

Generative AI in medicine and life science

Sonja Aits

Cell Death, Lysosomes and Artificial Intelligence Group,
Lund University

AI Lund – Coordinator for AI in Medicine and
Life Science

COMPUTE research school – Study director

sonja.aits@med.lu.se



Generative AI

= AI that focuses on **creating new data**

AI chatbots

ChatGPT

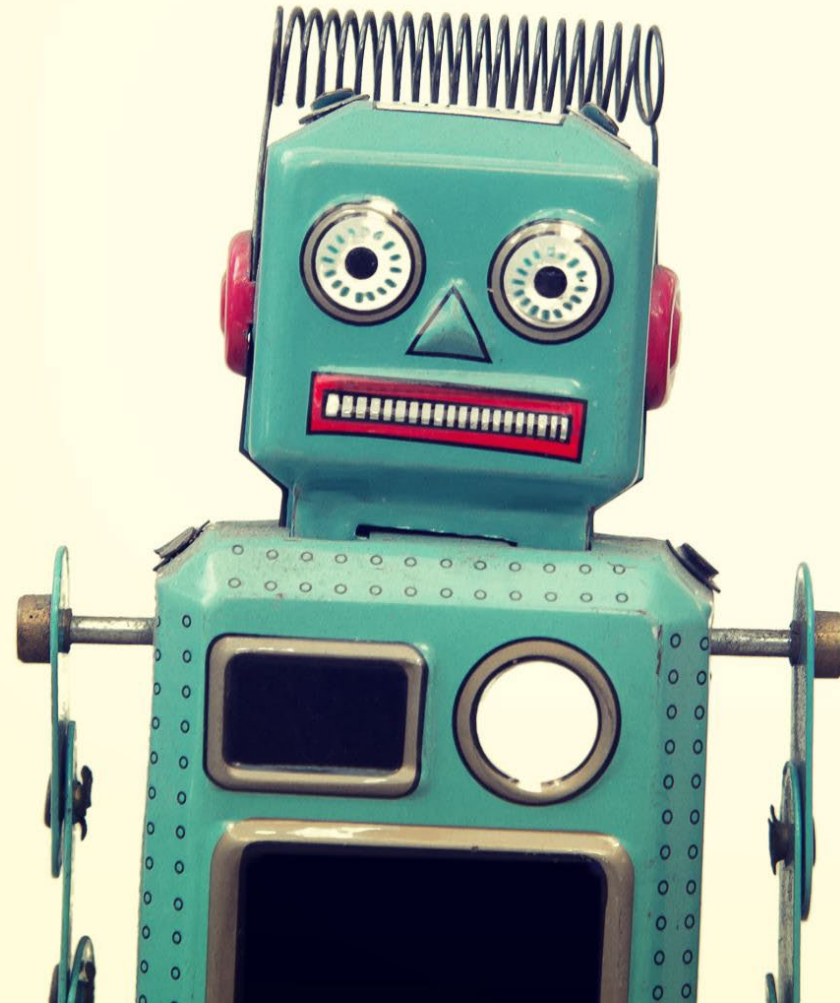
Google Gemini

HuggingChat

Microsoft Copilot

Claude

Perplexity



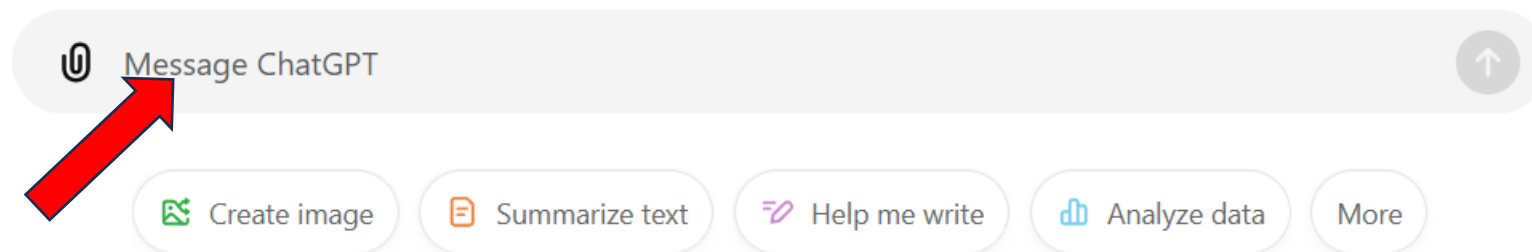


Application areas

AI chatbots for beginners - ChatGPT

<https://chatgpt.com/>

What can I help with?



