



南方科技大学
SOUTHERN UNIVERSITY OF SCIENCE AND TECHNOLOGY



检索增强生成 (RAG)

南方科技大学
唐 博
tangb3@sustech.edu.cn





内容提纲



- **RAG简介与应用**：记忆无痕，问道有源
- **RAG原理与流程**：横看成岭，侧见成峰
- **RAG向量数据库**：载舟亦水，活水为源
- **RAG上手与实战**：手中有剑，心中有数



内容提纲



- ✓ **RAG简介与应用**：记忆无痕，问道有源
- **RAG原理与流程**：横看成岭，侧见成峰
- **RAG向量数据库**：载舟亦水，活水为源
- **RAG上手与实战**：手中有剑，心中有数



大模型工作原理示例



大模型 $f()$



返回结果 y

什么样的学生适合报考南科大？

Thought for 20 seconds

南方科技大学（简称“南科大”）作为一所创新型、研究型大学，自2011年建校以来，凭借其独特的办学理念和优质资源吸引了众多学生。以下从多个维度分析适合报考南科大的学生类型，帮助判断是否与学校特点匹配：

一、学术兴趣与学科方向

- 1. 理工科优势领域**
南科大以理工医商为主，尤其在材料科学、计算机、电子工程、生物医学等前沿学科领域具有较强实力。例如，化学、材料科学在ESI排名中进入全球前1%。若学生对人工智能、新能源、生物技术等新兴领域有浓厚兴趣，南科大的科研平台（如粤港澳大湾区量子科学中心）能提供优质资源。
- 2. 交叉学科探索**
学校推行“通识+专业教育”，鼓励跨学科学习。例如“金融数学”“生物医学工程”等交叉专业，适合希望突破传统学科边界的学生。

二、学习模式适应性

- 1. 研究导向型学习**
本科生即可进入实验室参与课题（如大二开始的“URP计划”），2022年本科生发表SCI论文超百篇。适合热衷科研、希望早进实验室的学生。
- 2. 国际化培养**
全英文专业课程占比超60%，与麻省理工、剑桥等高校联合培养。需具备较强英语能力（高考英语120+/150分制较有竞争力），并有志于海外深造者（近年境外升学率约40%）。

提示词 x



DeepSeek, YYDS?



DeepSeek并不是无所不能：

懂一点，但可能乱说！

会一点，但可能不说！

做一点，但可能会错！



DeepSeek 可能乱说



南科大排名从国内第十到国内前三需要做哪些努力

Thinking... ^

好的，用户问南方科技大学要从国内第十提升到前三需要做哪些努力。首先，我需要了解南科大目前的状况。根据之前的对话，南科大是一所创新型、研究型大学，成立于2011年，以理工医商为主，采用综合评价录取，国际化程度高，科研平台不错，本科生有参与科研的机会，毕业生深造率也高。但可能还存在一些不足，比如文科规模小，传统名校光环不足。

用户现在关心的是排名提升，国内前三通常是清北复交浙这样的顶尖学府。南科大作为年轻大学，要进入前三挑战很大。需要分析排名指标，通常大学排名考虑学术声誉、科研产出、师资力量、国际化、学生质量、资金投入等。

首先，学术声誉和科研产出方面，南科大已经有一定基础，比如材料科学进入ESI前1%，但可能需要在更多学科领域取得突破，尤其是国际顶尖期刊论文的数量和影响力。可能需要引进更多顶尖学者，增加科研经费，建立国家级重点实验室。

其次，师资力量是关键。虽然南科大可能已经有一些优秀教师，但要达到前三，可能需要更多国际知名的教授，提高师生比，吸引诺奖级学者或者院士级别的教授加入。

国际化程度方面，南科大已经和麻省理工、剑桥有合作，但可能需要进一步扩大国际合作，增加国际学生和教师的比例，提升全英文课程的数量和质量，参与更多国际联合研究项目。

学生质量方面，虽然现在采用631模式，但可能需要提高录取标准，吸引更多高分考生，同时加强本科生的科研和实践机会，提升毕业生就业和深造的质量，尤其是进入世界顶尖学府的比例。

资金投入也是重要因素。南科大位于深圳，有地域优势，可能需要争取更多政府拨款和企业合作，增加校友捐赠，提升资金实力以支持科研和教学。



南方科技大学（简称“南科大”）若要从国内第十名左右跃升至前三名（对标清北复交浙等顶尖高校），需在学科建设、科研实力、师资队伍、人才培养、国际化、社会声誉等多维度实现跨越式突破。以下从具体策略角度展开分析：

一、学科布局：聚焦前沿领域，打造“不可替代性”

1. 集中资源突破“卡脖子”领域

- **聚焦方向**：在量子计算、人工智能、合成生物学、新能源材料等国家战略领域建立“学科高地”，例如建立国家级实验室（对标中国科大“量子信息国家实验室”）。
- **案例参考**：西湖大学以生物医学、人工智能为突破口，5年内建成10个世界级研究中心，快速提升国际影响力。

2. 交叉学科创新机制

- 设立跨学科研究院（如“智能科学与医学工程研究院”），打破院系壁垒，推动“AI+生物”“材料+能源”等融合方向，产出颠覆性成果。

二、科研实力：从“量”到“质”的跃升

1. 提升原创性研究比重

- 鼓励教授在《Nature》《Science》《Cell》等顶刊发表论文，设立专项奖励（如港科大对顶刊论文奖励50万/篇）。
- 数据对标：2023年清华在NS主刊发文58篇，南科大需从当前年均10篇左右向30篇以上突破。

2. 技术转化与产业赋能

- 依托深圳产业优势，与大疆、华为、比亚迪共建联合实验室，推动专利商业化（参考斯坦福-硅谷模式）。
- 目标：技术转化率从目前的15%提升至30%，孵化独角兽企业3-5家。





DeepSeek可能乱说



南科大排名从国内第十到国内前三需要做哪些努力



Thinking... ^

好的，用户问南方科技大学要从国内第十提升到前三需要做哪些努力。首先，我需要了解南科大目前的状况。根据之前



南方科技大学（简称“南科大”）若要从国内第十名左右跃升至前三名（对标清北复交浙等顶尖高校），需在学科建设、科研实力、师资队伍、人才培养、国际化、社会声誉等多维度实现跨越式突破。以下从具体策略角度展开分析：

一、学科布局：聚焦前沿领域，打造“不可替代性”

2023年清华大学在NS上发文数量再创新高

记者从清华大学获悉，2023年清华大学在《自然》（Nature）、《科学》（Science）期刊上发表论文数量再创新高，达到39篇，其中作为第一单位发表的论文21篇，位居全国部属高校第一。

其次，师资力量是关键。虽然南科大可能已经有一些优秀教师，但要达到前三，可能需要更多国际知名的教授，提高师生比，吸引诺奖级学者或者院士级别的教授加入。

国际化程度方面，南科大已经和麻省理工、剑桥有合作，但可能需要进一步扩大国际合作，增加国际学生和教师的比例，提升全英文课程的数量和质量，参与更多国际联合研究项目。

学生质量方面，虽然现在采用631模式，但可能需要提高录取标准，吸引更多高分考生，同时加强本科生的科研和实践机会，提升毕业生就业和深造的质量，尤其是进入世界顶尖学府的比例。

