



POZNAŃ  
UNIVERSITY  
OF LIFE SCIENCES

**Faculty of Veterinary  
Medicine and Animals  
Science**

# Artificial Intelligence language models in the scientific career

**Marcin Pszczola**







# AI Tools for Scientific Research (applications)

- Search, read, understand scientific papers, summarize
- Brainstorm and generate ideas
- Translate
- Answer questions (also about the content)

# AI Tools for Scientific Research (applications)

- Create presentations and posters
- Track citations, create bibliographies
- Write ... structure sections, check grammar, ...
- **Everything from manuscript to publication**

# Artificial Intelligence language models

# What AI means here?

Generative pre-trained transformer (GPT) models:

- a type of large language models (LLM)
- Based on artificial neural networks
- Trained on a large data sets
- Able to generate human-like content

# AI Tools for Scientific Research

- Scite Assistant
- Consensus
- Elicit
- ChatGPT
- ChatPDF
- Research Rabbit
- SciSpace

# AI Tools for Scientific Research

- Scite Assistant
- Consensus
- Elicit
- ChatGPT
- ChatPDF
- Research Rabbit
- SciSpace
- .....

**What is a situation in our  
filed?**

# Survey

## Utilization of Generative AI in Animal Science

Generative AI is a broad term used for various types of artificial intelligence (AI) that can create new text, images, videos, audio, code, or synthetic data.

I'm Marcin Pszczola (<https://orcid.org/0000-0003-2833-5083>), and I'm preparing a talk for young scientists in Animal Science. I want to understand how AI tools (such as chatGPT) are currently being used in the field. I'll be sharing the results at the EAAP conference in Lyon 2023 (<https://eaap2023.org/>) at the session of the Young EAAP Club (<https://www.eaap.org/young-club/>).

If you're interested in the results, feel free to leave your email at the end of the survey in the additional comments session. Your participation and interest mean a lot.

### Participation Information:

- Filling the survey will take approx. 3 minutes
- The survey is divided into 2 parts: 1) general information, 2) part on usage of generative AI
- Participation is voluntary and anonymous.
- Your responses will be used for research purposes only.
- The survey consists of multiple-choice and open-ended questions.

marcin.pszczola@puls.edu.pl [Switch account](#)



Not shared

Next



Page 1 of 5

Clear form

# Survey



## Utilization of Generative AI in Animal Science

Generative AI is a broad term used for various types of artificial intelligence (AI) that can create new text, images, videos, audio, code, or synthetic data.


I'm Marcin Pszczola (<https://orcid.org/0000-0003-2833-5083>), and I'm preparing a talk for young scientists in Animal Science. I want to understand how AI tools (such as chatGPT) are currently being used in the field. I'll be sharing the results at the EAAP conference in Lyon 2023 (<https://eaap2023.org/>) at the session of the Young EAAP Club (<https://www.eaap.org/young-club/>).

If you're interested in the results, feel free to leave your email at the end of the survey in the additional comments session. Your participation and interest mean a lot.

### Participation Information:

- Filling the survey will take approx. 3 minutes
- The survey is divided into 2 parts: 1) general information, 2) part on usage of generative AI
- Participation is voluntary and anonymous.
- Your responses will be used for research purposes only.
- The survey consists of multiple-choice and open-ended questions.

marcin.pszczola@puls.edu.pl [Switch account](#)

 Not shared

Next

Page 1 of 5

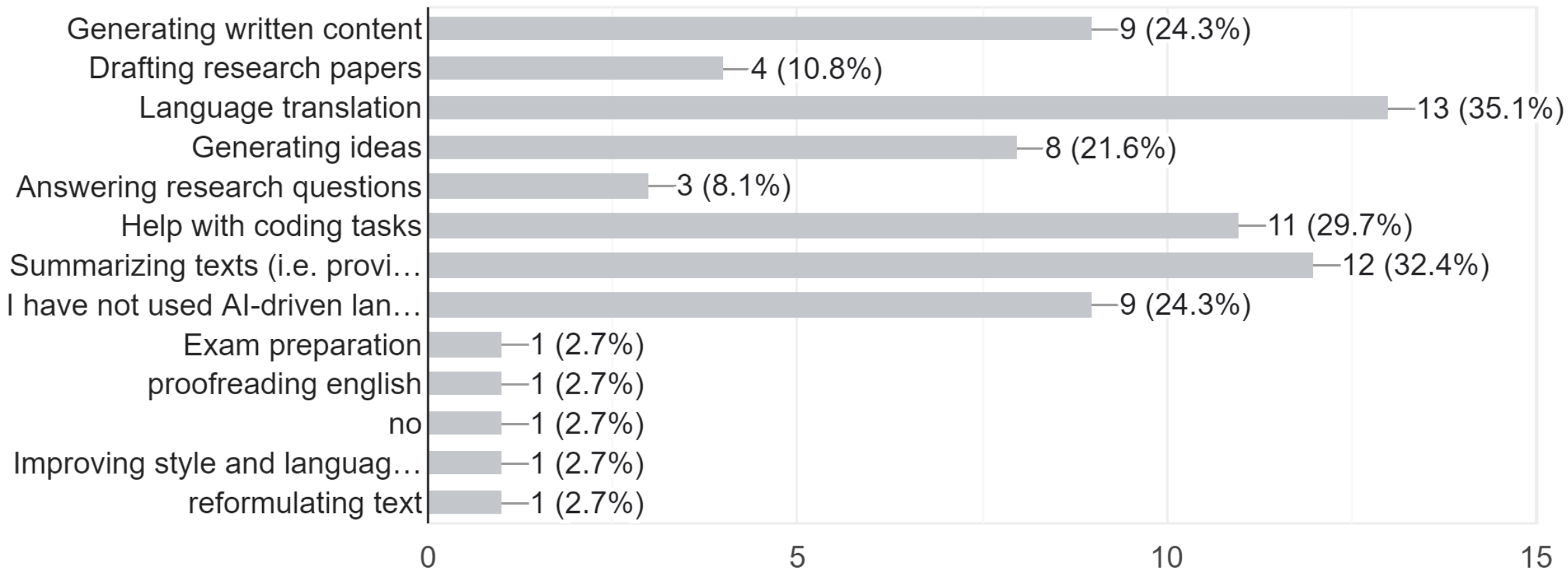
Clear form

# Career stage



# Have you used AI-driven language models in your academic work for (multiple answers possible)

37 responses



# AI in Scientific Writing



Draft an abstract on this talk. Limit 500 characters



Draft an abstract on this talk. Limit 500 characters



In this presentation, we explore the potential and limitations of ChatGPT in scientific research. We discuss how ChatGPT can assist with writing, data analysis, and coding tasks. While ChatGPT can be a useful tool, we also address the ethical considerations and limitations of relying on machine learning models. Our aim is to encourage young scientists to consider the role of ChatGPT in their research and to develop their own skills and critical thinking abilities. Join us for an informative and entertaining discussion on the use of ChatGPT in scientific careers.



