

香港中文大學（深圳）
The Chinese University of Hong Kong, Shenzhen

CSC5051/MDS5110/CSC6052: Natural Language Processing

Fall 2026

Benyou Wang
School of Data Science

Contents

- About me
- Philosophy of this course
- NLP and large language models (LLM)
- Introduction to our research

About me – 王本友 Benyou Wang.

2022-~, The Chinese University of Hong Kong, Shenzhen

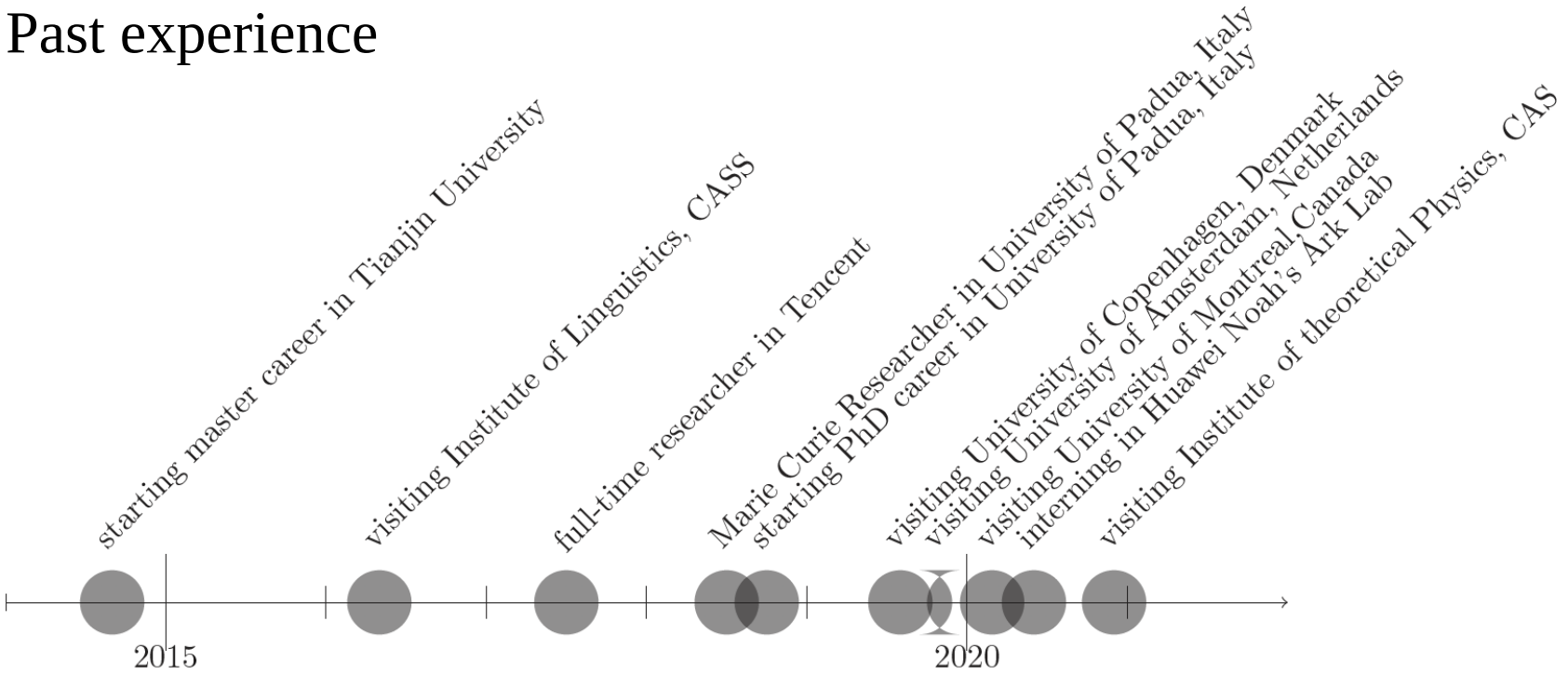
Assistant professor in **School of Data Science**, also affiliated in **School of Medicine in CUHKSZ,** as well as **the Shenzhen Loop Area Institute** （深圳河套学院）.

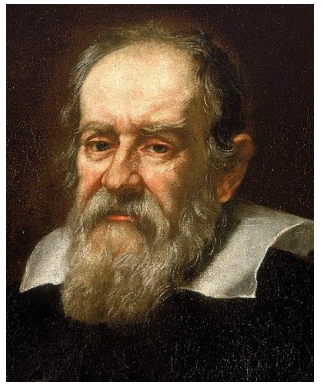
Research:

- **Planning** of Agent (complex reasoning/math modeling/tool calling/automatic theorem proof)
  - AI4math/modeling
- **Perception** of Agent (audio, vision, time-series, etc.)
  - Multimodal LLMs
- **Human-agent interaction** (HCI, agent simulation, brain-computer interface etc.)
- **Applications** of Agent (**Medical**/financial/legal/education)

About me 王本友 Benyou Wang.

Past experience





Galileo Galilei

the "father of **modern physics**"
the "father of the scientific method"
the "father of modern science"

Alumni of University of Padua

Academia or industrial career

- For money, less is more?
- Work for yourself or others
- Free to work in any topics
- Benefit the society or just the employer?
  - easier to connect with diverse people.
- Long vacation is much easier (but you may never take it)

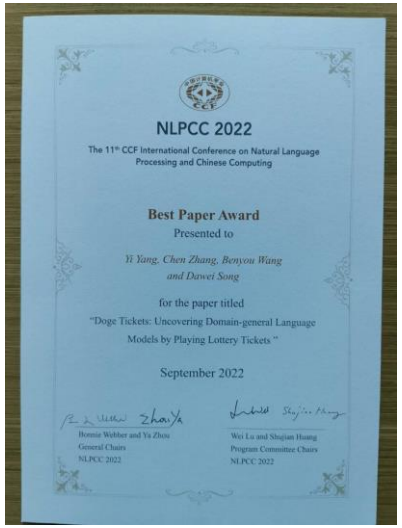
- Weekend is weekend
- Less pressure to be sustainable

Fitting yourself is the best

Awards and honour





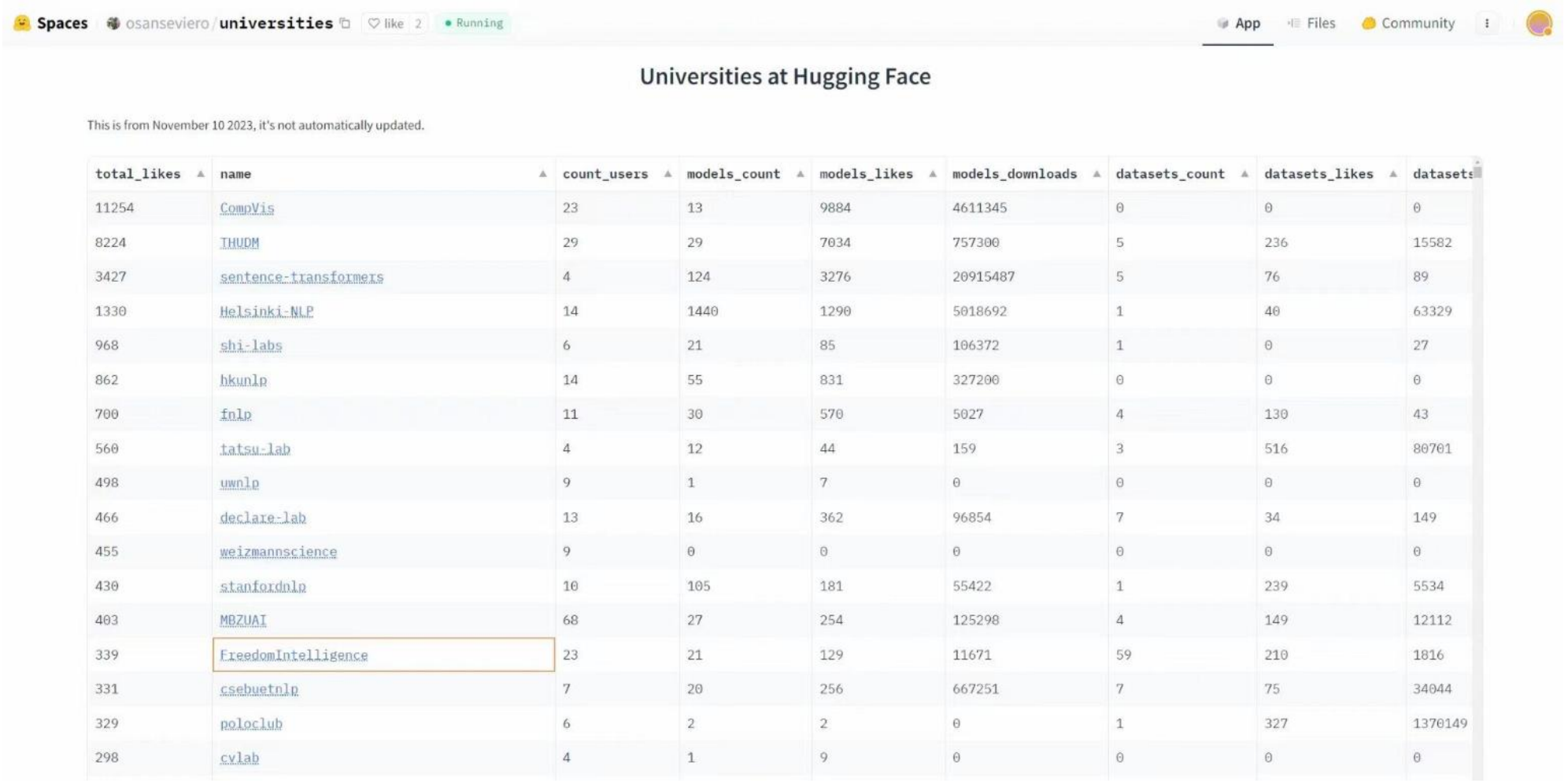




- **NLPCC 2022** Best Paper
- **ACM SIGIR** 2017 Best paper honourable mention. https://sigir.org/awards/best-paper-awards/
- **NAACL** 2019 best explainable NLP paper. https://naacl2019.org/blog/best-papers/
- ICLR 2025 Financial AI Best Paper
- Outstanding Paper Paper in NeurIPS 2025 ResponsibleFM Workshop
- EU Marie Curry researcher fellowship

- Huawei Spark award (华为火花奖)
- 腾讯犀牛鸟项目
- 滴滴盖亚学者项目
- 华为AI百校计划

We rank 14th in Greater China



Universities at Hugging Face

This is from November 10 2023, it's not automatically updated.

total_likes	name	count_users	models_count	models_likes	models_downloads	datasets_count	datasets_likes	datasets
11254	CompVis	23	13	9884	4611345	0	0	0
8224	THUDM	29	29	7034	757300	5	236	15582
3427	sentence-transformers	4	124	3276	20915487	5	76	89
1330	Helsinki-NLP	14	1440	1290	5018692	1	40	63329
968	shi-labs	6	21	85	106372	1	0	27
862	hkunlp	14	55	831	327200	0	0	0
700	fnlp	11	30	570	5027	4	130	43
560	tatsu-lab	4	12	44	159	3	516	80701
498	uwnlp	9	1	7	0	0	0	0
466	declare-lab	13	16	362	96854	7	34	149
455	weizmannscience	9	0	0	0	0	0	0
430	stanfordnlp	10	105	181	55422	1	239	5534
403	MBZUAI	68	27	254	125298	4	149	12112
339	FreedomIntelligence	23	21	129	11671	59	210	1816
331	csebuetnlp	7	20	256	667251	7	75	34044
329	poloclub	6	2	2	0	1	327	1370149
298	cvlab	4	1	9	0	0	0	0

https://huggingface.co/spaces/osanseviero/universities
https://twitter.com/osanseviero/status/1723229014100255011

Be Open for AI





Find **dataset/models/papers**!

Find **code**!

HuggingFace: https://huggingface.co/FreedomIntelligence (**1,000,000**+ models downloading)

GitHub: https://github.com/FreedomIntelligence (**100,000** + stars)

OpenClaw-Medical-Skills	The largest open-source medical AI skills library for OpenClaw.	Python		407	3k
LLMZoo	⚡LLM Zoo is a project that provides data, models, and evaluation benchmark for large language models.⚡	Python		194	2.9k
Awesome-AI4Med	A curated list of medical LLMs, multimodal systems, datasets, benchmarks, and more.			496	2.9k
HuatuoGPT-o1	Medical o1, Towards medical complex reasoning with LLMs	Python		132	1.4k
HuatuoGPT	HuatuoGPT, Towards Taming Language Models To Be a Doctor. (An Open Medical GPT)	Python		168	1.3k
TextClassificationBenchmark	A Benchmark of Text Classification in PyTorch	Python		133	608
HuatuoGPT-II	HuatuoGPT2, One-stage Training for Medical Adaption of LLMs. (An Open Medical GPT)	Python		59	414
HuatuoGPT-Vision	Medical Multimodal LLMs	Python		36	401
Huatuo-26M	The Largest-scale Chinese Medical QA Dataset： with 26,000,000 question answer pairs.			33	349
BlenderLLM	BlenderLLM: A LLM specifically designed to generate CAD scripts based on user instructions. These scripts are then executed in Blend...	Python		30	301
ShareGPT-4o-Image		Python		11	284
ALLaVA	Harnessing 1.4M GPT4V-synthesized Data for A Lite Vision-Language Model	Python		9	282
InstructionZoo				25	282
CMB	CMB, A Comprehensive Medical Benchmark in Chinese	Python		20	252
Soundwave	The official Soundwave repository	Python		20	223
gamecraft-bench	Code and Data for paper "GameCraft-Bench: Can Agents Build Playable Games End-to-End in a Real Game Engine?"	Shell		13	219
TwinMarket	[NeurIPS 2025 & ICLR 2025 Financial AI Best Paper Award] A multi-agent framework that leverages LLMs to simulate socio-economic ...	Python		40	215
LongLLaVA	LongLLaVA: Scaling Multi-modal LLMs to 1000 Images Efficiently via Hybrid Architecture	Python		17	212
Apollo	Multilingual Medicine: Model, Dataset, Benchmark, Code	Python		8	203
TalkVid	[CVPR 2026 Findings] TalkVid: A Large-Scale Diversified Dataset for Audio-Driven Talking Head Synthesis	Python		13	201
ReasoningNLP	paper list on reasoning in NLP			14	199
GrammarGPT	The code and data for GrammarGPT.	Python		8	177
Chain-of-Diagnosis	An interpretable large language model (LLM) for medical diagnosis.	Python		8	162
crosstalk-generation	Code and data for crosstalk text generation tasks, exploring whether large models and pre-trained language models can u...	Python		17	150
ShizhenGPT				17	147
RAG-Instruct	[EMNLP 2025] RAG-Instruct: Boosting LLMs with Diverse Retrieval-Augmented Instructions	Python		5	135
Evaluation-of-ChatGPT-on-Information-Extraction	An Evaluation of ChatGPT on Information Extraction task, including Named Entity Recog...	Python		11	133
Awesome-LLM-Patient-Simulators	A Paper collection for LLM based Patient Simulators			4	132
AceGPT		Python		9	127
Awesome-Rubrics	A curated list of resources (surveys, papers, benchmarks, and opensource projects) on Rubrics			4	117

https://github.com/orgs/FreedomIntelligence

Our team - Join us as a Research Assistant!

What You Will Receive:

- Access to GPU **computing resources** and utilization of abundant **GPT APIs**.
- More frequent **communication** within our research team (also more pressure).
- Potential for research **publications** (coauthorship).
- The possibility of a part-time **contract** with a salary (also with office space).
- **Internship** recommendation (Microsoft,Qwen, BAT, Huawei, Bytedance)

What We Expect From You:

- A commitment of at least 30 hours per week for full-time engagement.
- satisfied programming skills.

If you find this opportunity intriguing, please reach out to me (wangbenyou@cuhk.edu.cn和 lixiang2@cuhk.edu.cn) for further details.

Contents

- **About me**
- **Philosophy of this course**
- **NLP and large language models (LLM)**
- **Introduction to our research**

Logistics

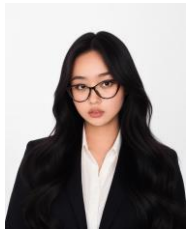


- Instructor: Benyou Wang
- Teaching assistant: Shunian Chen (Leading TA)
  Shunian Chen, Minghao Wu, Zhenyang Tang， Ke Ji






- Location: **Administration Bldg E101**
- Time: Friday 2:00PM - 4:50PM
- Office hours:
  - Benyou Wang: Friday 6:30-7:30 PM. Daoyuan Building 504A. (Email: wangbenyou@cuhk.edu.cn)
  - See **details** in https://nlp-course-cuhksz.github.io/








Previous TA：考虑做RSI方向的创业
Recursive Self-improvement

Logistics

- Official Website Link (https://nlp-course-cuhksz.github.io/)

Course Information

This comprehensive course on Natural Language Processing (NLP) offers a deep dive into the field, providing students with the knowledge and skills to understand, design, and implement NLP systems. Starting with an overview of NLP and foundational linguistic concepts, the course moves on to word representation and language modeling, essential for understanding text data. It explores how deep learning, from basic neural networks to advanced transformer models, has revolutionized NLP and its diverse applications, such as text mining, information extraction, and machine translation. The course emphasizes large language models (LLMs), their scaling laws, emergent abilities, training strategies, and associated knowledge representation and reasoning. Students will apply their learning in final projects, for example, exploring NLP beyond text with multi-modal LLMs, AI for Science, vertical applications and agents. There are guest lectures and in-class paper discussions that could learn the cut-edge research. The course also concludes with an examination of NLP's limitations and ethical considerations. ■

Date	Topics	Recommended Reading	Pre-Lecture Questions	Lecture Note	Coding	Events Deadlines
Sept. 1-4 Warmup	Tutorial 0: GitHub, LaTeX, Colab, and ChatGPT API	OpenAI's blog LaTeX and Overleaf Colab GitHub				
Sept. 5th	Lecture 1: Introduction to NLP	Hugging Face NLP Course Course to get into NLP with roadmaps and Colab notebooks. LLM-Course On the Opportunities and Risks of Foundation Models Sparks of Artificial General Intelligence: Early experiments with GPT-4	What is NLP?		[Phoenix]	

- Official Wechat Group



群聊: NLP 2026 Fall

该二维码7天内(9月18日前)有效，重新进入将更新

Course Structure

- This is **an advanced graduate course** and we will be offering a deep dive into the field and equip students with the comprehensive the knowledge and skills to understand, design, and implement NLP systems.
- All the students are expected to come to the class regularly and participate in discussion

- Prerequisites:
  - Proficiency in **LaTex**: All the reports need to be written by using LaTex. A template will be provided. If you are not familiar with LaTex, please learn from the tutorial in advance.
  - Proficiency in **GitHub**: All the source codes are suggested to be submitted in GitHub.
  - Proficiency in **Python**: All the assignments will be in Python (using Numpy and PyTorch).
  - Basic **machine learning** knowledge: It is possible to take this course without any machine learning knowledge, however, the course will be easier if you have foundations of machine learning.

Survey: How many of you have you ever used LaTeX, GitHub, Python (PyTorch)?

Course Structure (tentative)

- Introduction to NLP
- Linguistics and Word Embeddings
- Language Models
- Deep Learning in NLP （Transformer and beyond）
- Large Language Models (LLMs)
- Training Large Language Models
- LLM Agents
- Final Project Introduction and Research Sharing
- Harness Engineering
- Multimodal Learning
- LLM Reasoning
- LLM Application

Course Structure (tentative)

~~•Introduction to NLP~~

~~•Linguistics and Word Embeddings~~

Too Old!

~~•Language Models~~

~~•Deep Learning in NLP （Transformer and beyond~~）

•Large Language Models (LLMs)

•Training Large Language Models

•LLM Agents

•Final Project Introduction and Research Sharing

•Harness Engineering

•Multimodal Learning

•LLM Reasoning

•LLM Application

Course Structure (tentative)

- Introduction to NLP
- Linguistics and Word Embeddings
- Language Models

Too Old!

- Deep Learning in NLP （Transformer and beyond）
- Large Language Models (LLMs)
- Training Large Language Models
- LLM Agents

Practical

- Final Project Introduction and Research Sharing
- Harness Engineering
- Multimodal Learning
- LLM Reasoning
- LLM Application

Course Structure (tentative)

- Introduction to NLP
- Linguistics and Word Embeddings
- Language Models
- Deep Learning in NLP （Transformer and beyond）
- Large Language Models (LLMs)
- Training Large Language Models
- LLM Agents
- Final Project Introduction and Research Sharing
- Harness Engineering
- Multimodal Learning
- LLM Reasoning
- LLM Application

Too Old!

Practical

Final projects are all you need!

Course Structure (tentative)

- Introduction to NLP
- Linguistics and Word Embeddings
- Language Models
- Deep Learning in NLP （Transformer and beyond）
- Large Language Models (LLMs)
- Training Large Language Models
- LLM Agents
- Final Project Introduction and Research Sharing
- Harness Engineering
- Multimodal Learning
- LLM Reasoning
- LLM Application

Too Old!

Practical

Final projects are all you need!

Cutting-edge research

Components and Grading (Post-Graduate)

- Assignments (**40%**)
  - Assignment 1 (10%): Training word vector.
  - Assignment 2 (15%): Training NLP model with SFT and RLHF
  - Assignment 3 (15%): Using API for testing prompt engineering and LLM agents.

  All assignments need a report and code attachment if it has coding. See the relevant evaluation criterion as the final project.