The screenshot shows the ChatGPT interface. At the top, there is a text input field with the placeholder text "Message ChatGPT". A red arrow points to the paperclip icon on the left side of this field. Below the input field, there are five buttons: "Create image" (with a green icon), "Summarize text" (with an orange icon), "Help me write" (with a purple icon), "Analyze data" (with a blue icon), and "More" (with a grey icon).

Prompt (= your question or task description)

Writing assistant: Title suggestions

Give me 15 presentation title ideas for an article discussing the benefits of natural language processing for clinicians



Certainly! Here are 15 presentation title ideas highlighting the benefits of Natural Language Processing (NLP) for clinicians:

1. "Empowering Clinicians: Unveiling the Benefits of Natural Language Processing"
2. "Enhancing Clinical Care: The Transformative Power of NLP"
3. "From Data to Insights: NLP's Role in Revolutionizing Clinical Practice"
4. "Streamlining Healthcare: How NLP Benefits Clinicians and Patients"
5. "Unlocking Clinical Insights: Harnessing the Potential of Natural Language Processing"
6. "Precision Medicine and Beyond: Leveraging NLP to Improve Clinical Decision-Making"
7. "Efficiency and Accuracy: How NLP Is Reshaping Clinicians' Workflows"
8. "Improving Patient Outcomes: The Role of NLP in Informed Clinical Decisions"
9. "NLP in Healthcare: Bridging the Gap Between Data and Clinical Action"
10. "Clinical Narratives Unleashed: How NLP Transforms Text into Insights"

Writing assistant: Outline generation

Prompt: “write an outline for a scientific review on cell death”



I. Introduction to Cell Death

- A. Importance of cell death in development and homeostasis
- B. Types of cell death: apoptosis, necrosis, autophagy
- C. Significance of understanding cell death mechanisms

II. Apoptosis: Programmed Cell Death

- A. Molecular pathways and key regulators
 - 1. Caspase activation and execution
 - 2. Bcl-2 family proteins and mitochondrial pathway
 - 3. Death receptor-mediated pathway
- B. Physiological and pathological roles of apoptosis
 - 1. Tissue remodeling during development
 - 2. Immune system regulation
 - 3. Apoptosis in cancer and neurodegenerative diseases

III. Necrosis and Its Variants

- A. Morphological and biochemical characteristics of necrosis

Writing assistant: Production of complete text

Prompt:

“write a script for a 2 min video that explains the process of gene expression in a clear and concise way to biomedical university students ”

Hi everyone, and welcome to this video about gene expression. Gene expression is the process by which DNA is converted into proteins. Proteins are the building blocks of cells and tissues, and they play a role in almost every function of the body.

The first step in gene expression is transcription...

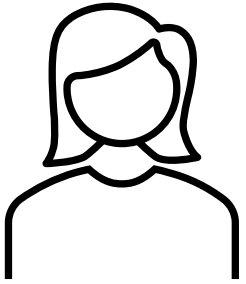
Writing assistant: Summarization + style change + translation



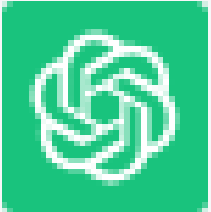
"Lysosomal membrane permeabilization (LMP) påverkar vävnadsförändringar, sjukdomar och cancerbehandling. Att upptäcka LMP har varit svårt. En ny känslig metod använder sig av galectin-punkter på läckande lysosomer. Galectin-1 och galectin-3 fungerar bra på grund av deras vanliga närvaro, snabba rörelse till läckande lysosomer och starka antikroppar. Galectin-färgning visar läckande lysosomer tidigt under lysosomala celldöd, vilket hjälper till att identifiera primära eller sekundära orsaker till celldöden. Metoden avslöjar att celler kan överleva viss LMP och snabbt bilda autofagiska strukturer vid galectin-punktplatser. Denna färgning fungerar även på vävnadsprover och upptäcker LMP i tumör-xenografts behandlade med vissa läkemedel och i involution av bröstkörtlar. Fluorescerande galectiner gör metoden användbar för automatisk screening av LMP i levande celler och djur. Galectin-punktester öppnar nya forskningsmöjligheter för LMP vid celldöd, autofagi, åldrande och inflammation."



