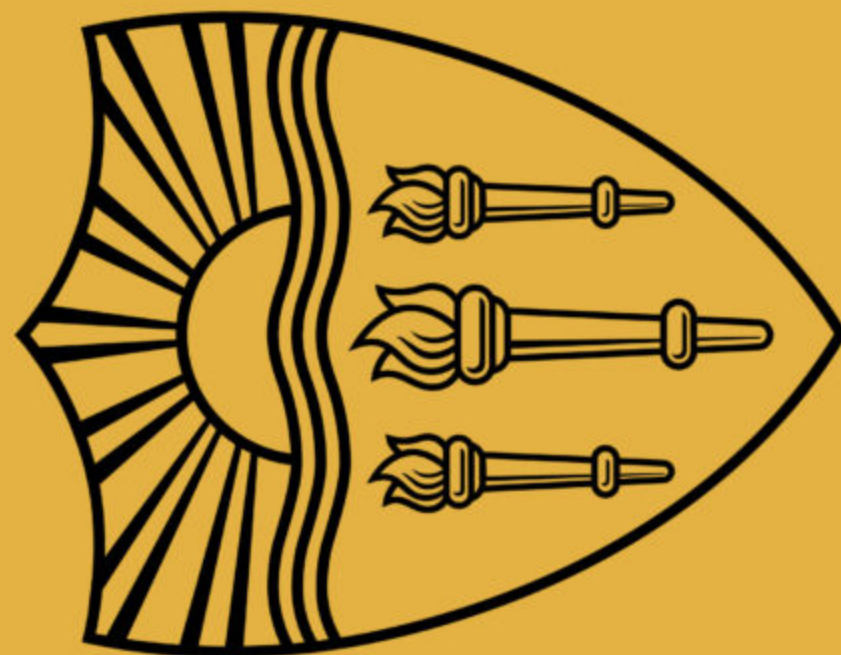


C  
S  
I

# Lecture 01: Introduction and Course Overview

*Swabha Swayamdipta*  
*USC CSCI 499, Spring 2024*  
*Jan 08, 2024*



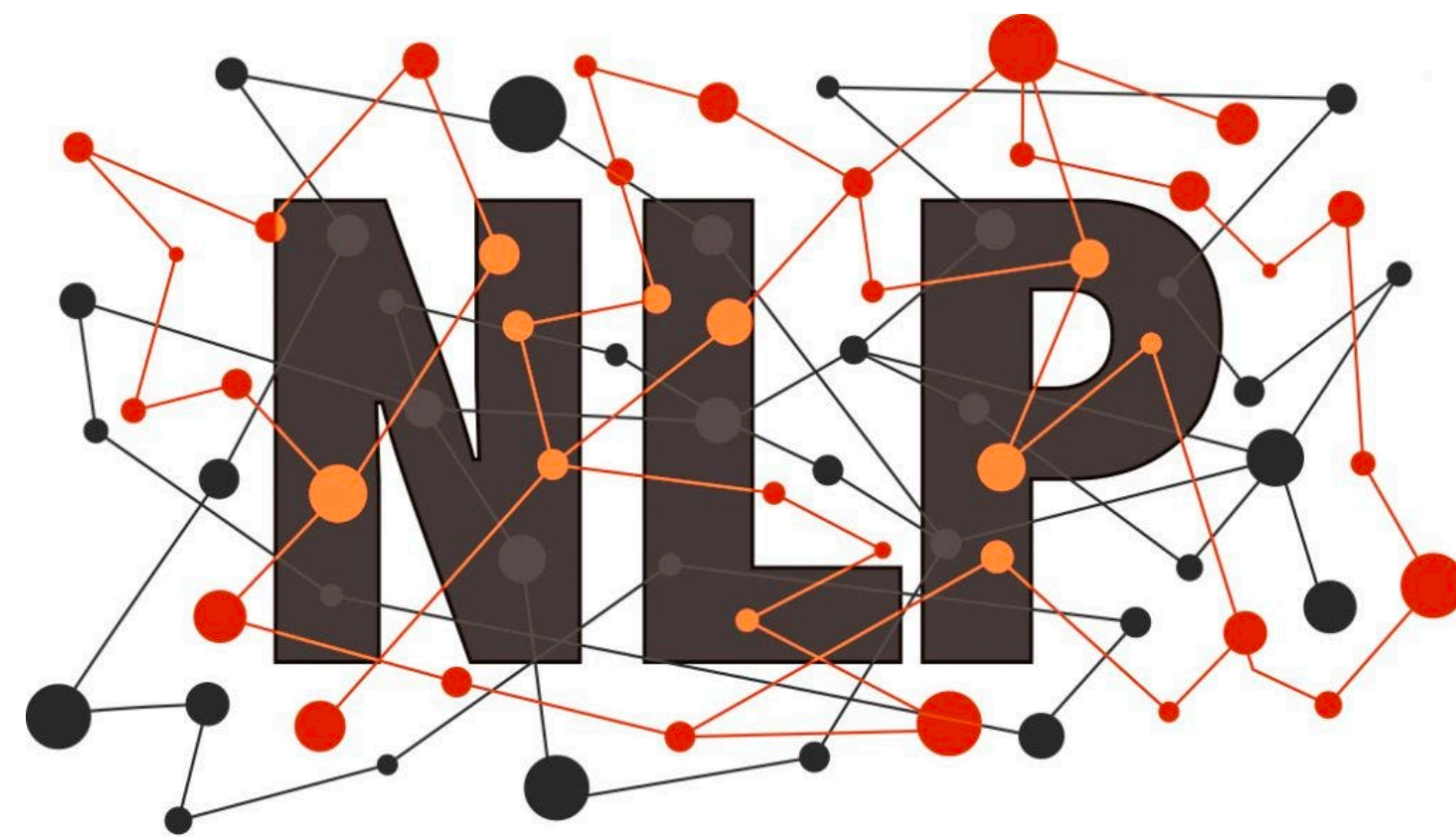
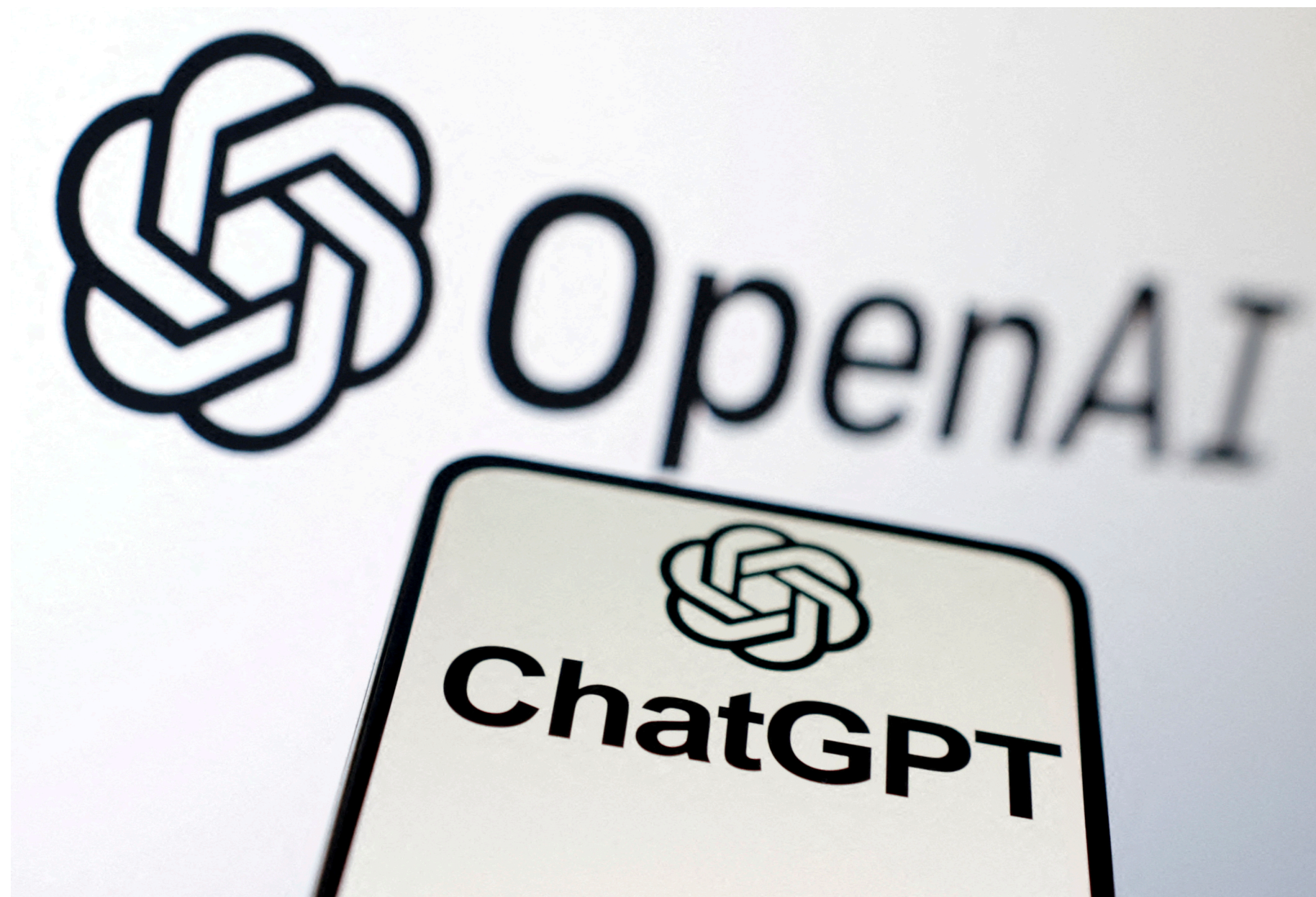


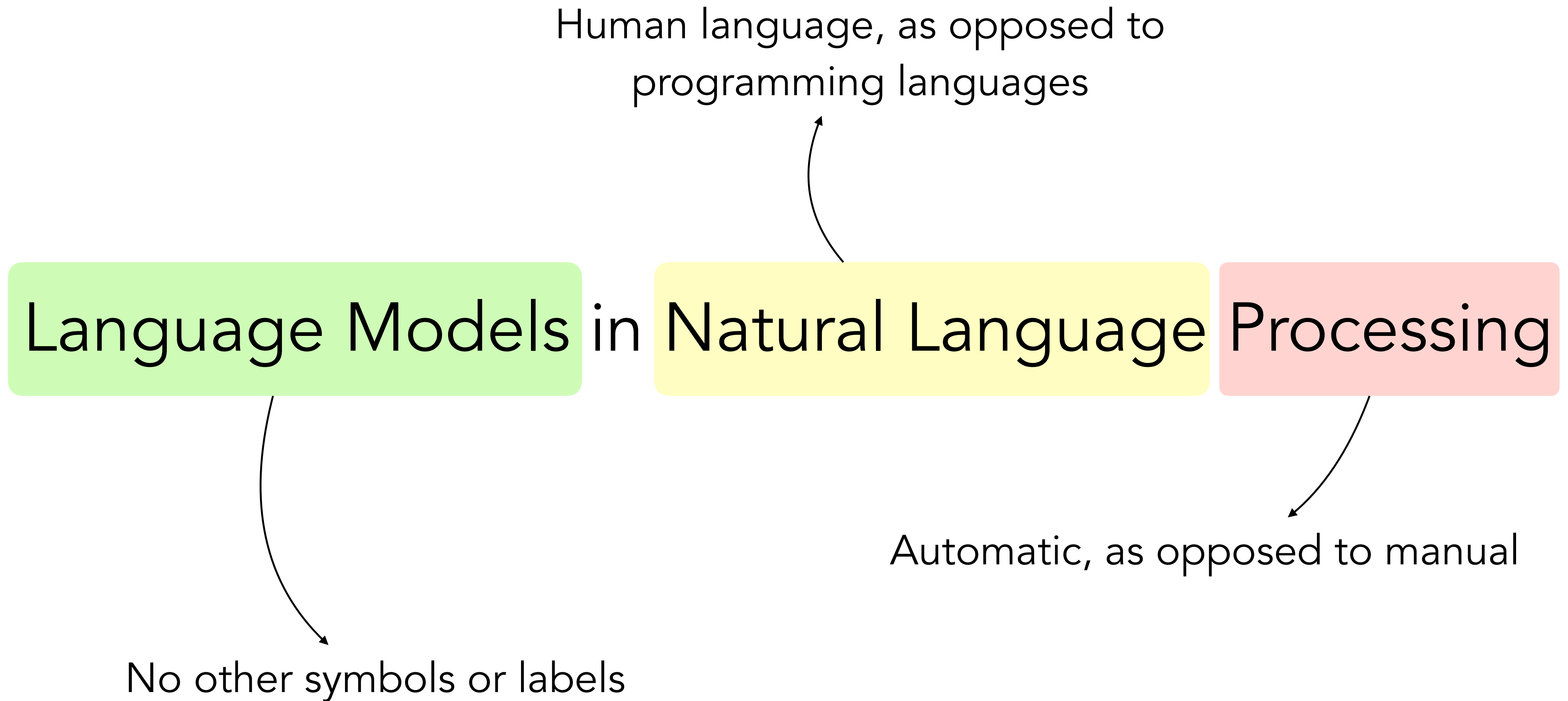
Image Source: Towards Data Science

**You**  
What is a good joke about large language models?

**ChatGPT**  
Sure, here's a lighthearted joke about large language models:  
  
Why did the language model bring a ladder to the conversation?  
  
Because it wanted to reach the next level of understanding!

**You**  
Can you make it funnier?

**ChatGPT**  
Certainly, let's try to add more humor:  
  
Why did the GPT model start a comedy career?  
  
Because its jokes were so well-structured, they always got a perfect "10<sup>24</sup>" laughs!

