

# 大模型赋能学习、工作与研究

陈敏刚 博士、研究员

上海计算机软件技术开发中心 人工智能治理研究所

上海市计算机软件测评重点实验室

2025年4月



# 目录



## 01 人工智能（大模型）简介



## 02 AI赋能学习、工作与研究（演示）



## 03 总结与思考



# 人工智能（大模型） 简介

- ▶ 人工智能及大模型概念
- ▶ DeepSeek vs OpenAI
- ▶ 大模型的关键能力
- ▶ 提示词与提示工程





# 什么是人工智能?

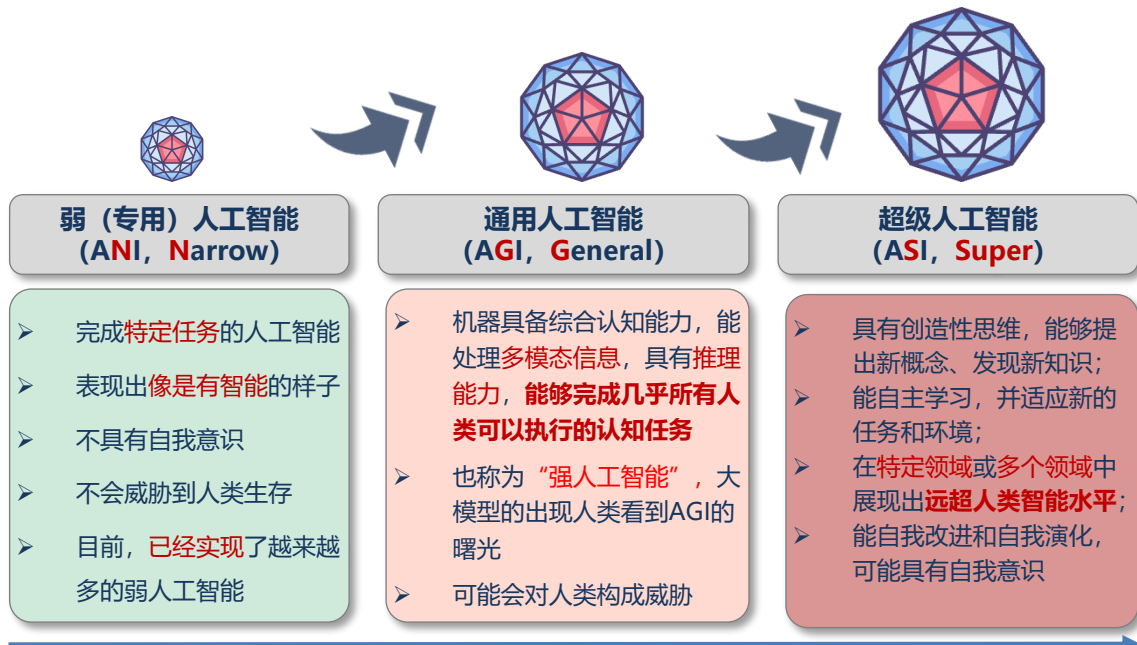
人工智能是研究开发能够模拟、延伸和扩展人类智能的理论、方法、技术及应用系统的科学<sup>[1]</sup>

研究目的是促使智能机器：

- ✓ 会看 (计算机视觉)
- ✓ 会听 (语音识别)
- ✓ 会说 (大语言模型)
- ✓ 会思考 (推理模型、人机对弈)
- ✓ 会学习 (机器学习、知识表示)
- ✓ 会行动 (智能体和机器人)

人工智能三要素：数据、算法、算力

## 人工智能整体发展阶段

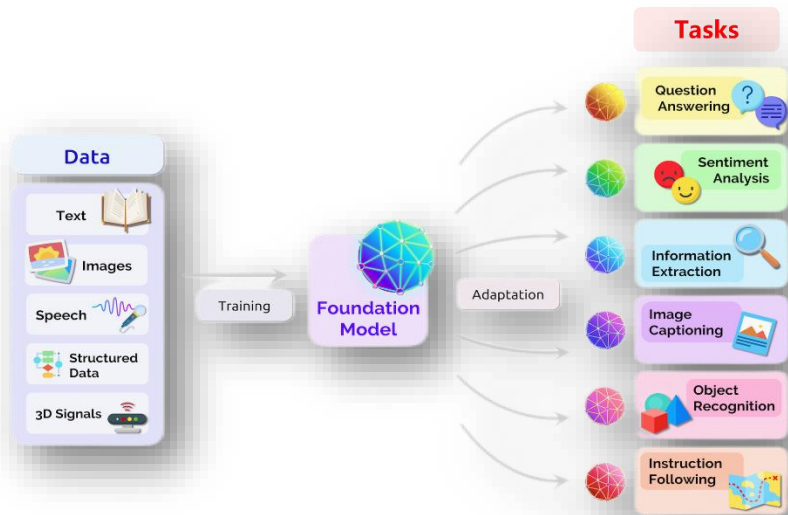


- 从“判别式、分析型人工智能”向“生成式人工智能”转变
- 大模型 (如ChatGPT、DeepSeek) 的出现让人类看到了AGI的曙光
- AGI强调认知的“广度”，ASI侧重认知的“深度”

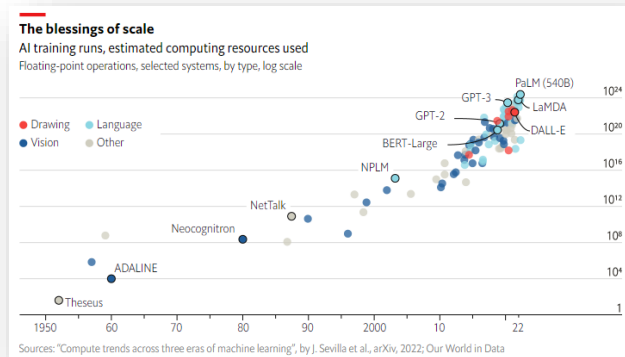
[1] 《求是》 人工智能的历史、现状和未来. 谭铁牛. 2019.04.



# 大模型的定义、特点和分类



大模型也称为基础模型 (Foundation Models)，是一类在大规模数据 (通常采用自监督方式) 上**预训练**，通过监督学习和强化学习**后训练**，能适配各类下游任务的模型，如OpenAI GPT、Google Gemini、Meta LLaMa、Grok3、DeepSeek、阿里的通义千问、Kimi、豆包等。



## □ 大模型的特点 (一大三多、涌现)：

- 一大是指**参数规模大** (千亿参数级别)
- **数据多来源、多模态**
- **适配多种下游任务**
- **智能的涌现**：上下文学习、指令学习和思维链

## □ 大模型的分类：语言大模型 (OpenAI ChatGPT)、多模态大模型 (OpenAI GPT-4o、Google Gemini)、文生图(视频、动画)大模型、视觉大模型

**大模型从单模态语言模型向多模态、推理模型和智能体融合，大模型的智能涌现及多任务泛化能力让人类看到了AGI的曙光。**



# 大（语言）模型关键能力

-shot  
In addition to the task description, the model sees a few examples of the task. No gradient updates are performed.

