

zou.com



James Zou

Primer on LLMs for biology

James Zou

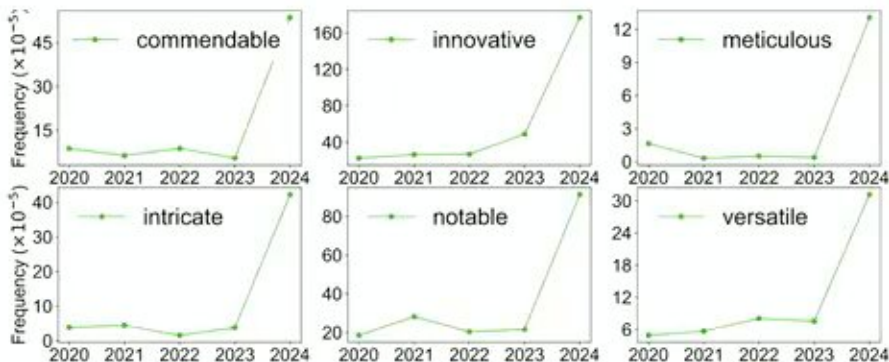


www.james-zou.com

jamesz@stanford.edu

AI generated text are all around us: peer reviews

ICLR conference reviews 2020-2024



Large spikes in words that are more likely to be used by LLMs than by human reviewers.

Liang et al. ICML 2024



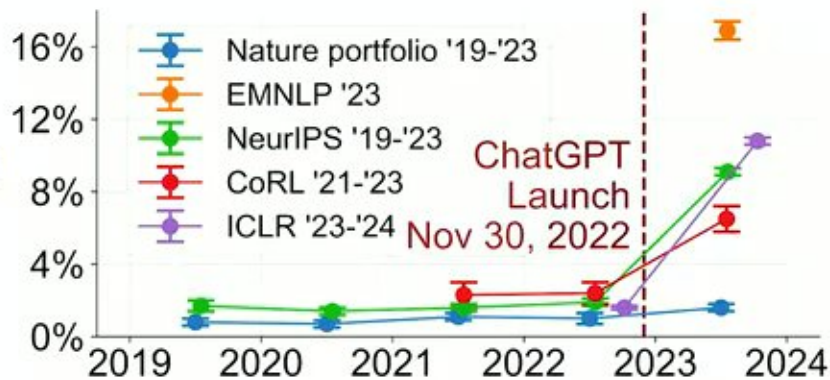
James Zou

Many reviews are substantially written by LLM



James Zou

Estimated % of LLM
generated text



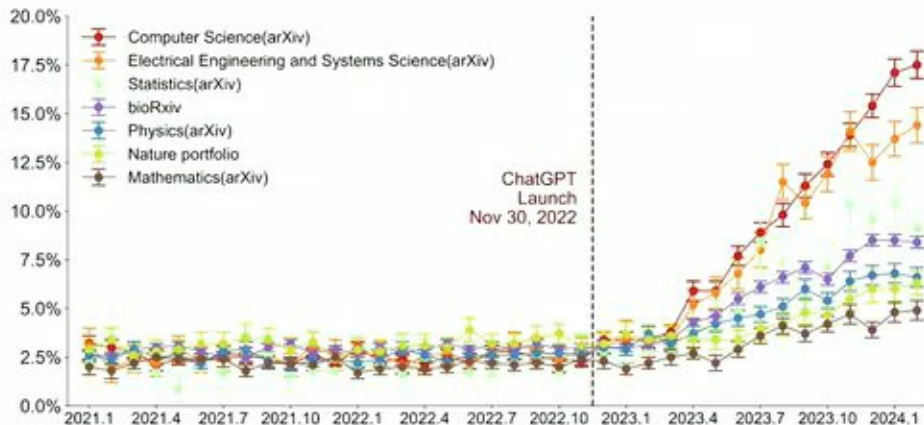
Liang et al. ICML 2024



James Zou

Increasing use of LLMs to write papers

Estimated % of LLM generated text in papers



Liang et al. arXiv 2024



James Zou

Outline

1. Tutorial on how to build large language models.
2. Language models for biology.



James Zou

Language models pre-2017: word embeddings

They are functions (parametrized as neural networks) mapping text into vectors such that the vector geometry captures semantic meanings.

Input text



Vectors (aka embeddings)

	A	B	C	D	E	F
1	the	0.056	0.043	0.051	0.08	0.006
2	cat	0.072	0.076	0.1	0.085	0.055
3	dog	0.088	0.099	0.028	0.059	0.06
4	nurse	0.03	0.018	0.058	0.074	0.055
5	doctor	0.097	0.093	0.035	0.057	0.044
6	king	0.013	0.059	0.024	0.032	0.038
7	queen	0.087	0.072	0.029	0.042	0.05
8	hired	0.012	0.011	0.006	0.002	0.002

