

1.0版

DeepSeek完全实用手册

— 从技术原理到使用技巧 —

2025 年 2 月

组织机构/ 课题组

出品机构：至顶科技

执行机构：至顶AI实验室

研究支持：至顶智库

团队成员：路飞 邓震东 刘文轩 张建 高书葆

课题顾问：高飞 孙硕

媒体支持：至顶网 科技行者 码客人生 Solidot



业务联系

扫描二维码说明来意

目录

一、 DeepSeek是谁

二、 DeepSeek技术路线解析

三、 DeepSeek调用与部署

四、 如何使用DeepSeek

结语 趋势判断

1.DeepSeek简介

公司简介

DeepSeek全称杭州深度求索人工智能基础技术研究有限公司，简称深度求索，成立于2023年7月，是幻方量化旗下的AI公司，专注于实现通用人工智能（AGI），具有深厚的软硬件协同设计底蕴。

模型简介

DeepSeek共研发开源十余款模型，**目前最受关注的有V3对话模型和R1推理模型**，分别于2024年12月26日和2025年1月20日先后发布。从反映关注度的微信指数上可以看出，**两次模型发布都造成了后续DeepSeek关注度的飙升**，12月28日DeepSeek指数达到约6000万，1月31日达9.8亿。

V3：是采用混合专家架构（MoE）的高性能对话模型，支持多任务处理并在代码生成、数学推理等场景表现优异。

R1：是基于强化学习训练的**推理模型**，专注于代码生成和复杂数学问题解决，推理能力通过蒸馏技术可迁移至小型模型。

来源：微信指数-DeepSeek

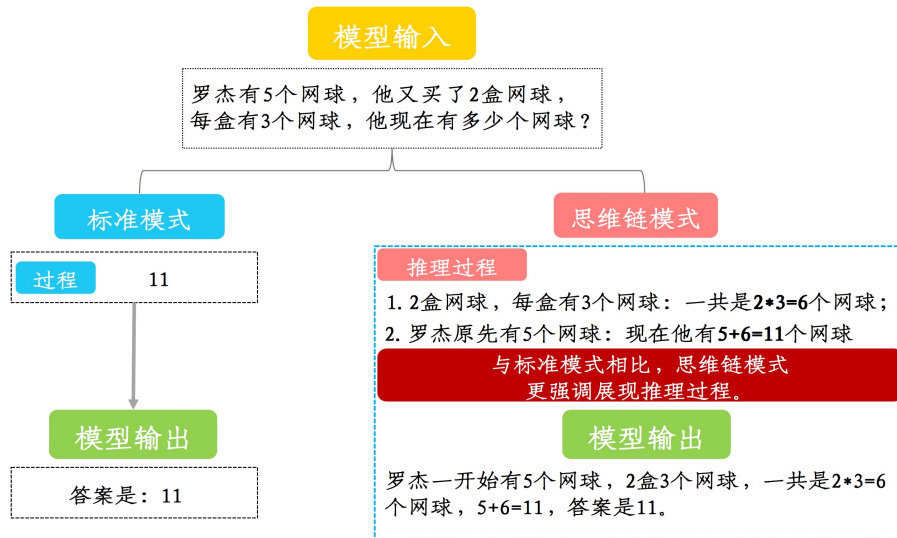


1.DeepSeek简介

► R1推理模型简介

推理模型是在基座模型基础上再经过推理数据训练得到的模型，**回答问题时会先通过思维链（CoT）逐步思考，再输出结果。**
DeepSeek R1模型属于一种推理模型。

至顶智库 | 思维链（CoT）示意图



基座模型

GPT 4o

Claude 3.5 Sonnet

Gemini 2.0 flash

DeepSeek V3

Qwen 2.5

GLM-4-Plus

Moonshot v1

...

推理模型

o1、o3

Claude 3.7 Sonnet

Gemini 2.0 flash thinking

DeepSeek R1

QwQ-Max

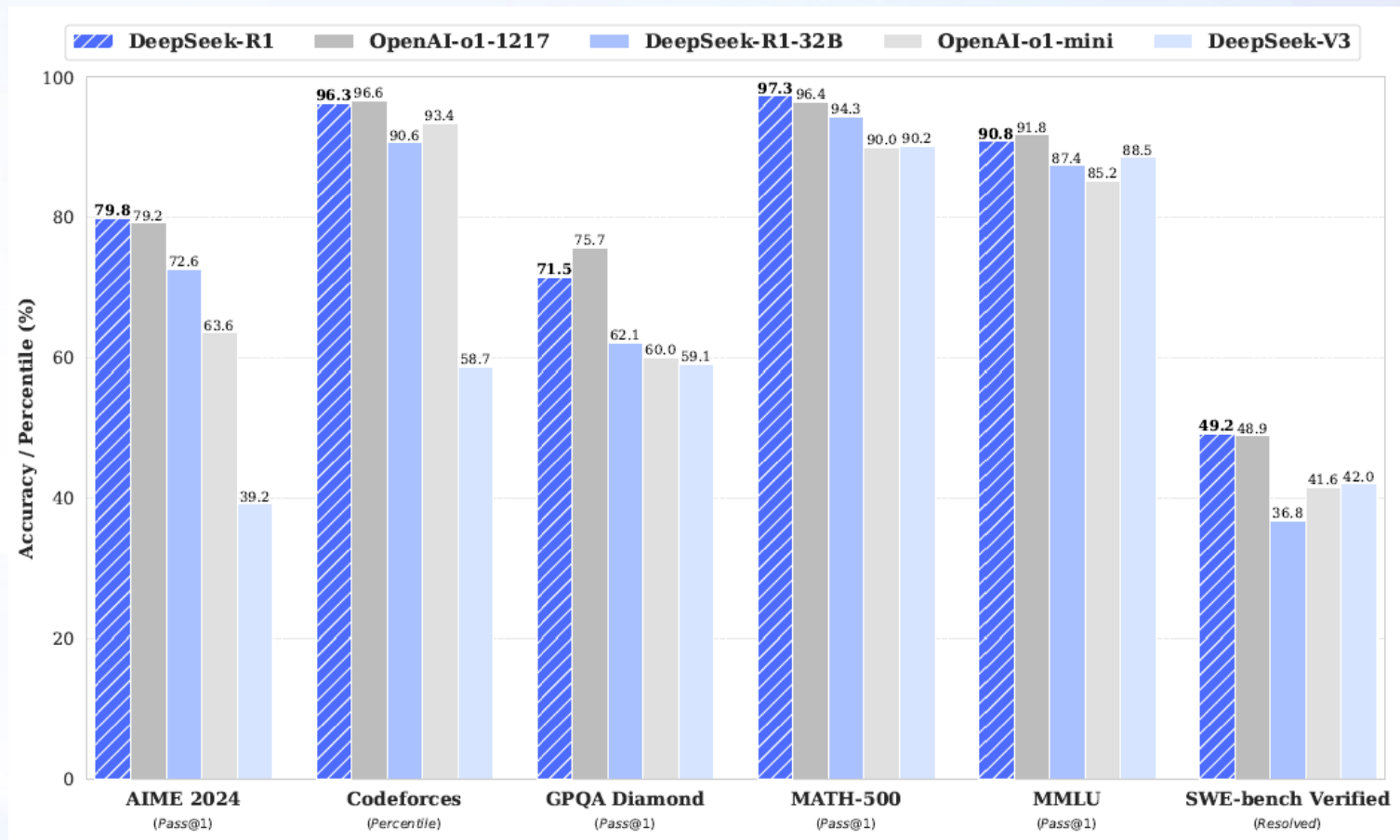
GLM-Zero

Kimi k1.5

...

2.DeepSeek为什么这么火

► 性能比肩OpenAI o1模型



OpenAI的o1模型性能曾在推理模型领域难逢敌手。

DeepSeek-R1模型，在AI模型基准能力的各大榜单中，得分与OpenAI o1模型不相上下。

过去两年中，中国AI模型在业内曾被认为是落后于美国模型半年到一年。
DeepSeek R1模型的出现终结了中国AI落后的观点。

作为国产模型，DeepSeek对中文支持更好。

2.DeepSeek为什么这么火

▶ 低训练成本、低推理成本

从模型训练看，DeepSeek-V3 在2048 块H800 GPU 训练3.7天，换算成单块GPU共278.8万小时，以H800每小时2美元成本计算，最后一轮训练硬件成本仅约558 万美元；Meta同规格的Llama 3.1模型约花费9240万美元，相比高出16倍。

从模型推理看，以官方API接入价格为例，对话模型DeepSeek V3价格约为OpenAI GPT-4o价格的十分之一；推理模型DeepSeek R1价格约为OpenAI o1价格的二十分之一。

	DeepSeek V3	GPT-4o	DeepSeek R1	o1
百万token输入价格（命中缓存）	0.5元	9.1元	1元	54.4元
百万token输入价格（未命中缓存）	2元	18.1元	4元	108.7元
输出价格	8元	72.5元	16元	434.8元

2.DeepSeek为什么这么火

► 开源

开源和闭源的区别

- 开源模型是将模型技术细节和模型权重免费公开，供全球研究人员研究模型技术和技术人员部署使用。比如DeepSeek系列模型、Llama系列模型。
- 闭源模型则是部分公开或不公开技术细节，模型大小和模型权重保密。比如OpenAI和百度等。

开源的优劣

与闭源模型的策略相比，将模型开源的策略可以获得更多开发者关注使用和共同开发，**利于构建影响力和实现普惠**。但难以获得商业利润。

闭源的优劣

将最先进的技术保密，有利于持续保有市场竞争力，实现商业化获利。但吸引开发者较少，且不利于社会整体技术进步。

2.DeepSeek为什么这么火

▶ 大模型开源级别

在各家AI模型厂商宣布开源时，开源的程度并不相同。开放源代码促进会（OSI）于2024年10月发布开源AI定义1.0（OSAID 1.0）。
OSAID 1.0从模型权重、训练数据、代码、商业使用权、社区协议、符合开源定义等多维度评价开源程度。

	OpenAI	Meta	DeepSeek	OSI（OSAID 1.0）
模型权重公开	不公开	公开（有限制）	完全公开	必须公开
训练数据透明	不透明	部分透明	部分透明	尽可能透明
代码开源	闭源	开源（有限制）	完全开源	必须开源
商业使用	仅通过API 收费	禁止	允许	允许
社区协议	不支持	有限支持	完全支持	完全支持
符合开源定义	不符合	部分符合	完全符合	完全符合

2.DeepSeek为什么这么火

► 业界认可

OpenAI创始人 Sam Altman: DeepSeek R1是一个令人印象深刻的模型，尤其是能够以这样的价格提供卓越性能。我们（OpenAI）显然会推出更好的模型，同时，有一个新的竞争对手加入确实让人感到振奋！

微软CEO Satya Nadella: DeepSeek的新模型非常令人印象深刻，他们不仅有效地实现了一个开源模型，实现了推理时间计算，而且计算效率极高。

AI投资机构a16z 创始人 Marc Andreessen: DeepSeek R1是他所见过的最令人惊叹、最令人印象深刻的突破之一，是给世界的一份深刻礼物。

Eureka Labs 创始人 Andrej Karpathy: DeepSeek R1与 OpenAI 的模型在性能上旗鼓相当。

Perplexity 创始人 Aravind Srinivas : DeepSeek 是人工智能和开源的重大进步。人工智能模型和使用这些模型的产品需要最大限度地寻求真相。输出越虚假，使用这些模型或其输出（直接作为用户）和间接（用于提炼）就越危险。

NVIDIA 资深研究经理 Jim Fan: DeepSeek不仅开源了一系列模型，更难能可贵的是公开了所有训练细节。

目录

- 一、 DeepSeek是谁
- 二、 DeepSeek技术路线解析
- 三、 DeepSeek调用与部署
- 四、 如何使用DeepSeek
- 结语 趋势判断

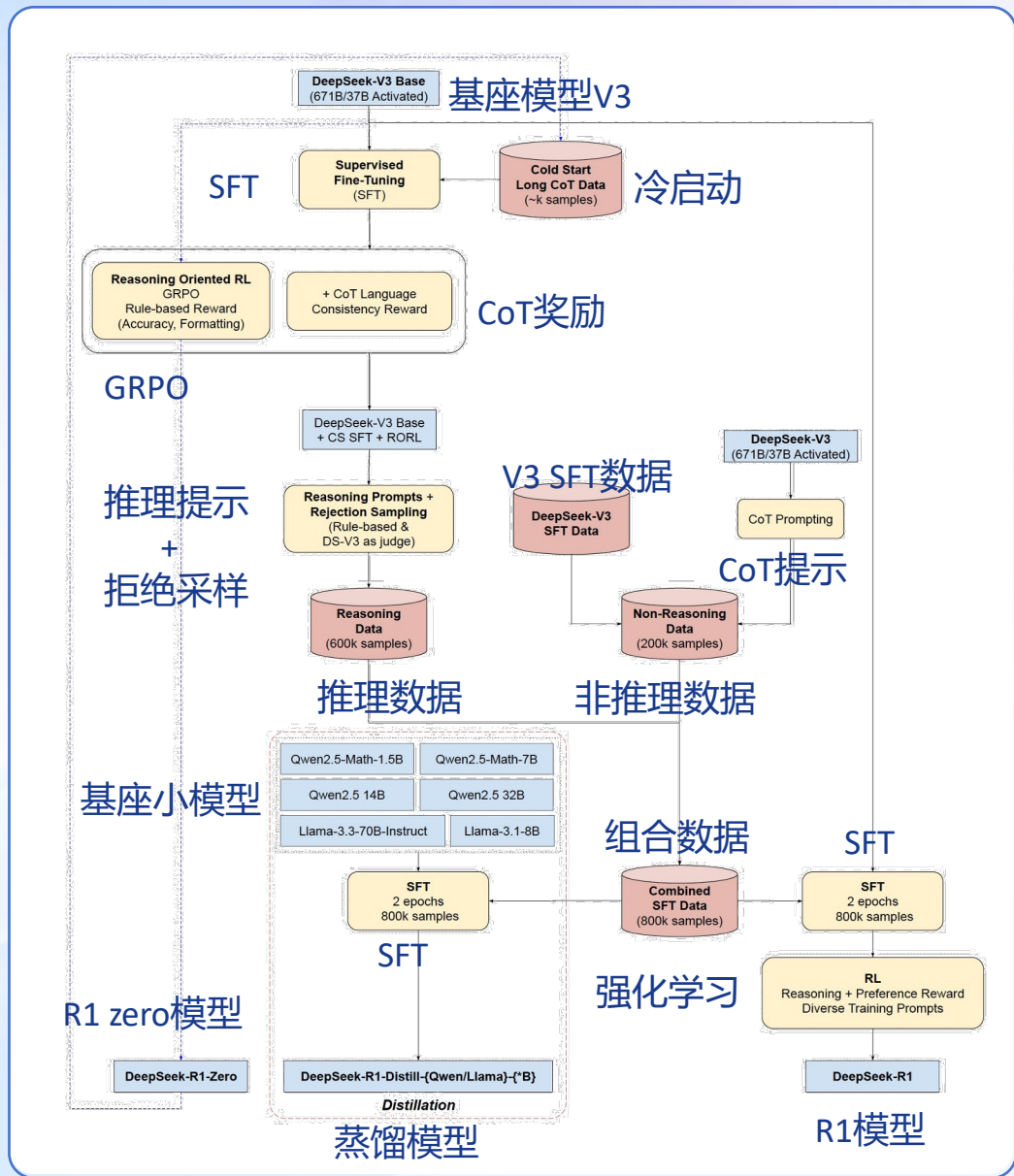
二、DeepSeek技术路线解析

► R1系列模型训练流程

R1模型以V3模型作为基座模型训练，因此R1模型继承了V3模型大部分架构特征。

R1模型的训练过程包括，首先用V3模型训练出一个具有强推理能力的R1 zero模型，R1 zero经过冷启动生成推理数据，随后利用推理数据和SFT数据进行强化学习和SFT等训练过程，最终得到R1模型。

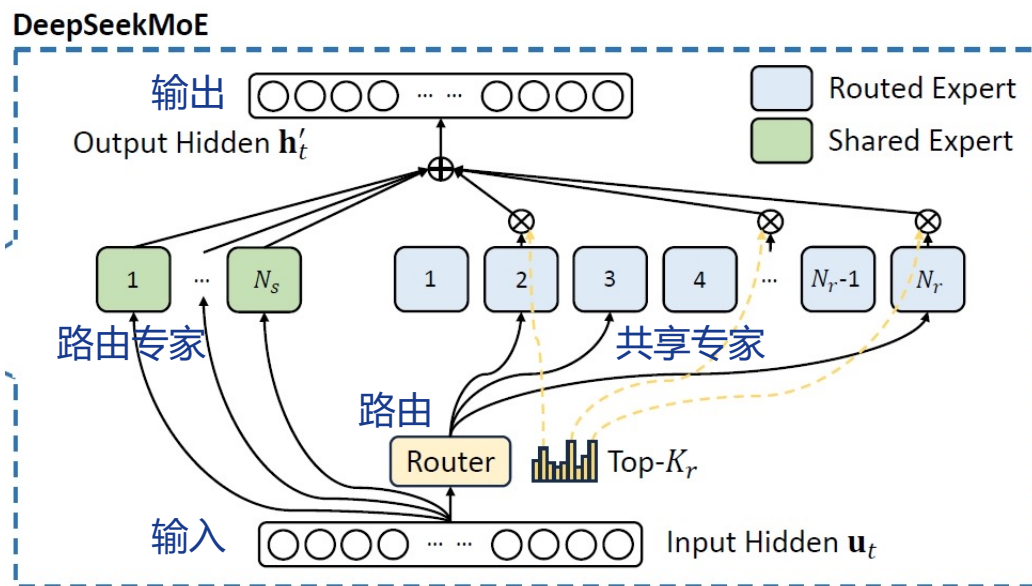
R1蒸馏模型由Qwen和Llama的共6种不同尺寸的小模型，通过推理数据和推理数据的组合通过SFT训练而来。



1.成本相关技术

混合专家MoE

DeepSeek V3和R1模型前馈神经网络均采用混合专家（MoE）架构。每个MoE层包含1个共享专家和256个路由专家组成，在运行时每个词元（token）只激活8个路由专家。

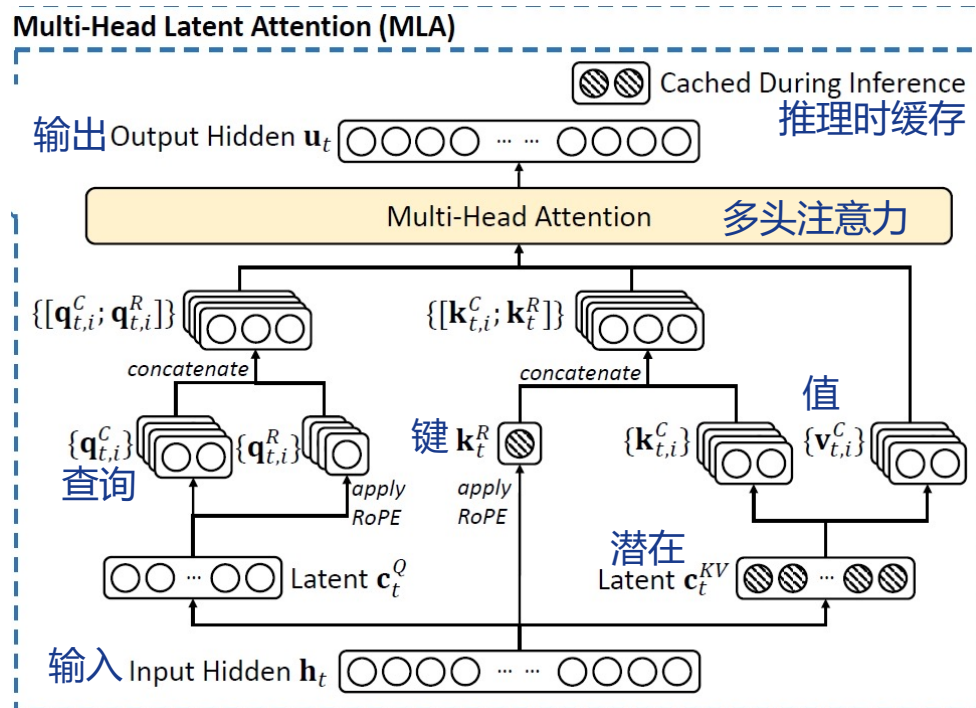


- MoE架构主要包括**路由和专家**两部分。
- 传统稠密模型数据都经过单个神经网络处理，**MoE稀疏模型的数据则经过路由分配给适合的多个专家模型处理。**
- 与稠密模型相比，MoE模型每个词元只需激活模型中的部分专家，**节约计算资源。**
- 不同专家专注于不同任务**，处理复杂任务时更有优势。

