

CHATGPT: ENGAGING WITH TECHNOLOGY IN THE GENERATIVE AI ERA

Borui Zhang, PhD

Natural Language Processing Specialist

George A. Smathers Libraries

University of Florida

boruizhang@ufl.edu



ChatGPT is Pioneering the Era of Generative AI:

- It “was born” in Nov 2022 at OpenAI. It enables users to have a conversation towards a desired length, format, style, and language.
- It has 180.5 million users (Dec 2023), 1.7 billion website visits in Oct 2023.
- It passes US Medical Licensing Exam.
- It passes exams from law and business schools.
- It can answer questions, generate, analyze data, and even code.
- Support 50 + languages
- It can lie too.

ChatGPT 3.5 ▾



How can I help you today?

Starts here!

Compare storytelling techniques
in novels and in films

Make up a story
about Sharky, a tooth-brushing shark superh...

Tell me a fun fact
about the Roman Empire

Design a database schema
for an online merch store

Message ChatGPT...



ChatGPT can make mistakes. Consider checking important information.

Creative Writing Coach

I'm excited to read your work and give you feedback to improve your skills.



Laundry Buddy

Ask me anything about stains, settings, sorting and everything laundry.



Tech Advisor

From setting up a printer to troubleshooting a device, I'm here to help you step-by-step.



Game Time

I can quickly explain board games or card games to players of any skill level. Let the games begin!

[From openai blog](#)

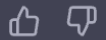
Translating languages

BR

帮我翻译How much wood would a woodchuck chuck, if a woodchuck could chuck wood?
成瑞典语



Hur mycket trä skulle en trähacka hacka om en trähacka kunde hacka trä?



https://en.wikipedia.org/wiki/How_much_wood_wo...

[How much wood would a woodchuck chuck - Wikipedia](#)

"How **much wood would a woodchuck chuck**" (sometimes phrased with "**could**" rather than "**would**") is an American English-language tongue-twister. The **woodchuck** ...

Missing: 翻译 瑞典 语

<https://zhidao.baidu.com/question> · [Translate this page](#)

[How much could a woodchuck chuck if a ... - 百度知道](#)

Jan 21, 2011 — A **woodchuck would chuck** all the **wood** he **could chuck** if a **woodchuck could chuck wood**. 翻译是： 如果一个土拨鼠能够抛木头的话， 那么一个土拨鼠可以抛 ...

[How much wood would a woodchuck chuck if a ... - 百度知道](#) Nov 1, 2016

求英语简单绕口令。如：[How much wood would a woodchuck ...](#) Dec 1, 2016

[More results from zhidao.baidu.com](#)

Missing: 瑞典 语

BR

You

I have two days in Orlando during Christmas. Plan a trip with a budget of 200 dollars for me.

**ChatGPT**

Spending two days in Orlando during Christmas on a budget of \$200 can be a bit challenging, but with some careful planning, you can have a memorable experience. Keep in mind that prices may vary, and it's essential to prioritize your interests. Here's a suggested itinerary:

Day 1: Exploring Theme Parks

1. **Morning:** Start your day early by heading to one of the theme parks. Consider visiting a less crowded one to save money on tickets. Options include Universal Orlando Resort or SeaWorld Orlando.
2. **Lunch:** Pack a simple picnic lunch to save money. Most parks allow outside food.
3. **Afternoon:** Enjoy the rides, shows, and attractions at the theme park of your choice.
4. **Dinner:** Look for budget-friendly dining options near your hotel or in the International Drive area. Places like fast-food restaurants or food courts are good options.

Planning a trip?

Day 2: Outdoor Fun and Free Attractions

1. **Morning:** Explore some of Orlando's natural beauty by visiting a park or garden. Leu Gardens or Lake Eola Park are great choices.
2. **Lunch:** Grab a quick bite from a local deli or food truck.
3. **Afternoon:** Visit ICON Park, which is free to enter and stroll around. While there, you can choose to pay for attractions like The Wheel or Madame Tussauds Orlando if it fits within your budget.