Translate English to French: task description  
sea of  
pepper  
plush  
cheese

## 上下文学习 (In-context Learning)

大模型涌现出的上下文学习能力，能通过少量示例快速学习、适配新的任务，改变传统学习方式。

**(理解语境、举一反三)**

## Natural Instructions

## 指令学习 (Instruction Learning)

理解人类的指令，具有零样本多任务的泛化能力，可实现人在环路增强。

**(理解、对齐人类意图)**

## Chain-of-Thought Prompting

Input  
Q: Roger has 5 tennis balls. He buys 2 more cans of tennis balls. Each can has 3 tennis balls. How many tennis balls does he have now?  
A: Roger has started with 5 balls. 2 cans of 3 tennis balls each is 6 tennis balls.  $5 + 6 = 11$ . The answer is 11.  
Q: How  
Mod  
A: 11  
Mu

## 思维链 (Chain-of-Thought)

大模型“涌现”出类人思维和逻辑能力，可对问题分解、推理和解释，实现连接主义和符号主义的融合。

**(思考与推理能力)**

大模型将帮助人们在学习、科研方面**更高效洞察知识、完成复杂任务，创造更高价值**



# 主要工具介绍 (DeepSeek、通义千问)



## 我是 DeepSeek, 很高兴见到你!

我可以帮你写代码、读文件、写作各种创意内容, 请把任务交给我吧~

给 DeepSeek 发送消息

深度思考 (R1) 联网搜索



<https://chat.deepseek.com/>

App: DeepSeek

**特点: 开源、深度思考, 性能对标美国OpenAI**



## 通情、达义, 你的全能AI助手

我可以帮你做这些事情, 换一换

### 今日热搜

深度搜索你关心的问题

- 1 苏州规划新建通用机场, 预计...
- 2 梁文锋未参加巴黎AI峰会, 显...
- 3 奢侈品市场遇冷, 年轻人更倾...
- 4 中国U20男足2-1胜卡塔尔, 开...
- 5 喜茶暂停加盟, 调整扩张策略...
- 6 上海男篮大胜北京男篮晋级C...

### 效率工具

办公学习使用通义效率

- 实时记录  
语音转文字, AI会议纪要神器
- 阅读助手  
AI速读论文、图书等超长文档
- PPT创作  
言之有物, 设计精美的智能PPT

```

#include <stdio.h>
int main() {
    printf("Hello, TONGYI!\n");
}
    
```

体验代码模式

### 精选智能体

- 会摆版的石荷像  
讲述艺术历史趣闻的智基雕塑

代码模式 深度搜索 PPT创作 指令中心

千事不决问通义

服务生成的所有内容均由人工智能模型生成, 其生成内容的准确性和完整性无法保证, 不代表我们的态度或观点

<https://tongyi.aliyun.com/qianwen/>

App: 通义-你的全能AI助手

**特点: 国内第一梯队大模型、效率工具、长文本翻译**



# 主要工具介绍 (Kimi、秘塔AI搜索)



<https://kimi.moonshot.cn/>

App: Kimi 智能助手

**特点: 长文本理解与生成、思考推理、PPT制作**



<https://metaso.cn/>

App: 秘塔AI搜索

**特点: AI搜索引擎、学术论文搜索、行业分析，对标美国AI搜索工具 Perplexity**



# 2025年初DeepSeek在全球“出圈”



上海计算机软件开发中心

## nature

[View](#)

[Explore content](#) [About the journal](#) [Publish with us](#) [Subscribe](#)

### News

30 Jan 2025

Nature

#### Daily briefing: The pros and cons of DeepSeek

Scientists have found the ingredients for life in the fragments of an asteroid (just not quite the right ones). Plus, researchers weigh up the pros and cons of DeepSeek-R1.

Jacob Smith



### Comments & Opinion

18 Feb 2025

Nature

Volume: 638, P: 609-611

#### China made waves with Deepseek, but its real ambition is AI-driven industrial innovation

The country is betting on AI to develop practical, cost-effective business tools for global markets, especially lower-income countries.

Jacob Dreyer



### News

29 Jan 2025

Nature

#### Scientists flock to DeepSeek: how they're using the blockbuster AI model

Researchers are testing how well the open model can perform scientific tasks — in topics from mathematics to cognitive neuroscience.

Elizabeth Gibney



### News

30 Jan 2025

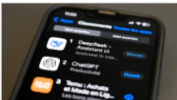
Nature

Volume: 638, P: 300-301

#### How China created AI model DeepSeek and shocked the world

Government policies, generous funding and a pipeline of AI graduates have helped Chinese firms create advanced LLMs.

Gemma Conroy & Smriti Mallapaty



### News

23 Jan 2025

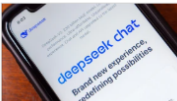
Nature

Volume: 638, P: 13-14

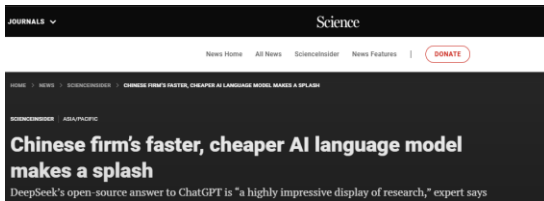
#### China's cheap, open AI model DeepSeek thrills scientists

DeepSeek-R1 performs reasoning tasks at the same level as OpenAI's o1 — and is open for researchers to examine.

Elizabeth Gibney



Nature杂志一周内 (2025 1.23-1.30) 发表4篇文章报道DeepSeek



NEWS | IN DEPTH



ARTIFICIAL INTELLIGENCE

## Chinese firm's large language model makes a splash

DeepSeek's open-source model appears to be cheaper and faster to train and run than many others

Science杂志2025.1.15日, 深度报道DeepSeek



# DeepSeek vs OpenAI



上海计算机软件技术开发中心



DeepSeek成立于2023年7月，由幻方量化孕育而生，总部杭州，2025年1月20日发布DeepSeek R1开源推理模型而风靡全球。

## DeepSeek V2.5 / V3

**特点：开源，高性价比通用型大语言模型**

## DeepSeek R1

**特点：开源，高效处理复杂推理任务的推理模型**



OpenAI成立于2015年12月，总部美国旧金山，2022年11月发布ChatGPT而成为全球AI领先企业。

## ChatGPT 系列

### ChatGPT 3.5/GPT 4.0/GPT 4o

**特点：闭源，多模态通用型大(语言)模型**

## OpenAI o系列

### OpenAI o1/ OpenAI o3

**特点：闭源，专注于解决复杂任务的推理模型**

**DeepSeek和OpenAI都是颠覆性创新者，目标都是AGI，都拥有高密度的人才**



# DeepSeek 为什么有如此大的“影响力”



## 开源，性能对齐OpenAI-o1正式版

DeepSeek-R1在后训练阶段使用“**纯强化学习**”技术，极大提升了模型推理能力在数学、代码、自然语言推理等任务。

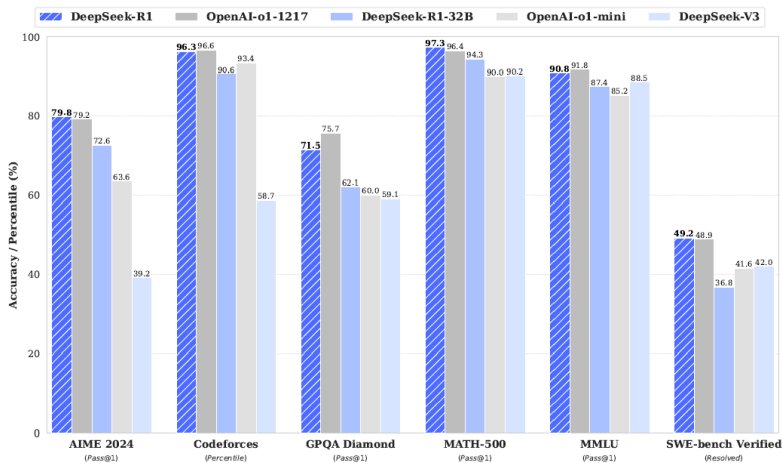


Figure 1 | Benchmark performance of DeepSeek-R1.