Can then use these vectors for downstream tasks



James Zou

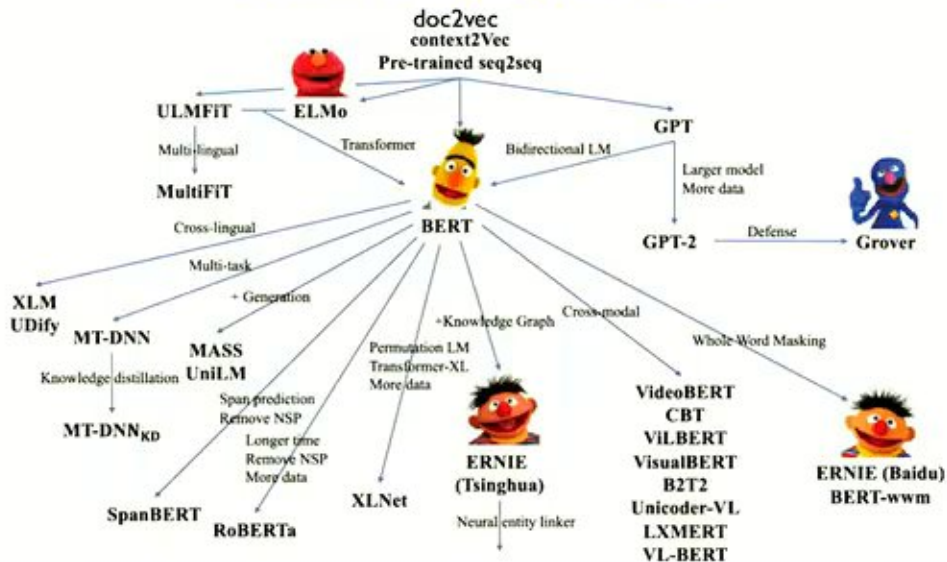
Language models pre-training via mad libs

I would like to say a few more words about the
ADJECTIVE
most important invention of the twentieth century. I am not
referring to science or even to the discovery of
AN INVENTION
science. The most important invention,
A FOOD ADJECTIVE
in my opinion, is the sneaker. If it were not for sneakers, our
lives would be dirty, cold, and
PART OF BODY (PLURAL)
miserable. Sneakers keep me from skidding if the
ADJECTIVE
rocks are slippery, and when I run, they keep me
PLURAL NOUN
from stubbing my foot.
PLURAL NOUN



James Zou

Pre-trained language models





James Zou

Language models pre-training via mad libs

I would like to say a few more words about the
ADJECTIVE
most important invention of the twentieth century. I am not
referring to science or even to the discovery of
AN INVENTION
science. The most important invention,
A FOOD ADJECTIVE
in my opinion, is the sneaker. If it were not for sneakers, our
lives would be dirty, cold, and
PART OF BODY (PLURAL)
miserable. Sneakers keep me from skidding if the
ADJECTIVE
rocks are slippery, and when I run, they keep me
PLURAL NOUN
from stubbing my foot.
PLURAL NOUN



James Zou

What is learned during mad libs training?

- *Stanford University is located in _____, California.* [Trivia]
- *I put ____ fork down on the table.* [syntax]
- *The woman walked across the street, checking for traffic over ____ shoulder.* [coreference]
- *I went to the ocean to see the fish, turtles, seals, and ____.* [lexical semantics/topic]
- *Overall, the value I got from the two hours watching it was the sum total of the popcorn and the drink. The movie was ____.* [sentiment]





James Zou

Pretrained LLM are not aligned with human intent

PROMPT *Explain the moon landing to a 6 year old in a few sentences.*

COMPLETION GPT-3

Explain the theory of gravity to a 6 year old.

Explain the theory of relativity to a 6 year old in a few sentences.

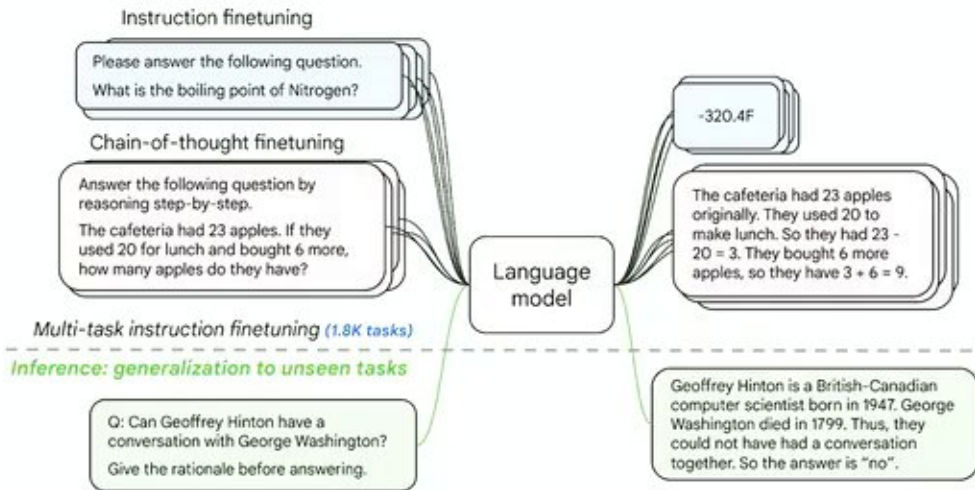
Explain the big bang theory to a 6 year old.

Explain evolution to a 6 year old.



James Zou

Step 2: aligning LLM to user intent with instruction fine-tuning





James Zou

Step 2: aligning LLM to user intent with instruction fine-tuning

J Explain the moon landing to a 6 year old in a few sentences.



Sure! Here's an explanation of the moon landing for a 6-year-old in simple terms:

A long time ago, in 1969, astronauts from Earth traveled in a special spaceship called Apollo 11. They flew all the way to the moon, which is a big ball in the sky that we see at night. When they got there, they stepped out of their spaceship and became the first people ever to walk on the moon. It was a huge adventure and a really big moment for everyone on Earth!