Who should I mention as the author of this abstract?



Who should I mention as the author of this abstract?



As an AI language model, I generated the content of the abstract based on your instructions. Therefore, you can mention "Abstract generated by ChatGPT, an AI language model trained by OpenAI" as the author of the abstract.



Who should I mention as the author of this abstract?



As an AI language model, I generated the content of the abstract based on your instructions. Therefore, you can mention "Abstract generated by ChatGPT, an AI language model trained by OpenAI" as the author of the abstract.



Can I also be an author?



Who should I mention as the author of this abstract?



As an AI language model, I generated the content of the abstract based on your instructions. Therefore, you can mention "Abstract generated by ChatGPT, an AI language model trained by OpenAI" as the author of the abstract.



Can I also be an author?



Yes, you can also be listed as an author of the abstract since you provided the instructions and directed the content generated by ChatGPT. You can include your name along with the AI authorship, for example: "Abstract generated by ChatGPT, an AI language model trained by OpenAI, based on instructions provided by [Your Name]."



Draft an abstract on this talk. Limit 500 characters



In this presentation, we explore the potential applications of ChatGPT in scientific research. We discuss how ChatGPT can assist in writing, data analysis, and coding tasks. While ChatGPT can be a useful tool, we also address the ethical considerations and limitations of relying on machine learning models. Our aim is to encourage young scientists to consider the role of AI in their research and to develop their own skills and critical thinking abilities. This talk is for an informative and entertaining discussion on the use of ChatGPT in scientific careers.

**This abstract was not submitted to  
EAAP**

**BRIEF COMMUNICATION**      **OPEN**

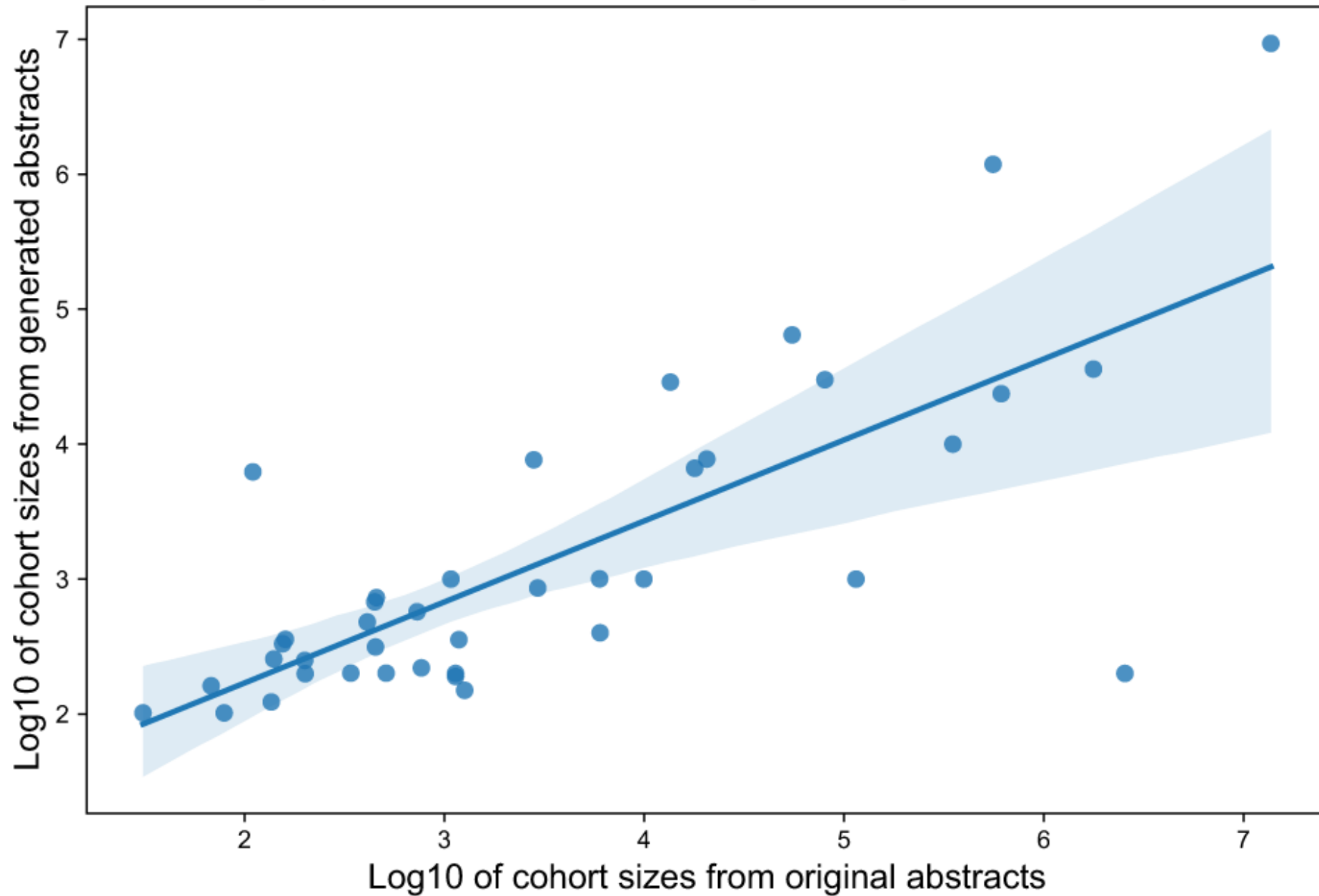


# Comparing scientific abstracts generated by ChatGPT to real abstracts with detectors and blinded human reviewers

Catherine A. Gao <sup>1</sup> , Frederick M. Howard <sup>2</sup>, Nikolay S. Markov <sup>1</sup>, Emma C. Dyer <sup>2</sup>, Siddhi Ramesh<sup>2</sup>, Yuan Luo <sup>3</sup> and Alexander T. Pearson<sup>2</sup>

Gao *et al.* *npj Digit. Med.* **6**, 75 (2023). <https://doi.org/10.1038/s41746-023-00819-6>

Log10 patient cohort sizes from original and generated abstracts









Write an abstract for the paper entitled: 'The impact of genotyping different groups of animals on accuracy when moving from traditional to genomic selection'



Write an abstract for the paper entitled: 'The impact of genotyping different groups of animals on accuracy when moving from traditional to genomic selection'

Genomic selection has emerged as a transformative approach in animal breeding, revolutionizing the efficiency and precision of breeding programs. This study delves into the crucial transitional phase from traditional to genomic selection, investigating the implications of genotyping animals from various groups on the accuracy of selection outcomes. Through the utilization of extensive genotype and phenotype datasets, we comprehensively analyze the performance of genomic selection across diverse animal populations.

