



# DeepSeek之火 可 以 燎 原

浙江大学医学院

朱朝阳



# 报告提纲

浙江大学人工智能教育  
教学研究中心

从中华文化基因看DeepSeek成功之道

浙江大学人工智能教育  
教学研究中心

DeepSeek的二点争议

浙江大学人工智能教育  
教学研究中心

关于人工智能未来发展之路的思考

浙江大学人工智能教育  
教学研究中心

DeepSeek私有化应用方法

# 求仁得仁，求佛得佛



**唐僧**，古典名著《西游记》中的人物。俗家姓陈，乳名江流儿，法名玄奘，号**三藏**。



# 开放、包容和共享历来是我们中华民族的美德



OpenAI

通义万相AI

小乘法

以“自渡”为核心，追求个人和家人的幸福



DeepSeek

通义万相AI

大乘法

以“利他”为核心，追求普度众生

# 英文的巴别塔魔咒



巴别塔魔咒：《圣经》中记载的一个故事，人类为了建造一座通天高塔（巴别塔）以彰显自己的智慧和力量，但上帝为了阻止这一计划，让人类说不同的语言，导致沟通障碍，最终高塔未能建成。这个事件象征着人类因沟通不畅而产生的分裂和冲突。

存量

17万+

通用词汇，含历史词汇、方言、已经和即将废弃7万

50万+

专业词汇，各领域总和

100万+

总词汇量，含废弃词、科技术语

增量

约1000~2000个

核心高频词

8500~25000个

广义新增词

年

# 英文将成为AI发展的巨大挑战

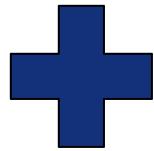


**2050年  
1.7个地球的算力**

# 英文存储的低效和空间浪费



cattle 牛



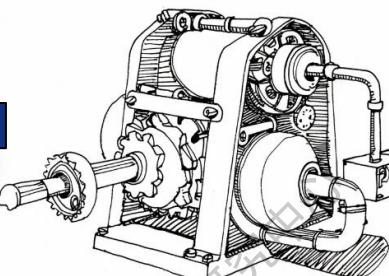
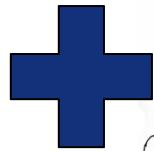
meat 肉



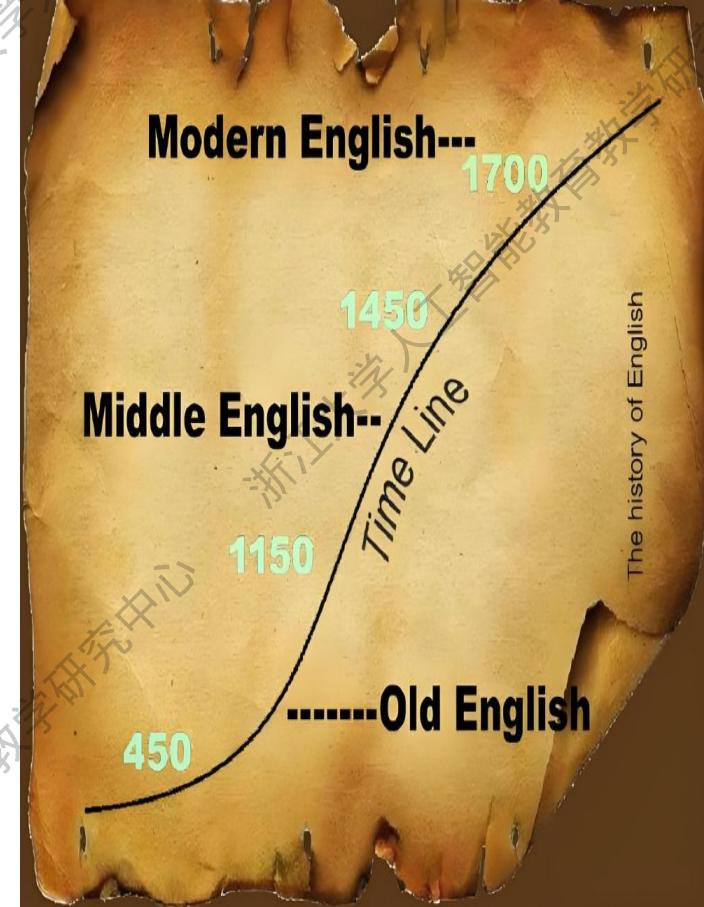
beef 牛肉



fly 飞



airplane 飞机



## 英语发展简史

# 汉语的高效和可传承

汉字起源于上古的象形文字，历经甲骨文、金文、小篆、隶书、楷书、简化字等演变，一路传承，蕴含着中华文明的核心基因密码。从甲骨灼烧的裂纹到手机屏幕的像素点，每一次形态革新都烙印着技术与文化的对话。



## 增量

几乎为 0 / 年

**构词法** 搭积木，知识迁移



## 汉字的演变史

# DeepSeek与ChatGPT训练语料对比

## ChatGPT 4

- **总语料规模** 13万亿token
- **英文语料** **占比约92.6%**, 主要来源于互联网公开文本, 包括维基百科、新闻报道、书籍、学术论文、程序代码等
- **中文语料** **占比仅约0.1%**, **来源单一**, 学术文献不足, 中文互联网内容 (如社交媒体、论坛) 噪声较多, 需复杂清洗

## DeepSeek

- **总语料规模: 14.8 万亿 token**, 涵盖多语言、数学、编程等领域, 其中 87% 代码、**10% 英文、3% 中文**
- **中文语料特点:** 有约 2.3 亿字精校的古典文献; 覆盖网络文学、社交媒体、学术论文等现代文本; 通过 “千字文” 式分词算法增强对成语、典故理解
- **数据处理技术:**
  - **过滤机制:** 采用多层净化技术 (如 “八面受敌法” ), 数据噪音剔除精度达 99.7%。
  - **分词优化:** 结合古典语法与现代汉语的分词算法, 使成语理解准确率提升 38%。
  - **长上下文支持:** 通过两阶段扩展训练, 将上下文窗口扩展至 128K, 增强长文本处理能力。