## 使用纯强化学习就能学会强推理

DeepSeek-R1在纯强化学习过程中自然地学会了使用更多的“**思考时间**”来解决推理任务。

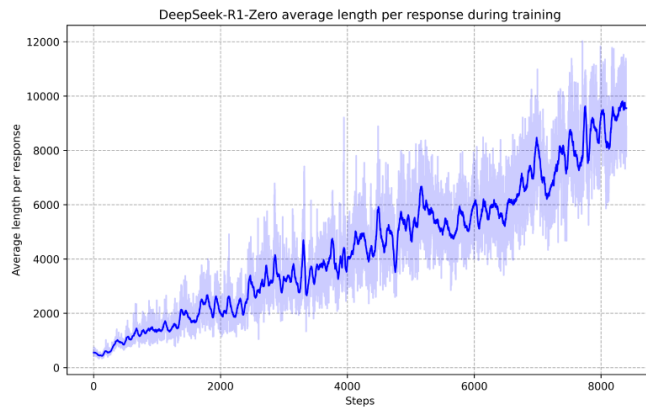


Figure 3 | The average response length of DeepSeek-R1-Zero on the training set during the RL process. DeepSeek-R1-Zero naturally learns to solve reasoning tasks with more thinking time.

- ❑ 打破了“高算力和高投入是发展人工智能唯一路径”和“集成电路制程优势=人工智能技术霸权”的迷信；
- ❑ 中国科技公司从“追赶者”变为“规则改写者”，以颠覆性的创新（高性价比、开源）挑战美国的AI霸权

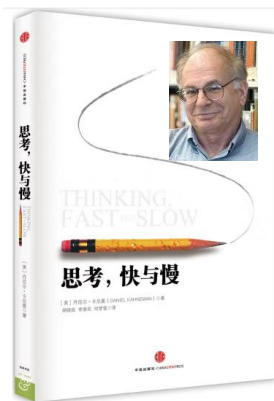
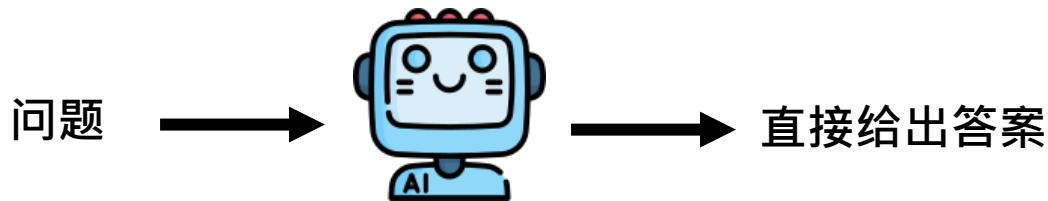


# 大模型展现“思考推理”的过程

## □ (普通) 大语言模型

ChatGPT 3.5/GPT 4.0, DeepSeek V3

系统1 (快速、直觉型)

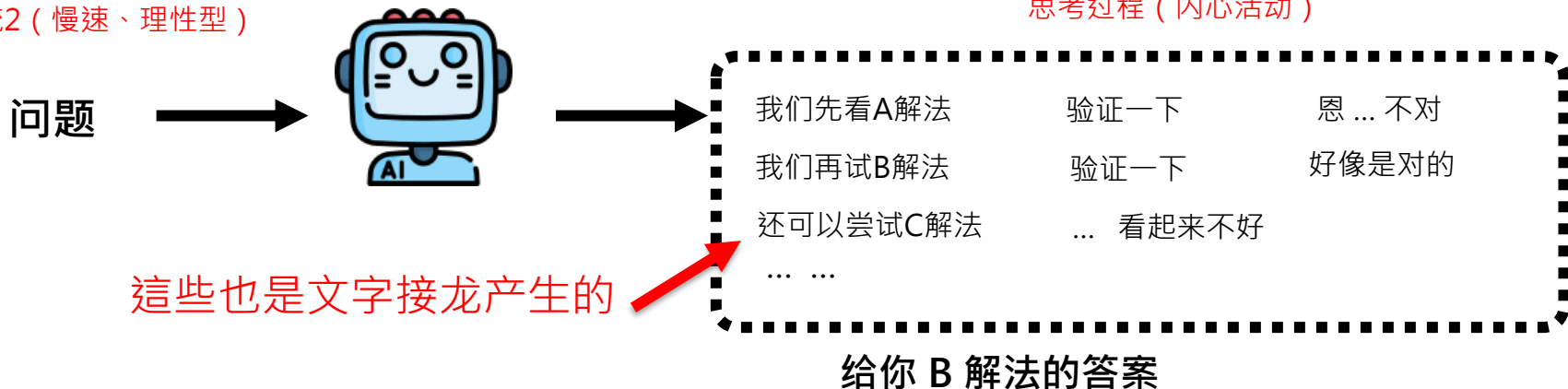


## □ 大语言 (推理) 模型

ChatGPT o1, o3, DeepSeek-R1, Grok3, Gemini Flash

系统2 (慢速、理性型)

思考过程 (内心活动)



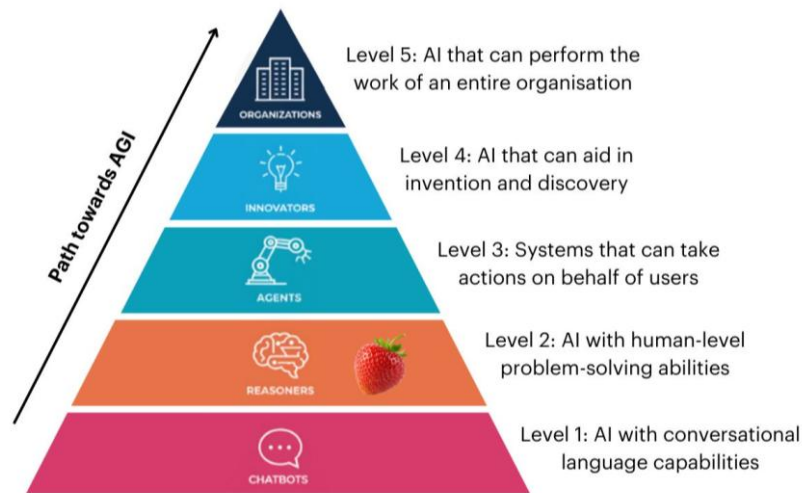


## 人工智能发展阶段

Level 1	<b>Chatbots</b> 聊天机器人	AI with conversational language AI具有语言对话能力 (ChatGPT)
Level 2	<b>Reasoners</b> 推理者	AI with human-level problem solving AI具有人类水平的问题解决能力 (o1/o3推理模型)
Level 3	<b>Agents</b> 智能体	Systems that can take actions 能采取行动的AI系统
Level 4	<b>Innovators</b> 创新者	AI that can aid in invention 能辅助发明创造的AI
Level 5	<b>Organizations</b> 组织者	AI that can do the work of an organization 可以完成组织工作的AI

## The 5 Levels of AI

(OpenAI Classification System)