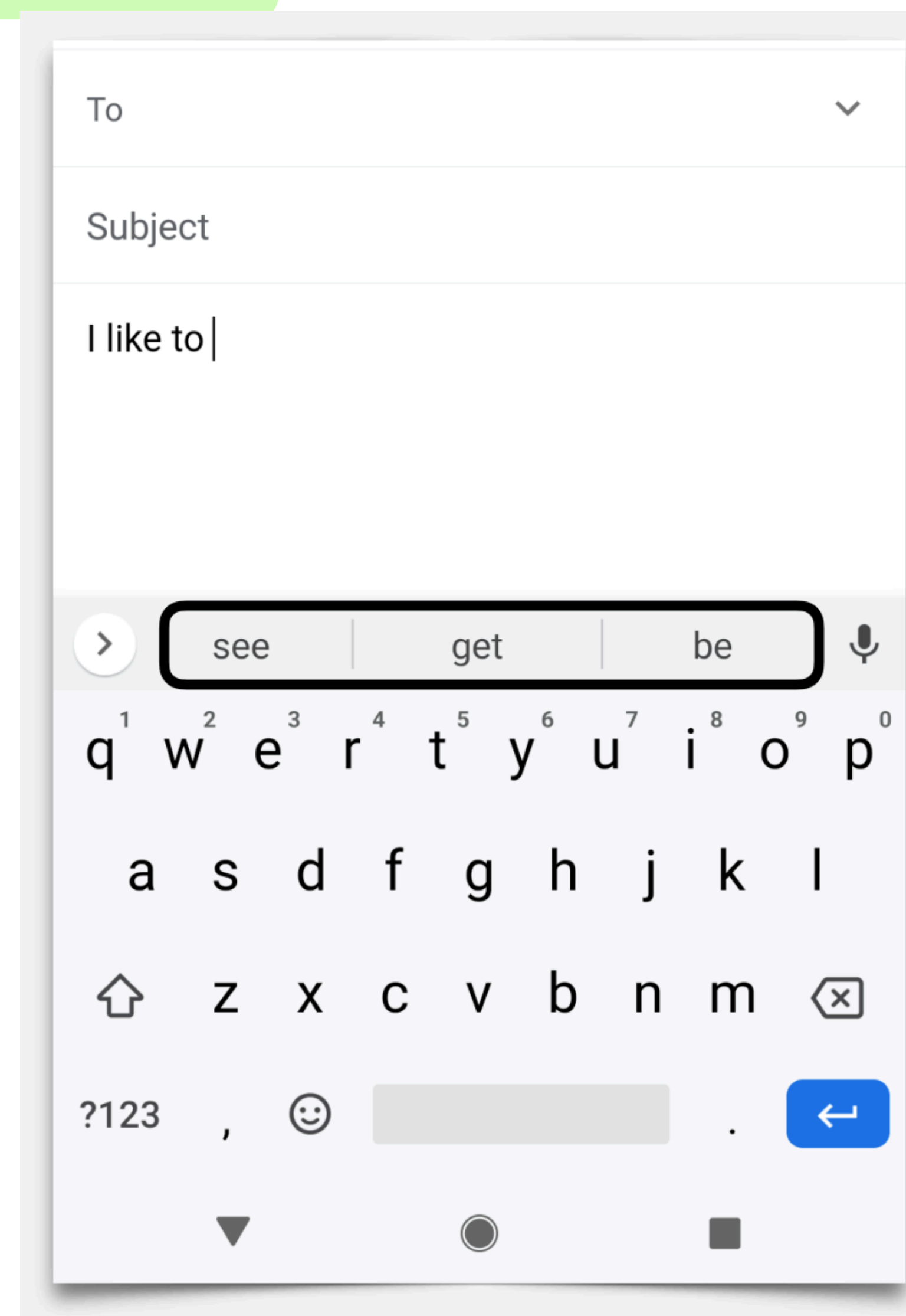


# What is Natural Language Processing ?

- Field at the intersection of computer science, AI (especially machine learning) and linguistics
- Goal: for computers to process human language, similar to human understanding, towards performing useful tasks
- Challenge: understanding and representing the meaning of language is something even humans struggle with
- Processing: produce outputs (Y) with language or text as input (X)
  - Outputs can be more language / text as well!

# Language Models

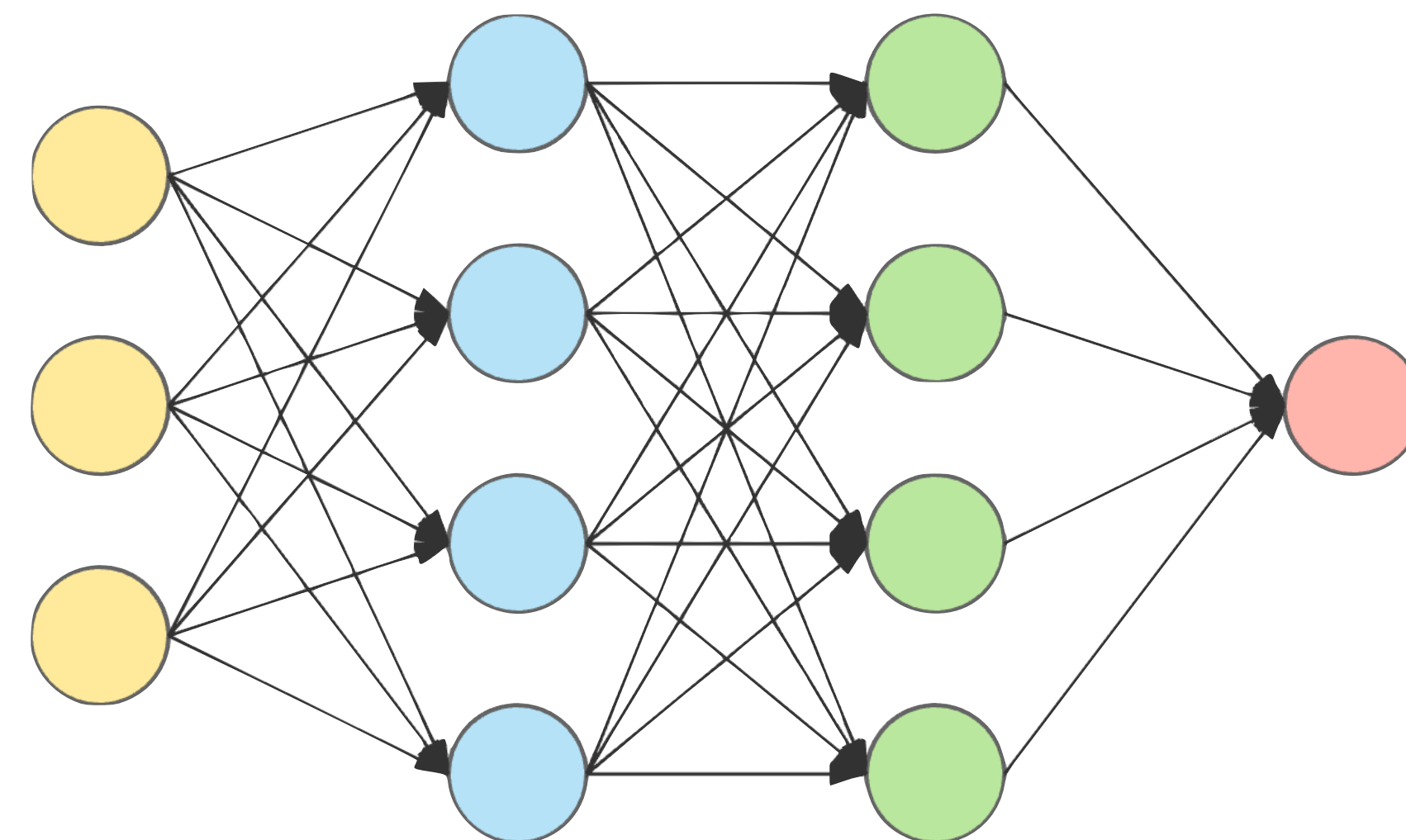
- Task: Given a sequence of words so far (**the context**), predict what comes next
  - Like autocomplete!
- We never know for sure what comes next, but we can still make good guesses!
- Question: what is X and what is Y here?
  - Input X = Language and Output Y = Language



# Language Model as Autocomplete

- Task: Given a sequence of words so far (the context), predict what comes next
- We never know for sure what comes next, but we can still make good guesses!

Dance the...



What words can follow this?



What is common to these words?

... night  
 ... charleston  
 ... cha-cha  
 ...atre

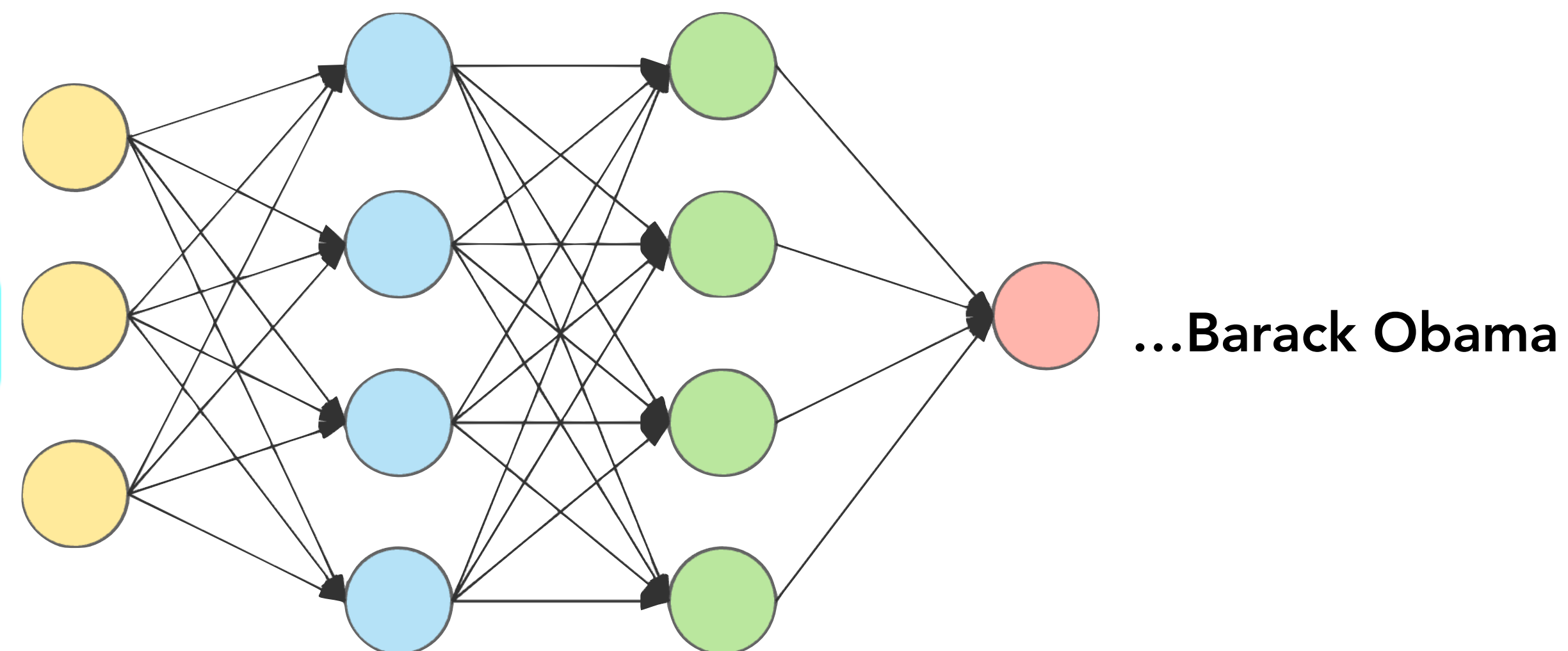
# Language Model as Autocomplete

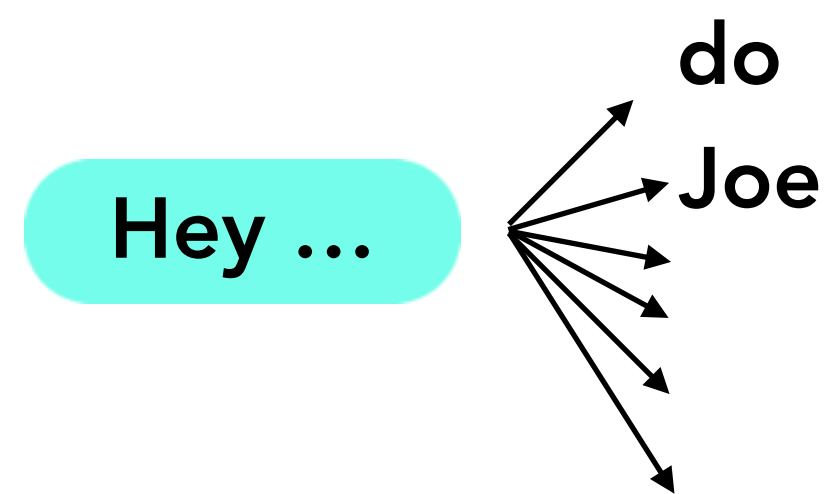
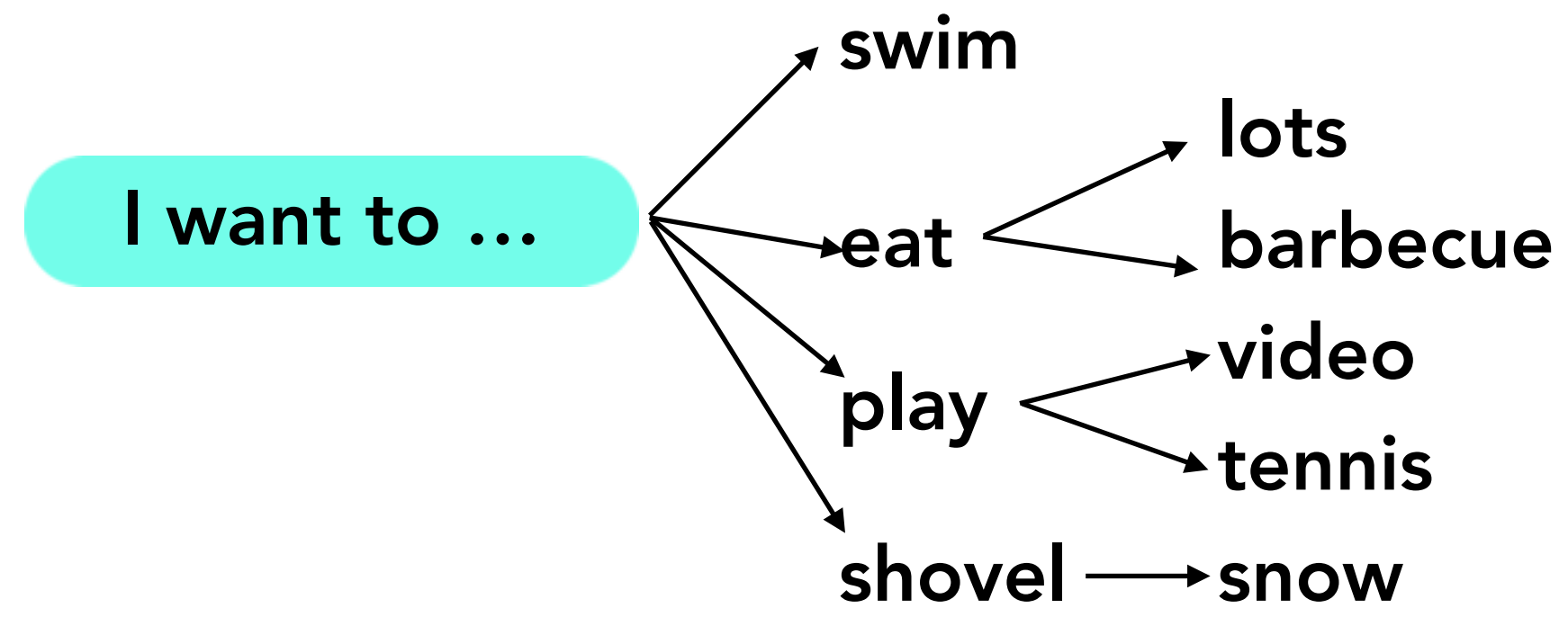
- Task: Given a sequence of words so far (the context), predict what comes next
- We never know for sure what comes next, but we can still make good guesses!