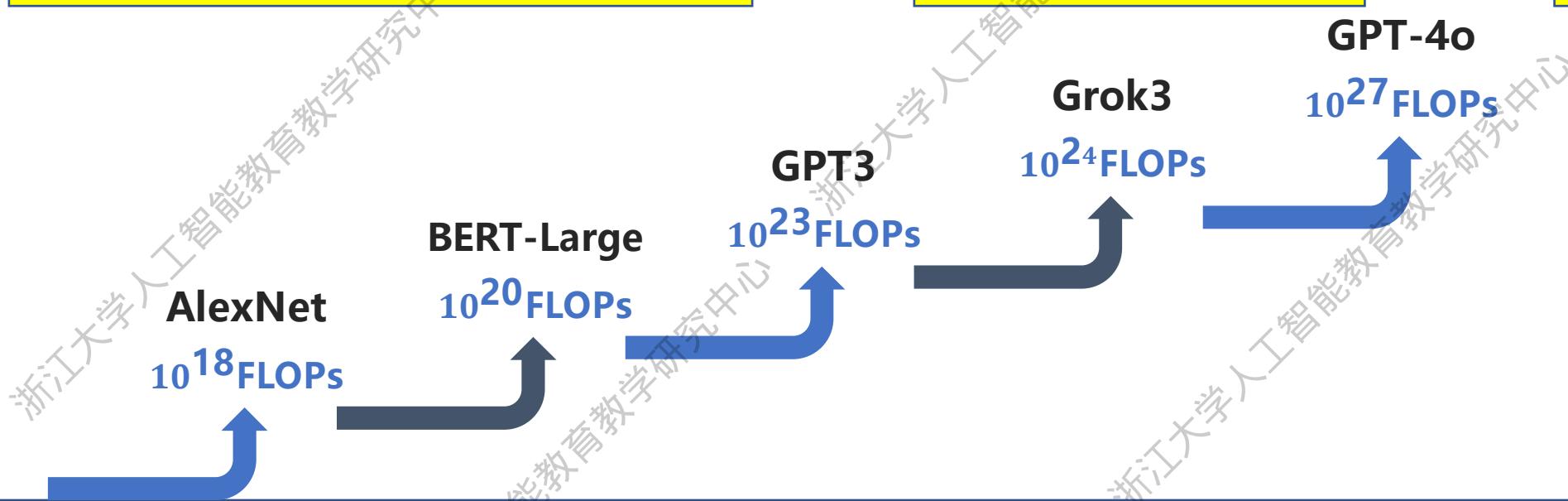
# 用汉字进行模型训练将在这场AI之争中胜出？



**中文:英文**

**DeepSeek 1 : 3.33**  
**ChatGPT 1 : 900**

# 大力出奇迹的范式



# 中华民族的不屈不挠精神

## DeepSeek的技术突围



人工智能  
三要素

算力优化

打破CUDA标准库壁垒，为非英伟达算力芯片提供机会

算法革新

采用思维链提升逻辑性能，实现模型的顿悟时刻

架构创新

知识蒸馏+MOE，证明小规模高质量数据远胜低质量大规模数据

# 报告提纲

关于人工智能未来发展之路的思考

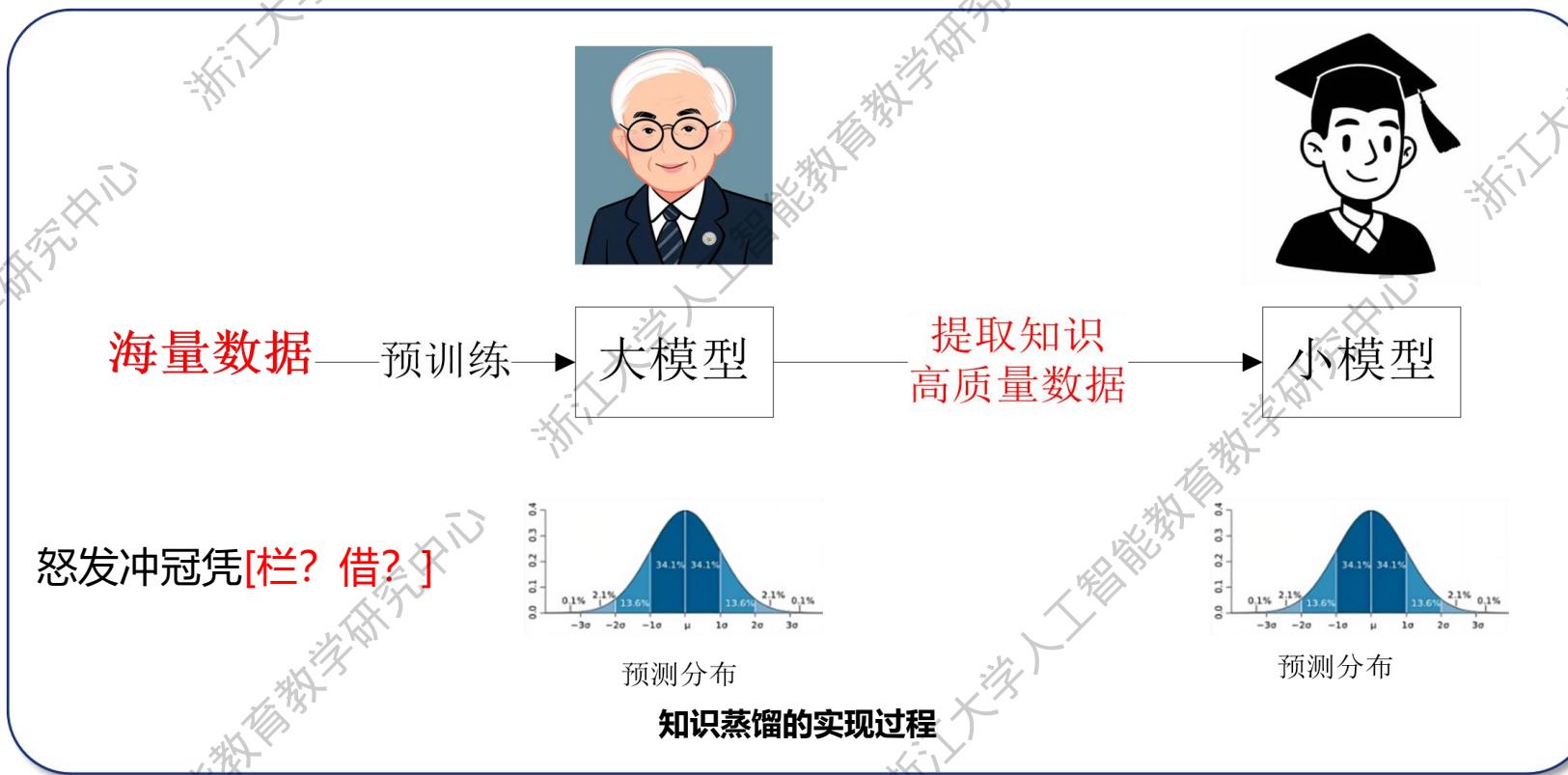
DeepSeek私有化应用方法

DeepSeek的二点争议

从中华文化基因看DeepSeek的成功之道

# 争议一：知识蒸馏是抄袭还是创新

知识蒸馏 (Knowledge Distillation) 是一种模型压缩和加速技术，旨在将大型模型 (教师模型) 的知识迁移到小型模型 (学生模型)，在减少计算资源消耗，缩短推理时间的同时，尽可能达到接近大型模型的性能。