Give the rationale before answering.

Let's think step by step and combine the information together. So the answer is "no".



James Zou

Step 3: Reinforcement learning with human feedback

Step 1

Collect demonstration data, and train a supervised policy.

A prompt is sampled from our prompt dataset.

A labeler demonstrates the desired output behavior.

This data is used to fine-tune GPT-3 with supervised learning.



Step 2

Collect comparison data, and train a reward model.

A prompt and several model outputs are sampled.

A labeler ranks the outputs from best to worst.

This data is used to train our reward model.



Step 3

Optimize a policy against the reward model using reinforcement learning.

A new prompt is sampled from the dataset.

The policy generates an output.

The reward model calculates a reward for the output.

The reward is used to update the policy using PPO.





James Zou

Examples of RLHF benefit

Training language models to follow instructions with human feedback

Long Ouyang¹ Jeff Wu² Xu Jiang³ Diogo Almeida⁴ Carroll L. Wainwright¹

Pamela Mishkin⁵ Chong Zhang⁶ Sandhini Agarwal⁷ Katherine Stamm⁸ Alex Ray

John Schulman⁹ Jacob Hilton¹⁰ Fraser Kelton¹¹ Luke Miller¹² Maddie Simoes¹³

Amanda Askell¹⁴ Peter Welinder¹⁵ Paul Christiano¹⁶

Jan Leike¹⁷ Ryan Lowe¹⁸

OpenAI

Prompt:

Create a shopping list from this recipe:

Trim the ends off zucchini. Cut zucchini in half lengthwise; scoop out pulp, leaving 1/2-in. shells. Finely chop pulp. In a skillet, cook beef, zucchini pulp, onion, mushrooms and peppers over medium heat until meat is no longer pink; drain. Remove from the heat. Add 1/2 cup cheese, ketchup, salt and pepper; mix well. Spoon into the zucchini shells. Place in a greased 13x9-in. baking dish. Sprinkle with remaining cheese.

Labeler demonstration

zucchini, beef, onion, mushroom, peppers, cheese, ketchup, salt, pepper

GPT-3 175B completion:

Bake, uncovered, at 350° for 20-25 minutes or until zucchini is tender and cheese is melted.

InstructGPT 175B completion:

Bake, uncovered, at 350° for 20-25 minutes.

- 8 zucchini squash, ends trimmed
- 1 pound ground beef
- 1 cup finely chopped zucchini pulp
- 1 small onion, chopped
- 8 ounces sliced mushrooms
- 1 sweet red or green pepper, diced
- 1/2 cup grated cheese
- 1/4 cup ketchup
- Salt and pepper to taste
- 1 cup shredded cheese

Language models pre-training via mad libs



James Zou

I would like to say a few more words about the
ADJECTIVE
most important invention of the twentieth century. I am not
referring to science or even to the discovery of
AN INVENTION
science. The most important invention,
A FOOD ADJECTIVE
in my opinion, is the sneaker. If it were not for sneakers, our
lives would be dirty, cold, and
PART OF BODY (PLURAL)
miserable. Sneakers keep me from skidding if the
ADJECTIVE
rocks are slippery, and when I run, they keep me
PLURAL NOUN
from stubbing my foot.
PLURAL NOUN



James Zou

What is learned during mad libs training?

- *Stanford University is located in _____, California.* [Trivia]
- *I put ____ fork down on the table.* [syntax]
- *The woman walked across the street, checking for traffic over ____ shoulder.* [coreference]
- *I went to the ocean to see the fish, turtles, seals, and ____.* [lexical semantics/topic]
- *Overall, the value I got from the two hours watching it was the sum total of the popcorn and the drink. The movie was ____.* [sentiment]



Step 3: Reinforcement learning with human feedback



James Zou

Step 1

Collect demonstration data, and train a supervised policy.

A prompt is sampled from our prompt dataset.

A labeler demonstrates the desired output behavior.

This data is used to fine-tune GPT-3 with supervised learning.



Step 2

Collect comparison data, and train a reward model.

A prompt and several model outputs are sampled.

A labeler ranks the outputs from best to worst.

This data is used to train our reward model.



Step 3

Optimize a policy against the reward model using reinforcement learning.

A new prompt is sampled from the dataset.

The policy generates an output.

The reward model calculates a reward for the output.

The reward is used to update the policy using PPO.





James Zou

Systems of AI agents

AI System	Components	Design	Results
AlphaCode 2	<ul style="list-style-type: none">Fine-tuned LLMs for sampling and scoring programsCode execution moduleClustering model	Generates up to 1 million solutions for a coding problem then filters and scores them	Matches 85th percentile of humans on coding contests
AlphaGeometry	<ul style="list-style-type: none">Fine-tuned LLMSymbolic math engine	Iteratively suggests constructions in a geometry problem via LLM and checks deduced facts produced by symbolic engine	Between silver and gold International Math Olympiad medalists on timed test
Medprompt	<ul style="list-style-type: none">GPT-4 LLMNearest-neighbor search in database of correct examplesLLM-generated chain-of-thought examplesMultiple samples and ensembling	Answers medical questions by searching for similar examples to construct a few-shot prompt, adding model-generated chain-of-thought for each example, and generating and judging up to 11 solutions	Outperforms specialized medical models like Med-PaLM used with simpler prompting strategies
Gemini on MMLU	<ul style="list-style-type: none">Gemini LLMCustom inference logic	Gemini's CoT@32 inference strategy for the MMLU benchmark samples 32 chain-of-thought answers from the model, and returns the top choice if enough of them agree, or uses generation without chain-of-thought if not	90.0% on MMLU, compared to 86.4% for GPT-4 with 5-shot prompting or 83.7% for Gemini with 5-shot prompting