Personal teacher



“explain what a Western Blot is”



A Western blot, also known as protein immunoblot, is a widely used laboratory technique in molecular biology and biochemistry. It is employed to detect and analyze specific proteins within a complex mixture of proteins extracted from cells or tissues. This technique provides information about the presence, size, and sometimes the quantity of a target protein.

The process involves several steps:

“how can lysosomal cell death be measured in cultured cancer cells?”
”which of these assays is best?”

Scientific discussion partner



The choice of assay for measuring lysosomal cell death in cultured cancer cells depends on various factors, including the specific goals of your research, the characteristics of the cells you are working with, and the resources available to you. Each assay has its advantages and limitations. Here's a breakdown of the pros and cons of some commonly used assays:

1. **Acridine Orange Staining:**

- Pros: Simple and cost-effective. Provides a visual indicator of lysosomal membrane permeabilization.
- Cons: May not provide quantitative data. Requires fluorescence microscopy or flow cytometry.

2. **Lysosomal Enzyme Release Assays:**

- Pros: Direct measurement of lysosomal enzyme leakage. Can provide quantitative data.
- Cons: Enzyme activity could be influenced by other factors. Requires specialized enzyme assays.

3. **Lysosome Tracker Dyes:**

- Pros: Specifically labels lysosomes. Provides real-time monitoring of lysosomal integrity.
- Cons: May not be as quantitative as enzymatic assays. Requires fluorescence-based analysis.

4. **Lysosome Permeabilization Assays:**

- Pros: Can directly measure the release of lysosomal contents using fluorescent substrates.

Regenerate

ChatGPT

Programming assistant

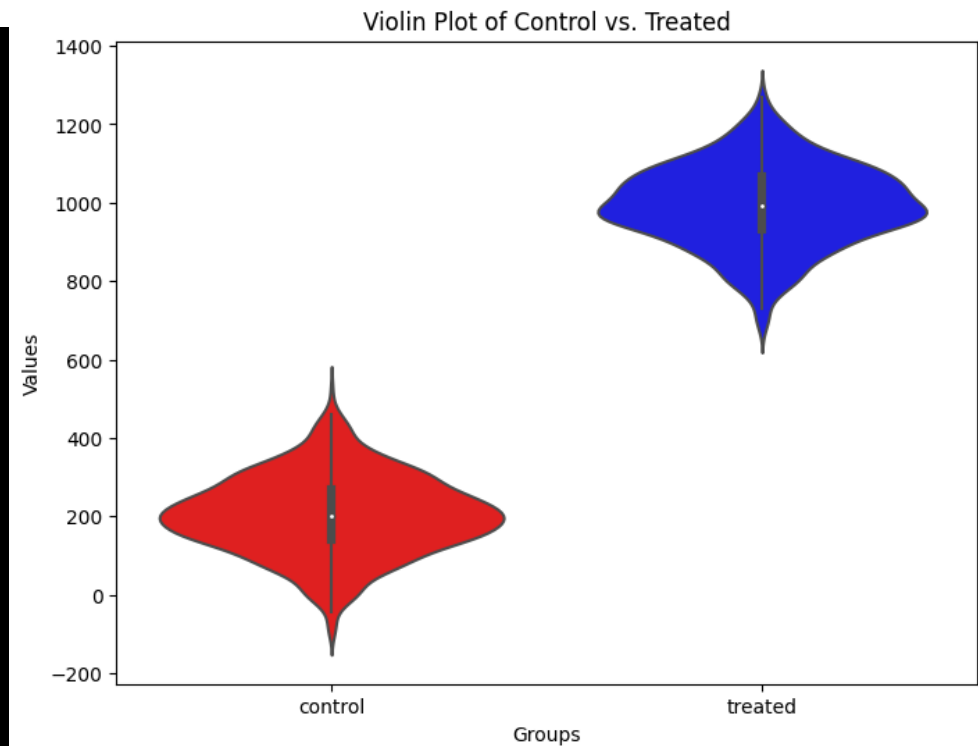
“generate a script that does the following: 1. loads the values from a pandas dataframe called df which has two columns called "control" and "treated". 2. plots the two columns as two data series in a violin plot with the control column values in red and the treated column values in blue”

```
# Set up the figure and axes for the plot
plt.figure(figsize=(8, 6))

# Create a violin plot using seaborn
sns.violinplot(data=df, palette={"control": "red", "treated": "blue"})

# Add labels and title
plt.xlabel("Groups")
plt.ylabel("Values")
plt.title("Violin Plot of Control vs. Treated")

# Show the plot
plt.show()
```