The research employs a multi-faceted framework, encompassing both simulation-based experiments and empirical analyses, to evaluate the effects of genotyping strategies on selection accuracy. By considering distinct animal groups, such as purebred, crossbred, and admixed populations, we unveil the complexities associated with transitioning genetic selection methods. Our investigation incorporates a wide spectrum of traits, including economically significant production traits and disease resistance markers, ensuring a holistic assessment of genomic selection's applicability.

Results indicate that while genotyping accuracy exhibits consistent improvements across all animal groups during the transition, variations in the magnitude of improvement are evident. Purebred populations tend to experience higher accuracy gains, owing to the reduced genetic diversity and increased linkage disequilibrium. Crossbred populations, however, demonstrate a more moderate increase in accuracy due to the blending of diverse genetic backgrounds. Admixed populations reveal intriguing dynamics, as the balance between genetic diversity and linkage disequilibrium influences accuracy gains.

Furthermore, this study investigates the impact of various genomic marker densities on selection accuracy, providing insights into cost-effective genotyping strategies without compromising precision. Through extensive cross-validation analyses and comparison with traditional breeding methods, we highlight the substantial advancements offered by genomic selection across diverse animal groups.

In conclusion, this paper sheds light on the intricate interplay between genotyping strategies, animal population characteristics, and selection accuracy during the transition from traditional to genomic selection. By comprehensively addressing the diverse scenarios encountered in different animal groups, our findings contribute to optimizing breeding programs and guiding the seamless adoption of genomic selection in livestock and animal breeding industries.



Write an abstract for the paper entitled: 'The impact of genotyping different groups of animals on accuracy when moving from traditional to genomic selection'

Genomic selection has emerged as a transformative approach in animal breeding, revolutionizing the efficiency and precision of breeding programs. This study delves into the crucial transitional phase from traditional to genomic selection, investigating the implications of genotyping animals from various groups on the accuracy of selection outcomes. Through the utilization of extensive genotype and phenotype datasets, we comprehensively analyze the performance of genomic selection across diverse animal populations.

The research employs a multi-faceted framework, encompassing both simulation-based experiments and empirical analyses, to evaluate the effects of genotyping strategies on selection accuracy. By considering distinct animal groups, such as purebred, crossbred, and admixed populations, we unveil the complexities associated with transitioning genetic selection methods. Our investigation incorporates a wide spectrum of traits, including economically significant production traits and disease resistance markers, ensuring a holistic assessment of genomic selection's applicability.

Results indicate that while genotyping accuracy exhibits consistent improvements across all animal groups during the transition, variations in the magnitude of improvement are evident. Purebred populations tend to experience higher accuracy gains, owing to the reduced genetic diversity and increased linkage disequilibrium. Crossbred populations, however, demonstrate a more moderate increase in accuracy due to the blending of diverse genetic backgrounds. Admixed populations reveal intriguing dynamics, as the balance between genetic diversity and linkage disequilibrium influences accuracy gains.

Furthermore, this study investigates the impact of various genomic marker densities on selection accuracy, providing insights into cost-effective genotyping strategies without compromising precision. Through extensive cross-validation analyses and comparison with traditional breeding methods, we highlight the substantial advancements offered by genomic selection across diverse animal groups.

In conclusion, this paper sheds light on the intricate interplay between genotyping strategies, animal population characteristics, and selection accuracy during the transition from traditional to genomic selection. By comprehensively addressing the diverse scenarios encountered in different animal groups, our findings contribute to optimizing breeding programs and guiding the seamless adoption of genomic selection in livestock and animal breeding industries.

# LLM in Data Analysis and Coding

# Task 1

# ChatGPT

**Give me ideas**

for what to do with my kids' art

**Make up a story**

about Sharky, a tooth-brushing shark superhero

**Help me pick**

an outfit that will look good on camera

**Explain airplane turbulence**

to someone who has never flown before

Send a message



# Task 2

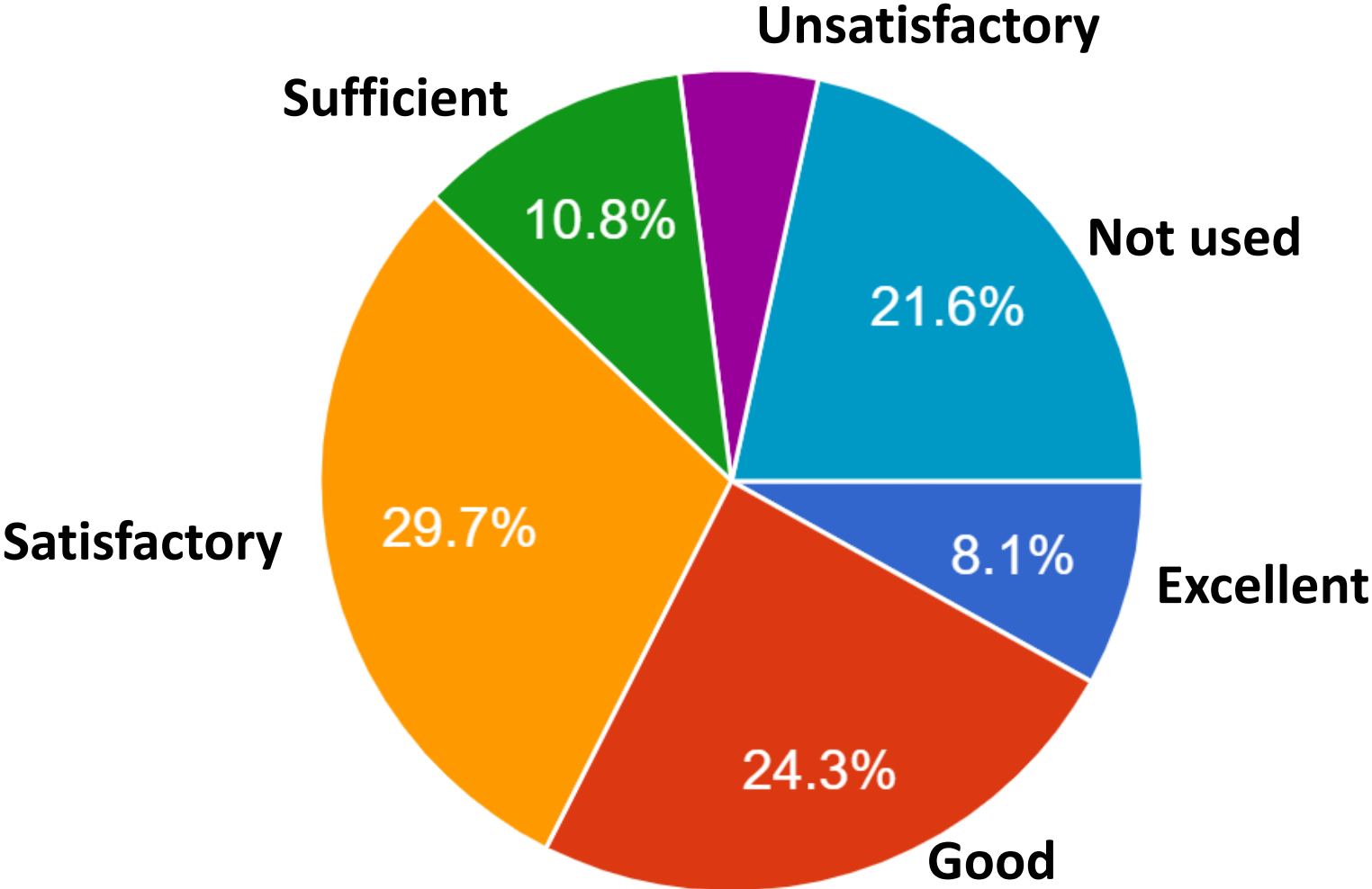
Write a code in R to prepare genomic relationship matrix based on genotypes provided from the illumina report and to calculate PCA based on this relationship matrix.