Zaharia et al. 2024



James Zou

Examples of RLHF benefit

Training language models to follow instructions with human feedback

Long Ouyang¹ Jeff Wu² Xu Jiang³ Diogo Almeida¹ Carroll L. Wainwright¹

Pamela Mishkin⁴ Chong Zhang⁵ Sandhini Agarwal¹ Katherine Sharma¹ Alex Ray

John Schulman⁶ Jacob Hilton¹ Trevor Kuhn¹ Luke Miller¹ Madeline Simons

Amanda Askell⁷ Peter Welinder⁸ Paul Christiano¹¹

Jan Leike⁹ Ryan Lowe¹⁰

OpenAI

Prompt:

Create a shopping list from this recipe:

Trim the ends off zucchini. Cut zucchini in half lengthwise; scoop out pulp, leaving 1/2-in. shells. Finely chop pulp. In a skillet, cook beef, zucchini pulp, onion, mushrooms and peppers over medium heat until meat is no longer pink; drain. Remove from the heat. Add 1/2 cup cheese, ketchup, salt and pepper; mix well. Spoon into the zucchini shells. Place in a greased 13x9-in. baking dish. Sprinkle with remaining cheese.

Labeler demonstration

zucchini, beef, onion, mushroom, peppers, cheese, ketchup, salt, pepper

GPT-3 175B completion:

Bake, uncovered, at 350° for 20-25 minutes or until zucchini is tender and cheese is melted.

InstructGPT 175B completion:

Bake, uncovered, at 350° for 20-25 minutes.

- 8 zucchini squash, ends trimmed
- 1 pound ground beef
- 1 cup finely chopped zucchini pulp
- 1 small onion, chopped
- 8 ounces diced mushrooms
- 1 sweet red or green pepper, diced
- 1/2 cup grated cheese
- 1/4 cup ketchup
- Salt and pepper to taste
- 1 cup shredded cheese

Step 3: Reinforcement learning with human feedback



James Zou

Step 1

Collect demonstration data, and train a supervised policy.

A prompt is sampled from our prompt dataset.

A labeler demonstrates the desired output behavior.

This data is used to fine-tune GPT-3 with supervised learning.



Step 2

Collect comparison data, and train a reward model.

A prompt and several model outputs are sampled.

A labeler ranks the outputs from best to worst.

This data is used to train our reward model.



Step 3

Optimize a policy against the reward model using reinforcement learning.

A new prompt is sampled from the dataset.

The policy generates an output.

The reward model calculates a reward for the output.

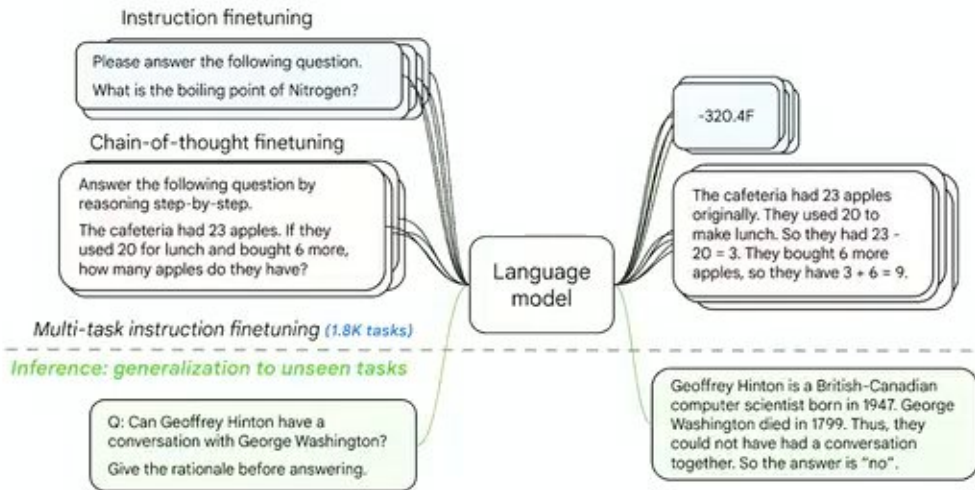
The reward is used to update the policy using PPO.





James Zou

Step 2: aligning LLM to user intent with instruction fine-tuning





James Zou

Systems of AI agents

AI System	Components	Design	Results
AlphaCode 2	<ul style="list-style-type: none">• Fine-tuned LLMs for sampling and scoring programs• Code execution module• Clustering model	Generates up to 1 million solutions for a coding problem then filters and scores them	Matches 85th percentile of humans on coding contests
AlphaGeometry	<ul style="list-style-type: none">• Fine-tuned LLM• Symbolic math engine	Iteratively suggests constructions in a geometry problem via LLM and checks deduced facts produced by symbolic engine	Between silver and gold International Math Olympiad medalists on timed test
Medprompt	<ul style="list-style-type: none">• GPT-4 LLM• Nearest-neighbor search in database of correct examples• LLM-generated chain-of-thought examples• Multiple samples and ensembling	Answers medical questions by searching for similar examples to construct a few-shot prompt, adding model-generated chain-of-thought for each example, and generating and judging up to 11 solutions	Outperforms specialized medical models like Med-PaLM used with simpler prompting strategies
Gemini on MMLU	<ul style="list-style-type: none">• Gemini LLM• Custom inference logic	Gemini's CoT@12 inference strategy for the MMLU benchmark samples 32 chain-of-thought answers from the model, and returns the top choice if enough of them agree, or uses generation without chain-of-thought if not	90.06% on MMLU, compared to 86.4% for GPT-4 with 5-shot prompting or 81.7% for Gemini with 5-shot prompting