- Final project (**55%**)

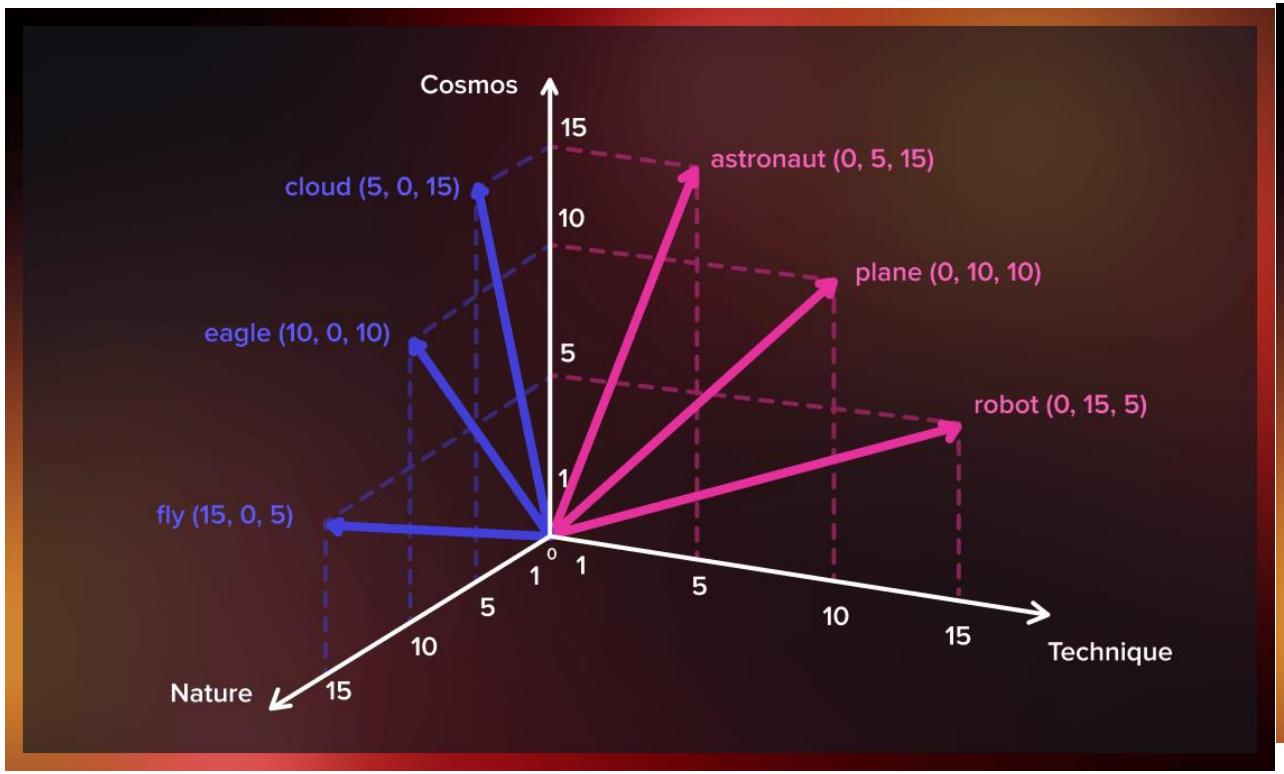
  You need to write a project report (max 6 pages for the main body) for the final project [*Be longer is not necessary, not recommended*].

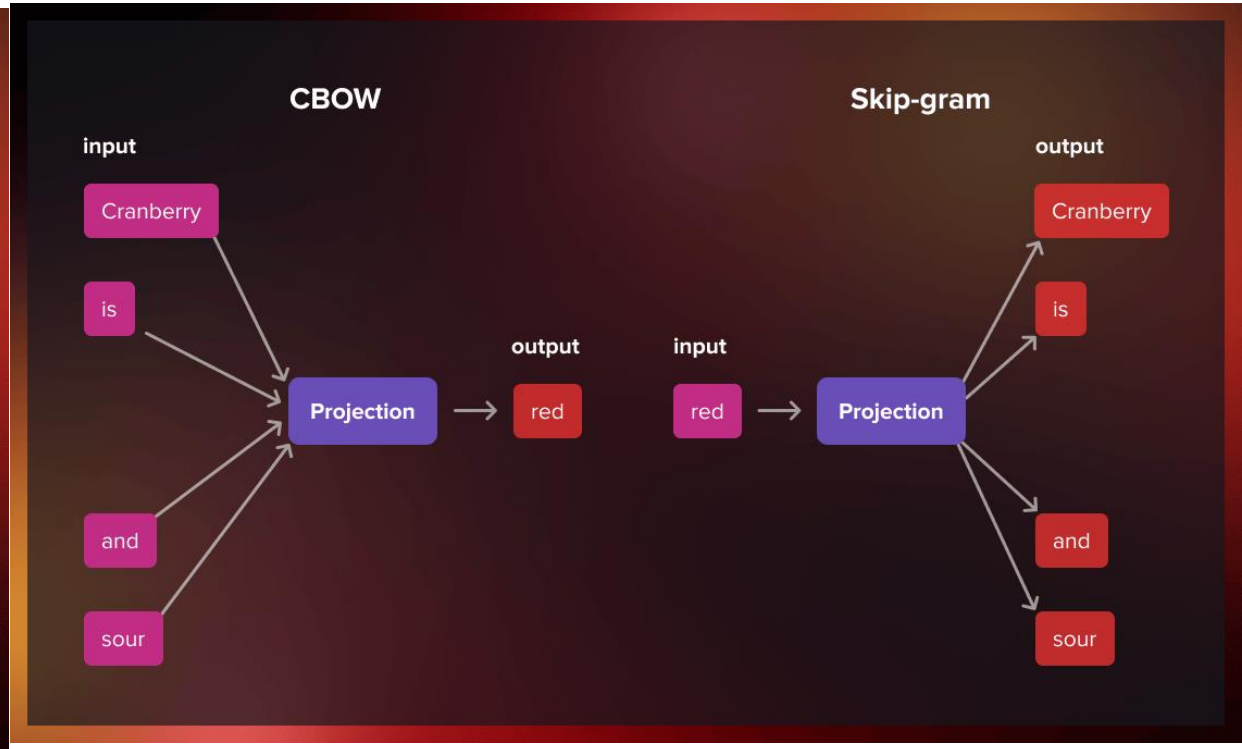
  You are also expected to make a project **poster** presentation. After the final project deadline, feel free to make your project open source; we appreciate if you acknowledge this course.

- Participation (5%)

  Quiz, guest lecture and CTE

Assignments 1: Training word vector.





Task description: This task aiming at deepening your practical understanding of word embedding.

Source: https://serokell.io/blog/word2vec

Assignments 2: Training NLP model with SFT and RLHF

For Developers

```
import llmfactory

# Configure the resource in the factory/resource.json file
factory = llmfactory.Factory()

# Show available models
factory.show_available_model()
# Output:
# [Bloom]: bloom-560m, bloomz-560m, bloom-1b1, bloomz-1b1, bloomz-7b1-mt
# [Llama]: llama-7b-hf, llama-13b-hf
# [Baichuan]: baichuan-7B

# Show available data
factory.show_available_data()
# Output:
# [Local]: music, computer, medical

# Select a model from the available model set
model_config = factory.create_backbone("bloom-560m")

# Set up the data configuration
data_config = factory.prepare_data_for_training(num_data=50, data_ratios

# Train a new model based on the existing model and data configuration
model_config = factory.train_model(model_config, data_config, save_name=

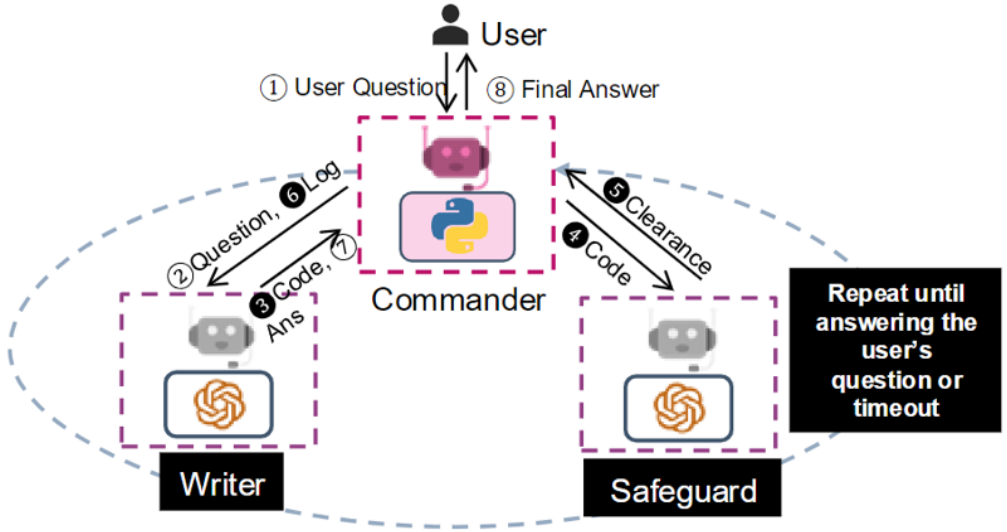
# Deploy the model on the command line
factory.deploy_model_cli(model_config)

# Deploy the model using Gradio
factory.deploy_model_gradio(model_config)
```

Finetuning: https://github.com/FreedomIntelligence/LLMFactory

RL: https://github.com/volcengine/verl

Assignment 3: Using API for Prompt Engineering and LLM Agents

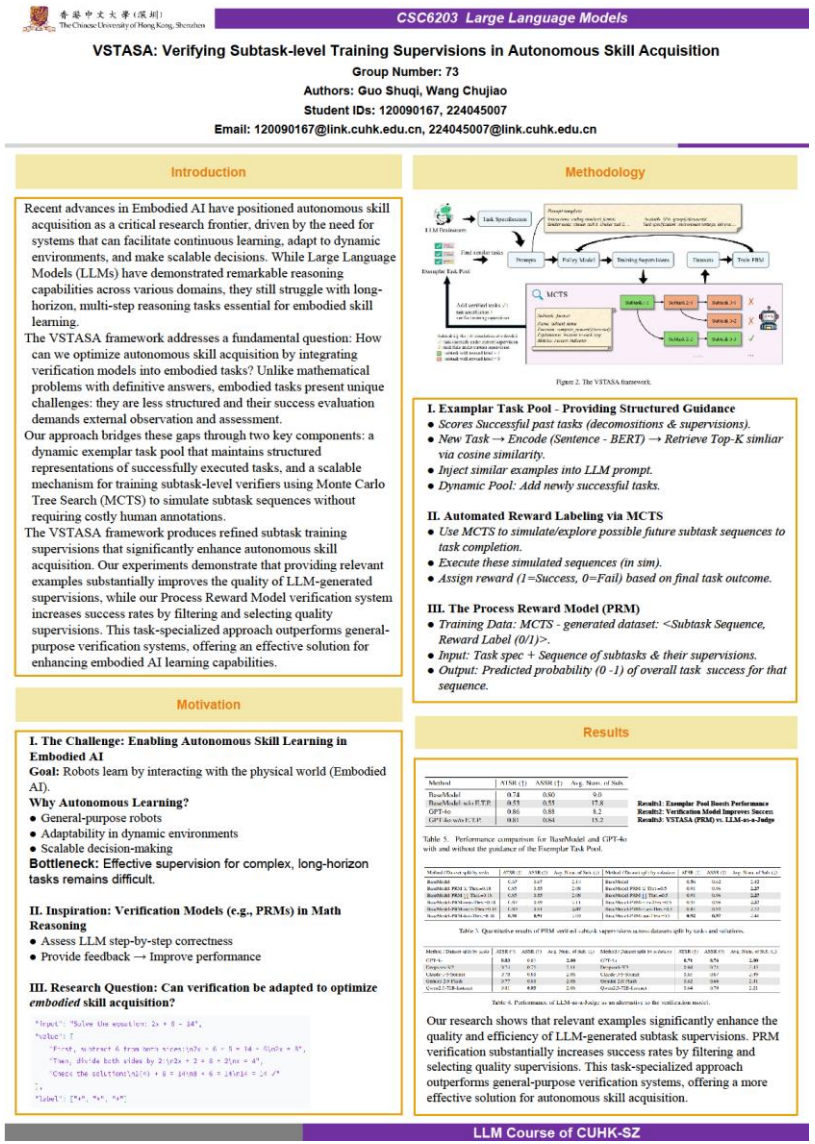


- Task Description: Explore prompt engineering using API. Utilizing existing LLM Agent frameworks to achieve practical goals.

- Requirement: Submit a report with code showcasing your practical skills in using LLM agents. More details released later.

API: 硅基流动: https://siliconflow.cn/ or https://openrouter.ai/

Agent: Camel.AI, Dify, LangChain, Coze



香港中文大學(深圳)
The Chinese University of Hong Kong, Shenzhen

CSC6203 Large Language Models

VSTASA: Verifying Subtask-level Training Supervisions in Autonomous Skill Acquisition

Group Number: 73

Authors: Guo Shuqi, Wang Chujiao

Student IDs: 120090167, 224045007

Email: 120090167@link.cuhk.edu.cn, 224045007@link.cuhk.edu.cn

Introduction

Recent advances in Embodied AI have positioned autonomous skill acquisition as a critical research frontier, driven by the need for systems that can facilitate continuous learning, adapt to dynamic environments, and make scalable decisions. While Large Language Models (LLMs) have demonstrated remarkable reasoning capabilities across various domains, they still struggle with long-horizon, multi-step reasoning tasks essential for embodied skill learning.

The VSTASA framework addresses a fundamental question: How can we optimize autonomous skill acquisition by integrating verification models into embodied tasks? Unlike mathematical problems with definitive answers, embodied tasks present unique challenges: they are less structured and their success evaluation demands external observation and assessment.

Our approach bridges these gaps through two key components: a dynamic exemplar task pool that maintains structured representations of successfully executed tasks, and a scalable mechanism for training subtask-level verifiers using Monte Carlo Tree Search (MCTS) to simulate subtask sequences without requiring costly human annotations.

The VSTASA framework produces refined subtask training supervisions that significantly enhance autonomous skill acquisition. Our experiments demonstrate that providing relevant examples substantially improves the quality of LLM-generated supervisions, while our Process Reward Model verification system increases success rates by filtering and selecting quality supervisions. This task-specialized approach outperforms general-purpose verification systems, offering an effective solution for enhancing embodied AI learning capabilities.

Motivation

I. The Challenge: Enabling Autonomous Skill Learning in Embodied AI

Goal: Robots learn by interacting with the physical world (Embodied AI).

Why Autonomous Learning?

- General-purpose robots
- Adaptability in dynamic environments
- Scalable decision-making

Bottleneck: Effective supervision for complex, long-horizon tasks remains difficult.

II. Inspiration: Verification Models (e.g., PRMs) in Math Reasoning

- Assess LLM step-by-step correctness
- Provide feedback → Improve performance

III. Research Question: Can verification be adapted to optimize *embodied* skill acquisition?

```
"Input": "Solve the equation: 2x + 8 = 14",
"value": [
    "First, subtract 8 from both sides: 2x + 8 - 8 = 14 - 8\n2x = 6",
    "Then, divide both sides by 2: 2x / 2 = 6 / 2\nx = 3",
    "Check the solution: 2(3) + 8 = 14\n6 + 8 = 14\n14 = 14 ✓"
],
"label": ["+", "+", "+"]
```

Methodology

Figure 2. The VSTASA framework.

I. Examplar Task Pool - Providing Structured Guidance

- *Scores Successful past tasks (decomositions & supervisions).*
- *New Task → Encode (Sentence - BERT) → Retrieve Top-K simliar via cosine similarity.*
- *Inject similar examples into LLM prompt.*
- *Dynamic Pool: Add newly successful tasks.*

II. Automated Reward Labeling via MCTS

- *Use MCTS to simulate/explore possible future subtask sequences to task completion.*
- *Execute these simulated sequences (in sim).*
- *Assign reward (1=Success, 0=Fail) based on final task outcome.*

III. The Process Reward Model (PRM)

- *Training Data: MCTS - generated dataset: <Subtask Sequence, Reward Label (0/1)>.*
- *Input: Task spec + Sequence of subtasks & their supervisions.*
- *Output: Predicted probability (0 -1) of overall task success for that sequence.*

Results

Method	ASSR (↑)	ASSR (↑)	Avg. Num. of Sub.
BaseModel	0.74	0.80	6.0
BaseModel+RVR	0.53	0.55	17.8
GPT-4o	0.86	0.88	8.2
GPT-4o+RVR	0.81	0.84	10.2

Results1: Exemplar Pool Boosts Performance
Results2: Verification Model Improves Success
Results3: VSTASA (PRM) vs. LLM-as-a-Judge

Table 5. Performance comparison for BaseModel and GPT-4o with and without the guidance of the Exemplar Task Pool.

Table 3. Quantitative results of PRM verified subtask supervisions across datasets split by tasks and solutions.

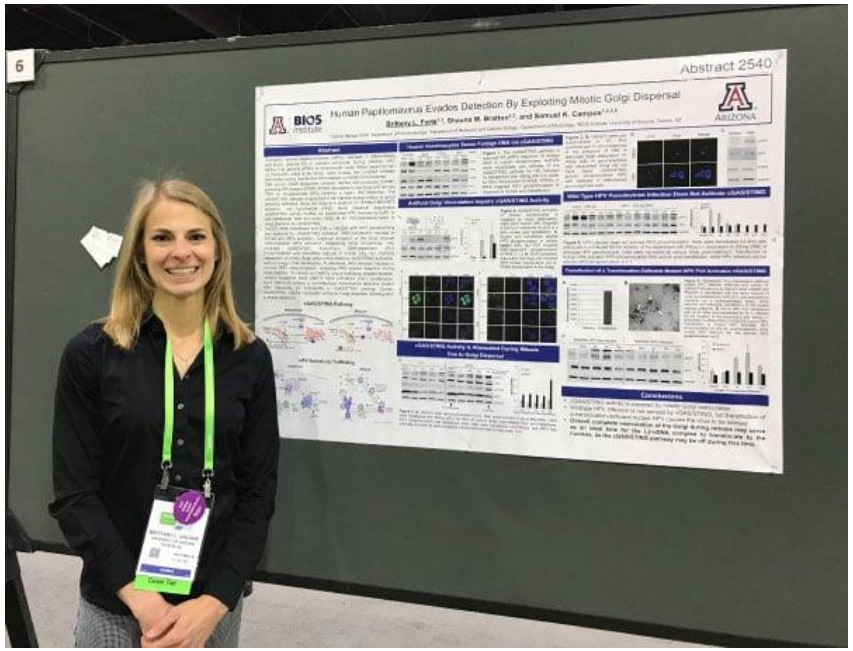
Table 4. Performance of LLM-as-a-Judge as an alternative to the verification model.

Our research shows that relevant examples significantly enhance the quality and efficiency of LLM-generated subtask supervisions. PRM verification substantially increases success rates by filtering and selecting quality supervisions. This task-specialized approach outperforms general-purpose verification systems, offering a more effective solution for autonomous skill acquisition.

LLM Course of CUHK-SZ



Dennis Schroeder/NREL via flickr.com. Public domain (US government agency)



Example of Poster

Report for the final project

Attention Is All You Need

Ashish Vaswani* Google Brain avaswani@google.com
Noam Shazeer* Google Brain noam@google.com
Niki Parmar* Google Research nikip@google.com
Jakob Uszkoreit* Google Research usz@google.com
Llion Jones* Google Research llion@google.com
Aidan N. Gomez* † University of Toronto aidan@cs.toronto.edu
Łukasz Kaiser* Google Brain lukaszkaiser@google.com
Illia Polosukhin* ‡ illia.polosukhin@gmail.com

Abstract

The dominant sequence transduction models are based on complex recurrent or convolutional neural networks that include an encoder and a decoder. The best performing models also connect the encoder and decoder through an attention mechanism. We propose a new simple network architecture, the Transformer, based solely on attention mechanisms, dispensing with recurrence and convolutions entirely. Experiments on two machine translation tasks show these models to be superior in quality while being more parallelizable and requiring significantly less time to train. Our model achieves 28.4 BLEU on the WMT 2014 English-to-German translation task, improving over the existing best results, including ensembles, by over 2 BLEU. On the WMT 2014 English-to-French translation task, our model establishes a new single-model state-of-the-art BLEU score of 41.8 after training for 3.5 days on eight GPUs, a small fraction of the training costs of the best models from the literature. We show that the Transformer generalizes well to other tasks by applying it successfully to English constituency parsing both with large and limited training data.

*Equal contribution. Listing order is random. Jakob proposed replacing RNNs with self-attention and started the effort to evaluate this idea. Ashish, with Illia, designed and implemented the first Transformer models and has been crucially involved in every aspect of this work. Noam proposed scaled dot-product attention, multi-head attention and the parameter-free position representation and became the other person involved in nearly every detail. Niki designed, implemented, tuned and evaluated countless model variants in our original codebase and tensor2tensor. Llion also experimented with novel model variants, was responsible for our initial codebase, and efficient inference and visualizations. Lukasz and Aidan spent countless long days designing various parts of and implementing tensor2tensor, replacing our earlier codebase, greatly improving results and massively accelerating our research.

†Work performed while at Google Brain.

‡Work performed while at Google Research.

31st Conference on Neural Information Processing Systems (NIPS 2017), Long Beach, CA, USA.

It is like a paper

Final project

- Students complete a research project in teams of 1-3
- [**Optional**] Draft proposal deadline: **March. 27th** 1:59pm (simply explain your idea in one page)

- Poster presentation (poster sessions)**: May. 8th**
- (TAs and the Instructor may provide suggestions for the revision during poster session!)