资金投入也是重要因素。南科大位于深圳，有地域优势，可能需要争取更多政府拨款和企业合作，增加校友捐赠，提升资金实力以支持科研和教学。



源”等融合方向，产出颠覆性成果。

二、科研实力：从“量”到“质”的跃升

1. 提升原创性研究比重

- 鼓励教授在《Nature》《Science》等顶刊发表论文，设立专项奖励（如港科大对顶刊论文奖励50万/篇）。

数据对标：2023年清华在NS主刊发文58篇，南科大需从当前年均10篇左右向30篇以上突破。

2. 技术转化与产业赋能

- 依托深圳产业优势，与大疆、华为、比亚迪共建联合实验室，推动专利商业化（参考斯坦福·硅谷模式）。
- 目标：技术转化率从目前的15%提升至30%，孵化独角兽企业3-5家。





DeepSeek可能不说



南方科技大学图书馆今年寒假的开放安排是？



DeepSeek可能不说



已搜索到 50 个网页 >

已深度思考 (用时 21 秒) ▾

根据您提供的搜索结果，目前未找到关于南方科技大学（南科大）图书馆2025年寒假开放安排的直接信息。所有搜索结果均涉及其他高校的图书馆安排，例如中国传媒大学、中国科学院大学南京学院、华东师范大学、中国科学技术大学等，但未提及南方科技大学的具体通知。

其他高校图书馆寒假安排参考

若需了解其他高校的寒假开放时间或服务调整，可参考以下信息（可能与南科大的安排存在差异）：

1. **开放时段**: 多数高校图书馆在寒假期间分阶段开放，例如：
 - 中国传媒大学：1月13日-1月27日、2月5日-2月18日工作日开放，2月24日恢复正常¹。
 - 华东师范大学：1月23日-2月13日部分区域开放，春节期间（1月28日-31日）闭馆⁵。



DeepSeek 可能做错



帮助我解答以下问题： Q.7 (12 points) Consider the numbers of the form $n^{13} - 2*n^7 + n$, where n is an integer. Determine for which values of n , the number $n^{13} - 2*n^7 + n$ is divisible by 98.



DeepSeek可能做错



收起思考过程 ^

思考已完成 292.474s

要确定哪些整数 n 使得数 $n^{13} - 2n^7 + n$ 能被 98 整除，首先分解 98 为 2×7^2 。我们需要该表达式同时被 2 和 49 整除。

1. 模 2 分析

表达式为 $n^{13} - 2n^7 + n$ 。模 2 下， $2n^7 \equiv 0 \pmod{2}$ ，因此只需考虑 $n^{13} + n \pmod{2}$ 。无论 n 是奇数还是偶数， $n^k \equiv n \pmod{2}$ ，故：

$$n^{13} + n \equiv n + n \equiv 2n \equiv 0 \pmod{2}$$

对所有整数 n 成立。因此，表达式总是被 2 整除。

2. 模 49 分析

表达式可因式分解为 $n(n^6 - 1)^2$ 。我们需要 $n(n^6 - 1)^2 \equiv 0 \pmod{49}$ 。分两种情况：

情况 1: $n \equiv 0 \pmod{7}$

若 $n = 7k$ ，则：

$$n(n^6 - 1)^2 = 7k \cdot ((7k)^6 - 1)^2$$

当 $k \equiv 0 \pmod{7}$ (即 $n \equiv 0 \pmod{49}$) 时，表达式被 49 整除。否则，第三项 $7k$ 仅被 7 整除，不满足条件。

情况 2: $n \not\equiv 0 \pmod{7}$

此时 n 与 7 互质。根据欧拉定理， $n^{42} \equiv 1 \pmod{49}$ 。我们需要 $(n^6 - 1)^2 \equiv 0 \pmod{49}$ ，即 $n^6 \equiv 1 \pmod{49}$ 。通过直接计算或利用模运算性质，满足 $n^6 \equiv 1 \pmod{49}$ 的解为：

$$n \equiv 1, 18, 19, 30, 31, 48 \pmod{49}$$

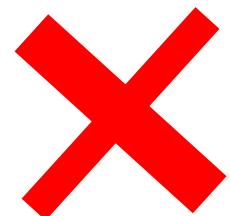
最终结论

当且仅当 n 满足以下条件时， $n^{13} - 2n^7 + n$ 被 98 整除：

$$n \equiv 0 \pmod{49} \quad \text{或} \quad n \equiv 1, 18, 19, 30, 31, 48 \pmod{49}$$

答案：

所有整数 n 满足 $n \equiv 0, 1, 18, 19, 30, 31, \text{ 或 } 48 \pmod{49}$ 。





检索增强生成(RAG)



DeepSeek并不是无所不能：

懂一点，但可能乱说！

会一点，但可能不说！

做一点，但可能会错！

检索增强生成(RAG)是解决这些问题的利剑！



RAG基本原理



OpenAI



LLaMA



deepseek



Qwen



智谱·AI



面壁智能



Midjourney



Colossal-AI

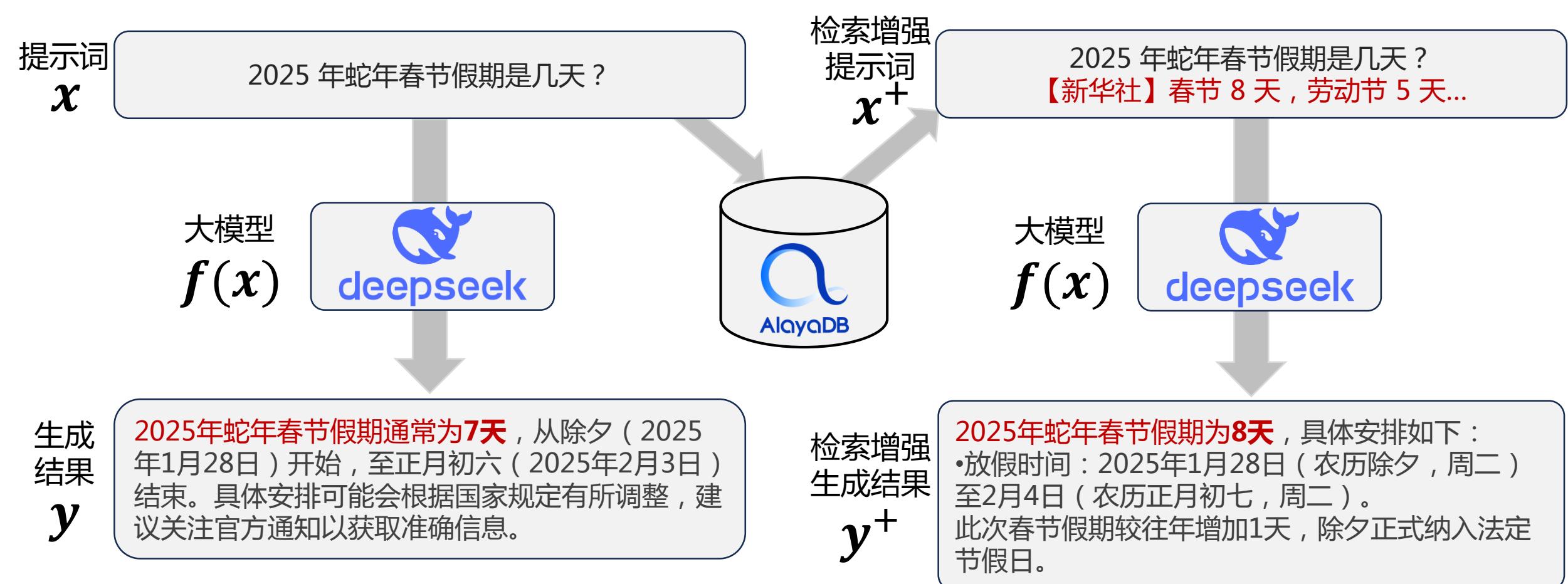
$$f(\mathbf{x}) \rightarrow y$$

检索增强生成 (RAG)