What words can follow this?

The 44th President of United States was ...







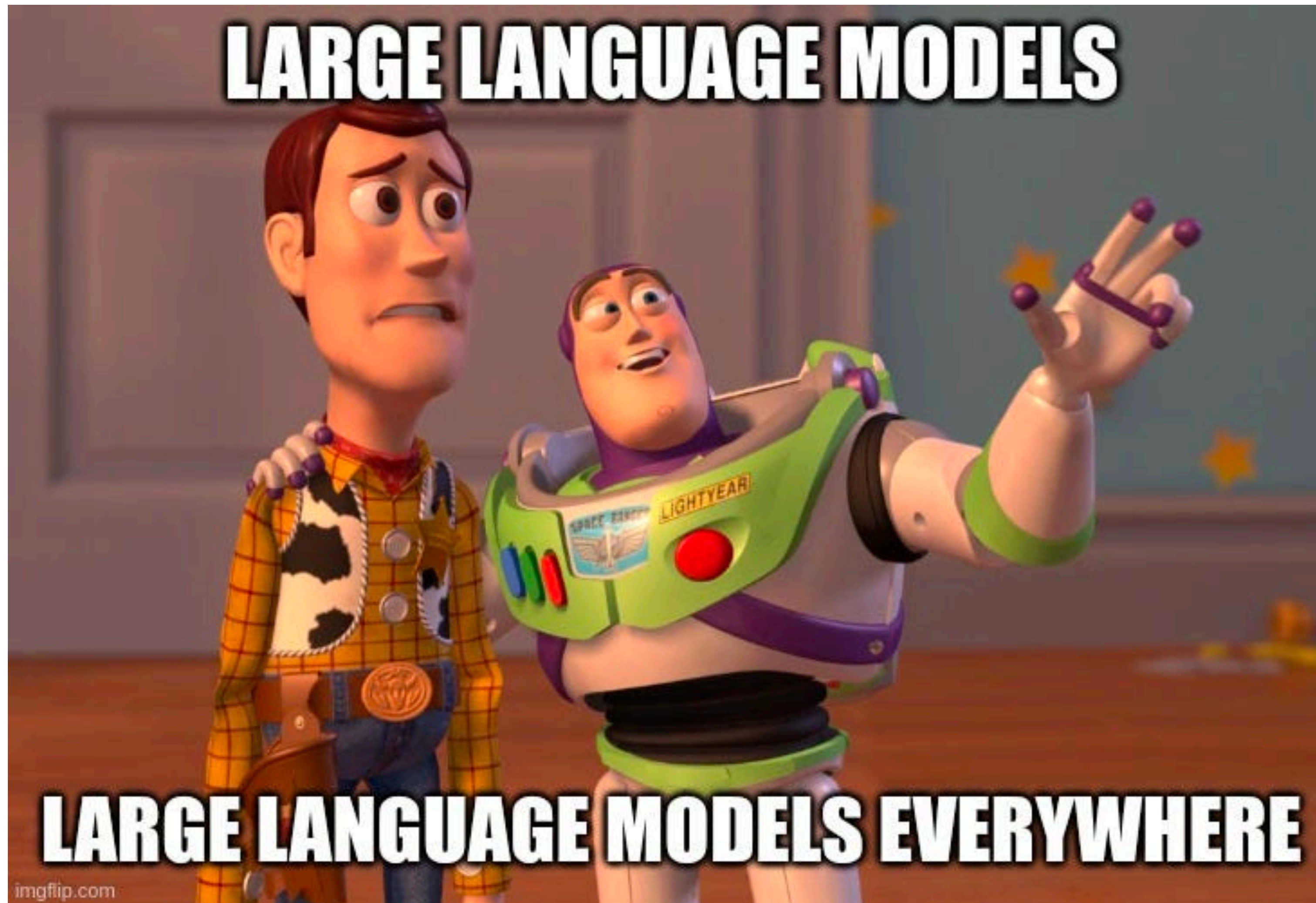
# Today's Lecture: Introduction to LMs in NLP

1. What are language models?
2. Why study language models?
3. LLMs - capabilities and risks
4. Class Syllabus
5. Class Logistics

CSCI499 Spring 2024: Language Models in NLP

 Spring 2024  Mon / Wed 4:00 - 5:50p  [KAP 134](#)

# Why study Language Models in Natural Language Processing?



# Language Models Are Everywhere



**Virtual Assistants**



**Translation**



**Content Creation**



**Code Generation**



**Transcription**



**Sentiment Analysis**

# Apple's Siri

- Understands the user
- Remembers what the user said earlier
- Can understand which alarm she is referring to



# Google Translate



- Detects language automatically
- Can reorder spans in text on the fly

中共中央政治局7月30日召开会议，会议分析研究当前经济形势，部署下半年经济工作。

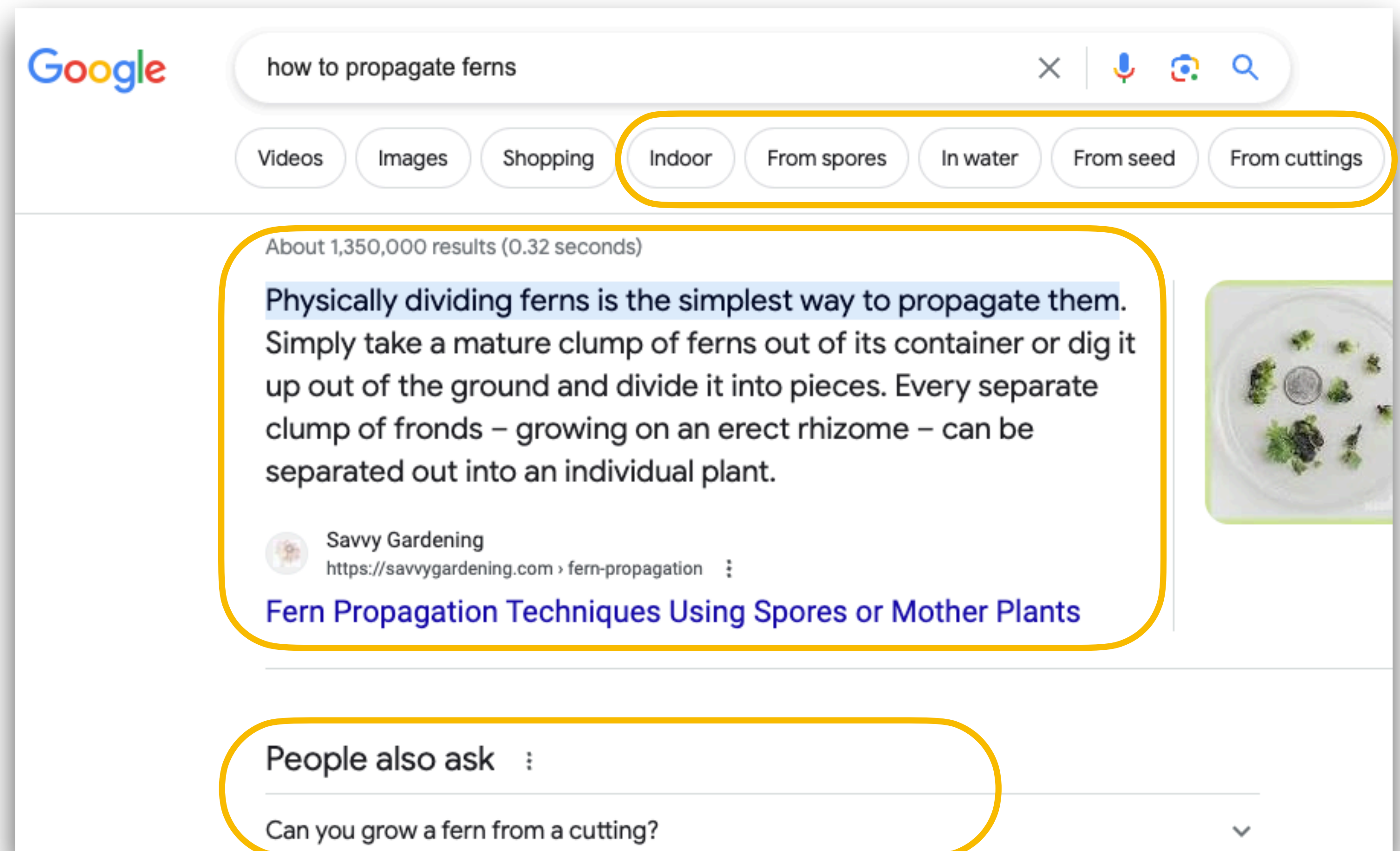
People's Daily, August 10, 2020

Translate

The Political Bureau of the CPC Central Committee held a meeting on July 30 to analyze and study the current economic situation and plan economic work in the second half of the year.

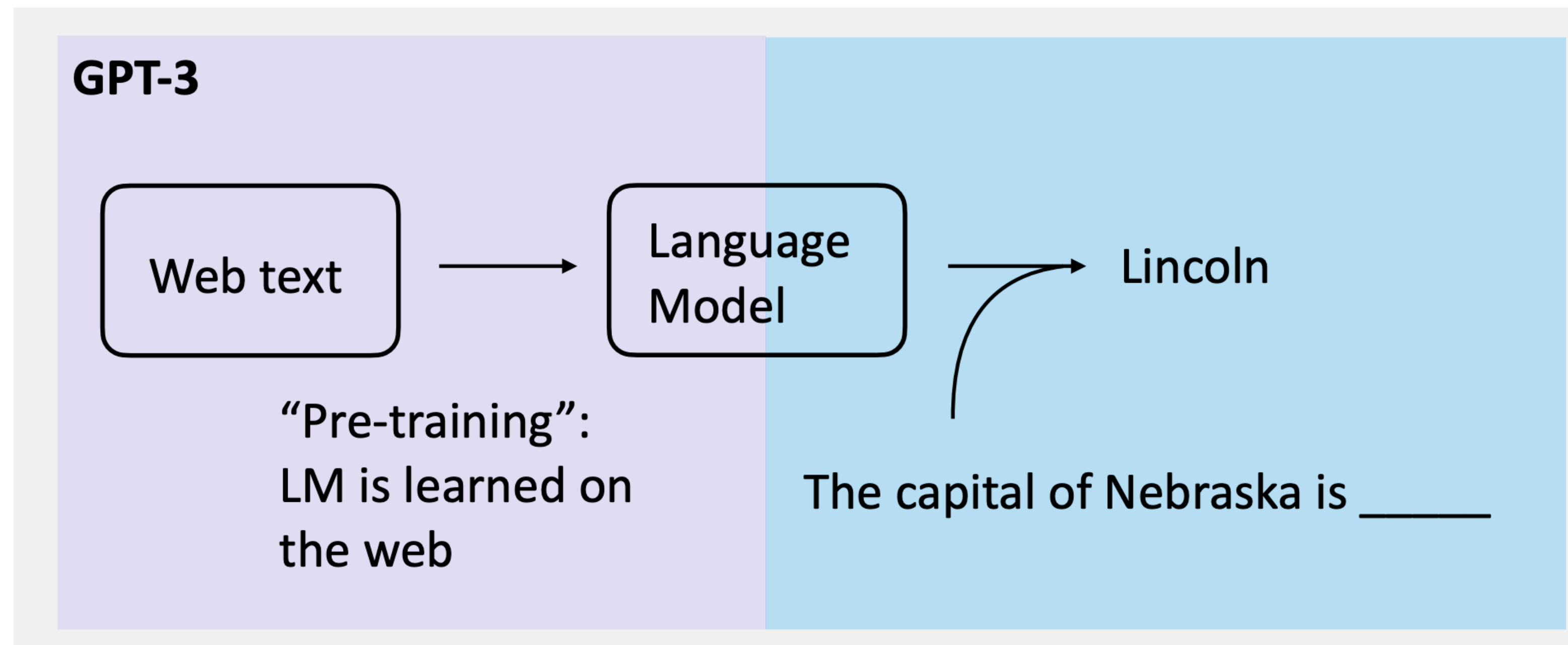
# Google Search

- Understands that a fern can be indoor, can be propagated either from seed or from cuttings
- Can find the exact passage in a webpage that answers the questions
- Can find related (in meaning) questions



# Large Language Models (LLMs)

- These models are trained over a ton of data (a curated scrape of the web). So they will have seen information about Nebraska and Lincoln.
- A big enough model can answer questions even without being trained to do so. What else can we get these models to do?





# GPT-4 Passes the Bar Exam!

## GPT Takes the Bar Exam

December 29, 2022

Michael Bommarito II<sup>1,2,3</sup>, Daniel Martin Katz<sup>1,2,3,\*</sup>

- 1 Illinois Tech - Chicago Kent College of Law (Chicago, IL USA)
- 2 Bucerius Law School (Hamburg, Germany)
- 3 CodeX - The Stanford Center for Legal Informatics (Stanford, CA USA)

\* Corresponding Author: dkatz3@kentlaw.iit.edu

### Abstract

Nearly all jurisdictions in the United States require a professional license exam, commonly referred to as “the Bar Exam,” as a precondition for law practice. To even sit for the exam, most jurisdictions require that an applicant completes at least seven years of post-secondary education, including three years at an accredited law school. In addition, most test-takers also undergo weeks to months of further, exam-specific preparation. Despite this significant investment of time and capital, approximately one in five test-takers still score under the rate required to pass the exam on their first try. In the face of a complex task that requires such depth of knowledge, what, then, should we expect of the state of the art in “AI?” In this research, we document our experimental evaluation of the performance of OpenAI’s TEXT-DAVINCI-003 model, often-referred to as GPT-3.5, on the multistate multiple choice (MBE) section of the exam. While we find no benefit in fine-tuning over GPT-3.5’s zero-shot performance at the scale of our training data, we do find that hyperparameter optimization and prompt engineering positively impacted GPT-3.5’s zero-shot performance. For best prompt and parameters, GPT-3.5 achieves a headline correct rate of 50.3% on a complete NCBE MBE practice exam, significantly in excess of the 25% baseline guessing rate, and performs at a passing rate for both Evidence and Torts. GPT-3.5’s ranking of responses is also highly-correlated with correctness; its top two and top three choices are correct 71% and 88% of the time, respectively, indicating very strong non-entanglement performance. While our ability to interpret these results is limited by nascent scientific understanding of LLMs and the proprietary nature of GPT, we believe that these results strongly suggest that an LLM will pass the MBE component of the Bar Exam in the near future.