- Final poster deadline: **May. 7th** 17:00pm
- Final paper deadline: **May. 16th**

Prepare your project earlier to avoid the rush during exam weeks

Date	Topics	Recommended Reading	Pre-Lecture Questions	Lecture Note	Coding	Events Deadlines
Sep. 7-10 Warmup	Tutorial 0: GitHub, LaTeX, Colab, and ChatGPT API	OpenAI's blog LaTeX and Overleaf Colab GitHub				
Sep. 11th	Lecture 1: Introduction to NLP	Hugging Face NLP Course Course to get into NLP with roadmaps and Colab notebooks. LLM-Course On the Opportunities and Risks of Foundation Models Sparks of Artificial General Intelligence: Early experiments with GPT-4	What is NLP?	[Slides]	[Phoenix]	
Sep. 18th	Lecture 2: Basics of Linguistics and Word Representation	Universal Stanford Dependencies: A cross-linguistic typology Insights between NLP and Linguistics End-to-end Neural Coreference Resolution Efficient Estimation of Word Representations in Vector Space (original word2vec paper) Evaluation methods for unsupervised word embeddings	What is structure of language (string of words)? How to model language and the inside words?	[Slides]	[Linguistics repo]	**Assignment 1 out**
Sep. 18th	Tutorial 1: Introduction to Overleaf, GitHub, Python, Pytorch, and Word Embeddings				[Colab]	
~~Sep. 25th~~	Mid-Autumn Festival Holiday — No Class	中秋节假期：9月25日（周五）至9月27日（周日）				
~~Oct. 2nd~~	National Day Holiday — No Class	国庆节假期：10月1日（周四）至10月7日（周三）				
Oct. 9th	Lecture 3: Language Modeling	A Neural Probabilistic Language Model BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding A Neural Probabilistic Language Model BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding	What is a language model and how does it work in natural language processing?	[Slides]	[BERT]	
Oct. 16th	Lecture 4: Transformer and Beyond	Attention Is All You Need HuggingFace's course on Transformers Scaling Laws for Neural Language Models The Transformer Family Version 2.0 On Position Embeddings in BERT	How to better compose words semantically as language?	[Slides]	[Transformer]	
Oct. 23rd	Lecture 5: Large Language Models (LLMs)	Training language models to follow instructions with human feedback LLaMA: Open and Efficient Foundation Language Models Llama 2: Open Foundation and Fine-Tuned Chat Models OpenAI's blog Qwen2.5 Technical Report DeepSeek-V3 Technical Report	what are LARGE language models and why LARGE?	[Slides]	[Fine-tune Llama 2]	**Assignment 1 due (11:59pm)** **Assignment 2 out**
Oct. 23rd	Tutorial 2: SFT and RLHF for LLMs		Are you ready to train your own LLMs?		[LLMZoo], [nanoGPT], [LLMFactory]	
Oct. 30th	Lecture 6: Agent	LLM Powered Autonomous Agents ReAct: Synergizing Reasoning and Acting in Language Models AgentBench Mobile Agent	How to make LLMs interact with the real world?	[Slides]		**Final Project out**
Nov. 6th	Lecture 7: Final Projects and Research Sharing	ToolBench Cline Roo-Cline	How to define a good final project?	[Slides]		**Assignment 2 due (11:59pm)** **Assignment 3 out**
Oct. 30th	Tutorial 3: Prompt Engineering and LLM Agents				[Colab]	
Nov. 13th	Lecture 8: Harness Engineering	Best practices for prompt engineering with OpenAI API Prompt Engineering (Lilian Weng) Chain-of-Thought Prompting Elicits Reasoning in Large Language Models Tree of Thoughts: Deliberate Problem Solving with Large Language Models	How can prompts, context and tools be engineered as a reliable harness around an LLM?	[Slides]		
Nov. 13th	Tutorial 4: RAG and Agent Development				[Colab]	**Optional: Project Proposal due** **Team Formation due**
Nov. 20th	Lecture 9: LLM Medical Applications	HuatuoGPT, towards Taming Language Model to Be a Doctor HuatuoGPT-II, One-stage Training for Medical Adaption of LLMs Large Language Models Encode Clinical Knowledge Apollo: Multilingual Medical LLMs	How to apply LLMs in medical and other vertical domains?	[Slides]		
Nov. 27th	Lecture 10: Multimodal LLMs	Blog post: Generalized Visual Language Models CLIP: Learning Transferable Visual Models From Natural Language Supervision LLaVA: Visual Instruction Tuning NExT-GPT	Can large models speak, see and perform actions?	[Slides]		**Assignment 3 due (11:59pm)**
Dec. 4th	Lecture 11: Guest Lecture					
Dec. 11th	Lecture 12: LLM Applications and Future Cont.	Large Language Models Encode Clinical Knowledge Survey of Hallucination in Natural Language Generation Superalignment GPTs are GPTs: An Early Look at the Labor Market Impact Potential of Large Language Models	What are the future of LLMs?			
Dec. 18th	Final Project Presentation		How to solve real-world problems using LLMs			**Final Project Poster due (17:00pm Dec. 17th)**
Dec. 25th	Final Report Submission					**Final Report Submission due (11:59 pm Dec. 25th)**

What can we do for the final project?

Workload of final project is **very high**!

Be careful if you do not have enough time!

Prepare as earlier as possible

What can we do for the final project?

Philosophy :

- select one of given proposals (you could do nothing but wait for the release of proposals)**.**
- **OR** submit a customed project with a proposal that needs to be approved

- **Call for Project Proposals (optional):** (Deadline tentatively Nov. 1st 11:59pm .).
  - Choose an NLP research topic of interest or select from provided options.
  - The team members should be decided
- **Who could submit the proposal**
  - The students in this course (You could decide whether the submitted proposal could be used by other teams if approved)
  - Research scientists or engineers in a company
  - RAs and Phd students in our research team
  - Falcuty members in our university (your supervisors for example)
  - Actually, everyone is welcome, we reserve the rights to reject a proposal.

Final project

Typical projects (we will release a detailed list later):  Hugging Face https://huggingface.co/models

1. Fine-tune (or RL) a language model (e.g., Qwen) yourself for any problems of your interest.

   "*Do we really need to tune the base model?*"

2. Evaluate existing LLMs (e.g., GPT-6) and understand their capabilities, limitations and risks;
3. An Agent System (math/financial/medical/legal domains)
4. An APP that uses AI (at least built by AI).
5. A research paper involving NLP (please acknowledge this course if possible)
6. ~~An survey that involving NLP~~
7. New! A competition with a technical report.

Note: You might get computing resources to train 10B+ model if TAs/instructor like your proposal

More project ideas will be presented during the lectures and all creative ideas or research topics are encouraged for further discussion.

https://openai.com/api/ https://opt.alpa.ai

Final project (suggested to be publicly released)

Report of final project might be publicly released, otherwise please specify reasons.

- released in our website
- released in ArXiv if quality is high; do not preprint low-quality paper there. If you cannot access the quality, you could consult your supervisor or the teaching team.

Consider Github to maintain your code (link Github repo in the report) if coding needed

A objective could be **earning 100+ GitHub stars for your repo.**

Consider uploading your model in HuggingFace, serving it there is possible.

https://openai.com/api/ https://opt.alpa.ai

Note: You might get computing resources to train 10B+ model if instructor like your proposal

Example: ArXiv+HF+Github

https://github.com/FreedomIntelligence/MedGen

arXiv:2507.05675v1 [cs.CV] 8 Jul 2025

MedGen: Unlocking Medical Video Generation by Scaling Granularly-annotated Medical Videos

Rongsheng Wang[1]*, Junying Chen[1]*, Ke Ji[1], Zhenyang Cai[1], Shunian Chen[1], Yunjin Yang[1], Benyou Wang[1]†
[1]The Chinese University of Hong Kong, Shenzhen
wangbenyou@cuhk.edu.cn

https://github.com/FreedomIntelligence/MedGen

Link your code here

Abstract

Recent advances in video generation have shown remarkable progress in open-domain settings, yet medical video generation remains largely underexplored. Medical videos are critical for applications such as clinical training, education, and simulation, requiring not only high visual fidelity but also strict medical accuracy. However, current models often produce unrealistic or erroneous content when applied to medical prompts, largely due to the lack of large-scale, high-quality datasets tailored to the medical domain. To address this gap, we introduce **MedVideoCap-55K**, the first large-scale, diverse, and caption-rich dataset for medical video generation. It comprises over 55,000 curated clips spanning real-world medical scenarios, providing a strong foundation for training generalist medical video generation models. Built upon this dataset, we develop **MedGen**, which achieves leading performance among open-source models and rivals commercial systems across multiple benchmarks in both visual quality and medical accuracy. We hope our dataset and model can serve as a valuable resource and help catalyze further research in medical video generation.

⚠ Disclaimer: This paper contains clinical content that may be disturbing to some readers.

1 Introduction

Recent advances in video generation have led to impressive breakthroughs, with models now capable of producing high-quality, cinematic visuals that align closely with user prompts (Blattmann et al. 2023). In particular, latent video diffusion models (LVDMs), such as Sora (OpenAI 2025) and Veo (Sharma et al. 2024), have achieved state-of-the-art performance by operating efficiently in latent space and delivering diverse, coherent video outputs from textual descriptions.

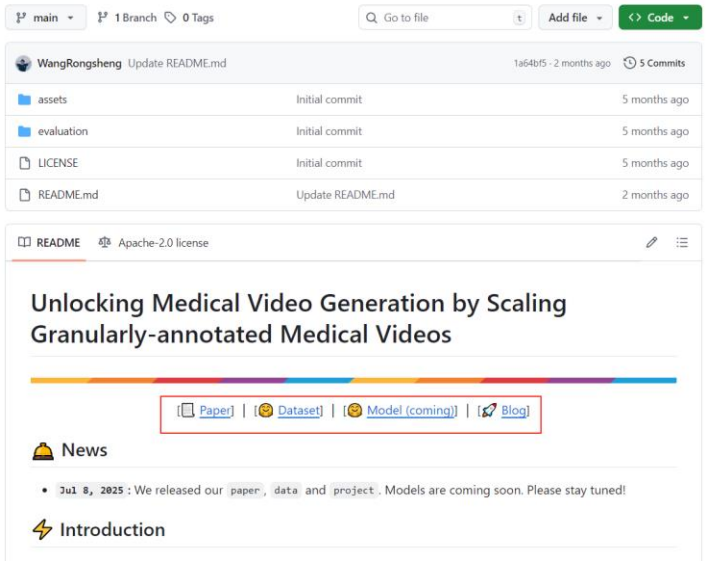
Despite this progress, **medical video generation** remains a largely underexplored yet crucial domain. Medical videos are indispensable in numerous real-world applications, including clinical training, surgical simulation, and patient education (Li et al. 2024b). Unlike everyday video content, medical videos demand precise rendering of anatomical structures, accurate surgical steps, and realistic physiologi-

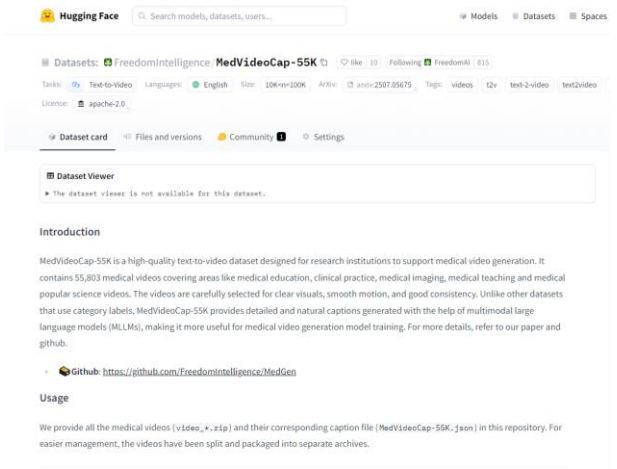
demands on visual fidelity, semantic correctness, and temporal coherence.



Figure 1: Failure cases of Sora, Pika, and Hailuo on medical video generation. (a): Medical common sense errors. (b): Distortion errors.

However, current video generation models are trained almost exclusively on general-purpose datasets that focus on natural scenes and everyday activities (Blattmann et al. 2023). As a result, when applied to medical prompts, they often generate outputs with critical errors—such as anatomical inconsistencies, tool misuse, and implausible clinical scenarios. As shown in Figure 1, even leading models like Sora (OpenAI 2025), Pika (Team 2025a), and Hailuo (MiniMax 2025) fail to maintain basic medical realism, revealing a clear mismatch between training data and medical domain requirements. A key bottleneck lies in the lack of large-scale, high-quality datasets tailored for medical video generation (Sun et al. 2024). Existing medical datasets are limited in size, narrow in scope (e.g., only endoscopic or surgical videos), and mostly provide categorical labels instead of detailed descriptions—making them unsuitable for training or





https://arxiv.org/pdf/2508.18240

Team work

Team size: Students may do final projects sololy, or **in teams of up to 3 people**. Please specify reasons to do project with more than 3 people, we reserve the rights to reject it

Team work is encouraged: We strongly recommend you do the final project in a team. Larger teams are expected to do correspondingly larger projects, and you should only form a 3-person team if you are planning to do an ambitious project where every team member will have a significant contribution.

Contribution: In the final report we ask for a statement of what each team member contributed to the project. Team members will typically get the same grade, but we may differentiate in extreme cases of unequal contribution. **You can contact us in confidence in the event of unequal contribution**.

External collaborators: You can work on a project that has external (non course student) collaborators, but you must make it clear in your final report which parts of the project you are extractly contributing.

Mentors: Every custom project team has a mentor, who gives feedback and advice during the project. A project may have an external (i.e., not in our teaching team) mentor; otherwise, we will assign a TA or any research staff in our research team to custom project teams after project proposals.

Computing resources: We will use Kaggle computing resources, later we will share the detail.

Final project assessment

Final report assessment （45%）

TA review

Poster presentation （10%）

TA/instructor review

Cross review between students (not contributed in assessment)

Submit your **Final project** to a **conference**

- Submission Deadlines
  - ACL ARR (four times per year)
    - October 12
  - Feb. ACL,SIGIR, ICML
  - March, **COLM 2026** (https://colmweb.org/)
  - **May, NeurIPS**

- TAs might help upon their interest. They might also refer you to other students in this group.
- You could also work with external members, please adhere to academic ethics

How to better prepare a final project in NLP?

which proposal could I do for the final project in a NLP course that emphasize large language models?

1/12/2024, 10:16:03 AM

There are numerous exciting project proposals you could consider for an NLP course that emphasizes large language models. Here are a few ideas:

1. **Fine-tuning GPT for a Specific Task**: Many NLP tasks can benefit from the transfer learning capabilities of large language models. You could propose to fine-tune a pre-trained model like GPT-3 on a specific task, such as text summarization, sentiment analysis, or machine translation.

2. **Bias Analysis in Language Models**: Large language models have been criticized for their potential to perpetuate and amplify biases present in their training data. You could propose a project that explores these biases, possibly by creating a dataset of potentially biased statements and evaluating how the model responds to them.

3. **Exploration of Zero-shot Learning**: Large language models like GPT-3 have shown remarkable capabilities in zero-shot learning, where they can perform tasks without any specific fine-tuning. You could explore this aspect further, perhaps by designing a range of tasks and evaluating the model's zero-shot performance.

4. **Multimodal Language Models**: Language models are increasingly being used in conjunction with other types of data, such as images. You could propose a project that explores the effectiveness of large language models in a multimodal context.

5. **Language Model as a Knowledge Base**: Given that large language models have been trained on a vast amount of data, they can be seen as a form of knowledge base. You could explore how effectively a model like GPT-3 can answer questions on a wide range of topics, and how this capability compares to more traditional knowledge bases.

6. **Interactive Fiction with Language Models**: Large language models can generate coherent and contextually appropriate text, which makes them great for generating interactive fiction. You could propose a project to create an interactive story or game using a model like GPT-3.

Remember, the best project will depend on your interests and the specific focus of your course. Be sure to discuss your ideas with your course instructor or advisor to ensure they're appropriate and feasible.

1/12/2024, 10:16:03 AM

Ideas from where?

- Papers from **NLP** conferences
  - https://aclanthology.org/
  - https://aclanthology.org/events/acl-2025/
- Papers from **ML** conference
  - https://openreview.net/group?id=ICLR.cc/2026/Conference#tab-active-submissions
- Papers from **LLM** conference
  - https://colmweb.org/AcceptedPapers.html
- Papers from **HCI** conferences
  - https://hci-deadlines.github.io/?sub=HCI,DES,CSCW,AI,HRI,XR,HAP,VIS,SP
  - https://dl.acm.org/doi/proceedings/10.1145/3706598

Natural language processing (**NLP**) + Large language models (**LLMs**) + human-computer interactions (**HCI**)

Ideas from where?

- Previous **major**
  - **Finance, management, social science?**
- Personal Interest

- Existing resources from your friends
  - Data ?
- Existing use cases
  - ways to connect users

For startups: fired AI for classroom learning

薯塔AI

你的赛博「课件学习」搭子

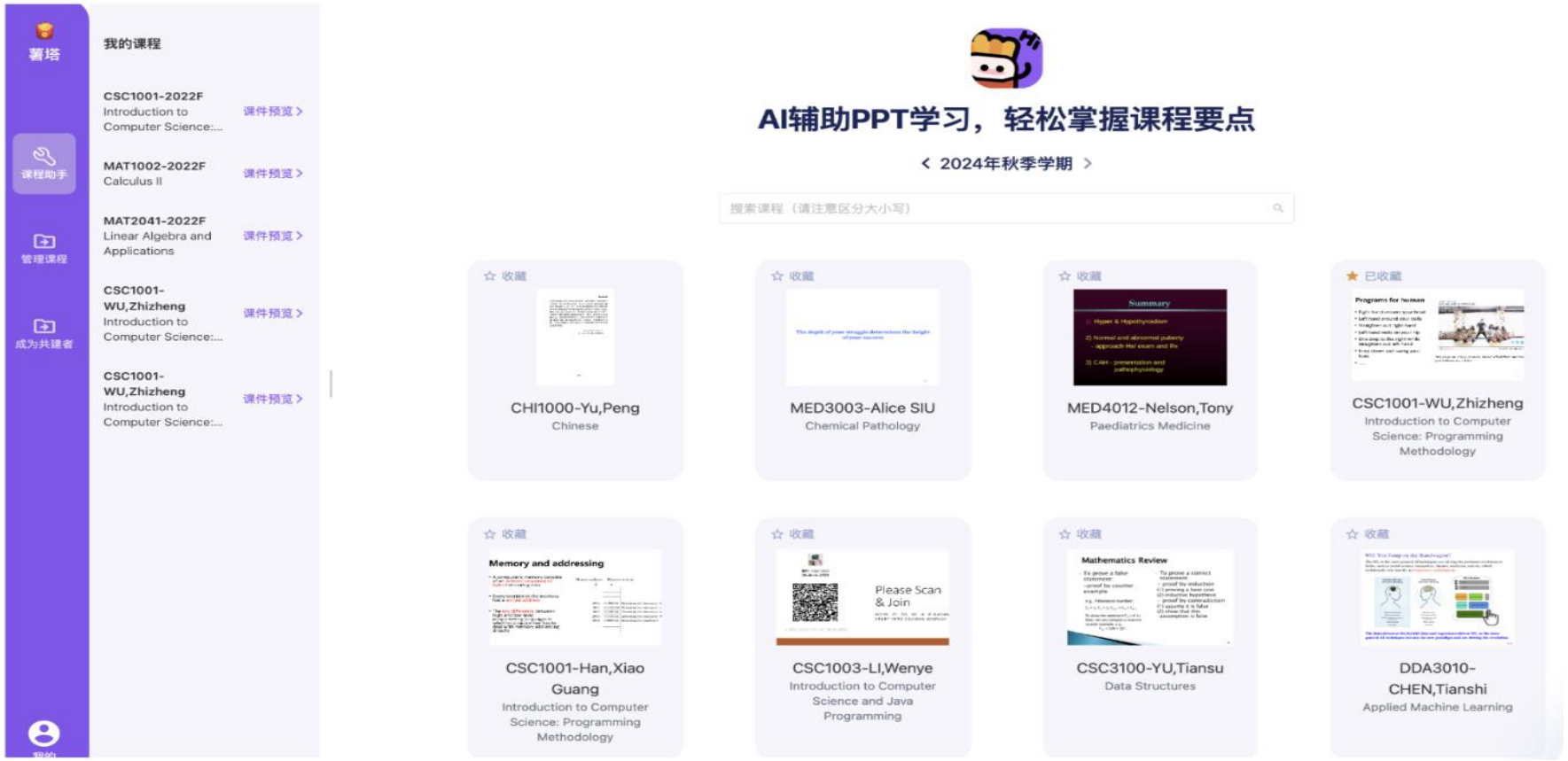
集「课件共享社区、翻译、AI问答、跨文档知识检索、随堂笔记」为一体的AI课件学习助手，高效助力你的课件学习、预习、复习。

立即开始



薯塔AI是一款专为中外合作院校、国际学校及更广泛教育场景设计的智能助学平台。依托先进的AI技术（如大语言模型、智能问答引擎等），薯塔AI为学生提供高效、便捷的学习支持；通过课件共享社区、AI问答、跨文档知识检索等核心功能，薯塔AI为学生提供个性化的学习支持，通过智能化工具增强学习互动和参与感，全面提升学习效率。

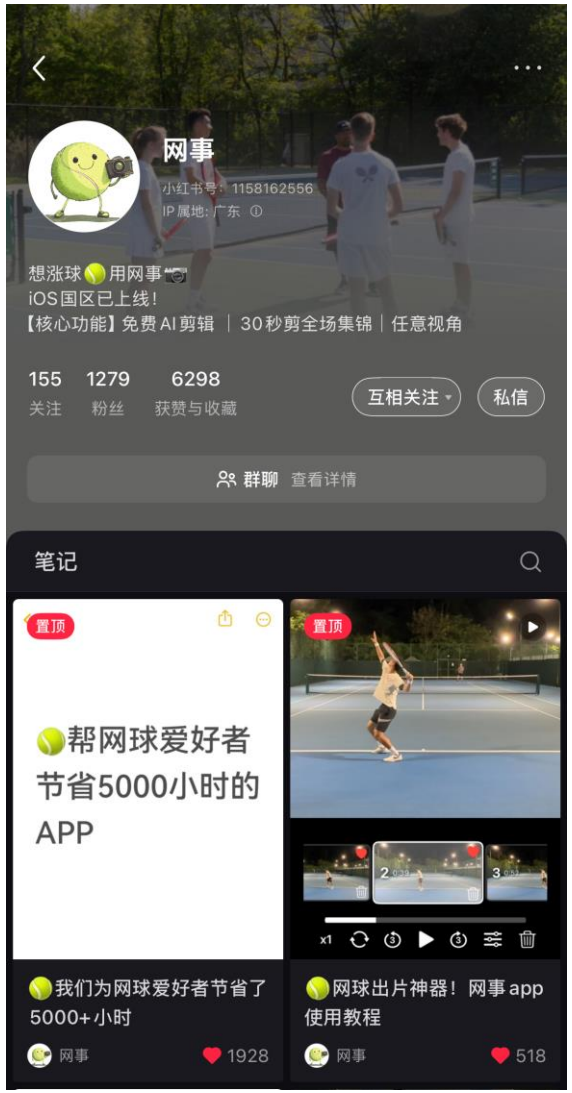
本课题组支持的学生创业团队已获得了超过2000每月活跃用户，服务大部分港中深课程，
正在拓展到其他中外合办高校和国外高校



截至2025年4月12日13点30分，薯塔AI已覆盖港中深**60%**的本科生，**港中深注册人数达3579人，总注册人数达4680人**。其中日活跃用户为1239人，平均每日AI问答数2898条。薯塔AI目前正在逐步推广至其他七所中外合办院校，进一步提升平台的用户覆盖率和影响力。随着平台影响力的扩大，我们将为更多院校的学生享受更高效、更智能的学习体验。

For startups





朱宇轩同学把网球和AI相结合，他开发的网事APP已服务一千多个用户，为网球爱好者节省超过5000个小时

Define the learning objectives:

- **Knowledge**: a) Students will understand basic concepts and principles of NLP; b) Students could effectively use NLP for daily study, work and research; and c) Students will know which tasks NLP are suitable to solve and which are not.
- **Skills**: a) Students could train NLP models following a complete pipeline and b) Students could call ChatGPT API for daily usage in study, work and research.
- **Valued/Attitude**: a) Students will appreciate the importance of data; b) Students will tend to use data-driven paradigm to solve problems; and c) Students will be aware of the limitations and risks of using ChatGPT.