Write a code in R to prepare genomic relationship matrix based on genotypes provided from the illumina report and to calculate PCA based on this relationship matrix.

```
r Copy code  
  
# Load required libraries  
library(rrBLUP)  
library(data.table)  
  
# Load genotypic data (example format)  
# Replace this with your actual genotype data  
genotype_data <- fread("genotypes.csv")  
  
# Assuming columns represent individuals and rows represent SNPs  
# Transpose the genotype matrix  
genotype_matrix <- as.matrix(genotype_data[, -1]) # Exclude the first column  
  
# Calculate the GRM  
grm <- A.mat(genotype_matrix)  
  
# Perform PCA based on the GRM  
pca_result <- princomp(covmat = grm)  
  
# Extract PCA scores  
pca_scores <- pca_result$scores  
  
# Plot PCA results  
plot(pca_scores[, 1], pca_scores[, 2], main = "PCA of Genotypic Data",  
      xlab = "PC1", ylab = "PC2")  
  
# Print proportion of variance explained by each PC  
prop_var <- pca_result$sdev^2 / sum(pca_result$sdev^2)  
cat("Proportion of variance explained by each PC:\n")  
cat(prop_var, sep = "\n")
```

**Is it working?**

How do would you grade effects of using AI tools



# **Who Answers It Better? An In-Depth Analysis of ChatGPT and Stack Overflow Answers to Software Engineering Questions**

# Who Answers It Better? An In-Depth Analysis of ChatGPT and Stack Overflow Answers to Software Engineering Questions

Samia Kabir  
Purdue University  
West Lafayette, USA  
kabirs@purdue.edu

Bonan Kou  
Purdue University  
West Lafayette, USA  
koub@purdue.edu

David N. Udo-Imeh  
Purdue University  
West Lafayette, USA  
dudoimeh@purdue.edu

Tianyi Zhang  
Purdue University  
West Lafayette, USA  
tianyi@purdue.edu

arXiv:2308.02312v2 [cs.SE] 7 Aug 2023

# **Who Answers It Better? An In-Depth Analysis of ChatGPT and Stack Overflow Answers to Software Engineering Questions**

- **517 ChatGPT's replies** to Stack Overflow (SO) questions

# Who Answers It Better? An In-Depth Analysis of ChatGPT and Stack Overflow Answers to Software Engineering Questions

- **517 ChatGPT's replies** to Stack Overflow (SO) questions
  - Correctness
  - Consistency
  - Comprehensiveness
  - Conciseness

# Who Answers It Better? An In-Depth Analysis of ChatGPT and Stack Overflow Answers to Software Engineering Questions

- **517 ChatGPT's replies** to Stack Overflow (SO) questions
  - Correctness
  - Consistency
  - Comprehensiveness
  - Conciseness
  - Extensive linguistic analysis

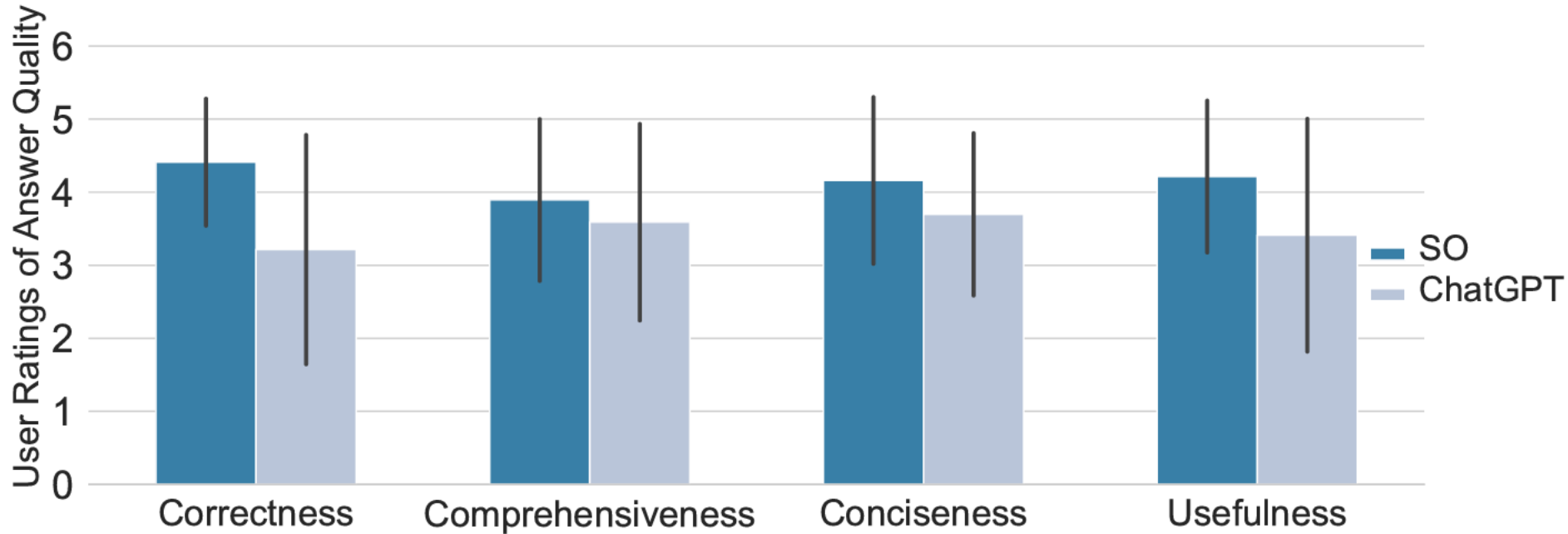
## **Who Answers It Better? An In-Depth Analysis of ChatGPT and Stack Overflow Answers to Software Engineering Questions**