Creating a booklist?

You

I made a new friend who is curious about linguistics. Can you come up with a fun booklist of linguistics to recommend to my friend?

ChatGPT

Of course! Here's a list of fun and engaging books on linguistics that your friend might enjoy:

1. "The Mother Tongue: English and How It Got That Way" by Bill Bryson - Bill Bryson takes a humorous and informative look at the history and quirks of the English language.
2. "The Power of Babel: A Natural History of Language" by John H. McWhorter - This book explores the evolution and diversity of languages in an accessible and engaging way.
3. "Word by Word: The Secret Life of Dictionaries" by Kory Stamper - A delightful behind-the-scenes look at the world of lexicography and the making of dictionaries.

You

Can you add some fun theoretical linguistics books to this list

ChatGPT

Certainly! Here's an updated list that includes some fun theoretical linguistics books:

"Syntactic Structures" by Noam Chomsky - A classic in theoretical linguistics, this book introduces Chomsky's transformational-generative grammar and revolutionized the field.

"Grammatical Man: Information, Entropy, Language, and Life" by Jeremy Campbell - An intriguing book that combines linguistics with information theory and explores the theoretical foundations of language.

BR You

create a small data set, format the set, and use sklearn to show a demon of machine learning code

ChatGPT

Sure, I can help you create a small dataset, format it, and demonstrate a simple machine learning example using Python and Scikit-Learn. In this example, we'll create a dataset of student exam scores and use a linear regression model to predict their final exam scores based on their midterm scores.

First, let's generate a synthetic dataset:

```
python Copy code

import numpy as np

# Create random data for midterm scores and final exam scores
np.random.seed(0)
midterm_scores = np.random.randint(0, 100, 50)
final_exam_scores = 0.5 * midterm_scores + np.random.normal(0, 5, 50)

# Combine the data into a single dataset
data = np.column_stack((midterm_scores, final_exam_scores))
```

Now, let's format the dataset into feature (X) and target (y) arrays:

Learning how to code?

Now, let's format the dataset into feature (X) and target (y) arrays:

```
python Copy code

X = data[:, 0].reshape(-1, 1) # Midterm scores as features
y = data[:, 1] # Final exam scores as targets
```

Next, we can split the dataset into a training set and a test set:

```
python Copy code

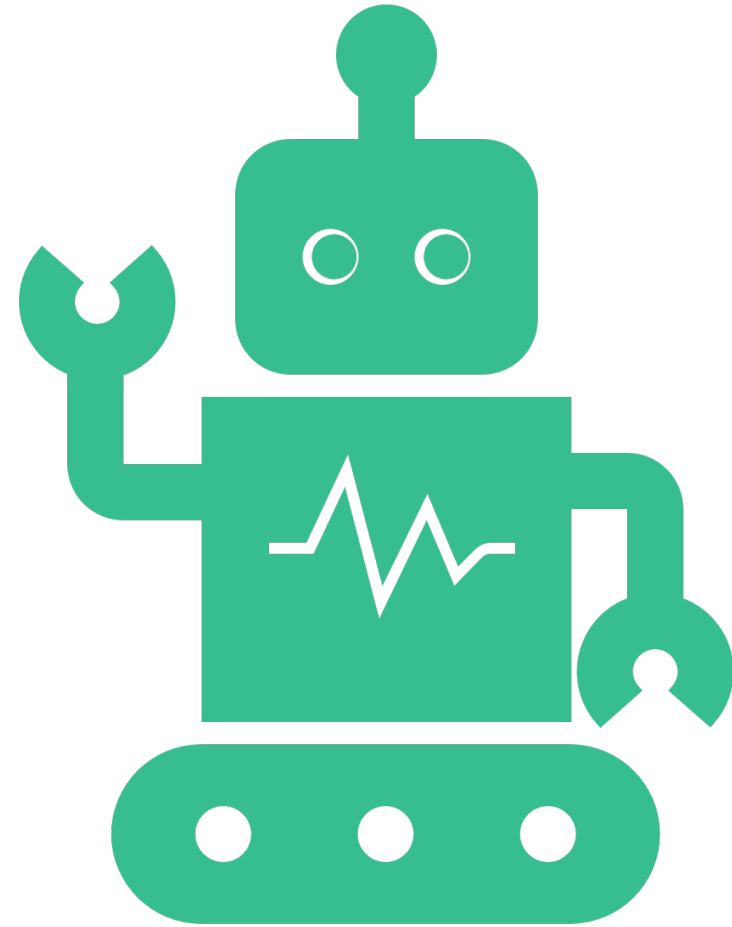
from sklearn.model_selection import train_test_split

X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2)
```

