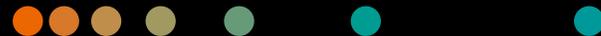


Dr. Ryan Butterfield, DrPH, MBA  
Siemens Healthineers,  
Senior Director, Biostatistics

# The Future of Statistical Analysis – to AI or not AI?

**Thanks to Siemens Healthineers  
leadership, colleagues, and staff for  
their support of this work.**

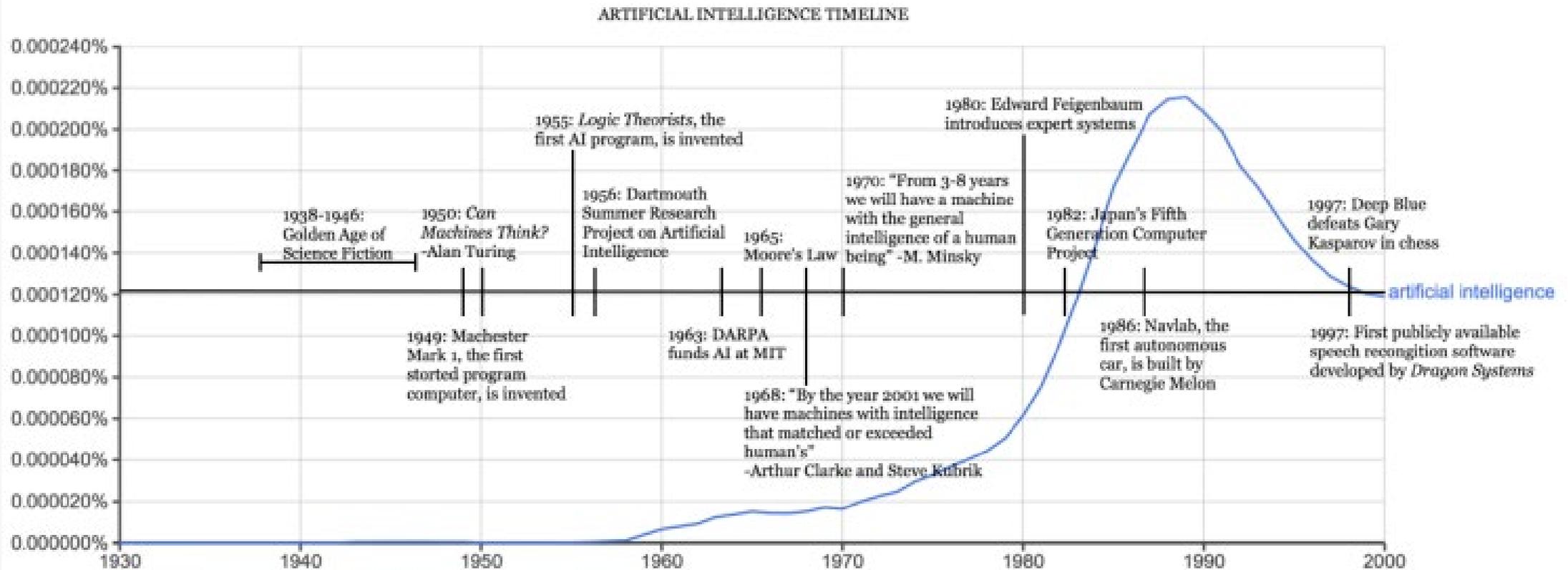
2023  
Siemens Healthineers



# Content of Presentation

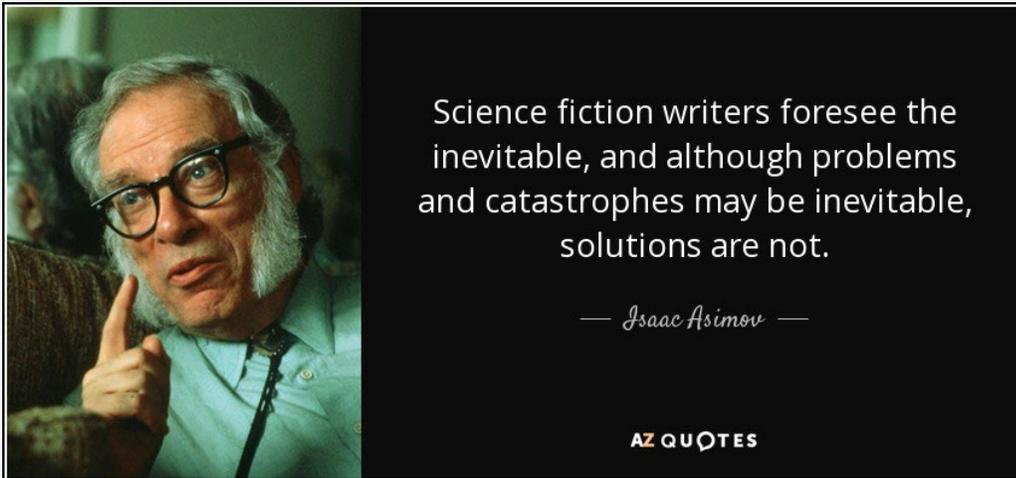
- Introduction/ Historical Reflection
- Summary of Current Developments
- Recent Literature on Utility of AI in the workplace
- Current Use of AI in Statistical Practice
  - Prompting Matters
  - Case Study - Uses in Biostat and DM at Siemens Healthineers
- Speculation on the Future Use of AI in Statistical Practice
- Conclusion
- Acknowledgments

# Introduction and Historical Reflection



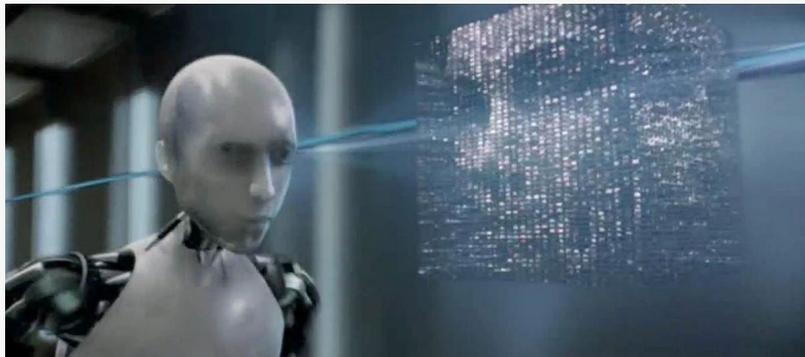
# Science Fiction and AI

- An interesting perspective as science has often reflected the imaginations of science fiction



# Science Fiction and AI

- AI Shows up throughout science fiction, Star Trek, Star Wars, Asimov's Universe, Douglas Adams, etc. just to name a few.