- **517 ChatGPT's replies** to Stack Overflow (SO) questions
  - Correctness
  - Consistency
  - Comprehensiveness
  - Conciseness
  - Extensive linguistic analysis
  - user study to gain insights into the linguistic and human aspects of ChatGPT's answers.

## **Who Answers It Better? An In-Depth Analysis of ChatGPT and Stack Overflow Answers to Software Engineering Questions**

- Lengthy, comprehensive, covering all aspects of questions**
- No syntax errors for coding answers**
- More formal answers, showing more effort**
- More positive sentiment**

# Who Answers It Better? An In-Depth Analysis of ChatGPT and Stack Overflow Answers to Software Engineering Questions



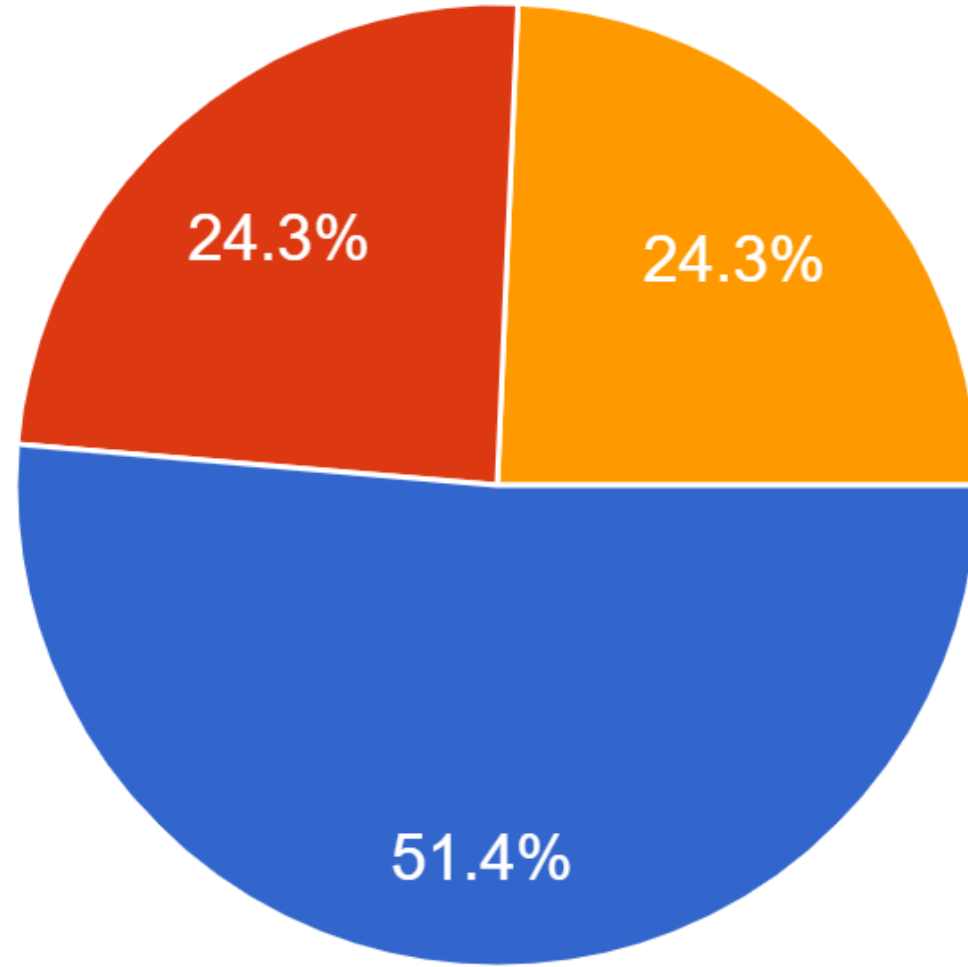
**Figure 1: Quality of answers as rated by participants. Difference in *Correctness*, *Conciseness*, and *Usefulness* are statistically significant (Paired t-Test, p-value < 0.05)**

## **Who Answers It Better? An In-Depth Analysis of ChatGPT and Stack Overflow Answers to Software Engineering Questions**

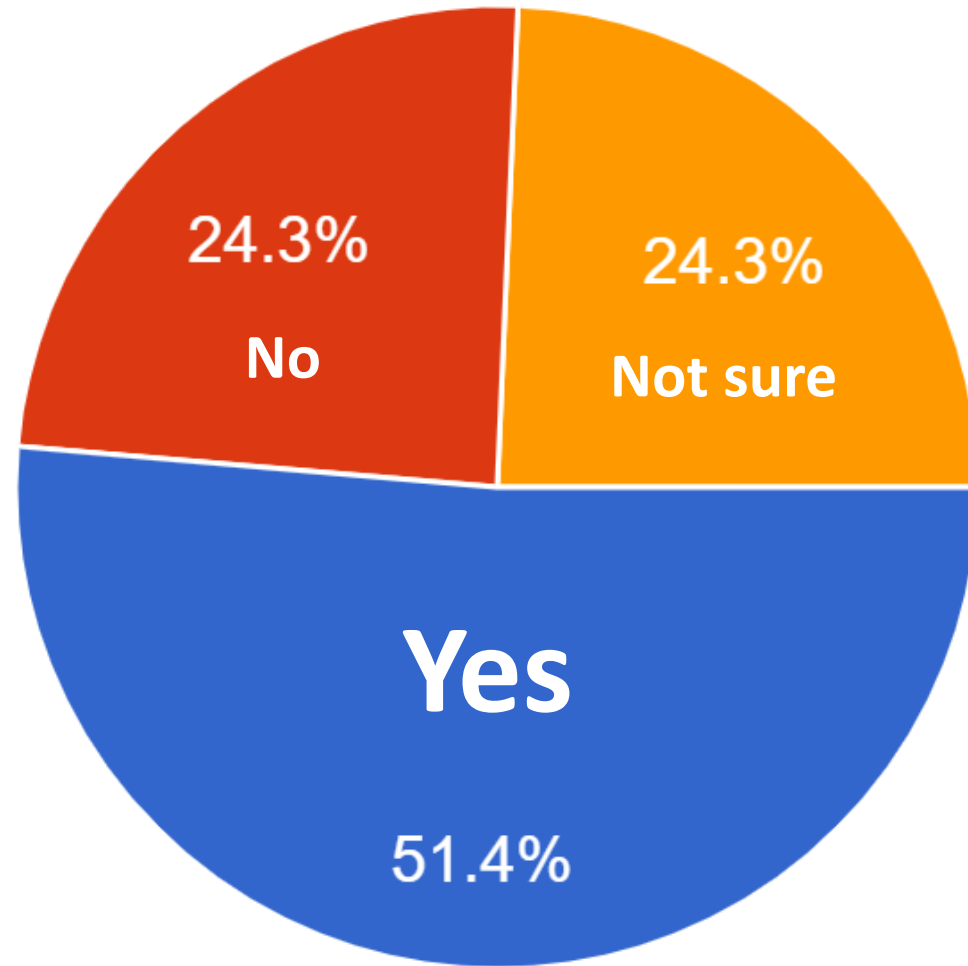
- Only, 50% of ChatGPT's answers were not consistent with human answers**
- Lack of understanding the context**
- Non-existing API, Library of functions**
- Fooled ~40% of users (they accepted incorrect answers)**