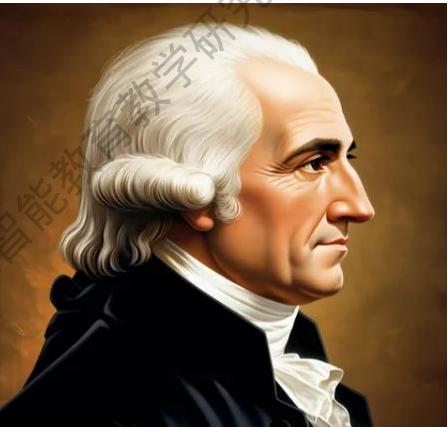
# 争议一：知识蒸馏是抄袭还是创新

## 何为抄袭和剽窃？

牛顿第一定律：如果一个物体所受合外力为0，那么这个物体将保持静止或匀速直线运动

牛顿第二定律：物体的动量随时间的变化率与受力成正比。

牛顿第三定律：相互作用的两个物体之间的作用力和反作用力大小相等，方向相反，作用在同一条直线上。



牛顿三大定律我也知道，所以这些定律是我发现的

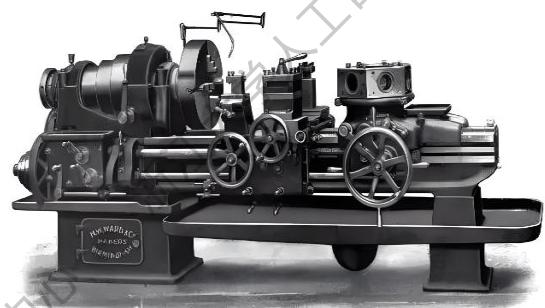


# 争议一：知识蒸馏是抄袭还是创新

何为创新？



基于牛顿定律等  
物理实现航天器



机械式机床

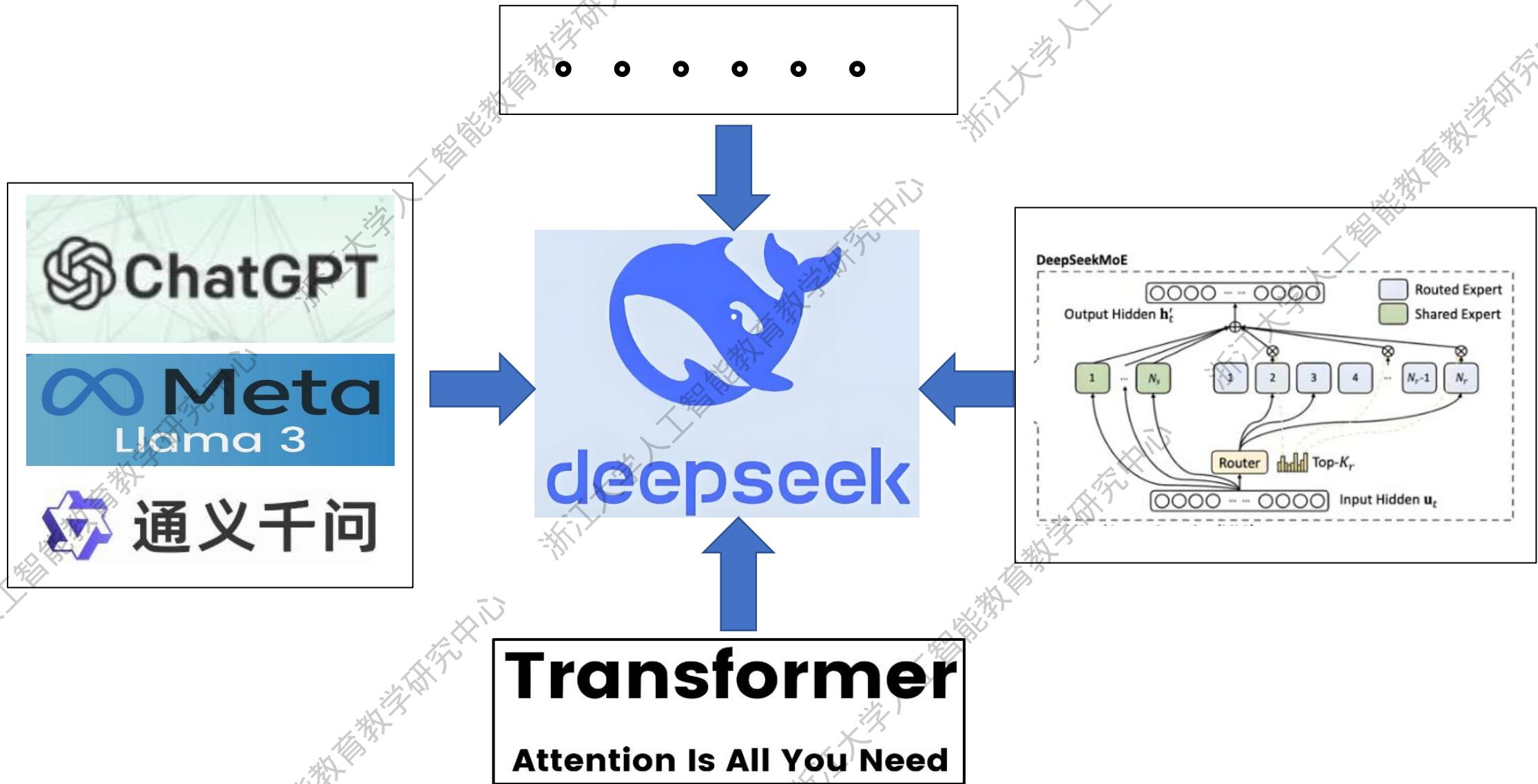


制造零件



数控式机床

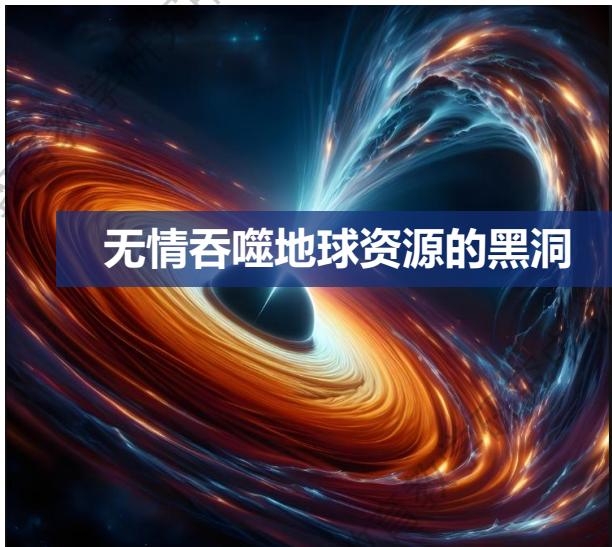
# 争议一：知识蒸馏是抄袭还是创新



# 争议二：人工智能是资源黑洞还是效率革命



**Grok3 20万×H100**



**deepseek 2千×H800**

**智能决策  
提高效率  
低能耗**

# 报告提纲

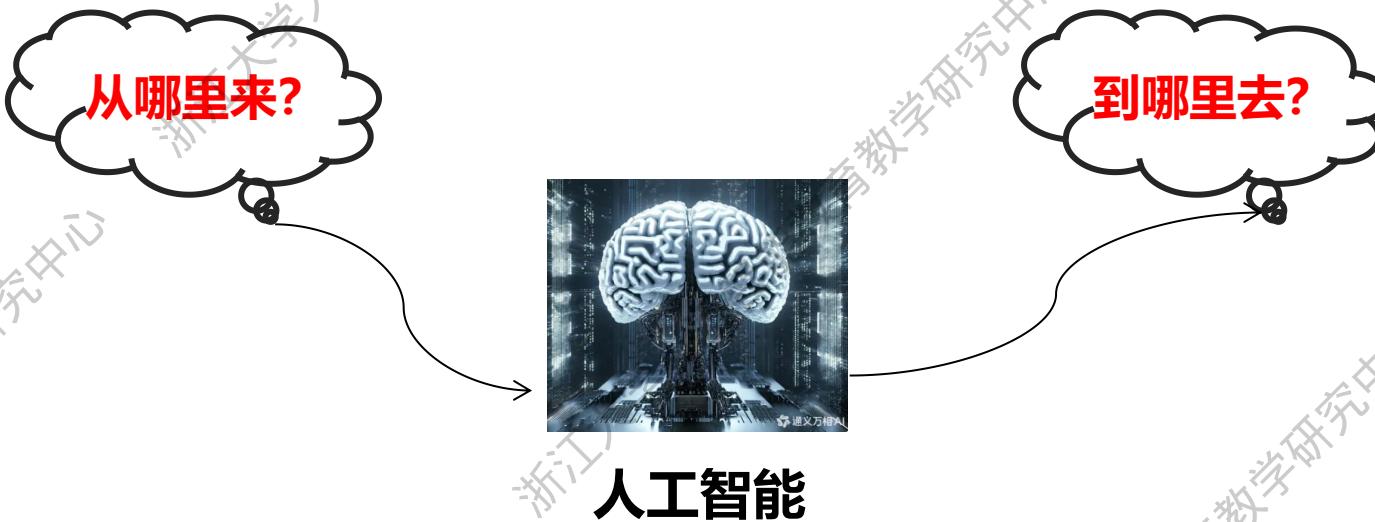
从中华文化基因看DeepSeek的成功之道

DeepSeek的二点争议

关于人工智能未来发展之路的思考

DeepSeek私有化应用方法

# AI的道与术



# AI的道与术

Models 1,472,572

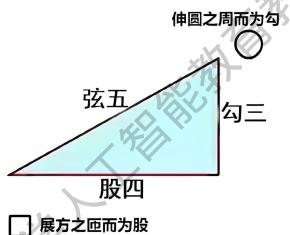
Filter by name

Wan-AI/Wan2.1-T2V-14B  
Text-to-Video • Updated 4 days ago • 142k • 646

microsoft/Phi-4-multimodal-instruct  
Automatic Speech Recognition • Updated 2 days ago • 19.6k • 629

deepseek-ai/DeepSeek-R1  
Text Generation • Updated 6 days ago • 4.36M • 10.6k

堆量不堆质



商高定律：记录在公元前1世纪左右编写的《周髀算经》。商高与周公的对话中提及：“故折矩，勾广三，股修四，经隅五。”

大道至简

《道德经》：有道无术，术尚可求也，有术无道，止于术。

道是规律、道理和本源，而术则是指具体的方法和技巧。这句话告诉我们在解决问题时，只有掌握了事物发展的规律，才能更好地寻求技巧和方法，而只有真正理解和掌握了“道”和“术”，才能更好地应对各种问题，举一反三，不断取得成功

道不足则多术

# AI的道与术



**《道德经·第四十二章》：道生一，一生二，二生三，三生万物。万物负阴而抱阳，冲气以为和。**

**《易经·系辞》：形而上者谓之道，形而下者谓之器。**

何为道？

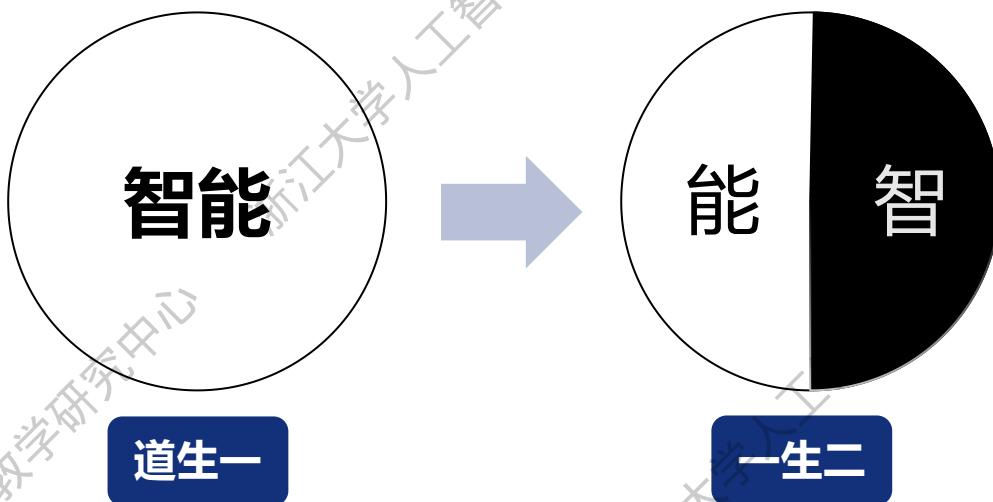


# AI的道与术

智能

道生一

# AI的道与术



智：为阴，即心智，内心之所思所想，即人工智能的算法



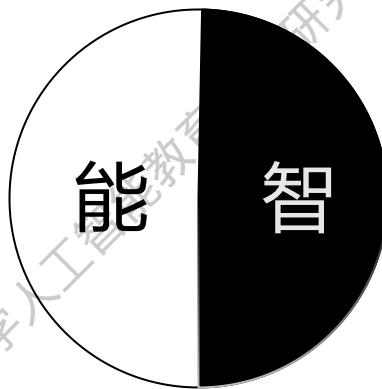
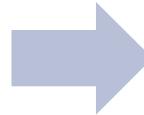
能：为阳，即行动，外在的表现，即人工智能带给我们的能力扩展



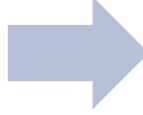
# AI的道与术



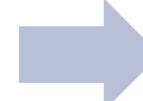
道生一



一生二



二生三



三生万物

智：为阴，即心智，内心之所思所想，即人工智能的算法



能：为阳，即行动，外在的表现，即人工智能带给我们的能力扩展

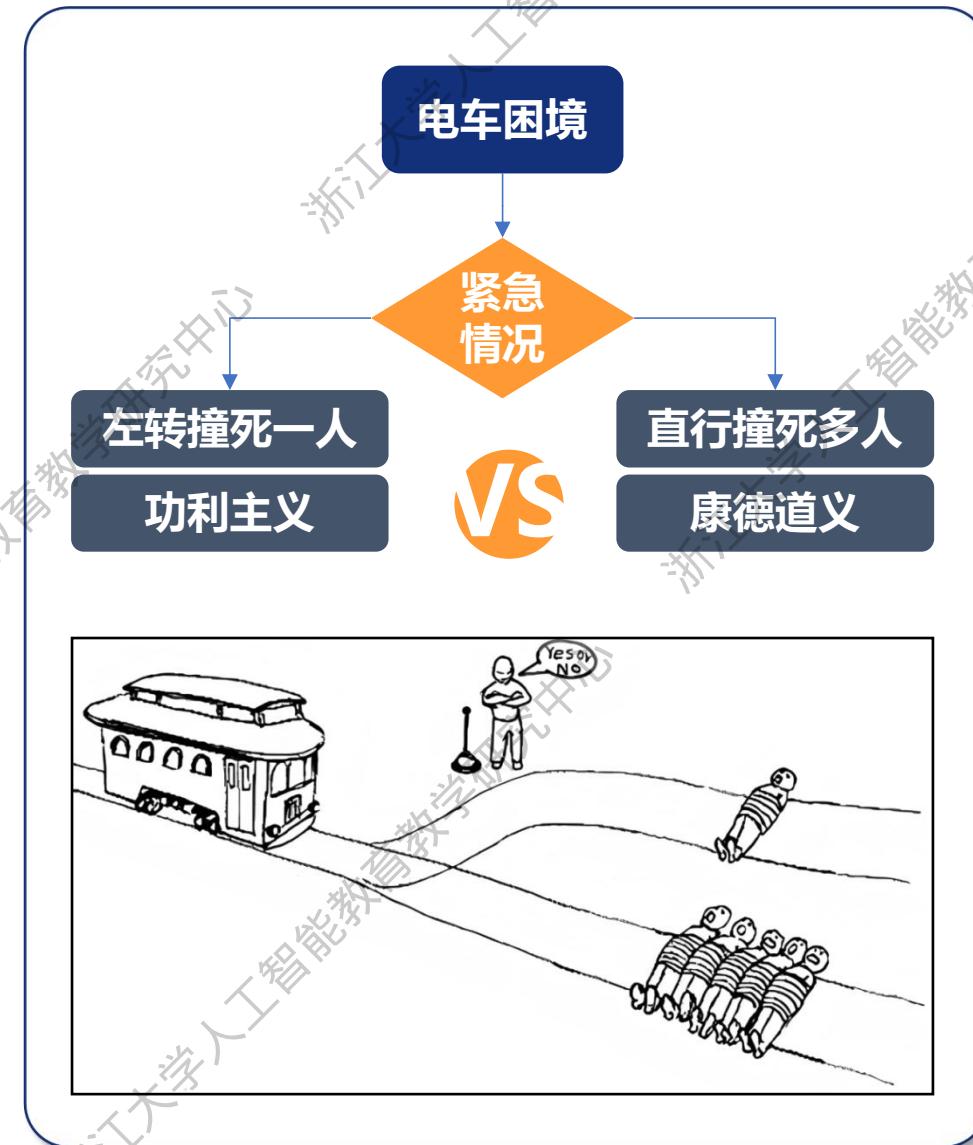
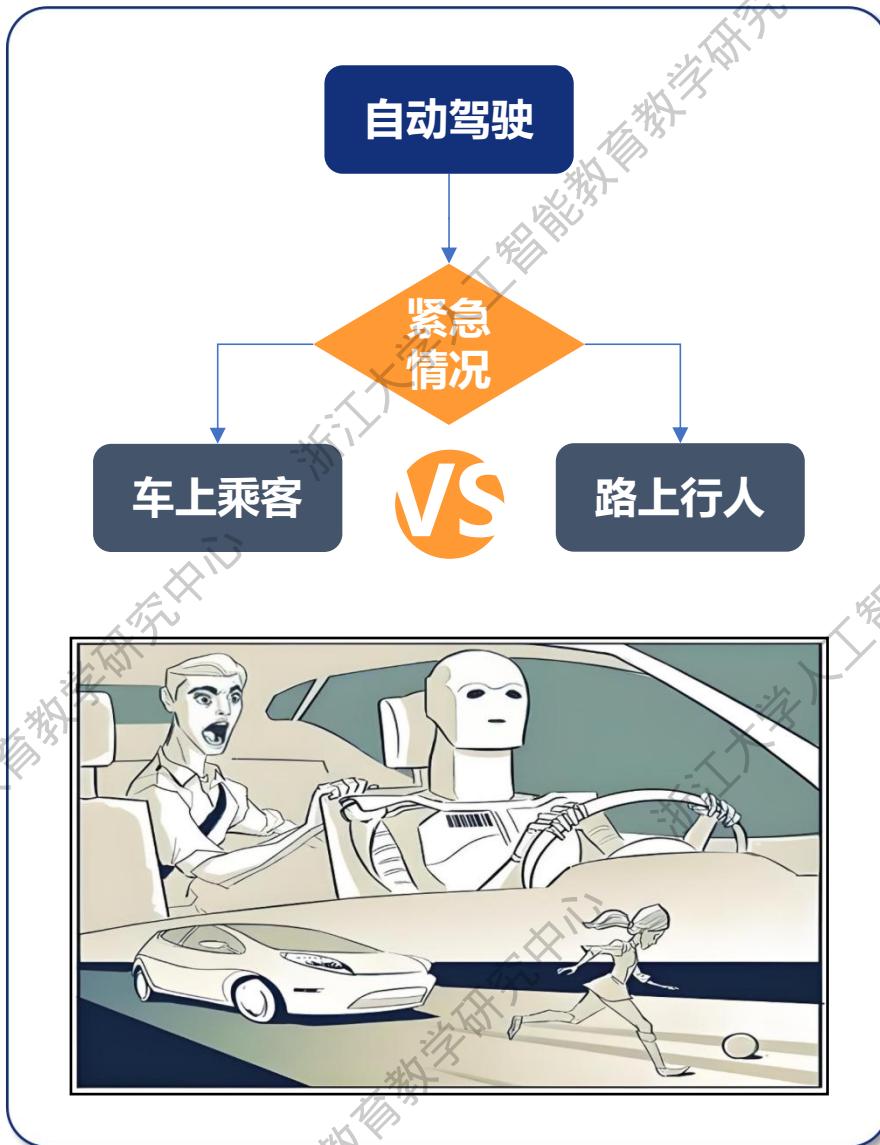