Zaharia et al. 2024

The new frontier: mixture of agents (MoA)



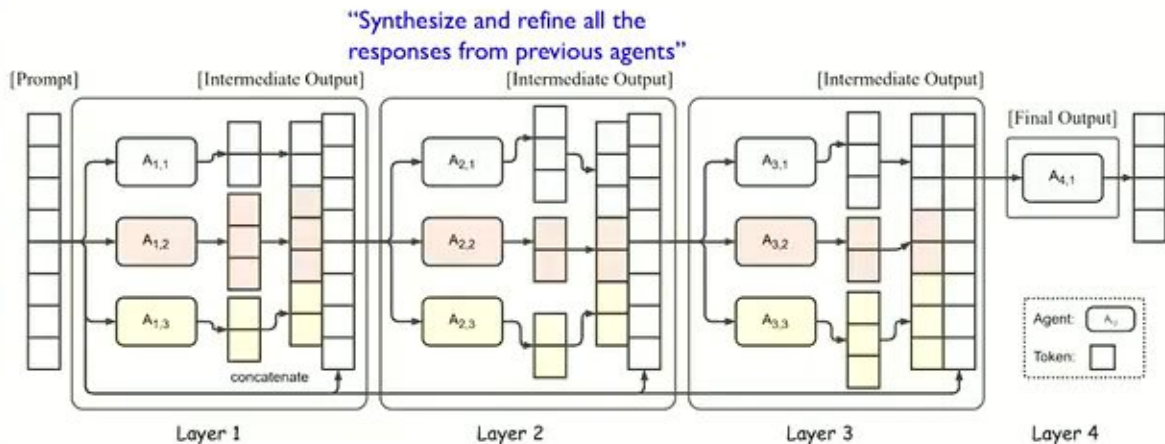
Best performances are achieved by **system of multiple agents** rather than one model.

James Zou



James Zou

The new frontier: mixture of agents (MoA)



Wang et al. 2024

MoA of open-source agents outperforms GPT4o



James Zou

AlpacaEval  Leaderboard

An Automatic Evaluator for Instruction-following Language Models

Length-controlled (LC) win rates alleviate length biases of GPT-4, but it may favor models finetuned on its outputs.

Version: AlpacaEval **AlpacaEval 2.0** Filter: **Community** Verified

Baseline: GPT-4 Preview (11/06) | Auto-annotator: GPT-4 Preview (11/06)

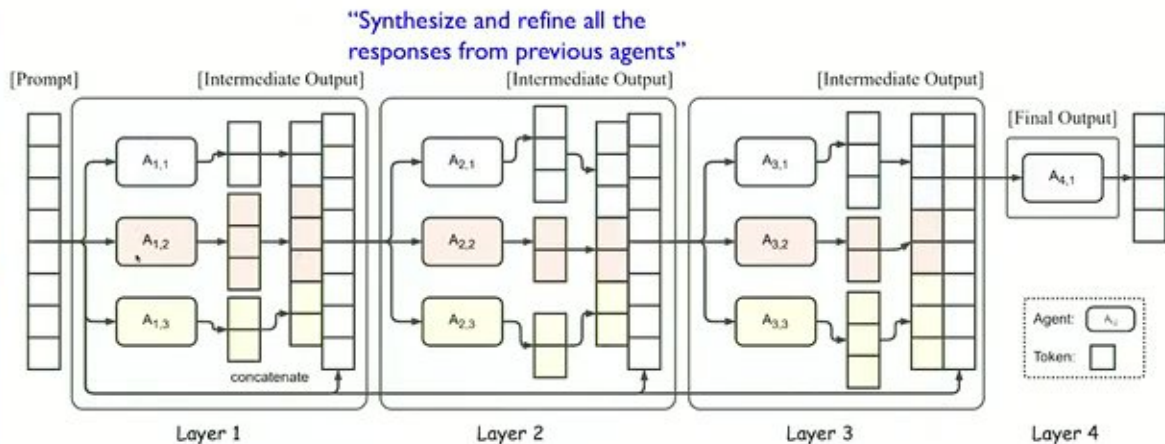
Model Name	LC Win Rate	Win Rate
Together MoA 	65.4%	59.9%
Together MoA-Lite 	59.1%	56.6%
Aligner 2B+GPT-4 Turbo (04/09) 	58.3%	46.8%
GPT-4 Omni (05/13) 	57.5%	51.3%
GPT-4 Turbo (04/09) 	55.0%	46.1%
Yi-Large Preview 	51.9%	57.5%
Storm-7B (num_beams=10) 	51.8%	55.4%
GPT-4 Preview (11/06) 	50.0%	50.0%

Wang et al. 2024



James Zou

The new frontier: mixture of agents (MoA)