Personal assistant

Create a table of the most common causes of death in the elderly, including the following columns: cause of death, number of deaths per year, risk factors. save as excel file, named causes-of-death. shade the header column in dark blue with white text and then alternate the columns in white and light blue shading with black text.

The Excel file with the table of the most common causes of death in the elderly, along with the specified formatting, has been created. You can download it using the link below:

[Download causes-of-death.xlsx](#) [↗]

Personal assistant

Create a table of the most common causes of death in the elderly, including the following columns: cause of death, number of deaths per year, risk factors. save as excel file, named causes-of-death. shade the header column in dark blue with white text and then alternate the columns in white and light blue shading with black text.

Cause of Death	Number of Deaths per Year	Risk Factors
Heart Disease	697000	High blood pressure, high cholesterol, smoking, obesity, physical inactivity
Cancer	601000	Age, family history, smoking, diet, alcohol consumption
Chronic Lower Respiratory Disease	160000	Smoking, exposure to pollutants, genetic factors
Stroke	150000	High blood pressure, high cholesterol, smoking, obesity
Alzheimer's Disease	120000	Age, family history, genetics, heart disease
Diabetes	85000	Obesity, inactivity, genetics, high blood pressure
Kidney Disease	51000	High blood pressure, diabetes, smoking, obesity
Pneumonia/Influenza	50000	Older age, weakened immune system, smoking, chronic diseases

Personal assistant

convert this to a powerpoint presentation with one disease per slide.
add a cartoon illustrating each disease to each slide

Cancer



Deaths per year: 601,000 per year

Risk Factors:

Age, family history, smoking, diet, alcohol consumption

Data analyst

S

Can you calculate the average number of deaths from this and make a histogram with the causes of death: Cause of DeathNumber of Deaths per YearHeart

