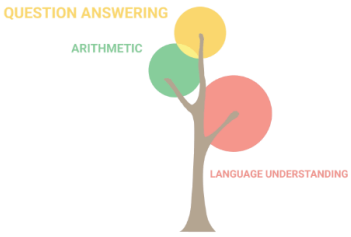
Awesome-LLM



🔥 Large Language Models(LLM) have taken the ~~NLP community~~ ~~AI community~~ **the Whole World** by storm. Here is a curated list of papers about large language models, especially relating to ChatGPT. It also contains frameworks for LLM training, tools to deploy LLM, courses and tutorials about LLM and all publicly available LLM checkpoints and APIs.

Trending LLM Projects

- Omost - a project to convert LLM's coding capability to image generation (or more accurately, image composing) capability.
- llama-fs - A self-organizing file system with llama 3.
- MiniCPM-V - A GPT-4V Level Multimodal LLM on Your Phone.
- fabric - an open-source framework for augmenting humans using AI.

Table of Content

Milestone Papers

Date	keywords	Institute	Paper	Publication
2017-06	Transformers	Google	Attention Is All You Need	NeurIPS citation 94979
2018-06	GPT 1.0	OpenAI	Improving Language Understanding by Generative Pre-Training	citation 8833
2018-10	BERT	Google	BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding	NAACL citation 76117
2019-02	GPT 2.0	OpenAI	Language Models are Unsupervised Multitask Learners	citation 16676

Date	keywords	Institute	Paper	Publication
2019-09	Megatron-LM	NVIDIA	Megatron-LM: Training Multi-Billion Parameter Language Models Using Model Parallelism	citation 1293
2019-10	T5	Google	Exploring the Limits of Transfer Learning with a Unified Text-to-Text Transformer	JMLR citation 14312
2019-10	ZeRO	Microsoft	ZeRO: Memory Optimizations Toward Training Trillion Parameter Models	SC citation 123
2020-01	Scaling Law	OpenAI	Scaling Laws for Neural Language Models	citation 2509
2020-05	GPT 3.0	OpenAI	Language models are few-shot learners	NeurIPS citation 26562
2021-01	Switch Transformers	Google	Switch Transformers: Scaling to Trillion Parameter Models with Simple and Efficient Sparsity	JMLR citation 1331

Date	keywords	Institute	Paper	Publication
2021-08	Codex	OpenAI	Evaluating Large Language Models Trained on Code	citation 2870
2021-08	Foundation Models	Stanford	On the Opportunities and Risks of Foundation Models	citation 2771
2021-09	FLAN	Google	Finetuned Language Models are Zero-Shot Learners	ICLR citation 2339
2021-10	T0	HuggingFace et al.	Multitask Prompted Training Enables Zero-Shot Task Generalization	ICLR citation 1327
2021-12	GLaM	Google	GLaM: Efficient Scaling of Language Models with Mixture-of-Experts	ICML citation 478
2021-12	WebGPT	OpenAI	WebGPT: Browser-assisted question-answering with human feedback	citation 753
2021-12	Retro	DeepMind	Improving language models by retrieving from trillions of tokens	ICML citation 653

Date	keywords	Institute	Paper	Publication
2021-12	Gopher	DeepMind	Scaling Language Models: Methods, Analysis & Insights from Training Gopher	citation 976
2022-01	COT	Google	Chain-of-Thought Prompting Elicits Reasoning in Large Language Models	NeurIPS citation 4047
2022-01	LaMDA	Google	LaMDA: Language Models for Dialog Applications	citation 1217
2022-01	Minerva	Google	Solving Quantitative Reasoning Problems with Language Models	NeurIPS citation 470
2022-01	Megatron-Turing NLG	Microsoft&NVIDIA	Using Deep and Megatron to Train Megatron-Turing NLG 530B, A Large-Scale Generative Language Model	citation 582
2022-03	InstructGPT	OpenAI	Training language models to follow instructions with human feedback	citation 6875