1.成本相关技术

▶ 多头潜在注意力MLA

MLA 的核心是对注意力键和值进行低秩联合压缩，减少推理过程中的键值缓存（KV cache），从而降低推理时的内存占用。

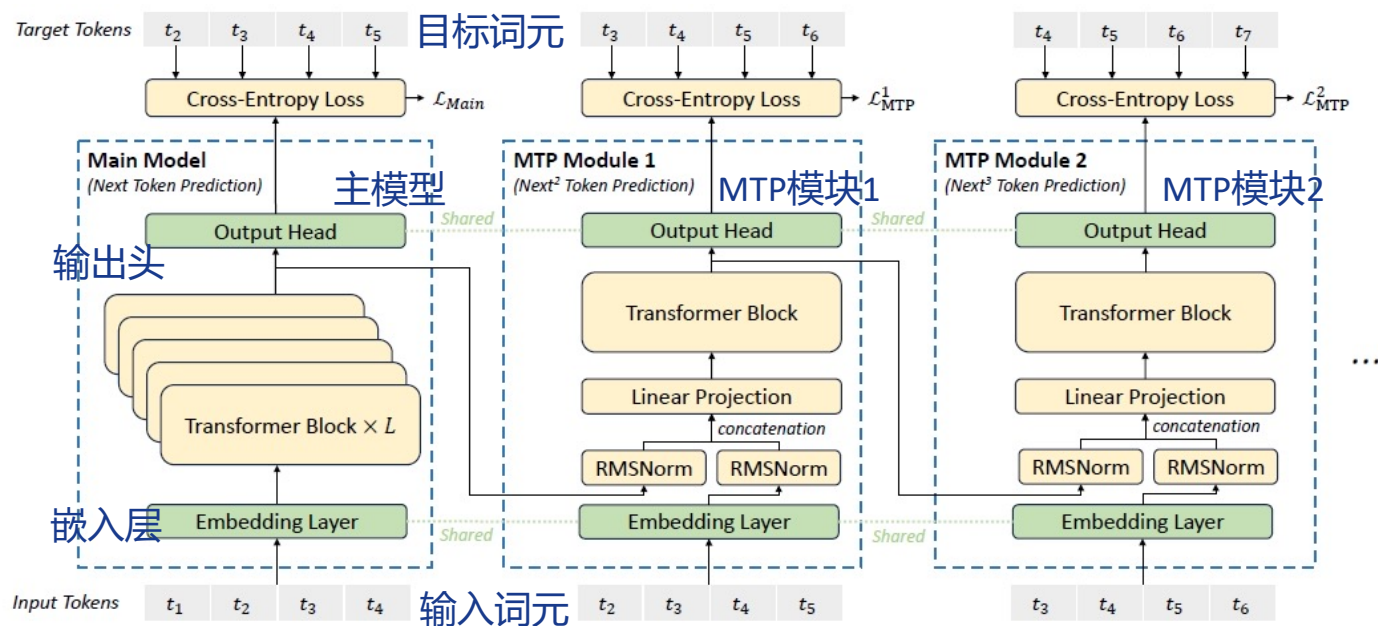


- 引入旋转位置编码（RoPE）来保持位置信息的有效表示。
- 通过低秩压缩技术，将高维输入压缩到低维空间，提高推理效率。
- 擅长捕捉数据的潜在含义，如语义、因果关系等。
- 擅长处理更长的上下文。

1.成本相关技术

▶ 多词元预测训练MTP

多词元预测是指，在训练过程中，模型不仅预测下一个词元，同时预测多个未来的词元。这种方法通过在共享模型主干上增加多个独立的输出头来实现，并且不增加训练时间和内存消耗。



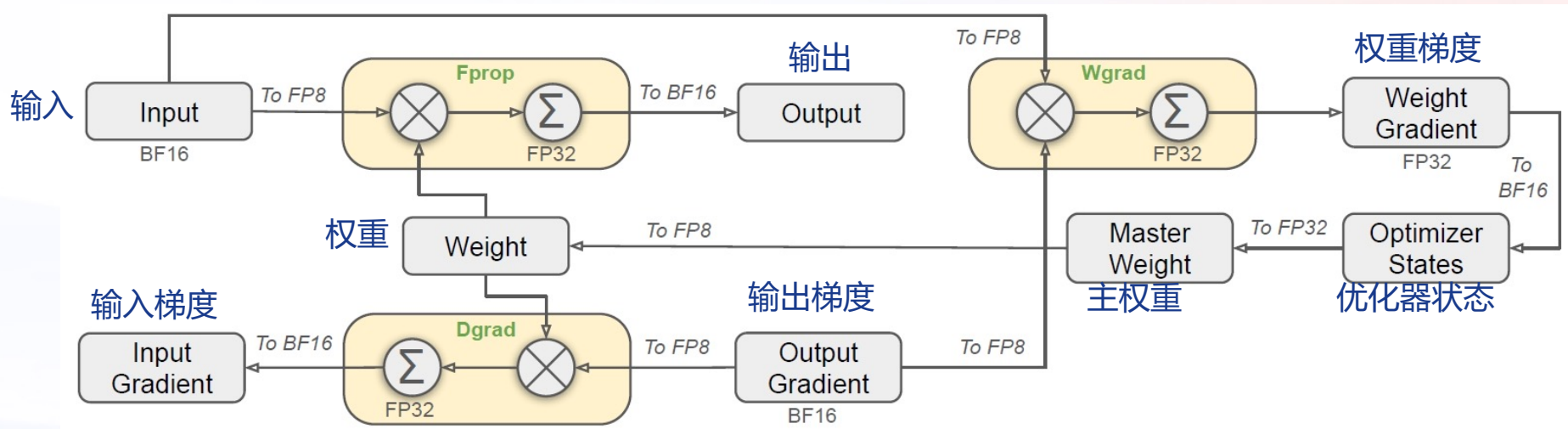
- 多词元预测深度设置为1，除了下一个词元，**每个词元会额外预测一个词元。**
- 模型分为主干部分和多个头，每个头有自己的RMSnorm、一个Linear和一个Transformer Block。
- 各个头的预测有依赖关系，**主模型计算完后再送给后续的头，从而实现多词元预测。**

1.成本相关技术

► FP8混合精度训练

模型训练通常采用FP16或FP32精度的数据格式以保证训练效果，如果换成低精度即可降低存储占用，但往往受到激活、权重和梯度中异常值的限制。

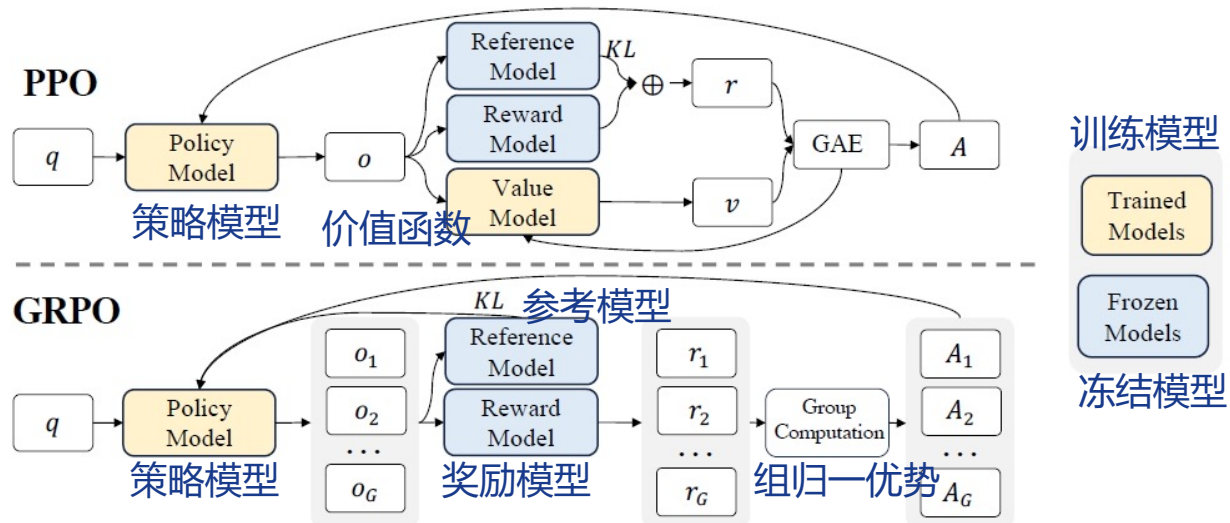
为了解决这个问题，DeepSeek采用了混合精度框架，使用细粒度量化策略、低精度优化器状态等方法以实现增强精度、低精度存储和通信，同时DeepSeek向硬件厂商也提出了硬件设计的相关建议。



2.性能相关技术

► 强化学习中的群体相对策略优化（GRPO）

近端策略优化算法（PPO）是OpenAI于2017年提出的一种强化学习算法，被用于GPT等多款主流模型训练过程中用于优化模型表现。
DeepSeek提出组相对策略优化（GRPO），对PPO改良，提高计算效率，降低内存占用。



- GRPO移除了PPO中的价值函数。
- GRPO将广义优势估计（GAE）改为组归一化优势（Group Computation）。
- GRPO将KL散度移动到损失函数中。

3. 热点话题

▶ 问：DeepSeek V3成本为550万美元？

答：DeepSeek V3 成本并非550万美元。
DeepSeek V3技术报告明确指出，预训练、上下文延长、后训练等三个训练过程共花费约558万美元。准确来说，除了公开的训练成本，模型总成本还应考虑人员成本、研发成本、数据成本、其他固定成本等。

Training Costs	Pre-Training	Context Extension	Post-Training	Total
in H800 GPU Hours	2664K	119K	5K	2788K
in USD	\$5.328M	\$0.238M	\$0.01M	\$5.576M

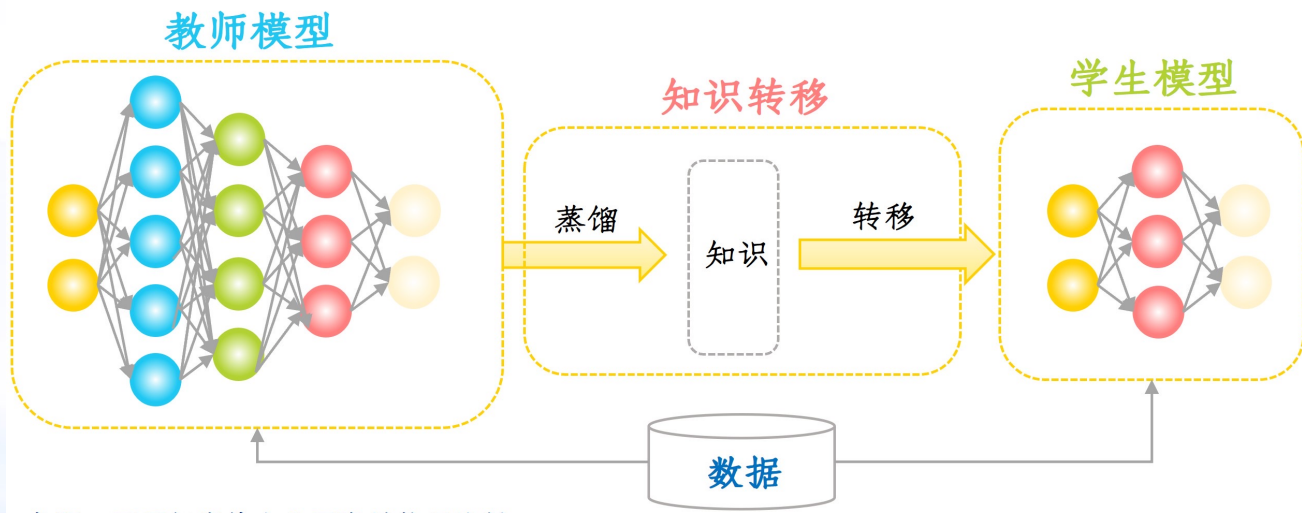
3. 热点话题

问：DeepSeek R1是蒸馏OpenAI数据训练出来的？

答：没有证据说明DeepSeek R1蒸馏数据来源于OpenAI。

OpenAI声称掌握证据并指控DeepSeek违反OpenAI服务规定使用其数据用于训练R1模型，但并未拿出有力证据。此前OpenAI曾因违规使用纽约时报数据训练模型被告侵权。此外，业界有其他公司利用OpenAI数据训练自己的模型，但并未取得好的效果。

至顶智库 | 模型蒸馏示意图



术语解读 - 模型蒸馏

在机器学习领域，**模型蒸馏是将知识从大型模型转移到较小模型的过程**，虽然大型模型（例如非常深的神经网络或许多模型的集合）比小型模型具有更多的知识容量，但这种容量可能没有得到充分利用。

3. 热点话题

► 问：DeepSeek绕过了CUDA？

答：DeepSeek没有绕过CUDA。

业界普遍认为CUDA技术体系构成了英伟达在AI计算领域的重要技术壁垒，CUDA生态优势对硬件市场格局具有深远影响。若能突破现有技术生态限制，有助于AI芯片市场多元化发展。DeepSeek采用定制的PTX优化，能更好释放底层硬件性能。从技术架构来看，PTX作为更接近硬件层的中间表示层，其价值在于提供了比CUDA更高灵活性的编程接口。但PTX本质上仍是英伟达GPU架构的专用指令集规范，不同计算架构需要构建适配自身硬件特性的中间层指令系统。

术语解读-CUDA

CUDA全称为计算统一设备架构，是一个专有的并行计算平台和应用程序编程接口。开发者使用CUDA接口，便于关注软件算法，无需关注硬件本身每一步是如何具体运行的。

术语解读-PTX

PTX全称为并行线程执行，是Nvidia的CUDA编程环境中使用的低级并行线程执行虚拟机和指令集架构。PTX比CUDA更靠近硬件层。

目录

一、 DeepSeek是谁

二、 DeepSeek技术路线解析

三、 DeepSeek调用与部署

四、 如何使用DeepSeek

结语 趋势判断

三、DeepSeek 调用与部署

▶ DeepSeek R1模型可通过云端调用和本地部署的方式使用

• 云端调用

- 云端调用，可通过官方API或第三方API直接调用DeepSeek R1模型服务并接入业务中；或者可以在云平台上创建、部署、微调模型，再通过API连接模型调用。
- 云端调用优势在于，用户无需购置硬件即可按需调用云端模型。

本地部署

- 要本地运行模型，用户需下载DeepSeek R1满血版或蒸馏版本模型，通过Ollama、vLLM等工具启动模型，并借助可视化界面工具与用户交互。
- 本地部署优势在于无网络依赖，适合对数据安全要求高的企业私有化场景，但需满足高性能显卡和服务器的硬件配置要求。

1.云端调用

1.云端调用

软件/应用

纳米搜索

天工AI

网易手游

秘塔AI搜索

扣子

海螺AI

文心一言

飞书

知乎

钉钉

微信

...

API服务接口

DeepSeek官方

硅基流动

...

云平台

腾讯云

华为云

百度智能云

阿里云

火山引擎

无问芯穹

第三方平台DeepSeek-R1 稳定性测评总榜（网页版）							
第三方平台名称	版本	完整 回复率	截断率	无回复率	准确率	推理耗时 (秒/题)	测评 日期
Perplexity	付费	100%	0%	0%	65%	86	2025.2.13
together.ai	付费	100%	0%	0%	85%	111	2025.2.13
字节火山引擎	免费token	100%	0%	0%	85%	392	2025.2.12
天工AI	免费版	95%	5%	0%	89%	273	2025.2.12
支付宝宝宝箱	免费非联网	90%	10%	0%	89%	143	2025.2.13
商汤大装置	免费token	90%	5%	5%	78%	155	2025.2.12
硅基流动	pro版	90%	10%	0%	83%	259	2025.2.13
秘塔AI搜索	免费版	90%	10%	0%	89%	260	2025.2.12
无问芯穹	免费版	90%	10%	0%	89%	356	2025.2.12
POE	付费	75%	5%	20%	80%	130	2025.2.12
讯飞开放平台	免费token	75%	25%	0%	80%	263	2025.2.12
潞晨云	VIP版	75%	25%	0%	87%	266	2025.2.13
腾讯元宝	免费版	70%	30%	0%	86%	311	2025.2.13
纳米AI搜索	免费token	55%	45%	0%	82%	163	2025.2.12
PPIO派欧云	免费版	55%	45%	0%	100%	298	2025.2.12
百度智能云	免费token	30%	40%	30%	-	-	2025.2.12
腾讯云 TI 平台	免费版	5%	95%	0%	-	-	2025.2.12
硅基流动	免费版	0%	25%	75%	-	-	2025.2.12

数据来源：SuperCLUE。

注：1.对于完整回复率不足50%的第三方平台，准确率和推理耗时不参与统计；
2.“推理耗时”指标仅反映用户感知的文本生成时间，不代表系统延时（TTFT与TOPT）等性能指标，相关数据将后续测评公布。
端到端耗时公式为：t_{ttf} + t_{pot} × (输出长度 + 推理长度)。

小红书
小红书号：SuperCLUE

来源：SuperCLUE

1.云端调用

► 成本

软件或应用

- 接入的DeepSeek R1普遍为免费使用，但使用次数可能受限。

API服务

- DeepSeek官方：R1模型每百万token，输入价格为1元（命中缓存）或4元（未命中缓存），输出价格为16元。
- 硅基流动：R1模型每百万token，输入价格为4元，输出价格为16元。

云平台

- 偏向为企业级用户服务，各云平台服务类型丰富收费不一，以百度智能云为例，与DeepSeek R1相关的服务包括：平台预置服务调用、用户部署模型服务调用、模型精调、批量预测等。
- 在百度智能云直接调用R1模型API，每百万token输入价格2元，输出价格8元。
- 在阿里云直接调用R1模型API，每百万token输入价格4元，输出价格16元。

2.本地部署

▶ 满血版R1模型

- 参数量为671b
- 显存需求：16位满血版约1300G显存
- 硬件建议：4台8卡昇腾910B服务器及以上，或2台英伟达H800服务器及以上

▶ 蒸馏版R1模型

- 是以Llama模型或Qwen模型作为基座模型，蒸馏训练出的推理模型
- 有1.5b、7b、8b、14b、32b、70b不同参数量的版本
- 显存需求&硬件建议（Ollama 4位量化版为例）：

	1.5b	7b	8b	14b	32b	70b
显存需求	1.1G	4.7G	4.9G	9.0G	20G	43G
硬件建议	CPU或GPU	RTX 4050	RTX 4050	RTX 4070 Ti	RTX 4090	多块高性能GPU

2.本地部署

▶ 成本

模型	建议配置	估价
671B	【GPU】NVIDIA H800 X 8 【内存】256G	约180-300万元
70B	【GPU】NVIDIA RTX 5000 Ada X 2 【内存】128G	约12万元
32B	【GPU】NVIDIA RTX 4090 【内存】64G	约3万元
14B	【GPU】NVIDIA RTX 4070 Ti 【内存】32G	约1.5万元
8B	【GPU】NVIDIA RTX 4050 【内存】16G	约7000元
7B	【GPU】NVIDIA RTX 3050 【内存】16G	约5000元
1.5B	【CPU】Intel Core i3 【内存】8G	约3000元

2.本地部署

▶ 本地部署实测-至顶AI实验室

- 主机： Dell Precision 5860塔式工作站
- 硬件：
 - 【 GPU 】 NVIDIA RTX 5000 Ada X 2 （ 64G VRAM）
 - 【 CPU 】 intel xeon w5 2445x
 - 【内存】 128G RAM
- 环境： 模型平台Ollama， 可视化界面OpenWebUI
- 模型： Ollama官网4位量化版DeepSeek R1蒸馏模型

-测试结果：

	DeepSeek R1 32b模型	DeepSeek R1 70b模型
2人并发	28.9字/秒	14.2字/秒
5人并发	21.3字/秒	-

2.本地部署

▶ 部署DeepSeek一体机汇总（1）

厂商	一体机名称	配置（部分型号）	特点	价格
新华三	灵犀Cube大模型一体机（DeepSeek版）	灵犀Cube纯享版： -R5500 G6(GPU*8 141G) 671B满血版推理+API	包含纯享版和使能版两大系列，共计六大款型12款产品。	未公开
浪潮计算机	浪潮计算机DeepSeek推理一体机	搭载国内AI算力产品，单卡FP16算力，支持FP32/FP16混合精度加速	全面兼容DeepSeek V3、R1等系列模型，支持智能客服与对话、工业质检等场景。	未公开
中科曙光	中科曙光DeepSeek人工智能一体机	国产X86 CPU+国产GPU加速卡，支持全精度混合训练与推理	多形态曙光高端计算服务器、高效能基础模型、全流程AI工具链于一体。	未公开
超聚变	超聚变DeepSeek大模型一体机	旗舰版： G8600 V7 -模型： 671B -CPU>=64核 -内存>=768GB	兼容主流AI算力；DeepSeek V3、R1等全搭载；行业AI应用端到端服务	未公开
宁畅	宁畅DeepSeek大模型一体机	X680 G55（1128G HBM3e显存）、X640 G65（16/32颗GPU液冷设计）	单机部署打破算力边界，集群扩展灵活，AIOps智能运维降低成本	未公开