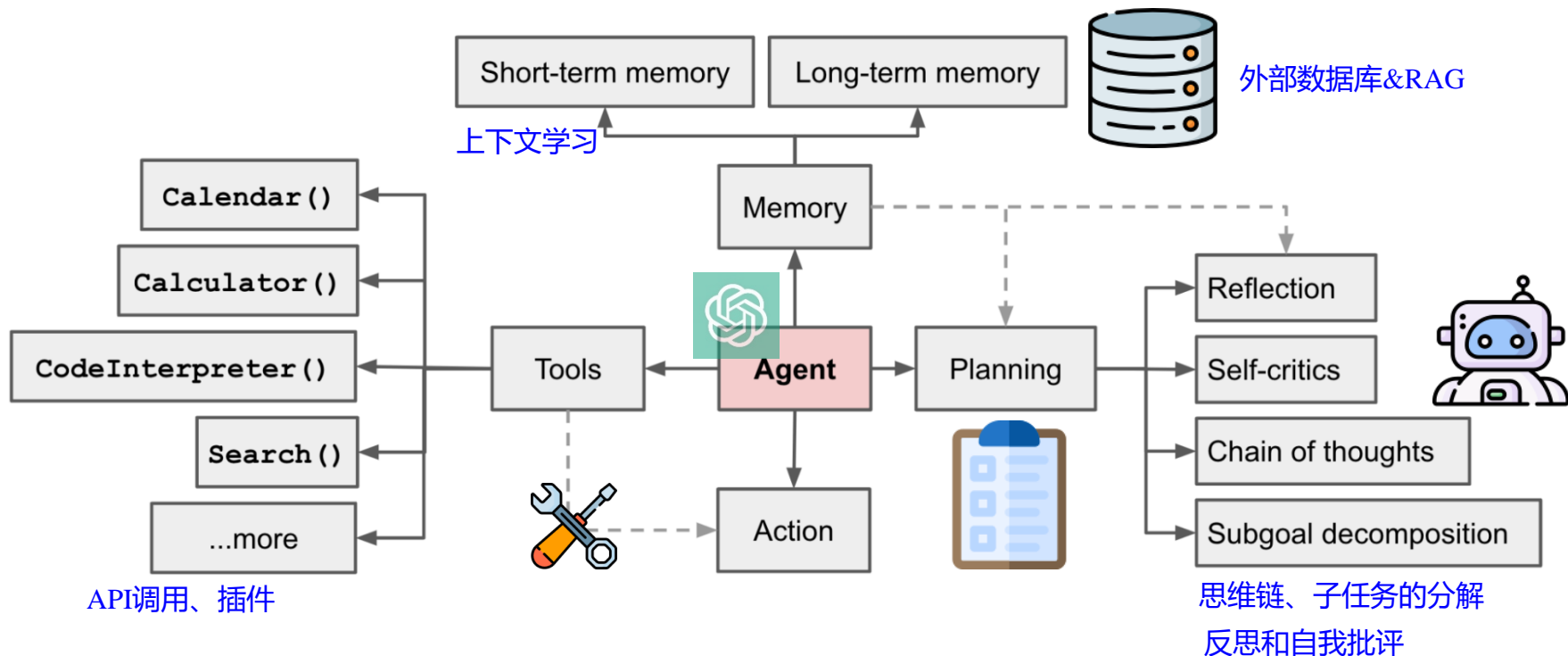
是否可以利用AI工具将左边的文字生成右边的概念图?



# 基于大语言模型的AI Agent

AI 智能体是有记忆、会推理规划、会使用工具、作出行动和决策的自主AI系统。

AI Agent = LLM + Memory + Planning Skills + Tool Use



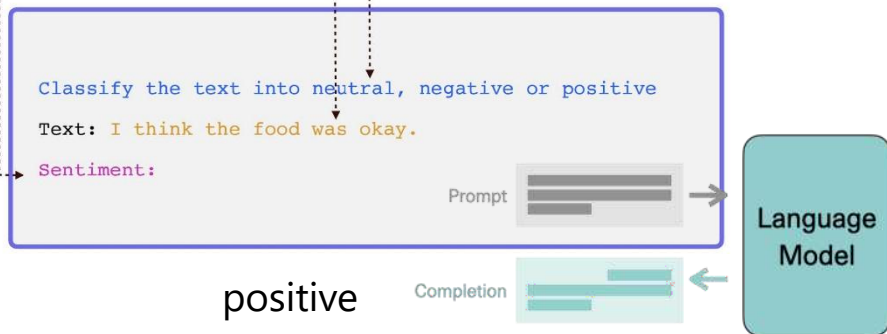


# 提示 (词) Prompt

- 我们将给大模型的输入称为**Prompt (提示)**，将大模型返回的输出称为**Completion (补全)**
- 提示 (词) 是用户和大模型交互的入口，实质上是 “用自然语言编写的与大模型交互的计算机指令”。

A prompt is composed with the following components:

- **Instructions**
- **Context**
- **Input data**
- **Output indicator**



- **Instruction(指令)**: 希望大模型执行的具体任务，任务必须明确说明。
- **Context (上下文/语境)**: 可引导大模型做出更好响应的附加信息或背景。例如，告诉大模型输出应在某个研究主题下生成研究的背景和意义。
- **Input Data(输入数据)**: 告诉大模型指令要处理的数据。向大模型清晰呈现这些数据可提高响应的质量。
- **Output indicator(输出指示)**: 告诉大模型期望输出的指示，比如输出的内容、格式、长度。

Demo: DeepSeek总结科技文章的主要观点



# DeepSeek三种使用模式对比



上海科学院  
SHANGHAI ACADEMY  
OF SCIENCE & TECHNOLOGY



上海计算机软件技术开发中心

使用模式	DeepSeek V3	深度思考 R1	联网搜索
定位	通用语言模型 (2024.12.26发布), 适合快速响应的一般问答	推理模型 (2025.1.20发布), 适合复杂推理和深度分析任务	适合需要结合网络信息回答问题 (DS训练语料更新至 <b>2024年7月</b> )
技术特点	响应低延时, 采用 <b>MTP</b> (Multi-Token Prediction)、 <b>MLA</b> (Multi-head Latent Attention)等创新技术	<b>纯强化学习</b> 驱动, 支持长 <b>思维链 (CoT)</b> 推理, 提供详细的思考和推理步骤	<b>RAG (检索增强生成)</b> , 实时网页数据抓取、智能语义理解、动态排序, 实现信息检索和答案生成
<b>适用的基本原则</b>	适用于快速问答、百科知识查询、“ <b>规范性</b> ” 文本生成任务 (明确的框架约束或模版要求)	适用于数理逻辑推理、编程及“ <b>开放性</b> ” 文本生成任务 (没有固定框架的约束)	适用于获取最新资讯、市场动态、学术研究成果等 <b>时效性强</b> 的信息
典型应用场景	常识问答、公文写作、新闻报道、活动策划、一般性内容创作	数学、程序设计、代码Debug、科学问题探索、思辨性文章的撰写	最新的行业分析报告撰写、文献综述撰写



# RTGO提示语 (Prompt) 结构

## Role (角色)

### 定义AI的角色:

- 人工智能领域的资深研究员;
- 经验丰富的数据分析师;
- 金融行业的分析师;
- ...

## Task (任务)

### 具体任务描述:

- 写一份关于XX事件的舆论分析报告  
(XX活动/事件相关背景信息如下.....)
- 写一份关于XXX活动的活动方案、  
品牌文案;
- ...

## Goal (目标)

### 希望达成的目标效果:

- 通过该报告为相关企业管理者  
提供.....策略支撑;
- 通过该文案吸引潜在客户, 促  
成消费.....
- ...

## Objective (操作要求)

### 输出的指示:

- 字数要求、段落结构、用词风格、  
内容要点、输出格式...
- 不希望输出的内容、讨论的主题
- ...