	GPT	GPT Top 2	GPT Top 3	NCBE
Evidence	63%	84%	98%	65%
Torts	62%	72%	93%	71%
Civil Procedure	52%	63%	79%	59%
Constitutional Law	49%	67%	87%	72%
Real Property	45%	72%	85%	65%
Contracts	45%	77%	86%	70%
Criminal Law & Procedure	35%	62%	86%	71%
AVERAGE	50%	71%	88%	68%

Table 2. Summary of performance by question category for GPT-3.5 and NCBE-Reported Students

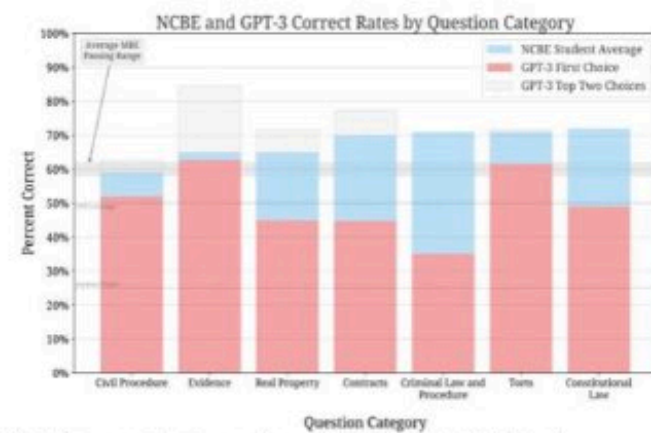


Fig 1. Summary of performance by question category for GPT-3.5 and NCBE-Reported Students

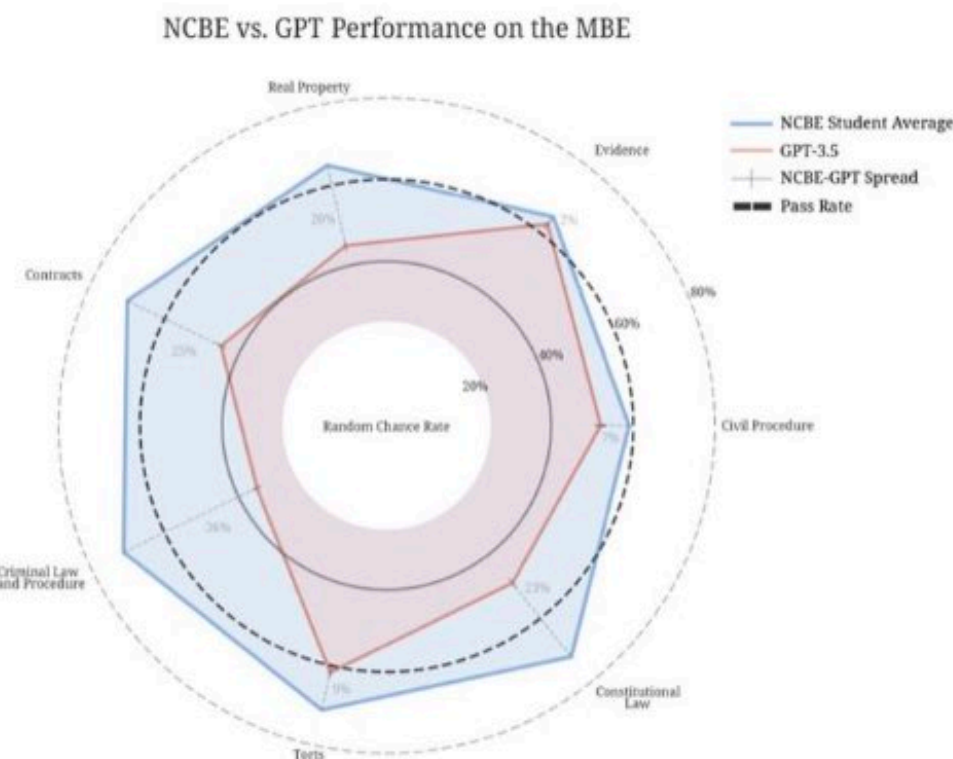
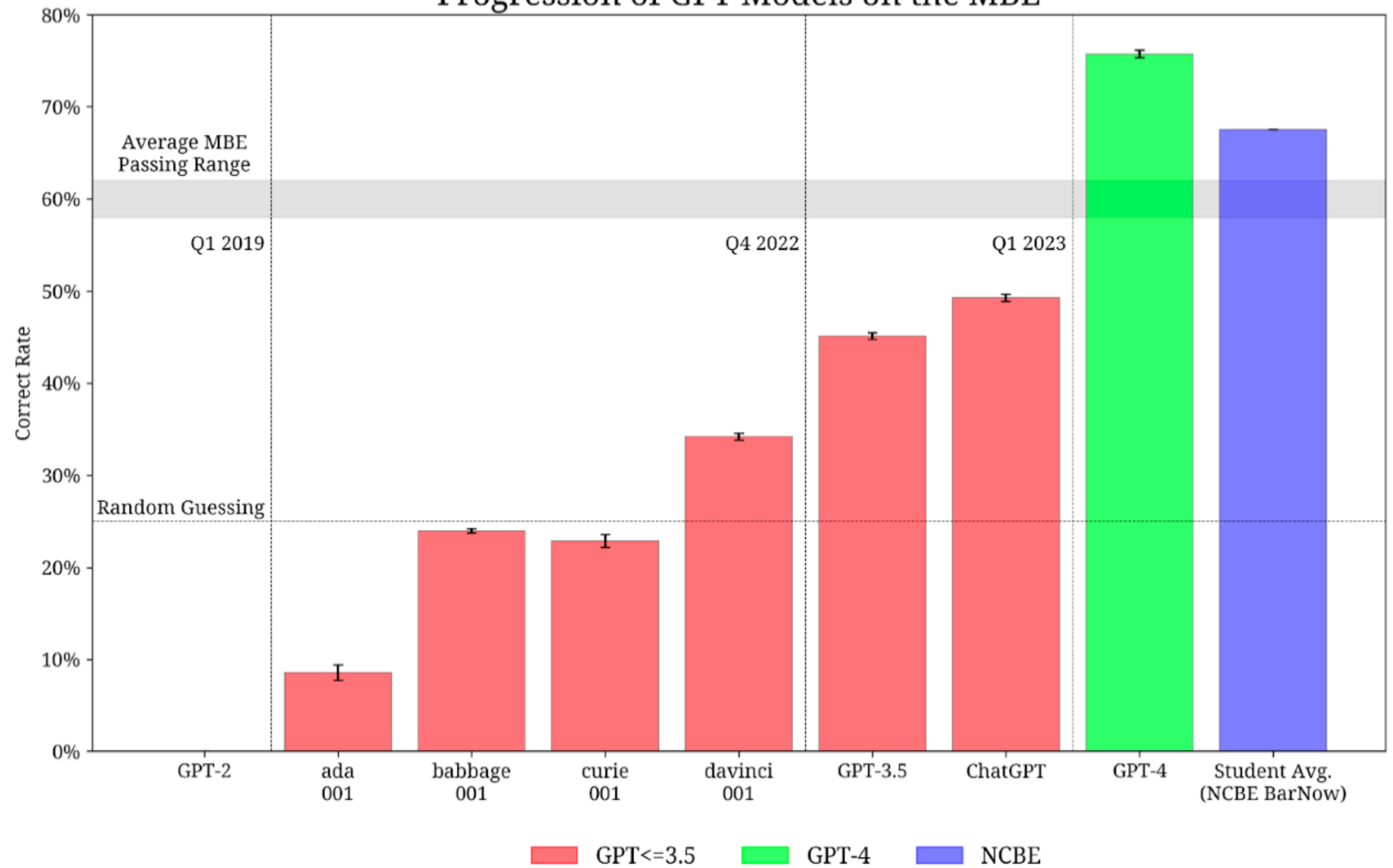


Fig 2. Accuracy by Question Category for GPT and Average Test-Takers

### Progression of GPT Models on the MBE



Legend: GPT<=3.5 (Red), GPT-4 (Green), NCBE (Blue)

# LLMs are being used to...

Code computer programs!

Summarize articles, podcasts or presentations

Draft emails

Script social media posts

Create a title for an article

Play games

Assist with job searches including resumes and cover letters

Ask trivia questions

Extremely powerful: can in many cases replace laborious manual efforts

Describe complex topics more simply

engine optimization.

Formulate product descriptions

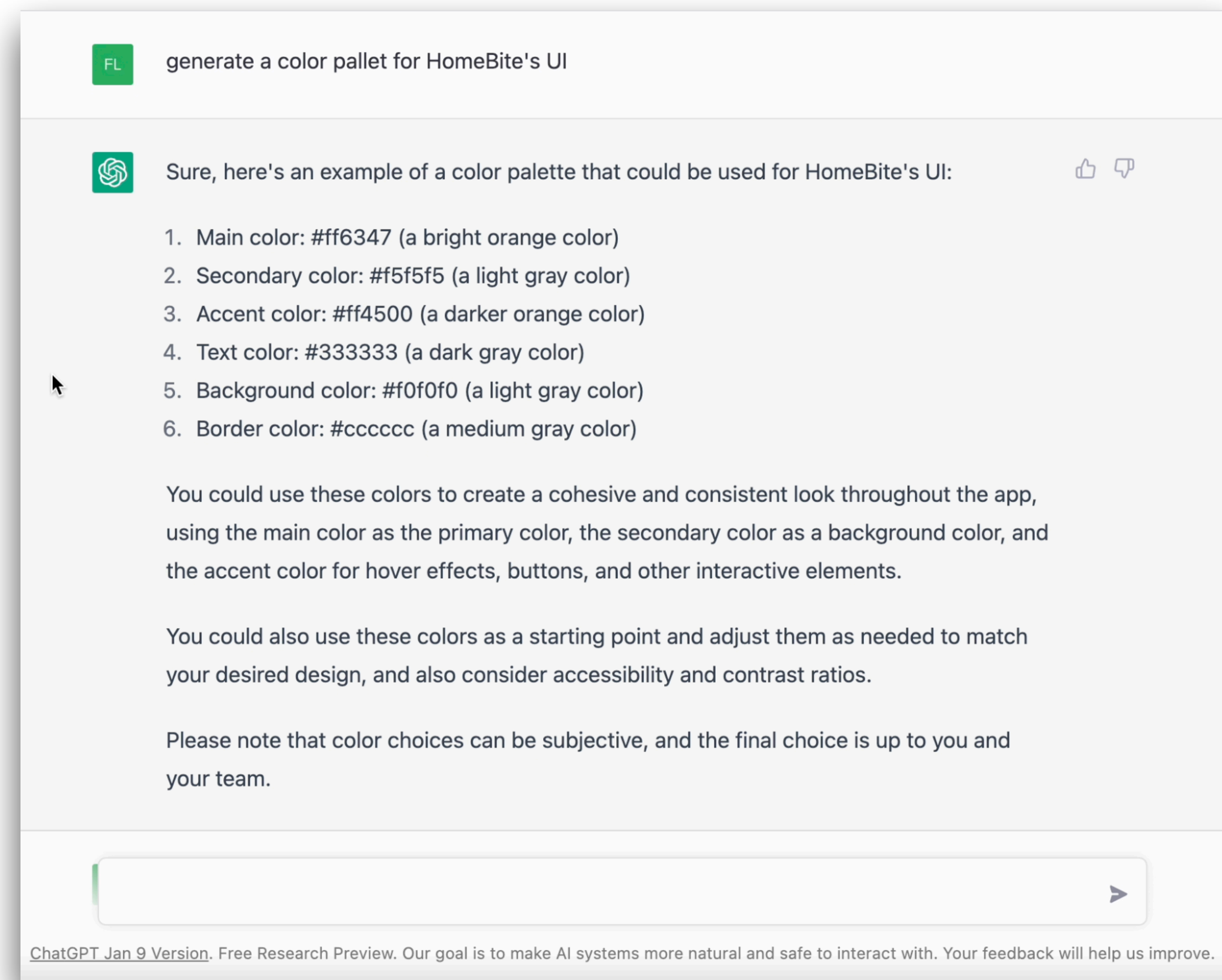
Solve math problems

Create articles, blog posts and quizzes for websites

Reword existing content for a different medium, such as a presentation transcript for a blog post

# Why should we care about language models?

- Foundational models in NLP today
  - Internal representations of models is learned exclusively from language models
  - Direct applications of LLMs are growing rapidly
- Classical importance
- Ubiquitous, so everyone has an opinion!
- Immense societal implications



# In your own words...

ethical implications of these technologies

theoretical details of large language models... looking forward to learn more about large language model architectures... hope to apply what I learn in this class to my own research in human genetics

dive deep on the math/algorithms behind how it works, so that I may engage in NLP related research and projects

challenge myself as well as explore my interest in NLP

learning about the theory behind it, I could also better understand its limitations, which is crucial when technology like ChatGPT is incorporating into all aspect of our life.

class that allows me to explore this overlap [between] ML and linguistics

risks from AI and AI safety issues

with each advance in AI there is a reflection cast onto our own intelligence and humanity

interested in artificial intelligence and human-machine "dialogue"