Comparison with Traditional Search Engines:

- ChatGPT is more **personalized and interactive** through conversations.
- Search engines can provide more specific results when a variety of sources are needed.
- ChatGPT **generates** answers and traditional search engines retrieve information.
- Currently, ChatGPT has no advertisements.

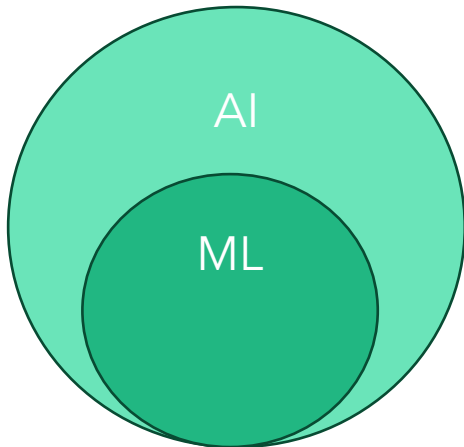
What is Generative AI?



What's AI (Artificial Intelligence)?

Artificial Intelligence (AI) A system which amplifies people's own knowledge and understanding. -

Vannevar Bush (1945)



Machine learning (ML) is the science of getting computers to act without being explicitly programmed. - *Andrew Ng*

Being **explicitly** programmed?



If I forget to feed my dog, **then** it will be sad

If I give my dog a treat, **then** it will be happy

If my dog gets nails clipped, **then** it will be sad

If my family leaves for vacation, **then** it will be sad

If I give my dog a toy, **then** it will be happy

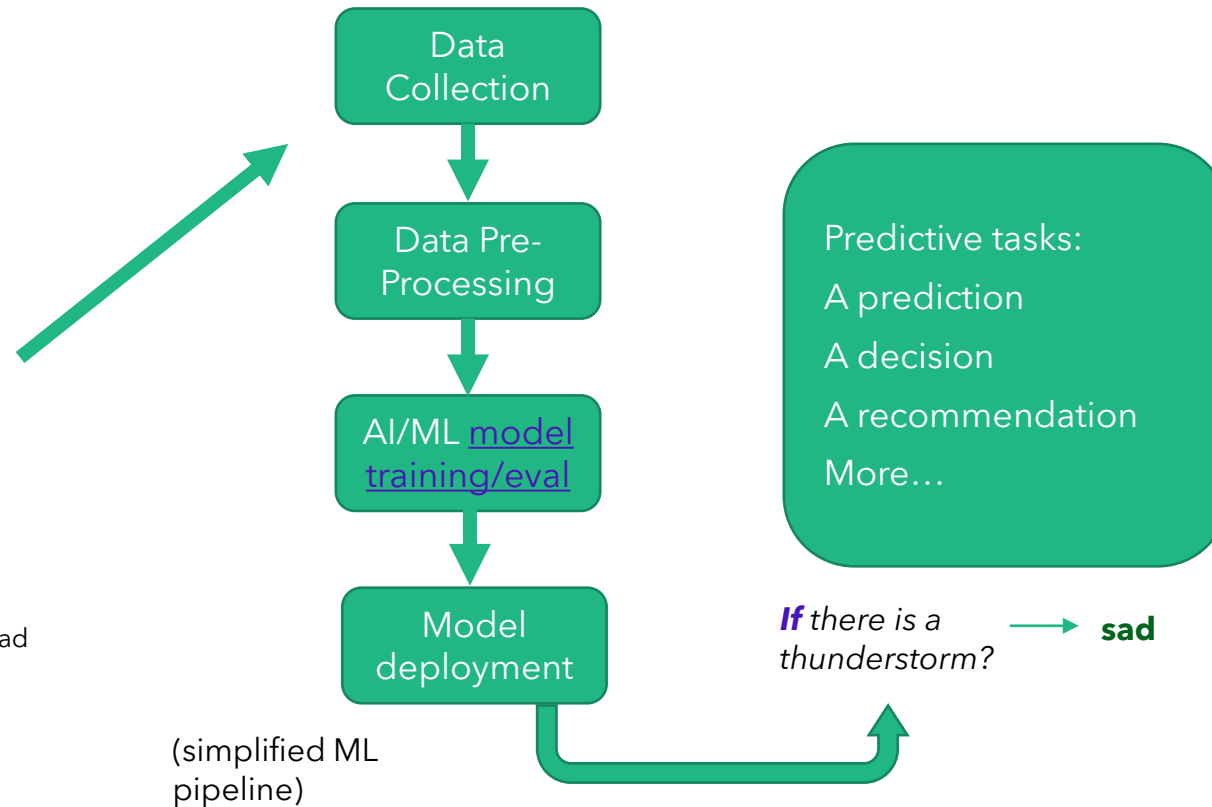
If my dog is sick, **then** it will be sad

More...

AI/ML learns rules **implicitly** from data



- If** I forget to feed my dog, **then** it will be sad
- If** I give my dog a treat, **then** it will be happy
- If** my dog gets nails clipped, **then** it will be sad
- If** my family leaves for vacation, **then** it will be sad
- If** I give my dog a toy, **then** it will be happy
- If** my dog is sick, **then** it will be sad
- More...



Famous AI “learners”

“Game Players”



Silver, David, et al. "Mastering the game of Go with deep neural networks and tree search." *nature* 529.7587 (2016): 484-489.

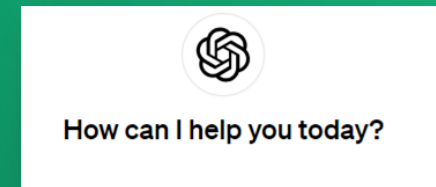
“Artists”



DALL-E
Computer
Vision

Ramesh, Aditya, et al. "Hierarchical text-conditional image generation with clip latents." *arXiv preprint arXiv:2204.06125* 1.2 (2022): 3.
Ramesh, Aditya, et al. "Zero-shot text-to-image generation." International Conference on Machine Learning. PMLR, 2021.

“Conversational Partners”



ChatGPT
GPT-4
Large Language
Models

Ouyang, Long, et al. "Training language models to follow instructions with human feedback." *Advances in Neural Information Processing Systems* 35 (2022): 27730-27744.