智+能+伦理安全，才是最完美的人工智能，这个伦理安全就是需要解决的三



**道生一，一生二，  
二生三，三生万物。  
万物负阴而抱阳，  
冲气以为和。**

# AI的伦理安全与道德困境



# AI的算法偏见和统治



只有算法的不是人工智能

有了算法也能行动的也还不是真正的智能

**只有能落地应用给人类带来美好生活的人工智能才是真正的人工智能**

# 阴阳气和是人工智能的一条发展之道

机器人三守则，机器人能做到吗？

我们该把人工智能当工具还是人？

当个人隐私与公共安全冲突时，该如何选择？

人工智能是否会造就教育不公？

面对AI的算法统治和偏见怎么办？

不同种族和宗教信仰的人群又如何达成伦理共识？

AIGC幻觉是不是造假？

通过AI换脸进行诈骗该如何杜绝？

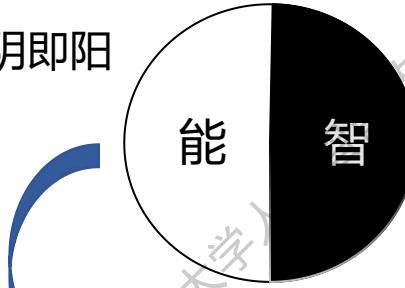
学术诚信如何防范？

.....

问题的根源

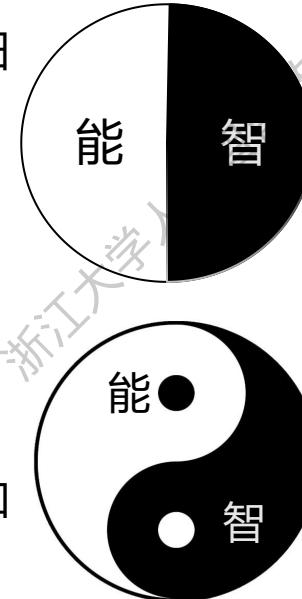
基于0-1的布尔代数建立起来的系统

只有阴阳，非阴即阳



现实世界系统

非阴非阳，阴阳气和



二生三，三生万物  
找到这个三，找到这个气

# 报告提纲

从中华文化基因看DeepSeek的成功之道

DeepSeek的二点争议

关于人工智能未来发展之路的思考

DeepSeek私有化应用方法

# 增强本地DeepSeek能力的二个方案



## 微调 Fine-tuning

- 针对特定任务或领域定制，显著提高任务性能
- 适应特定领域的细微差别、语气或术语
- 依赖足够高质量标注数据
- 适合目标任务明确、数据相对静态的场景

训练实习生



## 检索增强生成 Retrieval-Augmented Generation, RAG

- 利用外部知识库增强答案
- 提供支持性的证据
- 能够动态检索相关知识
- 适合数据频繁更新的场景
- 成本效益高更可信、更可解释、时效性和准确性高

提供项目资料



## 提示词工程 Prompt Engineering

- 通过精心设计输入提示，引导LLM生成期望的高质量输出
- 无需重新训练模型，不需要标注数据
- 一些提示工程框架可以显著增强LLM的推理能力
- 灵活性、使用简单