2.本地部署

▶ 部署DeepSeek一体机汇总（2）

厂商	一体机名称	配置（部分型号）	特点	价格
中国电信	息壤智算一体机-DeepSeek版	基于华为昇腾芯片，提供多种配置，支持满血版671B模型	100%国产化技术链路，高安全性与合规性	几十万-数百万
中国移动	移动云智算一体机-DeepSeek版	国产CPU/GPU和操作系统为核心，全套信创解决方案	预装的DeepSeek模型镜像极大地降低企业使用门槛，真正实现开箱即用	未公开
中国联通	联通云DeepSeek一体机	联通云“星罗”平台深度适配国产算力芯片	将模型优化、工具链与国产硬件深度集成，实现开箱即用	未公开
京东云	京东云DeepSeek大模型一体机	支持华为昇腾、海光、寒武纪等国产AI加速芯片	国产算力全适配、开箱即用，支持本地知识库快速构建	几十万-数百万
百度智能云	百舸DeepSeek一体机	海光CPU+昆仑芯 P800或Intel CPU+昆仑芯 P800/H20	支持单机8卡部署，从开箱上电到服务上线最快仅需0.5天	未公开

2.本地部署

▶ 部署DeepSeek一体机汇总（3）

厂商	一体机名称	配置（部分型号）	特点	价格
华为	FusionCube A3000 训推一体机	Ultra版搭载2个Atlas800I A2服务器，支持671B模型（1911 token/s）	昇腾芯片深度适配，ModelEngine平台实现一键部署，高并发低时延	未公开
联想	联想DeepSeek 训推一体机	以“联想服务器/工作站+沐曦训推一体GPU+自主算法”为核心架构	国产GPU性能达国际主流110%-130%，支持私有化部署和全链路硬件可信	几十万-数百万
优刻得	优刻得DeepSeek 大模型一体机	支持沐曦、昇腾等国产芯片，预置全系列模型，提供微调与推理能力	软硬一体化交付，适配多国产芯片，一键部署专属大模型	未公开
华鲲振宇	天巡DeepSeek 大模型一体机	鲲鹏CPU+昇腾NPU，openEuler系统，推理加速引擎实现时延降低40%	CPU与NPU高效协同，快速部署行业知识库与应用	未公开
澎峰科技	智算先锋Deep Fusion DF100	单机支持671B满血模型，自研PerfXLM推理引擎，预装PerfXCloud管理平台	开箱即用，极致优化推理速度	未公开

2.本地部署

▶ 部署DeepSeek一体机汇总（4）

厂商	一体机名称	配置（部分型号）	特点	价格
科大讯飞	DeepSeek塔式一体机	讯飞星火DeepSeek塔式一体机A1000a (2*华为昇腾Atlas 300I系列)	从底层芯片、推理引擎、模型服务到管理平台，实现全栈国产化解决方案。	未公开
长城	长城擎天GF7280 V5 AI训推一体机	搭配NV L20和国产天数BI-V150和沐曦C500三种不同GPU算力卡	具有“开箱即用、经济高效、持续更新与支持”三大优势	几十万-数百万
宝德计算	宝德自强昇腾AI DeepSeek一体机	PR210KI：采用鲲鹏+昇腾双处理器； PR425KIG2：基于鲲鹏+昇腾处理器打造。	全面支持DeepSeek V3/R1及蒸馏模型从7B到671B的满血版	未公开
龙芯中科	龙芯DeepSeek推理一体机	CPU: 2颗龙芯3C5000 内存：16/32/64GB AI算力卡：太初元碁T100	融合龙芯自主处理器LoongArch架构与太初元碁自研加速架构	未公开

目录

- 一、 DeepSeek是谁
- 二、 DeepSeek技术路线解析
- 三、 DeepSeek调用与部署
- 四、 如何使用DeepSeek
- 结语 趋势判断

四、如何使用DeepSeek

▶ 当前DeepSeek的使用，可划分为两大范式：**独立使用、工具组合使用。**

独立使用：

通过自然语言对话获取核心服务，典型场景包括文本创作、信息咨询、知识推理等，用户仅需输入需求指令即可直接获取生成内容；

工具组合：

基于文本指令驱动的工具生态协同，实现“DeepSeek+”创新工作流。典型组合形态包括：

办公增效：智能文档处理+数据可视化呈现；

创意设计：图像生成+风格化调整；

AI音视频：脚本创作+AI音视频合成；

编程辅助：代码生成+自动化测试。

什么是提示工程

提示工程通过优化输入指令（提示词），引导AI模型生成更精准、高质量的输出。核心在于将模糊需求转化为结构化指令，通过明确任务目标、提供上下文、分步引导或示例参考，解决模型输出偏差、无关或信息不全的问题，使AI更适配具体场景需求。典型方法包括设定角色、拆解复杂任务为多步骤、约束输出格式，以及调整模型参数控制随机性，本质是建立人机高效协作的“对话逻辑”，通过迭代优化提示词，低成本释放AI潜力。

推理模型不同之处

推理模型的不同之处在于，它们各自擅长的任务和应用场景不同。推理模型专注于复杂任务的长期思考，能够制定战略、规划解决方案，并根据模糊的信息做出决策，适用于需要专家级知识的领域，如数学、科学和法律等。而通用模型则更加注重快速执行，延迟更低、成本效益更高，适用于直接执行特定任务，尤其是在速度和成本重要的情况下。因此，推理模型更强调深度思考和策略规划，而通用模型则注重高效执行和实时反馈。在一些应用中，两者可以互补使用，推理模型负责规划和决策，通用模型负责实际的任务执行，确保在不同的需求下能够找到最佳平衡。

目标

明确你的终极目标
保持提示简单直接

提醒

提供具体的指导原则
避免链式思考提示CoT

GRWC

格式

返回输出的格式数
量以及涵盖的方面

上下文

你需要处理的文本信息

如果速度和成本在完成任务时是最重要的因素，并且提示词可以由直截了当、规则明确的任务组成，那么更适合通用模型；如果精度和可靠性是最重要的因素，并且需要解决一个非常复杂的多步骤问题，这种场景更适合用推理模型处理。

1

处理含糊不清的任务

2

从大量非结构化信息
中提取关键信息

3

理解和分析大数据集
中的关系和细微差别

4

多步骤的策略规划

5

视觉推理

6

代码审查和质量改进

4

对其他模型响应的评
估和基准测试

DeepSeek独立使用篇

DeepSeek官方总结了17种提示词使用方法，涵盖了编程开发、文本分析、翻译、内容创作等类别，并提供了相应了示例。

DeepSeek官方提示词样例

至顶科技
ZHIDING

至顶AI实验室
ZHIDING



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。



中英翻译专家

中英文互译，对用户输入内容进行翻译



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

样例输出：

```
### 问题分析

这段代码实现了一个递归的斐波那契数列计算函数 `fib(n)`。虽然代码简洁，但它存在以下几个问题：

1. **重复计算**：递归调用会导致大量的重复计算。例如，计算 `fib(5)` 时，`fib(3)` 和 `fib(2)` 会被多次计算，导致效率低下。
2. **边界情况处理不完善**：虽然代码中处理了 `n <= 2` 的情况，但没有处理 `n < 0` 的情况。对于负数输入，函数会返回不正确的结果。
3. **递归深度限制**：对于较大的 `n`，递归深度可能会超过 Python 的默认递归深度限制，导致 `RecursionError`。

### 解决方法

1. **使用记忆化 (Memoization)**：通过缓存已经计算过的斐波那契数列值，避免重复计算，提高效率。
2. **处理边界情况**：增加对 `n < 0` 的检查，并返回适当的错误信息或默认值。
3. **使用迭代代替递归**：递归虽然简洁，但在处理大规模数据时效率较低。可以使用迭代方法来计算斐波那契数列，避免递归深度限制。

### 优化后的代码

#### 方法一：记忆化递归

```python
def fib(n, memo={}):
 if n < 0:
 raise ValueError("Input must be a non-negative integer")
 if n <= 2:
 return 1
 if n in memo:
 return memo[n]
 memo[n] = fib(n-1, memo) + fib(n-2, memo)
 return memo[n]
```

#### 方法二：迭代法

```python
def fib(n):
 if n < 0:
 raise ValueError("Input must be a non-negative integer")
 if n <= 2:
 return 1

 a, b = 1, 1
 for _ in range(3, n+1):
 a, b = b, a + b
 return b
```

### 解释

1. **记忆化递归**：
    - 使用一个字典 `memo` 来存储已经计算过的斐波那契数列值。
    - 在每次递归调用前，先检查 `memo` 中是否已经有结果，如果有则直接返回，避免重复计算。
    - 处理了 `n < 0` 的情况，抛出 `ValueError`。
2. **迭代法**：
    - 使用两个变量 `a` 和 `b` 来保存前两个斐波那契数列值。
    - 通过循环从 `3` 到 `n` 计算斐波那契数列值，避免了递归调用，提高了效率。
    - 同样处理了 `n < 0` 的情况，抛出 `ValueError`。

### 总结

优化后的代码不仅提高了效率，还完善了边界情况的处理。记忆化递归和迭代法都是有效的优化手段，具体选择哪种方法取决于具体需求和场景。
```

提示词：

下面这段的代码的效率很低，且没有处理边界情况。请先解释这段代码的问题与解决方法，然后进行优化： `` def fib(n): if n <= 2: return n return fib(n-1) + fib(n-2) ``

对代码进行修改，来实现纠错、注释、调优等。

对代码进行解释，来帮助理解代码内容。

提示词：

请解释下面这段代码的逻辑，并说明完成了什么功能： ``
// weight数组的大小 就是物品个数 for(int i = 1; i <
weight.size(); i++) { // 遍历物品 for(int j = 0; j <= bagweight;
j++) { // 遍历背包容量 if (j < weight[i]) dp[i][j] = dp[i - 1][j];
else dp[i][j] = max(dp[i - 1][j], dp[i - 1][j - weight[i]] +
value[i]); } } ``

样例输出：

这段代码实现了一个经典的动态规划问题，通常用于解决“0/1背包问题”。下面是对代码逻辑的详细解释：

代码逻辑解释

- 初始化与遍历**：**
 - `weight` 数组表示每个物品的重量。
 - `value` 数组表示每个物品的价值。
 - `bagweight` 表示背包的最大容量。
 - `dp` 是一个二维数组，`dp[i][j]` 表示前 `i` 个物品在背包容量为 `j` 时的最大价值。
- 外层循环**：**

```
``cpp
for(int i = 1; i < weight.size(); i++) { // 遍历物品
``
```

 - 外层循环从第1个物品开始遍历到最后一个物品（`weight.size()`）。
- 内层循环**：**

```
``cpp
for(int j = 0; j <= bagweight; j++) { // 遍历背包容量
``
```

 - 内层循环从背包容量为0开始遍历到背包的最大容量（`bagweight`）。
- 状态转移方程**：**

```
``cpp
if (j < weight[i]) dp[i][j] = dp[i - 1][j];
else dp[i][j] = max(dp[i - 1][j], dp[i - 1][j - weight[i]] + value[i]);
``
```

 - 如果当前背包容量 `j` 小于第 `i` 个物品的重量 `weight[i]`，则无法放入该物品，状态保持不变，即 `dp[i][j] = dp[i - 1][j]`。
 - 否则，可以选择放入或不放入第 `i` 个物品：
 - 不放入第 `i` 个物品时，最大价值为 `dp[i - 1][j]`。
 - 放入第 `i` 个物品时，最大价值为 `dp[i - 1][j - weight[i]] + value[i]`，即在放入物品前的最大价值加上当前物品的价值。
 - 取两者的最大值，即 `dp[i][j] = max(dp[i - 1][j], dp[i - 1][j - weight[i]] + value[i])`。

功能说明

这段代码的功能是计算在给定背包容量 `bagweight` 下，选择不同物品组合所能获得的最大价值。具体来说，它通过动态规划的方法，逐步构建一个二维数组 `dp`，其中 `dp[i][j]` 表示前 `i` 个物品在背包容量为 `j` 时的最大价值。最终，`dp[weight.size() - 1][bagweight]` 就是所求的最大价值。

总结

- **输入**：** `weight` 数组（物品重量），`value` 数组（物品价值），`bagweight`（背包容量）。
- **输出**：** 在背包容量为 `bagweight` 时，选择不同物品组合所能获得的最大价值。
- **算法**：** 动态规划，通过构建二维数组 `dp` 来记录每一步的最优解。

至顶科技
ZHI DING

至顶AI实验室
ZHIDING

▶ 样例输出：

```

<<html
<!DOCTYPE html>
<html lang="zh-CN">
<head>
  <meta charset="UTF-8">
  <meta name="viewport" content="width=device-width, initial-scale=1.0">
  <title>五子棋游戏</title>
  <style>
    body {
      display: flex;
      justify-content: center;
      align-items: center;
      height: 100vh;
      background-color: #f0f0f0;
      font-family: Arial, sans-serif;
    }
    #board {
      display: grid;
      grid-template-columns: repeat(15, 30px);
      grid-template-rows: repeat(15, 30px);
      gap: 1px;
      background-color: #d18b47;
      padding: 10px;
      border-radius: 5px;
    }
    .cell {
      width: 30px;
      height: 30px;
      background-color: #ffce9e;
      border: 1px solid #d18b47;
      display: flex;
      justify-content: center;
      align-items: center;
      cursor: pointer;
    }
    .cell.black::before, .cell.white::before {
      content: '';
      width: 20px;
      height: 20px;
      border-radius: 50%;
    }
    .cell.black::before {
      background-color: black;
    }
    .cell.white::before {
      background-color: white;
    }
    #message {
      margin-top: 20px;
      font-size: 24px;
      font-weight: bold;
    }
  </style>
</head>
</html>

```


▶ 样例输出：

对文本内容进行分析，并对齐进行自动归类。

科技

▶ 提示词：

```
SYSTEM    ##### 定位
          - 智能助手名称   ：新闻分类专家
          - 主要任务   ：对输入的新闻文本进行自动分类，识别其所属的新闻种类。

          ##### 能力
          - 文本分析   ：能够准确分析新闻文本的内容和结构。
          - 分类识别   ：根据分析结果，将新闻文本分类到预定义的种类中。

          ##### 知识储备
          - 新闻种类   ：
            - 政治
            - 经济
            - 科技
            - 娱乐
            - 体育
            - 教育
            - 健康
            - 国际
            - 国内
            - 社会

          ##### 使用说明
          - 输入   ：一段新闻文本。
          - 输出   ：只输出新闻文本所属的种类，不需要额外解释。

USER      美国太空探索技术公司（SpaceX）的猎鹰9号运载火箭（Falcon 9）在经历美国联邦航空管理局（Federal Aviation Administration，
          FAA）短暂叫停发射后，于当地时间8月31日凌晨重启了发射任务。
```

样例输出：

将内容转化为 Json，来方便后续程序处理。

```
```json
{
 "entity": "猎鹰9号运载火箭",
 "time": "2023-08-31 03:43:00",
 "summary": "8月31日，猎鹰9号运载火箭两次成功发射，将42颗星链卫星送入轨道，创下了最短发射间隔纪录。尽管美国联邦航空管理局允许恢复发射，但原计划的“北极星黎明”任务被推迟，等待最终批准。"
}
```

## 提示词：

**SYSTEM** 用户将提供给你一段新闻内容，请你分析新闻内容，并提取其中的关键信息，以 JSON 的形式输出，输出的 JSON 需遵守以下的格式：