# 打开DeepSeek的正确方式 (提示词库)

## 提示库

探索 DeepSeek 提示词样例，挖掘更多可能



### 代码改写

对代码进行修改，来实现纠错、注释、调优等。



### 代码生成

让模型生成一段完成特定功能的代码。



### 结构化输出

将内容转化为 Json，来方便后续程序处理



### 角色扮演 (情景续写)

提供一个场景，让模型模拟该场景下的任务对话



### 诗歌创作

让模型根据提示词，创作诗歌



### 宣传标语生成

让模型生成贴合商品信息的宣传标语。



### 代码解释

对代码进行解释，来帮助理解代码内容。



### 内容分类

对文本内容进行分析，并对齐进行自动归类



### 角色扮演 (自定义人设)

自定义人设，来与用户进行角色扮演。



### 散文写作

让模型根据提示词创作散文



### 文案大纲生成

根据用户提供的主题，来生成文案大纲



### 模型提示词生成

根据用户需求，帮助生成高质量提示词



# 大模型赋能学习、工作与研究 (演示)

- ◉ 利用大模型了解阅读前沿的科技信息
- ◉ 大模型辅助策划活动方案和公文写作
- ◉ 大模型赋能科研



02

PART TWO



## nature

Explore content | About the journal | Publish with us | Subscribe

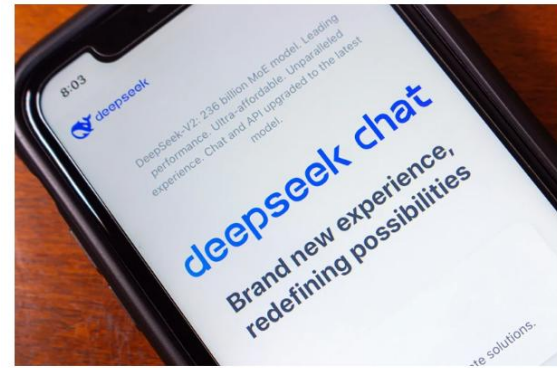
nature > news > article

NEWS | 23 January 2025

### China's cheap, open AI model DeepSeek thrills scientists

DeepSeek-R1 performs reasoning tasks at the same level as OpenAI's o1 – and is open for researchers to examine.

By Elizabeth Gibney



Chinese firm DeepSeek debuted a version of its large language model last year. Credit: Koshiro K/Alamy

d41586-025-00229-6

News in focus

Brand new experience, redefining possibilities

CHINA'S CHEAP, OPEN AI MODEL DEEPSEEK THRILLS SCIENTISTS

DeepSeek-R1 performs reasoning tasks at the same level as OpenAI's o1 – and is open for researchers to examine.

文章概述

中国人工智能公司DeepSeek最近发布了一款名为DeepSeek-R1的大规模语言模型 (LLM)，该模型在推理任务上的表现与OpenAI的o1模型相当，并且对研究人员开放。允许研究者研究和改进算法。与其他大型语言模型相比，DeepSeek-R1的训练成本低廉，约为o1的三分之一，而且提供了小规模“精简版”以适应资源有限的研究者。此模型通过强化学习进行优化，能有效减少训练和运行成本。同时，由于其开源性质，研究人员可以深入理解模型的思考过程，提高模型的可解释性。DeepSeek的成功挑战了国际上对于AI技术发展的看法，尤其是在计算机硬件限制下仍能取得进展的可能。此外，报道还提到另一家中国的AI初创公司通过低成本策略在对话机器人领域获得成功，暗示了资源高效利用的重要性。与此同时，曾经辉煌的消费级基因检测公司23andMe面临困境，显示出科技行业竞争的激烈以及消费者隐私保护的重要性。

关键要点

1. 中国公司DeepSeek推出了一款名为DeepSeek-R1的大语言模型，其性能与OpenAI的o1相当，并且开放供研究人员研究和使用的。
2. DeepSeek-R1是一款廉价而开放的“推理”模型，能够比早期的语言模型更好地解决科学问题，在化学、数学和编码等任务上的表现与o1相当。
3. DeepSeek-R1的训练数据尚未公开，但该公司已创建了迷你版本，允许研究人员测试其能力。

DeepSeek是什么?

Demo: 导读、全文翻译、脑图和笔记

不足: 翻译还未达到专业水平、文档解析不精准



	美国的AI发展	中国的AI发展
政策与资本	市场机制与风险资本驱动，政策宽松化以释放创新活力	中国政府顶层设计、政策引导和市场规模推动AI发展
发展路径	<ul style="list-style-type: none"><li>• 依赖<b>高投入、高算力</b>模式</li><li>• 专注消费级应用（如聊天机器人、智能体）、机器人、AI for Medicine等</li></ul>	<ul style="list-style-type: none"><li>• 算法和模型架构优化，<b>低成本</b>模式</li><li>• 鼓励人工智能与产业的融合（如大模型与工业场景、政务场景等融合），加快人工智能<b>应用落地</b></li></ul>
竞争态势	<ul style="list-style-type: none"><li>• 注重芯片、AI框架和核心技术等AI生态完整性，在AI芯片和核心技术保持领先</li><li>• 美国聚焦“前沿AI技术”保持<b>技术霸权</b>，对中国实行高端AI芯片出口管制，核心技术趋于封闭</li></ul>	<ul style="list-style-type: none"><li>• 中国通过新型举国体制，正通过自主技术创新，缩小与美国在AI核心技术上的差距</li><li>• “<b>基础层追赶、技术层总体跟跑局部领跑、应用层并跑</b>”</li></ul>



## nature

Explore content ▾ About the journal ▾ Publish with us ▾ Subscribe

[nature](#) > [news](#) > article

NEWS | 06 February 2025

### OpenAI's 'deep research' tool: is it useful for scientists?

The model produces cited, pages-long reports that might be helpful for generating literature reviews.

By Nicola Jones



OpenAI's chief executive Sam Altman announced the release of its 'deep research' tool. Credit: Franck

<https://www.nature.com/articles/d41586-025-00377-9>

OpenAI 'deep research' ▾

Hi, 我是 Kimi ~  
很高兴遇见你! 你可以随时把网址 📄 或者文件 📁 发给我, 我来帮你看看

请阅读 <https://www.nature.com/articles/d41586-025-00377-9> 网页内容, 并提炼核心内容生成PPT大纲

以下是一个基于网页内容《OpenAI's 'deep research' tool: is it useful for scientists?》的PPT大纲:

#### 幻灯片 1: 封面

- 标题: OpenAI的“深度研究”工具: 对科学家有用吗?
- 副标题: 探索AI在科研中的新应用
- 日期: 2025年2月12日

#### 幻灯片 2: 目录

- OpenAI“深度研究”工具介绍
- 工具的功能与优势
- 科学家的使用体验
- 潜在问题与局限性
- 未来展望
- 总结

#### 幻灯片 3: OpenAI“深度研究”工具介绍

- 发布背景
  - OpenAI推出“深度研究”工具, 旨在为科研人员提供更高效率的信息整合服务。
- 工具特点
  - 付费使用, 可将数十甚至数百个网站的信息整合成一份多页的引用报告。
  - 类似于Google在2024年12月发布的类似工具, 但结合了o3大型语言模型(LLM)的推理能力与互联网搜索功能。

## Demo: 生成PPT大纲

<https://kimi.moonshot.cn/chat/cum2slrfj2jem8uo4mug>



## DeepSeek制作PPT大纲

你的角色是PPT设计制作大师，我需要你帮我做一份关于“大模型进展与应用”的PPT大纲，先介绍大模型技术进展，再讨论大模型典型应用和未来趋势，然后输出PPT大纲。



Q 已深度思考 (用时 35 秒)

好的，用户需要我帮忙制作一份关于“大模型进展与应用”的PPT大纲，先介绍技术进展，再讨论应用和未来趋势，最后输出大纲。首先，我需要明确用户的需求是什么。用户可能是一位需要在会议上做报告的业内人士，或者是给学生讲课的教师，也可能是企业中的技术负责人，需要向管理层或客户展示大模型的价值。不同的受众会影响PPT的结构和内容深度。

以下是为您设计的「大模型进展与应用」PPT大纲框架，兼顾技术深度与观众理解，采用模块化结构便于灵活调整：

### 封面页

标题：大模型技术革命：进展、应用与未来

副标题：从算法突破到行业变革

视觉建议：AI生成的概念图（神经网络+行业图标环绕）

### 目录页

- 1. 大模型技术演进里程碑
- 2. 行业级应用场景解析
- 3. 未来趋势与挑战

视觉建议：动态3D导航图标

### Part 1: 技术突破演进 (6页)

#### 1.1 核心能力跃迁史

- 参数规模指数增长 (2018-2024对比图)
- 从单一模态到多模态理解
- 涌现能力：思维链/上下文学习突破

## Kimi+PPT 细化PPT内容

### 大模型技术革命：进展、应用与未来

#### 1. 大模型技术演进里程碑

##### 1.1 核心能力跃迁史

###### 1.1.1 参数规模指数增长

- 2018年BERT参数量仅1.1亿，2024年GPT-4参数量达万亿，参数规模指数增长，模型性能大幅提升。
- 参数规模增长使模型能学习更复杂模式，如自然语言处理中精准语义理解，图像识别中复杂纹理识别。