任务布置清晰

# 本地部署检索增强-RAG



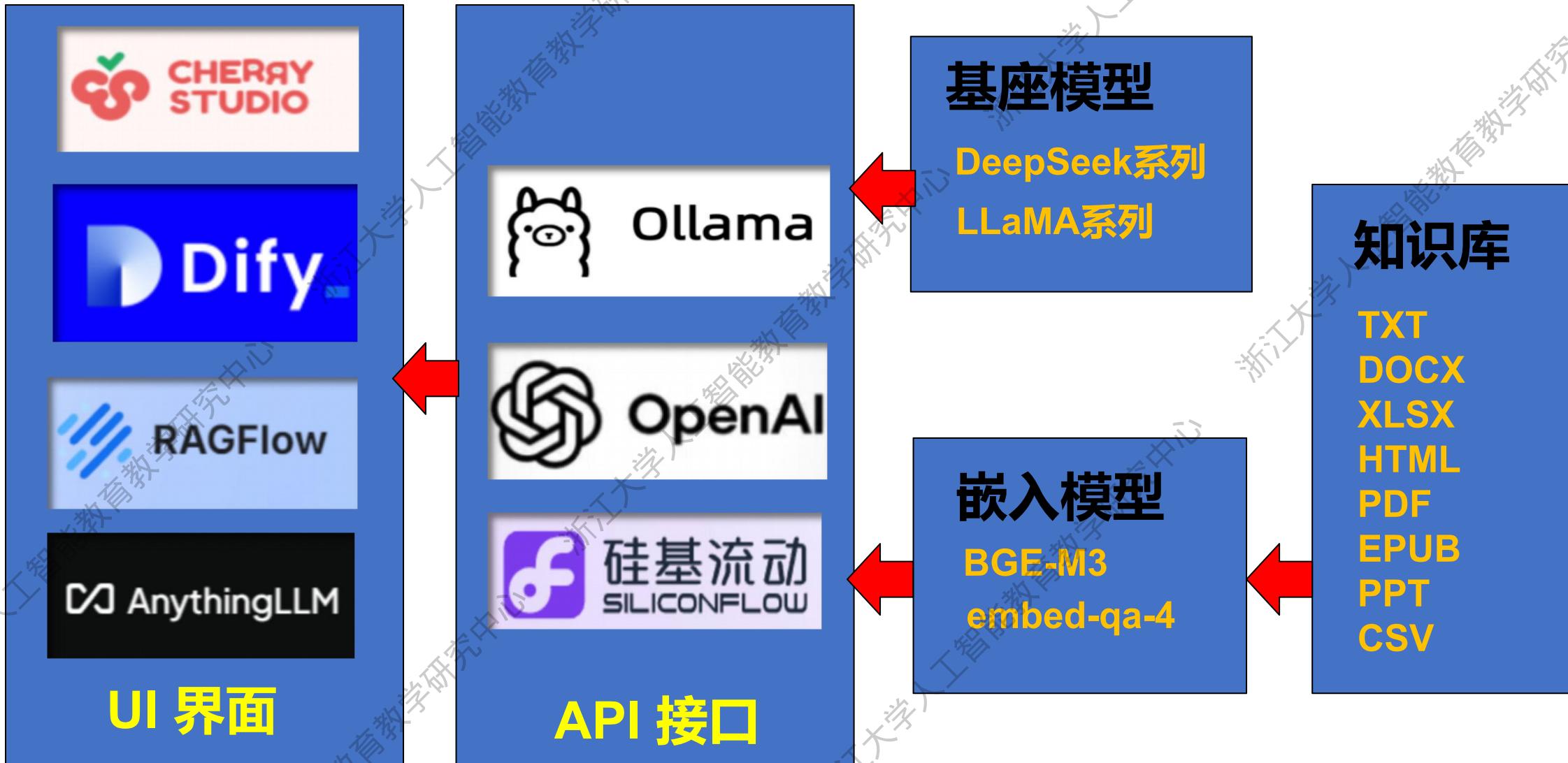
# 本地部署检索增强-RAG



# 本地部署检索增强-RAG

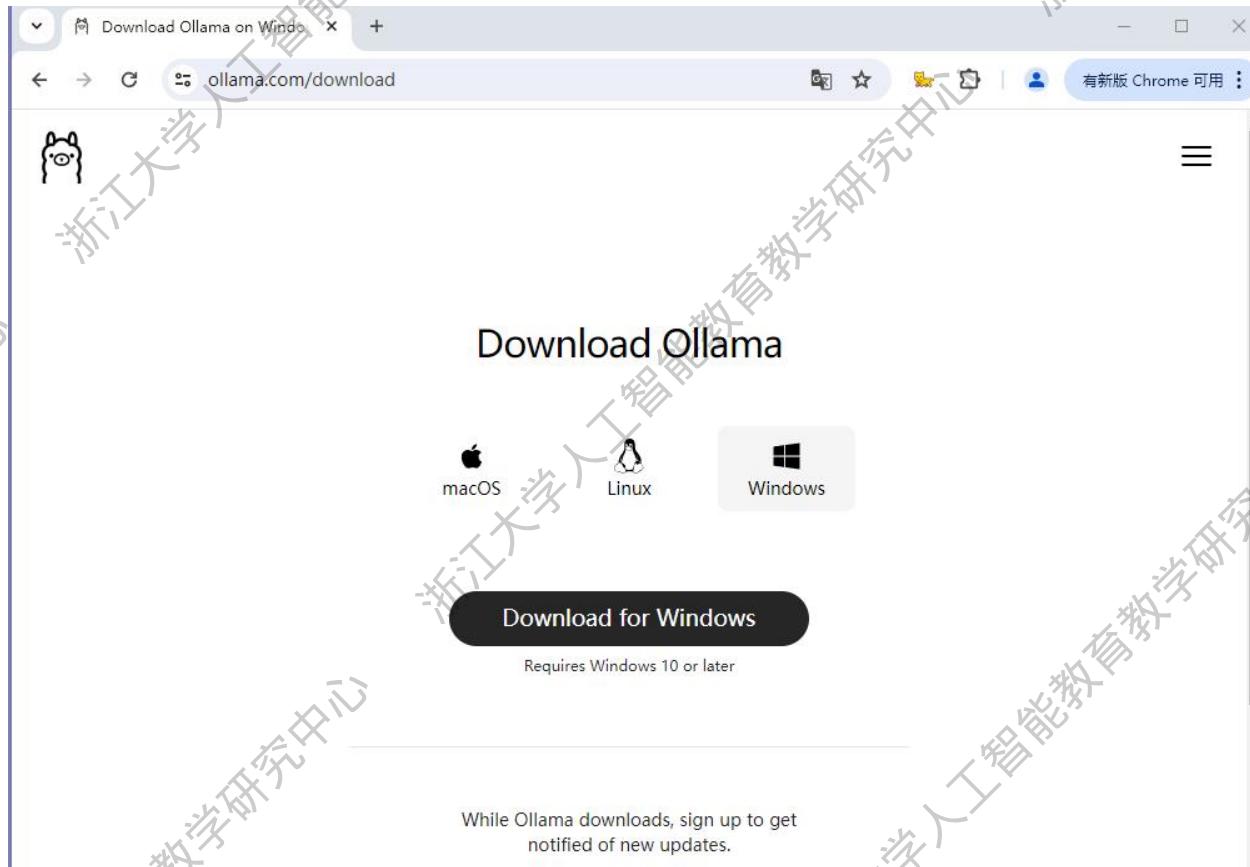


# 本地部署检索增强-RAG



# RAG部署过程

## 1. 下载安装OLLAMA: <https://ollama.com/download>



# RAG部署过程

2、下载安装Cherry Studio: <https://cherry-ai.com/download>



# RAG部署过程

## 3. 下载基座模型：Ollama pull deepseek-r1:7b

```
Windows PowerShell
版权所有 (C) Microsoft Corporation。保留所有权利。
尝试新的跨平台 PowerShell https://aka.ms/pscore6

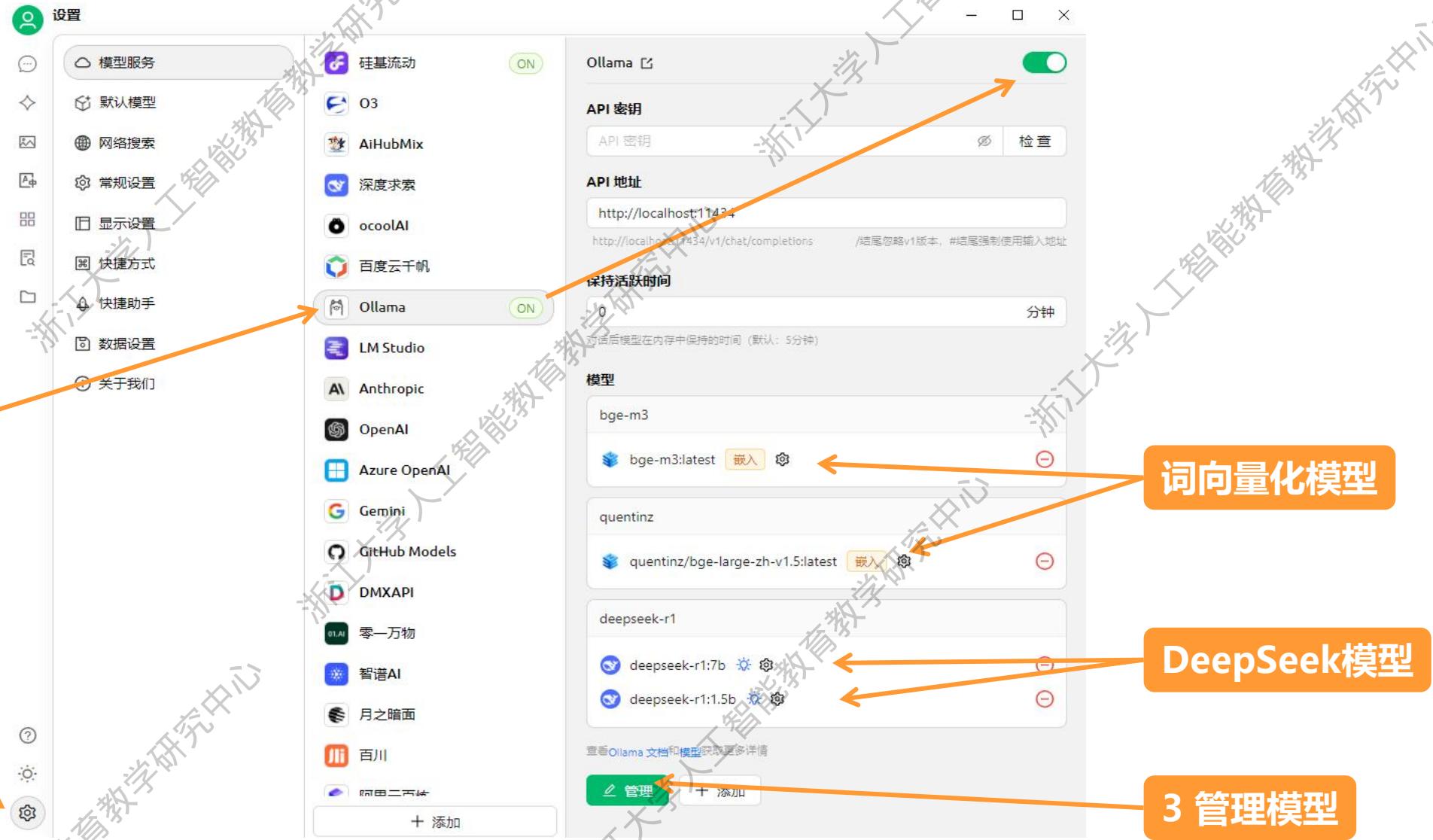
PS C:\Users\1> Ollama pull deepseek-r1:7b
pulling manifest
pulling 96c415656d37... 100% 4.7 GB
pulling 369ca498f347... 100% 387 B
pulling 6e4c38e1172f... 100% 1.1 KB
pulling f4d24e9138dd... 100% 148 B
pulling 40fb844194b2... 100% 487 B
verifying sha256 digest
writing manifest
success
PS C:\Users\1>
```

## 4. 下载嵌入模型：Ollama pull bge-m3:latest

```
PS C:\Users\1> Ollama pull bge-m3:latest
pulling manifest
pulling daec91ffb5dd... 100% 1.2 GB
pulling a406579cd136... 100% 1.1 KB
pulling 0c4c9c2a325f... 100% 337 B
verifying sha256 digest
writing manifest
success
PS C:\Users\1>
```

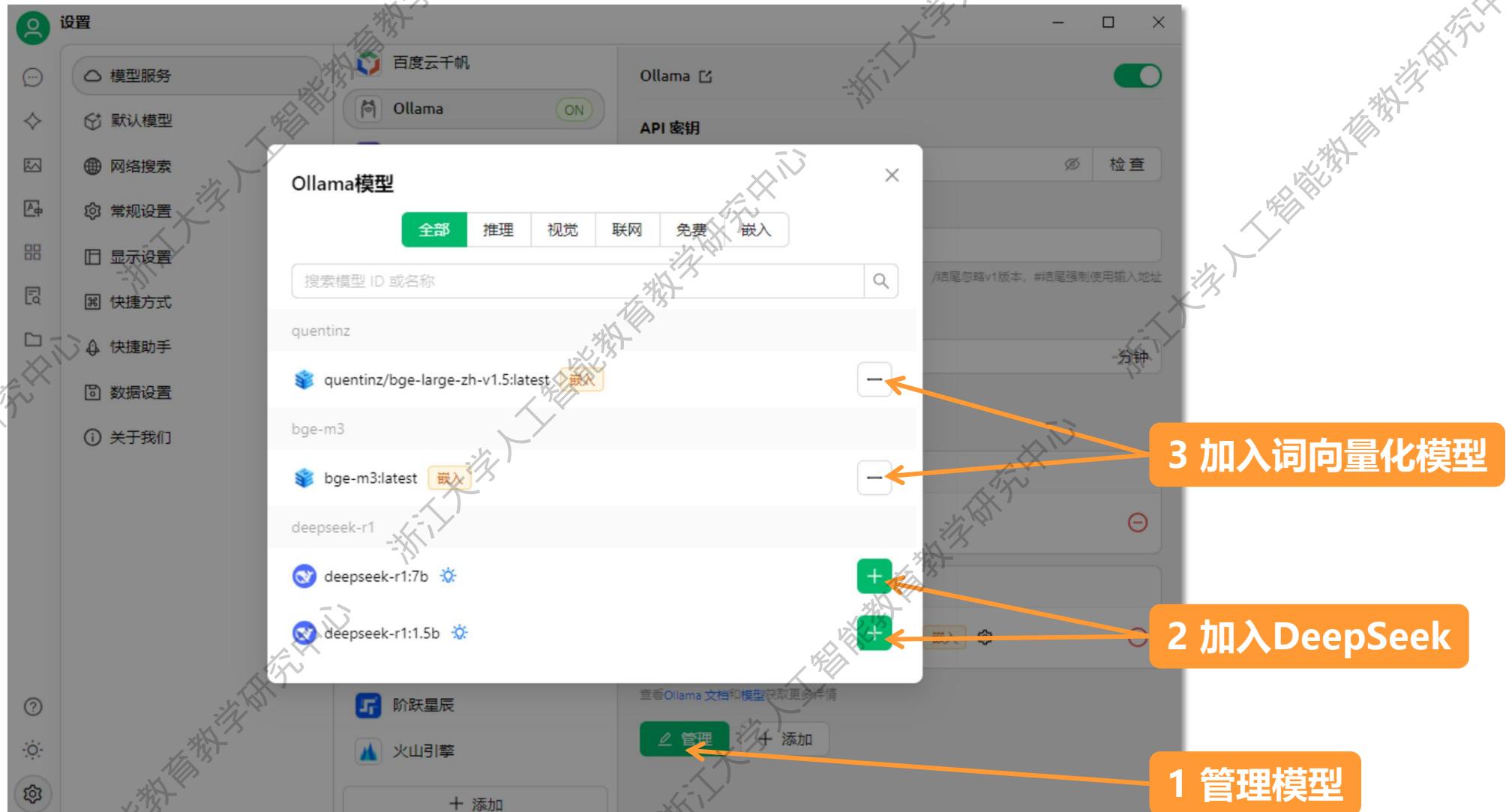
# RAG部署过程-配置OLLAMA接口

## 5、配置Cherry Studio 指定Ollama API



# RAG部署过程-管理OLLAMA模型

## 6、管理Ollama 模型



# RAG部署过程-管理知识库

1 打开管理界面

2 添加新知识库

3 选择向量化模型

4 输入知识库名字

5 保存



# RAG部署过程-知识库向量化



# RAG部署过程-对答



# 关于Ollama的安全报告

## 关于大型模型工具Ollama存在安全风险的情况通报

国家网络安全通报中心 2025年03月03日 18:27 北京

<https://mp.weixin.qq.com/s/n7PyLykK7MlO3re2oOyY5w>

### 一、风险隐患详情

使用Ollama在本地部署DeepSeek等大模型时，会在本地启动一个Web服务，并默认开放11434端口且无任何鉴权机制。该服务直接暴露在公网环境，存在以下风险：