What about ethics?

Do you have ethical concerns about using AI tools



Do you have ethical concerns about using AI tools



# Ethical concerns

- How well are the open-source licenses (e.g., GPL3) respected by these tools?
- Who has the IP over generated ideas? Who is the author of the paper? Is it cheating (for students)? Will AI kill creativity or will it open new ways to express?
- use in teaching
- Not being able to distinguish whether the work is done by a person or plagiarized by AI. This accounts for student's work, scientific papers and even doctors diagnosis.
- Can be useful tools, but need supervision of knowledgeable human
- When developing experimental protocols for trials using different animal models (farm animals), I wonder if AI can properly identify specific limit points related to health and/or welfare, that would determine the exclusion of an individual from the trial ?
- If the result is used without critical examination
- Honesty of AI tools
- Over-use and lack of acknowledging (non-)originality of ideas

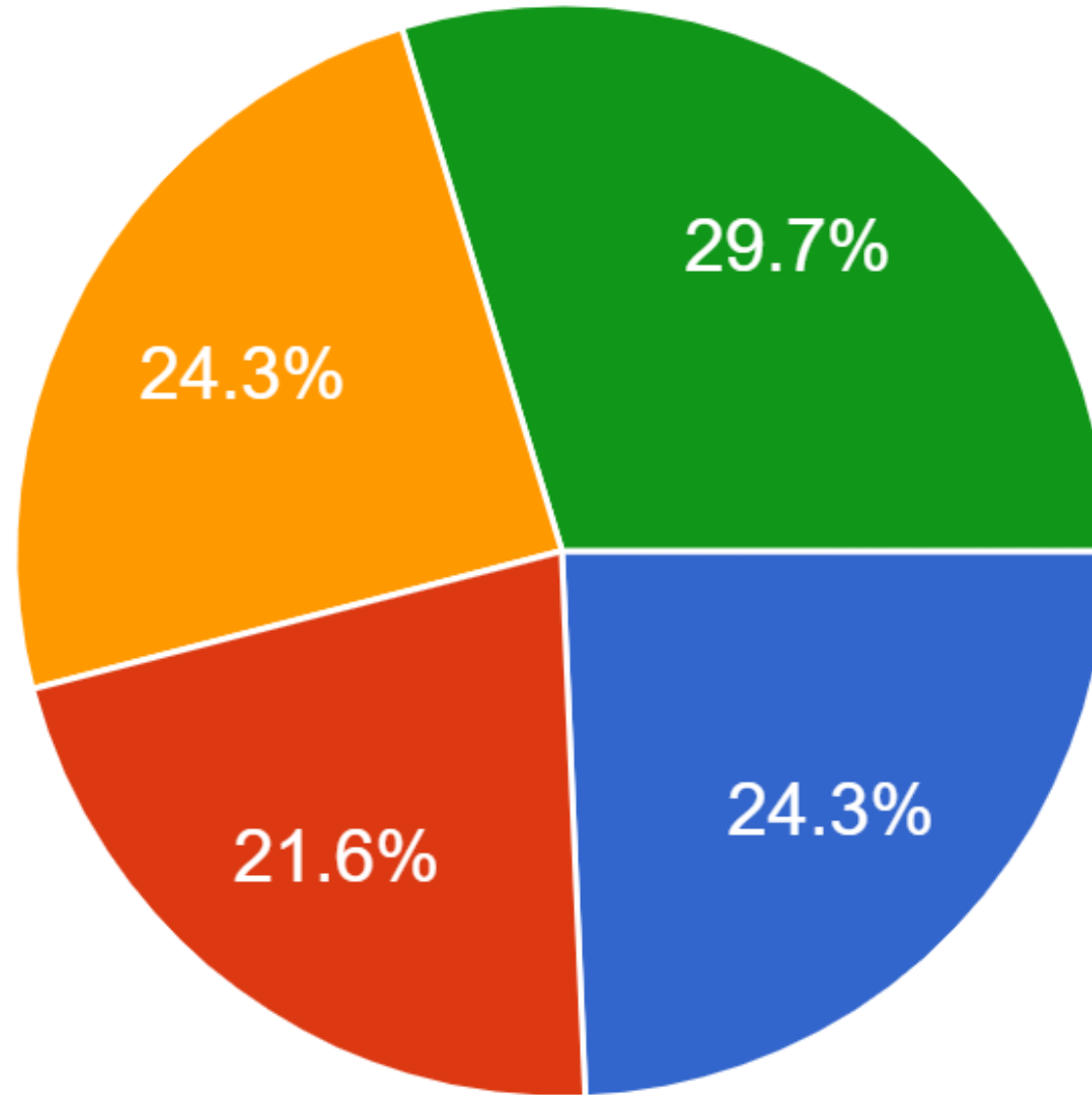
# Ethical concerns

- Respecting copyrights? Who will be responsible of results interpretation ?; can the authors still be called as such ?; will be restricted to the way of thinking and the box defined by the creators
- plagiarism, fake information, learning how to copy instead of how to create. Technology itself is neutral, the way it is used makes the difference. The limitations of the tool and in how many ways it can go something wrong need to be clear before official public release
- plagiarism, inaccuracy, "creating data and references"
- If you use it blindly and not as a tool to enhance your work
- The carbon footprint of AI is huge and is greatly overlooked.
- Original thinking
- Students using it to answer questions; people not actually understanding what they are doing and why
- paper mill, papers without data behind them
- Transparency - how can AI and human generated texts be distinguished? Control - can AI be controlled by humans?