# Introduction and (RECENT) Historical Reflection

Date	Development
2020	In February 2020, Microsoft introduces its Turing Natural Language Generation (T-NLG), which is the "largest language model ever published at 17 billion parameters." <sup>[127]</sup>
	In November 2020, <a href="#">AlphaFold 2</a> by DeepMind, a model that performs <a href="#">predictions of protein structure</a> , wins the <a href="#">CASP</a> competition. <sup>[128]</sup>
2022	<a href="#">OpenAI</a> introduces <a href="#">GPT-3</a> , a state-of-the-art autoregressive language model that uses <a href="#">deep learning</a> to produce a variety of computer codes, poetry and other language tasks exceptionally similar, and almost indistinguishable from those written by humans. Its capacity was ten times greater than that of the T-NLG. It was introduced in May 2020, <sup>[129]</sup> and was in beta testing in June 2020.
	<a href="#">ChatGPT</a> , an AI <a href="#">chatbot</a> developed by <a href="#">OpenAI</a> , debuts in November 2022. It is initially built on top of the <a href="#">GPT-3.5 large language model</a> . While it gains considerable praise for the breadth of its knowledge base, deductive abilities, and the human-like fluidity of its natural language responses, <sup>[130][131]</sup> it also garners criticism for, among other things, its tendency to " <a href="#">hallucinate</a> ." <sup>[132][133]</sup> a phenomenon in which an AI responds with factually incorrect answers with high confidence. The release triggers widespread public discussion on artificial intelligence and its potential impact on society. <sup>[134][135]</sup>
	A November 2022 class action lawsuit against <a href="#">Microsoft</a> , <a href="#">GitHub</a> and <a href="#">OpenAI</a> alleges that <a href="#">GitHub Copilot</a> , an AI-powered code editing tool trained on public GitHub repositories, violates the copyrights of the repositories' authors, noting that the tool is able to generate source code which matches its training data verbatim, without providing attribution. <sup>[136]</sup>
	By January 2023, <a href="#">ChatGPT</a> has more than 100 million users, making it the fastest growing consumer application to date. <sup>[137]</sup>



2023	On March 7, 2023, <a href="#">Nature Biomedical Engineering</a> writes that "it is no longer possible to accurately distinguish" human-written text from text created by large language models, and that "It is all but certain that general-purpose large language models will rapidly proliferate... It is a rather safe bet that they will change many industries over time." <sup>[145]</sup>
	In response to ChatGPT, <a href="#">Google</a> releases in a limited capacity its chatbot <a href="#">Google Bard</a> , based on the <a href="#">LaMDA</a> and <a href="#">PaLM</a> large language models, in March 2023. <sup>[146][147]</sup>
	On March 29, 2023, a petition of over 1,000 signatures is signed by Elon Musk, Steve Wozniak and other tech leaders, calling for a 6-month halt to what the petition refers to as an "an out-of-control race" producing AI systems that its creators can't "understand, predict, or reliably control." <sup>[148][149]</sup>
	In May 2023, Google makes an announcement regarding Bard's transition from LaMDA to PaLM2, a significantly more advanced language model. <sup>[150]</sup>

# Introduction and (RECENT) Historical Reflection

In the last week of May 2023, a [Statement on AI Risk](#) is signed by [Geoffrey Hinton](#), [Sam Altman](#), [Bill Gates](#), and many other prominent AI researchers and tech leaders with the following succinct message: "Mitigating the risk of extinction from AI should be a global priority alongside other societal-scale risks such as pandemics and nuclear war."<sup>[151][152]</sup>

On July 9, 2023, [Sarah Silverman](#) files a class action lawsuit against Meta and OpenAI for copyright infringement for training their large language models on millions of authors' copyrighted works without permission.<sup>[153]</sup>

In August, 2023, the New York Times, CNN, Reuters, the Chicago Tribune, Australian Broadcasting Corporation (ABC) and other news companies block OpenAI's GPTBot web crawler from accessing their content, while the New York Times also updates its terms of service to disallow the use of its content in large language models.<sup>[154]</sup>

On September 13, 2023, in a serious response to growing anxiety about the dangers of AI, the US Senate holds the inaugural bipartisan "[AI Insight Forum](#)", bringing together senators, CEOs, civil rights leaders and other industry reps, to further familiarize senators with the nature of AI and its risks, and to discuss needed safeguards and legislation.<sup>[155]</sup> The event is organized by Senate Majority Leader [Chuck Schumer](#) (D-NY),<sup>[156]</sup> and is chaired by U.S. Senator [Martin Heinrich](#) (D-N.M.), Founder and co-Chair of the Senate AI Caucus.<sup>[157]</sup> Reflecting the importance of the meeting, the forum is attended by over 60 senators,<sup>[158]</sup> as well as [Elon Musk](#) (Tesla CEO), [Mark Zuckerberg](#) (Meta CEO), [Sam Altman](#) (OpenAI CEO), [Sundar Pichai](#) (Alphabet CEO), [Bill Gates](#) (Microsoft co-founder), [Satya Nadella](#) (Microsoft CEO), [Jensen Huang](#) (Nvidia CEO), [Arvind Krishna](#) (IBM CEO), [Alex Karp](#) (Palantir CEO), [Charles Rivkin](#) (chairman and CEO of the MPA), [Meredith Stiehm](#) (president of the Writers Guild of America West), [Liz Shuler](#) (AFL-CIO President), and [Maya Wiley](#) (CEO of the [Leadership Conference on Civil and Human Rights](#)), among others.<sup>[155][156][158]</sup>