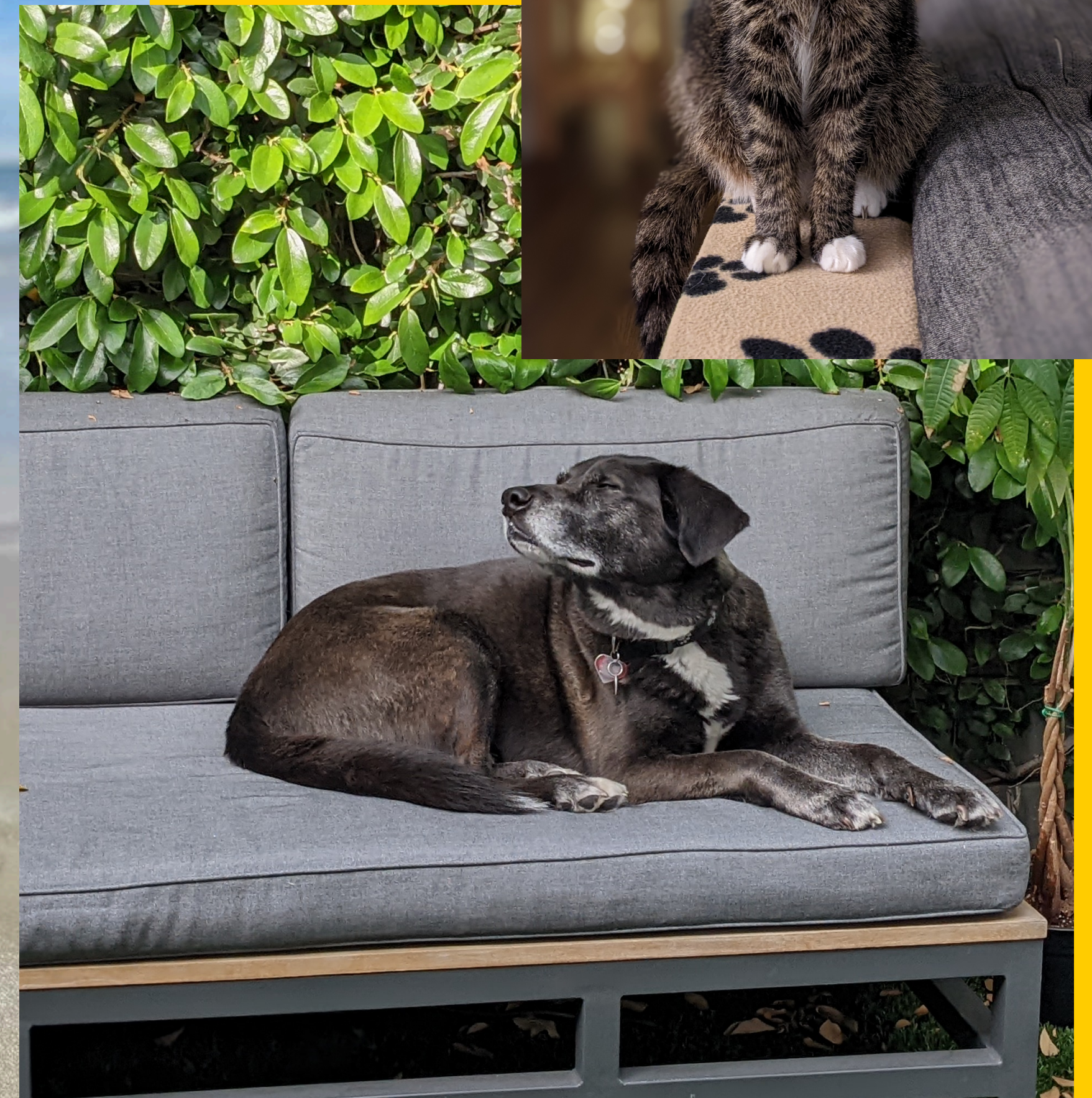
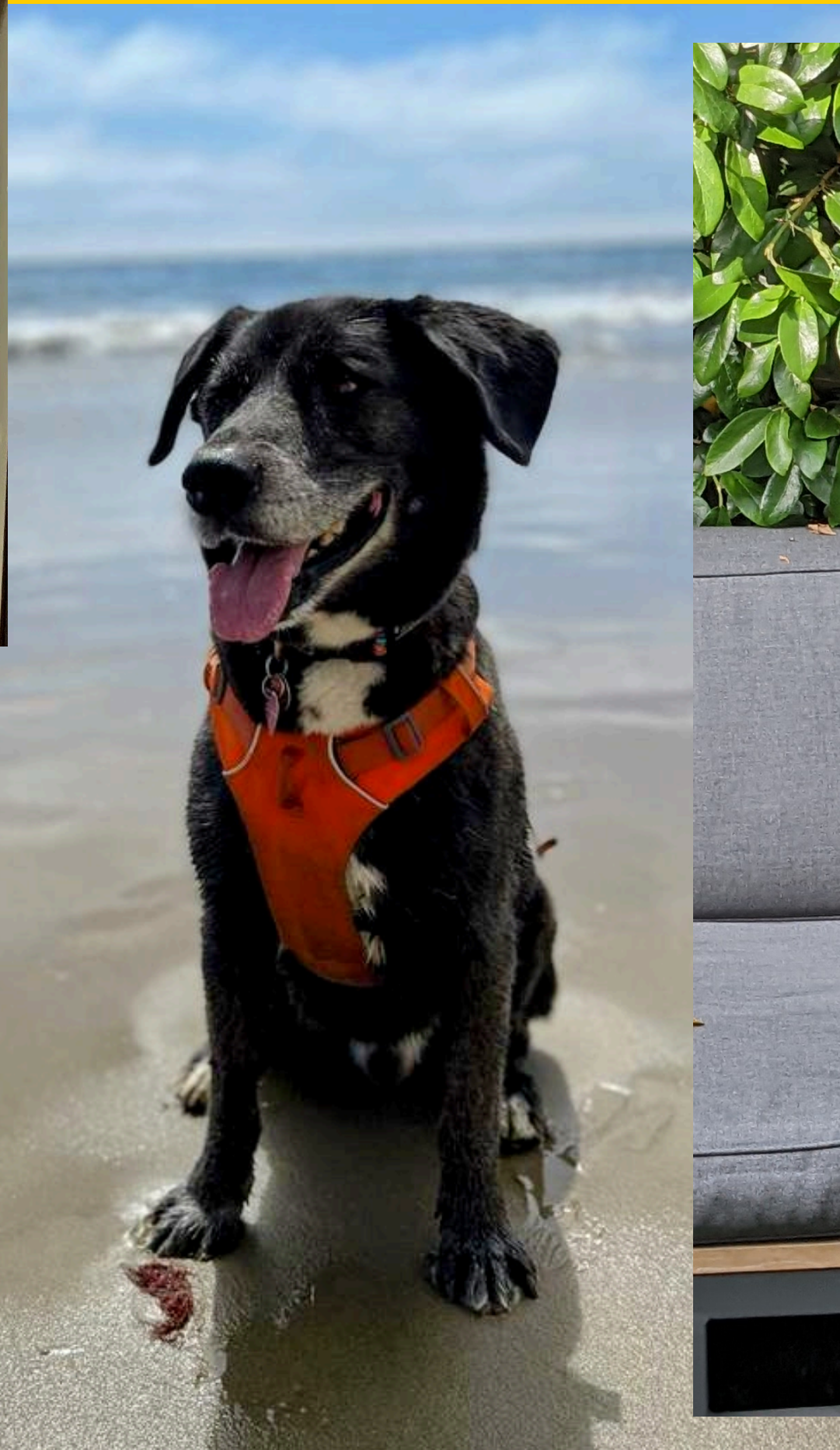
help with code refactoring/optimization in embedded systems.

exploring and assessing gender and racial bias in letters of recommendation

understanding some of the important concepts behind ChatGPT

# Class Introductions

Name, How long at USC?, 1 fun fact



# Why does this work?

## GPT Takes the Bar Exam

December 29, 2022

Michael Bommarito II<sup>1,2,3</sup>, Daniel Martin Katz<sup>1,2,3,\*</sup>

**1** Illinois Tech - Chicago Kent College of Law (Chicago, IL USA)  
**2** Bucerius Law School (Hamburg, Germany)  
**3** CodeX - The Stanford Center for Legal Informatics (Stanford, CA USA)

\* Corresponding Author: dkatz3@kentlaw.iit.edu

**Abstract**

Nearly all jurisdictions in the United States require a professional license exam, commonly referred to as "the Bar Exam," as a precondition for law practice. To even sit for the exam, most jurisdictions require that an applicant completes at least seven years of post-secondary education, including three years at an accredited law school. In addition, most test-takers also undergo weeks to months of further, exam-specific preparation. Despite this significant investment of time and capital, approximately one in five test-takers still score under the rate required to pass the exam on their first try. In the face of a complex task that requires such depth of knowledge, what, then, should we expect of the state of the art in "AI?" In this research, we document our experimental evaluation of the performance of OpenAI's TEXT-DAVINCI-003 model, often-referred to as GPT-3.5, on the multistate multiple choice (MBE) section of the exam. While we find no benefit in fine-tuning over GPT-3.5's zero-shot performance at the scale of our training data, we do find that hyperparameter optimization and prompt engineering positively impacted GPT-3.5's zero-shot performance. For best prompt and parameters, GPT-3.5 achieves a headline correct rate of 50.3% on a complete NCBE MBE practice exam, significantly in excess of the 25% baseline guessing rate, and performs at a passing rate for both Evidence and Torts. GPT-3.5's ranking of responses is also highly-correlated with correctness; its top two and top three choices are correct 71% and 88% of the time, respectively, indicating very strong non-entailment performance. While our ability to interpret these results is limited by nascent scientific understanding of LLMs and the proprietary nature of GPT, we believe that these results strongly suggest that an LLM will pass the MBE component of the Bar Exam in the near future.

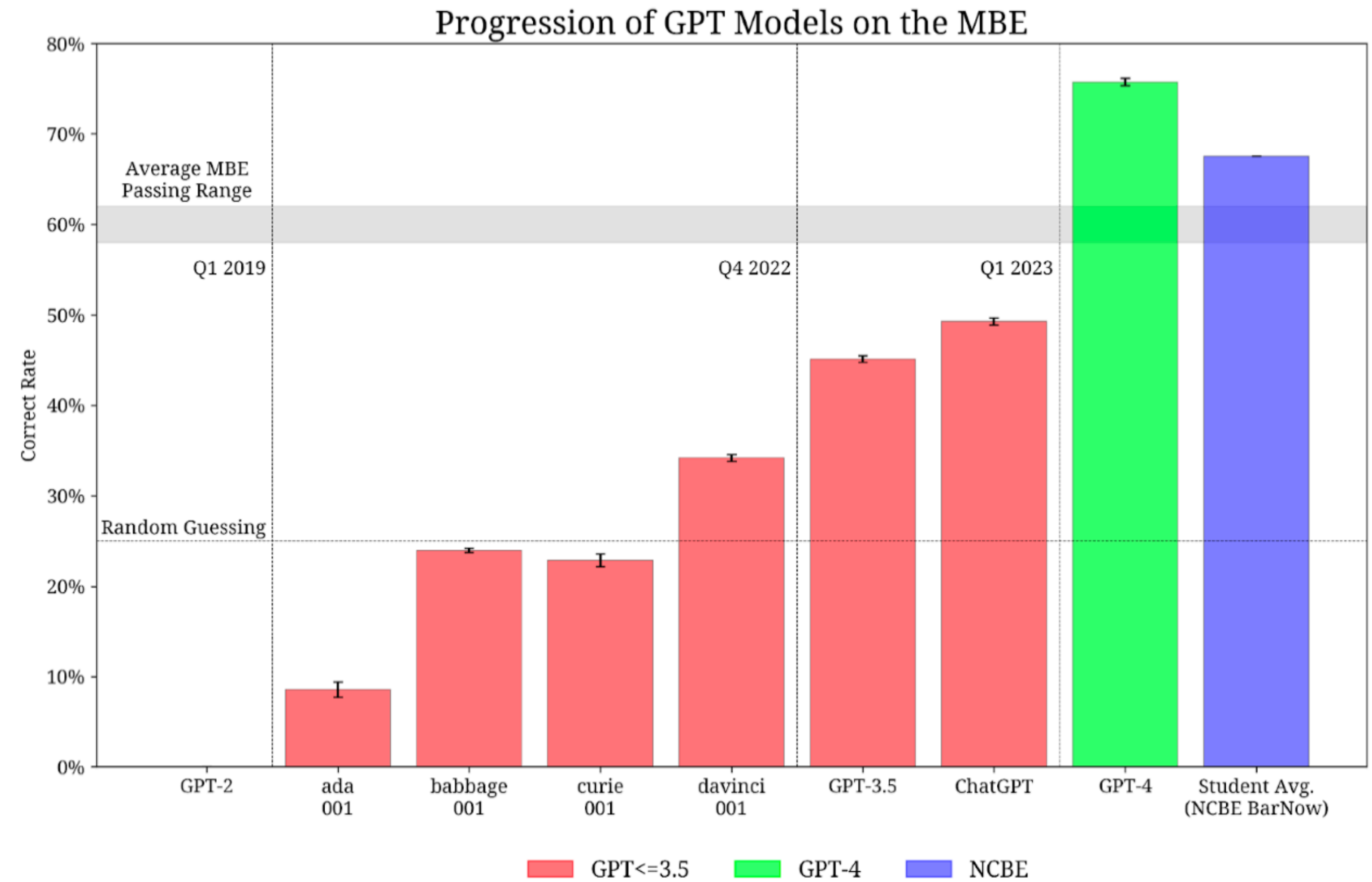
Fig 1. Summary of performance by question category for GPT-3.5 and NCBE-Reported Students

Question Category	NCBE Student Average	GPT-3.5	NCBE-GPT Spread
Civil Procedure	~65%	~65%	0%
Evidence	~65%	~65%	0%
Contracts	~65%	~65%	0%
Constitutional Law and Procedure	~65%	~65%	0%
Torts	~65%	~65%	0%
Criminal Law	~65%	~65%	0%

**Table 2. Summary of performance by question category for GPT-3.5 and NCBE-Reported Students**

	GPT	GPT Top 2	GPT Top 3	NCBE
Evidence	63%	84%	98%	65%
Torts	62%	72%	93%	71%
Civil Procedure	52%	63%	79%	59%
Constitutional Law	49%	67%	87%	72%
Real Property	45%	72%	85%	65%
Contracts	45%	77%	86%	70%
Criminal Law & Procedure	35%	62%	86%	71%
AVERAGE	50%	71%	88%	68%

Fig 2. Accuracy by Question Category for GPT and Average Test-Takers



The model has seen how to do a lot of tasks already when it was being built!



**Yann LeCun** @ylecun · 4d

AR-LLMs can pass the bar exam, medical licensing & MBA exams. But on the IIT entrance exams they perform badly on chemistry, horribly on physics, and terribly on math. They are good with rote learning & fluency but bad with building mental models & reasoning with them.