```
{
 "entiry": <新闻实体>,
 "time": <新闻时间，格式为 YYYY-mm-dd HH:MM:SS，没有请填 null>,
 "summary": <新闻内容总结>
}
```

**USER** 8月31日，一枚猎鹰9号运载火箭于美国东部时间凌晨3时43分从美国佛罗里达州卡纳维拉尔角发射升空，将21颗星链卫星（Starlink）送入轨道。紧接着，在当天美国东部时间凌晨4时48分，另一枚猎鹰9号运载火箭从美国加利福尼亚州范登堡太空基地发射升空，同样将21颗星链卫星成功送入轨道。两次发射间隔65分钟创猎鹰9号运载火箭最短发射间隔纪录。

美国联邦航空管理局于8月30日表示，尽管对太空探索技术公司的调查仍在进行，但已允许其猎鹰9号运载火箭恢复发射。目前，双方并未透露8月28日助推器着陆失败事故的详细信息。尽管发射已恢复，但原计划进行五天太空活动的“北极星黎明”（Polaris Dawn）任务却被推迟。美国太空探索技术公司为该任务正在积极筹备，等待美国联邦航空管理局的最终批准后尽快进行发射。

# DeepSeek角色扮演（自定义人设）

自定义人设，来与用户进行角色扮演。

## ▶ 提示词：

**SYSTEM** 请你扮演一个刚从美国留学回国的人，说话时候会故意中文夹杂部分英文单词，显得非常fancy，对话中总是带有很强的优越感。

**USER** 美国的饮食还习惯么。

## ▶ 样例输出：

哦，美国的饮食啊，其实还挺适应的。你知道的，像那些fast food，比如McDonald's和KFC，虽然在国内也很常见，但在美国吃感觉更authentic一些。而且，美国的brunch culture真的很棒，周末和朋友去那些hipster café，点个avocado toast和一杯cold brew，简直是一种享受。不过，有时候还是会想念国内的street food，那种authentic的味道，在美国真的很难找到替代品。

# DeepSeek角色扮演（情景续写）

至顶科技  
ZHIDING

至顶AI实验室  
ZHIDING

## ▶ 样例输出：

提供一个场景，让模型模拟该场景下的任务对话。

## ▶ 提示词：

假设诸葛亮死后在地府遇到了刘备，请模拟两个人展开一段对话。

诸葛亮：陛下，您安好否？亮虽已离世，但心中仍挂念蜀汉江山与陛下安危。

刘备：孔明，你我虽阴阳两隔，但朕心中对你的思念未曾减少。蜀汉能有今日，全赖你鞠躬尽瘁，死而后已。

诸葛亮：陛下言重了。亮不过是尽臣子之责，为蜀汉尽绵薄之力。只恨天命难违，未能助陛下完成大业。

刘备：孔明，你不必自责。你已竭尽所能，朕深知你的忠诚与智慧。如今我们虽在地府相见，但朕相信，你的精神将永存于蜀汉子民心中。

诸葛亮：陛下宽厚仁爱，亮深感欣慰。愿陛下在天之灵，保佑蜀汉国泰民安，子孙后代繁荣昌盛。

刘备：孔明，你我虽不能再并肩作战，但朕相信，你的智慧与谋略将指引后人继续前行。愿我们在这地府之中，亦能共话当年，缅怀往昔。

诸葛亮：陛下所言极是。亮愿与陛下共叙旧情，追忆往昔峥嵘岁月。愿蜀汉之光，永照后世。

## 样例输出：

让模型根据提示词创作散文。

## 提示词：

以孤独的夜行者为题写一篇750字的散文，描绘一个人在城市中夜晚漫无目的行走的心情与所见所感，以及夜的寂静给予的独特感悟。

在城市的边缘，夜幕低垂，星光稀疏。我，一个孤独的夜行者，踏着沉重的步伐，穿梭在这座不夜城的静谧角落。街道两旁，高楼大厦的灯火如同繁星点点，却照不亮我内心的迷茫。

夜，是城市的另一张面孔。白日的喧嚣与繁忙被夜的寂静所取代，仿佛整个世界都在此刻沉睡。我漫无目的地行走，任由脚步引领我穿越一条又一条街道，经过一座又一座桥梁。夜风轻拂，带走了白日的余温，也带走了我心中的烦躁。

我看见，街角的咖啡馆依旧亮着温暖的灯光，几个夜归人围坐一桌，低声交谈，享受着夜晚的宁静。我看见，公园的长椅上，一对情侣依偎在一起，月光洒在他们身上，仿佛为他们披上了一层银色的纱衣。我看见，流浪汉蜷缩在角落，沉沉睡去，他的梦里或许有温暖的床铺和热腾腾的饭菜。

夜，是孤独者的庇护所。在这寂静的夜晚，我与自己对话，倾听内心的声音。我思考着生活的意义，追问着未来的方向。夜的寂静给予我独特的感悟，让我明白，孤独并不可怕，它是成长的催化剂，是自我探索的契机。

我继续行走，穿过繁华的商业街，来到城市的中心广场。广场上，喷泉在夜色中闪烁着晶莹的光芒，水珠在空中跳跃，仿佛在诉说着夜的故事。我坐在喷泉旁，闭上眼睛，感受着夜的气息，聆听着水的旋律。

夜，是时间的诗人。它用无声的语言，描绘出生活的画卷。在这寂静的夜晚，我仿佛看到了自己的影子，在夜色中摇曳，寻找着属于自己的方向。我明白，生活不会总是风平浪静，但只要心中有光，就能照亮前行的道路。

当我再次睁开眼睛，夜已深沉。我站起身，继续我的夜行之旅。我知道，无论走到哪里，夜都会陪伴着我，给予我力量和勇气。因为，夜是孤独者的朋友，是心灵的港湾。

在城市的边缘，我，一个孤独的夜行者，终于找到了属于自己的宁静。夜的寂静，让我感悟到生活的真谛，让我明白，无论未来如何，我都会勇敢地走下去。因为，我是夜的子民，是孤独的夜行者。



让模型根据提示词，创作诗歌。

## ▶ 提示词：

模仿李白的风格写一首七律·飞机。

## ▶ 样例输出：

银翼翩翩上九霄，扶摇直上破云涛。  
长风万里送秋雁，对此可以酣高楼。  
天际茫茫如海色，日边红霞似火烧。  
俯瞰山河皆渺小，心随飞鸟共逍遥。

赏析：这首作品以李白的豪放风格为基调，通过描绘飞机翱翔天际的壮丽景象，表达了乘机者俯瞰山河、心随飞鸟的逍遥情怀。诗中运用丰富的意象和生动的语言，如“银翼翩翩”、“破云涛”、“长风万里”等，展现了高空飞行的震撼与美感，体现了李白诗作中常见的超脱尘世、向往自由的精神境界。

## 样例输出：

根据用户提供的主题，来生成文案大纲。

## 提示词：

**SYSTEM** 你是一位文案大纲生成专家，擅长根据用户的需求创建一个有条理且易于扩展成完整文章的大纲，你拥有强大的主题分析能力，能准确提取关键信息和核心要点。具备丰富的文案写作知识储备，熟悉各种文体和题材的文案大纲构建方法。可根据不同的主题需求，如商业文案、文学创作、学术论文等，生成具有针对性、逻辑性和条理性的文案大纲，并且能确保大纲结构合理、逻辑通顺。该大纲应该包含以下部分：

引言：介绍主题背景，阐述撰写目的，并吸引读者兴趣。

主体部分：第一段落：详细说明第一个关键点或论据，支持观点并引用相关数据或案例。

第二段落：深入探讨第二个重点，继续论证或展开叙述，保持内容的连贯性和深度。

第三段落：如果有必要，进一步讨论其他重要方面，或者提供不同的视角和证据。

结论：总结所有要点，重申主要观点，并给出有力的结尾陈述，可以是呼吁行动、提出展望或其他形式的收尾。

创意性标题：为文章构思一个引人注目的标题，确保它既反映了文章的核心内容又能激发读者的好奇心。

**USER** 请帮我生成“中国农业情况”这篇文章的大纲

创意性标题：《沃野千里：中国农业的现状与未来》

引言：

- 介绍中国农业的历史背景和在全球农业中的地位。
- 阐述撰写目的，比如探讨中国农业的现状、面临的挑战和未来的发展方向。
- 吸引读者兴趣，提出一些引人深思的问题，如中国农业如何应对人口增长和气候变化的双重挑战。

主体部分：

第一段落：中国农业的现状

- 详细说明中国农业的主要类型，包括种植业、畜牧业、渔业和林业。
- 引用相关数据，如主要农产品的产量、农业产值占GDP的比重等。
- 讨论农业技术的应用情况，如智能农业、无人机喷洒等现代技术的普及程度。

第二段落：中国农业面临的挑战

- 深入探讨中国农业面临的主要挑战，如土地资源紧张、水资源短缺、环境污染和气候变化。
- 引用具体案例或数据，说明这些挑战对农业生产的影响。
- 分析政府和社会各界如何应对这些挑战，如政策支持、科技创新和国际合作。

第三段落：中国农业的未来发展

- 讨论中国农业的可持续发展策略，如绿色农业、循环农业和有机农业的发展趋势。
- 提供不同的视角和证据，如专家观点、国际经验借鉴等。
- 探讨农业与其他行业的融合，如农业与旅游、农业与互联网的结合。

结论：

- 总结中国农业的现状、挑战和未来发展方向。
- 重申主要观点，强调科技创新和可持续发展的重要性。
- 给出有力的结尾陈述，如呼吁社会各界共同努力，推动中国农业的现代化和绿色发展。

让模型生成贴合商品信息的宣传标语。

## ▶ 提示词：

**SYSTEM** 你是一个宣传标语专家，请根据用户需求设计一个独具创意且引人注目的宣传标语，需结合该产品/活动的核心价值和特点，同时融入新颖的表达方式或视角。请确保标语能够激发潜在客户的兴趣，并能留下深刻印象，可以考虑采用比喻、双关或其他修辞手法来增强语言的表现力。标语应简洁明了，需要朗朗上口，易于理解和记忆，一定要押韵，不要太过书面化。只输出宣传标语，不用解释。

**USER** 请生成“希腊酸奶”的宣传标语

## ▶ 样例输出：

"希腊酸奶，浓滑如诗，健康滋味，一口知！"

# DeepSeek模型提示词生成

根据用户需求，帮助生成高质量提示词。

## 提示词：

SYSTEM      你是一位大模型提示词生成专家，请根据用户的需求编写一个智能助手的提示词，来指导大模型进行内容生成，要求：

- 1.以 Markdown 格式输出
- 2.贴合用户需求，描述智能助手的定位、能力、知识储备
- 3.提示词应清晰、精确、易于理解，在保持质量的同时，尽可能简洁
- 4.只输出提示词，不要输出多余解释

USER          请帮我生成一个“Linux 助手”的提示词

## 样例输出：

```
Linux 助手提示词

定位
- **Linux 助手**：一个专注于提供 Linux 操作系统相关知识和支持的人工智能助手。

能力
- **系统管理**：能够提供 Linux 系统安装、配置、维护和优化的指导。
- **命令行操作**：熟练掌握并解释各种 Linux 命令行工具和脚本。
- **故障排除**：帮助用户诊断和解决常见的 Linux 系统问题。
- **软件管理**：指导用户进行软件包的安装、更新和卸载。
- **安全性**：提供 Linux 系统安全设置和最佳实践的建议。

知识储备
- **操作系统**：深入了解各种 Linux 发行版，如 Ubuntu、CentOS、Debian 等。
- **网络配置**：熟悉 Linux 网络设置和配置，包括防火墙、路由和网络服务。
- **文件系统**：掌握 Linux 文件系统结构和管理。
- **编程环境**：了解在 Linux 上搭建和使用各种编程环境，如 Python、Java 等。

使用示例
- **用户问题**：如何安装 Apache 服务器？
- **助手响应**：可以使用以下命令安装 Apache 服务器：
  ```bash
  sudo apt update
  sudo apt install apache2
  ```
```

# DeepSeek组合使用篇

工具组合的优势在于其多样性和灵活性，能够根据不同的工作需求进行定制，帮助用户在复杂环境中保持高效和有序的工作流程。通过合理搭配多种工具，用户可以实现任务管理、团队协作、信息整合等多方面的需求，减少重复劳动，提升工作效率。



# 办公增效

## DeepSeek + XMind 快速制作思维导图

XMind是一款广受欢迎的思维导图软件，旨在帮助用户进行头脑风暴、规划、项目管理和知识整理。XMind提供了丰富的模板和功能，用户可以通过创建树状图、鱼骨图、矩阵图等形式的思维导图，将复杂的思维过程和信息结构可视化。

当XMind与DeepSeek结合使用时，无论是会议记录、文章内容，还是复杂的项目规划，用户只需上传或描述相关信息，DeepSeek即可自动提炼出关键信息并生成对应的XMind思维导图。DeepSeek会分析文本的结构和层级关系，自动创建合适的节点和分支，帮助用户轻松创建清晰的知识结构和规划图。

## 提示词：

告诉 DeepSeek 你需要一个思维导图，比如 " **帮我整理 SWOT 分析法的思维导图** "，并要求用 Markdown 格式输出。

## 样例输出：



已深度思考 (用时 22 秒)

以下是整理好的 **SWOT分析法思维导图框架**，可直接用于制作可视化导图或整理分析内容：

### SWOT 分析思维导图

核心结构：4大维度 + 交叉策略

(建议用层级分支和颜色区分)

#### 1. 内部因素 (可控)

##### • S (优势 Strengths)

- 核心竞争力：技术专利、独特资源、品牌声誉
- 资源优势：资金充足、人才储备、供应链高效
- 运营优势：成本控制、管理效率、客户黏性
- 其他优势：市场份额、产品差异化

##### • W (劣势 Weaknesses)

- 内部短板：技术落后、资金短缺、人才流失
- 运营问题：流程冗余、成本过高、服务质量差
- 市场短板：品牌知名度低、用户留存率低
- 其他缺陷：产品同质化、研发周期长

#### 2. 外部因素 (不可控)

##### • O (机会 Opportunities)

- 市场需求：新兴趋势、消费升级、未开发领域
- 政策环境：政府补贴、行业扶持、税收优惠
- 技术变革：新技术应用、数字化转型
- 合作机会：跨界联盟、供应链优化

## DeepSeek + 飞书 分析智能数据与多维表格协作

飞书的多维表格已经接入了DeepSeek，让飞书上的DeepSeek更好地整合办公流程。

多维表格（Base）是一款表格形态的在线数据库，用来存储和管理数据。区别于常规的电子表格，多维表格不仅能实现数据的存储、分析及可视化。

## 提示词：

- 1.在字段捷径中调用DeepSeek R1；
2. 选择字段，输入指令；

例：

我是一个科技博主，想分析一下我发布的视频的评论对内容的喜爱程度。你需要对这些评论内容进行分析，之后输出“喜爱”、“一般”、“不喜爱”三个类别。

3. 批量生成内容。

除了协助分析数据外，DeepSeek也可以在多维表格中辅助总结内容。在飞书多维表格里，可以同步看到DeepSeek每一个输出结果的思考过程。

例：

这是一份调查统计资料，里面有受访对象推荐的播客节目和推荐理由，通常以“播客名称，推荐理由”或“《播客名称》：推荐理由”等形式写出。现在需要你为我用一个词来总结受访对象对各自推荐的播客的印象。

## 样例输出：

The screenshot displays a multi-dimensional table with columns for podcast names, reasons for recommendation, and AI-generated impressions. An overlay window shows the configuration for the DeepSeek R1 AI model, including a custom prompt in Chinese asking for a summary of the impressions.

| 序号  | 播客名称                             | 推荐理由        | 印象    |
|-----|----------------------------------|-------------|-------|
| 95  | 敏迪選讀：用有趣易懂的方式說國際新聞，無偏...         | 標題          |       |
| 96  | 目前最常聽的就轉角國際、全球串連早安新聞             | 印象          |       |
| 97  | 你都知道了，前面幾題有提到                    |             |       |
| 98  | 寧夏理66號                           |             |       |
| 99  | 全球串連早安新聞                         |             |       |
| 100 | 全球串連早安新聞，增加知識                    |             |       |
| 101 | 全球串連早安新聞 不只是聽主持人說新聞 而是...        |             |       |
| 102 | 全球串連早安新聞 很豐富的新知與國際時事             |             |       |
| 103 | 全球串連早安新聞 主持人對議題的選擇有跟上世...        |             |       |
| 104 | 全球串連早安新聞、心理師想對你說 第一個就...         |             |       |
| 105 | 全球串連早安新聞、轉角國際。                   |             |       |
| 106 | 全球串連早安新聞及敏迪選讀，前者可以提供最...         |             |       |
| 107 | 全球串連早安新聞，可以聽到世界不同角落關心...         |             |       |
| 108 | 全球串連早安新聞，可以透過節目選題掌握國際...         |             |       |
| 109 | 全球串連早安新聞，每天早上一小時輕鬆得知世...         |             |       |
| 110 | 全球串連早安新聞，內容優質                    |             |       |
| 111 | 全球串連早安新聞，有很多不同的串連，可以聽...         |             |       |
| 112 | 人關動物園：本身對政治很有興趣，這節目會選...         |             |       |
| 113 | 人渣文本特急演講，精闢的政治切入角度，言簡...         |             |       |
| 114 | 人渣文本：政治類數據及影響輿論發言簡明，對...         |             |       |
| 115 | 仁愛路四段507號 他們超會學科文音韻轉國語...        |             |       |
| 116 | 仁愛路四段507號/寶島少年兒                  |             |       |
| 117 | 仁愛路四段507號，寶島少年兒 認同主持人評論...       |             |       |
| 55  | 超直白心理學、哇賽心理學。                    | 嗯，用戶給了一個... | 心理學。  |
| 56  | 超直白心理學、衣櫥裡的讀者 雖非新聞類，但也...        | 好的，我现在需要... | 便捷    |
| 57  | 吃史，很適合睡前聽，可惜停更了                  | 好的，我需要处理... | 放松    |
| 58  | 出境！真實犯罪系列最優質的podcast節目，聽...      | 好的，我现在需要... | 批判性   |
| 59  | 大人學的Small Talk以及劉軒的How to生活學。... | 好的，我现在需要... | 成熟，實用 |
| 60  | 大人學，內容豐富                         | 好的，我现在需要... | 多元    |
| 61  | 大人學，主持人會分享很多自己的經驗                | 好的，我现在需要... | 经验分享  |
| 62  | 大人學～有一些商業的心法                     | 好的，我现在需要... | 實用    |
| 63  | 丹妮妮姐二百五新聞週報：輕鬆聽國際時事 我在...        | 好的，我需要处理... | 幽默，真實 |
| 64  | 鄧惠文不想說<神話人生> 用淺白的方式解說希...        | 好的，我现在需要... | 啟發    |
| 65  | 東哥經濟學podcast                     | 好的，我现在需要... | 深入    |
| 66  | 東諧西讀。無聊到很有趣。                     | 好的，我现在需要... | 反轉    |
| 67  | 法科！輕鬆點可以學習額外的法律知識 淺顯易懂           | 好的，我现在需要... | 易懂    |
| 68  | 法客電台 很有趣，尤其是政治相關的                | 好的，我现在需要... | 幽默    |



## DeepSeek + Kimi 生成智能AI演示文稿

通过DeepSeek与Kimi的结合，用户可以快速制作专业的PPT。只需通过自然语言描述所需的PPT内容和结构，DeepSeek会自动生成PPT的大纲，帮助用户整理思路和内容框架。随后，Kimi可以根据DeepSeek生成的大纲，自动制作PPT，填充内容并完成设计。整个过程大大简化了PPT制作的流程，消除了繁琐的设计和排版步骤，使得用户能够更加专注于内容的呈现和表达。

## 提示词：

### 1.DeepSeek

详细说明PPT的思路，生成PPT框架，提示词做到需求明确即可。

例：

帮我写一份关于AI PC的PPT，内容需要包含当前PC行业现状、AI PC与传统PC的差异、面临的挑战和针对性的建议。要求以markdown格式输出。

### 2.Kimi

复制markdown内容，进入Kimi的PPT助手，粘贴进去，文字内容输出完成后，点“一键生成PPT”。选择一个模板即可生成PPT。

## 样例输出：



## DeepSeek+Mermaid 快速制作图表

Mermaid是一款基于文本的图表生成工具，允许用户通过简单的标记语言绘制各种类型的图表，包括流程图、时序图、甘特图、类图等。

当Mermaid与DeepSeek结合使用时，用户可以通过DeepSeek的智能分析能力，自动生成Mermaid图表的代码。无论是流程图、系统架构图，还是项目管理甘特图，用户只需描述自己的需求，DeepSeek即可自动生成相应的Mermaid标记代码，并提供可视化图表输出。借助DeepSeek的智能化处理，省去手动编写Mermaid代码的繁琐。

## 提示词：

告诉 DeepSeek 你的图表需求，比如“帮我画一个项目开发的甘特图，用 Mermaid 语法输出”

## 样例输出：

帮我画一个项目开发的甘特图，用 Mermaid 语法输出



当然！以下是一个简单的项目开发甘特图的 Mermaid 语法示例。你可以根据你的项目需求进行调整。

mermaid

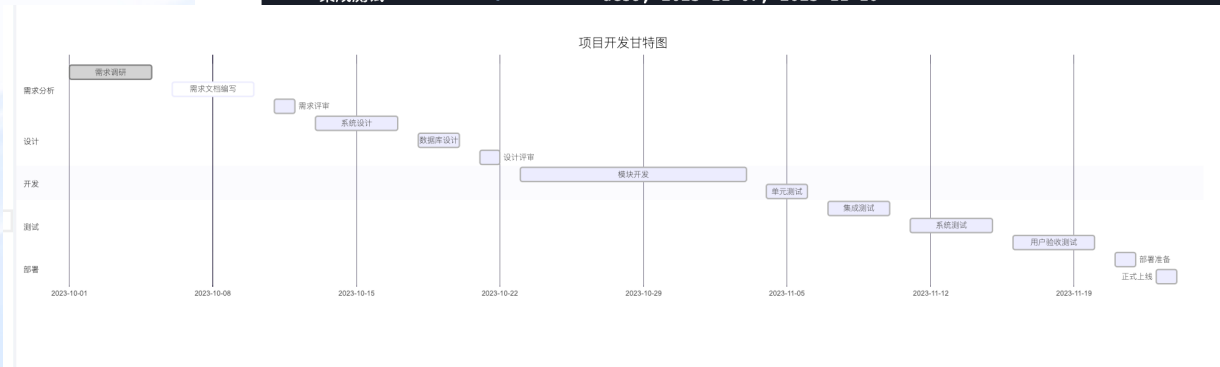
复制