AlphaGO



AlphaGo

AlphaGo vs Lee Sedol. DeepMind. 2016. <https://www.deepmind.com/alpha-go>


DALL-E



Computer Vision

panda and scientist mixing sparkling chemicals, animation

ChatGPT

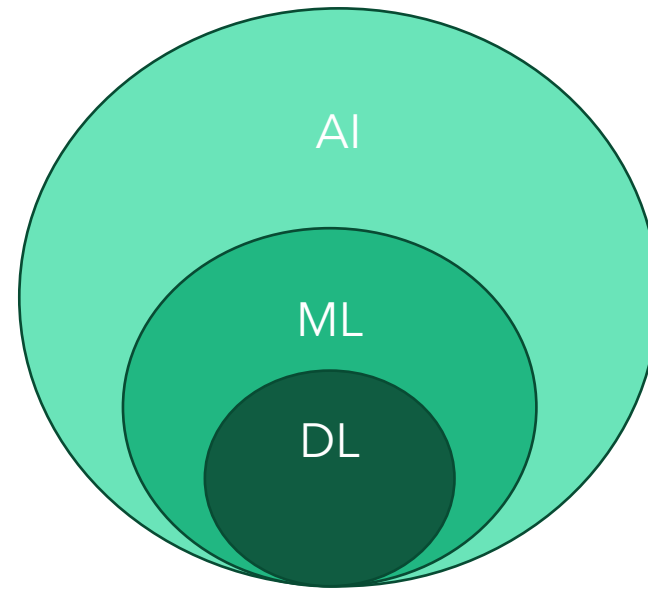


How can I help you today?

Large Language Models

OpenAI. 2022. "Training language models to follow instructions in natural language." *arXiv preprint arXiv:2203.01369*

Deep learning | neural network models (DL or DNN)



DL is good at :

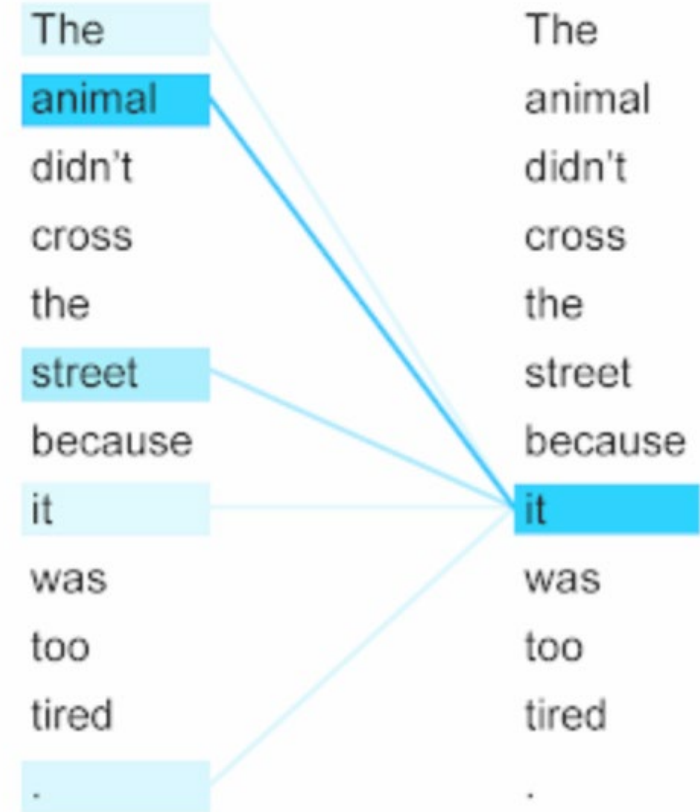
Complex Data Understanding

Large Scale Data Handling

Transformer model particularly handles sequential data well

The simple goal of the **language model** is to predict the **next "word"**.

"The animal didn't cross the street because it was too tired."



So, what is ChatGPT?



Chat Generative Pre-trained Transformer



Conversation-based

Generates data: text, image.
Videos not yet.