# But LLMs are far from perfect

## *The ChatGPT Lawyer Explains Himself*

In a cringe-inducing court hearing, a lawyer who relied on A.I. to craft a motion full of made-up case law said he “did not comprehend” that the chat bot could lead him astray.

Share full article 267



## *A.I.-Generated Content Discovered on News Sites, Content Farms and Product Reviews*

The findings in two new reports raise fresh concerns over how artificial intelligence may transform the misinformation landscape.

Hallucination leading to misinformation

## *An A.I. Hit of Fake ‘Drake’ and ‘The Weeknd’ Rattles the Music World*

A track like “Heart on My Sleeve,” which went viral before being taken down by streaming services this week, may be a novelty for now. But the legal and creative questions it raises are here to stay.

Give this article 215



## *This Tool Could Protect Artists From A.I.-Generated Art That Steals Their Style*

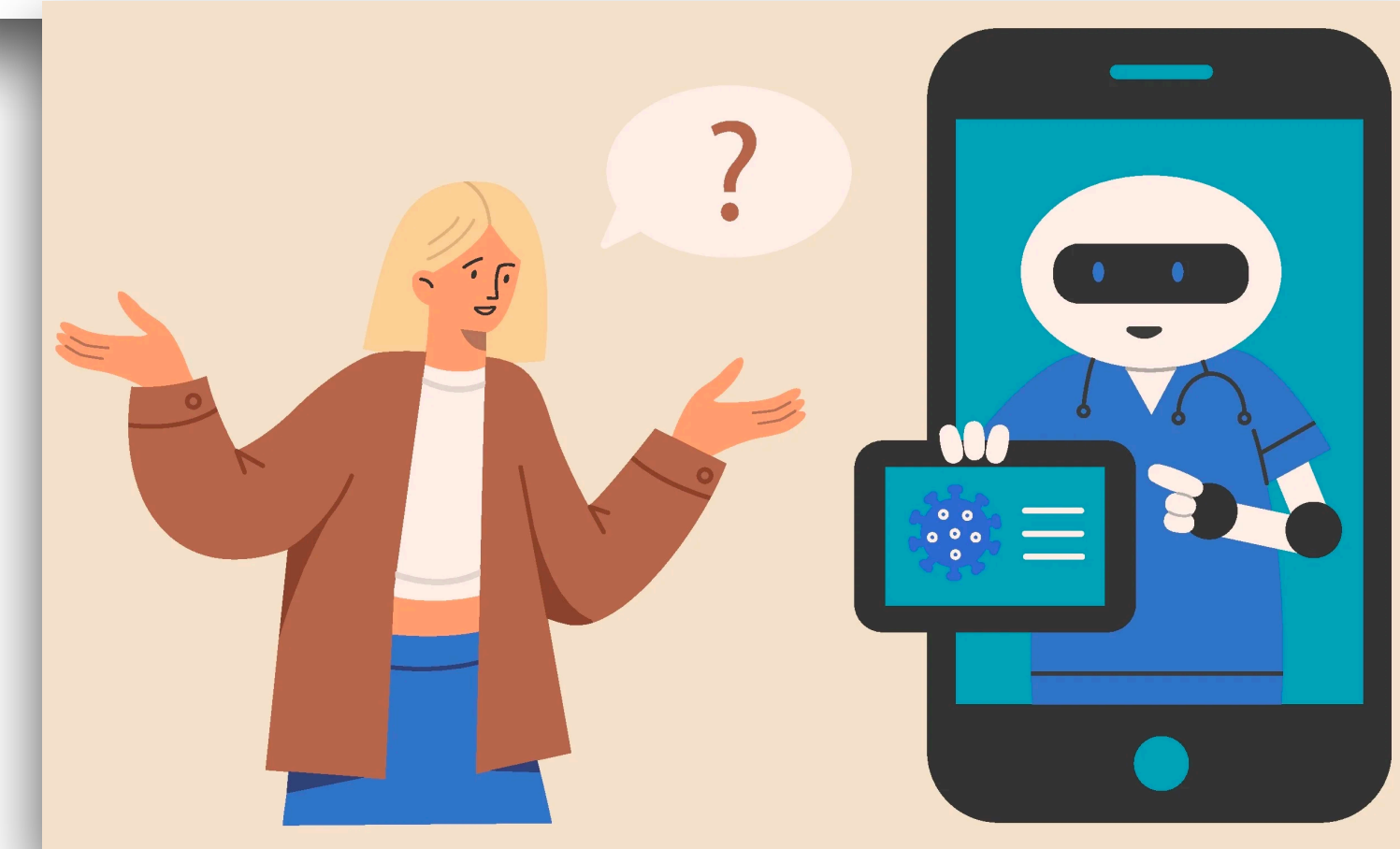
Artists want to be able to post their work online without the fear “of feeding this monster” that could replace them.

Privacy and Copyright Issues

Mar 8, 2023 - Technology

## Chatbot therapy, despite cautions, finds enthusiasts

Peter Allen Clark



Ethical Issues and Biases



# Class Syllabus

# The Pre-Neural Age of LMs (-2013)

- n-gram language models
  - The power of context
  - Smoothing
  - Evaluating LMs

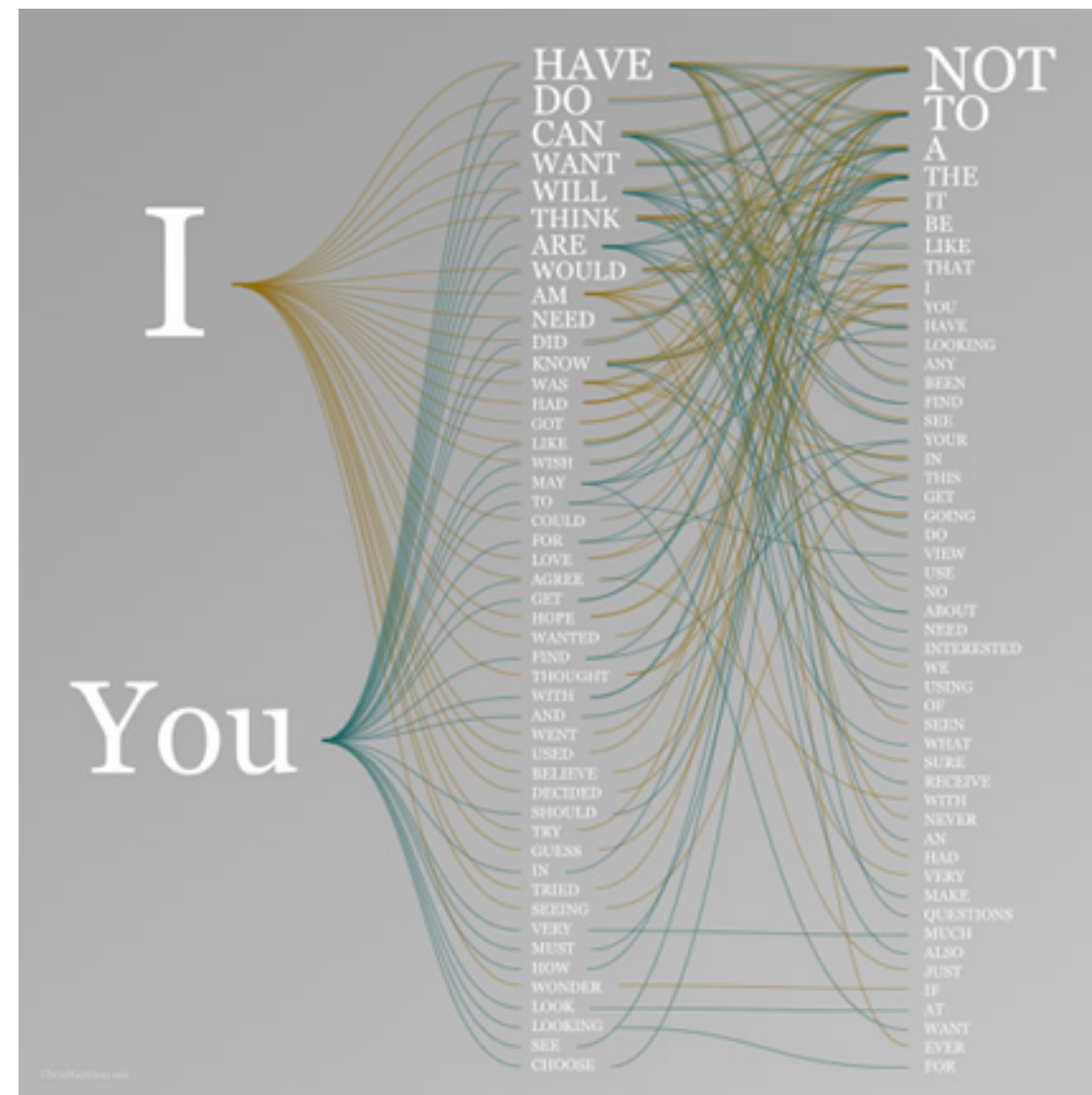
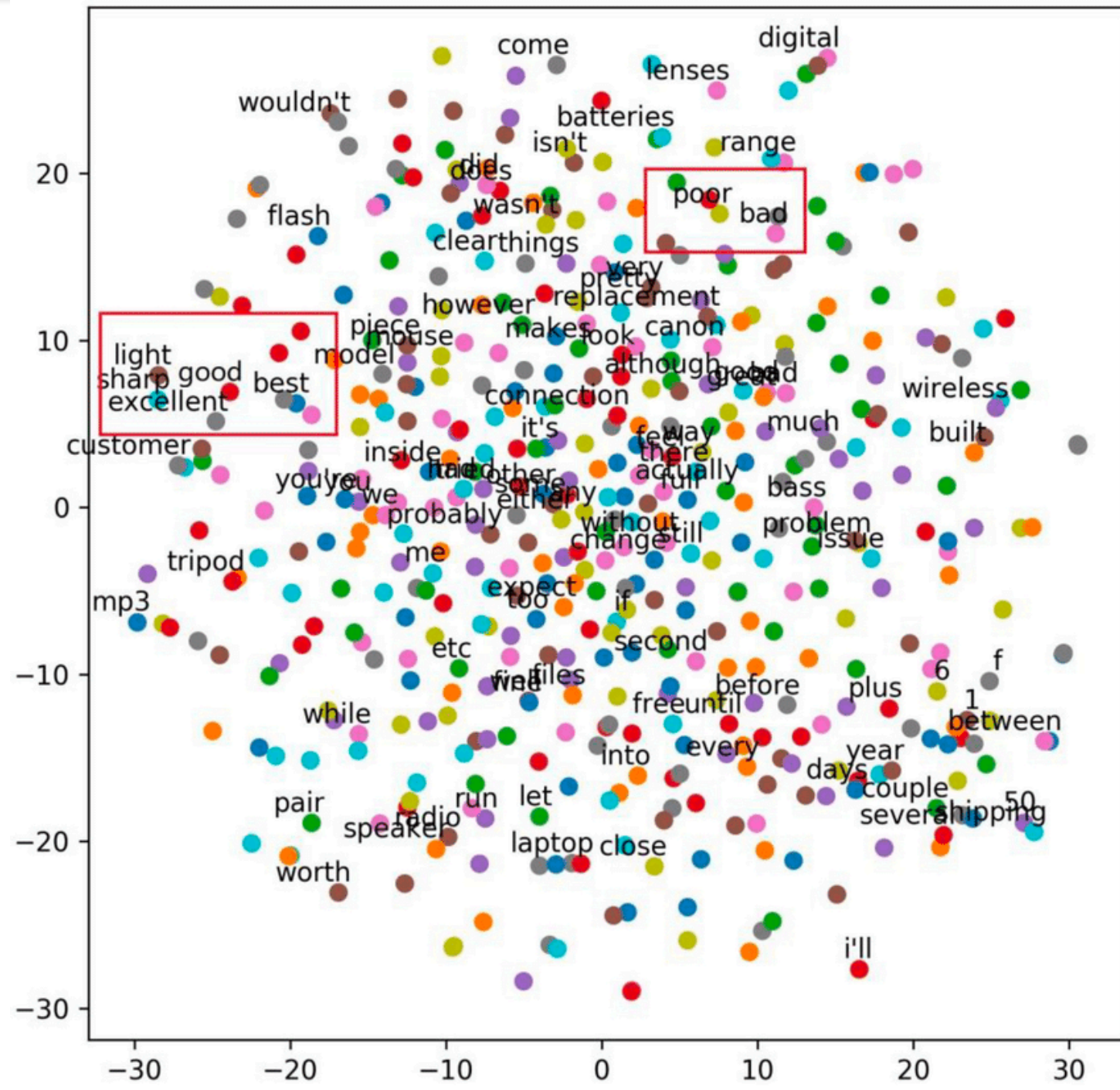
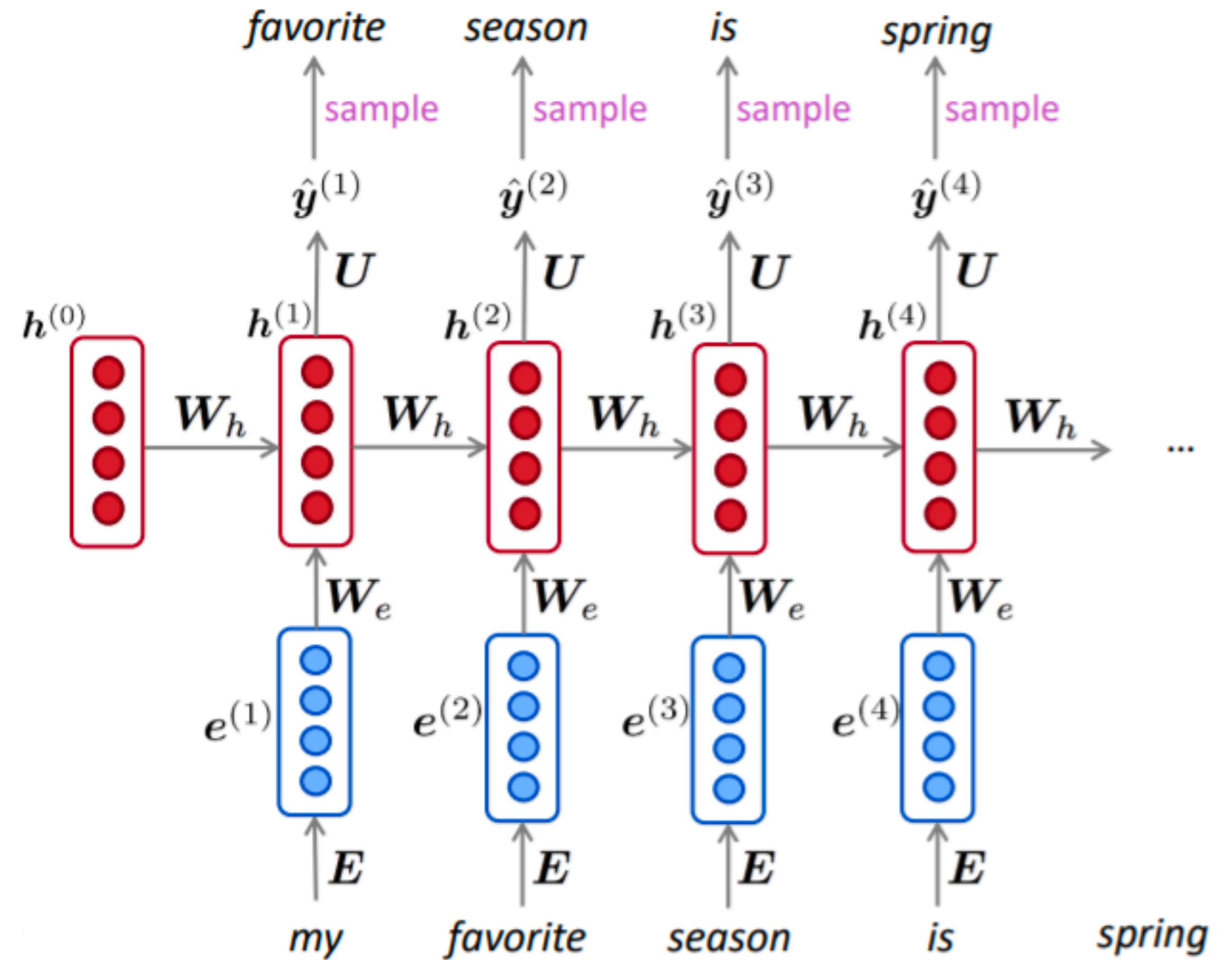