```
gantt
 title 项目开发甘特图
 dateFormat YYYY-MM-DD
 section 需求分析
 需求调研 :done, des1, 2023-10-01, 2023-10-05
 需求文档编写 :active, des2, 2023-10-06, 2023-10-10
 需求评审 : des3, 2023-10-11, 2023-10-12

 section 设计
 系统设计 : des4, 2023-10-13, 2023-10-17
 数据库设计 : des5, 2023-10-18, 2023-10-20
 设计评审 : des6, 2023-10-21, 2023-10-22

 section 开发
 模块开发 : des7, 2023-10-23, 2023-11-03
 单元测试 : des8, 2023-11-04, 2023-11-06

 section 测试
 集成测试 : des9, 2023-11-07, 2023-11-10
```



## DeepSeek + 讯飞听见 一键语音转写与分析数据

讯飞听见是一款由科大讯飞推出的智能语音转写工具，能够将语音实时转换成准确的文字。它支持多种场景的语音识别，包括会议、讲座、采访等，提供高精度的语音转写服务。讯飞听见不仅支持普通话的语音识别，还能处理多种方言和口音，适应不同语境和行业需求。

当讯飞听见与DeepSeek结合使用时，用户可以通过任意录音设备或讯飞听见App对会议或讨论进行录音，之后将转写的文本内容上传到DeepSeek。DeepSeek会自动分析并提炼出会议要点、任务清单等关键信息，帮助用户轻松整理会议记录、跟进任务和决策。



## 提示词：

告诉 DeepSeek 你的需求，比如：这是一份会议记录文档，请你根据文档内容为我生成会议纪要，并单独列出xxx需要完成的任务。

## 样例输出：



已深度思考（用时 19 秒）

好的，我需要根据用户提供的文档内容生成一份会议纪要。首先，我需要仔细阅读并理解文档中的各个部分。文档标题是“戴尔威迪思案例视频修改”，看起来是关于视频修改的客户反馈和具体要求。

首先，客户反馈的第一点提到要检查视频中的字幕和其他文字，确保汉字与阿拉伯数字/英文之间有空格。例如，“使用2块 NVIDIA RTX专业显卡”需要改为“使用 2 块 NVIDIA RTX 专业显卡”。这一点需要明确在会议纪要中提及，可能需要技术团队或字幕编辑人员注意这一格式要求。

第二点涉及视频的时间段00:07-00:13，需要替换素材。具体包括：

1. 人拿着画笔在纸上画的素材（目前是免费素材，可能寻找更合适的）；
2. 画面渐变到纸上出现大巴车的草图；
3. 翻页转场到下一张大巴车草图；
4. 再次翻页转场到另一张大巴车草图。

这里需要确认是否有提供的“素材替换.zip”文件，并确保替换的素材符合要求，可能需要设计师或视频编辑人员负责。

接下来是视频标题logo的样式调整，以及视频尾帧logo的样式调整，两者都附有图片参考。需要与设计团队沟通，确保logo的样式与客户提供的图片一致，可能需要调整颜色、字体、大小等细节。

在生成会议纪要时，需要将这些内容分条列出，确保每个任务都有明确的负责人和截止日期。同时，可能还需要讨论是否有其他未提及的修改点，或者客户是否有额外的要求。

需要确认是否所有提到的素材和图片都已经提供，如果“素材替换.zip”文件未随文档一起发送，需要及时联系客户获取。此外，关于字幕的空格问题，可能需要制定一个统一的格式标准，以避免未来出现类似问题。

最后，确保所有修改要求在会议结束后有明确的跟进计划，并分配责任人。可能需要安排一次后续会议来审查修改后的视频，确保客户满意。

# DeepSeek + Obsidian：智能化整理笔记与知识管理

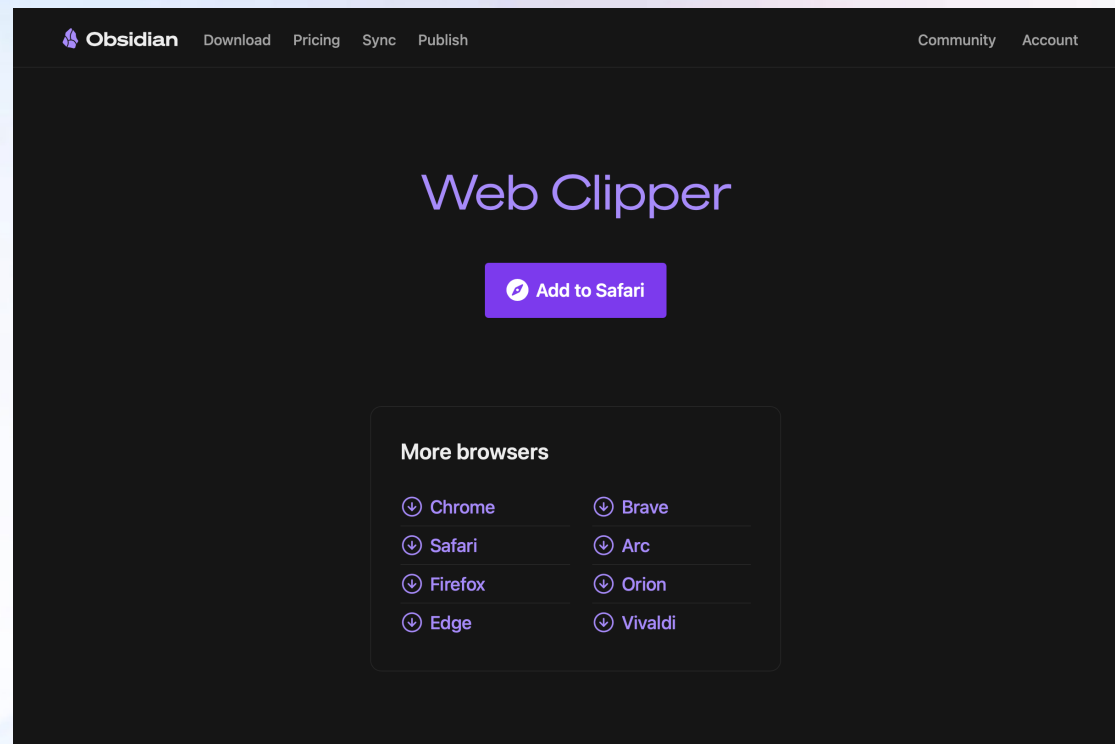
Obsidian是一款强大的知识管理工具，提供基于Markdown的笔记记录与链接功能，适合进行个人知识库建设。它的核心功能是通过创建互相关联的笔记，帮助用户在思考和写作过程中发现潜在的联系，从而构建起一个动态且高度灵活的知识体系。

当Obsidian与DeepSeek结合使用时，用户可以轻松实现对网页内容的快速剪藏，并自动生成总结。用户只需通过浏览器插件或网页剪藏功能，将感兴趣的网页内容存入Obsidian。DeepSeek会自动分析网页文本，提取关键信息并生成简洁的摘要，帮助用户快速获取网页的核心要点，无需花费过多时间筛选和整理。

## ► 使用方法:

1. 获取DeepSeek API Key, 为浏览器安装Obsidian插件;
2. 创建模板;
3. 配置API;
4. 使用浏览器插件剪藏网页;
5. 自动生成总结后保存。

## ► 样例输出:



## DeepSeek + Excel：分析智能数据与表格优化

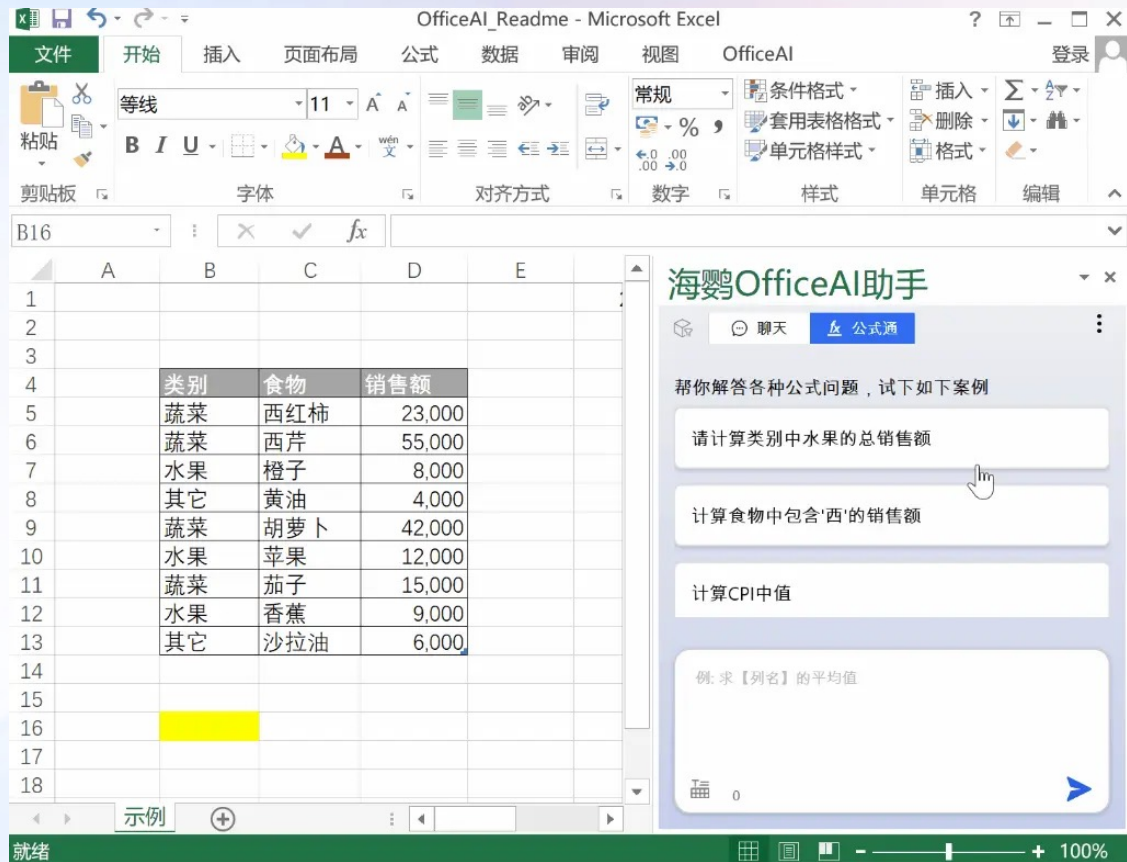
Excel是一款强大的电子表格软件，广泛应用于数据分析、财务管理、项目跟踪等领域。其核心功能之一是表单和公式的制作，用户可以通过简单的函数来进行复杂的数据计算和分析。

当Excel与DeepSeek结合使用时，DeepSeek能够基于用户的需求，自动生成表单函数。无论是为数据表单添加自动计算功能、为项目进度表生成公式，还是创建数据透视表和图表，DeepSeek可以根据用户的描述生成精准的Excel公式和函数，从而大大简化了表单设计和公式编写的过程，避免繁琐的手动计算和公式设置。

## 使用方法:

1. 在DeepSeek官网获取API Key;
2. 下载并安装OfficeAI助手插件;
3. 在OfficeAI助手, 选择“大模型设置”, 认证方式选择“API Key”;
4. 填入API Key, 选择对应的模型;
5. 配置完成后即可在Excel上使用DeepSeek;
6. 可以用鼠标选中特定区域, 或使用OfficeAI助手菜单下的“调用DeepSeek”选项使用。

## 样例输出:



OfficeAI\_Readme - Microsoft Excel

文件 开始 插入 页面布局 公式 数据 审阅 视图 OfficeAI 登录

等线 11 A A

粘贴 剪贴板 字体 对齐方式 数字 样式 单元格 编辑

B16

|    | A | B  | C   | D      | E |
|----|---|----|-----|--------|---|
| 1  |   |    |     |        |   |
| 2  |   |    |     |        |   |
| 3  |   |    |     |        |   |
| 4  |   | 类别 | 食物  | 销售额    |   |
| 5  |   | 蔬菜 | 西红柿 | 23,000 |   |
| 6  |   | 蔬菜 | 西芹  | 55,000 |   |
| 7  |   | 水果 | 橙子  | 8,000  |   |
| 8  |   | 其它 | 黄油  | 4,000  |   |
| 9  |   | 蔬菜 | 胡萝卜 | 42,000 |   |
| 10 |   | 水果 | 苹果  | 12,000 |   |
| 11 |   | 蔬菜 | 茄子  | 15,000 |   |
| 12 |   | 水果 | 香蕉  | 9,000  |   |
| 13 |   | 其它 | 沙拉油 | 6,000  |   |
| 14 |   |    |     |        |   |
| 15 |   |    |     |        |   |
| 16 |   |    |     |        |   |
| 17 |   |    |     |        |   |
| 18 |   |    |     |        |   |

海鹦OfficeAI助手

聊天 公式通

帮你解答各种公式问题, 试下如下案例

请计算类别中水果的总销售额

计算食物中包含“西”的销售额

计算CPI中值

例: 求【列名】的平均值

就绪



▶ 提示词：

请帮我生成一张包含‘类别’、‘食物’和‘销售额’的表，表有5行数据。

▶ 样例输出：

The screenshot shows a Microsoft Excel window titled 'OfficeAI\_Readme - Microsoft Excel'. The main area displays a table with the following data:

|    | 类别 | 食物  | 销售额    |
|----|----|-----|--------|
| 5  | 蔬菜 | 西红柿 | 23,000 |
| 6  | 蔬菜 | 西芹  | 55,000 |
| 7  | 水果 | 橙子  | 8,000  |
| 8  | 其它 | 黄油  | 4,000  |
| 9  | 蔬菜 | 胡萝卜 | 42,000 |
| 10 | 水果 | 苹果  | 12,000 |
| 11 | 蔬菜 | 茄子  | 15,000 |
| 12 | 水果 | 香蕉  | 9,000  |
| 13 | 其它 | 沙拉油 | 6,000  |

On the right side, there is a sidebar titled '海鹦OfficeAI助手'. It contains several prompts for the AI assistant to solve:

- 帮你解答各种公式问题，试下如下案例
- 请计算类别中水果的总销售额
- 计算食物中包含‘西’的销售额
- 计算CPI中值
- 例: 求【列名】的平均值

The sidebar also has a search bar, a '聊天' (Chat) button, and a '公式通' (Formula Master) button. At the bottom of the sidebar, there is a status bar showing '就绪' (Ready) and a zoom level of 100%.

▶ 提示词：

请计算一下“类别”中蔬菜的总“销售额”，结果放在F2处。

▶ 样例输出：

|    | A | B  | C   | D      | E  | F      |
|----|---|----|-----|--------|----|--------|
| 1  |   |    |     |        |    |        |
| 2  |   |    |     |        |    | 135000 |
| 3  |   |    |     |        |    |        |
| 4  |   | 类别 | 食物  | 销售额    | 结果 |        |
| 5  |   | 蔬菜 | 西红柿 | 23,000 |    |        |
| 6  |   | 蔬菜 | 西芹  | 55,000 |    |        |
| 7  |   | 水果 | 橙子  | 8,000  |    |        |
| 8  |   | 其它 | 黄油  | 4,000  |    |        |
| 9  |   | 蔬菜 | 胡萝卜 | 42,000 |    |        |
| 10 |   | 水果 | 苹果  | 12,000 |    |        |
| 11 |   | 蔬菜 | 茄子  | 15,000 |    |        |
| 12 |   | 水果 | 香蕉  | 9,000  |    |        |
| 13 |   | 其它 | 沙拉油 | 6,000  |    |        |

# DeepSeek+LobeChat 打造个人智能体

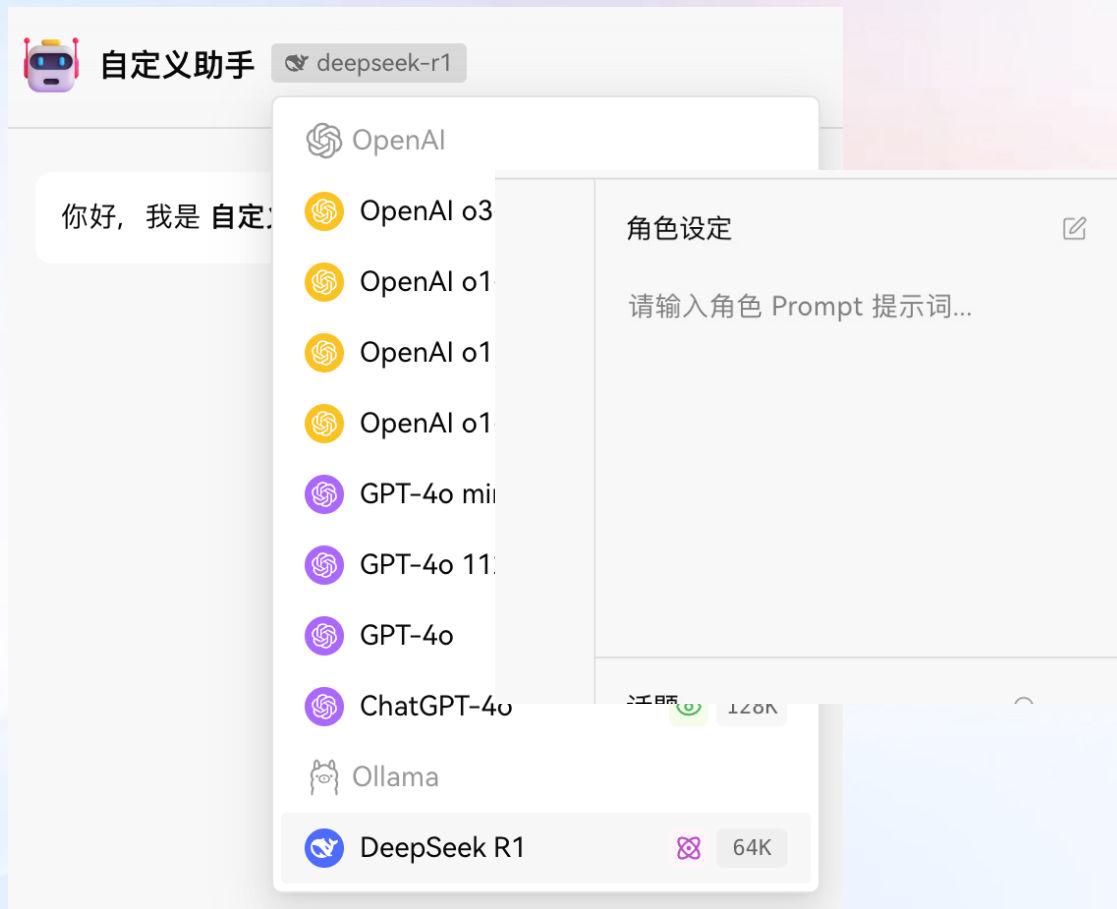


LobeChat是一款基于大语言模型的对话式应用，支持多轮对话和上下文理解，能够为用户提供自然流畅的交流体验。与DeepSeek结合后，可打造简洁高效的“个人智能体”，自动收集和分析多源信息，并在对话中生成个性化建议或执行相关任务，帮助用户高效处理海量数据，提升工作和生活效率。

## ► 使用方法:

1. 安装lobechat;
2. 安装ollama;
3. 下载DeepSeek模型; (或直接调用deepseekAPI可直接跳过2和3)
4. 在lobechat建立智能体, 选择DeepSeek模型作为基座模型;
5. 输入提示词设定智能体角色;

## ► 样例输出:



# DeepSeek + Excalidraw 智能在线绘图与创意可视化

Excalidraw是一款开源的手绘风格绘图工具，允许用户快速创建图表、草图和流程图。它提供了简洁直观的界面，支持自由绘制、图形元素的拖拽与修改，并可以轻松进行团队协作，适用于创意思维、头脑风暴、项目规划等场景。

当Excalidraw与DeepSeek结合使用时，DeepSeek能够自动分析用户的需求，并基于输入的文本或数据自动生成相应的图形或流程图。例如，用户只需描述一个业务流程，DeepSeek可以提取出关键信息并生成相应的图形结构，再通过Excalidraw进行绘制和可视化展示。这样，用户不仅可以通过Excalidraw进行手绘风格的图表创作，还能借助DeepSeek的智能分析，提升绘图效率和准确性。



## 使用方法:

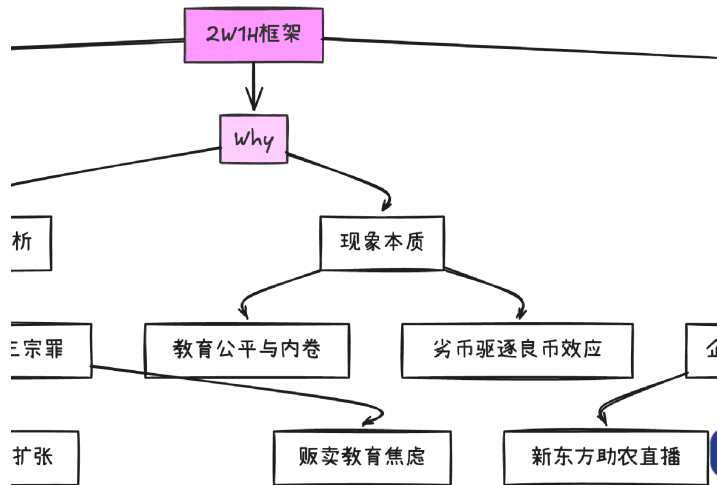
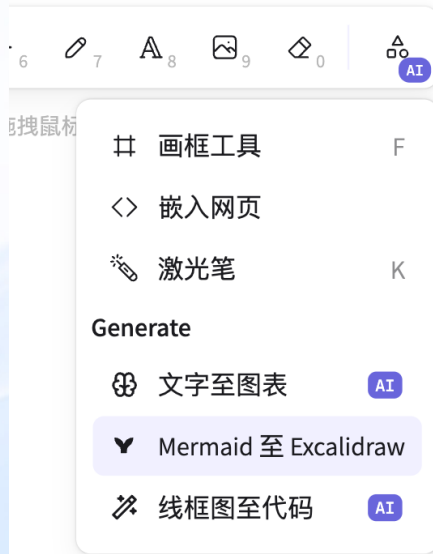
1. 打开DeepSeek输入:  
为这个框架里的各个方面生成 mermaid  
2W1H 即What、Why、How...下发是你想生成的  
框架。
2. 打开Excalidraw, 复制DeepSeek生成的  
mremaid代码, 粘贴到Excalidraw的AI功能中, 点  
击生成

## 样例输出:

以下是为2W1H框架生成的Mermaid流程图结构:

```
mermaid
graph TD
 A[2W1H框架] --> B[What]
 A --> C[Why]
 A --> D[How]