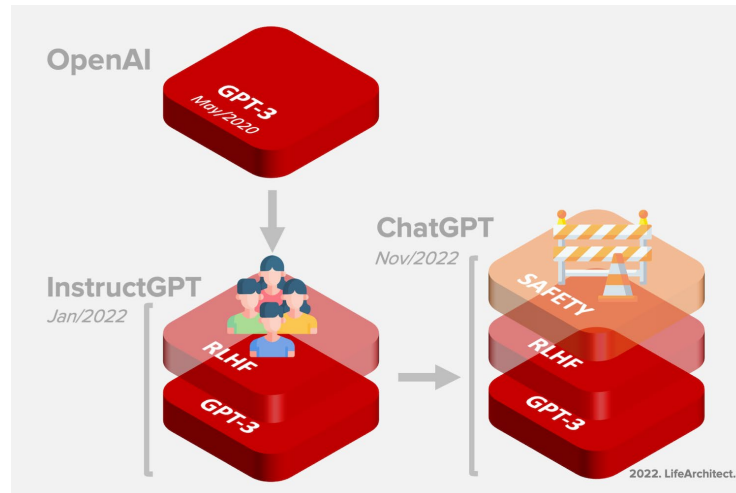
Language Model

Based on Transformer DL Architecture

GPT series

Features	GPT-1 (2018)	GPT-2 (2019)	GPT-3 (2020)	GPT-4 (2023)
Parameters	110 million	1.5billion	175 billion	More than GPT-3
Performance	Limitations in longer context understanding and coherence	Generating more coherent and contextually relevant text	Generates human-like text, struggles with sarcasm and idiomatic expressions	Expected to significantly improve upon GPT-3's performance
Potential Applications	Basic conversational agents, simple QA systems	More coherent and contextually relevant text	Chatbots, virtual assistants, content generation, research	Creative writing, art, improvement of existing applications