# Introduction and (RECENT) Historical Reflection

- History of Statistical Models Behind AI
  - Primarily Neural Networks
- Current Statistical Models
  - Large Language Models
  - Generative Adversarial Model
- Future State
  - AI embedded in Quantum Computing System, allows the computing and memory power to be sustainable for an AI to exist long term

## How do companies use quantum computing?

Quantum computing can revolutionize industries. We give some example use cases below:

### ML

Machine learning (ML) is the process of analyzing vast quantities of data to help computers make better predictions and decisions. Research in quantum computing studies the physical limits of information processing and is breaking new ground in fundamental physics. This research leads to advances in many fields of science and industry, such as chemistry, optimization, and molecular simulation. It is also a growing area of interest for financial services to predict market movements and for manufacturing to improve operations.

### Optimization

Quantum computing can improve research and development, supply-chain optimization, and production. For example, you could apply quantum computing to decrease manufacturing process-related costs and shorten cycle times by optimizing elements such as path planning in complex processes. Another application is the quantum optimization of loan portfolios so that lenders can free up capital, lower interest rates, and improve their offerings.

### Simulation

The computational effort required to simulate systems accurately scales exponentially with the complexity of drug molecules and materials. Even using approximation methods, current supercomputers cannot achieve the level of accuracy that these simulations demand. Quantum computation has the potential to solve some of the most challenging computational problems faced in chemistry, allowing the scientific community to do chemical simulations that are intractable today. For example, [Pasqal](#) built their QUBEC computational software to run chemistry simulations. QUBEC automates the heavy lifting necessary to run quantum computational tasks from automatic provisioning of the computing infrastructure to running pre- and post-processing classical calculations and performing error mitigation tasks.

# Summary of Current Developments

- Current State of Artificial Intelligence is really EI... Electronic Intelligence
  - Status of Sentience
  - EI may or may not evolve to become AI
  - EI have a constrained programming with heuristics to control the scope of capability
- The most advanced current systems are:
  - ChatGPT4
  - LaMDA, aka Bard



# Summary of Current Developments

- The First Wave of AI was about Classification (Past State)
  - The Second Wave of AI is Generative (Current State)
    - The Third Wave of AI will be Interactive (Future State)

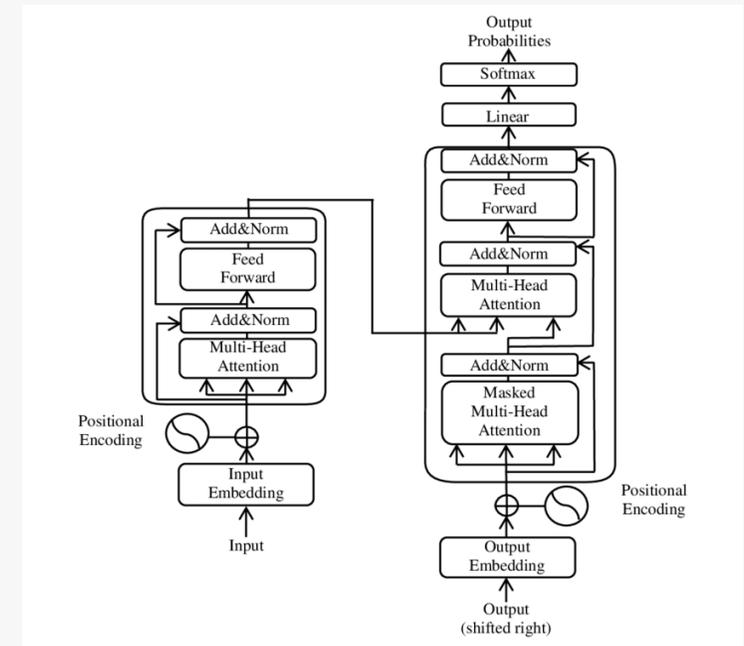


- Here are some specific examples of different kinds of AI in use today:
  - Reactive AI: Video game enemies, thermostats, and traffic lights
  - Limited memory AI: Self-driving cars, spam filters, and product recommendation systems
  - Theory of mind AI: (Not yet possible)
  - Self-aware AI: (Not yet possible)
  - Machine learning: Facial recognition, spam filtering, and product recommendation systems
  - Deep learning: Self-driving cars, medical diagnosis, and machine translation
- It is likely that we will see even more innovative and powerful AI systems in the future, i.e. AI creating AI (intelligence explosion).

# Summary of Current Developments

## • Composition

- Large Language Models i.e. ChatGPT, Bard (Language Model for Dialogue Applications, LaMDA)
- Reporting
  - Useful in producing written elements through mosaic approach of combining existing written prose
- Computer programming
- Report Drafting
- Bard is built off neural network-based transformer models



12

Yueling Jia, CC BY-SA 3.0  
 <<https://creativecommons.org/licenses/by-sa/3.0/>>, via Wikimedia Commons

# Summary of Current Developments

## • Computation

- Wolfram Alpha + ChatGPT
- Computationally intensive processes combined with large language model attributes

☰ Contents

See also:

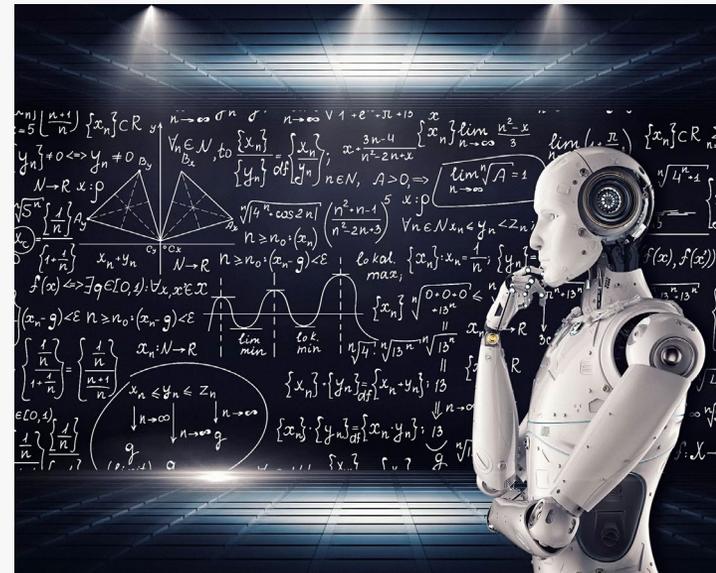
📄 ["What Is ChatGPT Doing ... and Why Does It Work?"](#) »

## ChatGPT Gets Its “Wolfram Superpowers”!

March 23, 2023

*This is part of an ongoing series about our LLM-related technology:*

- [ChatGPT Gets Its “Wolfram Superpowers”!](#)
- [Instant Plugins for ChatGPT: Introducing the Wolfram ChatGPT Plugin Kit](#)
- [The New World of LLM Functions: Integrating LLM Technology into the Wolfram Language](#)
- [Prompts for Work & Play: Launching the Wolfram Prompt Repository](#)
- [Introducing Chat Notebooks: Integrating LLMs into the Notebook Paradigm](#)



# Summary of Current Developments

- **Hybrid - Collective Superintelligence**
  - **Swarm Intelligence Approaches - Human Groups with AI intertwined**

## 'Swarm Intelligence' Predicts the Kentucky Derby's First 4 Finishers, Wins \$11000

A software platform that uses a human swarm has just beaten the 540-to-1 odds of the Kentucky Derby.

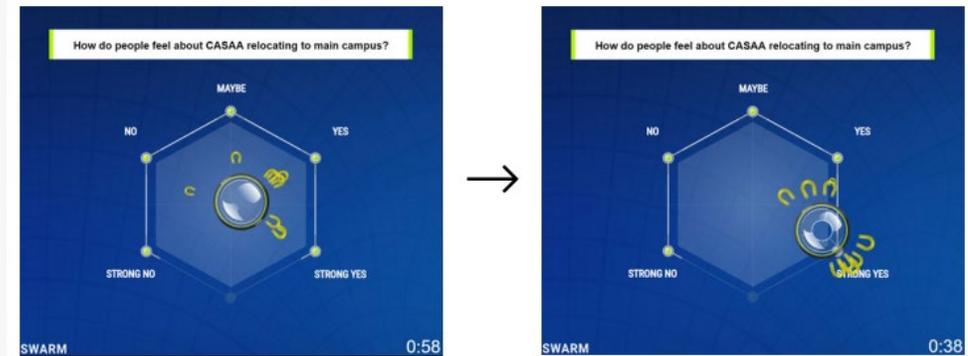
/AI&Intelligence / AI&Intelligence / HorseRacing / HumanSwarm



## Keeping up With A.I: Why Humanity Must Cultivate a 'Hive Mind'

Pooling our intellectual resources is how humanity will keep ahead of purely artificial intelligences.

/Op-Ed / AlphaGo / ArtificialIntelligence / HiveMind



**Figure 1:** Swarm deliberation at (left) 2 seconds into the swarm and (right) 22 seconds into the swarm, showing a transition from uncertainty and weak positivity in the initial stages to a consensus of strong positivity in the final stages. View the replay [here](#).

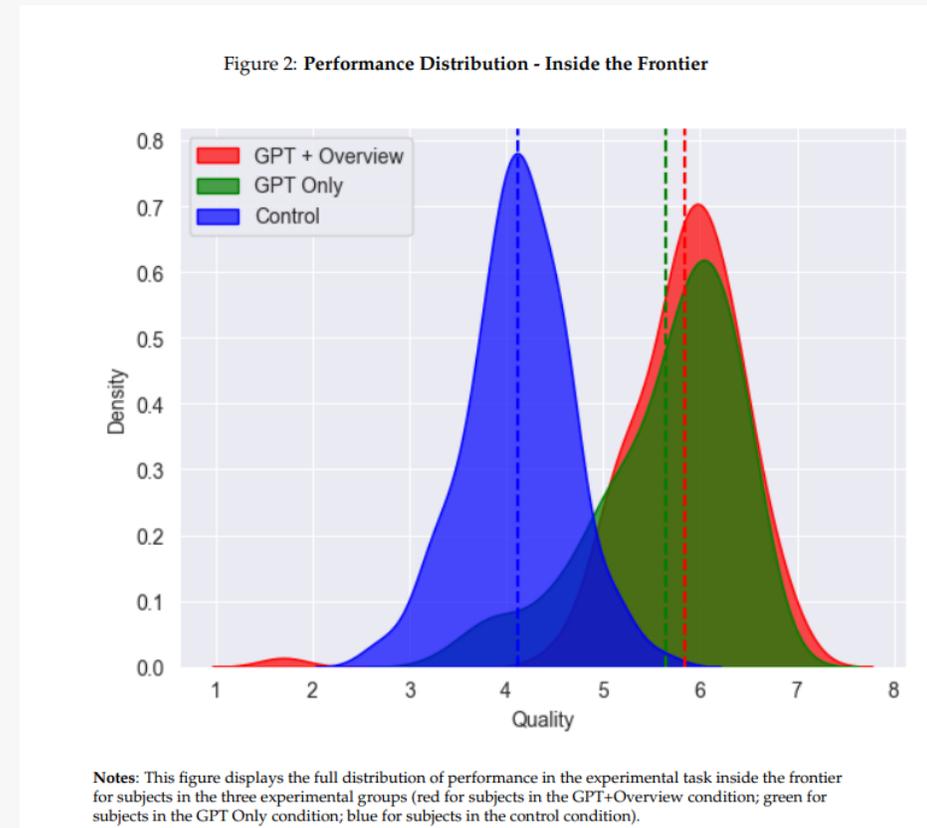
### Maximizing Zoom Meeting Effectiveness

Since the first experiment in using Swarm AI, Dr. Pearson's faculty group now regularly uses Swarm to streamline their meetings. They've found that their meetings are more engaging and effective: They're able to accomplish more in each meeting by identifying the topics they agree on and those they need to discuss more to reach an acceptable level of consensus.