$$f(\mathbf{x}^+) \rightarrow y^+$$



RAG基本原理





定制版南科AI



南科大图书馆今年寒假的开放安排是什么？

已搜索到 50 个网页 >

已深度思考（用时 21 秒） ▾

根据您提供的搜索结果，目前未找到关于南方科技大学（南科大）图书馆**2025年寒假开放安排**的直接信息。所有搜索结果均涉及其他高校的图书馆安排，例如中国传媒大学、中国科学院大学南京学院、华东师范大学、中国科学技术大学等，但未提及南方科技大学的具体通知。

其他高校图书馆寒假安排参考

若需了解其他高校的寒假开放时间或服务调整，可参考以下信息（可能与南科大的安排存在差异）：

- 开放时段：**多数高校图书馆在寒假期间分阶段开放，例如：
 - 中国传媒大学：1月13日-1月27日、2月5日-2月18日工作日开放，2月24日恢复正常¹。
 - 华东师范大学：1月23日-2月13日部分区域开放，春节期间（1月28日-31日）闭馆⁵。

南科大图书馆今年寒假的开放安排是什么？

南科AI

根据南方科技大学图书馆发布的《2025年寒假开放和服务安排》，2025年寒假的开放安排如下：

琳恩图书馆和涵泳图书馆

- **闭馆时间：**1月20日至2月12日。
- **建议：**请读者在闭馆前尽量借阅琳恩图书馆的藏书。

一丹图书馆

- **开放时间**（1月13日至2月14日，除春节八天假期外）：
 - 二、三、四楼：8:00-22:00。
 - 一楼24小时学习区：正常开放。
 - 服务台服务时间：维持不变。
 - 报告厅、放映厅、培训教室：暂停预约。
- **闭馆时间：**
 - 1月16日至1月19日（因馆舍清洁需要）。
 - 1月28日至2月4日（春节八天假期）：仅开放一楼学习区，开放时间调整为8:00-22:00，服务台暂停服务。

图书和京东阅读本延期归还



定制版南科AI



帮助我解答以下问题: Q.7 (12 points) Consider the numbers of the form $n^{13} - 2n^7 + n$, where n is an integer. Determine for which values of n , the number $n^{13} - 2n^7 + n$ is divisible by 98.

收起思考过程 ^
思考已完成

要确定哪些整数 n 使得数 $n^{13} - 2n^7 + n$ 能被 98 整除, 首先分解 98 为 2×7^2 。我们需要该表达式同时被 2 和 49 整除。

1. 模 2 分析

表达式为 $n^{13} - 2n^7 + n$ 。模 2 下, $2n^7 \equiv 0 \pmod{2}$, 因此只需考虑 $n^{13} + n \pmod{2}$ 。无论 n 是奇数还是偶数, $n^k \equiv n \pmod{2}$, 故:

$$n^{13} + n \equiv n + n \equiv 2n \equiv 0 \pmod{2}$$

对所有整数 n 成立。因此, 表达式总是被 2 整除。

2. 模 49 分析

表达式可因式分解为 $n(n^6 - 1)^2$ 。我们需要 $n(n^6 - 1)^2 \equiv 0 \pmod{49}$ 。分两种情况:

情况 1: $n \equiv 0 \pmod{7}$

若 $n = 7k$, 则:

$$n(n^6 - 1)^2 = 7k \cdot ((7k)^6 - 1)^2$$

当 $k \equiv 0 \pmod{7}$ (即 $n \equiv 0 \pmod{49}$) 时, 表达式被 49 整除。否则, 第三项 $7k$ 仅被 7 整除, 不满足条件。

情况 2: $n \not\equiv 0 \pmod{7}$

此时 n 与 7 互质。根据欧拉定理, $n^{42} \equiv 1 \pmod{49}$ 。我们需要 $(n^6 - 1)^2 \equiv 0 \pmod{49}$, 即 $n^6 \equiv 1 \pmod{49}$ 。通过直接计算或利用模运算性质, 满足 $n^6 \equiv 1 \pmod{49}$ 的解为:

$$n \equiv 1, 18, 19, 30, 31, 48 \pmod{49}$$

最终结论

当且仅当 n 满足以下条件时, $n^{13} - 2n^7 + n$ 被 98 整除:

$$n \equiv 0 \pmod{49} \quad \text{或} \quad n \equiv 1, 18, 19, 30, 31, 48 \pmod{49}$$

答案:

所有整数 n 满足 $n \equiv 0, 1, 18, 19, 30, 31, \text{ 或 } 48 \pmod{49}$ 。

帮助我解答以下问题: Q.7 (12 points) Consider the numbers of the form $n^{13} - 2n^7 + n$, where n is an integer. Determine for which values of n , the number $n^{13} - 2n^7 + n$ is divisible by 98.



南科AI-学习助手

To determine for which values of n the number $n^{13} - 2n^7 + n$ is divisible by 98, we need to show that it is divisible by both 2 and 49, since $98 = 2 \times 49$.

Step 1: Divisibility by 2

The expression $n^{13} - 2n^7 + n$ can be simplified in terms of parity (even or odd):

- n^{13} has the same parity as n because raising an integer to any power does not change its parity.
- $2n^7$ is always even because it is multiplied by 2.
- n has the same parity as n .

Thus, the expression $n^{13} - 2n^7 + n$ simplifies to $n - n \pmod{2}$, which is always even.

Therefore, $n^{13} - 2n^7 + n$ is always divisible by 2.

Step 2: Divisibility by 49

To check divisibility by 49, we use Fermat's Little Theorem, which states that if p is a prime number and n is not divisible by p , then $n^{p-1} \equiv 1 \pmod{p}$. For $p = 7$, this means $n^6 \equiv 1 \pmod{7}$ if n is not divisible by 7.

We factorize the polynomial:

$$n^{13} - 2n^7 + n = n(n^6 - 1)^2$$

Case 1: n is not a multiple of 7

If n is not a multiple of 7, then $n^6 \equiv 1 \pmod{7}$, so $(n^6 - 1) \equiv 0 \pmod{7}$. Therefore, $(n^6 - 1)^2 \equiv 0 \pmod{49}$, and hence $n(n^6 - 1)^2 \equiv 0 \pmod{49}$.

