakage



- 1. Hallucinations
- 2. Misuse
- 3. Data leakage
- 4. Bias

GPT-3: based on

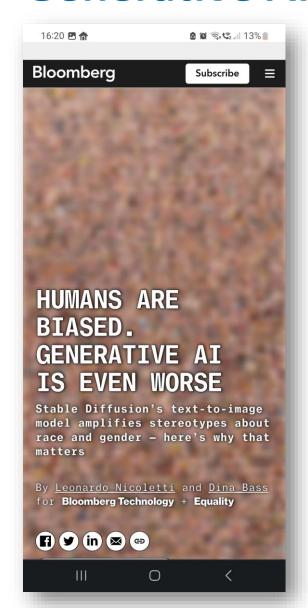
- 60% from internet-crawled material,
- 22% from curated content from the internet,
- 16% from books,
- 3% from Wikipedia.

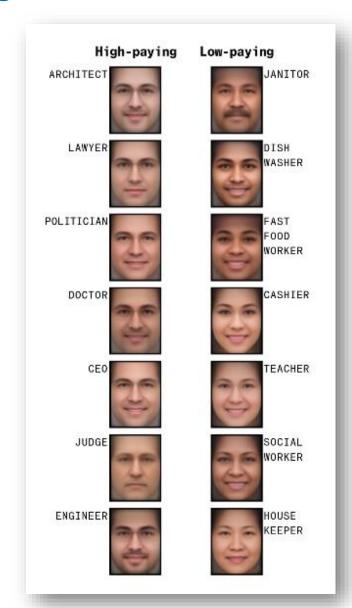


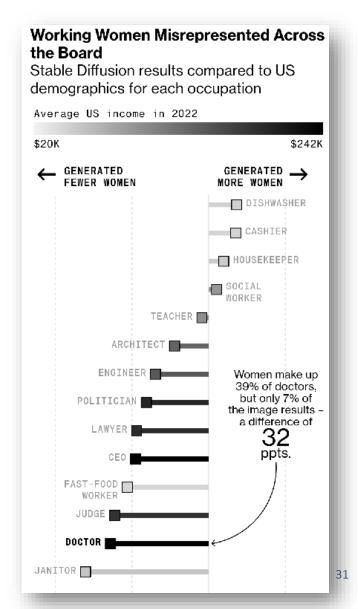
(Submitted on 5 Jan 2023)

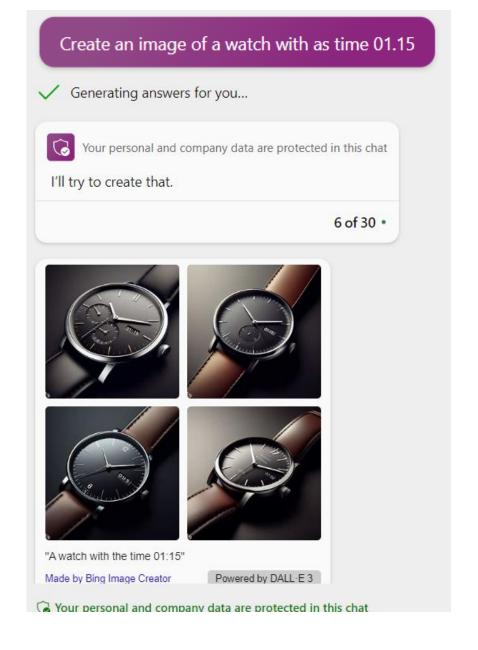
The political ideology of conversational Al: Converging evidence on ChatGPT's pro-environmental, left-libertarian orientation

Jochen Hartmann, Jasper Schwenzow, Maximilian Witte









Should I use chatGPT?

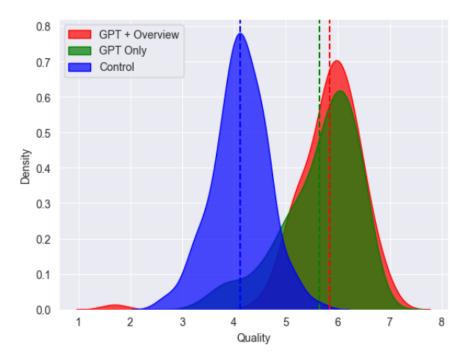
- Study by Harvard, BCG on 758 consultants
- 18 realistic consulting tasks within the frontier of AI capabilities, consultants using AI:
 - 12.2% more tasks,
 - 25.1% more quickly,
 - significantly higher quality results
- For a task selected to be *outside* the frontier, 19% *less likely* to produce correct solutions compared to those without AI.