Date	keywords	Institute	Paper	Publication
2022-04	PaLM	Google	PaLM: Scaling Language Modeling with Pathways	citation 4135
2022-04	Chinchilla	DeepMind	An empirical analysis of compute-optimal large language model training	NeurIPS citation 177
2022-05	OPT	Meta	OPT: Open Pre-trained Transformer Language Models	citation 2342
2022-05	UL2	Google	Unifying Language Learning Paradigms	ICLR citation 120
2022-06	Emergent Abilities	Google	Emergent Abilities of Large Language Models	TMLR citation 1461
2022-06	BIG-bench	Google	Beyond the Imitation Game: Quantifying and extrapolating the capabilities of language models	citation 1063
2022-06	METALM	Microsoft	Language Models are General-Purpose Interfaces	citation 81
2022-09	Sparrow	DeepMind	Improving alignment of dialogue agents	citation 368

Date	keywords	Institute	Paper	Publication
			via targeted human judgements	
2022-10	Flan-T5/PaLM	Google	Scaling Instruction-Finetuned Language Models	citation 1981
2022-10	GLM-130B	Tsinghua	GLM-130B: An Open Bilingual Pre-trained Model	ICLR citation 755
2022-11	HELM	Stanford	Holistic Evaluation of Language Models	citation rate limited by upstream service
2022-11	BLOOM	BigScience	BLOOM: A 176B-Parameter Open-Access Multilingual Language Model	citation 1570
2022-11	Galactica	Meta	Galactica: A Large Language Model for Science	citation 477
2022-12	OPT-IML	Meta	OPT-IML: Scaling Language Model Instruction Meta Learning through the Lens of Generalization	citation 190
2023-01	Flan 2022 Collection	Google	The Flan Collection: Designing Data and Methods for Effective	ICML citation 354

Date	keywords	Institute	Paper	Publication
			Instruction Tuning	
2023-02	LLaMA	Meta	LLaMA: Open and Efficient Foundation Language Models	citation 6091
2023-02	Kosmos-1	Microsoft	Language Is Not All You Need: Aligning Perception with Language Models	citation 333
2023-03	PaLM-E	Google	PaLM-E: An Embodied Multimodal Language Model	ICML citation 868
2023-03	GPT 4	OpenAI	GPT-4 Technical Report	citation 4809
2023-04	Pythia	EleutherAI et al.	Pythia: A Suite for Analyzing Large Language Models Across Training and Scaling	ICML citation 606
2023-05	Dromedary	CMU et al.	Principle-Driven Self-Alignment of Language Models from Scratch with Minimal Human Supervision	NeurIPS citation 165
2023-05	PaLM 2	Google	PaLM 2 Technical Report	citation 734
2023-05	RWKV	Bo Peng	RWKV: Reinventing	EMNLP citation 209

Date	keywords	Institute	Paper	Publication
			RNNs for the Transformer Era	
2023-05	DPO	Stanford	Direct Preference Optimization: Your Language Model is Secretly a Reward Model	Neurips 
2023-05	ToT	Google&Princeton	Tree of Thoughts: Deliberate Problem Solving with Large Language Models	NeurIPS 
2023-07	LLaMA 2	Meta	Llama 2: Open Foundation and Fine-Tuned Chat Models	
2023-10	Mistral 7B	Mistral	Mistral 7B	
2023-12	Mamba	CMU&Princeton	Mamba: Linear-Time Sequence Modeling with Selective State Spaces	
2024-03	Jamba	AI21 Labs	Jamba: A Hybrid Transformer-Mamba Language Model	

Other Papers

If you're interested in the field of LLM, you may find the above list of milestone papers helpful to explore its history and state-of-the-art. However, each direction of LLM offers a unique set of insights and contributions, which are essential to understanding the field as a whole. For a detailed list of papers in various subfields, please refer to the following link:

- Awesome-LLM-hallucination - LLM hallucination paper list.
- awesome-hallucination-detection - List of papers on hallucination detection in LLMs.
- LLMsPracticalGuide - A curated list of practical guide resources of LLMs
- Awesome ChatGPT Prompts - A collection of prompt examples to be used with the ChatGPT model.
- awesome-chatgpt-prompts-zh - A Chinese collection of prompt examples to be used with the ChatGPT model.
- Awesome ChatGPT - Curated list of resources for ChatGPT and GPT-3 from OpenAI.
- Chain-of-Thoughts Papers - A trend starts from "Chain of Thought Prompting Elicits Reasoning in Large Language Models.
- Awesome Deliberative Prompting - How to ask LLMs to produce reliable reasoning and make reason-responsive decisions.
- Instruction-Tuning-Papers - A trend starts from `Natrural-Instruction` (ACL 2022), `FLAN` (ICLR 2022) and `T0` (ICLR 2022).
- LLM Reading List - A paper & resource list of large language models.
- Reasoning using Language Models - Collection of papers and resources on Reasoning using Language Models.
- Chain-of-Thought Hub - Measuring LLMs' Reasoning Performance
- Awesome GPT - A curated list of awesome projects and resources related to GPT, ChatGPT, OpenAI, LLM, and more.
- Awesome GPT-3 - a collection of demos and articles about the OpenAI GPT-3 API.
- Awesome LLM Human Preference Datasets - a collection of human preference datasets for LLM instruction tuning, RLHF and evaluation.