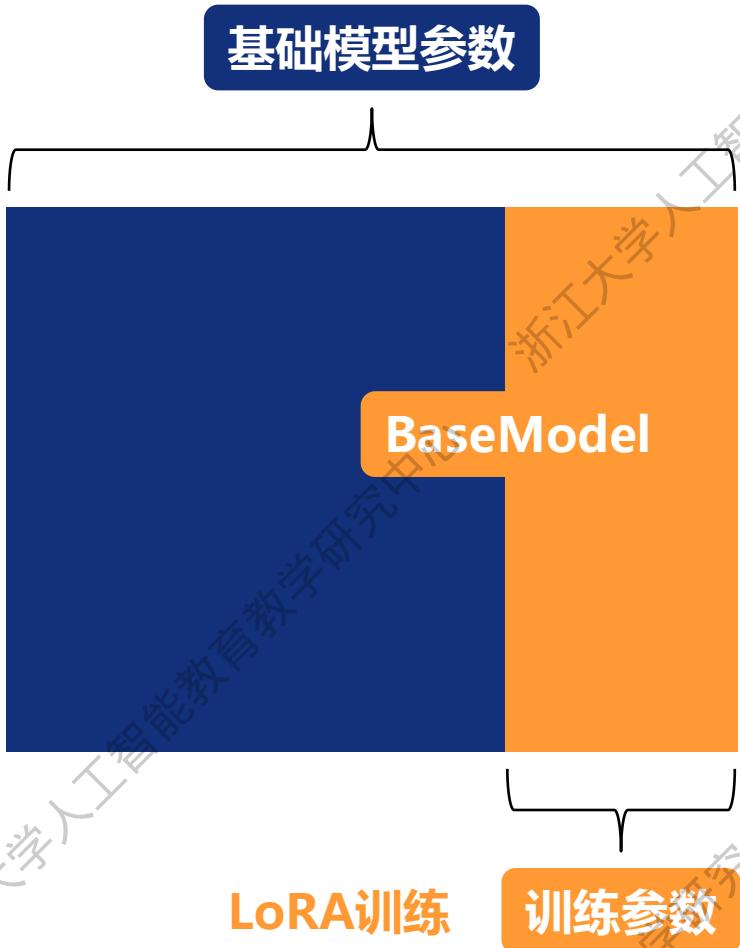
1. 未授权访问：未授权用户能够随意访问模型，并利用特定工具直接对模型及其数据进行操作，攻击者无需认证即可调用模型服务、获取模型信息，甚至通过恶意指令删除模型文件或窃取数据。
2. 数据泄露：通过特定接口可访问并提取模型数据，引发数据泄露风险。如：通过/api/show接口，攻击者能够获取模型的license等敏感信息，以及其他接口获取已部署模型的相关敏感数据信息。
3. 攻击者可利用Ollama框架历史漏洞（CVE-2024-39720/39722/39719/39721）。直接调用模型接口实施数据投毒、参数窃取、恶意文件上传及关键组件删除等操作，造成模型服务的核心数据、算法完整性和运行稳定性面临安全风险。

### 二、安全加固建议

1. 限制Ollama监听范围：仅允许11434端口本地访问，并验证端口状态。
2. 配置防火墙规则：对公网接口实施双向端口过滤，阻断11434端口的出入站流量。
3. 实施多层认证与访问控制：启用API密钥管理，定期更换密钥并限制调用频率。部署IP白名单或零信任架构，仅授权可信设备访问。
4. 禁用危险操作接口：如push/delete/pull等，并限制chat接口的调用频率以防DDoS攻击。
5. 历史漏洞修复：及时更新Ollama至安全版本，修复已知安全漏洞。

目前，已有大量存在此类安全隐患的服务器暴露在互联网上。建议广大用户加强隐患排查，及时进行安全加固，发现遭网络攻击情况第一时间向当地公安网安部门报告，配合公安网安部门开展调查处置工作。

# 大模型微调的二种技术——LoRA



LoRA (Low-Rank Adaptation of Large Language Models) 的核心思想是通过仅训练低秩矩阵，然后将这些参数注入到原始模型中，从而实现模型的微调，这种方法显著减少了所需的训练资源，特别适合在资源有限的环境中使用。

经过LoRA微调后的参数会覆盖原来基础模型的部分参数，因此模型的参数规模不会改变，但是输出已经跟专业数据对齐了

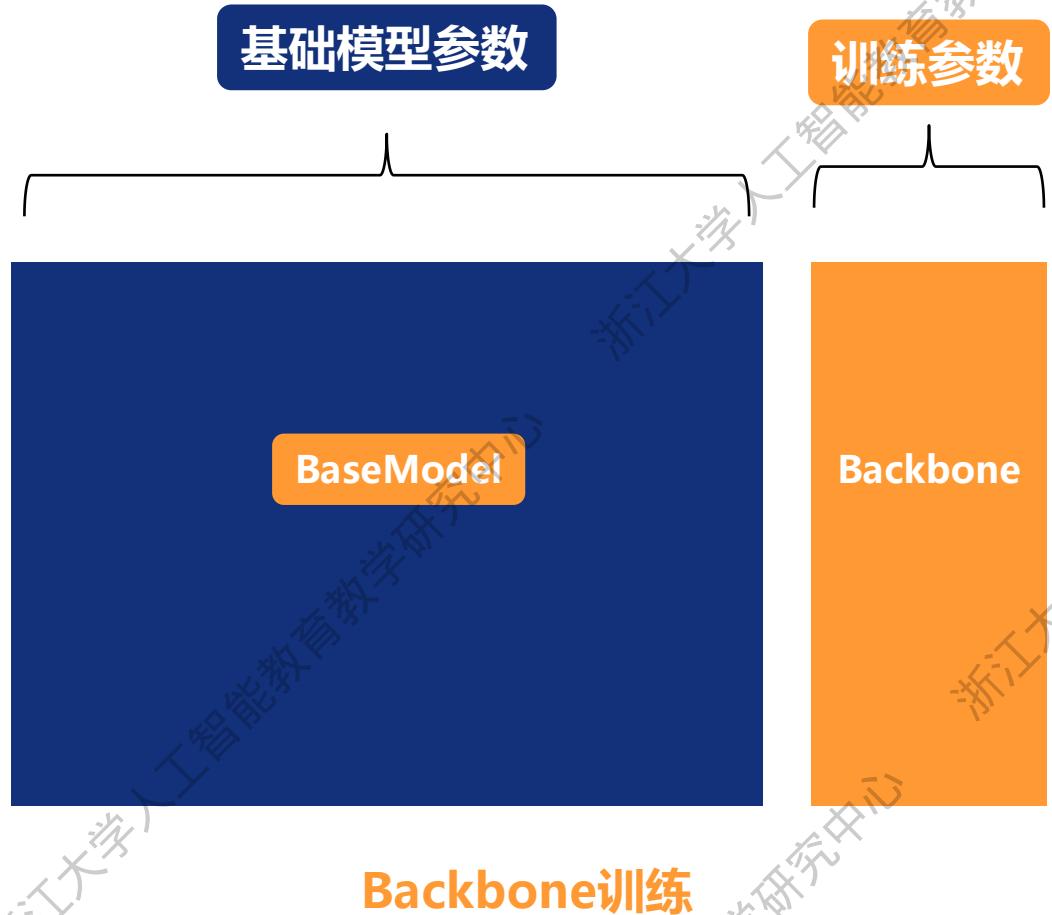
# 大模型微调的二种技术——LoRA



LoRA (Low-Rank Adaptation of Large Language Models) 的核心思想是通过仅训练低秩矩阵，然后将这些参数注入到原始模型中，从而实现模型的微调，这种方法显著减少了所需的训练资源，特别适合在资源有限的环境中使用。

经过LoRA微调后的参数会覆盖原来基础模型的部分参数，因此模型的参数规模不会改变，但是输出已经跟专业数据对齐了

# 大模型微调的二种技术——Backbone

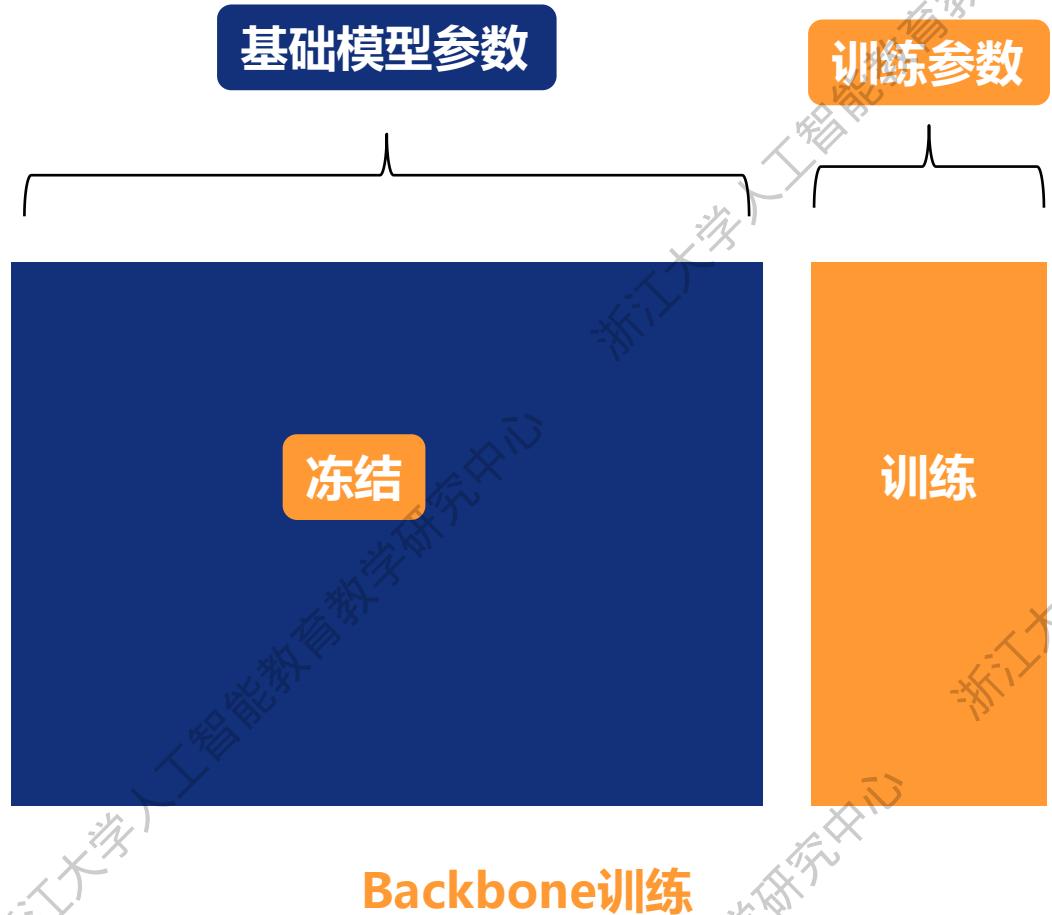


**Backbone**微调就是通过设计一个专业的骨架模型，这个骨架模型可以是不同类型的深度神经网络，如MLP、CNN等，但是参数较小。

将这个骨架模型装载在基础模型的输出端，在微调训练时，基础模型的参数冻结，只训练骨架模型。

这种微调的方法非常实用，不同的骨架模型完成不同的专业场景任务，推理的时候，只需要切换不同的骨架模型。就好比一个专业的摄影师通过更换不同的镜头和滤镜完成专业水平照片的拍摄。