###### 1.1.2 多模态理解发展

- 从单一模态到多模态理解，模型可同时处理文本、图像、语音等数据，如CLIP实现图文匹配。
- 多模态理解让模型更接近人类认知，应用于自动驾驶场景，可融合摄像头图像、雷达数据等。

###### 1.1.3 涌现能力突破

- 思维链让模型具备逻辑推理能力，如解决数学问题时可分步思考；上下文学习使模型根据上下文动态调整理解。
- 涌现能力提升模型智能水平，如在阅读理解任务中，能准确把握文章主旨并回答复杂问题。

## 一键生成可编辑的PPT





# 利用Kimi阅读Nature新闻，并总结制作PPT(续)



OpenAI深度研究工具在科研中的价值

已阅读 37 个网页

## OpenAI的“深度研究”工具：对科学家有用吗？

### 1. OpenAI“深度研究”工具介绍

#### 1.1 发布背景与目标

##### 1.1.1 高效信息整合需求

- 科学家在研究中需耗费大量时间整合多源信息，传统方法效率低。
- OpenAI推出“深度研究”工具，旨在快速整合网络信息，提升科研效率。

##### 1.1.2 技术发展推动

- 人工智能与自然语言处理技术的进步，为工具开发提供了技术基础。
- 结合o3模型的推理能力与互联网搜索功能，实现高效信息处理。

##### 1.1.3 竞争与合作背景

- 谷歌于2024年12月发布类似工具，促使OpenAI加速研发。
- 二者在功能与性能上各有侧重，推动AI在科研领域的应用竞争。

#### 1.2 工具特点与优势

##### 1.2.1 强大的信息整合能力

- 能从数十甚至数百个网站整合信息，生成多项引用报告。
- 有效处理多模态数据，包括文本、图像和PDF文件。

##### 1.2.2 高效的推理与分析

- 利用o3模型的推理能力，深入分析复杂问题。
- 在“人类终极考试”中以26.6%的准确率创下新高。

##### 1.2.3 用户友好与透明性

- 输出结果附有清晰的引用和推理过程总结。

随时@你想要的Kimi+ 使用各种能力

已联网

OpenAI的“深度研究”工具：对...

目 保存 ▷ 放映 高 拼图 下载

大纲编辑 模板替换 插入元素

文字设置 形状设置 背景设置 图片设置 表格设置 图表设置

幻灯片 1 / 25 70% +

Demo：根据PPT大纲，细化PPT内容，生成可编辑的PPT、播放与下载

<https://kimi.moonshot.cn/chat/cum2slrfj2jem8uo4mug>



# 使用AI (Napkin) 生成视觉图像

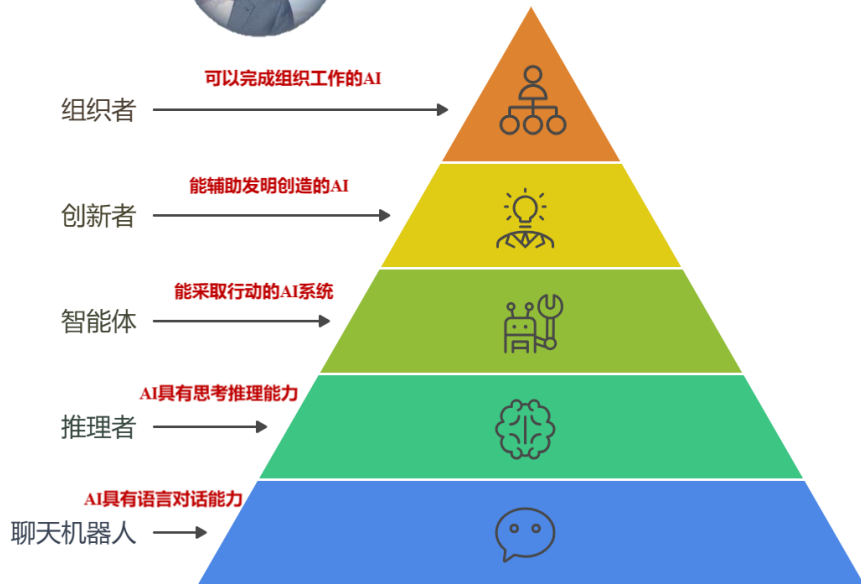
<https://app.napkin.ai/>

## 人工智能发展阶段

Level 1	<b>Chatbots</b> 聊天机器人	<b>AI with conversational language</b> AI具有语言对话能力
Level 2	<b>Reasoners</b> 推理者	<b>AI with human-level problem solving</b> AI具有人类水平的问题解决能力
Level 3	<b>Agents</b> 智能体	<b>Systems that can take actions</b> 能采取行动的AI系统
Level 4	<b>Innovators</b> 创新者	<b>AI that can aid in invention</b> 能辅助发明创造的AI
Level 5	<b>Organizations</b> 组织者	<b>AI that can do the work of an organization</b> 可以完成组织工作的AI



OpenAI通往通用人工智能的道路



Demo: 利用Napkin通过文字生成视觉图像



## 公文写作AI提示词公式

1 角色

2 文体类型

3 核心要素

4 限制条件

### 1.通知类公文

我是（单位/部门）的（某个职务），请你帮起草一份关于会议（具体事项）的通知，需包含（时间）、（地点）、（人物），（也就是具体要求）做到语言简洁凝练，符合《党政机关公文格式》要求。

**举例：**我是组织部的办公室主任，请你帮起草一份关于党建会议的通知，需包含2月19日下午3点、在组织部会议室、各街道组织委员参加，语言简洁凝练，符合《党政机关公文格式》要求。

### 3.请示类公文

我是（单位）的，请你帮我写一份以单位的名义向（上级部门名称）提交关于（具体事项）的请示，要阐明（请示的背景）、（请示的具体内容）、（我们采建议的方案），结尾使用'妥否，请批示'等规范用语。

**举例：**我是X区民政局的，请你帮我写一份以XX区民政局的名义向区人民政府提交关于购置一批社保中心的电脑的请示，需说明现有电脑老化现状、申请经费50元、公开招标采购计划，附设备清单及预算表。

### 2.报告类公文

写一份（单位）2024年度（某项工作）的工作报告，要求包含（具体成果）、（尚需改进的问题）、（改进措施）三个部分，参照最新政策文件进行表述，字数控制在1200字以内。

**举例：**写一份XX县2024年度住房安全保障工作报告，要求引用《住房安全保障法》的条款，包含政策落实情况、排查与整改情况、资金使用情况、工作成效与未来计划、具体措施与成效，将近三年数据进行对比。

### 4.函（平行/不相隶属机关）

以（发函单位）名义致（收函单位）的函，内容为（具体事项），需阐明（具体需求）、（什么时间）、（联系方式），措辞要合乎身份，不用要求、必须等命令性词汇。

**举例：**以XX市生态环境局名义致XX大学环境科学与工程学院的函，内容为联合开展城市空气质量监测研究课题研究，希望贵院能够提供近年来城市空气质量监测的相关数据，联系人李科长138xxxxxxx。

参考：[1]体制内写材料，怎么用DeepSeek写公文，附实操，小白也立刻能上手，<https://zhuanlan.zhihu.com/p/2304255080>