# Recent Literature on Utility of AI in the Workplace

- Navigating the Jagged Technological Frontier: Field Experimental Evidence of the Effects of AI on Knowledge Worker Productivity and Quality. Authors: Dell'Acqua, F., McFowland, E., ..., Lakhani, K. Working Paper 24-013
- N=758, randomized to three arms (no access, AI with access, AI with access plus training)
- Tasked to complete 18 realistic tasks within the AI capabilities:
  - Those two groups with AI completed:
    - 12.2% more tasks
    - 25.1 % more quickly when comparing completed tasks rates
    - With 40% higher quality
- Those with AI, who were baseline at lower skill rankings showed an increase of 43% than those where were higher ranked (17%).

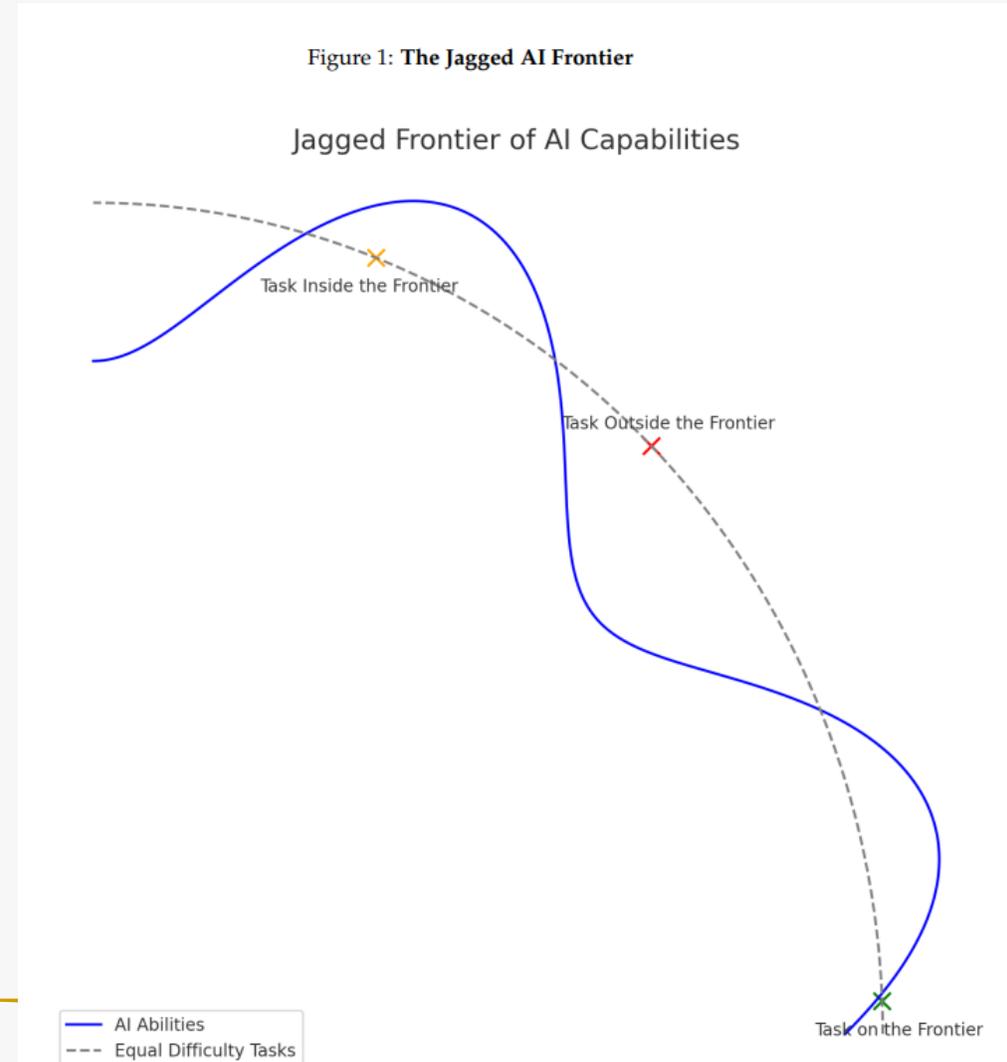
[https://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=4573321](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4573321)