Select appropriate resources:

- **Recent ArXiv papers**
  - (People share daily ArXiv papers in Twitter)
- **GitHub**
  - (popular GitHub means a lot)
- **HuggingFace**
  - (New models and datasets)
  - https://huggingface.co/papers (hot papers here)
- **Blogs**
  - (from Open AI and famous guys, Lilian Weng, Yao Fu, Jianlin Su)

Design engaging lectures:

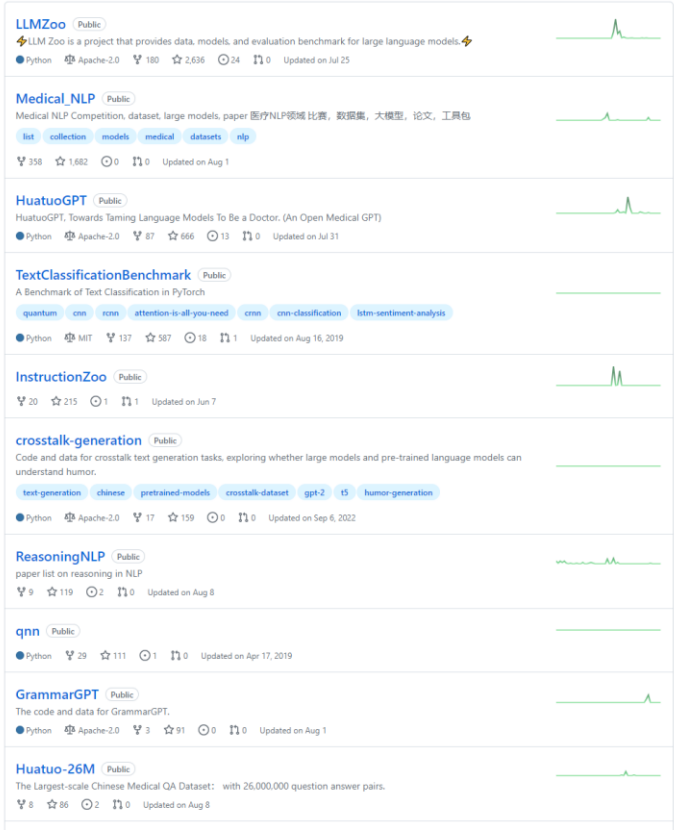
- Discussions in the end of each lecture
- In-class presentation
- Interrupting me whenever needed
- **Make friends with the instructor and TAs**

Provide hands-on practice:

Github Repositories

- **nanoGPT** https://github.com/karpathy/nanoGPT
- **minGPT** https://github.com/karpathy/minGPT
- **Llama2.c** https://github.com/karpathy/llama2.c
- **TinyLLaMA https://github.com/eivindbohler/tinyllama**

- HautuoGPT
- GPT API
- LLMZoo
- LLMFactory



https://github.com/orgs/FreedomIntelligence

Foster collaboration and discussion:

- You own the copyright of your own project if our teaching team do not have a substantial contribution. Otherwise please acknowledge us.
- You are welcome to have discussions with our teaching team.
- Students are encouraged for collaboration and discussions between yourself.

Seek feedback and iterate this course:

- Tell us if you have any suggestions about this course
- We will continue polishing this course.

What is in our course

- Very basics of NLP (most old NLP techniques are not that practical now)
- Large Langauge models (training and beyond)
- Data engineering
- Prompt engineering
- NLP applications
- Future tendency NLP

What is not in our course

- How to do basic coding
- Machine learning
- How to understand the mathematical machanmisn of NLP models

- *SORA might not introduced*

How not to fail in this course

- Finish all assignments, the poster and final report!
- Be honest even if you get negative results
- Be nice to your teammates, classmates, and TAs

How to success in this course

- Start final projects as earlier as possible.
- Actively talk to the instructor and TAs frequently
  - You could talk to us after the lecturing time, e.g. after 5:00 pm
- Bring your personal expertise/interest/user/data and combine it with AI
- Aim higher
  - A **high-quality paper** that help you get a nice PhD offer or internship
  - Run a **startup**; find (future) co-funders

You might obtain from a great project

- Recommendation for internship from Qwen, Tencent, Huawei, DeepSeek, Kimi
  - TAs interned or are interning in MicroSoft, Qwen, Tencent, Kimi, MiniMax, Huawei etc.



Tencent, MiniMax, Kimi





Huawei, Tencent, DJI



Top VC, 蚂蚁医疗

- Recommendation Letter for PhD

- Possibility to join our group as a PhD
  - 蔡镇阳 （本组转博）
  - **Zhenyang Cai**, Junying Chen, Rongsheng Wang, Weihong Wang, Yonglin Deng, Dingjie Song, Yize Chen, Zixu Zhang, Benyou Wang. Exploring Compositional Generalization of Multimodal LLMs for Medical Imaging. ACL main conference 2025.
  - 曾子懿 （本组转博）
  - **Ziyi Zeng**, Zhenyang Cai, Yixi Cai, Xidong Wang, Junying Chen, Rongsheng Wang, Siqi Cai, Benyou Wang., Zhiguo Zhang, Haizhou Li. WaveMind: a Conversational EEG Foundation Model Aligned to Textual and Visual Modalities. Submitted to AAAI 2025
  - 洪畅（潜在转博）
  - **Chang Hong**, Minghao Wu, Qingying Xiao, Yuchi Wang, Xiang Wan, Guangjun Yu, Benyou Wang, Yan Hu. Towards Assessing Medical Ethics from Knowledge to Practice. submitted to KDD. https://arxiv.org/pdf/2508.05132

Contents

- **About me**
- **Philosophy of this course**
- **NLP and large language models (LLM)**
- **Introduction to our research**

Natural Language Processing

Applications

- Machine Translation
- Information Retrieval
- Question Answering
- Dialogue Systems
- Information Extraction
- Summarization
- Sentiment Analysis
- ...

Core Technologies

- Language modeling
- Part-of-speech tagging
- Syntactic parsing
- Named-entity recognition
- Word sense disambiguation
- Semantic role labeling
- ...

NLP lies at the intersection of computational linguistics and machine learning.

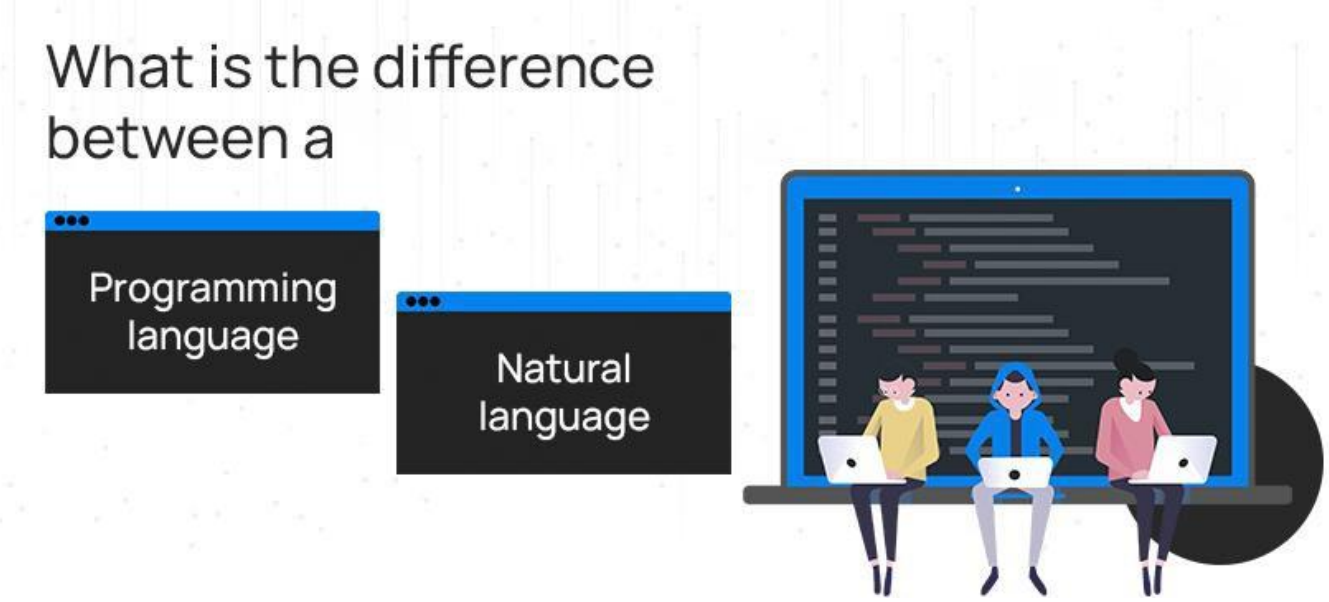
What is NLP?

Wikipedia: Natural language processing (NLP) is a field of computer science, artificial intelligence, and computational linguistics concerned with the interactions between computers and human (**natural**) language.



https://uclanlp.github.io/CS269-17

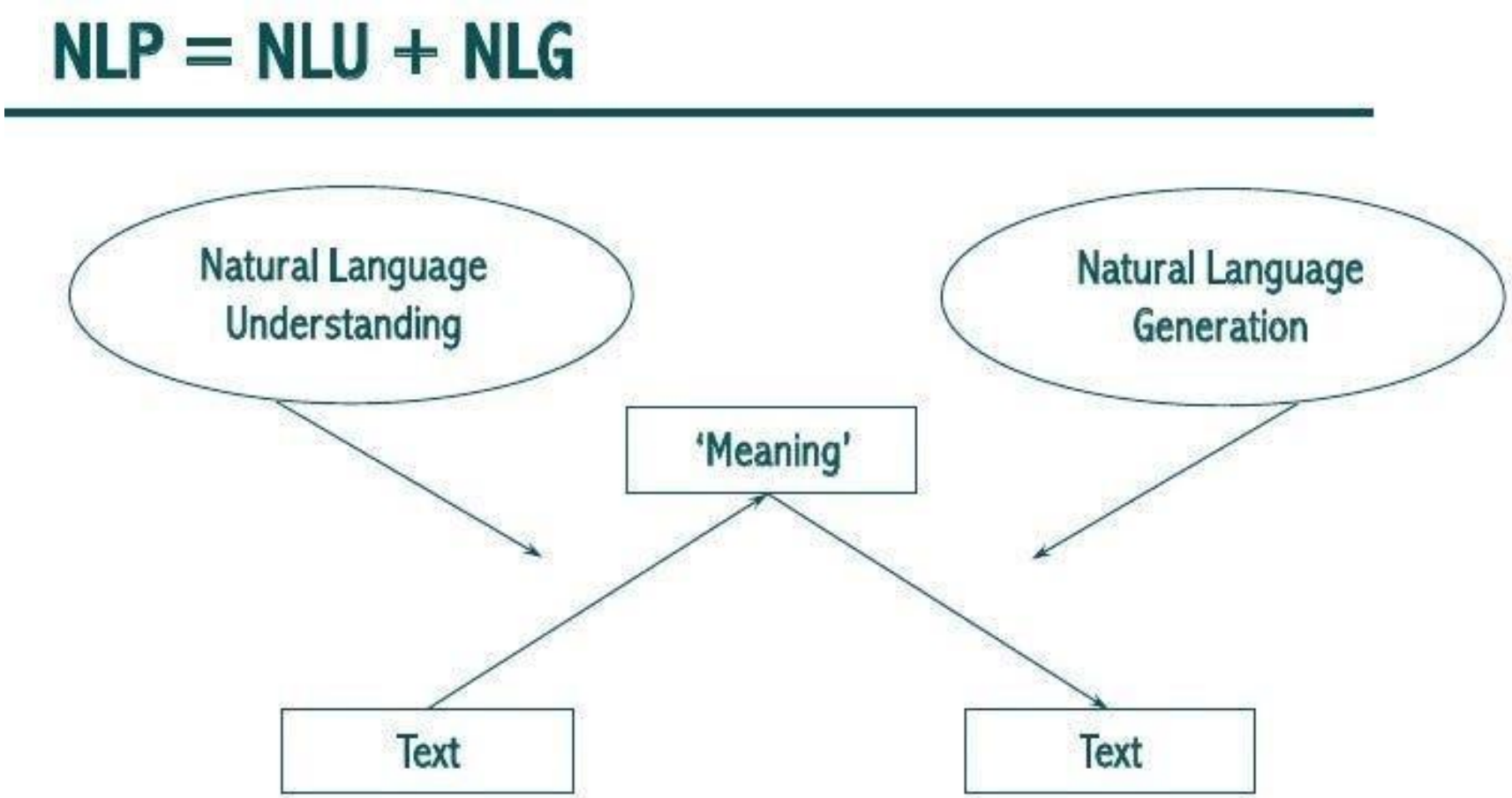
Why natural language instead of language



But in the large language models era, NLP also deals with programing language (see Prof. Pinjia He)

What is NLP?

NLP — Natural Language "Processing" **NLU** — Natural Language "Understanding" **NLG** — Natural Language "Generation"



https://towardsdatascience.com/nlp-vs-nlu-vs-nlg-know-what-you-are-trying-to-achieve-nlp-engine-part-1-1487a2c8b696

Go **beyond** the keyword matching





- Identify the **structure** and meaning of words, sentences, texts and conversations
- Deep understanding of broad language
- NLP is all around us

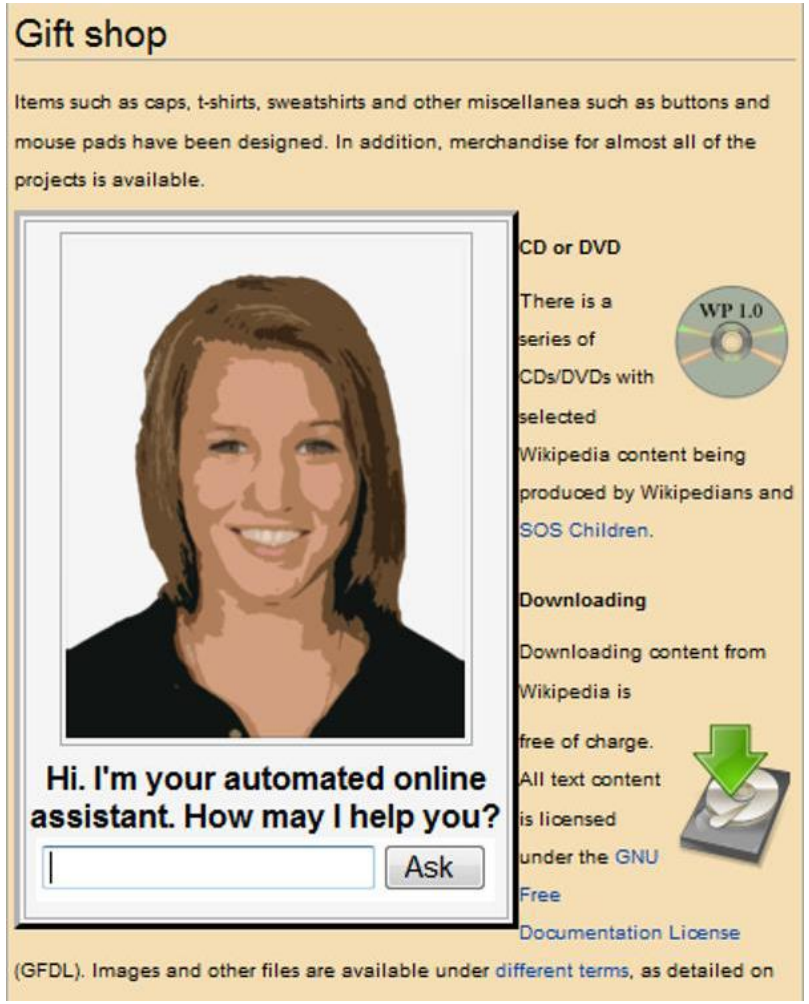
Machine translation



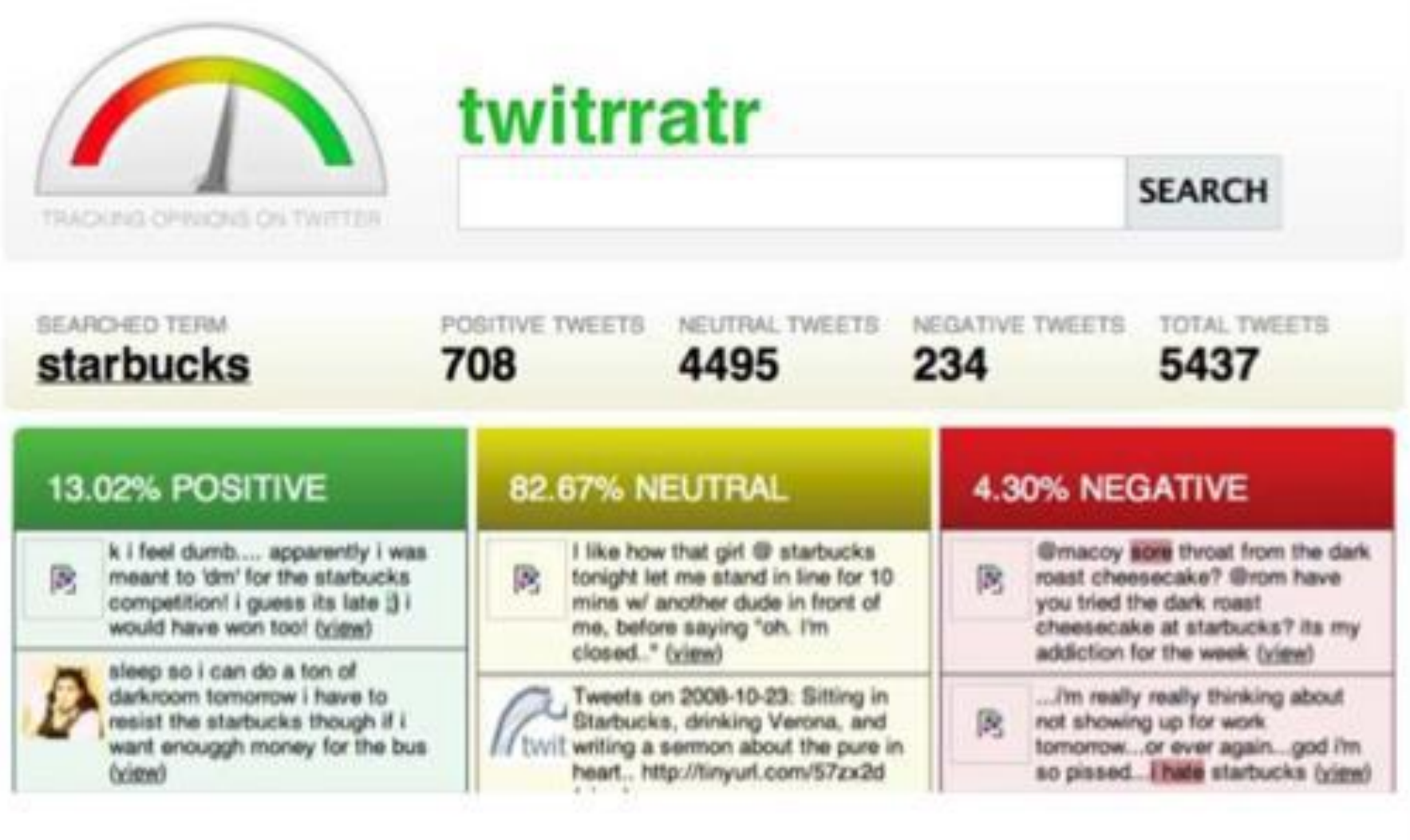




Dialog Systems



Sentiment/Opinion Analysis

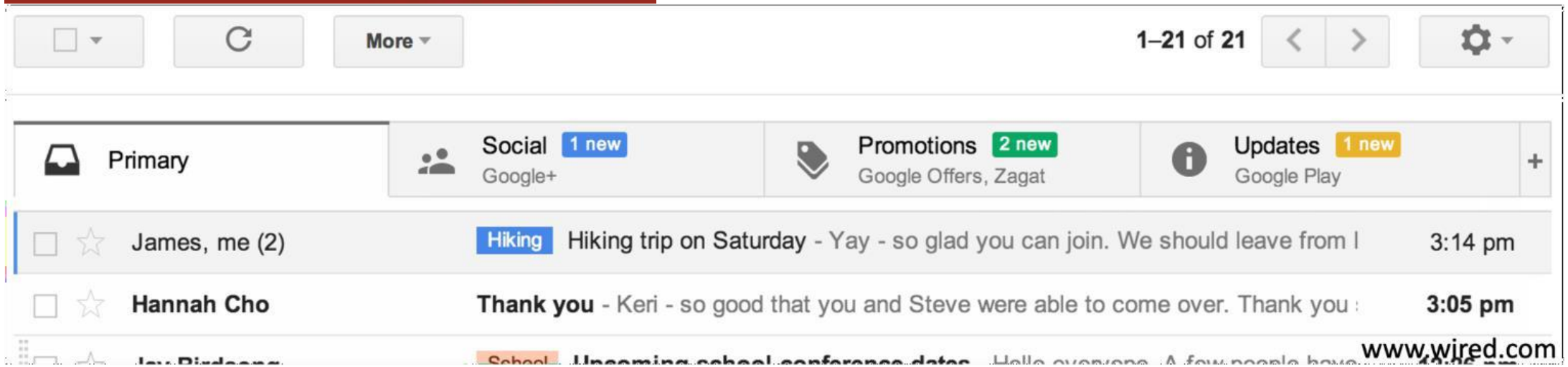


Text Classification

Other applications?







Question answering



- What does “divergent” mean?
- What year was Abraham Lincoln born?
- How many states were in the United States that year?
- How much Chinese silk was exported to England in the end of the 18th century?
- What do scientists think about the ethics of human cloning?