Case 2: n is a multiple of 7

If n is a multiple of 7, then $n \equiv 0 \pmod{7}$, and thus $n(n^6 - 1)^2 \equiv 0 \pmod{7}$. Since n is a multiple of 7, $n^{13} - 2n^7 + n$ is divisible by $7^2 = 49$.

Conclusion

The expression $n^{13} - 2n^7 + n$ is divisible by 98 if and only if n is either divisible by 49 or not divisible by 7.



定制版南科招生AI



什么样的学生适合报考南科大？



南科招生AI

Thought for 15 seconds

根据南方科技大学的招生政策和培养目标，以下学生更适合报考南科大：

学业成绩优秀：南科大注重学生的学术能力，尤其是理科方向（如物理、化学等）。学生需要在高考中取得优异成绩，同时具备扎实的学科基础。

学科特长和创新潜质：南科大特别关注在数学、物理、化学等学科上有特长的学生，或在科技创新、学术研究等方面有突出表现的学生。

独立思考和创新能力：南科大强调培养具有创新能力和批判性思维的学生，适合那些乐于探索、勇于尝试新事物的学生。

家国情怀和全球视野：南科大致力于培养具有家国情怀和全球视野的学生，适合那些关心社会问题、有志于服务国家，同时具有国际化视野的学生。

适应综合评价录取模式：南科大采用“631”综合评价录取模式，学生需要在高考、能力测试（机试和面试）以及高中平时成绩中均表现良好。

对南科大教育理念认同：南科大注重“创知、创新、创业”的教育特色，适合那些认同学校理念，愿意在南科大提供的学术环境中成长的学生。

如果你符合以上特点，南科大将是一个很好的选择！





内容提纲



- ✓ **RAG简介与应用**：记忆无痕，问道有源
- ✓ **RAG原理与流程**：横看成岭，侧见成峰
- **RAG向量数据库**：载舟亦水，活水为源
- **RAG上手与实战**：手中有剑，心中有数



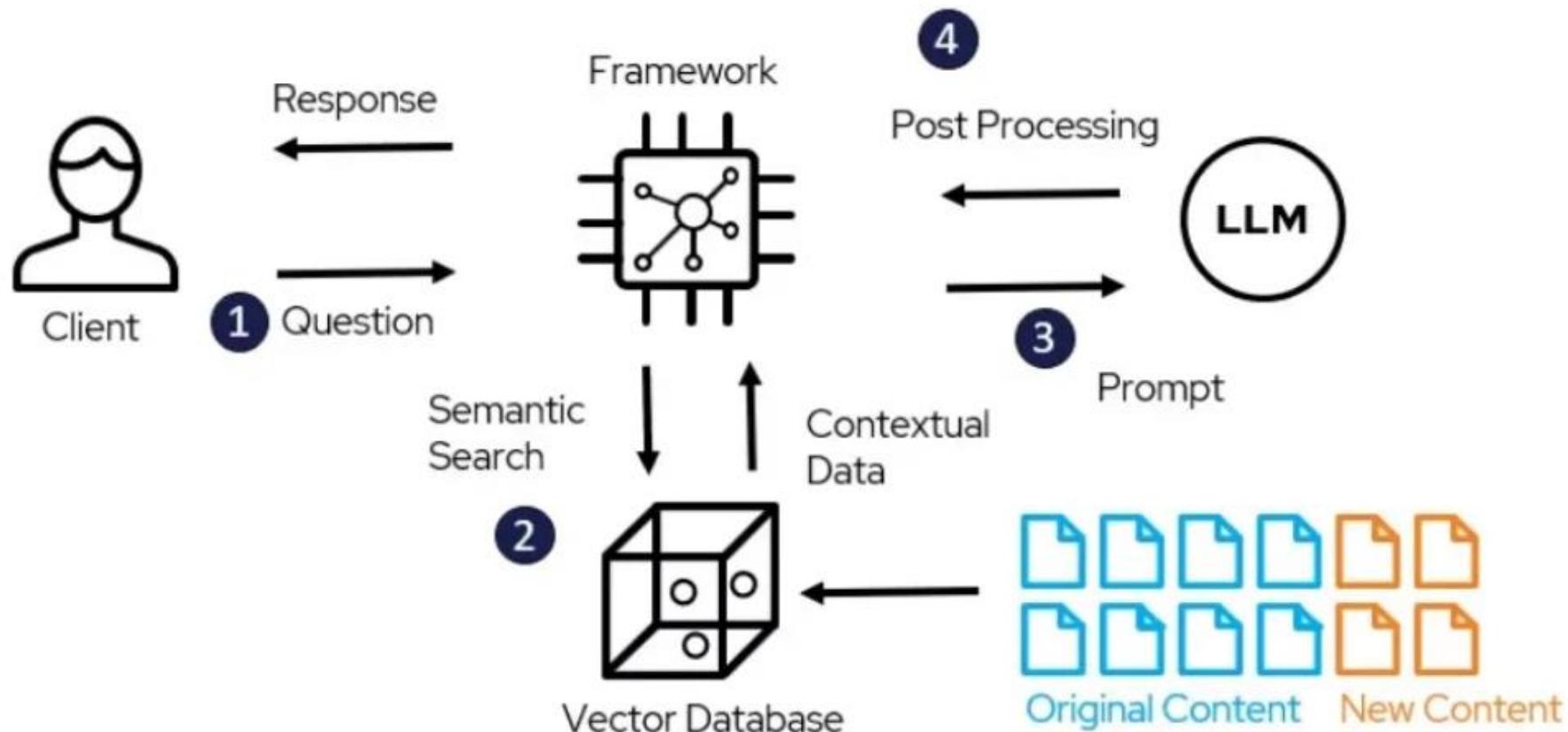
RAG原理与模式



你想拥有**为你定制服务的DeepSeek吗？**

RAG原理

□ 如何根据你的问题从你的数据中提取出最相关的内容 ?