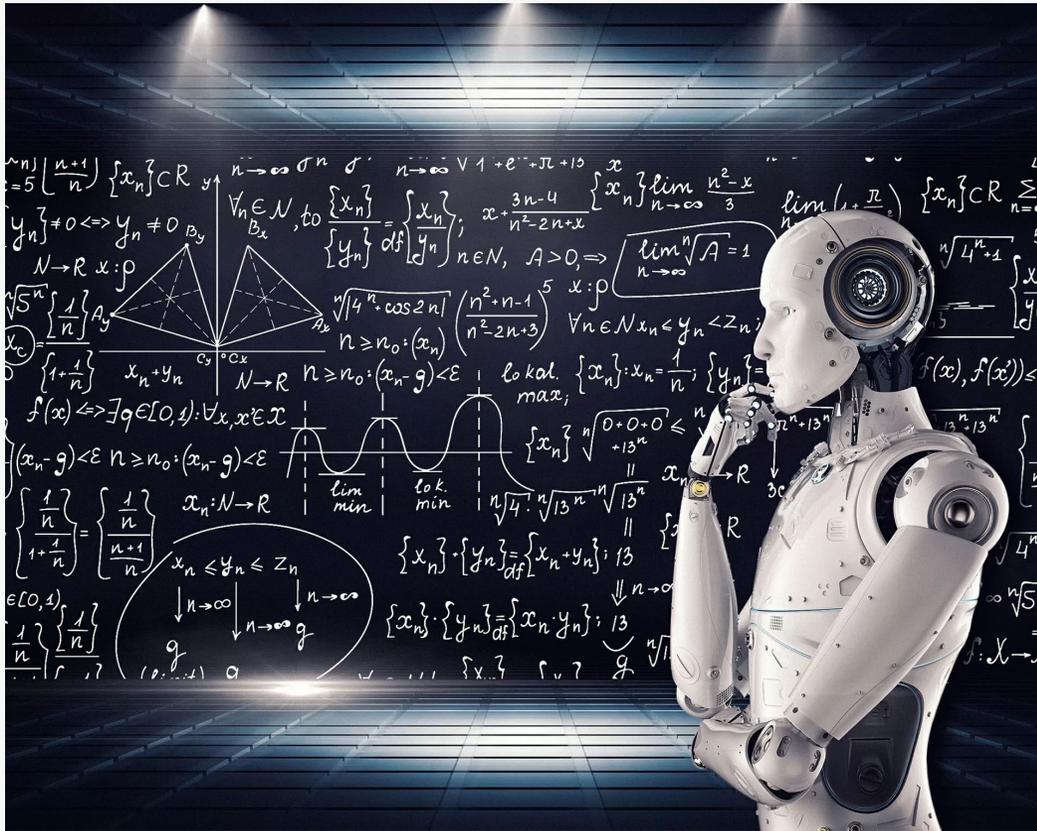
# Recent Literature on Utility of AI in the Workplace

- For tasks outside of AI capabilities, those with AI were 19% lower than those without
- Two approaches came out of this: the **"Centaur"** who were half horse/half human which reflects the approach of **dividing and delegating their solution-creation activities to the AI and Themselves** or
- **Cyborgs** - where there was a **complete integration of task flow between the AI and individual, with a continuous interaction with the technology.**

[https://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=4573321](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4573321)



# Current Use of AI in Statistical Practice at Siemens Healthineers



- Beta Launch of ChatGPT 4 within Siemens Healthineers
- Biostatistics and Data Management were assigned as Beta Testers
  - Goals of Project:
    - Demonstrate and develop use cases
    - Show benefit to worker skill and process improvement
    - Develop protocols for use of AI within a regulated clinical trial environment
    - Management and leadership uses in overseeing the execution of trials and business activities

# Current Use of AI in Statistical Practice at Siemens Healthineers— A ‘Prompt’ Experiment (Bard vs ChatGPT)

What are the key skills needed for a biostatistician to use with AI

A biostatistician who wants to work with AI would need a combination of skills in statistics, computer science, and domain knowledge in biology or medicine. Here are some key skills:

- 1. Statistical Analysis:** A strong understanding of statistical methods and principles is crucial. This includes knowledge of probability, hypothesis testing, regression models, and Bayesian statistics.
- 2. Machine Learning:** Familiarity with machine learning algorithms and principles, including both supervised and unsupervised learning methods. This includes understanding how to train, validate, and test models, as well as how to prevent overfitting.
- 3. Programming:** Proficiency in programming languages commonly used in data science and AI, such as Python or R. This includes understanding how to use libraries and packages for data analysis, machine learning, and data visualization.
- 4. Data Management:** Skills in handling, cleaning, and processing large datasets. This includes knowledge of databases and query languages like SQL, as well as experience with data preprocessing techniques.
- 5. Domain Knowledge:** Understanding of biology, medicine, or whatever field the biostatistician is working in. This is important for understanding the data, formulating relevant questions, and interpreting the results of analyses.
- 6. Ethics:** Understanding of ethical considerations in AI, such as issues of privacy, bias, and fairness. This is particularly important in healthcare, where decisions can have significant impacts on people's lives.
- 7. Communication:** Ability to communicate complex results to non-technical stakeholders. This includes visualizing data and results in a clear and understandable way.
- 8. Problem-Solving:** Ability to formulate and solve complex problems, often with incomplete or messy data.
- 9. Continuous Learning:** AI is a rapidly evolving field, so a commitment to continuous learning and staying up-to-date with the latest research and techniques is important.

What are the key skills needed for a biostatistician to use with AI

View other drafts

Key skills needed for a biostatistician to use with AI include:

- Knowledge of AI and machine learning:** Biostatisticians need to have a basic understanding of AI and machine learning concepts, such as supervised learning, unsupervised learning, and deep learning. This will allow them to identify and select the appropriate AI techniques for their work.
- Programming skills:** Biostatisticians need to be able to program in order to implement AI techniques and analyze data. Popular programming languages used in biostatistics and AI include Python, R, and SAS.
- Data science skills:** Biostatisticians need to have strong data science skills, such as data cleaning, data preprocessing, and data visualization. This will allow them to prepare their data for AI analysis and interpret the results of their AI models.
- Statistical knowledge:** Biostatisticians need to have a strong foundation in statistical theory and methods. This will allow them to evaluate the performance of their AI models and ensure that their results are valid and reliable.
- Domain knowledge:** Biostatisticians need to have a deep understanding of the biological or medical domain in which they are working. This will allow them to ask the right questions and interpret the results of their AI models in a meaningful way.

In addition to the above skills, biostatisticians who use AI also need to have strong communication and collaboration skills. They need to be able to communicate their findings to both technical and non-technical audiences, and they need to be able to collaborate effectively with other scientists and engineers.

Here are some specific examples of how biostatisticians can use AI:

- Develop new statistical methods:** Biostatisticians can use AI to develop new statistical methods for analyzing complex data. For example, AI can be used to develop new methods for identifying patterns in data, detecting outliers, and predicting outcomes.
- Improve the efficiency and accuracy of data analysis:** AI can be used to automate and streamline the data analysis process. This can free up biostatisticians to focus on more complex and strategic tasks.
- Develop new insights into biological and medical systems:** AI can be used to analyze large and complex datasets in ways that would not be possible using traditional statistical methods. This can lead to new insights into the causes and mechanisms of diseases, as well as the development of new diagnostic and therapeutic strategies.