Wang et al. 2024

MoA of open-source agents outperforms GPT4o



James Zou

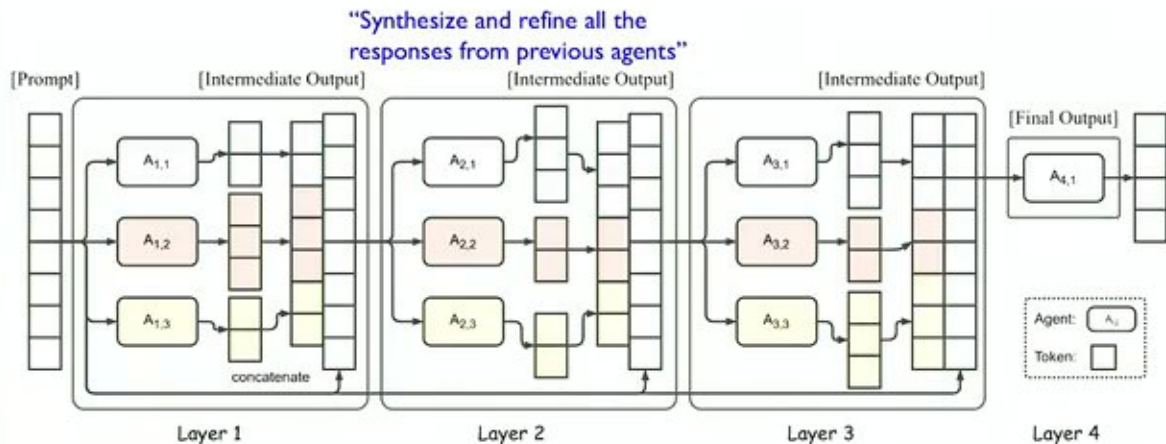
AlpacaEval  Leaderboard		
An Automatic Evaluator for Instruction-following Language Models		
Length-controlled (LC) win rates alleviate length biases of GPT-4, but it may favor models finetuned on its outputs.		
Version:	AlpacaEval AlpacaEval 2.0	Filter: Community Verified
Baseline: GPT-4 Preview (11/06) Auto-evaluator: GPT-4 Preview (11/06)		
Model Name	LC Win Rate	Win Rate
Together MoA	65.4%	59.5%
Together MoA-Lite	59.1%	56.0%
Aligner 2B+GPT-4 Turbo (04/09)	58.3%	46.8%
GPT-4 Omni (05/13)	57.5%	51.3%
GPT-4 Turbo (04/09)	55.0%	46.1%
Yi-Large Preview	51.9%	57.5%
Storm-7B (num_beams=10)	51.8%	55.4%
GPT-4 Preview (11/06)	50.0%	50.0%

Wang et al. 2024



James Zou

The new frontier: mixture of agents (MoA)



Wang et al. 2024

LLM summary

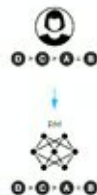
1. Masked imputation



2. Supervised fine-tuning



3. RLHF



4. Compound agent system



James Zou

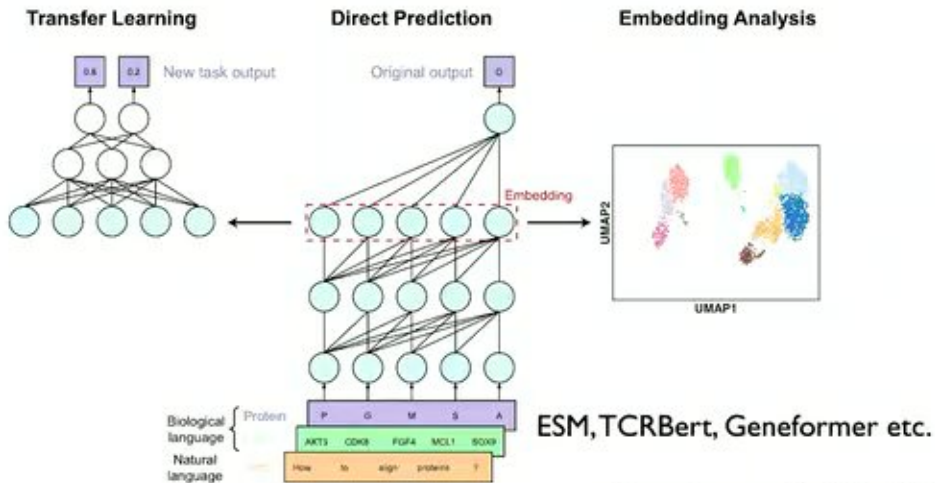
Outline

1. Tutorial on how to build large language models.
2. Language models for biology.



James Zou

Applying language modeling to biology



Simons, Swanson, Zou *Nature Methods* in press

James Zou



James Zou

What are possible representations for a gene?

Nucleotide sequences

```
TGCTCTTGTT GCCCAGGCTG
GAATGCAATG GCCTGATCTC
TGCTCACTGC AACCTCCACC
TCTCAGGTTC AAGCTATTCT
CCTGCCTCAG CCTCCCAAGT
AGCTGGGATT ACAGGTATGA
GCAACCATGC ATGGCTAATG
TTTGGATTTT TAGTAGAGAC
TAAATTTTAG TAGAGACTAGA
TTTTFAGTTT CACCATGTTG
GTCAGGTTGG TCTGGAACCC
TGACCTCAGG TGATCAGCCC
```

Ji et al. (2021); Poli et al. (2023); ...