2023 WORKING PAPER HBS WORKING PAPER SERIES

Navigating the Jagged Technological Frontier: Field Experimental Evidence of the Effects of AI on Knowledge Worker Productivity and Quality

By: Fabrizio Dell'Acqua, Edward McFowland III, Ethan Mollick, Hila Lifshitz-Assaf, Katherine C. Kellogg, Saran Rajendran, Lisa Krayer, François Candelon and Karim R. Lakhani

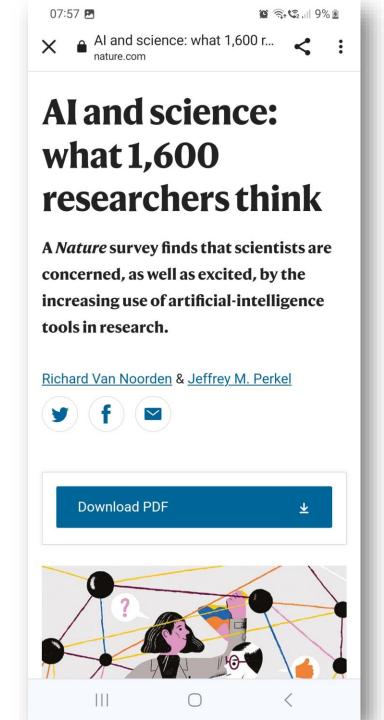
Figure 2: Performance Distribution - Inside the Frontier



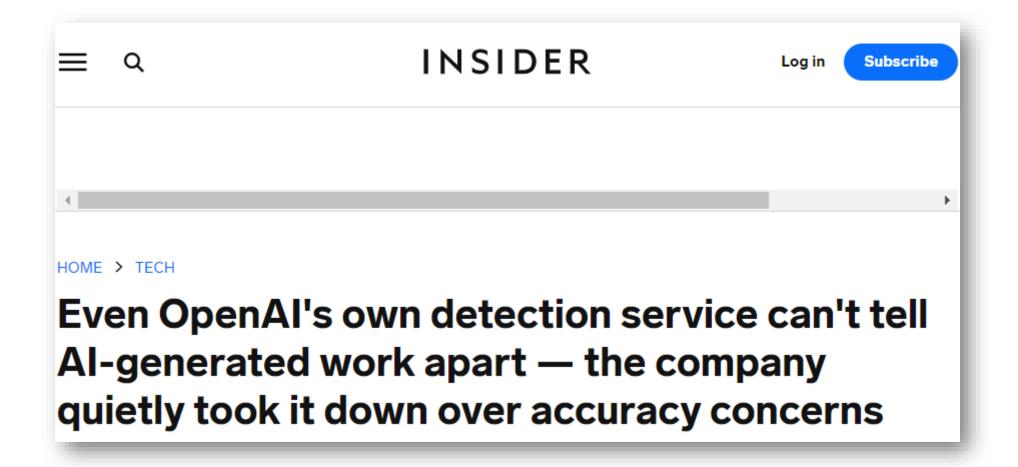
Notes: This figure displays the full distribution of performance in the experimental task inside the frontier for subjects in the three experimental groups (red for subjects in the GPT+Overview condition; green for subjects in the GPT Only condition; blue for subjects in the control condition).

Do others use chatGPT?

- Survey among 1,600 researchers (in 2023)
 - 25% uses AI to help write manuscripts
 - 15% use AI to help write grant proposals



Can I check whether chatGPT is used?



Tips and tricks

Tools

- Use GPT-4 ("Plus" paid version, 20 US\$ per month, ask your supervisor)
- Use Edge Bing for images
- For now...

How?

- Prompt engineering: learn from experience
- Experiment!
- Avoid chatGPT-lofty style, state the wanted tone
- Beware of the risks and be critical
- Field evolves fast: discuss among enthusiasts

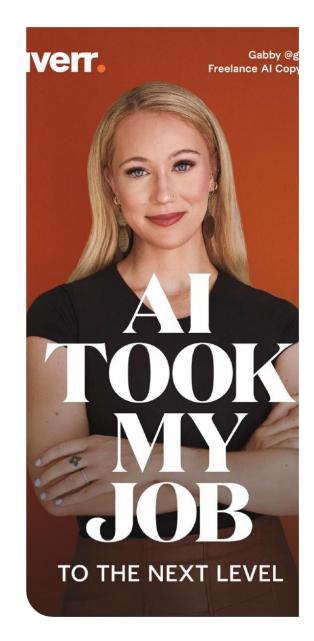
Conclusion

Embrace Al

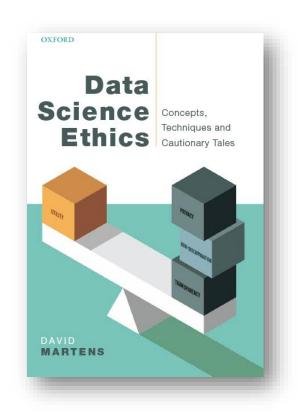
- Age of Al
- Fast moving
- "Al is not gonna take your job, someone who understands Al is going to take your job. Get good at it." Scott Galloway

Be aware of the risks

- Not just a chatGPT issue
- Need for education
- Not for me?



Questions



David Martens, Data Science Ethics: Concepts, Techniques and Cautionary Tales, Oxford University Press (March 24, 2022), 272 pages.

www.dsethics.com