Image Courtesy: Chris Harrison

# Early Neural LMs (2013-2018)

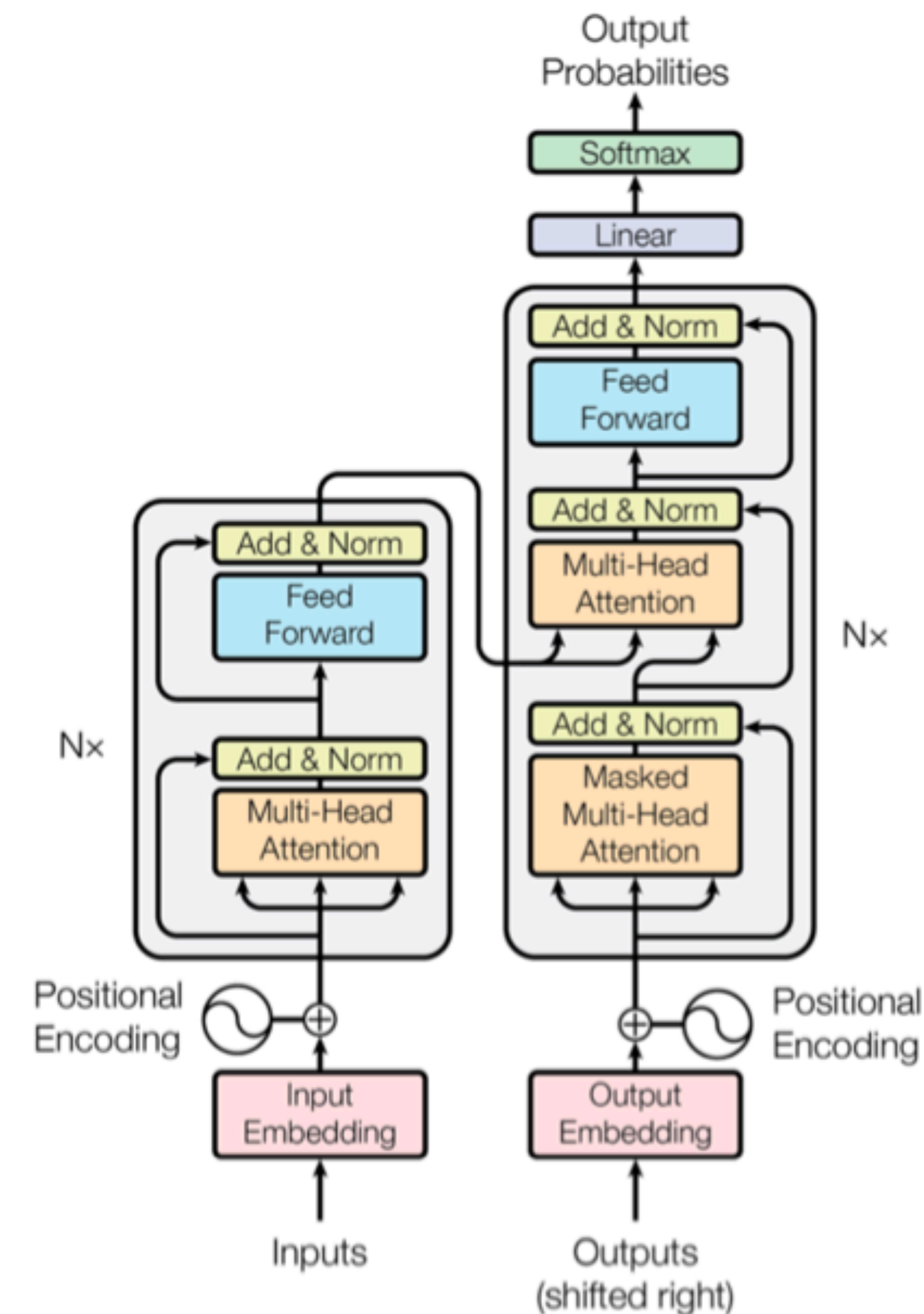
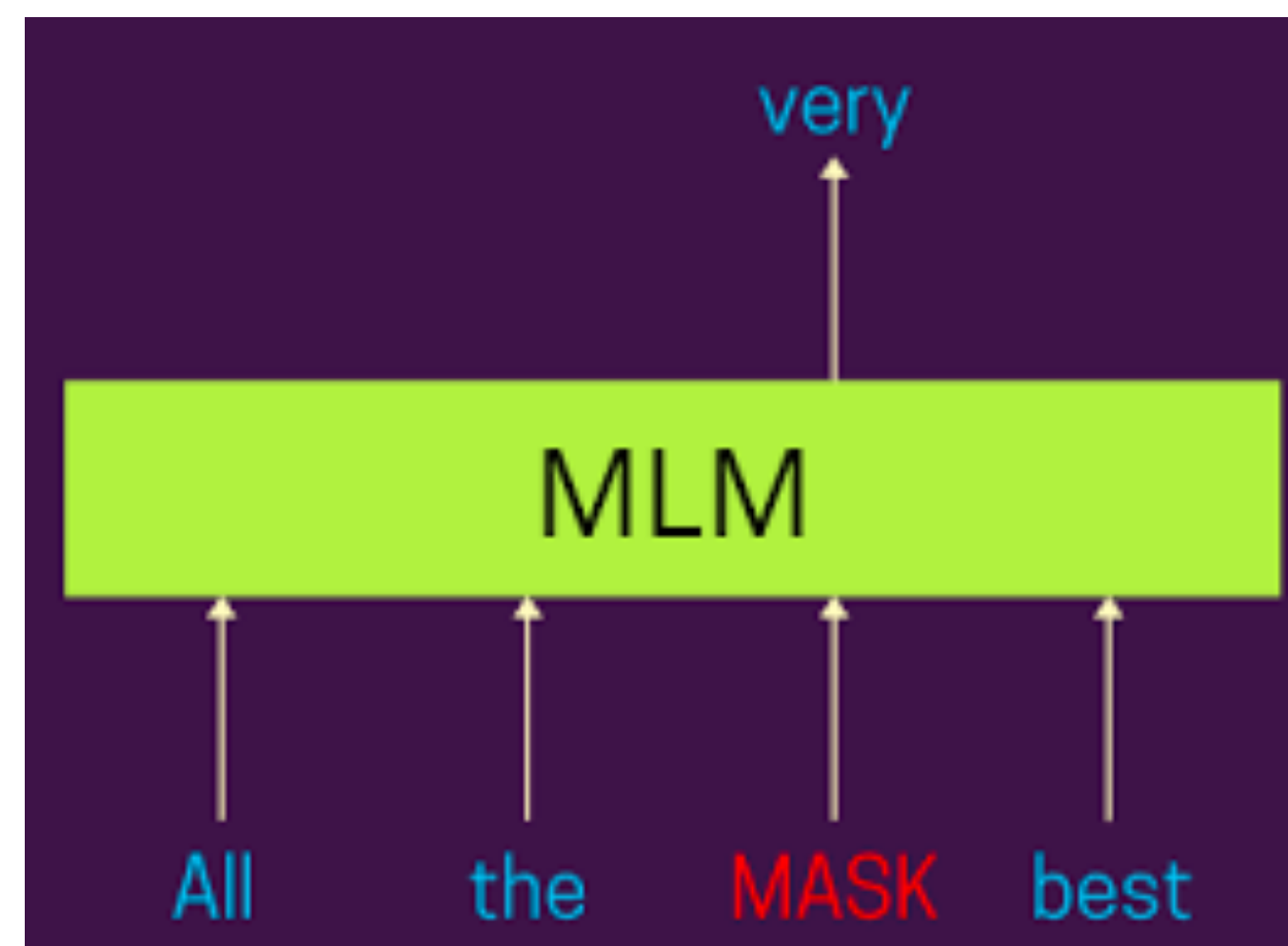


- (ML Basics) Logistic Regression
- Word Embeddings
- Feed-forward Networks
- Recurrent Neural Net Language Models
- Basics of Neural Nets: Backpropagation
- Encoder-Decoder Models



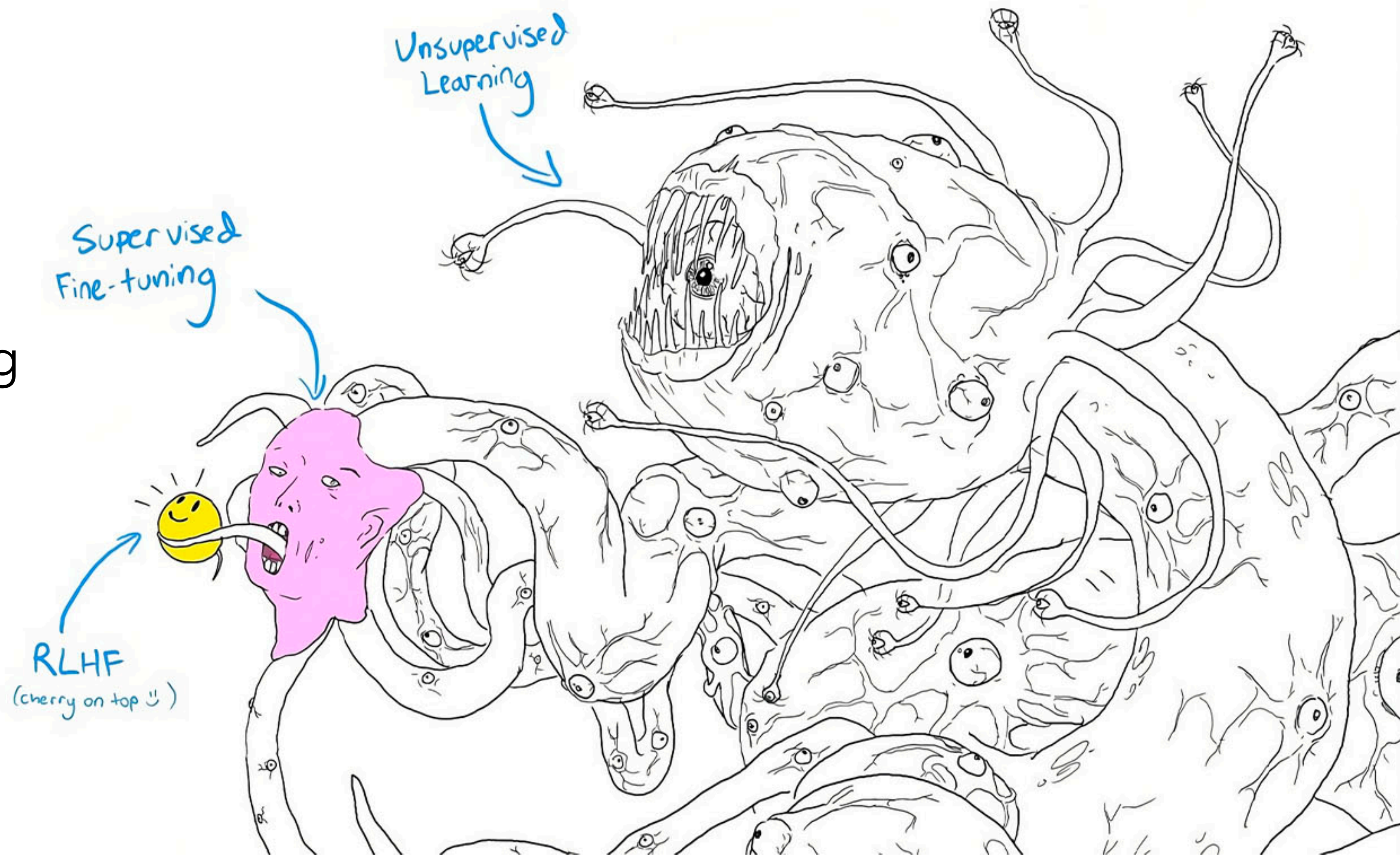
# Modern Neural Language Models (2018 - present)

- Attention
- Transformers
  - Self-Attention Networks
- Masked Language Models
- Decoder-only Models



# Large Language Models

- Pretraining and Finetuning
- Generation from LLMs
- Prompting and Instruction Tuning
- Preference Alignment using Human Feedback
- Outstanding Technical Issues (e.g., hallucinations) and Socio-Technical Issues (e.g., privacy)



Shoggoth with smiley face meme (image credit @anthrupad) - thanks to Justin Cho

# Concrete Outcomes

- Fundamentals of language modeling
- Build a language model - homework and / or project
- Learn the connections between this language model and models such as OpenAI's ChatGPT / GPT-4 models
- Current capabilities and outstanding issues with LLMs
- Exciting new problems