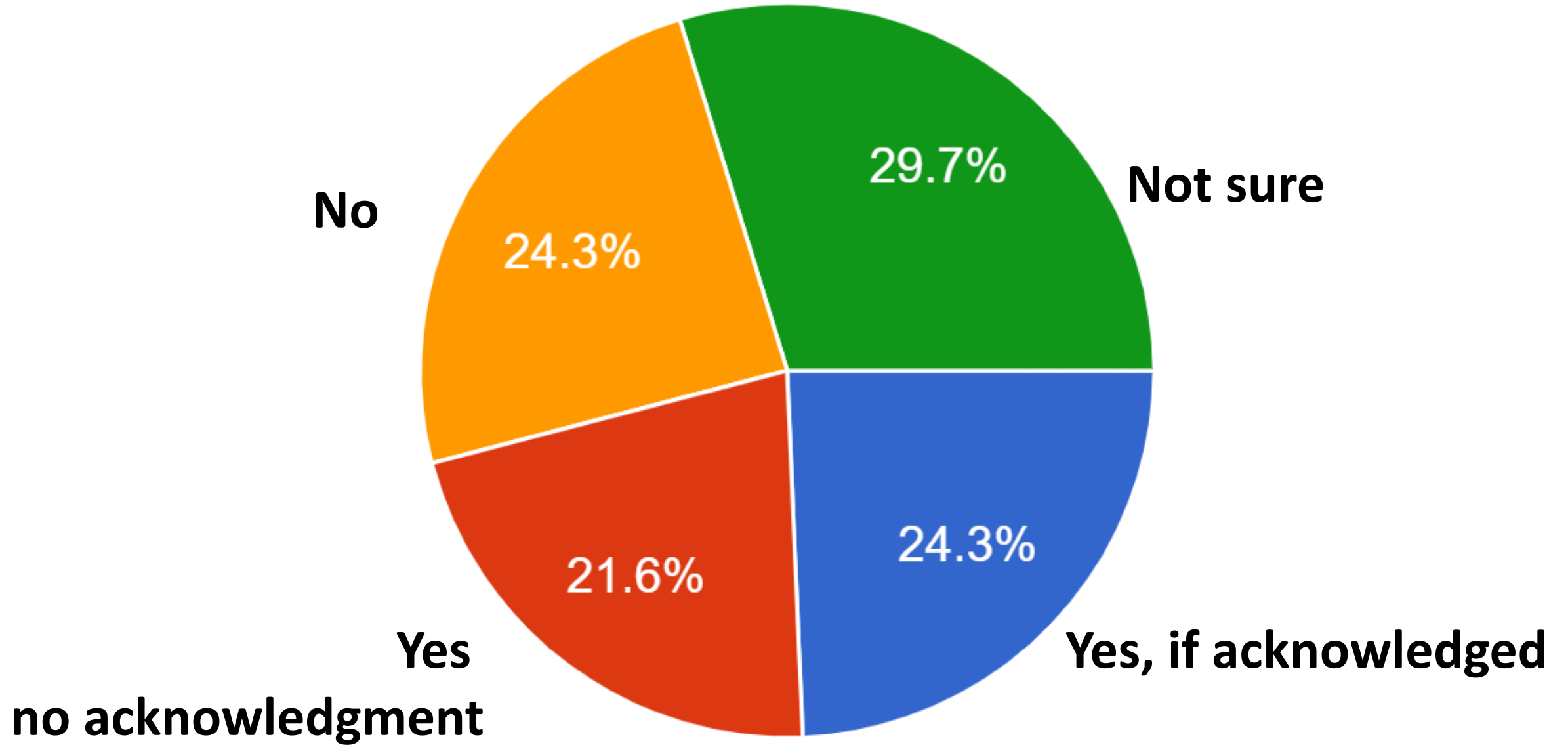
# Ethical concerns

- Ethically risky, not at all comfortable with the idea that machines replace humans at this point; we need to put limits to this crazy dynamics!
- AI tools could be VERY beneficial, but if people solely rely on them, or accept the answers without a second thought, these answers could be faulty. Also the authorship of the text, if large proportions are generated by AI is questionable. On the other hand, there are parts of texts, e.g. in research grants, that do not require critical thinking, just descriptions of the environment, or available facilities. Here, although the text is generated by AI, I would not see it as ethical concern, as it frees up time that would be used to construct this obligatory, but frankly many times not so useful general text.

Do you think that using AI-tools should be allowed in the process of writing scientific papers