- RWKV-howto - possibly useful materials and tutorial for learning RWKV.
- ModelEditingPapers - A paper & resource list on model editing for large language models.
- Awesome LLM Security - A curation of awesome tools, documents and projects about LLM Security.
- Awesome-Align-LLM-Human - A collection of papers and resources about aligning large language models (LLMs) with human.
- Awesome-Code-LLM - An awesome and curated list of best code-LLM for research.
- Awesome-LLM-Compression - Awesome LLM compression research papers and tools.
- Awesome-LLM-Systems - Awesome LLM systems research papers.
- awesome-llm-webapps - A collection of open source, actively maintained web apps for LLM applications.
- awesome-japanese-llm - 日本語LLMまとめ - Overview of Japanese LLMs.
- Awesome-LLM-Healthcare - The paper list of the review on LLMs in medicine.
- Awesome-LLM-Inference - A curated list of Awesome LLM Inference Paper with codes.
- Awesome-LLM-3D - A curated list of Multi-modal Large Language Model in 3D world, including 3D understanding, reasoning, generation, and embodied agents.
- LLMDatahub - a curated collection of datasets specifically designed for chatbot training, including links, size, language, usage, and a brief description of each dataset
- Awesome-Chinese-LLM - 整理开源的中文大语言模型，以规模较小、可私有化部署、训练成本较低的模型为主，包括底座模型，垂直领域微调及应用，数据集与教程等。
- LLM4Opt - Applying Large language models (LLMs) for diverse optimization tasks (Opt) is an emerging research area. This is a collection of references and papers of LLM4Opt.

LLM Leaderboard

- Chatbot Arena Leaderboard - a benchmark platform for large language models (LLMs) that features anonymous, randomized battles in a crowdsourced manner.

- AlpacaEval Leaderboard - An Automatic Evaluator for Instruction-following Language Models using Nous benchmark suite.
- Open LLM Leaderboard - aims to track, rank and evaluate LLMs and chatbots as they are released.
- OpenCompass 2.0 LLM Leaderboard - OpenCompass is an LLM evaluation platform, supporting a wide range of models (InternLM2,GPT-4,LLaMa2, Qwen,GLM, Claude, etc) over 100+ datasets.

Open LLM

- Meta
  - Llama 3-8|70B
  - Llama 2-7|13|70B
  - Llama 1-7|13|33|65B
  - OPT-1.3|6.7|13|30|66B
- Mistral AI
  - Mistral-7B
  - Mixtral-8x7B
  - Mixtral-8x22B
- Google
  - Gemma-2|7B
  - RecurrentGemma-2B
  - T5
- Apple
  - OpenELM-1.1|3B
- Microsoft
  - Phi1-1.3B
  - Phi2-2.7B
  - Phi3-3.8|7|14B
- AllenAI
  - OLMo-7B
- xAI
  - Grok-1-314B-MoE
- Cohere
  - Command R-35B
- DeepSeek
  - DeepSeek-Math-7B
  - DeepSeek-Coder-1.3|6.7|7|33B
  - DeepSeek-VL-1.3B|7B

    - DeepSeek-MoE-16B
    - DeepSeek-v2-236B-MoE
- Alibaba
    - Qwen-1.8|7|14|72B
    - Qwen1.5-1.8|4|7|14|32|72|110B
    - CodeQwen-7B
    - Qwen-VL-7B
- 01-ai
    - Yi-34B
    - Yi1.5-6|9|34B
    - Yi-VL-6B|34B
- Baichuan
    - Baichuan-7|13B
    - Baichuan2-7|13B
- BLOOM
    - BLOOMZ&mT0
- Zhipu AI
    - GLM-2|6|10|13|70B
    - CogVLM2-19B
- OpenBMB
    - MiniCPM-2B
    - OmniLLM-12B
    - VisCPM-10B
    - CPM-Bee-1|2|5|10B
- RWKV Foundation
    - RWKV-v4|5|6
- ElutherAI
    - Pythia-1|1.4|2.8|6.9|12B
- Stability AI
    - StableLM-3B
    - StableLM-v2-1.6|12B
    - StableCode-3B
- BigCode
    - StarCoder-1|3|7B
    - StarCoder2-3|7|15B
- DataBricks
    - MPT-7B
- Shanghai AI Laboratory
    - InternLM2-1.8|7|20B