RAG原理



□ 如何从**你的数据**中提取出**最相关的提示词**增强**你**的问题？

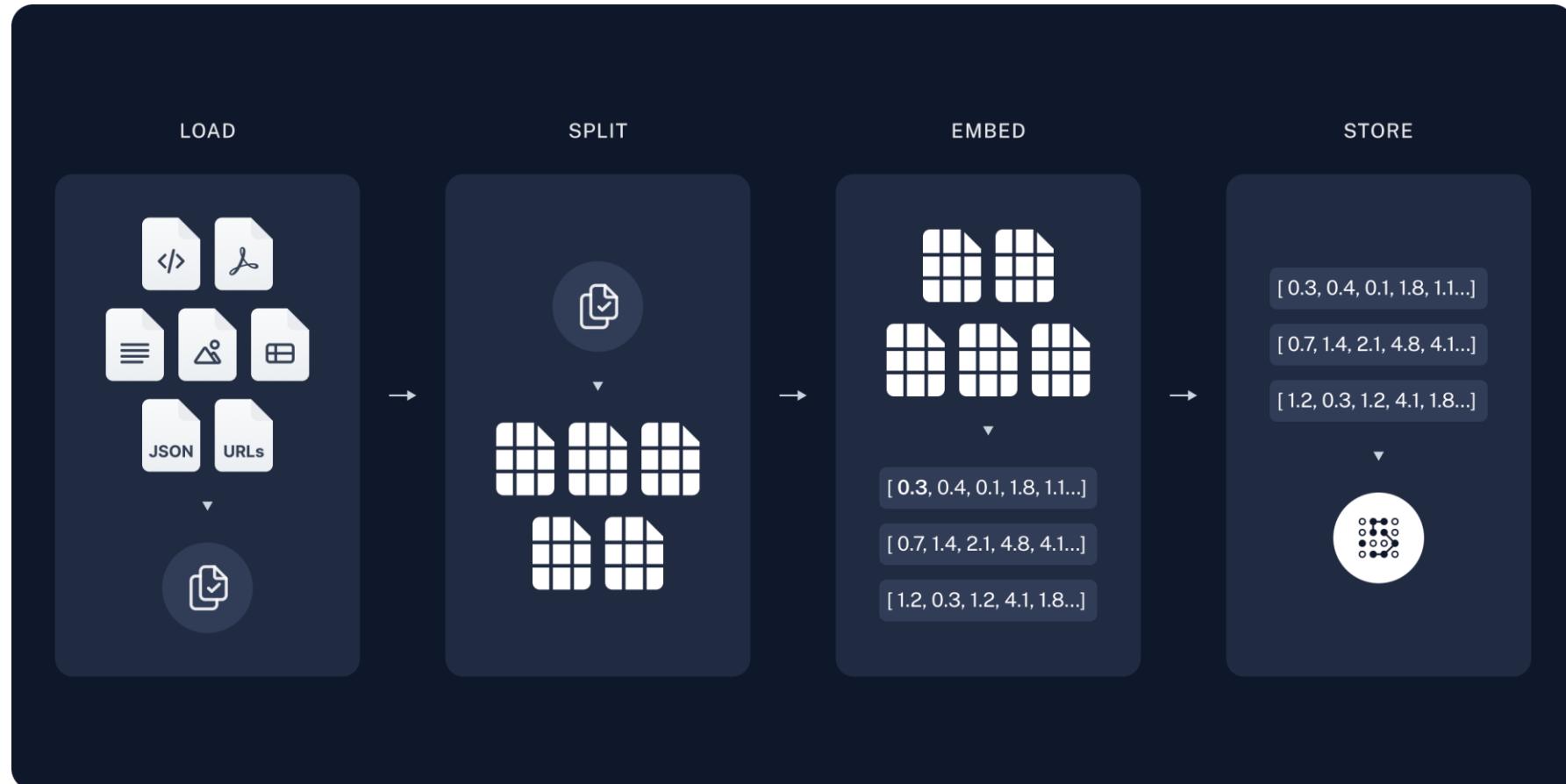
- **你的数据** → 如何**存储** → 向量化
- **最相关的** → 如何**检索** → 相似度
- **你**的问题 → 如何**增强** → 拼起来



RAG数据处理流程



□ 数据处理 : 分片 (Chunking), 嵌入 (Embedding)



Pic from: <https://python.langchain.com/docs/tutorials/rag/>

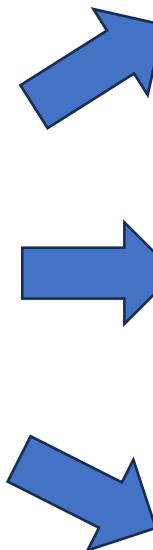


Chunking



□ 分片 (Chunking) 就是划分段落

南方科技大学(简称:南科大)是深圳在中国高等教育改革发展的时代背景下创建的一所高起点、高定位的公办新型研究型大学。2022年2月14日,教育部等三部委公布第二轮"双一流"建设高校及建设学科名单,南方科技大学及数学学科入选"双一流"建设高校及建设学科名单。学校借鉴世界一流理工科大学的学科设置和办学模式,以理、工、医为主,兼具商科和特色人文社科的学科体系,在本科、硕士、博士层次办学,在一系列新的学科方向上开展研究,使学校成为引领社会发展的思想库和新知识、新技术的源泉。



Chunk 1

南方科技大学(简称:南科大)是深圳在中国高等教育改革发展的时代背景下创建的一所高起点、高定位的公办新型研究型大学。

Chunk 2

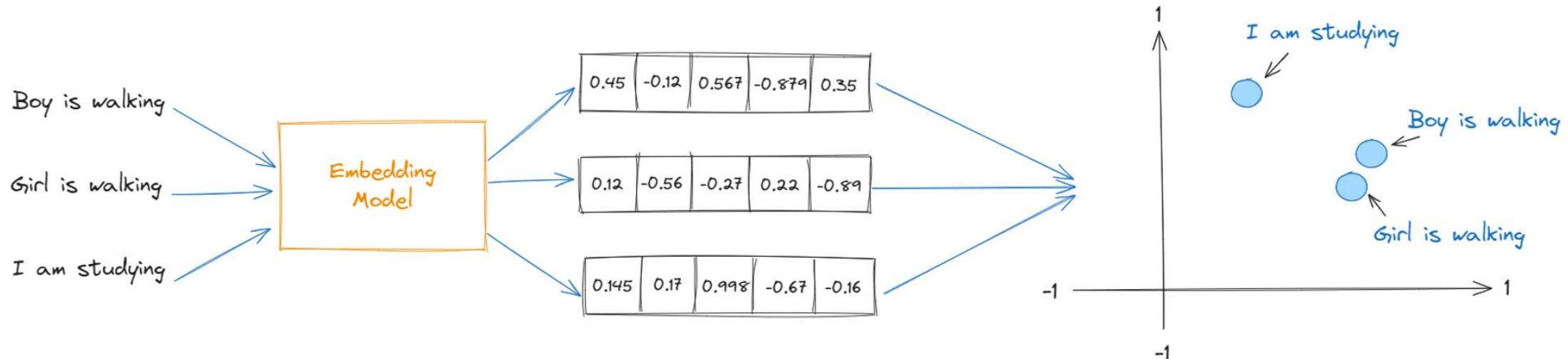
2022年2月14日,教育部等三部委公布第二轮"双一流"建设高校及建设学科名单,南方科技大学及数学学科入选"双一流"建设高校及建设学科名单。

Chunk 3

学校借鉴世界一流理工科大学的学科设置和办学模式,以理、工、医为主,兼具商科和特色人文社科的学科体系,在本科、硕士、博士层次办学,在一系列新的学科方向上开展研究,使学校成为引领社会发展的思想库和新知识、新技术的源泉。