Disease697000Cancer601000Chronic Lower Respiratory

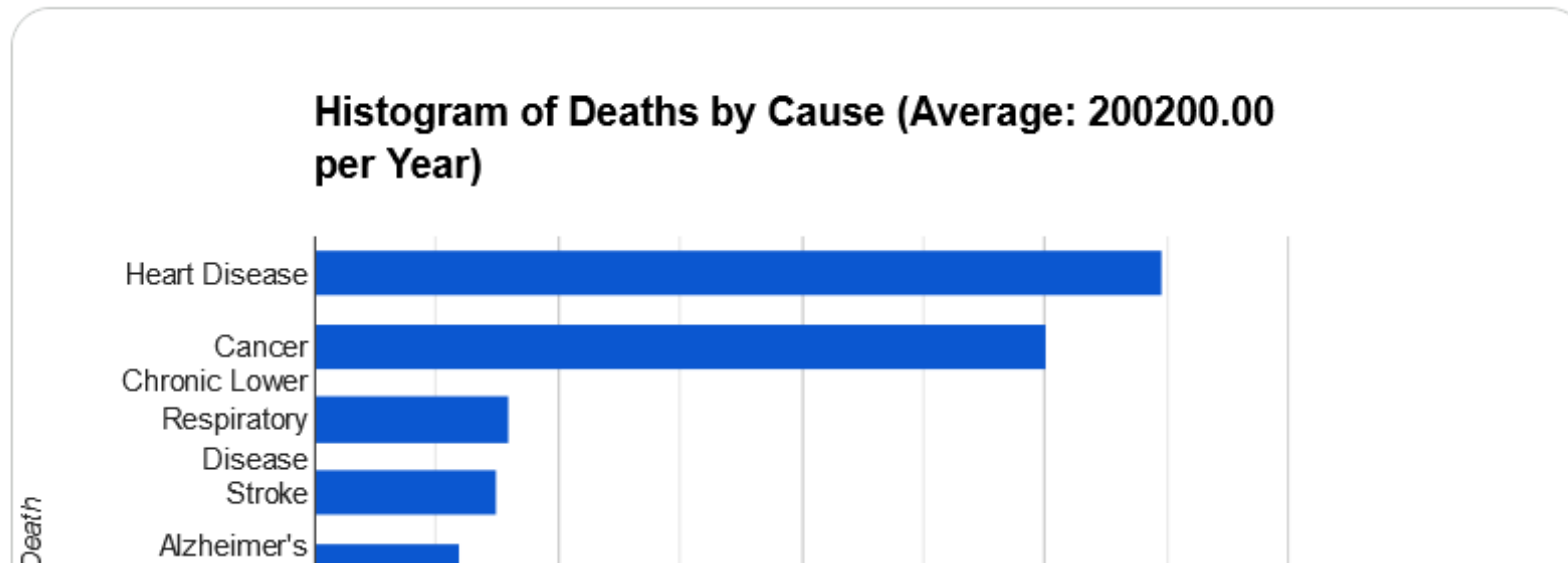
Disease160000Stroke150000Alzheimer's Disease120000Diabetes85000Kidney

Disease51000Pneumonia/Influenza50000Accidents46000Liver Disease42000

Show drafts



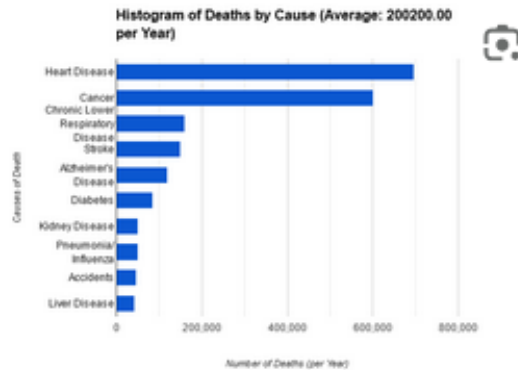
Sure, here is the graph:



Data analyst

S

Describe this graph in three sentences



Certainly! Here is a description of the graph in three sentences:

The histogram illustrates the leading causes of death in the United States, with heart disease and cancer accounting for the highest number of deaths per year. The remaining causes show a significant decrease in the number of deaths, with liver disease contributing to the lowest number of fatalities annually.

Image designer

what images or graphics could you use to convey the concept of healthy ageing in a presentation?

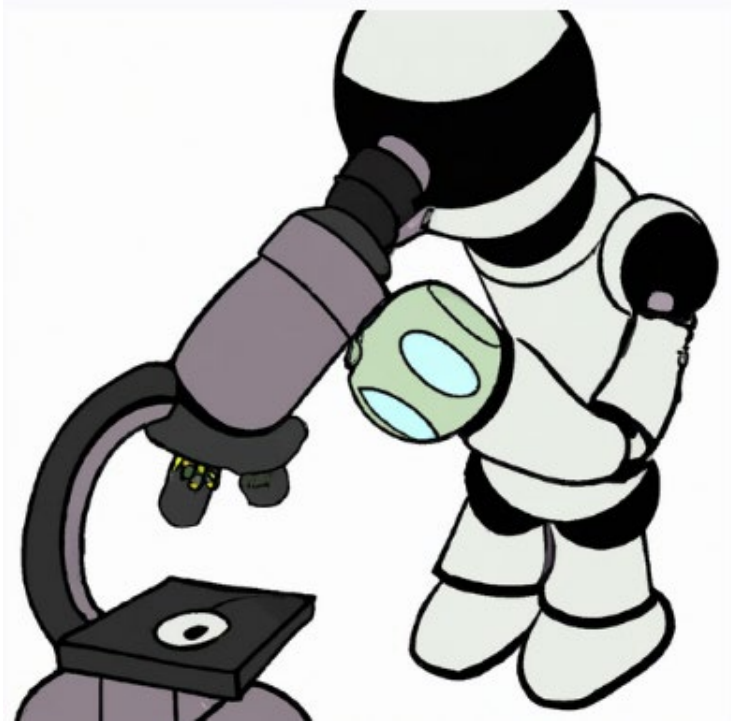
To visually convey **healthy aging** in a presentation, you might consider the following imagery and graphics:

1. **Active Seniors:** Photos or illustrations of older adults engaging in various activities like walking, yoga, dancing, or gardening. These show physical activity, vitality, and engagement.
2. **Balanced Nutrition:** Pictures of balanced, colorful meals with fruits, vegetables, lean proteins, and whole grains, symbolizing a healthy diet as a cornerstone of aging well.
3. **Social Connections:** Images of seniors interacting with family, friends, or community groups, showing the importance of social support and engagement.
4. **Brain Health:** Icons or images of brain-related activities, like puzzles, reading, or learning, to emphasize cognitive health and mental stimulation.
5. **Medical Check-Ups and Self-Care:** Visuals of routine health check-ups or self-care activities, such as an older adult at a doctor's office or engaging in personal care practices.

Image designer

DALL-E2

“a cartoon image of a white robot looking into a microscope”



AI chatbots can support you in many ways

- Writing assistant
 - Title suggestions
 - Content suggestions
 - Outline generation
 - Rephrasing/Correction/Shortening
 - Summarization/Style change/Translation
 - Production of complete text
- Discussion partner
- Personal teacher
- Programming assistant
- Personal assistant
- Data analyst
- Image designer



Best practices

Be clear and specific

- Write in English (other languages possible)
 - Give context
 - Be specific
 - **Bad:** *How can you measure DNA damage?*
 - **Good:** *Explain different methods for assessing DNA damage in a cancer cell culture to assist with the design of new experiments. For each method, summarize the underlying principles, advantages and disadvantages/limitations. Present the results in bullet points*
- Clarity ≠ Brevity
- Give examples of desired output
 - No rambling
 - No polite phrases

Add structure

- Delimiters, e.g. “ “, “”” “””, < >

Summarize this: “”” Your lengthy text containing many quotes designated with “....” and many paragraphs. “””

- Numbered sequence

Write python code that does the following: 1. Loads all tsv files in the folder XYZ, 2. Generate a scatter plot with column A as x and column B as y, 3. Save the scatter plot in folder XYZ in jpeg format and use the name of the tsv file + _plot as file name. 4. ...

Don't waste time on fixing minor errors

- Spelling
- Grammar
- Phrasing
- Capitalization

what is the abcxe steps fo first aid assessment?

< 2/2 >

The **ABCDE** steps for first aid assessment provide a structured approach to assessing and managing a patient, particularly in emergencies. Each letter stands for a critical component to check:

Use more than one chatbot

ChatGPT

Google Gemini

HuggingChat

Microsoft Copilot

Claude

Perplexity

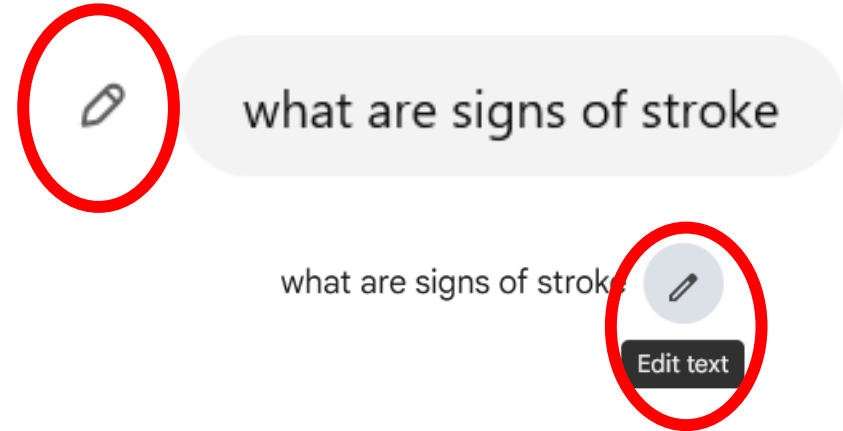
...