  - InternLM-Math-7B|20B
  - InternLM-XComposer2-1.8|7B
  - InternVL-2|6|14|26

LLM Data

- LLMDataHub

LLM Evaluation:

- lm-evaluation-harness - A framework for few-shot evaluation of language models.
- lighteval - a lightweight LLM evaluation suite that Hugging Face has been using internally.
- OLMO-eval - a repository for evaluating open language models.
- instruct-eval - This repository contains code to quantitatively evaluate instruction-tuned models such as Alpaca and Flan-T5 on held-out tasks.
- simple-evals - Eval tools by OpenAI.
- Giskard - Testing & evaluation library for LLM applications, in particular RAGs
- LangSmith - a unified platform from LangChain framework for: evaluation, collaboration HITL (Human In The Loop), logging and monitoring LLM applications.
- Ragas - a framework that helps you evaluate your Retrieval Augmented Generation (RAG) pipelines.

LLM Training Frameworks

- DeepSpeed - DeepSpeed is a deep learning optimization library that makes distributed training and inference easy, efficient, and effective.
- Megatron-DeepSpeed - DeepSpeed version of NVIDIA's Megatron-LM that adds additional support for several features such as MoE model training, Curriculum Learning, 3D Parallelism, and others.
- torchtune - A Native-PyTorch Library for LLM Fine-tuning.
- torchtitan - A native PyTorch Library for large model training.
- Megatron-LM - Ongoing research training transformer models at scale.
- Colossal-AI - Making large AI models cheaper, faster, and more accessible.
- BMTrain - Efficient Training for Big Models.
- Mesh Tensorflow - Mesh TensorFlow: Model Parallelism Made Easier.
- maxtext - A simple, performant and scalable Jax LLM!

- Alpa - Alpa is a system for training and serving large-scale neural networks.
- GPT-NeoX - An implementation of model parallel autoregressive transformers on GPUs, based on the DeepSpeed library.

LLM Deployment

> Reference: llm-inference-solutions

- vLLM - A high-throughput and memory-efficient inference and serving engine for LLMs.
- TGI - a toolkit for deploying and serving Large Language Models (LLMs).
- exllama - A more memory-efficient rewrite of the HF transformers implementation of Llama for use with quantized weights.
- llama.cpp - LLM inference in C/C++.
- ollama - Get up and running with Llama 3, Mistral, Gemma, and other large language models.
- Langfuse - Open Source LLM Engineering Platform Tracing, Evaluations, Prompt Management, Evaluations and Playground.
- FastChat - A distributed multi-model LLM serving system with web UI and OpenAI-compatible RESTful APIs.
- MindSQL - A python package for Txt-to-SQL with self hosting functionalities and RESTful APIs compatible with proprietary as well as open source LLM.
- SkyPilot - Run LLMs and batch jobs on any cloud. Get maximum cost savings, highest GPU availability, and managed execution -- all with a simple interface.
- Haystack - an open-source NLP framework that allows you to use LLMs and transformer-based models from Hugging Face, OpenAI and Cohere to interact with your own data.
- Sidekick - Data integration platform for LLMs.
- QA-Pilot - An interactive chat project that leverages Ollama/OpenAI/MistralAI LLMs for rapid understanding and navigation of GitHub code repository or compressed file resources.
- Shell-Pilot - Interact with LLM using Ollama models(or openAI, mistralAI)via pure shell scripts on your Linux(or MacOS) system, enhancing intelligent system management without any dependencies.
- LangChain - Building applications with LLMs through composability
- Floom AI gateway and marketplace for developers, enables streamlined integration of AI features into products
- Swiss Army Llama - Comprehensive set of tools for working with local LLMs for various tasks.
- LiteChain - Lightweight alternative to LangChain for composing LLMs