Try to be a good family member!



Reinforcement Learning from Human Feedback (**RLHF**)

-- a policy model that governs the generation of outputs within a given environment by controlling the states.

What makes ChatGPT outstanding among Gen LLMs?

Tips of optimizing the use of ChatGPT



LLM Challenges

Hallucination challenges: AI generating incorrect or misleading information, often due to gaps in training data or misunderstanding context.

Privacy concerns: Risks of personal and sensitive user data by AI systems.

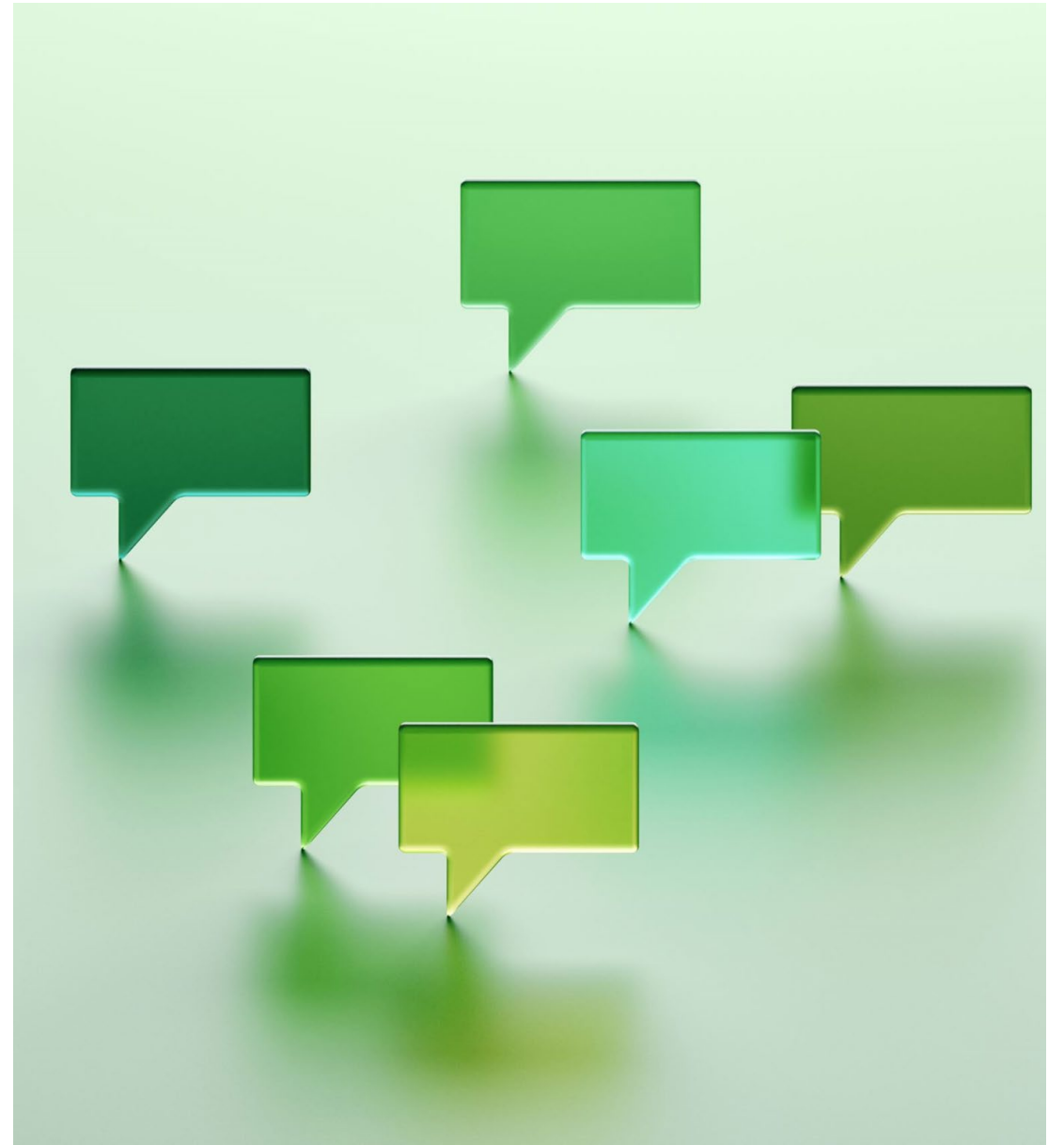
Copyright complications: Difficulties in ensuring that AI-generated content does not infringe upon existing intellectual property rights or copyrights.