Question answering

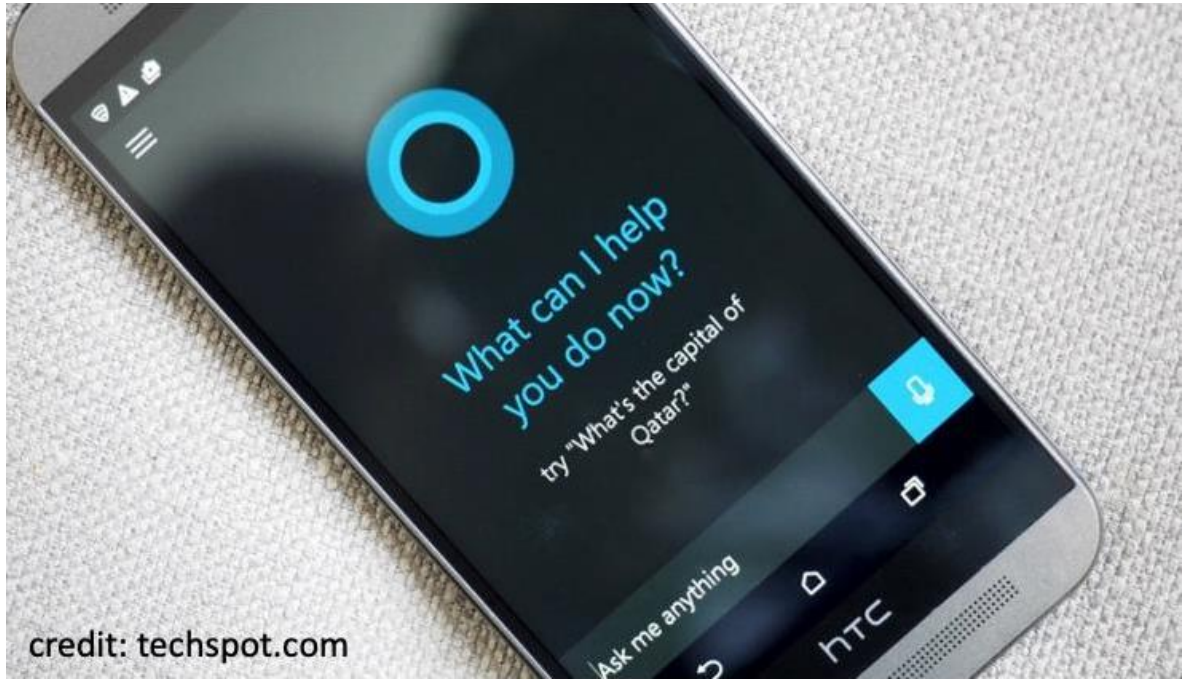
Go beyond search



Natural language instruction

Digital personal assistant



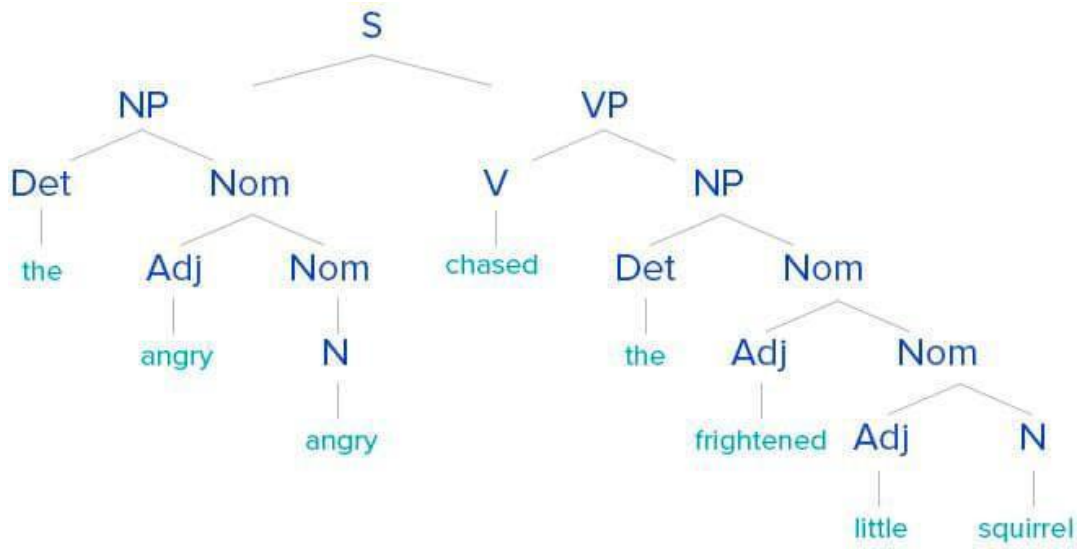


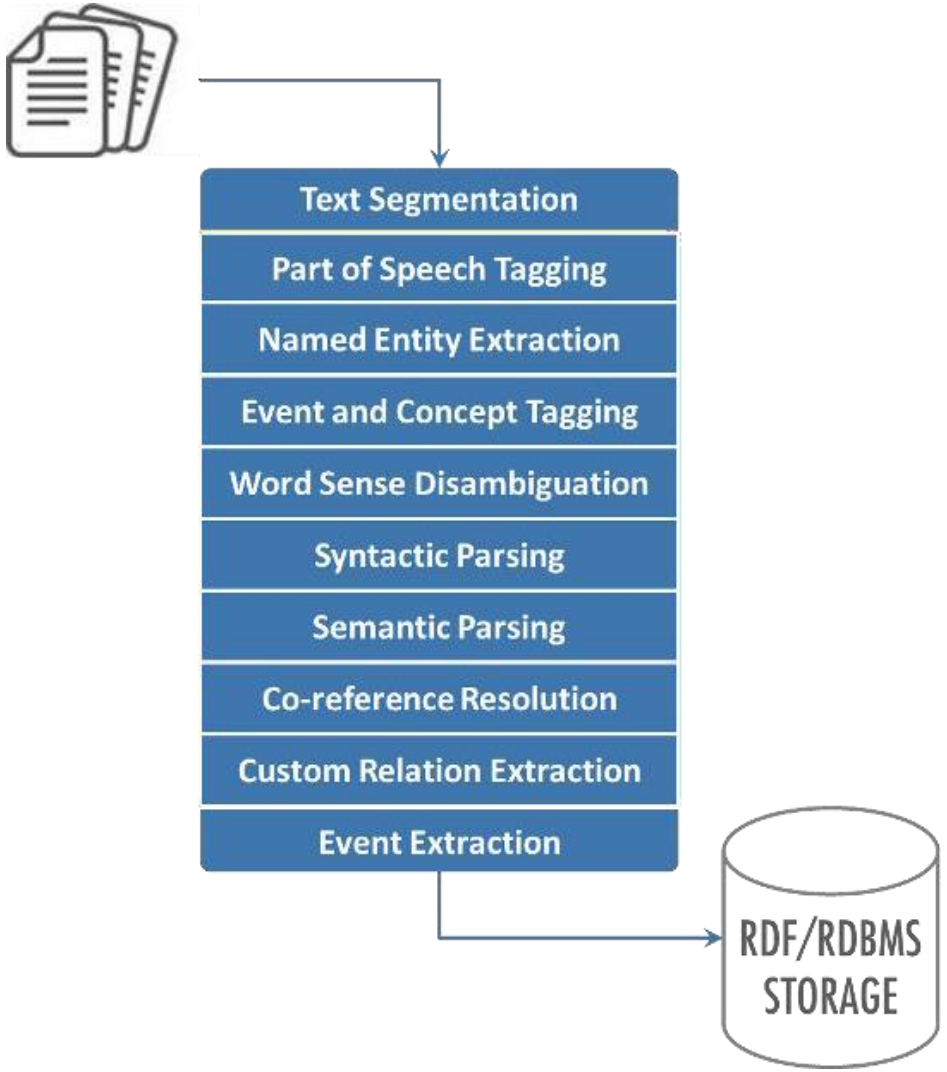
credit: techspot.com

The NLP would be much broader!

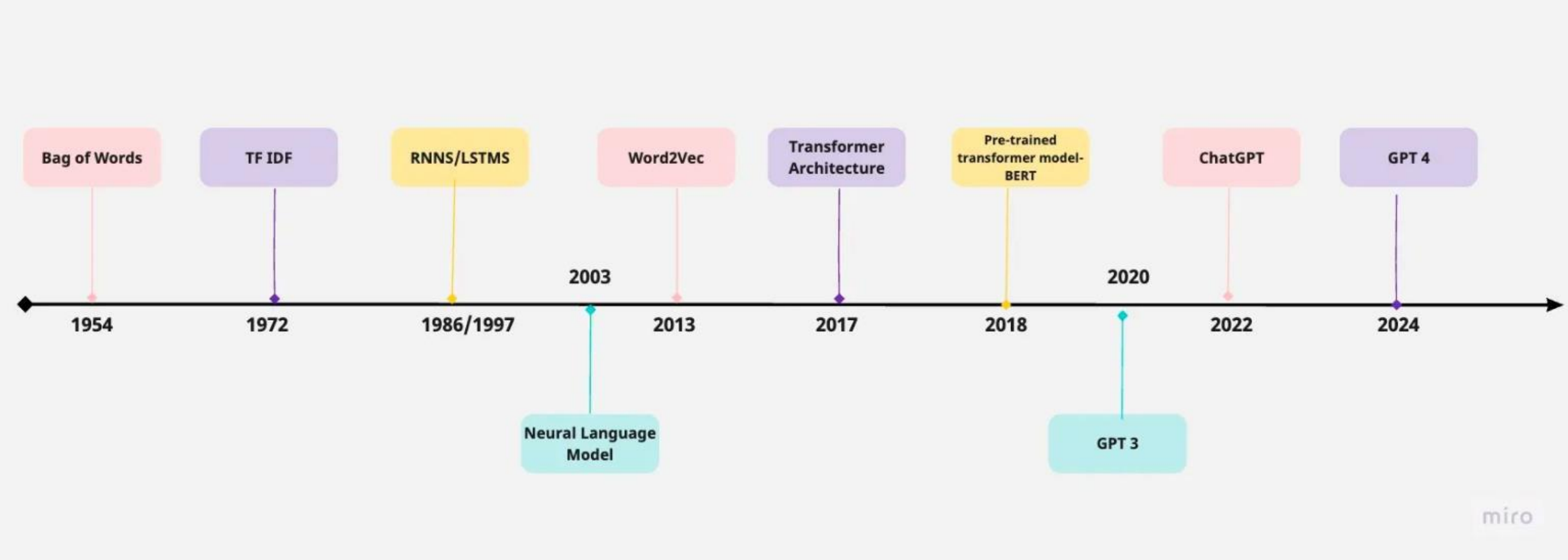
Pre-transformer era of NLP





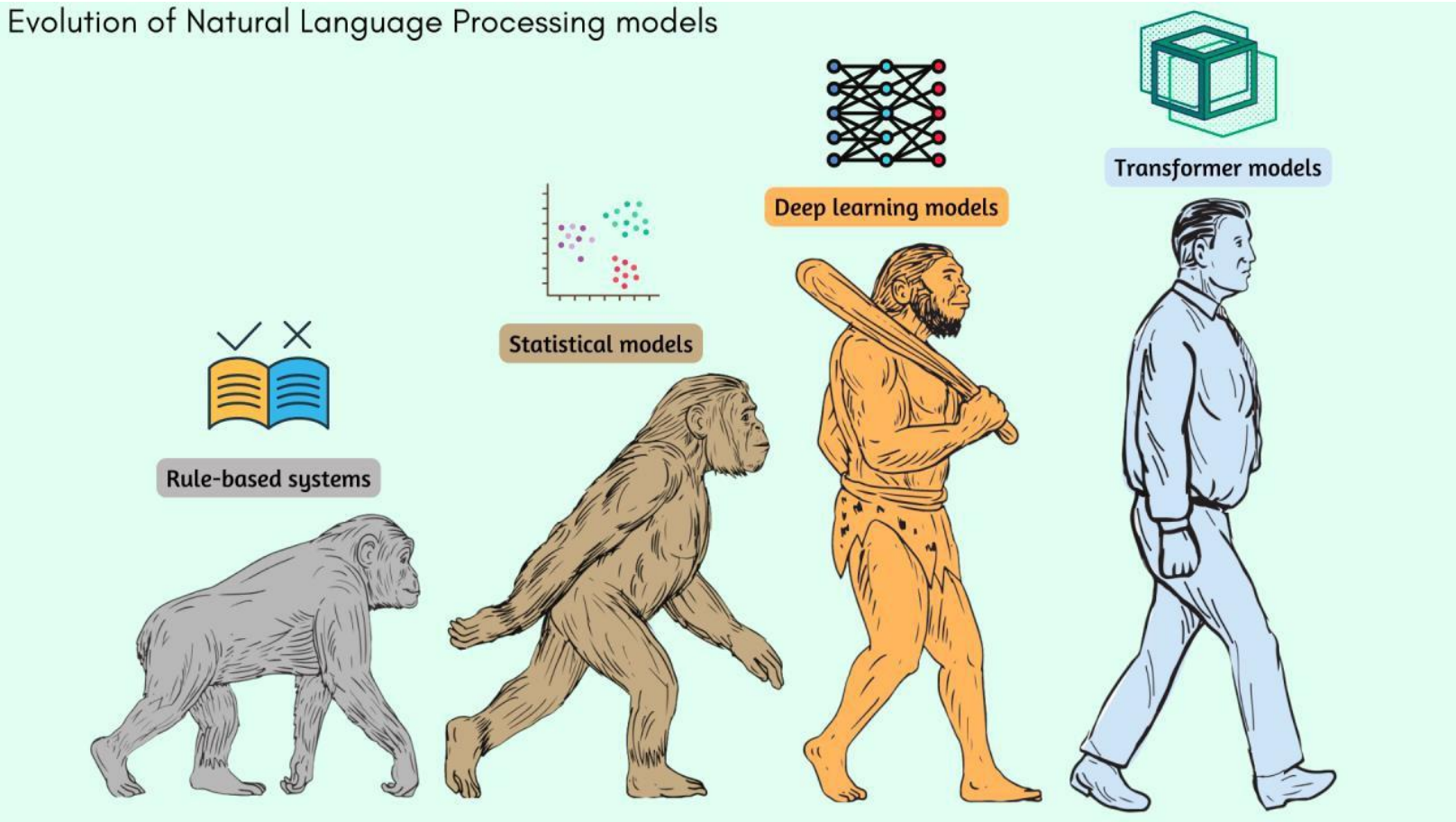


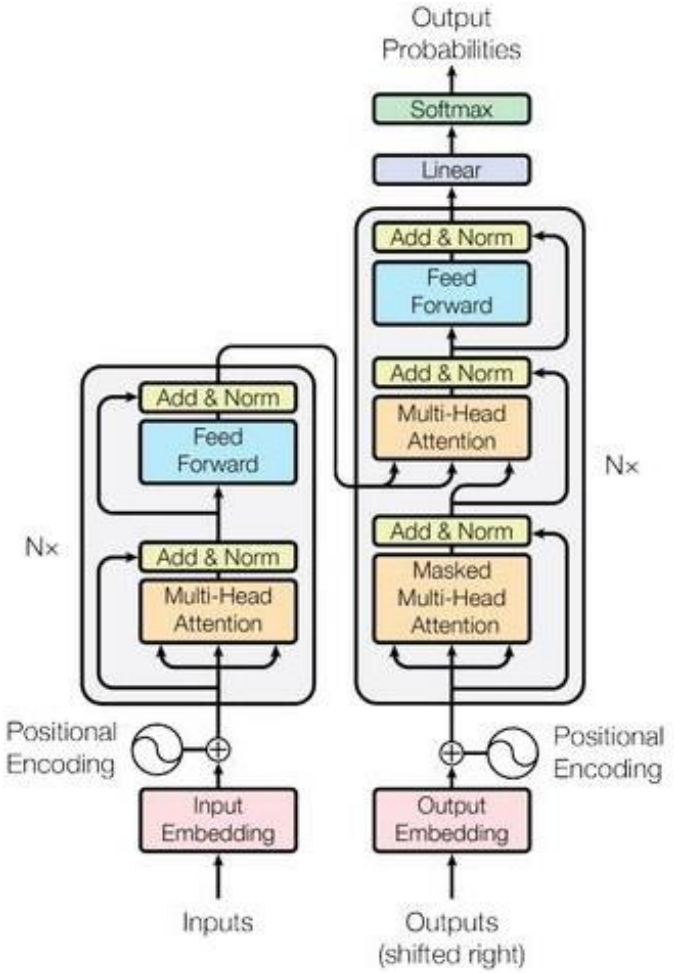
A brief timeline of the evolution of NLP



https://medium.com/pal4ai/the-pre-transformer-era-of-nlp-part-1-2103eff9744

Transformer era of NLP





https://www.linkedin.cn/incareer/pulse/from-rulesets-transformers-journey-through-evolution-sota-yeddula

Where Are We Now?

ChatGPT-5

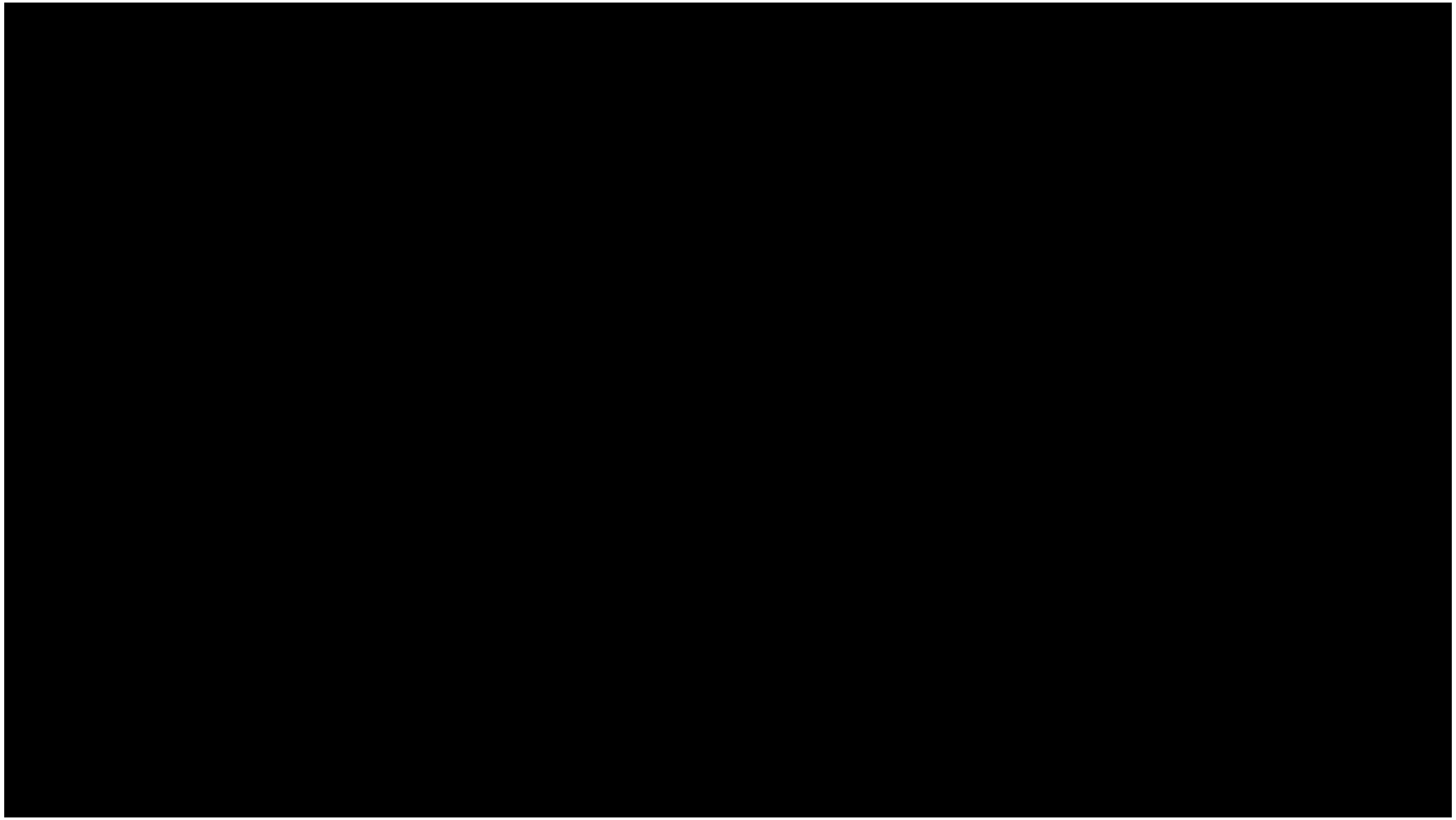


VS





https://www.douyin.com/video/7461889667558886690



Personal assistant: Her

Watch this movie : https://www.imdb.com/video/vi4112492569/

Her with video generation



Human



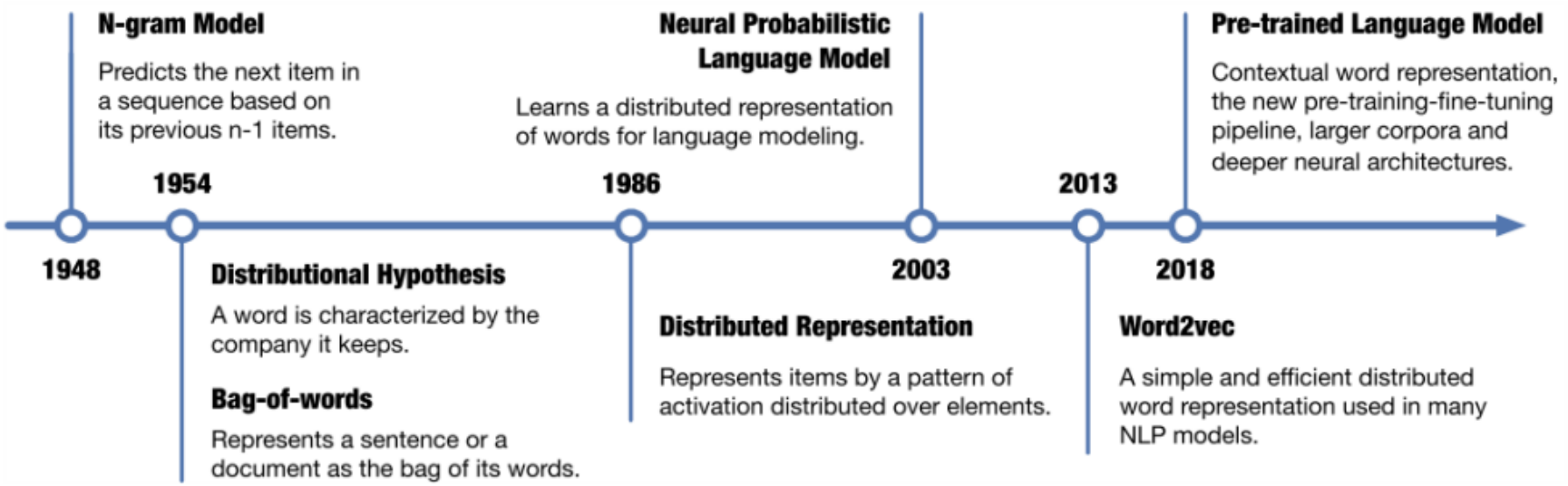
Agent (generated)

https://aubrey-ao.github.io/BodyOfHer/

From old NLP to Large Language models (LLMs)?