Embedding

□ AI的世界：存在即向量



向量空间中，
距离越近内容越相关

RAG 中的向量检索

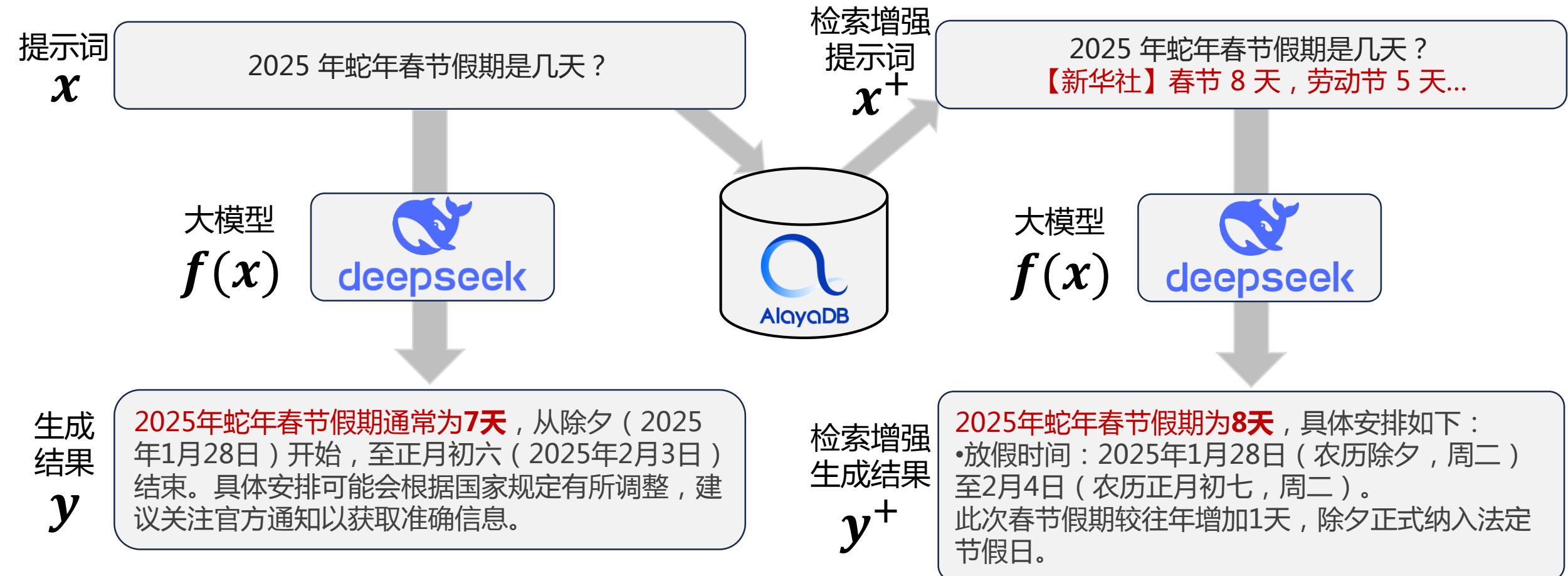
- 向量检索：找到向量空间中最近的K个向量

Vector search returns similar items **based on their semantic meaning**, rather than exact term matches.



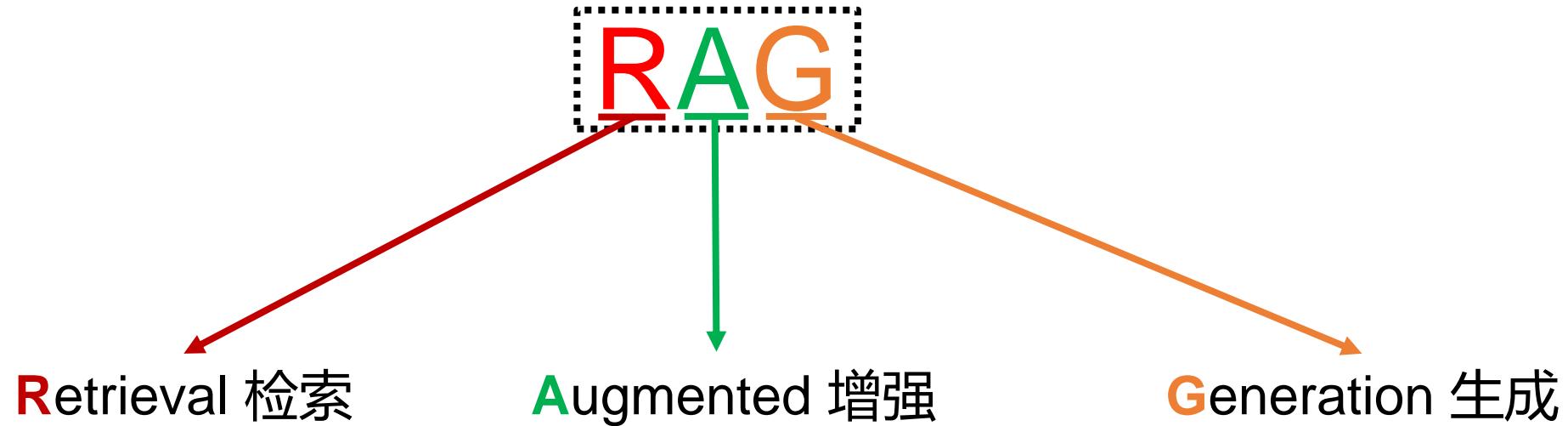


RAG 中的增强



增强：将向量空间中距离最近的向量所对应的文本添加到提示词后

检索增强生成 RAG



- **检索**：从向量空间中找到距离最近的向量
- **增强**：把距离最近向量对应的文本添加到你的提示词后面
- **生成**：大模型根据增强后的提示词生成答案