- magentic - Seamlessly integrate LLMs as Python functions
- wechat-chatgpt - Use ChatGPT On Wechat via wechaty
- promptfoo - Test your prompts. Evaluate and compare LLM outputs, catch regressions, and improve prompt quality.
- Agenta - Easily build, version, evaluate and deploy your LLM-powered apps.
- Serge - a chat interface crafted with llama.cpp for running Alpaca models. No API keys, entirely self-hosted!
- Langroid - Harness LLMs with Multi-Agent Programming
- Embedchain - Framework to create ChatGPT like bots over your dataset.
- CometLLM - A 100% opensource LLMOps platform to log, manage, and visualize your LLM prompts and chains. Track prompt templates, prompt variables, prompt duration, token usage, and other metadata. Score prompt outputs and visualize chat history all within a single UI.
- IntelliServer - simplifies the evaluation of LLMs by providing a unified microservice to access and test multiple AI models.
- OpenLLM - Fine-tune, serve, deploy, and monitor any open-source LLMs in production. Used in production at BentoML for LLMs-based applications.
- DeepSpeed-Mii - MII makes low-latency and high-throughput inference, similar to vLLM powered by DeepSpeed.
- Text-Embeddings-Inference - Inference for text-embeddings in Rust, HFOIL Licence.
- Infinity - Inference for text-embeddings in Python
- TensorRT-LLM - Nvidia Framework for LLM Inference
- FasterTransformer - NVIDIA Framework for LLM Inference(Transitioned to TensorRT-LLM)
- Flash-Attention - A method designed to enhance the efficiency of Transformer models
- Langchain-Chatchat - Formerly langchain-ChatGLM, local knowledge based LLM (like ChatGLM) QA app with langchain.
- Search with Lepton - Build your own conversational search engine using less than 500 lines of code by LeptonAI.
- Robocorp - Create, deploy and operate Actions using Python anywhere to enhance your AI agents and assistants. Batteries included with an extensive set of libraries, helpers and logging.
- LMDeploy - A high-throughput and low-latency inference and serving framework for LLMs and VLs
- Tune Studio - Playground for devs to finetune & deploy LLMs
- LLocalSearch - Locally running websearch using LLM chains
- AI Gateway — Gateway streamlines requests to 100+ open & closed source models with a unified API. It is also production-ready with support for caching,

fallbacks, retries, timeouts, loadbalancing, and can be edge-deployed for minimum latency.

- talkd.ai dialog - Simple API for deploying any RAG or LLM that you want adding plugins.
- Wllama - WebAssembly binding for llama.cpp - Enabling in-browser LLM inference

LLM Applications

- YiVal — Evaluate and Evolve: YiVal is an open-source GenAI-Ops tool for tuning and evaluating prompts, configurations, and model parameters using customizable datasets, evaluation methods, and improvement strategies.
- Guidance — A handy looking Python library from Microsoft that uses Handlebars templating to interleave generation, prompting, and logical control.
- LangChain — A popular Python/JavaScript library for chaining sequences of language model prompts.
- FLAML (A Fast Library for Automated Machine Learning & Tuning): A Python library for automating selection of models, hyperparameters, and other tunable choices.
- Chainlit — A Python library for making chatbot interfaces.
- Guardrails.ai — A Python library for validating outputs and retrying failures. Still in alpha, so expect sharp edges and bugs.
- Semantic Kernel — A Python/C#/Java library from Microsoft that supports prompt templating, function chaining, vectorized memory, and intelligent planning.
- Prompttools — Open-source Python tools for testing and evaluating models, vector DBs, and prompts.
- Outlines — A Python library that provides a domain-specific language to simplify prompting and constrain generation.
- Promptify — A small Python library for using language models to perform NLP tasks.
- Scale Spellbook — A paid product for building, comparing, and shipping language model apps.
- PromptPerfect — A paid product for testing and improving prompts.
- Weights & Biases — A paid product for tracking model training and prompt engineering experiments.
- OpenAI Evals — An open-source library for evaluating task performance of language models and prompts.
- LlamaIndex — A Python library for augmenting LLM apps with data.
- Arthur Shield — A paid product for detecting toxicity, hallucination, prompt injection, etc.