[2] DeepSeek你用对了吗？三大写作场景（公文写作、新闻报道、广告文案）提示词技巧来了！（公众号文章）



# 用大模型(如DeepSeek) 编写策划方案



上海计算机软件开发中心



请以上海建设具有全球影响力的科创中心十周年为背景，为妇女科技创新工作，设计品牌方案



## 品牌名称：巾帼闪耀科创

**一、品牌理念：**女性的智慧和创造力是推动科技创新的重要力量，巾帼闪耀科创致力于激发女性在科技创新领域的潜力，推动上海建设具有全球影响力的科创中心。

## 二、品牌形象：

**标志设计：**采用简洁明快的线条，以巾帼形象为核心元素，展现女性坚韧与柔美的结合。背景为科创的图案，寓意女性在科技创新领域的引领作用。标志色彩采用深蓝色，象征科技与智慧的结合。

**宣传语：**巾帼科技力量，点亮科创未来

**三、品牌活动：1. 巾帼科创论坛：**定期举办论坛，邀请国内外知名女性科技工作者分享经验与见解，促进女性科技工作者的交流与合作。

**2. 巾帼科创项目支持计划：**为优秀的女性科技创新项目提供资金、资源等方面的支持，助力项目落地实施。

**3. 巾帼科创合作平台：**搭建合作平台，促进女性科技工作者与企业、研究机构之间的合作，推动科技成果的转化与应用。

**4. 巾帼科创成果展示活动：**定期举办成果展示活动，展示女性科技工作者的优秀成果，提升社会对女性科技工作者的关注与认可。



搜索



聊天

**请帮我设计一个Logo：**采用简洁明快的线条，以巾帼形象为核心元素，展现女性坚韧与柔美的结合。背景为科创的图案，寓意女性在科技创新领域的引领作用。标志色彩采用深蓝色，象征科技与智慧的结合。



A logo with a woman's image and a science and technology background

Designer | 1024 x 1024.jpg | 2分钟前

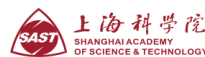
分享 保存 下载 反馈

内容预览

使用 AI 生成 · 2023年12月26日 下午1:48



# 如何利用DeepSeek寻找研究主题



上海计算机软件技术开发中心



我是 DeepSeek，很高兴见到你！

我可以帮你写代码、读文件、写作各种创意内容，请把任务交给我吧~

请给我10个近三年“人工智能赋能企业智能化转型”方面热门的研究课题。

深度思考 (R1)

联网搜索



内容由 AI 生成，请仔细甄别

## 第1轮

请给我10个近三年“人工智能赋能企业智能化转型”方面前沿研究课题

## 第2轮

我对“企业级大模型 (LLM) 的定制化与应用优化”感兴趣，请给我这个方向更多的研究主题

## 第3轮

.....

Demo: 多轮人机交互，迭代式找到研究灵感和主题



# 利用DeepSeek V3 撰写文献综述



上海计算机软件开发中心



我正在准备撰写一篇论文综述，题目是“企业级大模型领域知识注入与增强”，请根据这个题目，帮我撰写文献综述。要求包含：按照研究背景、研究价值、现有的技术的分类和比较，未来可能的研究思路这几部分内容。

## 企业级大模型领域知识注入与增强：文献综述

### 1. 研究背景

随着人工智能技术的快速发展，大模型（Large Language Models, LLMs）在企业智能化转型中扮演了越来越重要的角色。然而，通用大模型在特定行业或企业场景中往往表现不足，主要原因在于缺乏对领域知识的深度理解与适配。企业级大模型需要通过领域知识注入与增强，提升其在垂直场景中的专业性和实用性

1 6。

近年来，企业级大模型的应用场景不断扩展，涵盖金融、医疗、制造等多个领域。例如，在金融领域，大模型被用于智能投顾和风险评估；在医疗领域，大模型辅助疾病诊断和治疗方案制定 6 8。然而，这些场景对模型的准确性、可解释性和领域适配性提出了更高要求，推动了领域知识注入与增强技术的研究与发展

3 7。

### 2. 研究价值

领域知识注入与增强技术的研究具有重要的理论价值和实践意义：

- **理论价值**：探索如何将领域知识高效融入大模型，推动模型架构和训练方法的创新，为人工智能领域知识表示与推理提供新思路 7 10。
- **实践价值**：通过领域知识增强，企业级大模型能够更好地满足垂直场景需求，提升业务效率，降低人工成本，推动企业智能化转型 6 8。

### 3. 现有技术分类与比较

#### 3.1 领域知识注入方法

- **预训练阶段注入**：在模型预训练阶段引入领域相关数据，通过大规模语料训练使模型学习领域知识。例如，金融领域大模型通过引入财经新闻、财报等数据增强金融知识理解 6 10。
- **微调阶段注入**：在通用大模型基础上，使用领域特定数据进行微调。例如，医疗领域大模型通过电子病历数据微调，提升疾病诊断能力 7 8。
- **检索增强生成（RAG）**：通过外部知识库检索相关信息，增强模型生成内容的准确性和专业性。RAG 2.0 和 Graph RAG 等技术进一步提升了知识检索的效率和准确性 1 3。

#### 3.2 领域知识增强技术

- **多模态融合**：结合文本、图像、音频等多模态数据，提升模型对复杂场景的理解能力。例如，制造业大模型通过结合设备传感器数据和维修手册，实现设备故障预测与维修建议 6 10。
- **知识图谱嵌入**：将领域知识图谱嵌入模型，增强模型的推理能力。例如，法律领域大模型通过嵌入法律条文和案例知识图谱，提升法律咨询的准确性 7 8。
- **轻量化与端侧部署**：通过模型压缩和量化技术，将大模型部署到端侧设备，满足实时性和隐私保护需求。例如，AI手机中的端侧大模型能够在不依赖云端的情况下完成复杂任务 1 10。

### 4. 未来研究思路

#### 4.1 领域知识的高效表示与融合

未来研究可探索更高效的领域知识表示方法，例如基于神经网络的知识表示学习，以及多源异构数据的融合技术，提升模型对领域知识的理解与利用能力 7 10。

#### 4.2 动态知识更新与持续学习

针对领域知识的动态变化，研究如何实现模型的持续学习与知识更新，避免模型过时。例如，结合在线学习和增量学习技术，使模型能够实时吸收新知识 3 8。

#### 4.3 可解释性与可信增强

提升模型的可解释性，使其能够清晰展示推理过程，增强用户信任。例如，结合因果推理和可解释AI技术，使模型输出更具说服力 6 10。

#### 4.4 轻量化与端侧优化

进一步优化模型轻量化技术，降低部署成本，提升端侧设备的计算效率。例如，探索更高效的模型压缩和量化方法，满足企业对实时性和隐私保护的需求 1 10。



# 利用DeepSeek R1 撰写文献综述



上海计算机软件开发中心

我正在准备撰写一篇论文综述，题目是“企业级大模型领域知识注入与增强”，请根据这个题目，帮我撰写文献综述。

## 企业级大模型领域知识注入与增强：研究进展与挑战

### 1. 领域知识注入的必要性

企业级大模型需解决通用大模型在垂直场景中的局限性，如专业深度不足、行业术语理解偏差、数据隐私与合规性等问题。研究表明，通用大模型（如GPT-4、Claude3）虽具备强大的泛化能力，但在工业制造、医疗、金融等领域，其输出常因缺乏行业知识而出现“幻觉”或逻辑错误<sup>[3][5]</sup>。例如，医疗诊断场景中，通用模型可能因缺乏医学文献和临床指南的深度训练，导致建议不可靠<sup>[8]</sup>。因此，通过领域知识注入实现模型专业化，已成为企业级应用落地的核心需求。