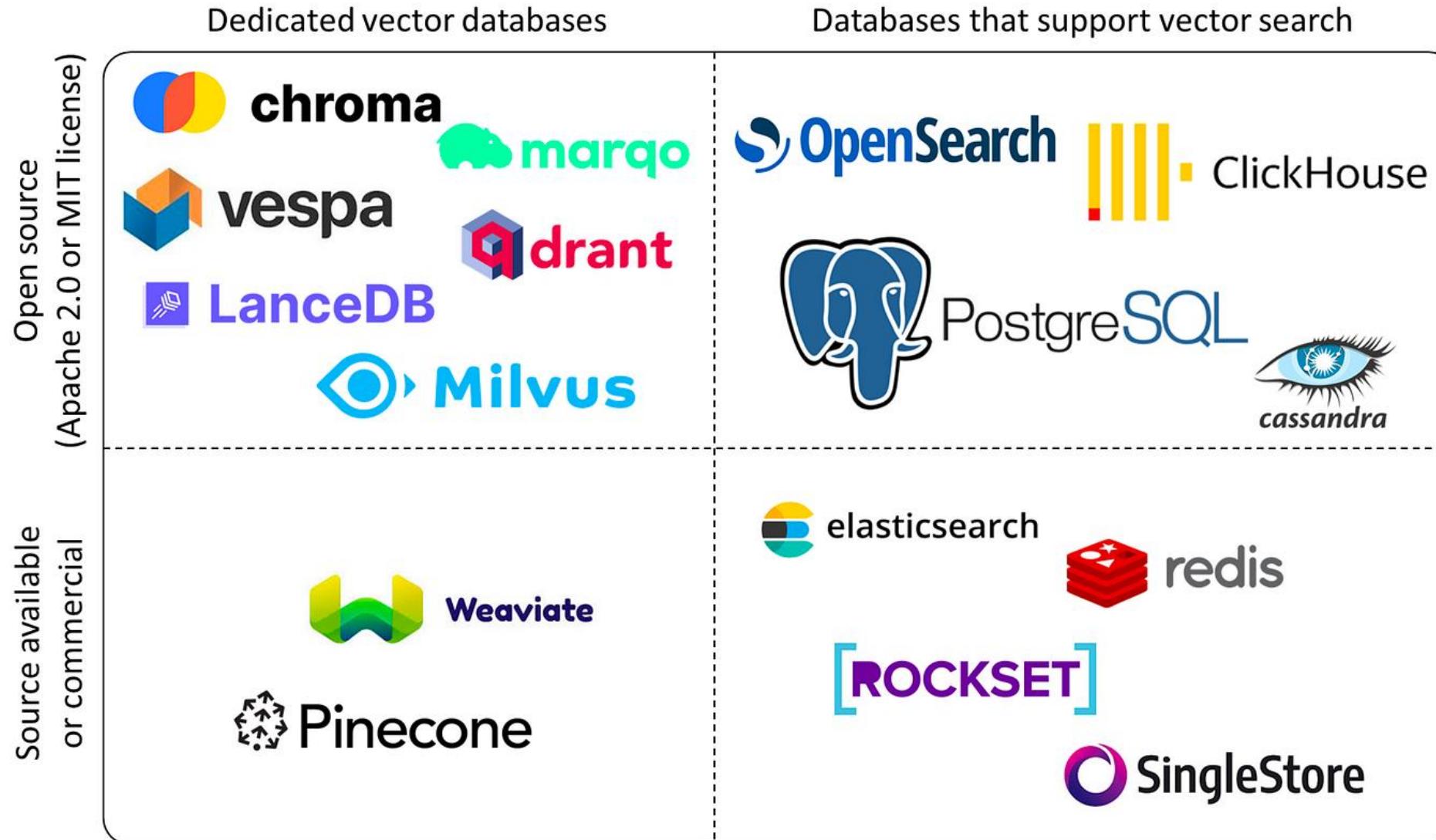
内容提纲



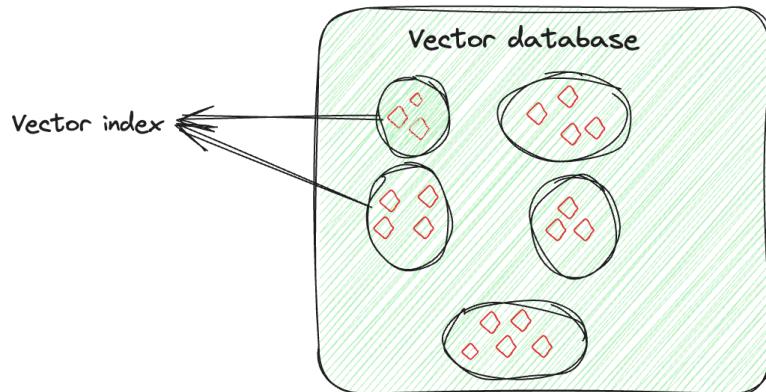
- ✓ **RAG简介与应用**：记忆无痕，问道有源
- ✓ **RAG原理与流程**：横看成岭，侧见成峰
- ✓ **RAG向量数据库**：载舟亦水，活水为源
- **RAG上手与实战**：手中有剑，心中有数



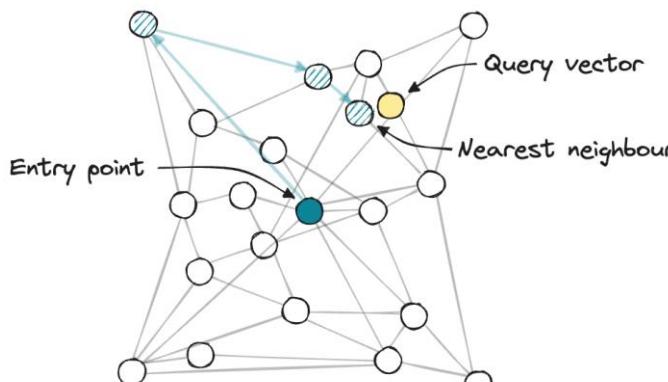
向量数据库：大模型的海马体



向量数据库核心能力



索引构建与更新



查询处理与优化

Compute Unit

CPU, FPGA, GPU, DSA

New Storage

DRAM, SSD, PMEM

Network

RDMA, CXL, Ethernet



AlayaLite: 更快,更准,更好



AlayaLite – A Fast, Flexible Vector Database for Everyone.
Seamless Knowledge, Smarter Outcomes.

Release Ver0.1.0a1 pypi v0.1.0a1 last commit today

Features

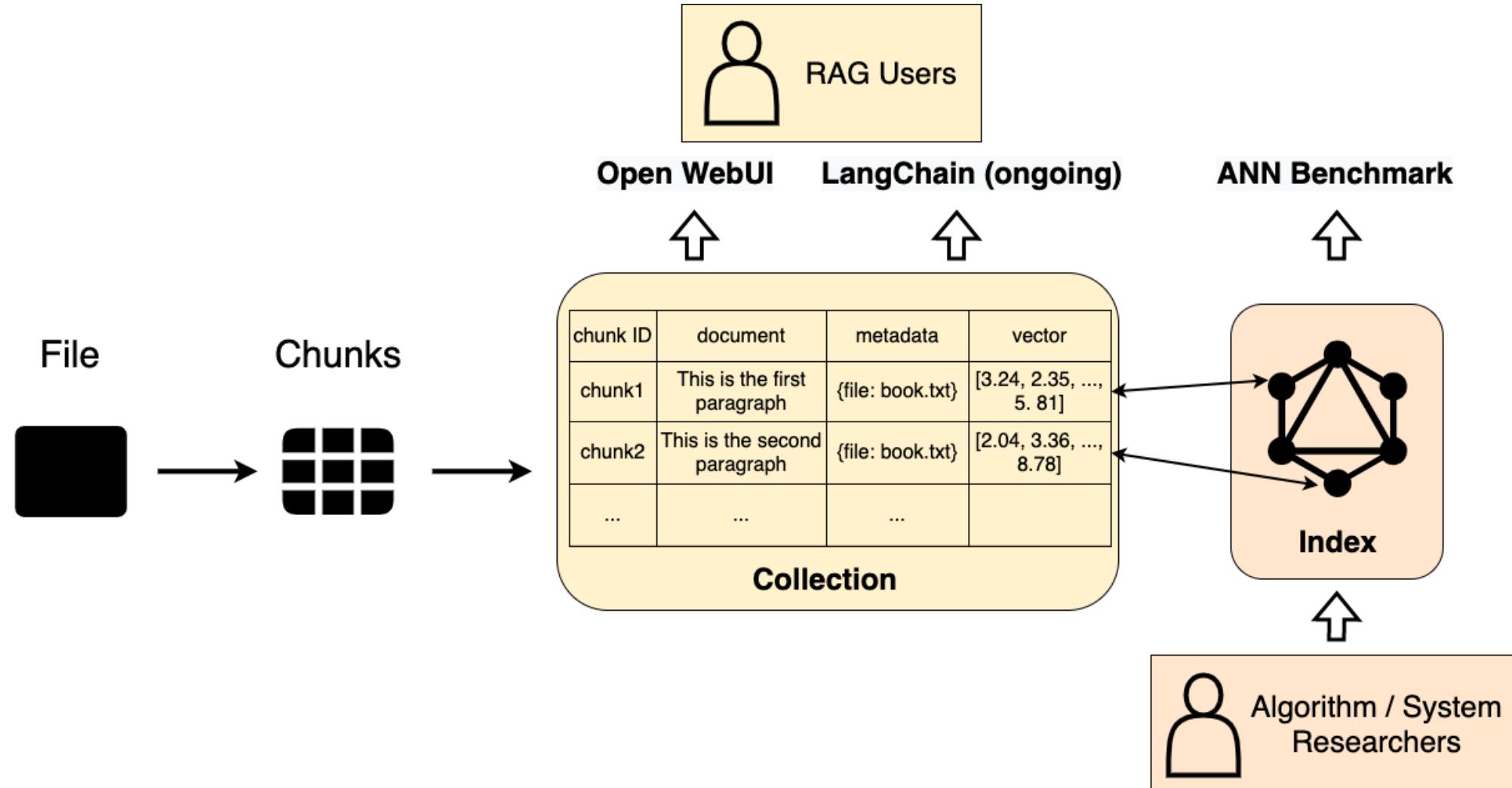
- **High Performance:** Modern vector techniques integrated into a well-designed architecture.
- **Elastic Scalability:** Seamlessly scale across multiple threads, which is optimized by C++20 coroutines.
- **Adaptive Flexibility:** Easy customization for quantization methods, metrics, and data types.
- **Ease of Use:** [Intuitive APIs](#) in Python.