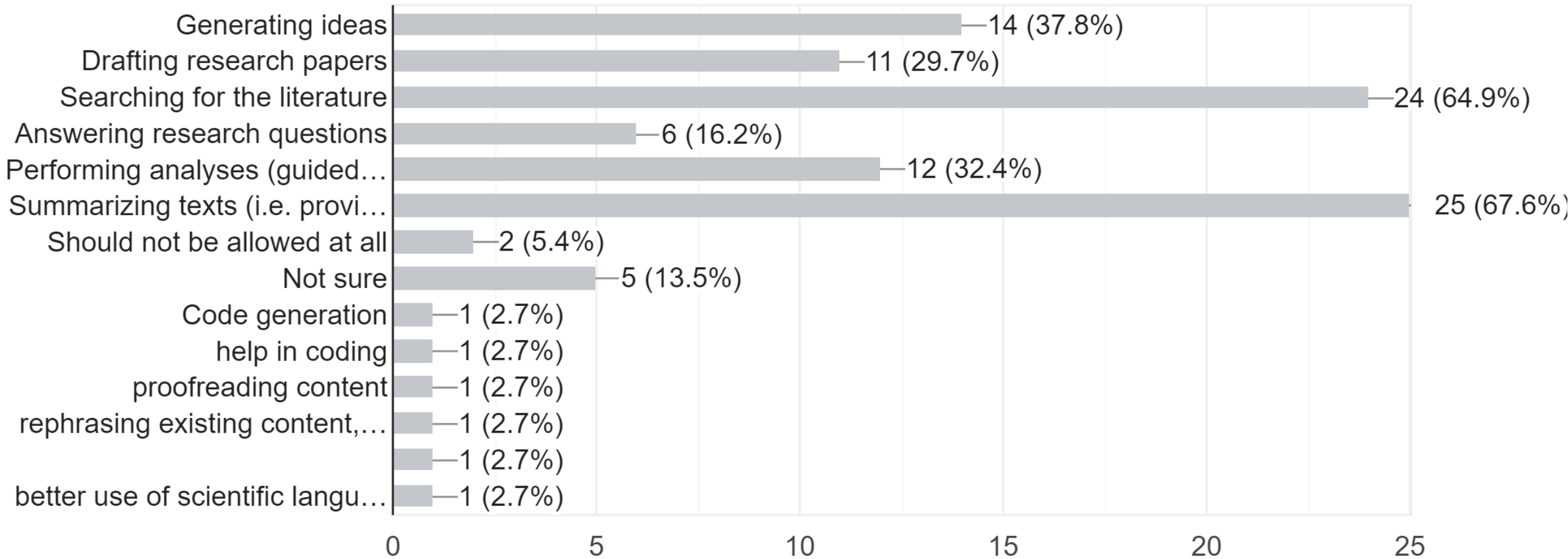


Do you think that using AI-tools should be allowed in the process of writing scientific papers



# In which aspect using AI-driven language models should be allowed in academia (research/teaching)? (multiple answers possible)

37 responses



**If you think that it should be allowed to write a scientific paper with help of AI-tools, to what extend, this help would be acceptable to you?**

- **Draft the "Abstract", "materials and methods", "Results", generate the "References". The "Discussion" and "Conclusions" should be properly done by the authors.**
- **Writing a paper? Not yet there... unsupervised AI is a mess!**
- **Helping yes but only for a draft, not to analyse data and interpret the data. What would be the role of the researcher then?**
- **Ideas must come from researcher, they also have responsibility ultimately.**
- **To the extent where results and conclusions were not affected by AI at all.**
- **Full if it allows to do better**



We can use AI in many ways  
to help us in scientific work

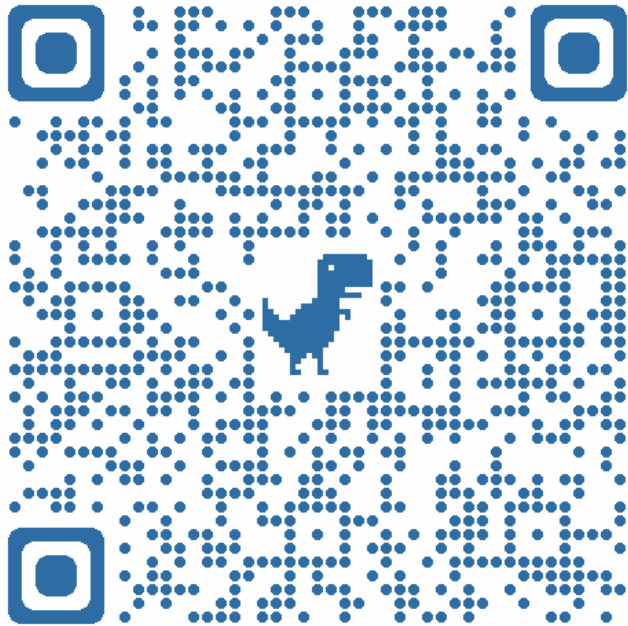
Should we

~~We can~~ use AI in many ways  
to help us in scientific work ???

Should we  
~~We can~~ use AI in many ways  
to help us in scientific work ???

Certainly not blindly!!!

# Thank you!



<https://forms.gle/2KeHZFX2AJQojByJA>

## Participate in the survey

[marcin.pszczola@puls.edu.pl](mailto:marcin.pszczola@puls.edu.pl)



EAAP Lyon 2023 presentation for Young EAAP scientists about chatGTP, can you help with outline of the presentation? The presentation should be funny, regarding using chatgtp in scientific career - writting, data analyses, coding. It should touch moral aspects as well.



EAAP Lyon 2023 presentation for Young EAAP scientists about chatGTP, can you help with outline of the presentation? The presentation should be funny, regarding using chatgtp in scientific career - writting, data analyses, coding. It should touch moral aspects as well.



Sure, I can help you with an outline for your presentation about using ChatGPT in scientific career. Here's a possible outline:

### I. Introduction

- A. Greet the audience and introduce yourself and your topic
- B. Explain what ChatGPT is and how it works
- C. Preview what you'll be talking about in your presentation

### II. ChatGPT for Writing

- A. Discuss how ChatGPT can help with writing scientific papers
- B. Give examples of how ChatGPT can help with writing abstracts, introductions, and conclusions
- C. Talk about the potential downsides of relying too heavily on ChatGPT for writing

### III. ChatGPT for Data Analysis

- A. Explain how ChatGPT can assist with data analysis in scientific research
- B. Provide examples of how ChatGPT can help with data visualization and interpretation
- C. Discuss the importance of human interpretation in data analysis and the limitations of ChatGPT



EAAP Lyon 2023 presentation for Young EAAP scientists about chatGTP, can you help with outline of the presentation? The presentation should be funny, regarding using chatgtp in scientific career - writting, data analyses, coding. It should touch moral aspects as well.



Sure, I can help you with an outline for your presentation about using ChatGPT in scientific career. Here's a possible outline:

#### I. Introduction

- A. Greet the audience and introduce yourself and your topic
- B. Explain what ChatGPT is and how it works
- C. Preview what you'll be talking about in your presentation

#### II. ChatGPT for Writing

- A. Discuss how ChatGPT can help with writing scientific papers
- B. Give examples of how ChatGPT can help with writing abstracts, introductions, and conclusions
- C. Talk about the potential downsides of relying too heavily on ChatGPT for writing

#### III. ChatGPT for Data Analysis

- A. Explain how ChatGPT can assist with data analysis in scientific research
- B. Provide examples of how ChatGPT can help with data visualization and interpretation
- C. Discuss the importance of human interpretation in data analysis and the limitations of ChatGPT

#### IV. ChatGPT for Coding

- A. Explore how ChatGPT can help with coding in scientific research
- B. Give examples of how ChatGPT can assist with tasks such as debugging and generating code snippets
- C. Discuss the importance of developing coding skills and the limitations of relying on ChatGPT for coding

#### V. Moral Aspects of Using ChatGPT in Scientific Research

- A. Talk about the ethical considerations of using ChatGPT in scientific research
- B. Discuss issues such as bias and transparency in using machine learning models like ChatGPT
- C. Encourage the audience to consider the ethical implications of their own use of ChatGPT

#### VI. Conclusion

- A. Summarize the key points of your presentation
- B. Encourage the audience to continue exploring the possibilities and limitations of ChatGPT in their scientific careers
- C. End with a funny or inspiring anecdote about your own experience using ChatGPT.