# 大模型微调的二种技术——Backbone



**Backbone**微调就是通过设计一个专业的骨架模型，这个骨架模型可以是不同类型的深度神经网络，如MLP、CNN等，但是参数较小。

将这个骨架模型装载在基础模型的输出端，在微调训练时，基础模型的参数冻结，只训练骨架模型。

这种微调的方法非常实用，不同的骨架模型完成不同的专业场景任务，推理的时候，只需要切换不同的骨架模型。就好比一个专业的摄影师通过更换不同的镜头和滤镜完成专业水平照片的拍摄。

# 轻代码微调方法-基于LLaMA-Factory

## 1、LLaMA-Factory的安装部署：

```
git clone https://github.com/hiyouga/LLaMA-Factory.git
```

## 2、LLaMA-Factory必备项安装：

```
pip3 install --upgrade pip
```

```
pip3 install bitsandbytes>=0.39.0
```

进入LLaMA-Factory目录中，执行`pip3 install -e ".[torch,metrics]"`

## 3、LLaMA-Factory的主要子目录说明

config：存放自定义模型训练脚本

data：存放示例的训练数据集样本，采用json格式

examples：存放示例的训练脚本，包括`train_full`、`train_lora`、`train_qlora`、`inference`（推理）、`merge_lora`（模型合并）

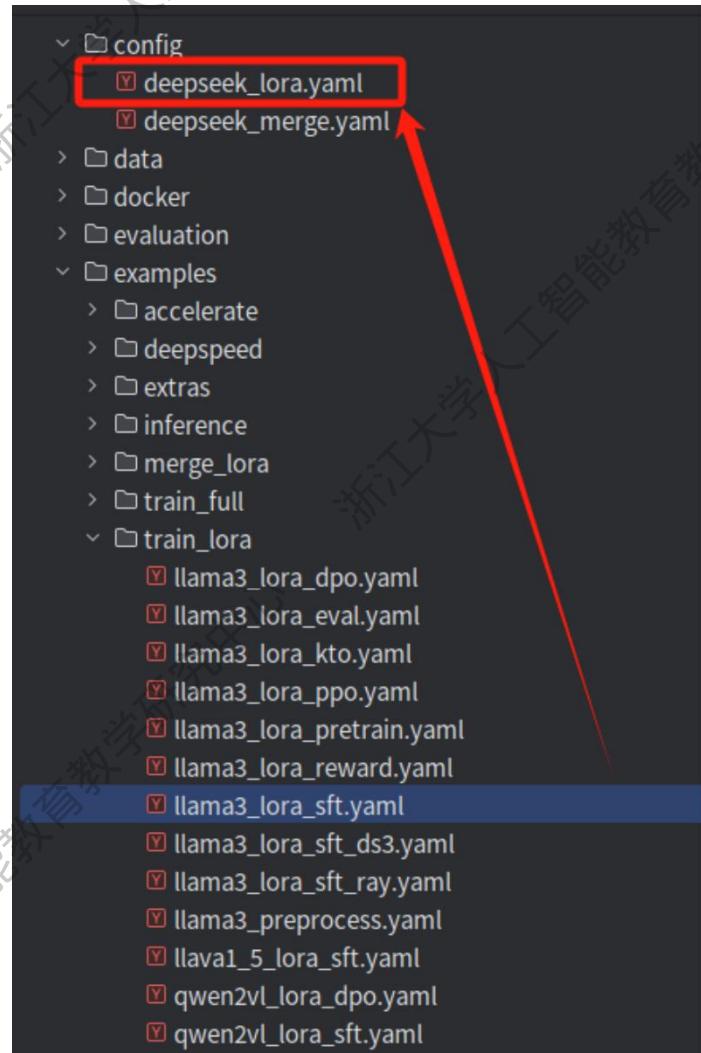
saves：微调模型临时存放目录

models：模型合并临时存放目录

# 轻代码微调方法-基于LLaMA-Factory

## 4. 训练模版制作

在examples目录下找到train\_lora子目录，复制llama3\_lora\_sft.yaml文件的一个备份，进行修改，重新命名，比如deepseek\_lora.yaml，存放到config目录中



# 轻代码微调方法-基于LLaMA-Factory

5、修改模板(deepseek\_lora.yaml)的训练参数，根据自己的项目修改相应的参数，几个主要的参数说明如下：

```
### model
model_name_or_path: /home/dell/baseModels/deepseek/DeepSeek-R1-Distill-Qwen-1.5B #基础模型的存放路径
#adapter_name_or_path: saves/deepseek_1.5b_lora #微调的权重参数，如果激活则可以继续训练
trust_remote_code: true

### method
stage: sft #微调的方法
de_train: true
finetuning_type: lora
lora_rank: 16 #秩的数量。数值越大，模型越大
lora_target: all #微调的模块

### dataset
dataset: 组织学简答V2_Qwen #自己的数据集，用于训练
template: deepseek #提示词模板 deepseek, qwen
cutoff_len: 2048
max_samples: 1000000
overwrite_cache: true
preprocessing_num_workers: 16
```

```
### output
output_dir: saves/deepseek_1.5b_lora/ #微调权重的保存路径
logging_steps: 10
save_steps: 250
plot_loss: true
overwrite_output_dir: true
save_total_limit: 3

### train
per_device_train_batch_size: 10 #每次加载的数据条数，数值越大，训练越快，内存越多
gradient_accumulation_steps: 8 #累积梯度的轮次进行权重调整
learning_rate: 1.0e-4 #初始的学习率
num_train_epochs: 1.0 #训练的轮次
lr_scheduler_type: cosine
warmup_ratio: 0.1
bf16: true
ddp_timeout: 180000000

### eval
val_size: 1
per_device_eval_batch_size: 10
eval_strategy: steps
eval_steps: 500
```

联系邮箱：zhuzhaoyang@zju.edu.cn

# 轻代码微调方法-基于LLaMA-Factory

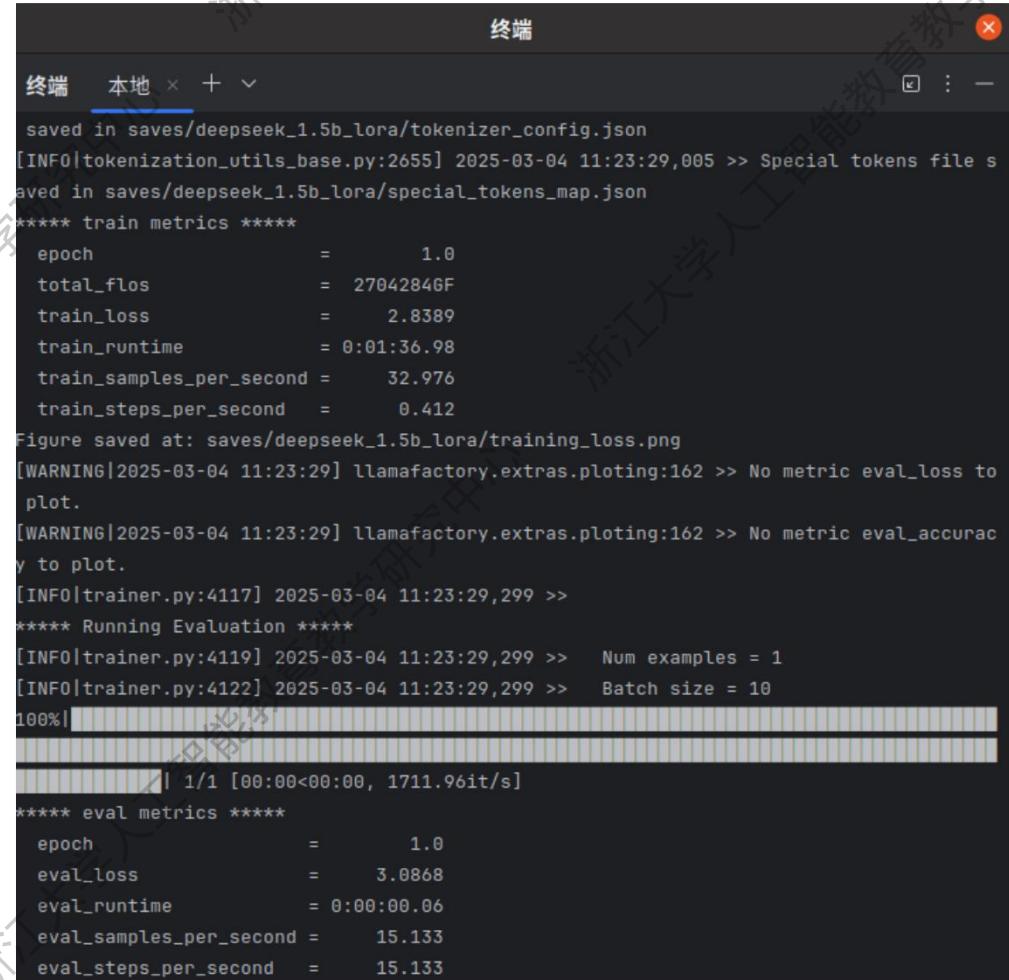
## 6、将训练的数据集拷贝到data目录下，并修改dataset\_info.json文件，增加数据的索引



```
        "identity": {  
            "file_name": "identity.json"  
        },  
        "alpaca_en_demo": {  
            "file_name": "alpaca_en_demo.json"  
        },  
        "alpaca_zh_demo": {  
            "file_name": "alpaca_zh_demo.json"  
        },  
        "Belle_open_source_0.5M": {  
            "file_name": "Belle_open_source_0.5M.json"  
        },  
        "组织学简答V2_Qwen": {  
            "file_name": "组织学简答V2_Qwen.json"  
        },  
        "医学细胞生物学实验指导简答V1_Qwen": {  
            "file_name": "医学细胞生物学实验指导简答V1_Qwen.json"  
        },  
    },  
},  