Ask for structured output and files

- Numbered lists
- Bullet points
- Tables
- Specific file formats
 - text documents (e.g. docx, pdf, LaTeX)
 - data files (e.g. json, csv, xlsx)
 - web pages (e.g. html, css)
 - scripts (e.g. py, sh)
 - presentations (e.g. pptx, Google slides)
 - images (e.g. jpg, tiff)

Keep trying but know when to stop

- Refine your prompts
- Follow-up questions
- Point out errors/desired improvements
- For scripts: copy and paste error code



- Stop after multiple tries failed to deliver the desired outcome

Keep in mind that you are sharing information!

- No sensitive data
- No confidential information



Keep you brain turned on!

A chain-link fence is in the foreground, slightly out of focus. The background shows a green field and distant hills under a hazy sky.

Limitations

Watch out!

- Made-up results (“hallucinations”)
- No control or information on training data
 - Biases
 - Intellectual property issues
- Check journal policies for AI
- Environmental footprint
- Ethical/legal/societal issues

AI **for** sustainability



Sustainability **of** AI

Lack of traceability

Can you give the source of the cause of death numbers you listed?



The figures for causes of death among the elderly I included were based on commonly cited statistics from public health sources such as:

1. **Centers for Disease Control and Prevention (CDC):** The CDC frequently publishes data on causes of death in the U.S. in their annual "Leading Causes of Death" reports. They provide age-specific data, including for older adults.
2. **World Health Organization (WHO):** The WHO publishes global and region-specific mortality data, including causes of death by age group, through their Global Health Estimates.
3. **National Center for Health Statistics (NCHS):** This organization, part of the CDC, offers detailed mortality data that includes the leading causes of death for different age groups.

Would you like the most recent numbers from these sources? If so, I can retrieve the latest available data for you.

Lack of traceability

No, I want to know where the numbers you used in the table came from



The numbers I initially provided for the table were estimated based on general trends from past CDC and WHO data rather than from a specific, up-to-date source. For exact and current statistics on causes of death, I would recommend consulting the latest **CDC Leading Causes of Death Report** or **WHO Global Health Estimates**, which provide verified and specific annual figures for different causes of death, especially in older populations.

If you need the exact current numbers, I can look up the latest available data for you.

Journals infiltrated with 'copycat' papers that can be written by AI

Tools such as ChatGPT can be used to generate almost-identical research papers that pass standard plagiarism checks. Hundreds are thought to have been published.

What else can you do with
generative AI?

Example applications for generative AI

- Designing new molecules: drugs, proteins, RNA/DNA sequences, ...
- Artificial data – software evaluation, model training, data augmentation
- Educational materials (e.g. simulate patients)
- Custom-made knowledge bases

Article

Learning the natural history of human disease with generative transformers

<https://doi.org/10.1038/s41586-025-09529-3>

Received: 18 May 2024

Artem Shmatko^{1,2,3,13}, Alexander Wolfgang Jung^{2,4,5,6,13}, Kumar Gaurav^{2,13}, Søren Brunak^{4,7},
Laust Hvas Mortensen^{5,7,8}, Ewan Birney^{2✉}, Tom Fitzgerald^{2✉} & Moritz Gerstung^{1,2,9,10,11,12✉}

- Trained on data from 0.4 million UK Biobank participants
- predicts rates of more than 1,000 diseases, conditional on individual disease history

Summary

- AI chatbots are fantastic support tools for scientists
- AI chatbots lack transparency
- Standard AI Chatbots are not suitable for sensitive/confidential data
- Generative AI can be used for many other life science applications and also with other types of data