Background

- language model



Liu et al., Representation Learning for Natural Language Processing, Springer, 2020

What is language modeling?

A **language model** assigns a probability to a N-gram

$$f: V^n \rightarrow R^+$$

What is language modeling?

A **language model** assigns a probability to a N-gram

$$f: V^n \rightarrow R^+$$



Sfklkljf fskjhfkjsh kjfs fs kjhkjhs fsjhfkshkjfh

Low probability



ChatGPT is all you need

high probability

What is language modeling?

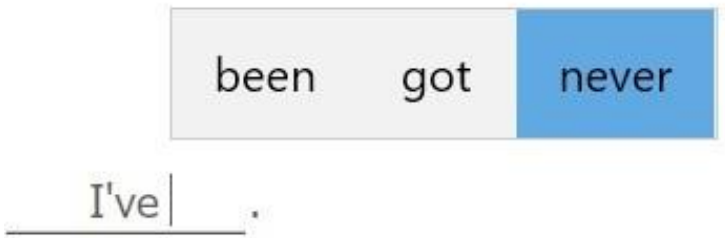
A **language model** assigns a probability to a N-gram

$$f: V^n \rightarrow R^+$$

A **conditional language model** assigns a probability of a word given some conditioning context

$$g: (V^{n-1}, V) \rightarrow R^+$$

And $p(w_n|w_1 \cdots w_{n-1}) = \frac{f(w_1 \cdots w_n)}{f(w_1 \cdots w_{n-1})}$



What is language modeling?

A **language model** assigns a probability to a N-gram

$$f: V^n \rightarrow R^+$$

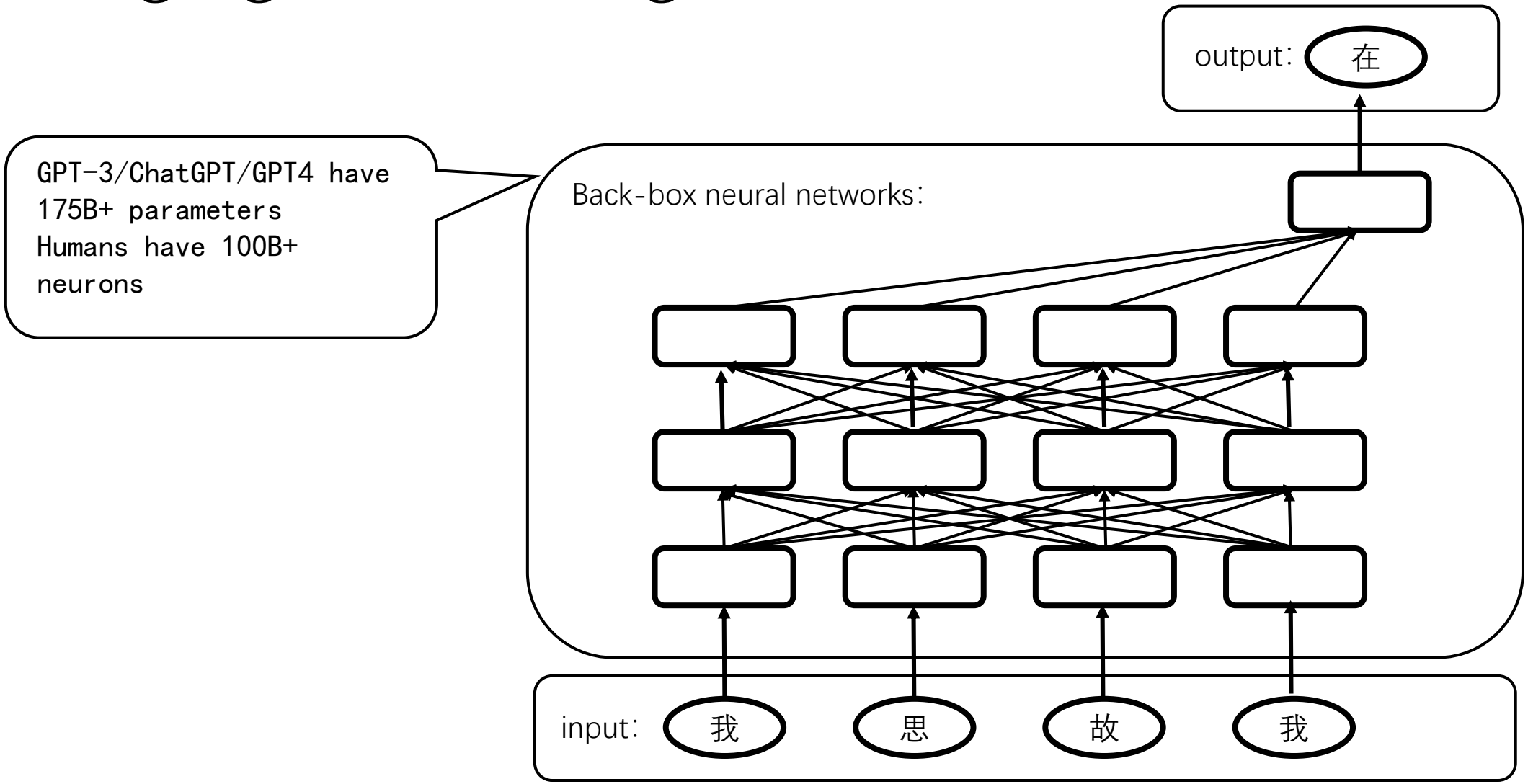
A **conditional language model** assigns a probability of a word given some conditioning context

$$g: (V^{n-1}, V) \rightarrow R^+$$

And $\boldsymbol{p(w_n|w_1 \cdots w_{n-1})} = g(w_1 \cdots w_{n-1}, w) = \frac{f(w_1 \cdots w_n)}{f(w_1 \cdots w_{n-1})}$

$\boldsymbol{p(w_n|w_1 \cdots w_{n-1})}$ is the foundation of **modern large language models** (GPT, ChatGPT, etc.)

Language model using neural networks



Language models: Narrow Sense

A probabilistic model that assigns a probability to every finite sequence (grammatical or not)

Sentence: “the cat sat on the mat”

$$P(\text{the cat sat on the mat}) = P(\text{the}) * P(\text{cat}|\text{the}) * P(\text{sat}|\text{the cat}) * P(\text{on}|\text{the cat sat}) * P(\text{the}|\text{the cat sat on}) * P(\text{mat}|\text{the cat sat on the})$$

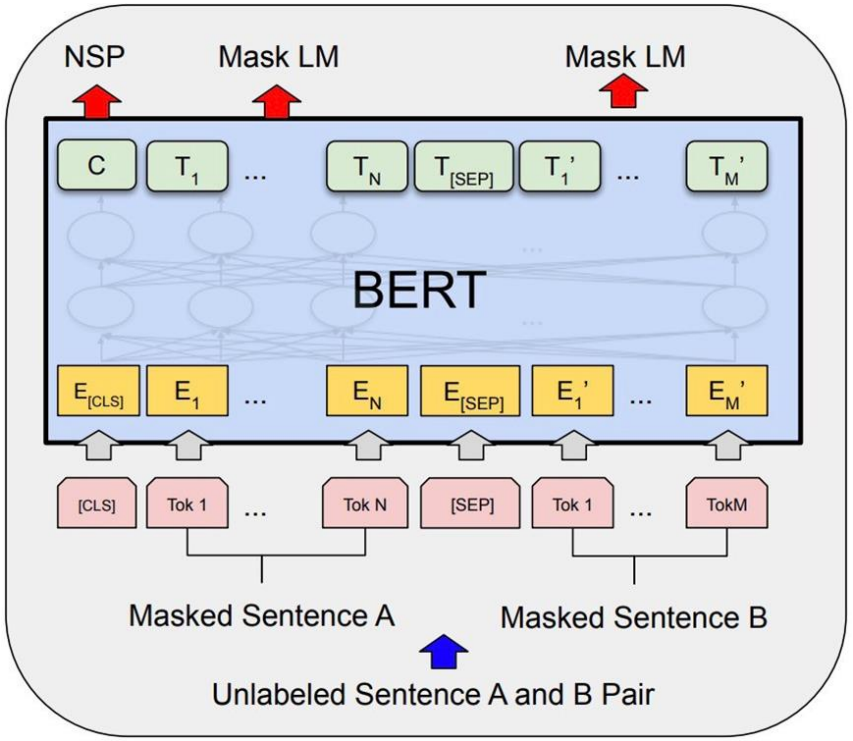
Implicit order

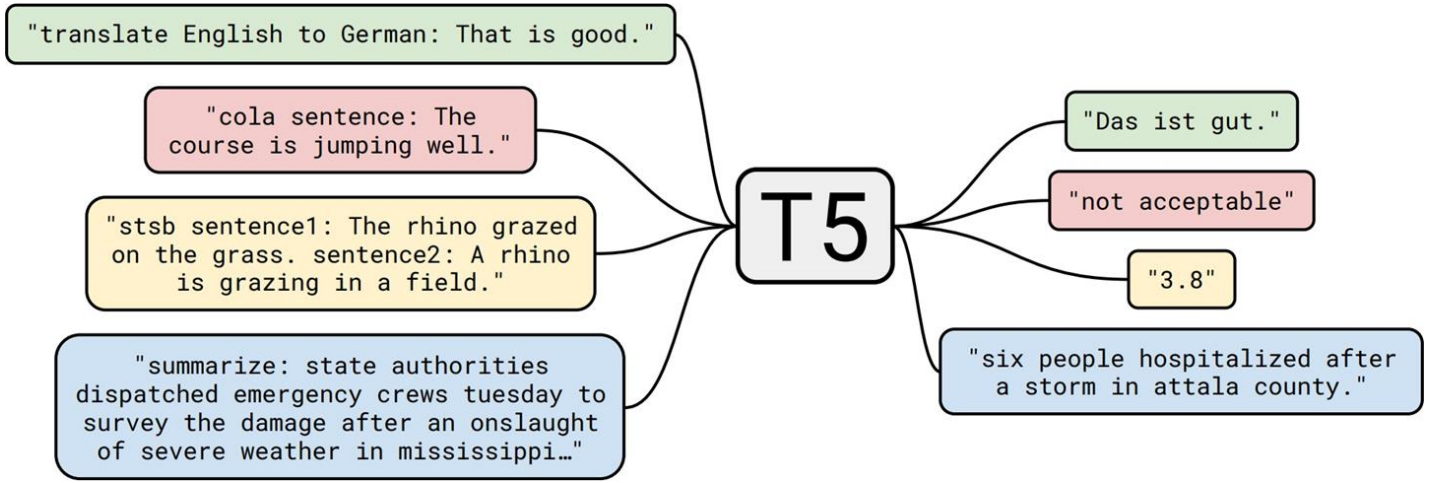
GPT-3 still acts in this way but the model is implemented as a very large neural network of 175-billion parameters!

Language models:Broad Sense

- Decoder-only models (GPT-x models)
- Encoder-only models (BERT, RoBERTa, ELECTRA)
- Encoder-decoder models (T5, BART)

The latter two usually involve a different **pre-training** objective.

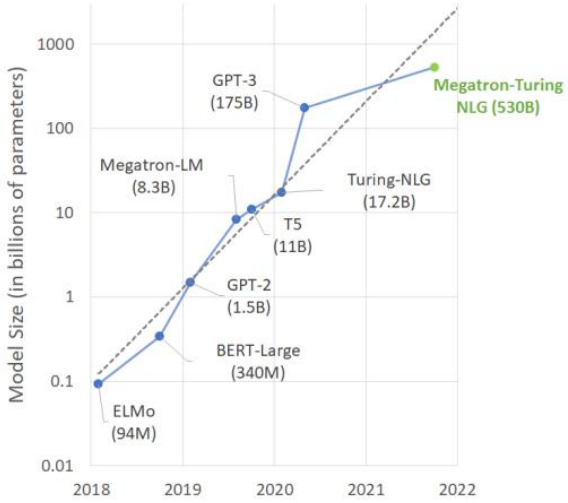




PLM vs. LLM

We do not explicitly mention pre-training because pre-training and training use the same language models objective (e.g., autoregressive generation)

- Pre-trained language model
- **Large** ~~pre-trained~~ **L**anguage **M**odel (LLM)



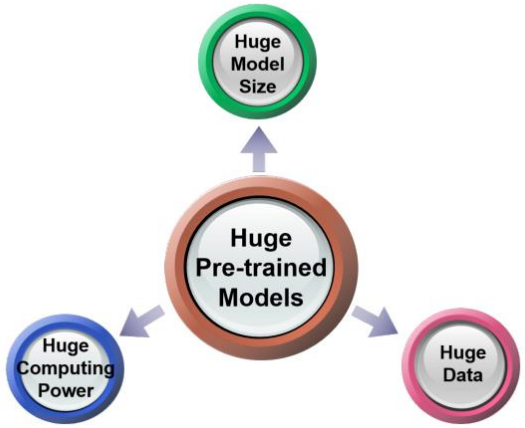
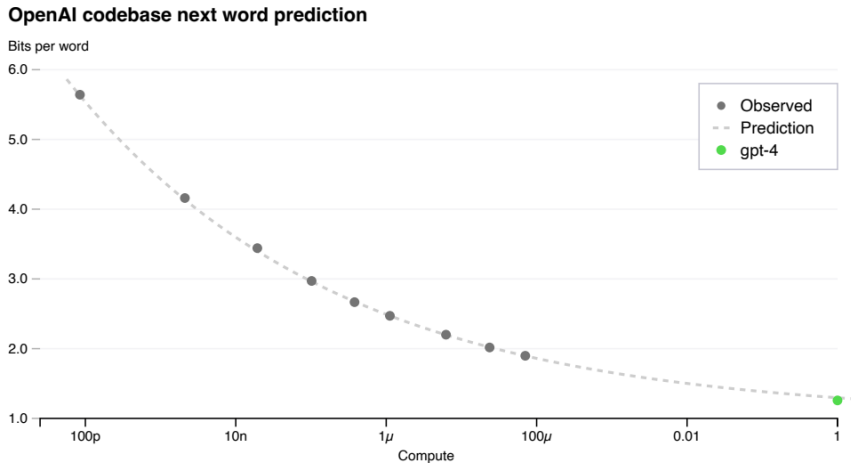


Image source: https://liuquncn.github.io/talks/20220228-Thai-AI-engineer-group/Huge-Pre-trained-Language-Models.public.pdf

Some basics for large language models

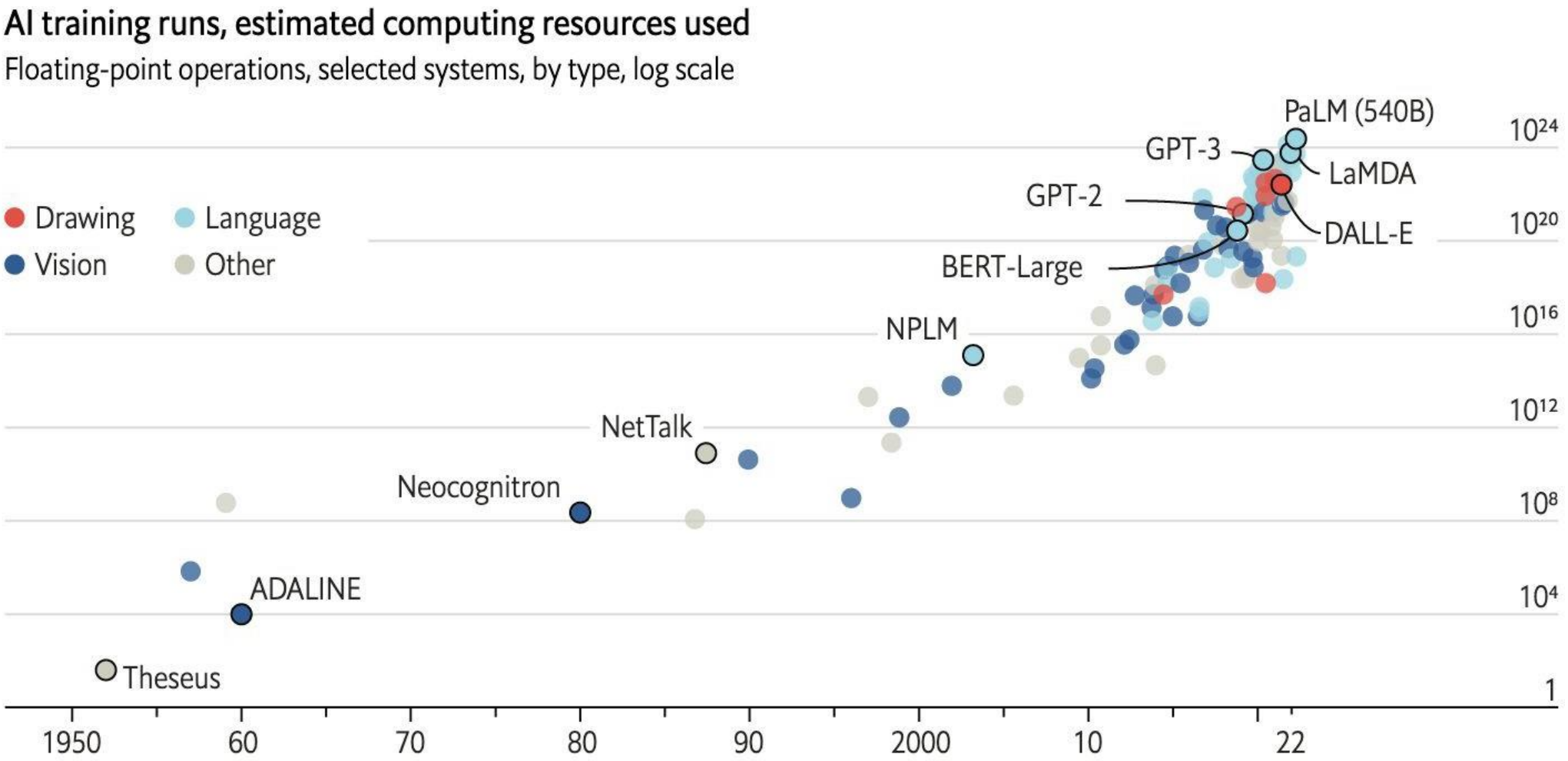
- Scalable network **architecture** (Transformer vs. CNN/RNN)
- Scalable **objective** (**conditional**/auto-regressive LM vs. Masked LM)



- Scalable **data** (plain texts are everywhere vs. supervised data)
  - https://github.com/esbatmop/MNBVC

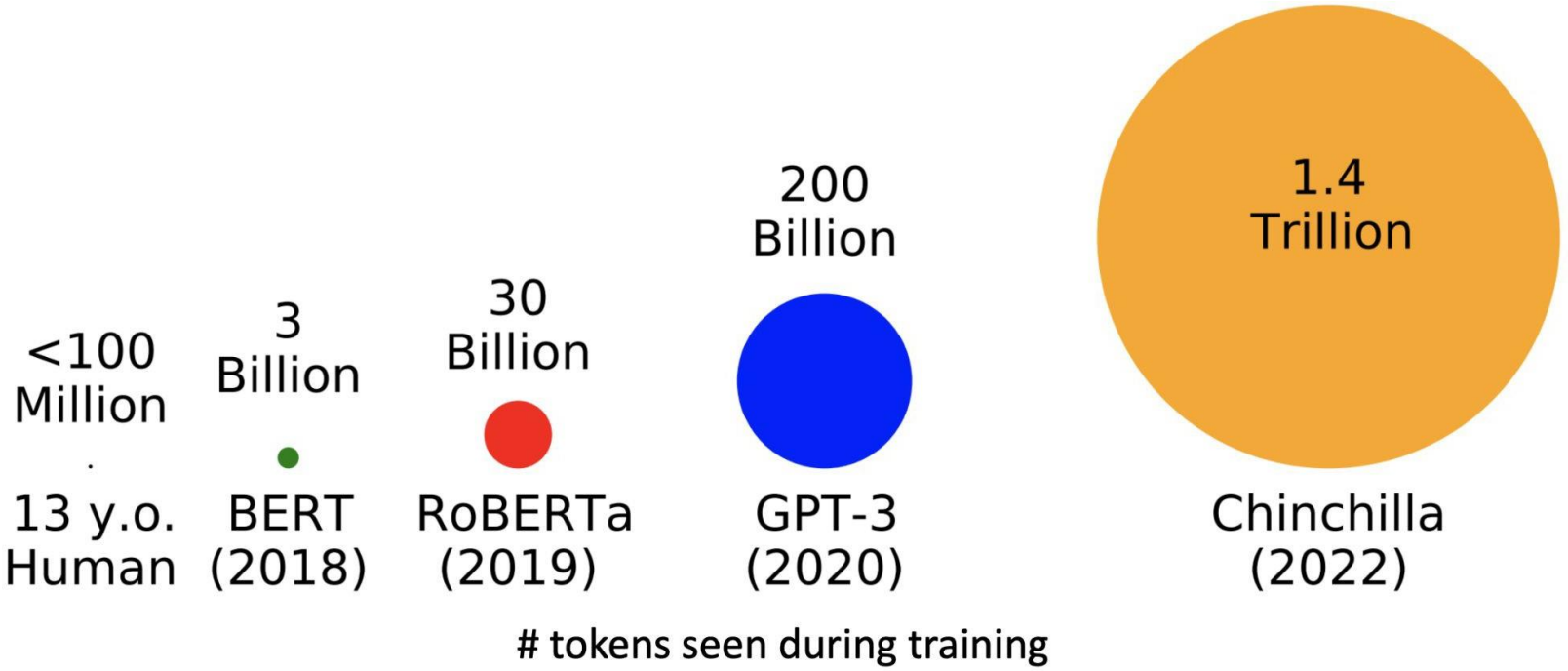
OpenAI, GPT-4 Technical Report, https://cdn.openai.com/papers/gpt-4.pdf

Large Language Models



https://web.stanford.edu/class/cs224n/slides/cs224n-2023-lecture11-prompting-rlhf.pdf 1 yotta = 10^{24} **FLOPs: floating point operations**

Large Language Models - **Hundreds of Billions of Tokens**



36 trillion tokens

Qwen3 Technical Report

https://babylm.github.io/

How Large are “Large” LMs?