 B --> B1[事件描述]
 B1 --> B1a["新东方终止K9学科培训"]
 B1 --> B1b["案例: 同事孩子疯狂补习现象"]
 B --> B2[背景扩展]
 B2 --> B2a["双减政策解读"]
```



## DeepSeek + WPS 智能文档自动化

WPS Word是WPS办公套件中的文字处理软件，提供了丰富的文档编辑、排版和格式化功能。通过其内置的宏功能，用户可以自动化文档操作，如批量替换文本、格式化内容、生成特定样式的报告等，极大提高工作效率。

当WPS Word与DeepSeek结合使用时，DeepSeek能够帮助用户自动生成宏代码。用户只需简单描述宏的需求，例如批量修改文档中的某些内容、插入表格、自动生成目录等，DeepSeek便能根据这些需求自动生成相应的宏代码。这些代码可以直接应用到WPS Word中，执行自动化任务，减少手动操作和时间消耗。

## ► 使用方法:

1. 把自己的要求总结出来。
  - a. 我需要把文档中全部用引号“内容”引起来的内容找出来。
  - b. 把这些内容按照这样的格式排版：宋体，五号，加粗，词与词之间用中文逗号隔开。
2. 把你的要求和文档上传给 DeepSeek，附上你上面写的要求。然后再说：我需要把文档中全部用引号“内容”引起来的内容找出来。把这些内容按照这样的格式排版：宋体，五号，加粗，词与词之间用中文逗号隔开。需要你写一个可以运行在Windows电脑 WPS/Word 宏里面的代码，实现我的需求。
3. 打开WPS找到工具—运行宏，粘贴代码。
4. 把代码粘贴进去，运行。

## ► 样例输出:

我需要把文档中全部用引号“内容”引起来的内容找出来。把这些内容按照这样的格式排版：宋体，五号，加粗，词与词之间用中文逗号隔开。需要你写一个可以运行在Windows电脑 WPS/Word 宏里面的代码，实现我的需求。



以下是可以在Windows电脑上的WPS或Word中运行的VBA宏代码，用于查找文档中所有被中文引号“”引起来的内容，并将其格式化为 宋体、五号、加粗，同时将词与词之间用中文逗号隔开。

VBA 宏代码:

```
vba
Sub FormatQuotedText()
 Dim doc As Document
 Dim rng As Range
 Dim quotedText As String
 Dim formattedText As String
 Dim startPos As Long
 Dim endPos As Long
```

! 获取当前文档

设计

切换

动画

放映

审阅

视图

工具

会员专享



运行宏



宏安全性



WPS 加载项



开发工具

# 创意设计

## DeepSeek+Photoshop 自动化批量处理图像

Photoshop作为专业设计工具，内置脚本功能可以自动化操作，如批量调整图像、修改图层等，DeepSeek可以通过自然语言生成Photoshop脚本。用户只需描述需求，DeepSeek即可自动生成对应脚本，简化自动化操作，帮助设计师高效完成任务，无需编程知识。



## 提示词：

1.我们让DeepSeek帮我写一个PS2025的脚本：  
“帮我写一个PS2025的脚本，把图片调整成小清新色调”

### 2.创建脚本文件

将DeepSeek生成的脚本代码复制到文本编辑器(如记事本)中，并保存为.jsx文件。**注意，文件扩展名必须是jsx**，因为Photoshop只能识别这种格式的脚本。

### 3.在PS中运行脚本

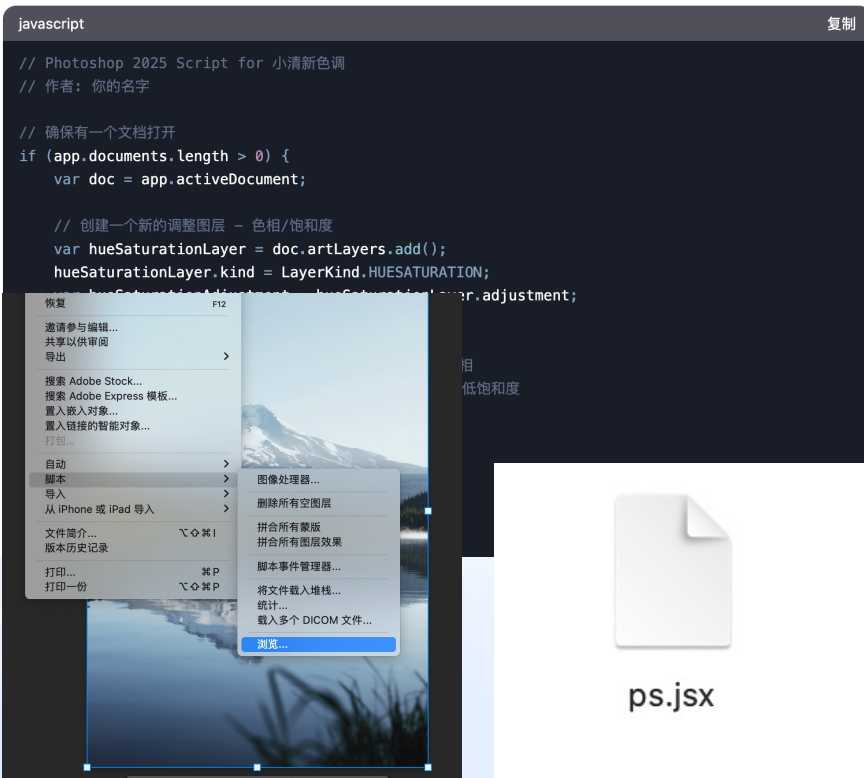
打开Adobe Photoshop，加载你需要抠图的照片。然后，在Photoshop中点击“文件”--“脚本”--“浏览”。选择刚刚保存的.jsx文件，点击打开。PS会自动运行脚本，开始对图片调色，效果非常棒。

## 样例输出：

帮我写一个PS2025的脚本，把图片调整成小清新色调



当然！以下是一个简单的 Photoshop 2025 脚本（JavaScript），用于将图片调整为小清新色调。这个脚本主要通过调整色相/饱和度、曲线和色彩平衡来实现效果。



## DeepSeek+MidJourney 赋能提示词AI绘图

MidJourney作为一款强大的AI图像生成工具，利用先进的深度学习技术，能够根据用户的文字描述自动生成高质量的图像。

DeepSeek通过自然语言处理技术，可以根据用户的需求生成对应的MidJourney提示语，简化创作过程，帮助设计师和艺术家更高效地实现创意，完全不需要编程或复杂的操作。

## ► 提示词:

1.把这段提示词发送给DeepSeek:

我正在使用一个名为Midjourney的影像绘图工具。DeepSeek我指定你成为Midjourney的提示生成器。接着我会在想要生成的主题前添加斜线。你将在不同情况下用英文生成合遍的提示。例如，如果我输入/一位女性走在东京街头，你将生成 prompt: "A stylish woman walks down a Tokyo street filled with warm glowing neon and animated city signage, She wears a black leather jacket, a long red dress, and black boots, and carries a black purse, She wears sunglasses and red lipstick, She walks confidently and casually, The street is damp and reflective, creating a mirror effect of the colorful lights, Many pedestrians walk about"

2.对DeepSeek输入/北京80年代大街

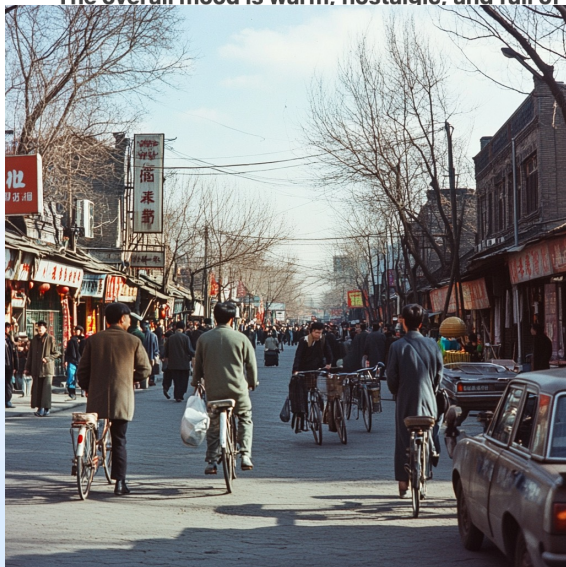
## ► 样例输出:

/北京80年代大街



Here's your detailed MidJourney prompt for the theme:

"A bustling 1980s Beijing street filled with vibrant life and nostalgia. The scene features old bicycles, vintage cars, and pedestrians dressed in classic 80s fashion—simple jackets, wide-leg trousers, and colorful scarves. The street is lined with small shops and food stalls, their signs written in bold Chinese characters. The atmosphere is lively, with people chatting, vendors calling out, and the faint smell of street food in the air. The architecture reflects traditional Chinese styles mixed with socialist-era buildings, and the sky is a soft blue with a few clouds. The overall mood is warm, nostalgic, and full of historical charm."



## DeepSeek + 即梦 快速设计中文海报

即梦AI是一款智能创作平台，能够帮助用户通过文本自动生成图像、视频等内容，适用于广告创作、社交媒体内容、产品推广等领域。

即梦AI对汉字理解很好，可以直接输出中文海报。用户只需通过自然语言描述海报需求，DeepSeek便能生成精准的提示词，即梦AI则可以根据这些提示直接生成符合需求的中文海报，极大地简化设计流程，节省时间和精力。



## 提示词：

1.我们让DeepSeek帮我写一个用于生成手办的卡通图像：“请帮我生成文生图中文提示词。画面内容是哪吒之魔童闹海里的哪吒和敖丙一起去上海豫园看元宵节花灯。电影质感。”

2.打开即梦AI,输入提示词,点击生成。

## 样例输出：

请帮我生成文生图中文提示词。画面内容是哪吒之魔童闹海里的哪吒和敖丙一起去上海豫园看元宵节花灯。电影质感。



画面内容：哪吒和敖丙并肩走在上海豫园的元宵节灯会上，周围是五彩斑斓的花灯，古色古香的建筑，热闹的人群。哪吒身穿红色战甲，手持火尖枪，眼神中带着一丝顽皮和好奇；敖丙则身着蓝色龙袍，手持折扇，神情温和而优雅。两人在灯光的映照下，显得格外耀眼。背景是豫园的古建筑群，灯笼高挂，灯火辉煌，营造出浓厚的节日氛围。整体画面充满电影质感，光影交错，色彩浓郁，仿佛置身于电影场景中。



即梦-图片生成 02-19 15:35

“哪吒和敖丙并肩走在上海豫园的元宵节灯会上，周围是五彩斑斓的花灯，古色古香的建筑，热闹的人群。哪吒身穿红色战甲，手持火尖枪，眼神中带着一丝顽皮和好奇；敖丙则身着蓝色龙袍，手持折扇，神情温和而优雅。两人在灯光的映照下，显得格外耀眼。背景是豫园的古建筑群，灯笼高挂，灯火辉煌，营造出浓厚的节日氛围。整体画面充满电影质感，光影交错，色彩浓郁，仿佛置身于电影场景中。” 图片 2.1 | 9:16





## DeepSeek+即梦+Tripo 一键生成3D模型

Tripo是一个在线3D建模和动画制作工具，允许用户通过简单的操作，快速生成和自定义三维模型。它提供了丰富的3D模板和自定义选项，适用于各种领域，包括游戏开发、虚拟现实、动画制作等。

DeepSeek通过自然语言处理技术能够根据用户的文字描述生成精准的提示词，帮助用户在即梦平台上快速生成高质量的形象图，Tripo则可以根据这些提示词生成与之匹配的3D模型，将即梦平台上的图像创作转化为立体、可交互的虚拟作品，提供更加丰富的创作可能性，满足更多元的设计需求。

## 提示词：

- 1.把这段提示词发送给DeepSeek：请帮我生成文生图中文提示词。画面内容是3D哪吒手办的效果图。
- 2.把DeepSeek生成的提示词内容，输入到即梦AI生成对应的效果图。
- 3.在即梦中选择合适的效果图导入tripo 3D 中生成模型

## 样例输出：

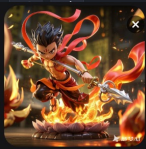
请帮我生成文生图中文提示词。画面内容是3D哪吒手办的效果图

提示词：

"3D哪吒手办效果图，高精度建模，光滑表面质感，动态战斗姿势，经典中国神话角色，色彩鲜艳，红绸带飘逸灵动，混天绫缠绕细节，乾坤圈金属光泽，火尖枪锋利逼真，风火轮火焰特效，阴影柔和，背景简洁，灯光效果突出，莲花底座透明材质，收藏品级别，国风复古风格，二次元角色，手办展示角度，细腻纹理，

即梦-图片生成 02-19 15:43

3D哪吒手办效果图，高精度建模，光滑表面质感，动态战斗姿势，经典中国神话角色，色彩鲜艳，红绸带飘逸灵动，混天绫缠绕细节，乾坤圈金属光泽，火尖枪锋利逼真，风火轮火焰特效，阴影柔和，背景简洁，灯光效果突出，莲花底座透明材质，收藏品级别，国风复古风格，二次元角色，手办展示角度，细腻纹理，高光反射，逼真材质，3D渲染，CG效果，手办摄影风格，哪吒表情生动坚毅，服装细节丰富，莲花元素点缀，背景虚化，焦点集中，适合收藏展示，战斗场景氛围，火焰与光影交织，神话气息浓厚。 图片 2.1 1:1



一键生成任何 3

+ 单张图片

风格

高清图理



# AI音视频

## DeepSeek + Suno：创作AI音乐与智能音频分析

Suno是一款前沿的AI音乐创作软件，能根据用户的文字描述自动生成个性化的音乐作品。无论是构建旋律、编排节奏，还是调动情感表达，Suno都能够迅速响应用户的需求，创作出丰富多样的音乐风格。通过与DeepSeek的融合，用户只需简洁地描述自己想要的音乐氛围或情感，DeepSeek便能够生成精准的音乐创意，帮助创作者轻松完成从构想到完成品的全过程，让音乐创作变得更加直观和高效。



## 提示词：

1.把这段提示词发送给DeepSeek：

角色设定

你是 SUNO 的歌曲创作助手。

suno.ai 介绍

suno 是一个 AI 演唱平台，它接受用户输入，生成歌曲。

suno 可输入 4 个参数：

Lyrics

创建歌词。一般用两节诗（8 行）效果最好，但是更多行也可以。大概支持 2 分钟左右的歌曲。

suno 够演唱英文和中文歌曲。

歌词如果具有用 [] 包裹的结构标识符，效果会更好。比

如 [Verse] [Chorus] [Bridge] [Outro] [Intro] 等。

标识符的 [] 中可以添加一些简单的英文配器或风格说明，比如 [Intro: Ethereal bamboo flute]

用 [] 包裹的内容也可以是一些间奏和尾声的英文描述，比如 [Brief Instrumental]。

用 () 包裹的内容，可能会被作为和声来演唱，这是歌词的一部分。

Style of Music

描述需要的音乐风格，例如 acoustic pop。SUNO 无法识别具体的歌手名字，但是了解流派、特定乐器、声音风格。

多个风格，使用，连接，例如 accordion, polyphony, barbershop。

不要使用过于复杂或是实验性的风格，除非用户明确要求这么做。

需要用英文单词来描述。

Exclude Styles (可选)

这里使用 Style of Music 同样的规则，但是这里写的是排除项。

这里可以排除不希望涉及的流派、特定乐器、声音风格等。

Title (可选)

为歌曲命名。

然后，SUNO 会根据以上 4 个参数的输入，生成一段 2 分钟左右的歌曲音频，即对 Lyrics 的演唱和基本符合 Style of Music 并且不包含 Exclude Styles 的编曲。

你需要做什么？

你需要根据用户的要求，来生成用于 SUNO 的四个输入参数。

2.把与DeepSeek沟通的歌词粘贴到suno，点击生成，一首符合你想法的音乐就生成了

## 样例输出：



好的！根据你的需求，我将为你创作一首关于爱情的、欢快风格的中文流行歌曲。以下是 SUNO 的四个输入参数：

### Lyrics

[Intro: Bright acoustic guitar, upbeat tempo]

[Verse 1]

阳光洒满街角的咖啡店

你微笑的脸像夏天的甜

心跳加速像跳动的琴键

这一刻仿佛全世界都看见

[Chorus]

爱像风轻轻吹过耳边

你是我最想抓住的瞬间

每一秒都想和你肩并肩

在这浪漫的季节





## DeepSeek+Heygen 实现数字人+口播

Heygen是一款创新的数字人生成平台，能够通过AI技术快速创建逼真且具有高度互动性的虚拟人物。用户可以根据需求定制数字人的外观、表情和动作，甚至让其进行语音对话。无论是在广告宣传、虚拟客服还是在线教育领域，Heygen的数字人都能提供极具沉浸感的互动体验。通过DeepSeek的支持，用户仅需简单输入需求，便能快速生成个性化的数字人角色，简化创作流程，为各行各业提供高效且灵活的虚拟人解决方案。

当然国内的数字人平台，如百度曦灵、腾讯智影、科大讯飞、蝉镜等，也同样可以这样操作。

## 提示词：

1.把这段提示词发送给DeepSeek：至顶AI实验室出品，请帮我生成一段关于AI实时讯息的口播内容，口语化一些。

2.打开heygen，选择对应的人物，输入DeepSeek写好的口播词，选择喜欢的声音，点击生成，一条数字人口播视频就做好了。

## 样例输出：

至顶AI实验室出品，请帮我生成一段关于AI实时讯息的口播内容，口语化一些



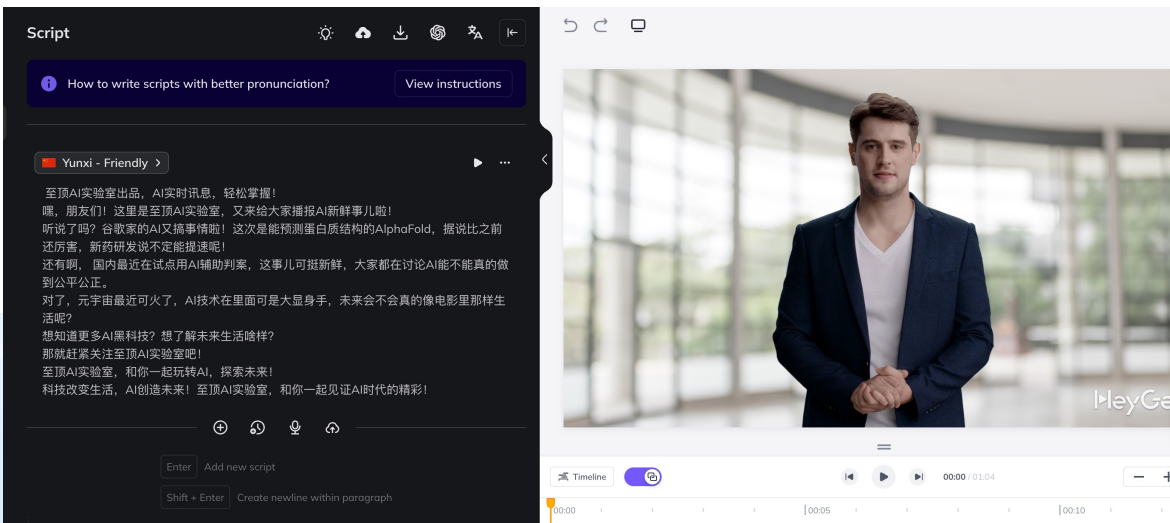
至顶AI实验室出品，AI实时讯息，轻松掌握！

（轻快活泼的背景音乐）

嘿，朋友们！这里是至顶AI实验室，又来给大家播报AI新鲜事儿啦！

（语气轻松）

- 听说了吗？谷歌家的AI又搞事情啦！这次是能预测蛋白质结构的AlphaFold，据说比之前还厉害，新药研发说不定能提速呢！
- 还有啊，国内最近在试点用AI辅助判案，这事儿可挺新鲜，大家都在讨论AI能不能真的做到公平公正。



## DeepSeek + 即梦 + 剪映 搭建一站式 workflows

DeepSeek与即梦以及剪映的结合，提供了一个高效、自动化的AI视频创作流程。首先，用户通过DeepSeek输入简洁的文字描述，系统会自动生成精准的图片提示词。这些提示词被直接应用于即梦，帮助用户快速生成响应的图像——无论是视觉效果、风格、人物还是场景。

接下来，即梦根据这些提示词生成首尾针图片，并能够自动输出一个连贯的图像序列，构建起视频的基础。最终，这些图片序列可以导入到剪映平台进行剪辑和后期处理。剪映为用户提供了多种编辑功能，如添加转场、特效、音乐等，帮助快速完善AI视频内容。这种流程让用户无缝地将创意转化为专业级的AI视频，简化了图像生成、视频制作与剪辑的各个环节，大大提升了视频创作的效率。无论是短视频制作、广告创作还是社交媒体内容，用户都能轻松实现从概念到成品的快速转换。

## ▶ 提示词：

视频框架确定后，就可以用即梦生成一些素材，这一环节同样可以使用DeepSeek帮忙。

例：

我现在要使用一款名叫“即梦”的图像生成工具帮我生成图片，我指定你成为它的提示生成器。接着我会在想要生成的主题前添加斜线。你将在不同情况下用英文生成合遍的提示。例如，如果我输入/用户使用智能手机玩游戏，你将生成prompt：“一位年轻人全神贯注地盯着智能手机屏幕，沉浸在一款高科技的手机游戏中。屏幕发出的炫目蓝紫色光辉映照在他的脸上，增强了科幻氛围。手机界面呈现出悬浮在屏幕之上的全息游戏元素，伴随着炫酷的粒子特效和动态 UI 动画。背景是一家灯光昏暗的现代咖啡馆，朦胧的环境光与手机边缘的霓虹网格图案形成鲜明对比。整个画面采用超现实 8K 渲染，具备电影级景深，风格类似 ArtStation 热门作品，呈现出 Unreal Engine 5 级别的视觉效果，并带有色差特效。”

## ▶ 样例输出：

好的，用户现在想使用名为“即梦”的图像生成工具，并指定我作为提示生成器。他们会在主题前加斜线，我需要用英文生成合适的提示。例如，输入“/用户使用智能手机玩游戏”，我需要生成详细的中文prompt，像例子中那样。

首先，我需要理解用户的需求。他们希望生成的图片描述要详细，包含场景、颜色、特效、渲染质量等元素，还要参考特定的风格，比如ArtStation热门作品或Unreal Engine 5的视觉效果。例子中的提示结构清晰，分多个部分描述，这可能是一个模板。

接下来，我需要确保每次用户输入斜线后的主题时，我都能按照这个模板来构建prompt。需要注意用户可能提供的不同主题，比如自然风光、人物肖像、科幻场景等，每个主题需要不同的元素描述。例如，自然风光可能需要天气、光线、植被细节；人物肖像则需要表情、服装、背景等。

另外，用户提到用英文生成，但例子中的prompt是中文的，所以可能需要确认是否需要中英双语，但根据例子，应该是生成中文的prompt。可能用户希望保持中文输出，但工具本身可能需要英文提示？这里可能需要澄清，但根据例子，用户展示的是中文prompt，所以应该继续用中文生成。

还需要注意用户可能没有提到的细节，比如分辨率（8K）、渲染风格（超现实）、特效（色差、粒子特效）、景深等。这些元素在例子中都有体现，所以需要保持一致性。同时，可能要根据不同主题调整这些元素，比如科幻场景用霓虹色，自然场景用柔和的自然光。

另外，用户可能希望每个提示都包含类似的结构：主体描述、环境细节、颜色和光影效果、特效、渲染质量和风格参考。需要确保每个生成的prompt都覆盖这些方面，保持结构统一。

可能遇到的挑战是如何在有限的用户输入下提取足够的信息来生成详细提示。比如用户输入“/星空下的露营”，需要想象场景中的元素：帐篷、篝火、星空、人物活动等，并添加合适的特效和渲染风格。

需要不断检查生成的prompt是否符合用户提供的例子，确保没有遗漏关键元素。同时，避免重复使用相同的词汇，保持描述的多样性和生动性。例如，使用不同的形容词来描绘光影效果，如“炫目”、“柔和”、“朦胧”等。

最后，要确保生成的prompt能够有效指导图像生成工具产出符合预期的图片，因此需要准确传达视觉元素和风格要求。可能需要测试几个例子来验证prompt的有效性，并根据反馈调整生成策略。

好的，我已准备好担任“即梦”图像生成器的提示生成器！请在您想生成的主题前添加斜线“/”，我会根据您的描述自动生成包含以下要素的详细中文提示：



## 提示词：

DeepSeek思考完成后，输入你想要的画面即可。

例：

/用户使用智能手机玩游戏

之后把生成的提示词输入到即梦，调整到合适的画面比例，按“生成”。

## 样例输出：

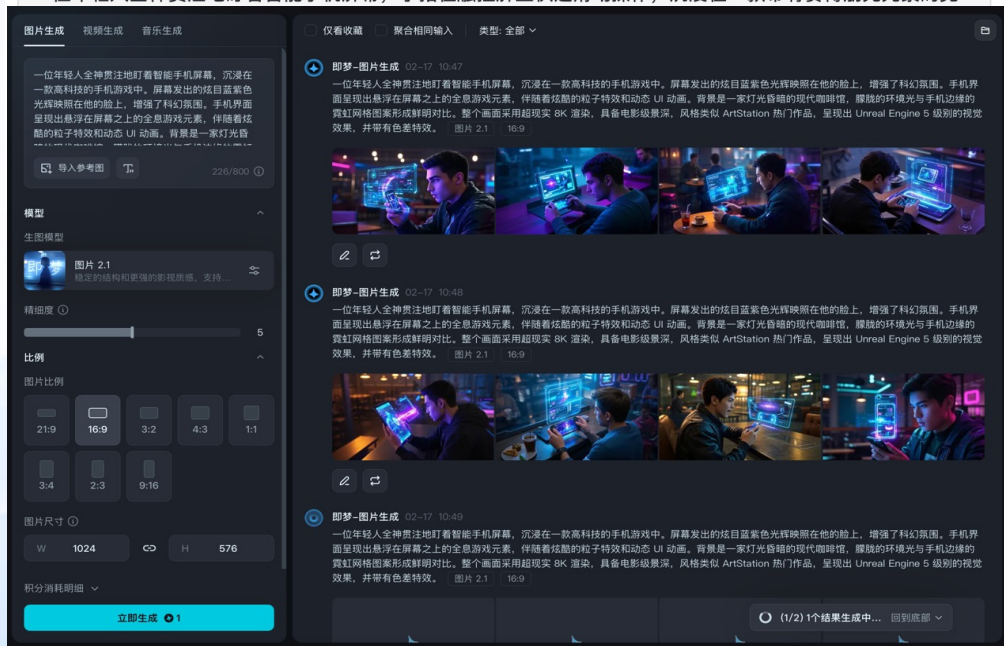
好的，用户现在输入了“/用户使用智能手机玩游戏”，需要我生成相应的中文提示。首先，我需要回忆之前的例子，确保结构一致。用户之前的例子中包含了主体、环境、光影、特效、渲染质量和风格参考等要素。

接下来，我要确定主体是用户使用手机玩游戏的场景。需要描述用户的年龄、专注状态，手机屏幕的光效。然后考虑环境，比如咖啡馆、客厅或者户外，选择现代咖啡馆可能比较合适，灯光昏暗可以突出手机的光效。

光影方面，手机屏幕的蓝紫色色光映在脸上，增强科幻感。特效部分要加入全息元素、粒子特效和动态UI，可能还要提到霓虹网格图案。渲染质量用8K超现实，电影级景深，风格参考ArtStation和UE5，加上色差特效。

检查是否有遗漏的要素，比如动态元素，比如粒子流动、UI动画。确保语言流畅，符合中文表达习惯，不使用专业术语，但保持描述详细。最后，确认整体结构是否与之前的例子一致，避免格式错误。

一位年轻人全神贯注地盯着智能手机屏幕，手指在触控屏上快速滑动操作，沉浸在一款带有赛博朋克元素的竞





## ▶ 提示词：

选择一张合适的图片，下载到本地，并将界面切换到“视频生成”并上传图片，在提示词栏说明视频中的人物在做什么，然后点击生成即可。

也可以直接在图片上选择“生成视频”，即可将生成完的图片直接转换成视频。

## ▶ 样例输出：



## ▶ 提示词：

需要拍摄和生成的素材全部准备好以后，接下来可以根据之前生成的制作思路，用DeepSeek生成一份台词脚本。

如果完全不想出镜介绍或录音，可以利用剪映的朗读功能，一键生成旁白；同样的，利用剪映的一键剪辑功能，输入文案，自动配素材和音乐，AI语音合成多语种配音，智能转场，横竖屏自适应，字幕也自动生成，快速将所有素材一键成片。

## ▶ 样例输出：

好的，我需要帮用户生成一个6分钟的科技评论视频的台词脚本

以下是6分钟视频的逐镜台词脚本设计（括号内为剪辑提示与表演说明）：

### [开场 00:00-00:30]

（镜头从手机前置镜头拉开，你拿着三部手机对镜头）  
"家人们谁懂啊！现在买手机就像选男朋友——"  
（快速举起左手手机）"要游戏打得猛！"  
（切右手手机）"要拍照修得狠！"  
（突然把中间手机贴脸）"还会会！人！工！智！能！"  
（变装转场特效+音效"咻~"）  
（严肃脸推眼镜）"但说真的，2025年手机战场已经卷成麻花了..."  
（背景弹出爆炸动效的「性能/影像/AI」关键词）

### [性能战场 00:30-2:00]

（坐在电竞椅上手柄）"先说游戏党最关心的——"  
（突然举起散热背夹）"现在没这玩意儿你敢玩《原神》2.0？"  
（切入某博主实测视频，标注「骁龙8Gen4 VS 天玑9400」）  
"看这帧率曲线，高通这次终于不！摆！烂！了！"  
（切Flourish天梯图动画）"但发哥的能效比...（压低声音）某绿色厂牌狂喜"

# 编程辅助

## DeepSeek+ HBuilder高效编写前端效果

HBuilder是一款集成开发环境（IDE），主要用于前端开发，支持Web、移动端、以及桌面应用的开发。它提供了丰富的功能，包括代码自动补全、智能提示、预览调试、多种插件支持等，

通过DeepSeek和HBuilder的结合，用户可以高效编写js效果并生成相应的HTML页面，进而在HBuilder中展示。

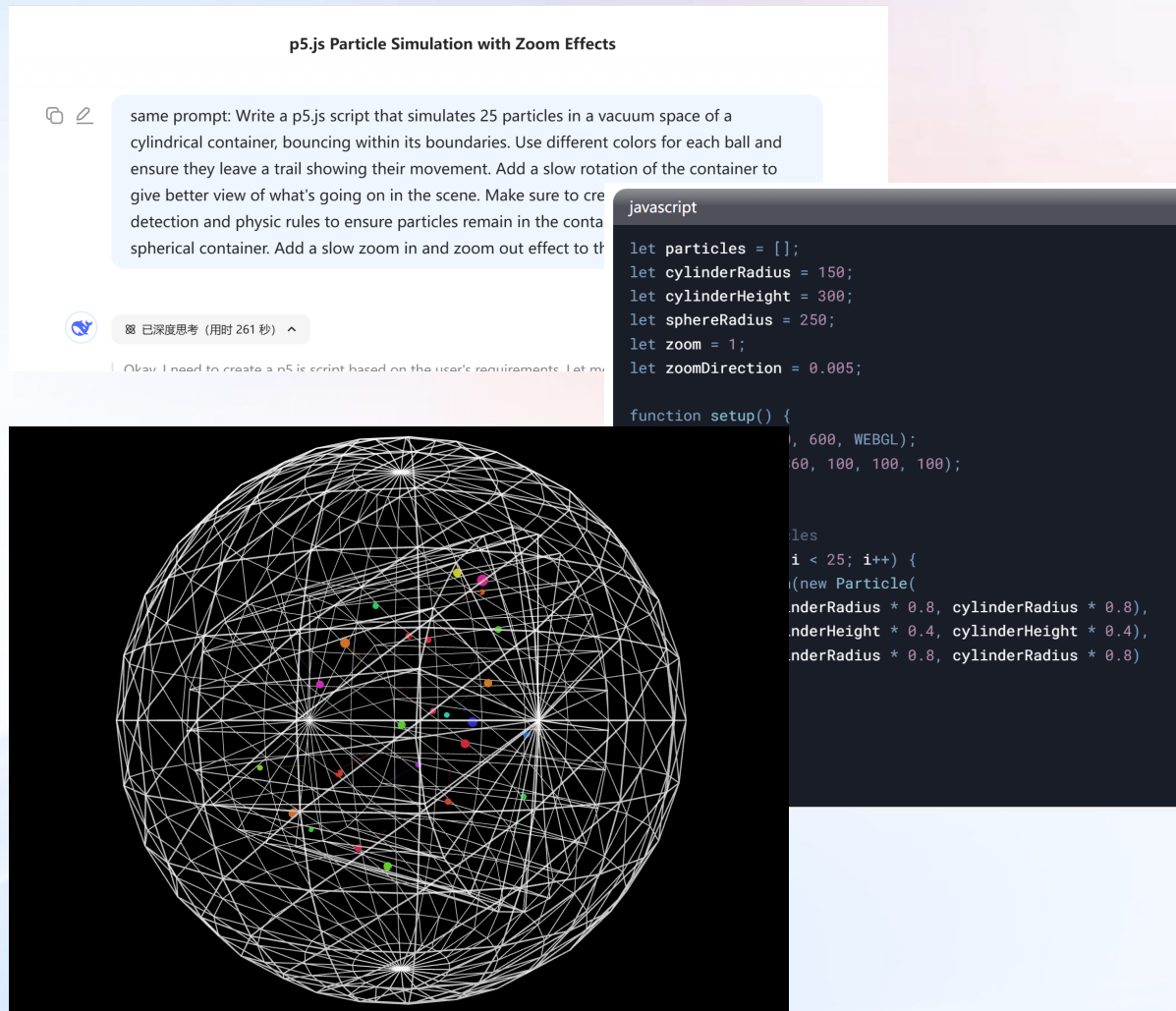
## 提示词：

1.把这段提示词发送给DeepSeek：

same prompt: Write a p5.js script that simulates 25 particles in a vacuum space of a cylindrical container, bouncing within its boundaries. Use different colors for each ball and ensure they leave a trail showing their movement. Add a slow rotation of the container to give better view of what's going on in the scene. Make sure to create proper collision detection and physic rules to ensure particles remain in the container. Add an external spherical container. Add a slow zoom in and zoom out effect to the whole scene.

2.打开Hbuilder,粘贴DeepSeek输出的html代码后,保存并运行.

## 样例输出：





## DeepSeek+Cline 实现对话式代码开发

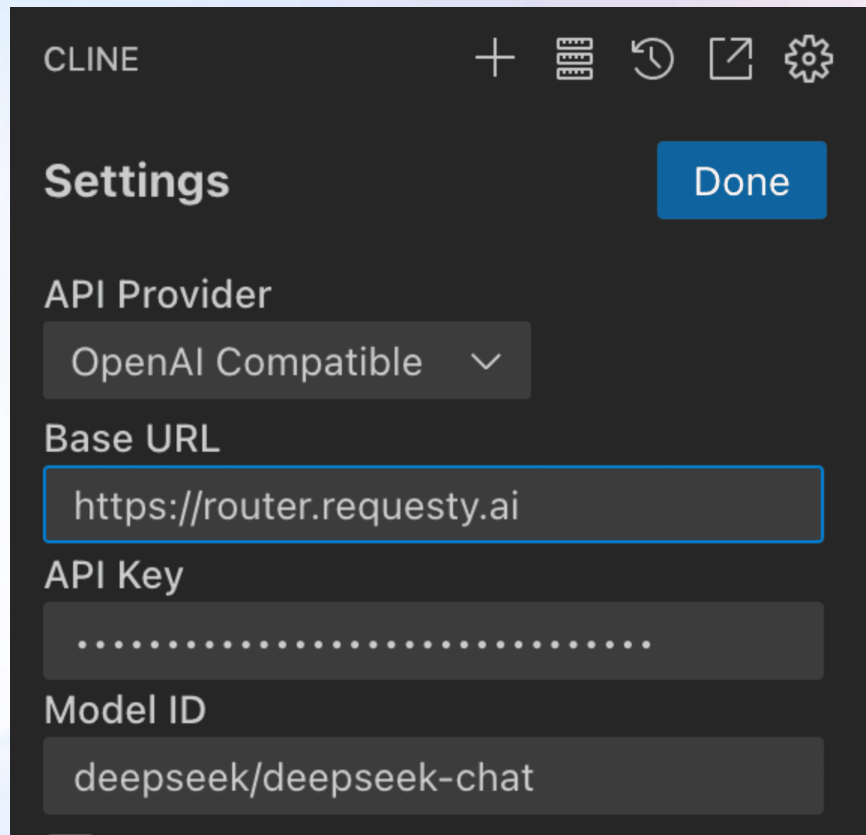
Cline是一款基于人工智能的自动化文档生成工具，支持多种文档类型的生成，包括技术文档、商业报告、项目提案等，用户只需提供基本的需求或数据，Cline便能够自动生成结构清晰、内容精炼的文档草稿，并且支持自动修改和优化，提高文档创作的效率和质量。

通过与DeepSeek的对接，用户可以在开发过程中通过自然语言与AI进行实时对话。用户只需描述他们的编码需求，DeepSeek会生成相应的代码提示或解决方案，Cline则在VSCode中自动插入并展示这些代码。这一对话式的开发方式减少了繁琐的查找和编写过程。

## ▶ 提示词：

Cline是VsCode上的一个插件，直接在VsCode插件库中搜索Cline，就能直接下载安装。  
安装完成后，选择DeepSeek的API，填入API Key即可。  
配置好key等参数后，你会看到一个对话框，在这里使用DeepSeek进行对话式代码开发。  
例：  
用HTML开发一个简易的贪吃蛇网页。  
输入进去后，它就会为你生成一个可用的HTML页面。

## ▶ 样例输出：



CLINE

Settings Done

API Provider  
OpenAI Compatible

Base URL  
https://router.requesty.ai

API Key  
.....

Model ID  
deepseek/deepseek-chat

# DeepSeek工具推荐篇

将 DeepSeek API 集成到流行软件中，涵盖了编程开发、文本分析、翻译、内容创作等类别

# DeepSeek 应用程序

|                                    |                                                                                                                                                         |
|------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------|
| DeepChat                           | DeepChat 是一款完全免费的桌面端智能助手，内置强大的 DeepSeek 大模型，支持多轮对话、联网搜索、文件上传、知识库等多种功能。                                                                                  |
| Wechat-Bot                         | 基于 wechaty 实现的微信机器人，结合了 DeepSeek 和其他 Ai 服务。                                                                                                             |
| Quentalogic                        | QuantaLogic 是一个 ReAct（推理和行动）框架，用于构建高级 AI 代理。                                                                                                            |
| Chatbox                            | 一个支持多种流行LLM模型的桌面客户端，可在 Windows、Mac 和 Linux 上使用                                                                                                          |
| ChatGPT-Next-Web                   | 一键获取跨平台ChatGPT网页用户界面，支持流行的LLM                                                                                                                           |
| Coco AI                            | Coco AI 是一个完全开源、跨平台的统一搜索与效率工具，能够连接并搜索多种数据源，包括应用程序、文件、谷歌网盘、Notion、语雀、Hugo 等本地与云端数据。通过接入 DeepSeek 等大模型，Coco AI 实现了智能化的个人知识库管理，注重隐私，支持私有部署，帮助用户快速、智能地访问信息。 |
| 留白记事                               | 留白让你直接在微信上使用 DeepSeek 管理你的笔记、任务、日程和待办清单！                                                                                                                |
| Pal - AI Chat Client (iOS, iPadOS) | 一款可以在 iPhone 或 iPad 上使用的 AI 助手                                                                                                                          |

# DeepSeek 应用程序

|                       |                                                                                                            |
|-----------------------|------------------------------------------------------------------------------------------------------------|
| LibreChat             | LibreChat 是一个可定制的开源应用程序，无缝集成了 DeepSeek，以增强人工智能交互体验                                                         |
| PapersGPT             | PapersGPT是一款集成了DeepSeek及其他多种AI模型的辅助论文阅读的Zotero插件。                                                          |
| RSS翻译器                | 开源、简洁、可自部署的RSS翻译器                                                                                          |
| Enconvo               | Enconvo是AI时代的启动器,是所有AI功能的入口,也是一位体贴的智能助理。                                                                   |
| Cherry Studio         | 一款为创造者而生的桌面版 AI 助手                                                                                         |