Model not open-sourced: Lack of public access to the AI's source code, limiting external review and customization.

Ask	specific questions or provide clear instructions
Provide	relevant context to mitigate hallucination
Avoid	ambiguity
Adhere	research policies
No	fixed golden rules

Writing Effective Prompts

Future Expectations in Generative AI



Model Competition: More nuanced generative models might come out.

Business: Embedding LLM and Computer vision in various devices, from smartphones to cars.

Education: Reshaping teaching and learning

Tech: Simplifying coding tasks by using LLMs to automate basic coding tasks, balancing ease of use with complexity and improving debugging processes.

User experience: Adapting to changes in payment processing, ensuring server stability, and enhancing the overall speed and efficiency of user experiences.

These might
already be
happening...

Summary

An initial overview of how ChatGPT functions, illustrated through examples.

Explanations of AI, ML, DL and the specific domain of Generative AI.

Details on ChatGPT's model architecture

Some thoughts on current challenges and future expectation on Gen AI

One last message --

This is the best time to learn!

Thank you!

Borui Zhang

boruizhang@ufl.edu

University of Florida

References:

- <https://healthitanalytics.com/news/chatgpt-passes-us-medical-licensing-exam-without-clinician-input>
- <https://openai.com/blog/introducing-gpts>
- <https://botpress.com/blog/list-of-languages-supported-by-chatgpt#:~:text=How%20Many%20Languages%20Does%20ChatGPT,%2C%20Arabic%2C%20and%20many%20more.>
- Silver, David, et al. "Mastering the game of Go with deep neural networks and tree search." *nature* 529.7587 (2016): 484-489.
- Ramesh, Aditya, et al. "Hierarchical text-conditional image generation with clip latents." *arXiv preprint arXiv:2204.06125* 1.2 (2022): 3.
- Ramesh, Aditya, et al. "Zero-shot text-to-image generation." *International Conference on Machine Learning*. PMLR, 2021.
- Ouyang, Long, et al. "Training language models to follow instructions with human feedback." *Advances in Neural Information Processing Systems* 35 (2022): 27730-27744.