- LMQL — A programming language for LLM interaction with support for typed prompting, control flow, constraints, and tools.
- ModelFusion - A TypeScript library for building apps with LLMs and other ML models (speech-to-text, text-to-speech, image generation).
- Flappy — Production-Ready LLM Agent SDK for Every Developer.
- GPTRouter - GPTRouter is an open source LLM API Gateway that offers a universal API for 30+ LLMs, vision, and image models, with smart fallbacks based on uptime and latency, automatic retries, and streaming. Stay operational even when OpenAI is down
- QAnything - A local knowledge base question-answering system designed to support a wide range of file formats and databases.
- OneKE — A bilingual Chinese-English knowledge extraction model with knowledge graphs and natural language processing technologies.
- llm-ui - A React library for building LLM UIs.
- Wordware - A web-hosted IDE where non-technical domain experts work with AI Engineers to build task-specific AI agents. We approach prompting as a new programming language rather than low/no-code blocks.
- Wallaroo.AI - Deploy, manage, optimize any model at scale across any environment from cloud to edge. Let's you go from python notebook to inferencing in minutes.

LLM Tutorials and Courses

- llm-course - Course to get into Large Language Models (LLMs) with roadmaps and Colab notebooks.
- UWaterloo CS 886 - Recent Advances on Foundation Models.
- CS25-Transformers United
- ChatGPT Prompt Engineering
- Princeton: Understanding Large Language Models
- CS324 - Large Language Models
- State of GPT
- A Visual Guide to Mamba and State Space Models
- Let's build GPT: from scratch, in code, spelled out.
- minbpe - Minimal, clean code for the Byte Pair Encoding (BPE) algorithm commonly used in LLM tokenization.
- femtoGPT - Pure Rust implementation of a minimal Generative Pretrained Transformer.
- Neurips2022-Foundational Robustness of Foundation Models
- ICML2022-Welcome to the "Big Model" Era: Techniques and Systems to Train and Serve Bigger Models

- GPT in 60 Lines of NumPy

LLM Books

- Generative AI with LangChain: Build large language model (LLM) apps with Python, ChatGPT, and other LLMs - it comes with a GitHub repository that showcases a lot of the functionality
- Build a Large Language Model (From Scratch) - A guide to building your own working LLM.
- BUILD GPT: HOW AI WORKS - explains how to code a Generative Pre-trained Transformer, or GPT, from scratch.

Great thoughts about LLM

- Why did all of the public reproduction of GPT-3 fail?
- A Stage Review of Instruction Tuning
- LLM Powered Autonomous Agents
- Why you should work on AI AGENTS!
- Google "We Have No Moat, And Neither Does OpenAI"
- AI competition statement
- Prompt Engineering
- Noam Chomsky: The False Promise of ChatGPT
- Is ChatGPT 175 Billion Parameters? Technical Analysis
- The Next Generation Of Large Language Models
- Large Language Model Training in 2023
- How does GPT Obtain its Ability? Tracing Emergent Abilities of Language Models to their Sources
- Open Pretrained Transformers
- Scaling, emergence, and reasoning in large language models

Miscellaneous

- Arize-Phoenix - Open-source tool for ML observability that runs in your notebook environment. Monitor and fine tune LLM, CV and Tabular Models.
- Emergent Mind - The latest AI news, curated & explained by GPT-4.
- ShareGPT - Share your wildest ChatGPT conversations with one click.
- Major LLMs + Data Availability
- 500+ Best AI Tools

- Cohere Summarize Beta - Introducing Cohere Summarize Beta: A New Endpoint for Text Summarization
- chatgpt-wrapper - ChatGPT Wrapper is an open-source unofficial Python API and CLI that lets you interact with ChatGPT.
- Open-evals - A framework extend openai's Evals for different language model.
- Cursor - Write, edit, and chat about your code with a powerful AI.
- AutoGPT - an experimental open-source application showcasing the capabilities of the GPT-4 language model.
- OpenAGI - When LLM Meets Domain Experts.
- EasyEdit - An easy-to-use framework to edit large language models.
- chatgpt-shroud - A Chrome extension for OpenAI's ChatGPT, enhancing user privacy by enabling easy hiding and unhiding of chat history. Ideal for privacy during screen shares.

Contributing

This is an active repository and your contributions are always welcome!

I will keep some pull requests open if I'm not sure if they are awesome for LLM, you could vote for them by adding 👍 to them.

If you have any question about this opinionated list, do not hesitate to contact me chengxin1998@stu.pku.edu.cn.