# Current Use of AI in Statistical Practice at Siemens Healthineers— A ‘Prompt’ Experiment (Bard vs ChatGPT)

graph a linear regression line example with 100 data points where y is bmi and x is height and weight

As an AI text-based model, I'm unable to create visual content such as graphs. However, I can guide you on how to do it using a programming language like Python with libraries such as pandas, numpy, and matplotlib.

Here's a simple example of how you might do it:

```
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
from sklearn.linear_model import LinearRegression

# Assuming you have a DataFrame 'df' with columns 'height', 'weight', and 'bmi'
df = pd.DataFrame(np.random.randint(0,100,size=(100, 3)), columns=)

# Create a new feature combining 'height' and 'weight'
df = df + df

# Create a Linear Regression model
model = LinearRegression()

# Reshape your data since sklearn requires you to do it for the features
X = df.values.reshape(-1,1)
y = df.values.reshape(-1,1)
```

graph a linear regression line example with 100 data points where y is bmi and x is height and weight

View other drafts

To graph a linear regression line example with 100 data points where y is BMI and x is height and weight, we can use the following steps:

1. Generate 100 random data points for height and weight.
2. Calculate the BMI for each data point.
3. Create a scatter plot of the height and weight data, with BMI on the y-axis.
4. Perform a linear regression on the data.
5. Plot the linear regression line on the scatter plot.

Here is an example of a Python code to generate the data and plot the graph:

```
Python

import numpy as np
import matplotlib.pyplot as plt

# Generate 100 random data points for height and weight
height = np.random.randint(150, 200, 100)
weight = np.random.randint(50, 100, 100)

# Calculate the BMI for each data point
bmi = weight / (height / 100) ** 2

# Create a scatter plot of the height and weight data
plt.scatter(height, bmi)
```

# Current Use of AI in Statistical Practice at Siemens Healthineers— A ‘Prompt’ Experiment (Bard vs ChatGPT)

```
# Fit your model
model.fit(X, y)

# Get your coefficients
intercept = model.intercept_
coef = model.coef_

# Create a line using the coefficients
line = coef * X + intercept

# Plot the points
plt.scatter(X, y, color='blue')

# Plot the line
plt.plot(X, line, color='red')

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

Please note that this is a very basic example and might not be suitable for your specific use case. The 'height\_weight' feature is simply the sum of 'height' and 'weight', which might not be the best way to combine these features. You might want to consider other ways of combining them, or even using a different kind of model that can handle multiple independent variables.

Citations: 1. 'height', 'weight', 'bmi' 2. 'height\_weight' 3. 'height' 4. 'weight' 5. 'bmi'

```
# Create a scatter plot of the height and weight data
plt.scatter(height, bmi)

# Perform a linear regression on the data
model = np.polyfit(height, bmi, 1)

# Plot the linear regression line on the scatter plot
plt.plot(height, model[0] * height + model[1], color='red')

# Set the axis labels and title
plt.xlabel('Height (cm)')
plt.ylabel('BMI')
plt.title('Linear Regression of BMI vs Height')

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

Use code with caution. [Learn more](#)

The resulting graph will show a scatter plot of the height and weight data points, with the linear regression line plotted in red. The slope of the line represents the change in BMI for every unit change in height.

It is important to note that this is just an example graph, and the actual linear regression line will vary depending on the specific data set used.

# Current Use of AI in Statistical Practice at Siemens Healthineers— A ‘Prompt’ Experiment (Bard vs ChatGPT)



Show an example of a longitudinal hierarchical random effects model using R and SAS

## Example of a longitudinal hierarchical random effects model using R

### Code snippet

```
# Load the necessary libraries
library(lme4)

# Create a data frame of the longitudinal data
data <- data.frame(
  id = c(1, 1, 1, 2, 2, 2),
  time = c(1, 2, 3, 1, 2, 3),
  y = c(10, 12, 14, 4, 14, 13)
)

# Fit a longitudinal hierarchical random effects model
model <- lmer(y ~ time + (1 | id), data)

# Print the summary of the model
summary(model)
```

Use code with caution. [Learn more](#)

This model will estimate the fixed effects of time and the random effect of individual (id). The random effect of individual accounts for the fact that the individuals have different intercepts and slopes.

## Example of a longitudinal hierarchical random effects model using SAS

### SAS

```
proc mixed data=data;
  class id time;
  model y = time / random intercept = id;
run;
```

Use code with caution. [Learn more](#)

This model is equivalent to the R model above.

### Interpretation of the results

The output of the model will include estimates of the fixed effects and the random effects. The fixed effects represent the average change in the outcome variable (y) for a one-unit change in the predictor variable (time). The random effects represent the individual differences in the intercepts and slopes.

For example, if the fixed effect of time is estimated to be 1, this means that the outcome variable is expected to increase by 1 unit for every one-unit increase in time. If the random effect of individual is estimated to be 2, this means that some individuals have intercepts that are 2 units higher than the average intercept, and other individuals have intercepts that are 2 units lower than the average intercept.

Longitudinal hierarchical random effects models are a powerful tool for analyzing longitudinal data. They can be used to estimate the effects of predictors on the outcome variable, as well as to account for individual differences in the intercepts and slopes.

# Current Use of AI in Statistical Practice at Siemens Healthineers— A ‘Prompt’ Experiment (Bard vs ChatGPT)

Sure, here are examples of how to fit a longitudinal hierarchical random effects model using both R and SAS.

## R Example:

```
# Load necessary library
library(lme4)

# Assuming you have a dataset 'df' with a response 'y', a time variable 'time', a subject identifier 'subject', and a predictor 'x'
# df &lt;- read.csv("your_data.csv")

# Fit the model
model &lt;- lmer(y ~ time + x + (1|subject), data = df)

# Print the model summary
summary(model)
```

In this example, `lmer` from the `lme4` package is used to fit a linear mixed-effects model. The formula `y ~ time + x + (1|subject)` specifies that the response `y` is modeled as a function of `time` and `x`, with a random intercept for each `subject`.