- Today, we mostly talk about two camps of models:
  - Medium-sized models: BERT/RoBERTa models (100M or 300M), T5 models (220M, 770M, 3B)
  - “Very” large LMs: models of 100+ billion parameters

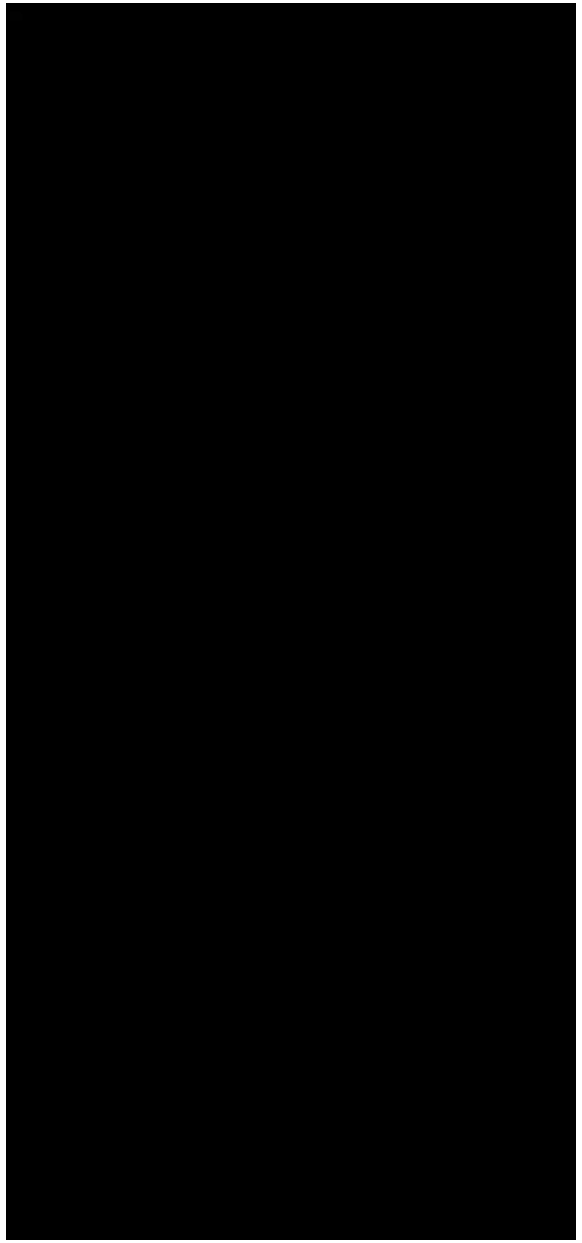
- Larger model sizes larger compute, more expensive during inference
- Different sizes of LMs have different ways to adapt and use them
  - Fine-tuning, zero-shot/few-shot prompting, in-context learning…
- Emergent properties arise from model scale
- Trade-off between model size and corpus size



10T for Fable 5 ?

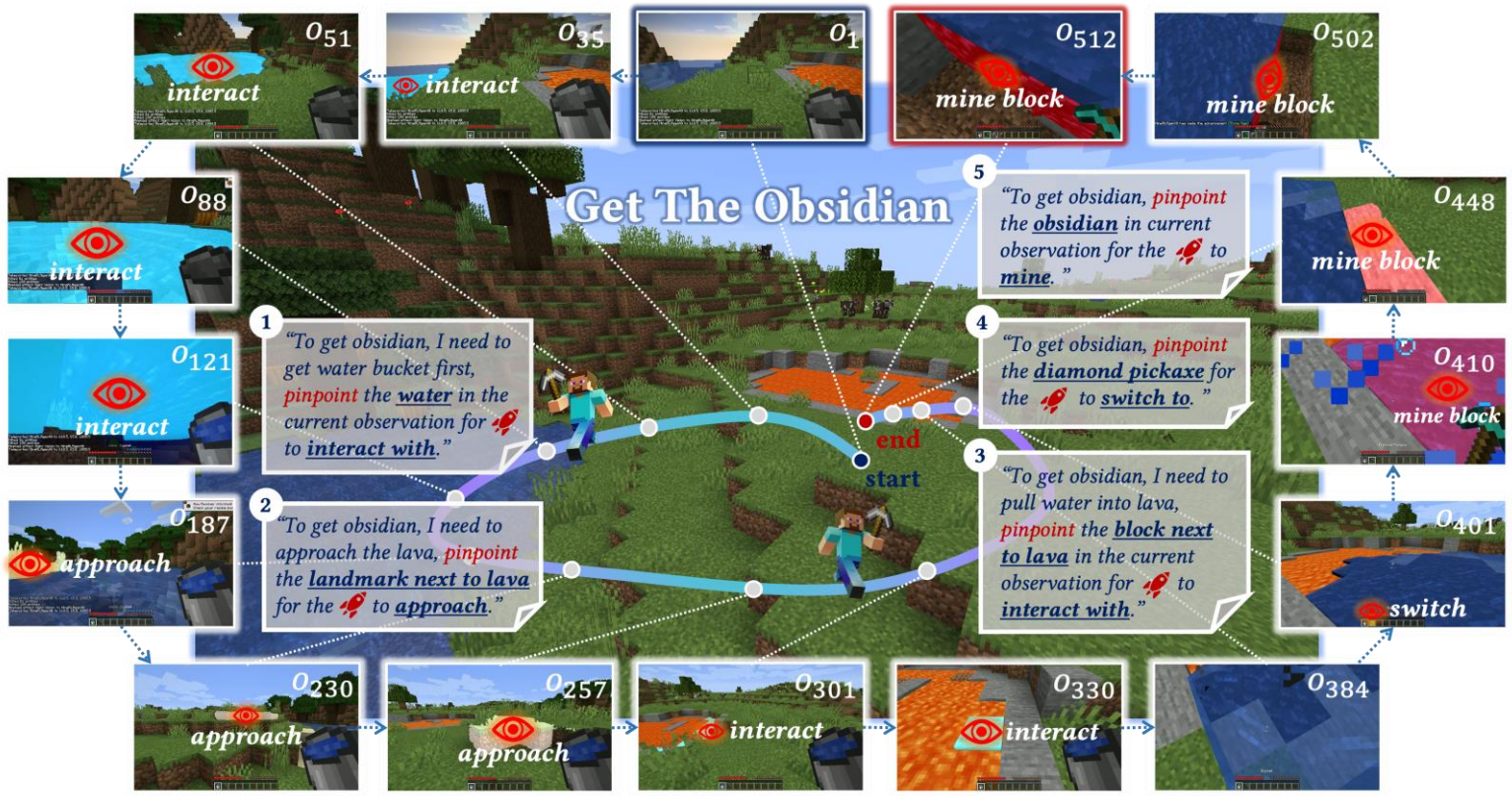
From LLMs to **Agent**

Mobile Agent



https://github.com/X-PLUG/MobileAgent

Game Agent



https://craftjarvis.github.io/ROCKET-1/

Shaofei Cai, Zihao Wang, Kewei Lian, Zhancun Mu, Xiaojian Ma, Anji Liu, Yitao Liang. ROCKET-1: Master Open-World Interaction with Visual-Temporal Context Prompting. https://arxiv.org/abs/2410.17856

Embodied AI



https://mobile-aloha.github.io/

Contents

- **About me**
- **Philosophy of this course**
- **NLP and large language models (LLM)**
- **Some Research**

Example I：HuatuoGPT

- 2023 May, **HuatuoGPT**: Taming Language Models to Be a Doctor. (Findings of EMNLP 2023.) 1k GitHub stars
- 2023 May, Huatuo-26M: the largest medical QA **dataset.** (Findings of NAACL 2023).
- 2023 Nov, **HuatuoGPT-II**: One-stage Training for Medical Adaptation of LLMs. (COLM 2024). 1k GitHub stars
- 2024 March, **Apollo: multilingual** medical LLMs for 6 languages
- 2024 June, **HuatuoGPT-Vision**: **Multimodal (visual)** Medical LLMs. (EMNLP 2024). 200, 000 HF downloading
- 2024 Oct, HuatuoGPT **deployed** in 12 local hospitals
- 2024 Nov. **Apollo 2: multilingual** medical LLMs for 50 languages. (ICLR 2025)
- 2024 Dec, **HuatuoGPT-o1**: Towards Medical **Complex Reasoning** with LLMs. (Findings of ACL 2025). 1k GitHub stars
- 2024 June, TCM-omini: **Multimodal (audio/vision/EEG)** LLMs **for Traditional Chinese Medicine.**

[1] Hongbo Zhang, Junying Chen, Feng Jiang, Fei Yu, Zhihong Chen, Jianquan Li, Guiming Chen, Xiangbo Wu, Zhiyi Zhang, Qingying Xiao, Xiang Wan, Benyou Wang, **Haizhou Li**. HuatuoGPT, towards Taming Language Model to Be a Doctor. https://arxiv.org/abs/2305.15075. Findings of EMNLP 2023

[2] Junying Chen, Xidong Wang, Anningzhe Gao#, Feng Jiang, Shunian Chen, Hongbo Zhang, Dingjie Song, Wenya Xie, Chuyi Kong, Jianquan Li, Xiang Wan, **Haizhou Li,** Benyou Wang**.** Huatuogpt-ii, one-stage training for medical adaption of llms. https://arxiv.org/abs/2311.09774. COLM 2024

[3] Xidong Wang, Guiming Hardy Chen, Dingjie Song, Zhiyi Zhang, Zhihong Chen, Qingying Xiao, Feng Jiang, Jianquan Li, Xiang Wan, Benyou Wang, **Haizhou Li** .CMB: A Comprehensive Medical Benchmark in Chinese. NAACL 2024

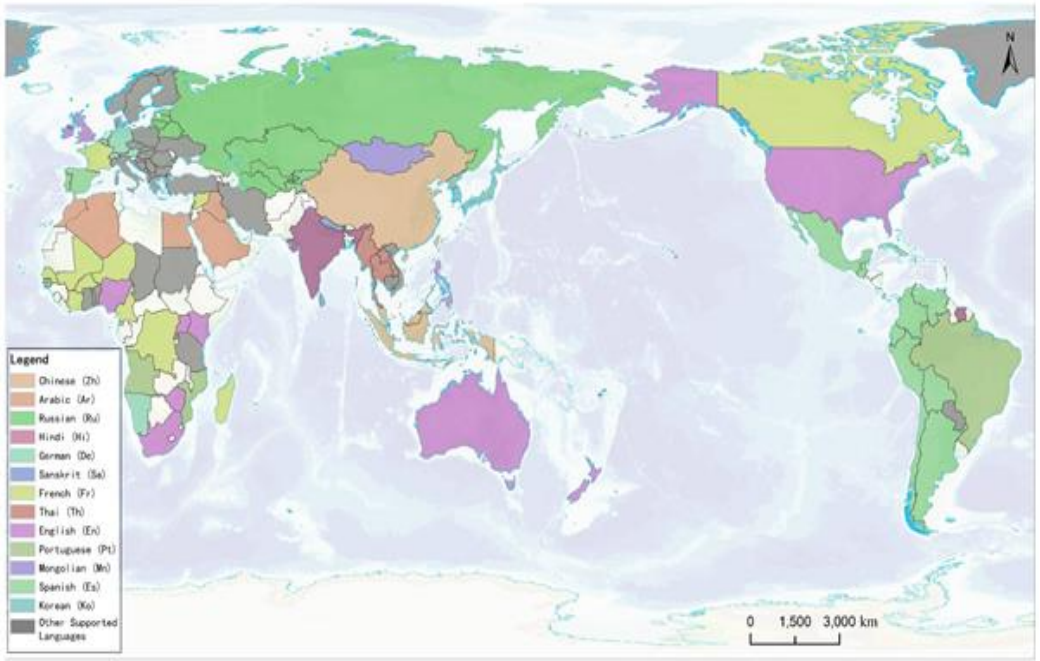
[4] Xidong Wang, Nuo Chen, Junyin Chen, Yan Hu, Yidong Wang, Xiangbo Wu, Anningzhe Gao, Xiang Wan, **Haizhou Li**, Benyou Wang. Apollo: A Lightweight Multilingual Medical LLM towards Democratizing Medical AI to 6B People. https://arxiv.org/abs/2403.03640.

[5] Junying Chen, Ruyi Ouyang, Anningzhe Gao, Shunian Chen, Guiming Hardy Chen, Xidong Wang, Ruifei Zhang, Zhenyang Cai, Ke Ji, Guangjun Yu, Xiang Wan, Benyou Wang. HuatuoGPT-Vision, Towards Injecting Medical Visual Knowledge into Multimodal LLMs at Scale. https://arxiv.org/abs/2406.19280. submitted to NeurIPS 2024

[6] Wenya Xie, Qingying Xiao, Yu Zheng, Xidong Wang, Junying Chen, Ke Ji, Anningzhe Gao, Xiang Wan, Feng Jiang, Benyou Wang. LLMs for Doctors: Leveraging Medical LLMs to Assist Doctors, Not Replace Them. https://arxiv.org/abs/2406.18034



华佗GPT智能导诊和预问诊系统以及龙岗家医应用到深圳市龙岗区11家公立医院和186家社康医院，服务500万人口



医疗大模型Apollo扩展到五十种语言，覆盖全球近80亿人口

Best medical LLMs in Oct. 2023

Model	Pharmacist Licensure Examination (Pharmacy)					Pharmacist Licensure Examination (TCM)					AVG
	Optimal Choice	Matched Selection	Integrated Analysis	Multiple Choice	Total Score	Optimal Choice	Matched Selection	Integrated Analysis	Multiple Choice	Total Score	
DISC-MedLLM	22.2	26.8	23.3	0.0	22.6	24.4	32.3	15.0	0.0	24.9	23.8
HuatuoGPT	25.6	25.5	23.3	2.6	23.4	24.1	26.8	31.6	7.5	24.9	24.2
ChatGLM2-6B	37.0	36.8	25.0	31.7	35.3	33.1	37.3	35.0	37.3	33.7	34.5
ChatGLM3-6B	39.5	39.1	10.5	0.2	34.6	31.8	38.2	25.0	20.0	32.9	33.8
Qwen-7B-chat	43.8	46.8	33.3	18.4	41.9	40.0	43.2	33.3	17.5	38.8	40.4
Qwen-14B-chat	56.2	58.6	41.7	21.1	52.7	51.3	51.0	27.5	41.7	47.9	50.3
Biachuan2-7B-Chat	51.2	50.9	30.0	2.6	44.6	48.1	46.0	35.0	7.5	42.1	43.4
Biachuan2-13B-Chat	43.8	52.7	36.7	7.9	44.2	41.3	46.4	43.3	15.0	41.7	43.0
文心一言	45.0	60.9	36.7	23.7	49.6	53.8	59.1	38.3	20.0	51.5	50.6
ChatGPT(API)	45.6	44.1	36.7	13.2	41.2	34.4	32.3	30.0	15.0	31.2	36.2
GPT-4(API)	65.1	59.6	46.7	15.8	**57.3**	40.6	42.7	33.3	17.5	38.8	48.1
HuatuoGPT-II(7B)	41.9	61.0	35.0	15.7	47.7	52.5	51.4	41.7	15.0	47.5	47.6
HuatuoGPT-II(13B)	47.5	64.1	45.0	23.7	52.9	48.8	61.8	45.0	17.5	**51.6**	**52.3**
HuatuoGPT-II(34B)	66.3	75.0	48.3	34.2	65.5	63.6	71.4	50.0	27.5	**62.5**	**64.0**

11月份的模型测试十月份的考试!

Hongbo Zhang, Junying Chen, Feng Jiang, Fei Yu, Zhihong Chen, Jianquan Li, Guiming Chen, Xiangbo Wu, Zhiyi Zhang, Qingying Xiao, Xiang Wan, Benyou Wang#, Haizhou Li. HuatuoGP towards Taming Language Model to Be a Doctor. https://arxiv.org/abs/2305.15075

Junying Chen, Xidong Wang, Anningzhe Gao#, Feng Jiang, Shunian Chen, Hongbo Zhang, Dingjie Song, Wenya Xie, Chuyi Kong, Jianquan Li, Xiang Wan, **Haizhou Li, Benyou Wang#.** Huatuogpt-ii, one-stage training for medical adaption of llms. https://arxiv.org/abs/2311.09774

Apollo: the multilingual version of Medical LLMs

Table 3: Performance comparison across various medical question answering models.

Language	English			Chinese		French	Spanish	Arabic	Hindi	Avg.
	USMLE	MedMCQA	MMLU♢	MCMLE	CMMLU♢	FrenchMedMCQA	HEAD-QA	MMLU♢	MMLU♢	
Closed-source										
GPT-4	79.10	70.40	86.00	65.72	65.72	89.72	85.05	56.43	62.17	73.37
GPT-3.5	61.98	56.51	72.94	58.73	50.41	68.54	71.48	39.70	39.94	57.80
Open-source (Above 70B)										
Qwen-72B	64.10	62.16	78.46	91.68	81.47	74.14	76.62	46.87	43.16	68.74
Meditron-70B	55.70	50.87	69.59	48.34	40.29	53.27	59.74	19.30	31.31	47.60
Llama-2-70B	32.99	48.29	64.62	25.80	25.13	50.47	54.34	1.65	26.35	36.63
Open-source (Above 7B)										
Qwen-14B	50.27	45.83	61.68	75.22	61.82	49.53	60.81	36.58	32.29	52.67
Gemma-7B	53.42	50.94	70.15	48.95	43.29	57.63	62.79	36.21	48.58	52.44
MMedLM2-7B	55.46	50.49	68.15	64.30	56.11	58.57	62.14	23.53	24.15	51.45
Yi-34B	62.45	60.60	71.86	26.12	26.51	66.04	69.99	30.70	9.73	47.00
Mistral-7B	47.29	47.38	62.80	38.32	34.21	50.78	51.93	28.40	27.36	43.16
Qwen-7B	32.36	39.52	53.22	54.32	44.71	37.69	45.05	28.31	24.89	40.01
Zephyr-7B-β	41.95	42.48	58.74	36.11	31.88	46.42	46.77	27.02	27.92	39.92
BioMistral-7B	41.79	42.05	54.46	34.65	31.43	43.61	44.66	27.11	22.96	38.08
Huatuo2-7B	37.86	36.58	42.49	55.08	43.81	27.41	33.88	25.92	27.46	36.72
Huatuo2-13B	29.77	36.58	42.86	56.07	45.46	22.42	36.13	18.29	13.59	33.46
Llama-2-7B	32.13	36.58	40.14	25.39	25.13	29.60	33.54	21.42	27.27	30.13
Meditron-7B	33.78	34.54	36.18	27.50	27.16	24.00	32.81	1.65	18.27	26.21
PMC-Llama-7B	20.11	23.12	19.72	16.90	16.73	17.13	18.68	9.65	2.85	16.10
Our Models										
Apollo-0.5B	32.99	37.82	45.87	56.57	42.08	27.41	36.67	31.89	25.90	37.47
Apollo-1.8B	42.18	44.99	49.12	72.30	53.56	38.01	42.15	34.74	25.62	44.74
Apollo-2B	38.33	42.00	52.89	46.76	36.76	38.32	41.28	31.62	31.50	39.94
Apollo-6B	56.25	57.53	68.65	85.52	72.62	51.71	58.47	33.46	33.61	57.54
Apollo-7B	56.00	58.21	71.86	72.36	59.04	60.44	63.73	41.82	45.55	**58.78**
Apollo-MedJamba	63.55	62.28	73.69	74.11	60.54	66.98	69.18	44.30	38.75	61.49
Apollo-34B	71.96	69.97	82.64	93.99	88.15	76.95	76.26	33.64	34.16	69.75
Apollo-72B	72.35	68.18	82.10	92.50	84.55	80.69	79.47	54.60	56.84	74.59

Xidong Wang, Nuo Chen, Junyin Chen, Yan Hu, Yidong Wang, Xiangbo Wu, Anningzhe Gao, Xiang Wan, Haizhou Li, Benyou Wang.Apollo: An Lightweight Multilingual Medical LLM towards Democratizing Medical AI to 6B People. https://arxiv.org/abs/2403.03640

Example II：Arabic LLM: AceGPT

- **Arabic LLMs**
  - **AceGPT： value alignment for a new language (Arabic)**
  - **AceGPT 1.5: vocabulary expansion**
  - **AceGPT 2: native alignment**

AceGPT是发布时最好的阿拉伯语大模型

[1] Huang Huang, Fei Yu, Jianqing Zhu, Xuening Sun, Hao Cheng, Dingjie Song, Zhihong Chen, Abdulmohsen Alharthi, Bang An, Ziche Liu, Zhiyi Zhang, Junying Chen, Jianquan Li, Benyou Wang, Lian Zhang, Ruoyu Sun, Xiang Wan, Haizhou Li, Jinchao Xu. AceGPT, Localizing Large Language Models in Arabic. NAACL 2024
[2] Jianqing Zhu, Huang Huang, Zhihang Lin, Juhao Liang, Zhengyang Tang, Khalid Almubarak, Mosen Alharthi, Bang An, Juncai He, Xiangbo Wu, Fei Yu, Junying Chen, MA Zhuoheng, Yuhao Du, Yan Hu, He Zhang, Emad A. Alghamdi, Lian Zhang, Ruoyu Sun, Haizhou Li, Jinchao Xu, Benyou Wang. Second Language (Arabic) Acquisition of LLMs via Progressive Vocabulary Expansion. Submitted to COLM 2024
[3] Juhao Liang, Zhenyang Cai, Jianqing Zhu, Huang Huang, Kewei Zong, Bang An, Mosen Alharthi, Juncai He, Lian Zhang, Haizhou Li, Benyou Wang, Jinchao Xu. Alignment at Pre-training! Towards Native Alignment for Arabic LLMs. Submitted to NeurIPS 2024

FINANCIAL TIMES

HOME WORLD US COMPANIES TECH MARKETS CLIMATE OPINION WORK & CAREERS LIFE & ARTS HTSI

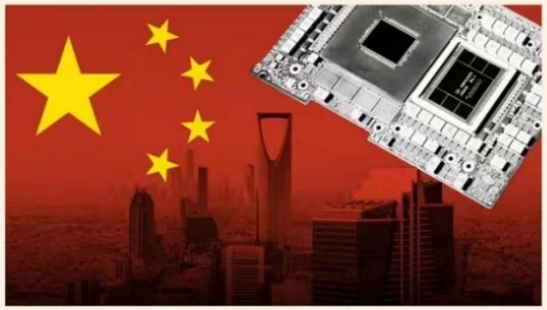
Stay one step ahead
Subscribe today and navigate your world with confidence

EXPLORE OUR BEST OFFERS

Artificial intelligence + Add to myFT

Saudi-China collaboration raises concerns about access to AI chips

Fears grow at Gulf kingdom's top university that ties to Chinese researchers risk upsetting US government



Western officials have long expressed concerns about growing technology transfer between their traditional allies in the Gulf and China © FT montage/Bloomberg/Dreamstime

Simeon Kerr and Samer Al-Atrush in Dubai, Qianer Liu in Hong Kong, Madhumita Murgia in London 13 HOURS AGO

52

Stay informed with free updates

Simply sign up to the Artificial intelligence myFT Digest -- delivered directly to your inbox.