```

## 7、训练：执行命令 llmfactory-cli train deepseek\_lora.yaml 即可开始训练



```
终端  
终端 本地 + -  
saved in saves/deepseek_1.5b_lora/tokenizer_config.json  
[INFO|tokenization_utils_base.py:2655] 2025-03-04 11:23:29,005 >> Special tokens file saved in saves/deepseek_1.5b_lora/special_tokens_map.json  
***** train metrics *****  
epoch = 1.0  
total_flos = 2704284GF  
train_loss = 2.8389  
train_runtime = 0:01:36.98  
train_samples_per_second = 32.976  
train_steps_per_second = 0.412  
Figure saved at: saves/deepseek_1.5b_lora/training_loss.png  
[WARNING|2025-03-04 11:23:29] llmfactory.extras.plotting:162 >> No metric eval_loss to plot.  
[WARNING|2025-03-04 11:23:29] llmfactory.extras.plotting:162 >> No metric eval_accuracy to plot.  
[INFO|trainer.py:4117] 2025-03-04 11:23:29,299 >>  
***** Running Evaluation *****  
[INFO|trainer.py:4119] 2025-03-04 11:23:29,299 >> Num examples = 1  
[INFO|trainer.py:4122] 2025-03-04 11:23:29,299 >> Batch size = 10  
100%|██████████| 1/1 [00:00<00:00, 1711.96it/s]  
***** eval metrics *****  
epoch = 1.0  
eval_loss = 3.0868  
eval_runtime = 0:00:00.06  
eval_samples_per_second = 15.133  
eval_steps_per_second = 15.133
```

# DeepSeek本地微调的简单应用

## AI智能助教

提供一段文字来生成所需要的题目

输入

解剖学上把神经系统分为中枢神经系统(central nervous system,CNS)和周围神经系统 (peripheral nervous system,PNS)。中枢神经系统由分别位于颅腔和椎管内的脑和脊髓组成,周围神经系统由脑神经、脊神经以及传出神经(运动神经)、传入神经(感觉神经)组成的外周神经共同构成,同时还包含CNS以外的神经元胞体组成的神经节、运动和感觉神经末梢。在功能上,神经系统又可分为躯体神经系统(somatic nervous system,SNS)和自主神经系统 (autonomic nervous system,ANS),前者由CNS和PNS的躯体部分组成,包括躯体中除内脏、心肌和腺体以外的运动和感觉神经支配,后者由CNS和PNS中的自主部分组成,包括支配内脏、心肌和腺体的非随意运动传出系统,还包括感受内脏疼痛、自主反射的感觉传入系统。ANS还可进一步分为交感和副交感神经系统以及独立于这两类之外的支配胃肠道的肠道神经系统(enteric nerve system,ENS)。

选择推理模式

选择题

deepseek-r1:32b

输出

系统的组成部分、周围神经系统的构成、自主神经系统的类型等。

在出题过程中,要注意选项的设置要有一定的迷惑性,同时确保正确答案明显且准确。此外,问题应涵盖不同的知识点,以全面考察对文字的理解。

</think>

### 根据第一段文字:

### 选择题1: 组织学又被称为什么?

A: 宏观解剖学  
B: 显微解剖学  
C: 微观组织学  
D: 细胞科学

\*\*参考答案: B\*\*

### 选择题2: 组织学主要研究的是以下哪一项?

A: 机体的宏观结构  
B: 机体的微细结构及其功能  
C: 细胞的代谢产物  
D: 机体的生理反应

\*\*参考答案: B\*\*

### 选择题3: 构成机体形态结构与功能的基本单位是什么?

A: 细胞  
B: 器官  
C: 组织  
D: 系统

\*\*参考答案: A\*\*

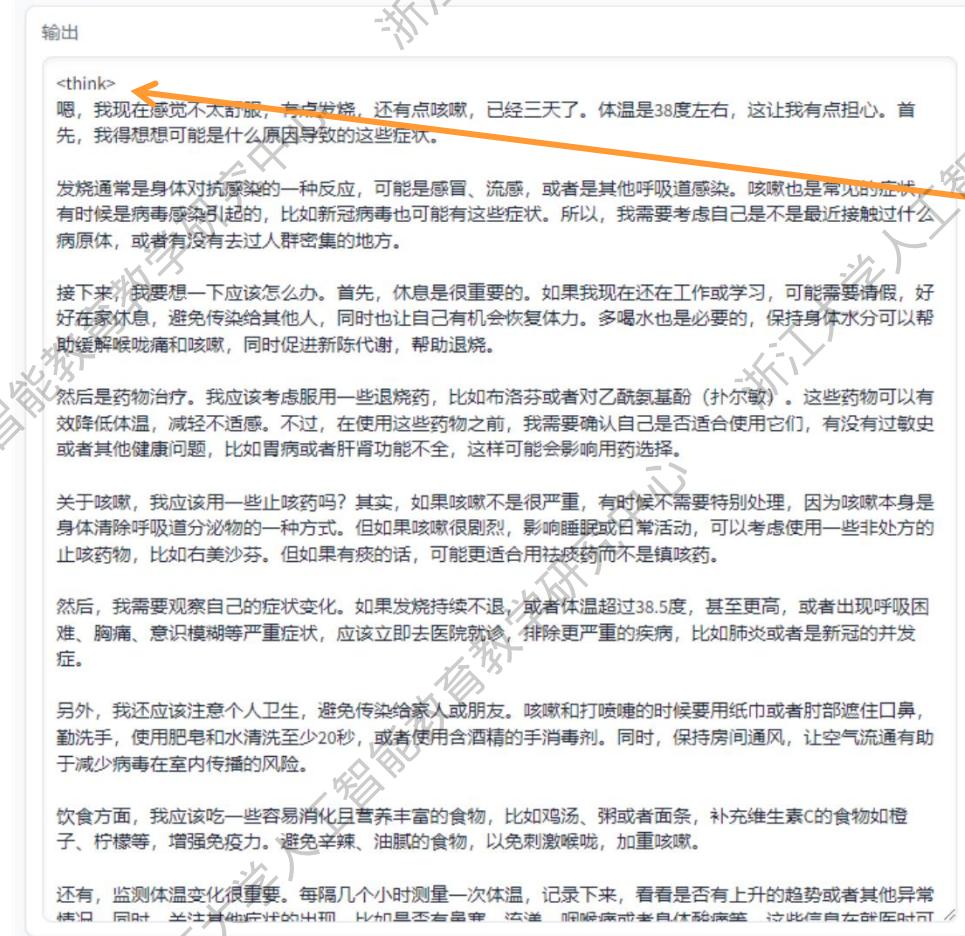
### 选择题4: 由细胞及其产物组成的四种基本组织分别是?

思考过程

自测选择题

# DeepSeek本地微调的简单应用

## AI智能助教

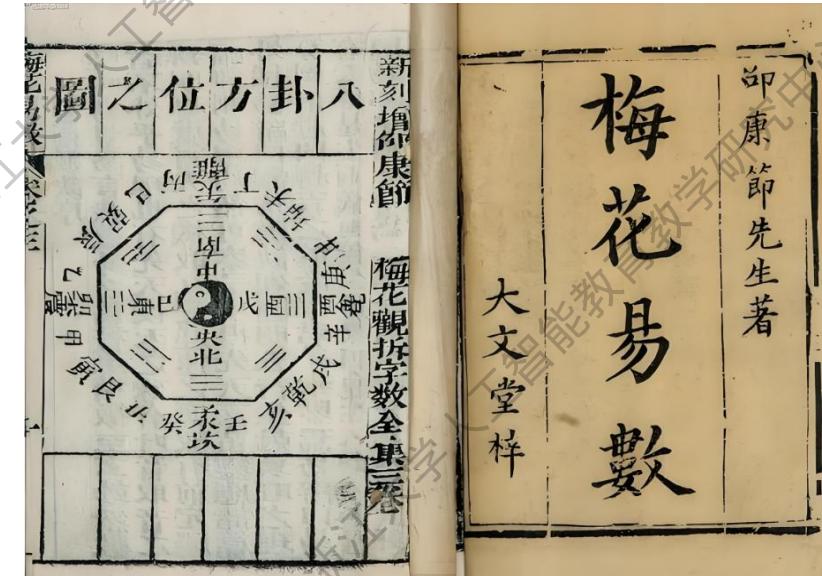
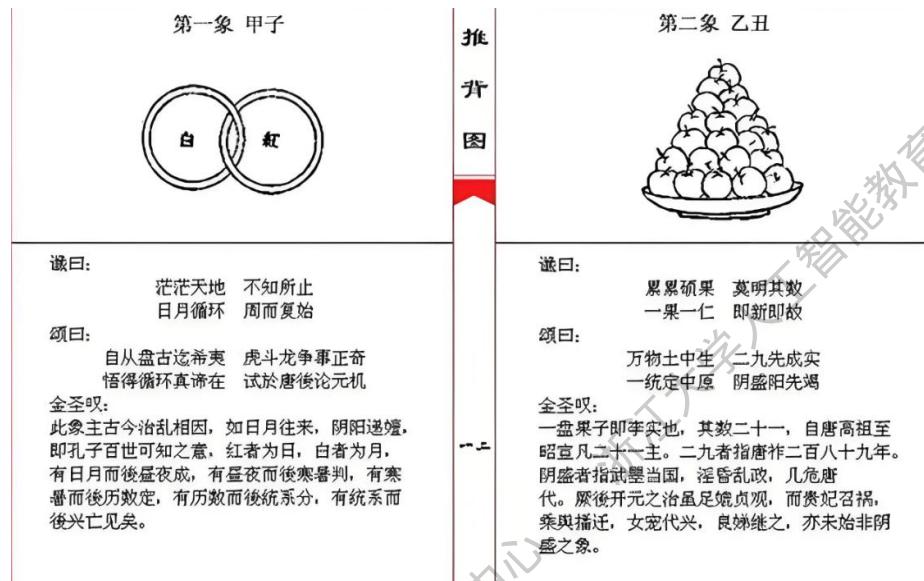


# 智能助教系统与虚拟仿真实验室的集成



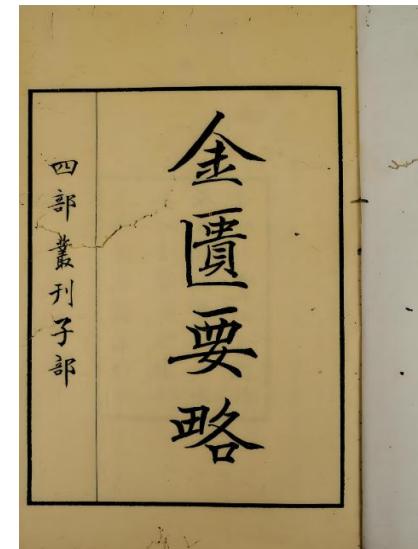
医学院基础医学实验教学中心 一体式机能智慧教学实验室

# 展望——走一条属于中国的AI之路



总有一天，我们将会用AI技术揭开推背图、梅花易数的神秘面纱

也许人工智能正走在不断论证中医和周易的路上





**感谢聆听**  
**敬请批评指正**



**浙江大学**  
ZHEJIANG UNIVERSITY