# What the class will NOT cover

- Detailed discussion on NLP classification tasks, e.g. question answering
- Classical NLP algorithms for structured prediction
  - e.g. logical semantics and lambda calculus
  - sequence tagging tasks
- In-depth discussion of linguistics

Discourse

Semantics

Syntax: Constituents

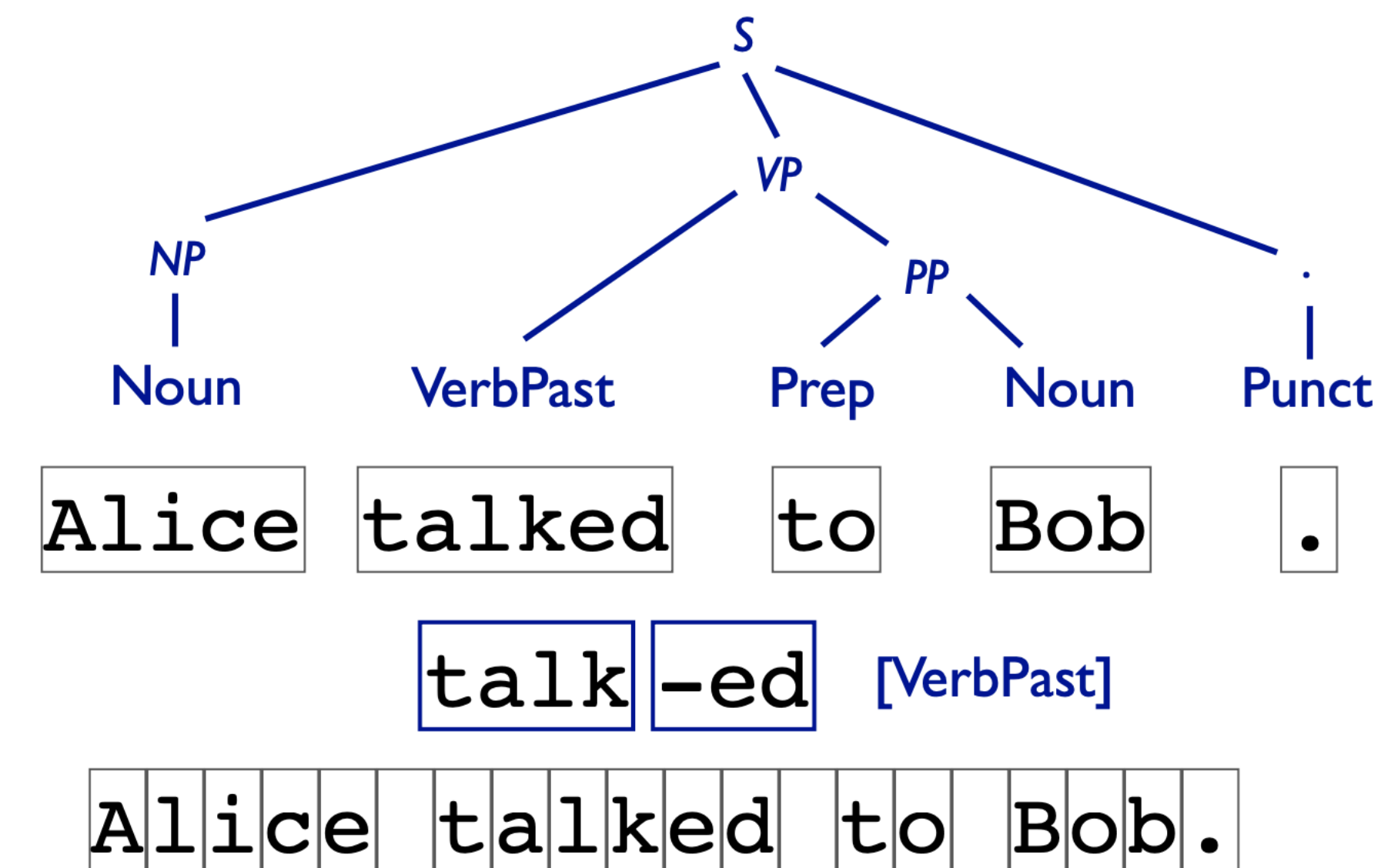
Syntax: Part of Speech

Words

Morphology

Characters

CommunicationEvent(e) SpeakerContext(s)  
 Agent(e, Alice) TemporalBefore(e, s)  
 Recipient(e, Bob)



# Class Logistics





# Staff and Website



Instructor: [Swabha Swayamdipta](#)

[swabhas@usc.edu](mailto:swabhas@usc.edu)

Office Hours: Tuesdays 9-10am; SAL 238



Teaching Assistant: [Sayan Ghosh](#)

[ghoshsay@usc.edu](mailto:ghoshsay@usc.edu)

Office Hours: Friday 10-11 am; Location TBD



Course Producer: [Xinyue Cui](#)

[xinyuecu@usc.edu](mailto:xinyuecu@usc.edu)

Lectures, Readings,  
Announcements



<https://swabhs.com/sp24-csci499-lm4nlp/>

# Pre-Requisites and D-Clearance

Students are required to have taken

- CSCI-270 Introduction to Algorithms and Theory of Computing, and
- CSCI-360 Introduction to AI / CSCI-467 Introduction to Machine Learning / equivalent experience

Fluency with python programming is recommended!

Please email me for special circumstances or specific clarifications

# Deliverables

- Homeworks - 40%
- Class Project - 40%
- Quizzes and Class Participation - 20%
- 6 total late days for homeworks and projects
  - No fractions
  - Max 3 per assignment / project deliverable
  - No late days for quizzes
- All work due by 11:59pm PT

# Homeworks

- 4 homework assignments - 10% each
  - Mix of programming and essay-type questions
- Topics will be based on modules:
  - n-gram language models (releasing soon!)
  - Word embeddings + RNN-LMs
  - Attention in Transformers
  - Advanced Topics in LLMs
- 10-14 days per homework



**Subject to change**

# Quizzes and Participation

20% of the grade: quizzes (15%) and participation (5%)

Quizzes will be in class and are designed to encourage you to do your readings:

- Released in class, to be submitted in class (allowed only for class attendees)
- Announced in the previous class / week

Class participation - This class is meant to be a discussion-oriented intro to research:

- Attend class - do inform of extenuating circumstances that prevent you from attending
- Ask questions in class, **and** volunteer to answer my questions!

# Class Project

- Styled like a research paper
- Teams of 2-3 students
  - We expect to see ~9 groups, based on enrollment
- We're working on getting the class CARC access
- Grading:
  - 5%: **project pitch** (3 mins) +
  - 5%: **project proposal** (1 page) +
  - 10%: **progress report** (3 pages) +
  - 10%: **presentation** of main findings (20 mins) +
  - 10%: **final report** (6-8 pages)
- 1 page project proposal **due by 2/5**

Jan 22:	<b>PROJECT PITCHES</b>	
Jan 24:	Logistic Regression	
Jan 29:	<b>Logistic Regression II</b>	HW1 Due 1/26
Jan 31:	Word Embeddings I	
Feb 5:	<b>Word Embeddings II</b>	<b>PROPOSAL DUE 2/5</b> HW2 Released
Feb 7:	Feedforward Neural Nets and Backprop	
Feb 12:	Recurrent Neural Network LMs	

## Modern Neural Language Models

Feb 14:	Sequence-To-Sequence and Attention	
<del>Feb 19:</del>	No Class	<b>PRESIDENT'S DAY</b>
Feb 21:	<b>Transformers - Building Blocks</b>	HW2 Due
Feb 26:	Transformers - Building Blocks II	
Feb 28:	<b>PROJECT DISCUSSIONS</b>	<b>PROGRESS REPORT DUE 3/1</b>

Apr 8: **PROJECT DISCUSSIONS**

## Outro and Project Presentations

Apr 10:	<b>Putting it all together</b>	No Additional Readings
Apr 15:	<b>PROJECT PRESENTATIONS</b>	
Apr 17:	<b>PROJECT PRESENTATIONS</b>	
Apr 22:	<b>PROJECT PRESENTATIONS</b>	
Apr 24:	<b>PROJECT PRESENTATIONS</b>	
<del>Apr 29:</del>	No Class	<b>STUDY WEEK</b>
May 1:	<b>PROJECT FINAL REPORT</b>	

# Class Project Topics - Examples

Detoxifying Language Model with Context Distillation

Haiku Generation with Large Language Models

Legal-SBERT: Creating Embeddings for the Legal Domain and Generating Summaries

Prompting for Diverse Responses: Making Large Language Models More Truthful

Forage: ML Generated Recipes

Learning the Language of Wine

Machine Translation from Inuktitut to

Creativity in choosing new and interesting problems often get rewarded!

Office Hours for Project Feedback

Relationships in Social

Commentary

Relationship Attribution with Limited Text

Was it Written?

See more: [Stanford CS224n Projects](#)

See more: [Stanford CS229 Machine Learning](#)

# Fall 23 CSCI 499 Class Projects

- Math Solver
- Enhancing Model Performance through Explanation-Guided Few-Shot Learning
- Music Score Generation
- Enhancing Debugging Skills of LLMs with Prompt Engineering
- Code Switch Language Model
- RhymeGPT
- BiasedCuisineNLP
- MotivationGPT
- Anti-ZeroGPT
- Conflict Resolution in IR





# Academic Conduct

**Plagiarism** – presenting someone else’s ideas as your own, either verbatim or recast in your own words – is a serious academic offense with serious consequences. Please familiarize yourself with the discussion of plagiarism in SCampus in [Part B, Section 11, “Behavior Violating University Standards”](#). Other forms of academic dishonesty are equally unacceptable. See additional information in SCampus and university policies on Research and Scholarship Misconduct.

# Textbooks

- **Jurafsky and Martin. "Speech and Language Processing." 3rd Ed.** This textbook contains chapters on the fundamentals of natural language processing.
- **Eisenstein. "Natural Language Processing."** This textbook contains an overview of machine learning approaches for NLP.
- **Goldberg. "Neural Network Methods for Natural Language Processing."** This textbook provides a deep learning perspective towards NLP.

# Anonymous Questions and Comments

- An anonymous feedback form will be made available through Piazza
- Feedback is welcome, but please be reasonable :) We will try our best to address the most important issues
- Piazza can be used to ask course / homework-related clarifying questions as well

The logo for Piazza, featuring the word "piazza" in a lowercase, blue, sans-serif font.

Welcome

# CSCI499 Spring 2024: Language Models in NLP

🌸 Spring 2024

🕒 Mon / Wed 4:00 - 5:50p

📍 KAP 134

- TODOs for you
  - Start brainstorming project ideas and preparing a pitch
- Next Class
  - n-gram Language Models

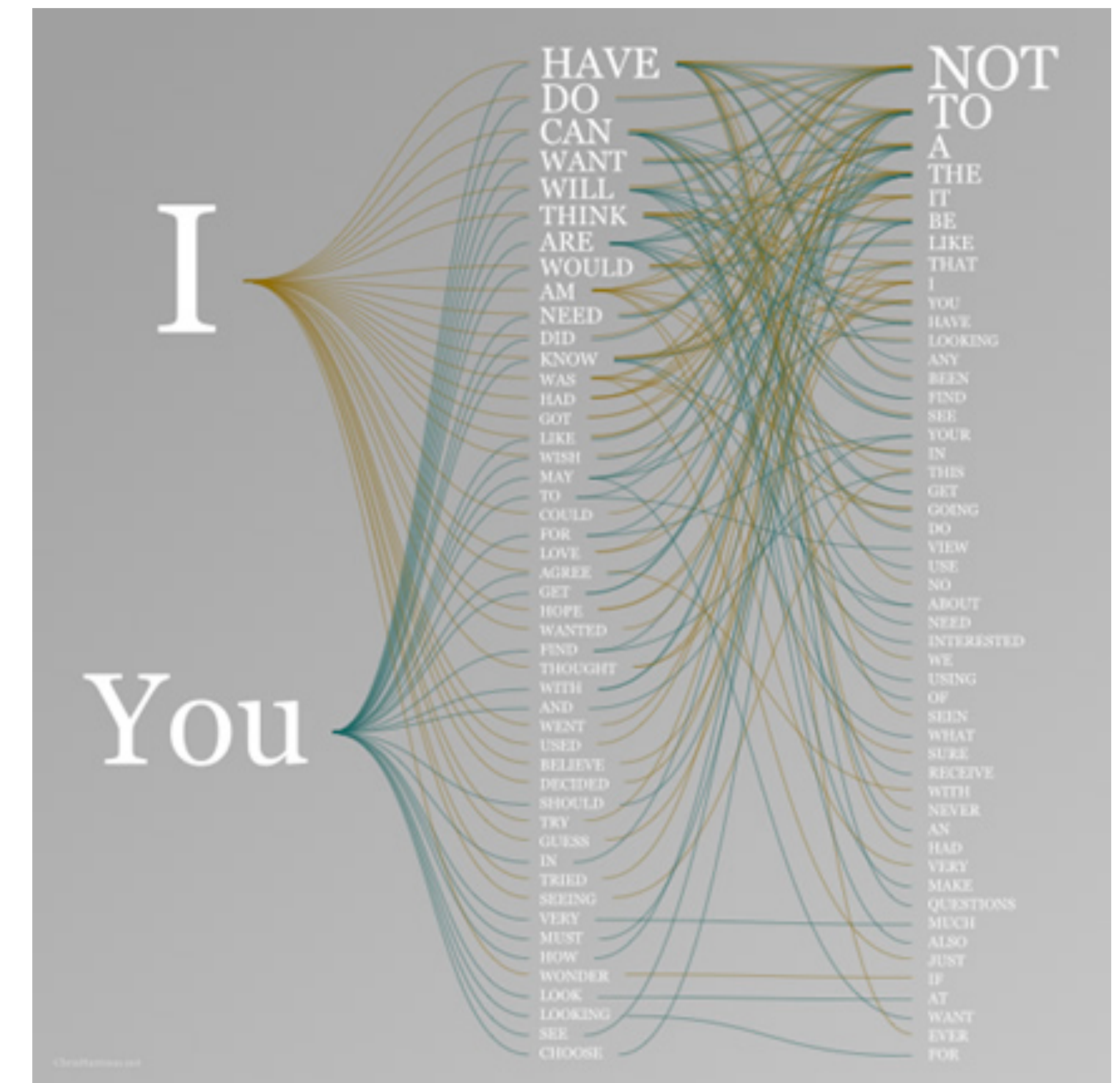


Image Courtesy: Chris Harrison