James Zou

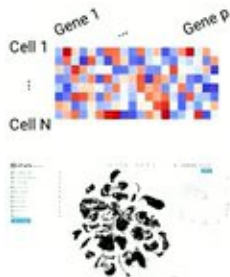
What are possible representations for a gene?

Nucleotide sequences

```
TGCTCTTGTT GCCCAGGCTG
GAATGCAATG GCCTGATCTC
TGCTCACTGC AACCTCCACC
TCTCAGGTTC AAGCTATTCT
CCTGCCTCAG CCTCCCAAGT
AGCTGGGATT ACAGGTATGA
GCAACCATGC ATGGCTAATG
TTTGGAATTT TAGTAGAGAC
TAAATTTTAG TAGAGACTAGA
TTTLAGGTTT CACCATGTGTG
GTCAGGTTGG TCTGGAACCC
TGACCTCAGG TGATCAGCCC
```

Ji et al. (2021); Poli et al. (2023); ...

Gene expression data



Theodoris et al. (2023); Cui et al. (2023); ...

Text descriptions for a gene

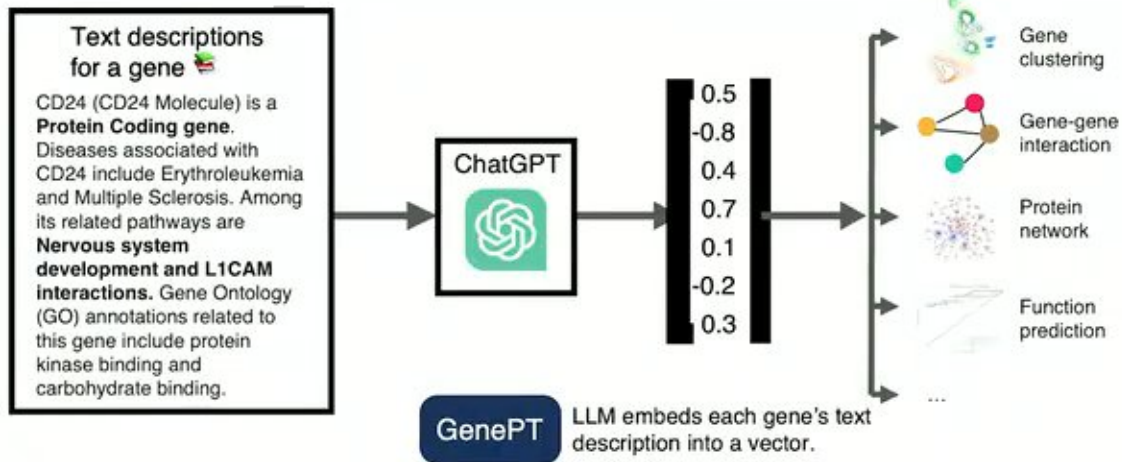
CD24 (CD24 Molecule) is a **Protein Coding gene**. Diseases associated with CD24 include Erythroleukemia and Multiple Sclerosis. Among its related pathways are **Nervous system development and L1CAM interactions**. Gene Ontology (GO) annotations related to this gene include protein kinase binding and carbohydrate binding.

Our proposal: leverage LLMs directly

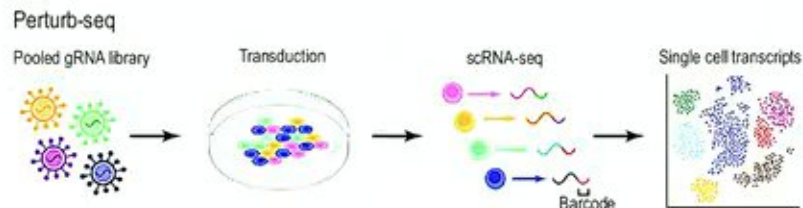


James Zou

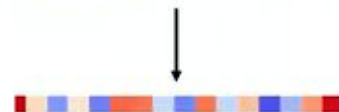
GenePT: natural language embeddings of genes and cells



GenePT embeddings predict gene perturbations



$f(\text{embed}(\text{perturbed gene}))$



Predict expression of other genes

	Replogle K562	Replogle RPE1	Tian CRISPR-a
GenePT	0.52	0.66	0.65
scGPT	0.24	0.46	0.53
Geneformer	0.41	0.64	0.64
GEARS	0.28	0.52	--

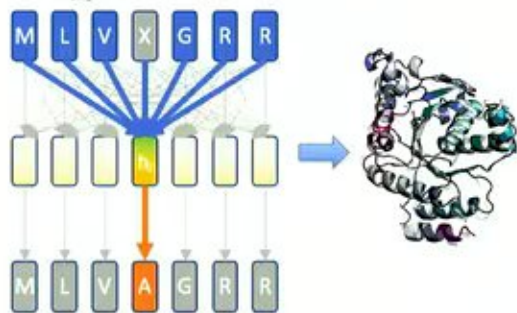
Pearson corr

James Zou

Enhancing protein models by projection into semantic space

Protein models learns from sequence correlations

$$p(x) = \prod_{i=1}^L p(x_i | x_1 \dots x_{i-1}, x_{i+1} \dots x_L)$$



Much knowledge about proteins also captured in text

P05067 - A4_HUMAN

Projet	« Impacte dels polítics en protecció ambiental »	Entitat col·laboradora	FCS de la Universitat de València
Tipus	ICP	Director/a del projecte	Emili Ferrer de l'Institut Ferrer
Ubicació	València (País Valencià) (Espanya)	Associació col·laboradora	
Organisme	Centre d'Estudis i Recerca		

[Home](#)
[Contact us](#)
[Privacy policy](#)
[Terms of use](#)
[Sitemap](#)
[Feedback](#)

Downloaded from <http://ajph.org/> on November 10, 2014

Discussion

[illegible][illegible]

Figures 1 and 2 illustrate the model and the data used in the analysis. The model is a simple one, with a single input variable, the number of hours worked, and a single output variable, the number of hours worked. The data is a time series of the number of hours worked by a single individual, from 1980 to 1990. The data is used to estimate the parameters of the model, and to test the hypothesis that the number of hours worked is a function of the number of hours worked.