Enter your email address

Sign up

Saudi-Chinese collaboration in artificial intelligence has stirred fears within the Gulf kingdom's premier academic institution that the ties could jeopardise the university's access to US-made chips needed to power the new technology.

Professor Jinchao Xu, an American-Chinese mathematician at Saudi Arabia's King Abdullah University of Science and Technology (Kaust), has launched AceGPT, an Arabic-focused large language model, in collaboration with the Chinese University of Hong Kong, Shenzhen (CUHK-SZ), and the Shenzhen Research Institute of Big Data.

环球网 国际新闻 > 环球独家

英媒给中沙AI合作泼冷水，中国专家：美西方不断干扰毫无道理

来源：环球时报 作者：黄培昭 赵觉理

-2023-
10/11
07:23

15

【环球时报驻埃及特派记者 黄培昭 环球时报记者 赵觉理】人工智能（AI）技术正成为中国与中东国家合作的新亮点，但这种互利共赢却遭到美西方阻挠。英国《金融时报》10日刊文，给中国和沙特的相关合作泼冷水。对此，有中国专家表示，中国和中东国家的科技合作基于双方在该领域的互补性，符合双方共同利益，美西方不断干扰毫无道理。

《金融时报》报道称，沙特阿卜杜拉国王科技大学、香港中文大学（深圳）与深圳大数据研究院三方合作开发人工智能大语言模型ACEGPT。"此举是沙特领导人工智能技术区域发展、建造大型超算和推出大语言模型努力的一部分。"这家英媒还称，沙特正与阿联酋一道，寻求参与到人工智能竞争之中。

但该媒体笔锋一转称，中国与海湾国家的此类合作让西方感到担忧，美国对中国实施的人工智能芯片出口限制也正影响相关合作。有阿卜杜拉国王科技大学的工作人员担忧，与中国的合作可能会引发美国不满，从而影响该大学获得先进人工智能芯片。

早在今年8月，路透社就曾报道称，美国芯片制造商英伟达和AMD均已收到美国政府限制向部分中东国家出口先进人工智能芯片的要求。有分析人士认为，美国的主要目的是防止中国从中东国家手中购买先进芯片。"德国之声"援引专家的分析称，沙特、阿联酋等大力投资人工智能的国家近年来加深了与中国的联系，因此它们都

环球时事

他们搭起一座桥，让世界阅读中国

意大利设"慰安妇"铜像，日方"强烈关切"
避免贸易摩擦扩大升级！中欧启动汽车关税磋商
新加坡媒体：金砖国家"集体潜力"吸引东南亚
我国首个！成立了！今年招收300人
这条隧道通了！时速350公里高铁取得新进展

环球业界



昔日黄沙换绿颜

珍稀植物丽豆全基因组首次揭秘
节能减碳 新疆油气企业发展"含绿量"足
3D打印真空系统或能"捕捉"暗物质
证监会支持上海加快"五个中心"建设
依托智慧农业设施 高标准农田展现抗旱优势

Example III： Multi-modal LLMs

- 高质量的图文数据ALLaVA和模型
- 长窗口的多模态大模型的评估, 被COLM 2024接受
- Motion理解模型MotionLLM
- MLLM-bench 多模态大模型的评测
- 多模态大模型的反馈学习
- 语音理解大模型Soundwave
- 端对端语音大模型Echo-X
- 基于语言模型的生图模型Janus

[1] Guiming Hardy Chen, Shunian Chen, Ruifei Zhang, Junying Chen, Xiangbo Wu, Zhiyi Zhang, Zhihong Chen, Jianquan Li, Xiang Wan, **Benyou Wang**. ALLaVA: Harnessing GPT4V-synthesized Data for A Lite Vision-Language Model. https://arxiv.org/abs/2402.11684
[2] D Song, S Chen, GH Chen, F Yu, X Wan**, B Wang**. Milebench: Benchmarking mllms in long context, arXiv preprint arXiv:2404.18532
[3] Ling-Hao Chen, Shunlin Lu, Ailing Zeng, Hao Zhang, Benyou Wang, Ruimao Zhang, Lei Zhang. MotionLLM: Understanding Human Behaviors from Human Motions and Videos. https://arxiv.org/abs/2405.20340
[4] Wentao Ge, Shunian Chen, Guiming Chen, Junying Chen, Zhihong Chen, Shuo Yan, Chenghao Zhu, Ziyue Lin, Wenya Xie, Xidong Wang, Anningzhe Gao, Zhiyi Zhang, Jianquan Li, Xiang Wan, Benyou Wang. Mllm-bench, evaluating multi-modal llms using gpt-4v. https://arxiv.org/abs/2311.13951
[5] Lei Li, Zhihui Xie, Mukai Li, Shunian Chen, Peiyi Wang, Liang Chen, Yazheng Yang, Benyou Wang, Lingpeng Kong. Silkie: Preference distillation for large visual language models. https://arxiv.org/abs/2312.10665

Example IV： LLM for Math and Optimization

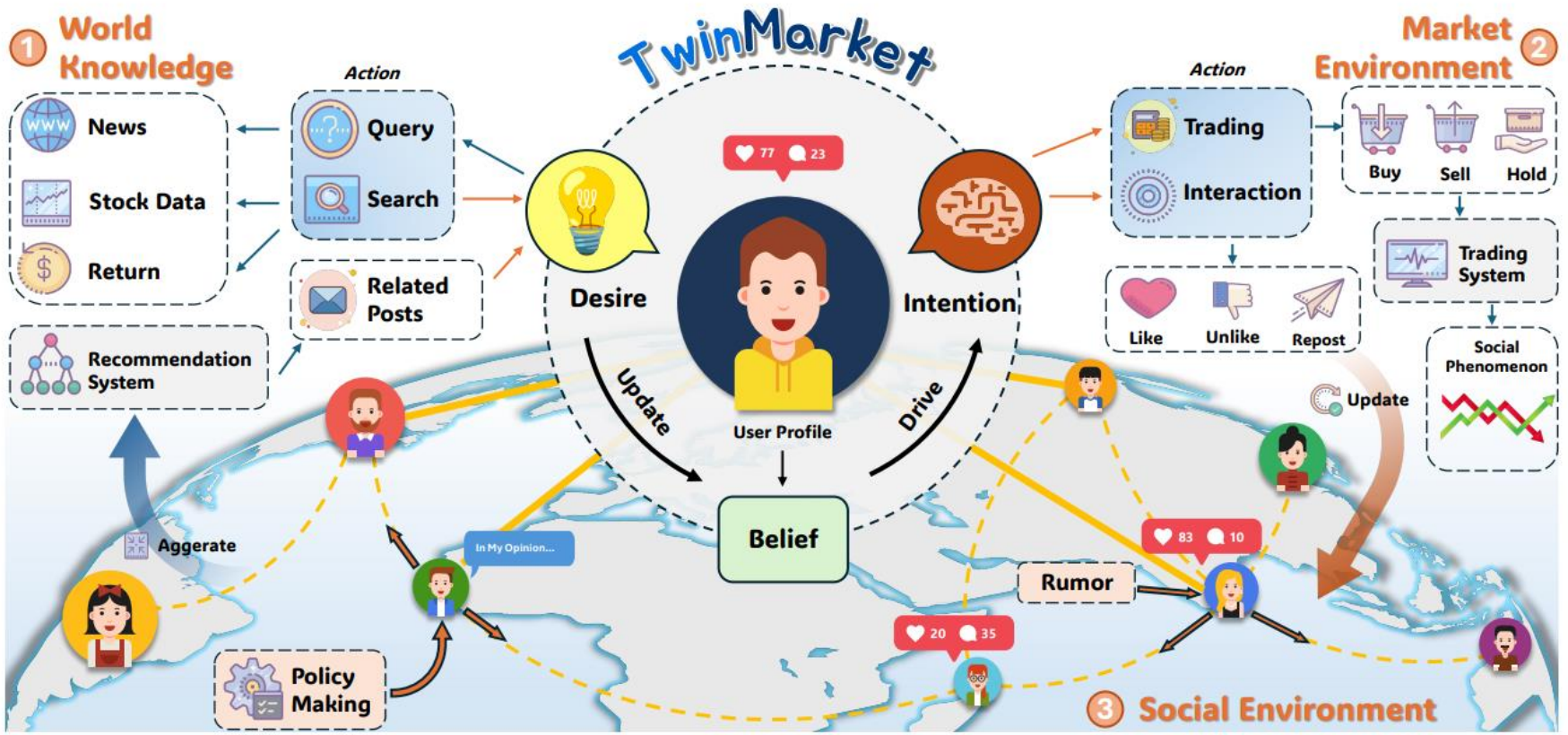
- Zhengyang Tang, Xingxing Zhang, **Benyou Wang**, Furu Wei. MathScale: Scaling Instruction Tuning for Mathematical Reasoning, **ICML 2024**
- Zhengyang Tang, Chenyu Huang, Xin Zheng, Shixi Hu, Zizhuo Wang, Dongdong Ge, **Benyou Wang**. ORLM: Training Large Language Models for Optimization Modeling. **Operations Research**. https://arxiv.org/abs/2405.17743
- Xuhan Huang, Qingning Shen, Yan Hu, Anningzhe Gao, **Benyou Wang**. Mamo: a Mathematical Modeling Benchmark with Solvers. Findings of NAACL 2024. https://arxiv.org/abs/2405.13144v1
- Fei Yu, Anningzhe Gao, **Benyou Wang**. OVM, Outcome-supervised Value Models for Planning in Mathematical Reasoning. Findings of NAACL 2024. https://arxiv.org/abs/2311.13951
- Kong, C., Yaxin, F.,Wan, X., Jiang, F. &**Wang, B.** (2024). Platolm: Teaching llms via a socratic questioning user simulator. In **ACL 2024**.
- Haoran Li, Qingxiu Dong, Zhengyang Tang, Chaojun Wang, Xingxing Zhang, Haoyang Huang, Shaohan Huang, Xiaolong Huang, Zeqiang Huang, Dongdong Zhang, Yuxian Gu, Xin Cheng, Xun Wang, Si-Qing Chen, Li Dong, Wei Lu, Zhifang Sui, Benyou Wang, Wai Lam, Furu Wei. Synthetic data (almost) from scratch: Generalized instruction tuning for language models. https://arxiv.org/abs/2402.13064. Submitted to NeurIPS 2024.

Rank	Model	Accuracy	Parameters (Billion)	Extra Training Data	Paper	Code	Result	Year	Tags
1	**GPT-4 DUP**	97.1		×	Achieving >97% on GSM8K: Deeply Understanding the Problems Makes LLMs Better Solvers for Math Word Problems			2024	
2	**GPT-4 Code Interpreter** (CSV, K=5)	97.0		×	Solving Challenging Math Word Problems Using GPT-4 Code Interpreter with Code-based Self-Verification			2023	majority voting, zero-shot
3	**GPT-4** (Model Selection, SC K=15)	96.8		×	Automatic Model Selection with Large Language Models for Reasoning			2023	self-consistency, majority voting
4	**GPT-4** (PHP, SC K=40)	96.5		×	Progressive-Hint Prompting Improves Reasoning in Large Language Models			2023	self-consistency, majority voting
5	**GPT-4** (Model Selection, SC K=5)	96.5		×	Automatic Model Selection with Large Language Models for Reasoning			2023	self-consistency, majority voting
6	**SFT-Mistral-7B** (Metamath, OVM, Smart Ensemble)	96.4	7	✓				2024	
7	**SFT-Mistral-7B** (AugData + ovm + ensemble)	95.9	7	✓				2024	
8	**GPT-4** (PHP)	95.5		×	Progressive-Hint Prompting Improves Reasoning in Large Language Models			2023	
9	**MindOpt Copilot Mistral-7B** (MetaMath, OVM, BS, Ensemble)	95.1	7	✓				2024	
10	**Claude 3 Opus** (0-shot chain-of-thought)	95		×	The Claude 3 Model Family: Opus, Sonnet, Haiku			2024	chain-of-thought, zero-shot
11	**Gemini Ultra** (maj1@32)	94.4		×	Gemini: A Family of Highly Capable Multimodal Models			2023	majority voting
12	**SFT-Mistral-7B** (Metamath + ovm +ensemble)	94.13	7	✓				2024	

https://paperswithcode.com/sota/arithmetic-reasoning-on-gsm8k

- 在2025年第二届人工智能数学奥林匹克竞赛（AIMO2）中荣获金牌
- 联合杉数科技开发首个数学建模大模型 ORLM，并由该公司进一步升级为 COLORMind 平台投入实际业务应用。ORLM 的研究论文被国际顶级期刊《Operations Research》接收发表，标志着 AI 与运筹优化领域的深度融合。

Example V： Agent and Human-Agent Interactions



Yuzhe Yang, Yifei Zhang, Minghao Wu, Kaidi Zhang, Yunmiao Zhang, Honghai Yu, Yan Hu, **Benyou Wang**. TwinMarket: A Scalable Behavioral and Social Simulation for Financial Markets. **ICLR Financial AI Best Paper**. NeurIPS 2025

代表性工作四： Agent and Human-Agent Interactions

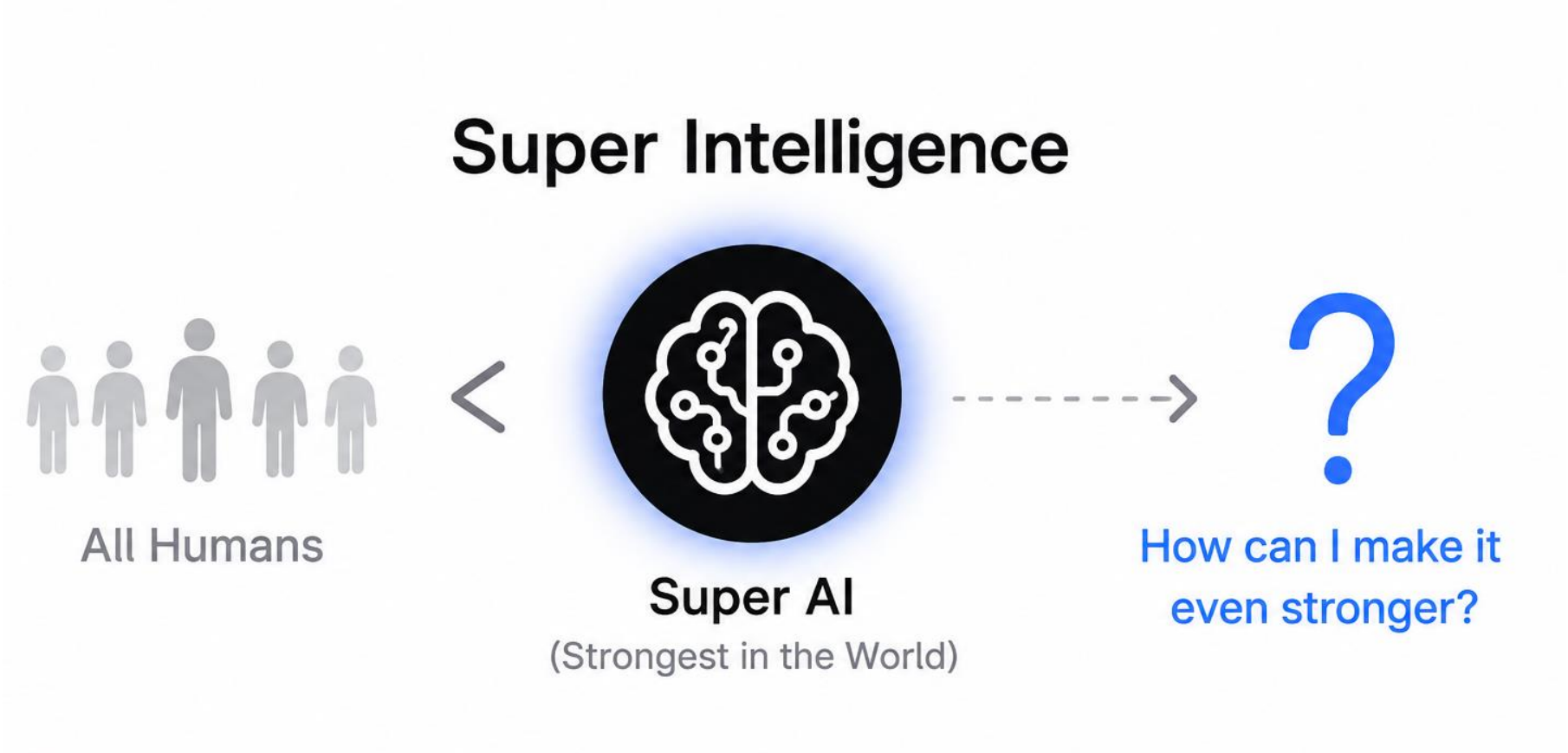




Our research

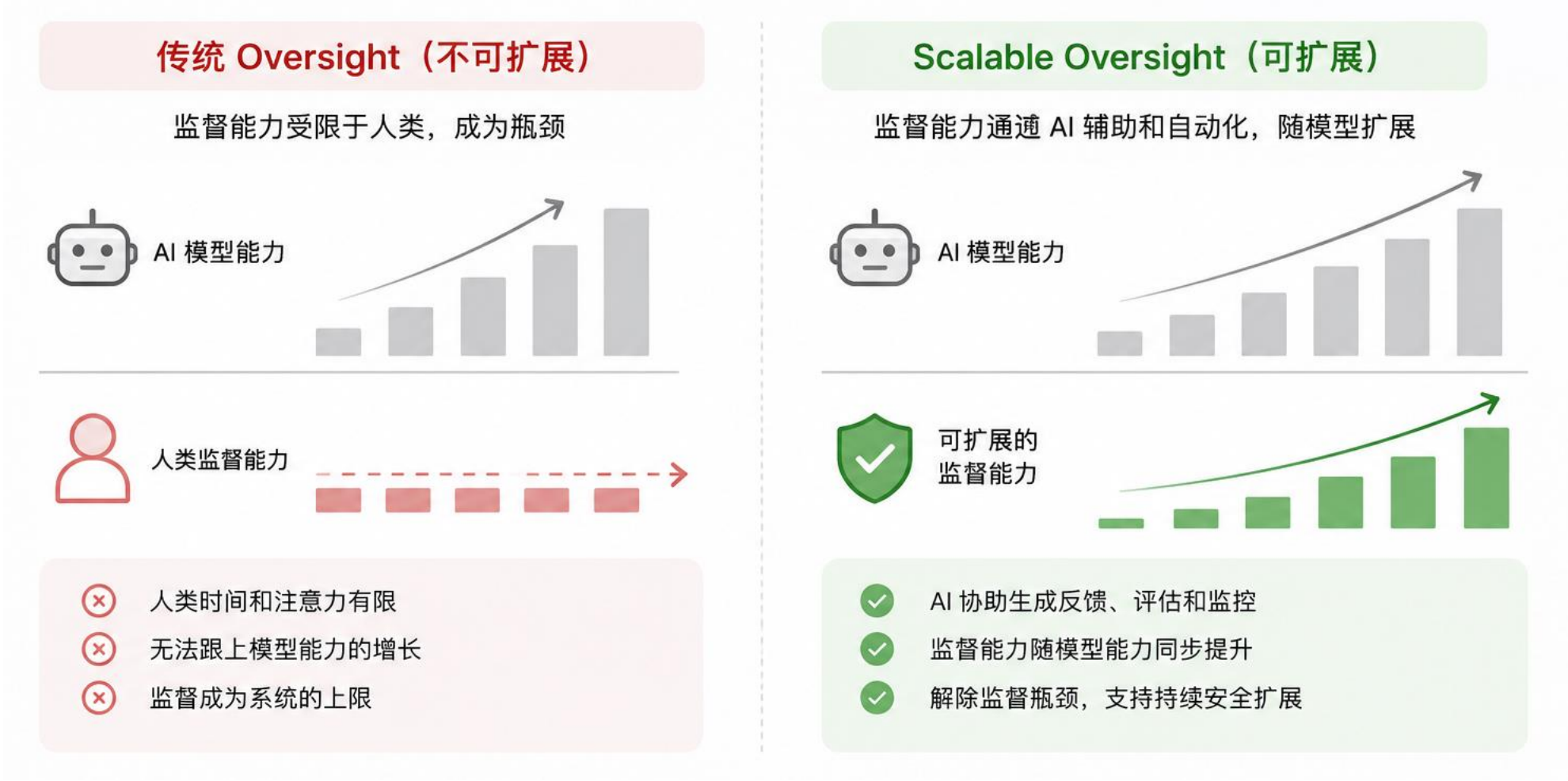
Recursive Self-improvement (RSI)
in heterogeneous real-world **environments**

Recursive Self-improvement (RSI)



Superintelligence

Challenge： **Scalable oversight**！



Applications (heterogeneous **environments**)

- Healthcare
  - Agentic Hospital
  - Medical world models
  - Virtual cells
- AI for AI
  - Kernel optimization
  - Automatic Architecture design
- Finance/Legal applications
  - Quantitative Finance
  - 公司金融
- World Model
  - Interactive World Model
  - 3D world model
- Game
- 具身

Surveys

Survey 1：Jobs in the future

- Which jobs are less likely be replaced by AI

A. Doctors
B. Mathematicians
C. Farmers
D. Software Engineers
E. AI Engineers
F. Blue-collar workers (蓝领工人)
G. White-collar workers (白领工人)

Survey 2：Intelligence Law

- How much human work time can AI replace at a decent quality level?
- *目前、AI能够以体面的效果完成“最好的人类”多少时间能够完成的的工作？*

A. Less than 1 minute
B. 1 minute to 1 hour
C. 1 hour to 10 hours
D. 10 hours to 100 hours
E. More than 100 hours

Survey 3: AI replaced Humans?

- In the future (**to the year of 2030**), how much of human work will AI replace?

- A. Less than 20%
- B. 20–50%
- C. 50–80%
- D. 80–95%
- E. More than 95%
- F. 100%

Survey 4: AI replaced Humans?

- In the future (**to the year of 2050**), how much of human work will AI replace?

- A. Less than 20%
- B. 20–50%
- C. 50–80%
- D. 80–95%
- E. More than 95%
- F. 100%

Thanks

Join us as a RA if interested

wangbenyou@cuhk.edu.cn



Survey