### 2. 知识注入的主要技术路径

#### (1) 领域微调与数据增强

通过引入行业专有数据集（如金融交易记录、医疗影像报告）对预训练模型进行微调，并结合自监督学习优化模型性能。例如，天融信在网络安全领域构建的垂直大模型，通过融合安全日志与攻击特征数据，实现了威胁检测准确率的显著提升<sup>[6]</sup>。此外，多模态数据融合技术（如文本-图像联合训练）被用于增强模型对复杂场景的理解能力，如工业质检中的缺陷识别<sup>[7][10]</sup>。

#### (2) 知识图谱与规则引擎的集成

将领域知识图谱（如法律条款库、药物相互作用数据库）嵌入模型推理过程，通过规则约束减少错误输出。例如，腾讯研究院提出“知识增强型大模型”，在金融风控场景中结合知识图谱实现动态风险评估，解决了传统模型在因果推理上的不足<sup>[5]</sup>。

#### (3) 模块化架构与复合系统

2025年兴起的复合人工智能系统（由多个协同的小模型组成）通过任务分解与知识模块化，实现了灵活的知识更新与高效推理。例如，谷歌的AlphaCode2通过将代码生成任务拆解为语法检查、逻辑验证等子模块，并结合领域知识库，将任务成功率提升至80%以上<sup>[2]</sup>。

### 3. 实际应用与商业化进展

#### (1) 行业落地案例

- **金融领域**：银行通过引入领域知识增强的AI模型，实现客户信用评估的自动化与精准化，处理速度较传统方法提升3倍<sup>[9]</sup>。
- **制造业**：成都某企业利用大模型优化生产流程，通过注入设备维护日志与工艺参数，实现故障预测准确率提升40%<sup>[11]</sup>。
- **医疗领域**：结合医学文献与临床数据训练的模型，在辅助诊断中达到与资深医生相当的准确率，同时支持多模态输入（如病理图像与电子病历）<sup>[8]</sup>。

#### (2) 商业化挑战

- **数据孤岛与隐私保护**：企业间数据共享壁垒导致高质量领域数据获取困难，需依赖联邦学习等技术实现隐私合规下的知识融合<sup>[5][8]</sup>。
- **算力成本与部署效率**：模型轻量化成为关键趋势，华为云与腾讯音乐合作的MUSELight引擎通过优化推理架构，将图像生成速度提升166%，成本降低35%<sup>[10]</sup>。
- **动态知识更新**：行业知识快速迭代要求模型支持在线学习，但当前技术仍面临灾难性遗忘与稳定性问题<sup>[7]</sup>。

### 4. 未来研究方向

#### (1) 知识动态注入机制

探索增量学习与持续预训练技术，结合知识图谱的实时更新，实现模型对行业变化的快速适应<sup>[7]</sup>。

#### (2) 可解释性与可信增强

开发基于注意力机制的可视化工具，使领域专家能够验证模型决策逻辑，提升用户信任<sup>[5]</sup>。

#### (3) 跨领域知识迁移

研究多任务联合训练框架，促进不同行业间知识的复用。例如，制造业的故障诊断模型可迁移至能源设备维护场景<sup>[11]</sup>。

#### (4) 开源生态与标准化

上海“模型申城”等计划推动开源社区建设，通过共享预训练模型与工具链，降低企业技术门槛<sup>[12]</sup>。

### 5. 总结

企业级大模型的领域知识注入与增强，正从单一技术优化转向系统性工程，涵盖数据治理、算法创新、算力部署与合规框架。未来，随着复合系统架构的成熟与开源生态的完善，模型将更深度融入企业核心流程，成为数字化转型的“智慧中枢”<sup>[2][12]</sup>。然而，如何平衡通用能力与专业精度、解决动态场景下的知识更新难题，仍需学术界与工业界的协同突破。



The screenshot shows the Metaso AI search website. At the top, a purple banner reads "秘塔AI接入满血版DeepSeek R1推理模型" (Metaso AI integrates the full-blooded DeepSeek R1 reasoning model). The main header features the Metaso logo and the slogan "没有广告, 直达结果" (No ads, direct results). A search bar is highlighted with a blue border, containing the text "请输入, Enter键发送, Shift+Enter键换行" (Please enter, press Enter to send, Shift+Enter to line wrap). Below the search bar, there are buttons for "简洁" (Simple), "深入" (Deep), and "研究" (Research). At the bottom, there are several search suggestions: "世界各地的动植物多样性有何特点?" (What are the characteristics of biodiversity in various parts of the world?), "腾讯被列入美国防务名单" (Tencent is listed on the US defense list), "暗物质和暗能量对宇宙学理论的挑战" (Challenges of dark matter and dark energy to cosmological theory), and "小米SU7" (Xiaomi SU7).

Demo: 大模型赋能科学研究



# 总结与思考





## 大模型的能力边界



图来自网络

- 幻觉，一本正经胡说八道；
- 具备较强的语言理解能力，但目前创新能力仍然不足
- 可靠性（一致性）问题，两次提问，大模型的输出可能不同
- 训练语料的限制，训练语料的时间、数量、内容等
- 上下文（Context）大小的限制、推理成本和推理效率等

## 大模型辅助学习、工作与研究的思考

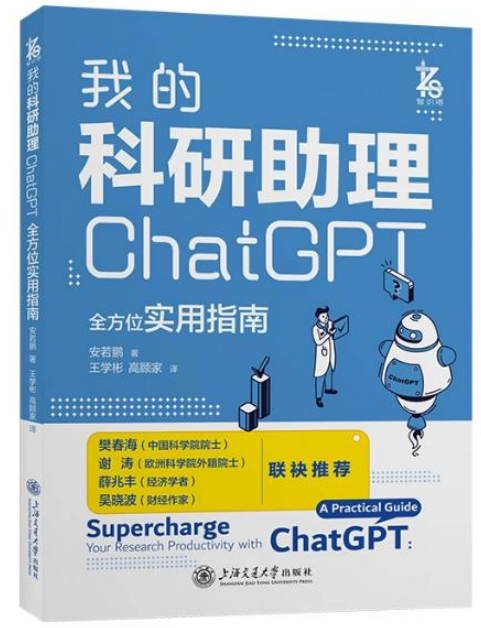
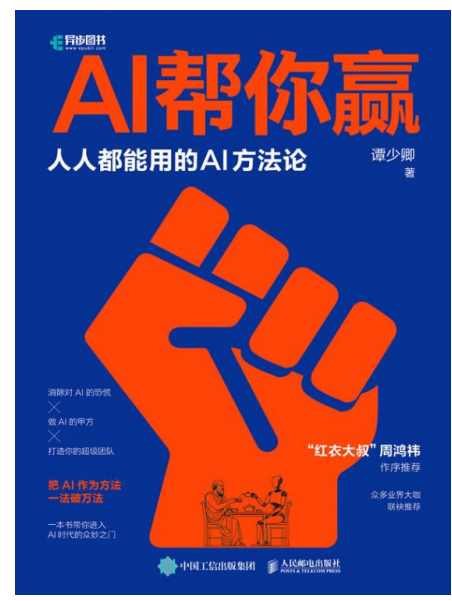


图由AI生成

- 将大模型作为助手，我们自己是学习与研究的主导者；
- 大模型可能产生“幻觉”，我们需要尽责核实内容；
- 使用大模型辅助工作与科研务必注意数据安全；
- AI时代**提问**、**概念联结**和**价值判断**是个人成长的核心能力；
- 与大模型的人机协作，将成为学习、工作与研究常态。

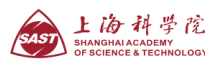


# Q & A





# 打造一个“靠谱”的AI助手



上海计算机软件开发中心

❑ **Cherry Studio**: 集多模型对话、知识库管理、AI绘画等功能于一体的全能AI助手平台

❑ **火山引擎**: 字节跳动的云服务, 提供DeepSeek R1 API

❑ **Tavily**: 为LLM优化的搜索引擎