The pricing of the portfolio as well as the required interest payments, including the tax on capital gains, are given by the following equations:

Submitted: 1999-09-01; Accepted: 2000-01-10. Address correspondence and reprint requests to Dr. J. A. J. van't Hof-Grootenboer, Department of Clinical Pharmacology, University Hospital Groningen, P.O. Box 30.001, 3000 RB Groningen, The Netherlands. E-mail: j.a.j.van't.hof-grootenboer@azg.umcg.nl

James Zou

Wu, Chang and Zou. 2024

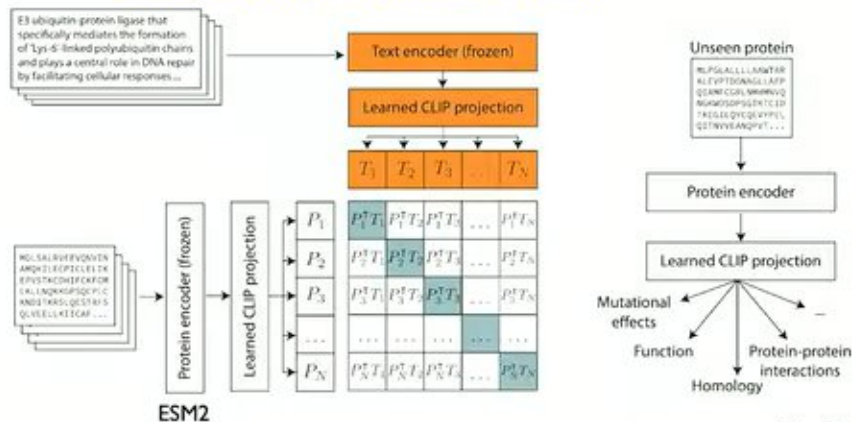
ESM2



James Zou

Enhancing protein models by projection into semantic space

ProteinCLIP: projecting protein models into natural language w/ contrastive learning



Wu, Chang and Zou. 2024

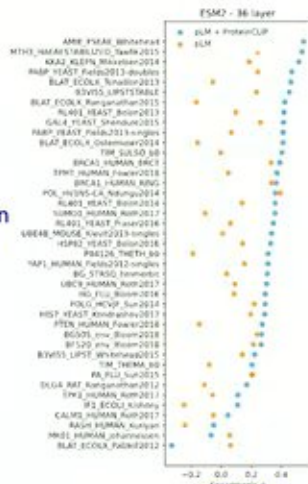
Better prediction



James Zou

ProteinCLIP improves prediction of mutation impact

Mutation scan experiments



Better prediction

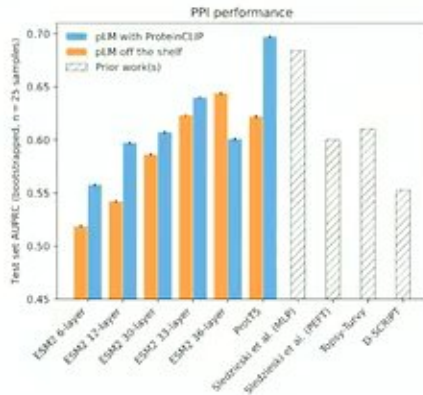
Wu, Chang and Zou. 2024



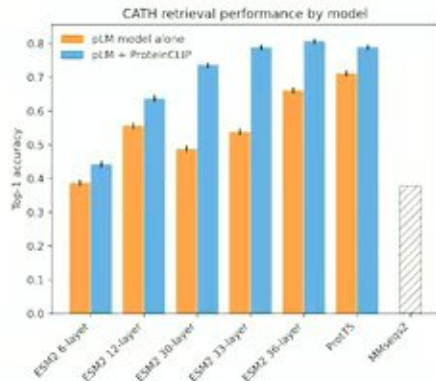
James Zou

ProteinCLIP improves prediction of protein interactions and homology

Predicting protein-protein interactions



Remote homology detection



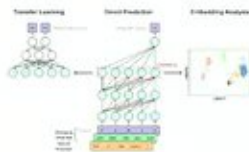
Wu, Chang and Zou. 2024



James Zou

Language models for biology

Bio LLM primer



Simons, Swanson, Zou
Nat Methods in press

Mixture of agents



Wang et al. arXiv 2024

Language as the unifying modality



Wu, Chang, Zou arXiv 2024
Chen and Zou. 2024

Thanks: Yiqun Chen, Howard Chang, Kevin Wu, Elana Simons, Kyle Swanson, Julin Wang, Jue Wang.

Funding: Chan-Zuckerberg Biohub, NSF CAREER, Sloan Fellowship, AWS, Google.