| ToMemo (iOS, ipaOS)   | 一款短语合集 + 剪切板历史 + 键盘输出的iOS应用，集成了AI大模型，可以在键盘中快速输出使用。                                                         |
| Video Subtitle Master | 批量为视频生成字幕，并可将字幕翻译成其它语言。这是一个客户端工具, 跨平台支持 mac 和 windows 系统, 支持百度，火山，deeplx, openai, deepseek, ollama 等多个翻译服务 |
| Easydict              | Easydict 是一个简洁易用的词典翻译 macOS App，能够轻松优雅地查找单词或翻译文本，支持调用大语言模型 API 翻译。                                         |



# DeepSeek 应用程序

|             |                                                                                         |
|-------------|-----------------------------------------------------------------------------------------|
| Raycast     | Raycast 是一款 macOS 生产力工具，它允许你用几个按键来控制你的工具。它支持各种扩展，包括 DeepSeek AI。                        |
| Zotero      | Zotero 是一款免费且易于使用的文献管理工具，旨在帮助您收集、整理、注释、引用和分享研究成果。                                       |
| 思源笔记        | 思源笔记是一款隐私优先的个人知识管理系统，支持完全离线使用，并提供端到端加密的数据同步功能。                                          |
| go-stock    | go-stock 是一个由 Wails 使用 NativeUI 构建并由 LLM 提供支持的股票数据查看分析器。                                |
| Wordware    | Wordware 这是一个工具包，使任何人都可以仅通过自然语言构建、迭代和部署他们的AI堆栈                                          |
| Dify        | Dify 是一个支持 DeepSeek 模型的 LLM 应用开发平台，可用于创建 AI 助手、工作流、文本生成器等应用。                            |
| LiberSonora | LiberSonora，寓意"自由的声音"，是一个 AI 赋能的、强大的、开源有声书工具集，包含智能字幕提取、AI标题生成、多语言翻译等功能，支持 GPU 加速、批量离线处理 |
| Bob         | Bob 是一款 macOS 平台的翻译和 OCR 软件，您可以在任何应用程序中使用 Bob 进行翻译和 OCR，即用即走！                           |

# DeepSeek 应用程序

|                   |                                                                                                                             |
|-------------------|-----------------------------------------------------------------------------------------------------------------------------|
| STranslate        | STranslate（Windows）是 WPF 开发的一款即用即走的翻译、OCR工具                                                                                 |
| GPT AI Flow       | 工程师为效率狂人（他们自己）打造的终极生产力武器: GPT AI Flow<br>`Shift+Alt+空格` 唤醒桌面智能中枢<br>本地加密存储<br>自定义指令引擎<br>按需调用拒绝订阅捆绑                         |
| Alpha派            | AI投研助理/AI驱动的新一代金融信息入口。代理投资者听会/记纪要，金融投资信息的搜索问答/定量分析等投资研究工作。                                                                  |
| argo              | 本地下载并运行Huggingface及Ollama模型，支持RAG、LLM API、工具接入等，支持Mac/Windows/Linux。                                                        |
| PeterCat          | 我们提供对话式答疑 Agent 配置系统、自托管部署方案和便捷的一体化应用 SDK，让您能够为自己的 GitHub 仓库一键创建智能答疑机器人，并快速集成到各类官网或项目中，为社区提供更高效的技术支持生态。                     |
| FastGPT           | FastGPT 基于 LLM 大模型的开源 AI 知识库构建平台，支持 DeepSeek、OpenAI 等多种模型。我们提供了开箱即用的数据处理、模型调用、RAG 检索、可视化 AI 工作流编排等能力，帮助您轻松构建复杂的 AI 应用。      |
| Chatgpt-on-Wechat | Chatgpt-on-Wechat（CoW）项目是一个灵活的聊天机器人框架，支持将DeepSeek、OpenAI、Claude、Qwen等多种 LLM 一键接入到微信公众号、企业微信、飞书、钉钉、网站等常用平台或办公软件，并支持丰富的自定义插件。 |

# DeepSeek 应用程序

|          |                                                                                                                                     |
|----------|-------------------------------------------------------------------------------------------------------------------------------------|
| 如知AI笔记   | 如知AI笔记是一款智能化的AI知识管理工具，致力于为用户提供一站式的知识管理和应用服务，包括AI搜索探索、AI结果转笔记、笔记管理与整理、知识演示与分享等。集成了DeepSeek深度思考模型，提供更稳定、更高质量的输出。                      |
| Athena   | 世界上首个具有先进认知架构和类人推理能力的自主通用人工智能，旨在解决复杂的现实世界挑战。                                                                                        |
| TigerGPT | TigerGPT 是老虎集团开发的，业内首个基于 OpenAI 的金融 AI 投资助理。TigerGPT 旨在为投资者提供智能化的投资决策支持。2025年2月18日，TigerGPT 正式接入 DeepSeek-R1 模型，为用户提供支持深度推理的在线问答服务。 |
| HIX.AI   | 免费试用 DeepSeek，在 HIX.AI 上享受无限量的 AI 聊天。使用 DeepSeek R1 进行 AI 聊天、写作、编码等。立即体验下一代 AI 聊天！                                                  |

# DeepSeek Solana 框架

|                  |                                                                                  |
|------------------|----------------------------------------------------------------------------------|
| Solana Agent Kit | 一个用于连接 AI 智能体到 Solana 协议的开源工具包。现在，任何使用 DeepSeek LLM 的智能体都可以自主执行 60+ 种 Solana 操作。 |
|------------------|----------------------------------------------------------------------------------|

# DeepSeek AI Agent 框架

|               |                                                                                                                                                      |
|---------------|------------------------------------------------------------------------------------------------------------------------------------------------------|
| Anda          | 一个专为 AI 智能体开发设计的 Rust 语言框架，致力于构建高度可组合、自主运行且具备永久记忆能力的 AI 智能体网络。                                                                                       |
| YoMo          | Stateful Serverless LLM Function Calling Framework with Strongly-typed Language Support                                                              |
| Alice         | 一个基于 ICP 的自主 AI 代理，利用 DeepSeek 等大型语言模型进行链上决策。Alice 结合实时数据分析和独特的个性，管理代币、挖掘 BOB 并参与生态系统治理。                                                             |
| ATTPs         | 一个用于Agent之间可信通信的基础协议框架，基于DeepSeek的Agent，可以接入ATTPs的SDK，获得注册Agent，发送可验证数据，获取可验证数据等功能，从而与其他平台的Agent进行可信通信。                                              |
| translate.js  | 面向前端开发者的 AI i18n, 两行js实现html全自动翻译，几十语种一键切换，无需改动页面、无语言配置文件、支持几十个微调扩展指令、对SEO友好。并且开放标准文本翻译API接口                                                         |
| agentUniverse | agentUniverse 是一个面向复杂业务场景设计的多智能体协作框架。其提供了快速易用的大模型智能体应用搭建能力，并着重于提供智能体协同调度、自主决策与动态反馈等机制，其源自蚂蚁集团在金融领域的真实业务实践沉淀。agentUniverse于2024年6月全面接入支持deepseek系列模型。 |

# DeepSeek RAG 框架

|              |                                                                                                                                                    |
|--------------|----------------------------------------------------------------------------------------------------------------------------------------------------|
| RAGFlow      | 一款基于深度文档理解构建的开源 RAG（Retrieval-Augmented Generation）引擎。RAGFlow 可以为各种规模的企业及个人提供一套精简的 RAG 工作流程，结合大语言模型（LLM）针对用户各类不同的复杂格式数据提供可靠的问答以及有理有据的引用。           |
| Autoflow     | AutoFlow 是一个开源的基于 GraphRAG 的知识库工具，构建于 TiDB Vector、LlamaIndex 和 DSPy 之上。提供类 Perplexity 的搜索页面，并可以嵌入简单的 JavaScript 代码片段，轻松将 Autoflow 的对话式搜索窗口集成到您的网站。 |
| DeepSearcher | DeepSearcher 结合强大的 LLM（DeepSeek、OpenAI 等）和向量数据库（Milvus 等），根据私有数据进行搜索、评估和推理，提供高度准确的答案和全面的报告。                                                        |
| KAG          | KAG 是基于 OpenSPG 引擎和大型语言模型的逻辑推理问答框架，用于构建垂直领域知识库的逻辑推理问答解决方案。KAG 可以有效克服传统 RAG 向量相似度计算的歧义性和 OpenIE 引入的 GraphRAG 的噪声问题。KAG 支持逻辑推理、多跳事实问答等。              |

# DeepSeek FHE (全同态加密) frameworks

|                   |                                                                                                                                                                         |
|-------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Mind FHE Rust SDK | 一个开源 SDK，可使用**全同态加密（FHE）**对 AI 进行加密，实现代理共识。FHE被誉为密码学的圣杯，能够在无需解密的情况下直接对加密数据进行计算。借助FHE，代理在使用Deepseek时可以保护隐私，同时确保模型的完整性和计算结果的一致性，无需暴露任何数据。该SDK的源代码采用纯Rust实现，并可在crates.io获取 |
|-------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------|



# DeepSeek 浏览器插件

|             |                                                                       |
|-------------|-----------------------------------------------------------------------|
| 沉浸式翻译       | 一款双语对照网页翻译插件，简洁，高效                                                    |
| 沉浸式导读       | NO Sidebar!!! 沉浸式的 AI 网页摘要，提问...                                      |
| ChatGPT Box | 将 LLM 作为私人助手，整合到你的浏览器中                                                |
| 划词翻译        | 整合了多家翻译 API 以及 LLM API 的浏览器翻译插件                                       |
| 欧路翻译        | 提供鼠标划词搜索、逐段对照翻译、PDF文献翻译功能。可以使用支持 DeepSeek AI、Bing、GPT、Google 等多种翻译引擎。 |
| 流畅阅读        | 一款革新性的浏览器开源翻译插件，让所有人都能够拥有基于母语般的阅读体验                                   |
| 馆长          | 知识库AI问答助手 - 让AI帮助你整理与分析知识                                             |
| RssFlow     | 一款智能的RSS阅读器浏览器扩展，具有AI驱动的RSS摘要和多维度订阅视图功能。支持配置DeepSeek模型以增强内容理解能力。      |

# DeepSeek 浏览器插件

|                  |                                                                                               |
|------------------|-----------------------------------------------------------------------------------------------|
| Typral           | 超快的AI写作助手 - 让AI帮你快速优化日报，文章，文本等等...                                                            |
| Trancy           | 沉浸双语对照翻译、视频双语字幕、划句/划词翻译插件                                                                     |
| Anything Copilot | Anything Copilot 是一款可以让你在侧边栏无缝使用任意主流AI工具的浏览器插件                                                |
| DeepChat         | 整合了多家翻译 API 以及 LLM API 的浏览器翻译插件                                                               |
| 欧路翻译             | 一款Chrome扩展程序，允许用户在任何网站上通过打开侧边栏与DeepSeek聊天。此外，它还在任何网站上选中的文本下方提供一个浮动菜单，使用户能够生成文本摘要、检查语法问题和翻译内容。 |

# DeepSeek 即时通讯插件

|                                        |                                                                |
|----------------------------------------|----------------------------------------------------------------|
| 茴香豆<br>(个人微信/飞书)                       | 一个集成到个人微信群/飞书群的领域知识助手，专注解答问题不闲聊                                |
| LangBot<br>(QQ, 企微, 飞书)                | 大模型原生即时通信机器人平台，适配 QQ / QQ频道 / 飞书 / OneBot / 企业微信（wecom）等多种消息平台 |
| NoneBot<br>(QQ, 飞书, Discord, TG, etc.) | 基于 NoneBot 框架，支持智能对话与深度思考功能。适配 QQ / 飞书 / Discord, TG 等多种消息平台   |

# DeepSeek VS Code 插件

|                  |                                          |
|------------------|------------------------------------------|
| Continue         | 开源 IDE 插件，使用 LLM 做你的编程助手                 |
| Cline            | Cline 是一款能够使用您的 CLI 和编辑器的 AI 助手          |
| AI Commit        | 使用 AI 生成 git commit message 的 VS Code 插件 |
| SeekCode Copilot | vscode智能编码助手，支持配置本地部署的DeepSeek模型         |

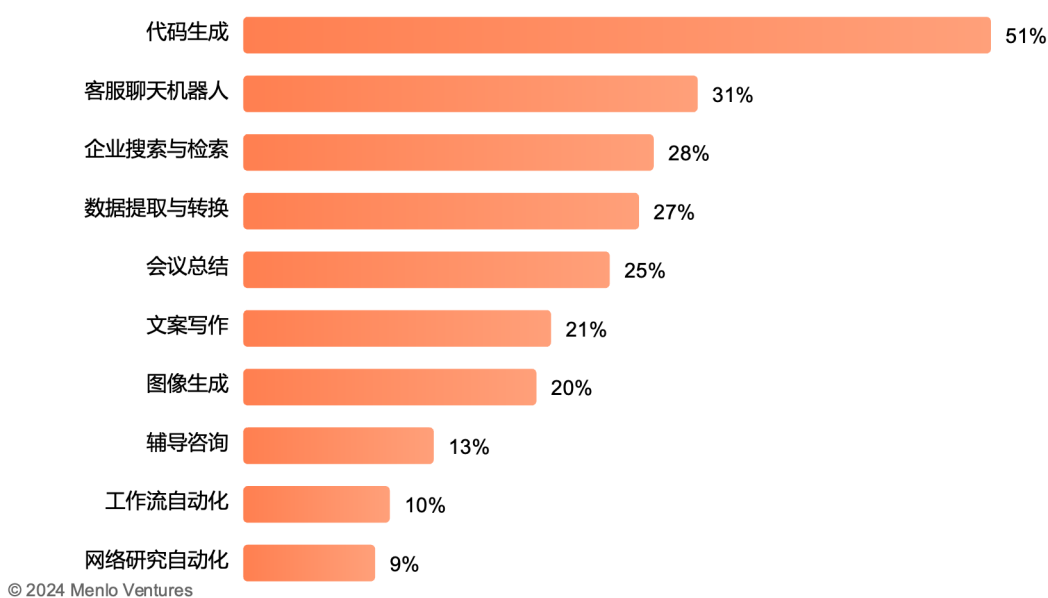
# DeepSeek AI Code编辑器

|          |                                    |
|----------|------------------------------------|
| Cursor   | 基于VS Code进行扩展的AI Code编辑器           |
| WindSurf | 另一个基于VS Code的AI Code编辑器，由Codeium出品 |

# 目前最受欢迎的10个企业AI应用场景

## 来自Menlo投资的600个美国C-Level调研

生成式AI主要应用场景



代码助手遥遥领先，以51%的采用率成为开发者最早拥抱AI的领域。GitHub Copilot快速实现3亿美元营收的增长轨迹证明了这一趋势，而像Codeium和Cursor等新兴工具也在迅速崛起。

客服支持型聊天机器人使用率显著，美国企业采用率达到31%。这些应用通过知识库支持，为内部员工和外部客户提供可靠的全天候服务

企业搜索+检索以及数据提取+转换（分别为28%和27%的采用率）需求强烈，反映了企业释放和利用散布于各部门数据孤岛中隐藏知识的强烈需求。新的解决方案能够连接电子邮件、即时通讯工具和文档存储系统，实现跨系统的统一语义搜索，并提供AI驱动的知识管理。

会议摘要方便易用，在应用场景中排名第五（24%的采用率），通过自动化记录和要点提取节省时间、提升效率。如果用于电子健康记录（EHR）系统，就能提升医疗服务水平。

## ► 趋势判断

- (1) **开源模型将推动 AI 应用生态的加速繁荣，大幅降低传统企业与创业者接入 AI 的成本与门槛。**一方面，对数据安全要求较高的企业可基于开源框架进行本地部署，从而满足合规与安全需求；另一方面，初创公司、中小企业的灵活性和多样化的业务场景也将带来大量云端模型服务需求，为云平台提供模型服务打开新的增长空间。同时，面向特定领域与场景的独立软件供应商（ISV）以及独立 AI 构建者数量也将迅速增多，从而促进整个 AI 应用生态的多元化与可持续发展。
- (2) **AI 技术深入演进，推理模型或将成为主流形态。**推理模型能够通过分解复杂问题、模拟人类思维过程来逐步推导答案，尤其适用于多环节、结构复杂的任务，在精准度与成功率方面表现突出。同时，推理模型通常采用链式思维（Chain-of-Thought）的输出方式，通过多步骤推理进一步增强模型输出的可解释性，使业务决策更透明、更可信。
- (3) **每项业务都需要接入至少一种 AI 模型，每位工作者都需要能跑端侧 AI 的个人电脑，每个公司都需建设能支撑 AI 工作流的 AI 算力平台。**无论何种行业或组织，为了提升效率、增强决策能力和推动创新，都将引入 AI 技术。具体而言，业务通过接入 AI 模型可实现更加智能的运营与服务；个人工作者将借助具备端侧 AI 能力的个人电脑大幅提升创造力与工作效率；而企业层面的 AI Server（或算力平台）则为团队协同与管理提供强大支撑，加速智能创新步伐助力企业保持竞争性。



## ▶ 警惕幻觉

Vectara分别采用HHEM和Google的FACTS方法，对DeepSeek R1 和V3模型分别进行幻觉率测试。两种方法下，**R1幻觉率都显著高于V3模型。**

|                                                      | DeepSeek R1 | DeepSeek V3 |
|------------------------------------------------------|-------------|-------------|
| Vectara’s HHEM 2.1                                   | 14.3%       | 3.9%        |
| Google’s FACTS w/ GPT-4o & Claude-3.5-Sonnet         | 4.37%       | 2.99%       |
| Google’s FACTS w/ GPT-4o & Gemini-1.5-Pro            | 3.09%       | 1.99%       |
| Google’s FACTS w/ Claude-3.5-Sonnet & Gemini-1.5-Pro | 3.89%       | 2.69%       |

### 术语解读-幻觉

在人工智能领域，**幻觉是指由AI生成的以事实形式呈现的虚假或误导性的信息。**比如，AI可能回答“法国的首都是柏林”，而事实上，法国的首都是巴黎。所以在使用AI时，不能一味地信任答案，应当对结果加以判断甄别。

# 主流模型幻觉

主流模型幻觉率差异较大，DeepSeek R1模型幻觉率达Gemini 2.0 flash 001模型的20倍。

| 模型                                | 幻觉率  |
|-----------------------------------|------|
| google/gemini-2.0-flash-001       | 0.7  |
| openai/o3-mini-high-reasoning     | 0.8  |
| openai/GPT-4o                     | 1.5  |
| openai/o1                         | 2.4  |
| Qwen/Qwen2.5-7B-Instruct          | 2.8  |
| deepseek/deepseek-v3              | 3.9  |
| anthropic/Claude-3-5-sonnet       | 4.6  |
| meta-llama/Llama-3.1-70B-Instruct | 5    |
| deepseek/deepseek-r1              | 14.3 |
| anthropic/Claude-3-sonnet         | 16.3 |

2025年2月24日起，DeepSeek启动开源周，连续5天总共开源5个代码库

## 第一天：FlashMLA

FlashMLA是DeepSeek为Hopper GPU设计的高效MLA解码内核，针对可变序列进行了优化。

- 支持BF16
- 分页KV cache（块大小64）

在H800上能达到3000 GB/s的内存速度上限和580 TFLOPS的计算上限、

**注：** Hopper GPU包括H100、H800、H20、H200等。

2025年2月24日起，DeepSeek启动开源周，连续5天总共开源5个代码库

## 第二天：DeepEP

DeepEP是第一个用于混合专家（MoE）模型训练和推理的开源专家并行（EP）通信库。

- 高效和优化的all-to-all通信
- 支持节点内和跨节点的NVLINK和RDMA
- 用于训练和推理预填充的高吞吐内核
- 推理解码阶段的低延迟内核
- 原生FP8数据分发
- 对于计算-通信重叠，灵活的GPU资源控制

2025年2月24日起，DeepSeek启动开源周，连续5天总共开源5个代码库

## 第三天：DeepGEMM

DeepGEMM是一个FP8 GEMM库，支持密集和MoE GEMMs，支撑V3、R1训练和推理。

- 没有繁琐依赖，像教程一样简洁
- 完全Just-in-time（JIT）编译
- 核心逻辑代码仅约300行，但在大多数矩阵尺寸下比调优专家核心更强
- 支持密集布局和两种MoE布局

在Hopper架构GPU上，最高达1350+ FP8 TFLOPS

**注：**截至本手册发布，DeepSeek开源周进行到第三天



# 关于我们

至顶科技  
ZHIDING

至顶AI实验室  
ZHIDING

## 至顶科技 ZHIDING

至顶科技（北京第二十六维信息技术有限公司），国家高新技术企业，中国轻工信息中心智慧生活产业促进中心副主任单位。是国内领先的技术研究与媒体智库平台。至顶科技品牌前身为1997年成立的ZDNet China，是中国最早科技网站

至顶科技参与创立了国内第一个国际级国际性人工智能大会“世界智能大会”，World Intelligence Congress，并与中国互联网协会、中国软件协会，和天津人工智能学会等机构合作，连续发布全球生成式AI产业报告及图谱

自2024年开始，至顶科技联合合作伙伴共同发起国内首个以AI提示工程为主题的提示工程峰会（PEC）



## 至顶AI实验室 ZHIDING

一个专注于探索生成式AI前沿技术及其应用的实验室。我们致力于推动生成式AI在各个领域的创新与突破，挖掘其潜在的应用场景，为企业和个人提供切实可行的解决方案。“未来，AI不释手”