## SAS Example:

```
/* Assuming you have a dataset 'df' with a response 'y', a time variable 'time', a subject identifier 'subject', and a predictor 'x' */
/* DATA df; SET your_data; RUN; */

PROC MIXED DATA=df;
  CLASS subject;
  MODEL y = time x / SOLUTION;
  RANDOM INTERCEPT / SUBJECT=subject;
RUN;
```

In this SAS example, `PROC MIXED` is used to fit the model. The `CLASS` statement specifies that `subject` is a categorical variable, and the `MODEL` statement specifies the fixed effects. The `RANDOM` statement specifies a random intercept for each `subject`.

Please replace `"your_data.csv"` and `"your_data"` with your actual data file name. Also, replace `y`, `time`, `x`, and `subject` with your actual column names.

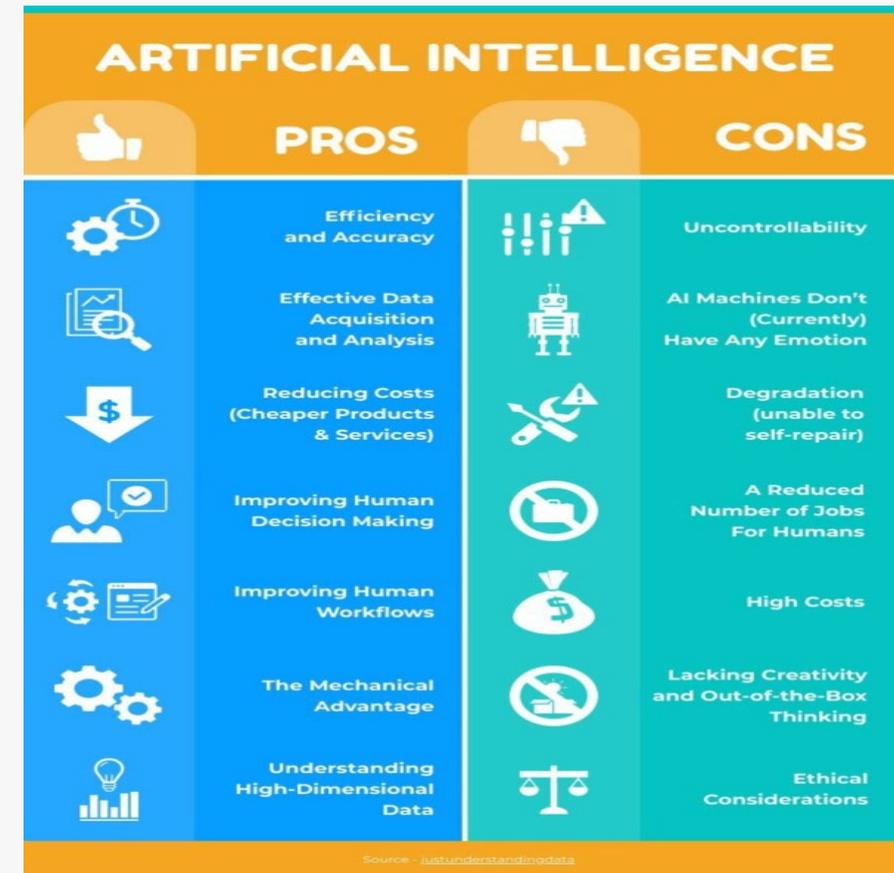
Remember, these are basic examples and might not be suitable for your specific use case. You might need to adjust the model specifications depending on the structure of your data and the specific research question you are trying to answer.

# Current Use of AI in Statistical Practice at Siemens Healthineers— Beta Testing in Biostatistics and Data Management

- New vs Experienced Employees
- Uses varied:
  - Program generating
  - Information Gathering
  - Training on theory, concepts, programmatic execution
  - Study Design
  - Ad Hoc Fix-It programs
  - Comparison of Statistical Techniques
  - Used as a Research Assistant
- What happened:
  - Use of ChatGPT Demonstrated a 10-15% increase in time to completion of regular activities within Data Management at both Director and Staff levels dependent on complexity level of tasks
  - Biostatistics tasks showed similar results with use of AI for construction of initial project setup activities i.e. customized analysis templates

# Current Use of AI in Statistical Practice at Siemens Healthineers— Beta Testing in Biostatistics and Data Management

- Concerns:
  - Over-Dependency or Reliance on AI i.e. lack of diligence on accuracy of deliverables
  - Bias potential
  - Data privacy concerns
  - Outside influence, can it be hacked?
- Pros:
  - Efficiency in programming for both functions
  - Improved time to onboarding completion of new hires and getting them into projects and contributing
  - Research and thought contribution



# Speculation on the Future

- Future Proofing
- Can we control the level of intelligence of an AI?
  - Mundane tasks
  - Computational Tasks
  - Society Changing Breakthroughs
  - Is it a One-size-fits all?
- Is the future one of AI dominance or digital replacement?
- Or
- A hybrid of human/AI in a Augmented Synergistic Collective Approach
- When AI start make decisions for themselves, they may or may not have constraints or even the possibility of constraints...i.e. Free Will?
  - This is the singularity that is anticipated... and perhaps feared.



# Conclusion

- From ChatGPT:
  - Overall, AI is a powerful tool that can be used by biostatisticians to improve the efficiency, accuracy, and impact of their work. By developing the skills listed above, **biostatisticians can and should position themselves to be leaders in this rapidly evolving field.**

# Acknowledgments

- Dr. Peace for this conference and support of students at Georgia Southern University,
- Ruth Whitworth for the behind-the-scenes support,
- Special Thanks to the BASS Planning Committee and Dr. Laura Gunn for the invitation,
- Siemens Healthineers for Sponsoring the Speaker.