DBG Group开源系统: <https://github.com/AlayaDB-AI/AlayaLite>

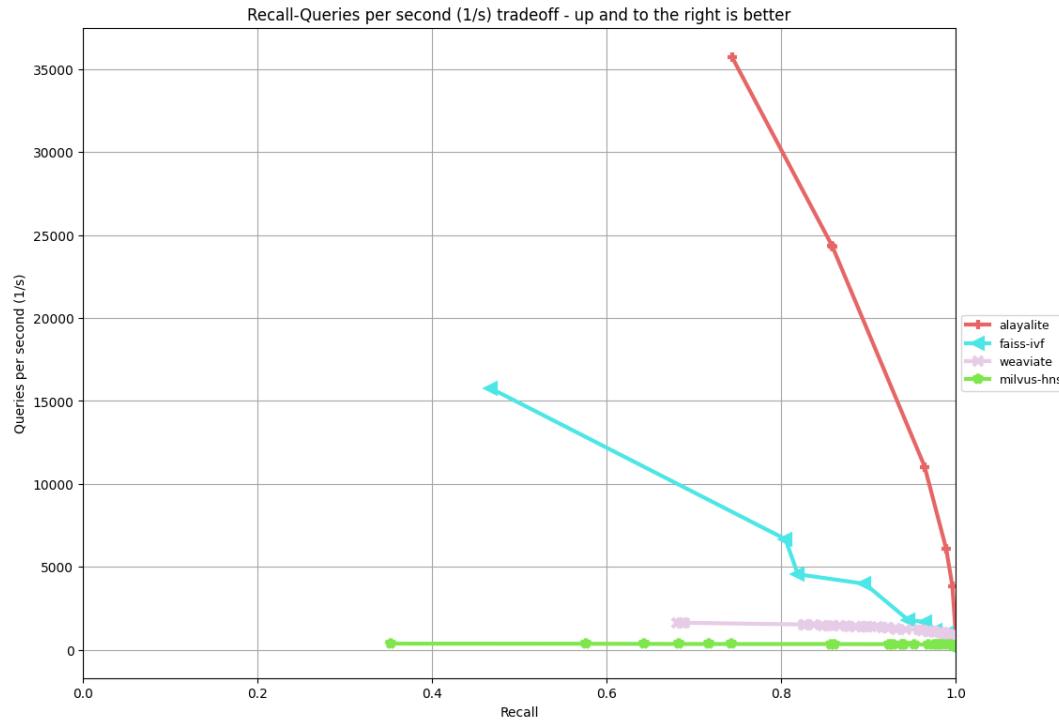
更快、更准、更好的向量数据库系统！



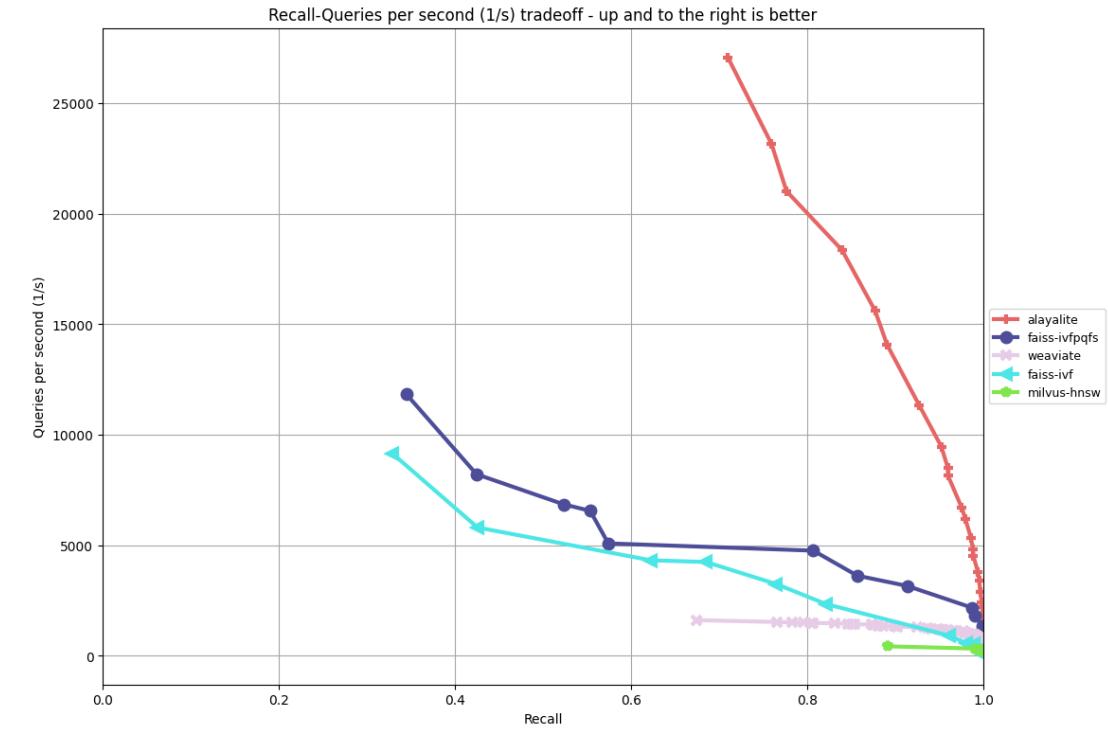
AlayaLite 用户群体



AlayaLite性能表现



Glove-25-angular

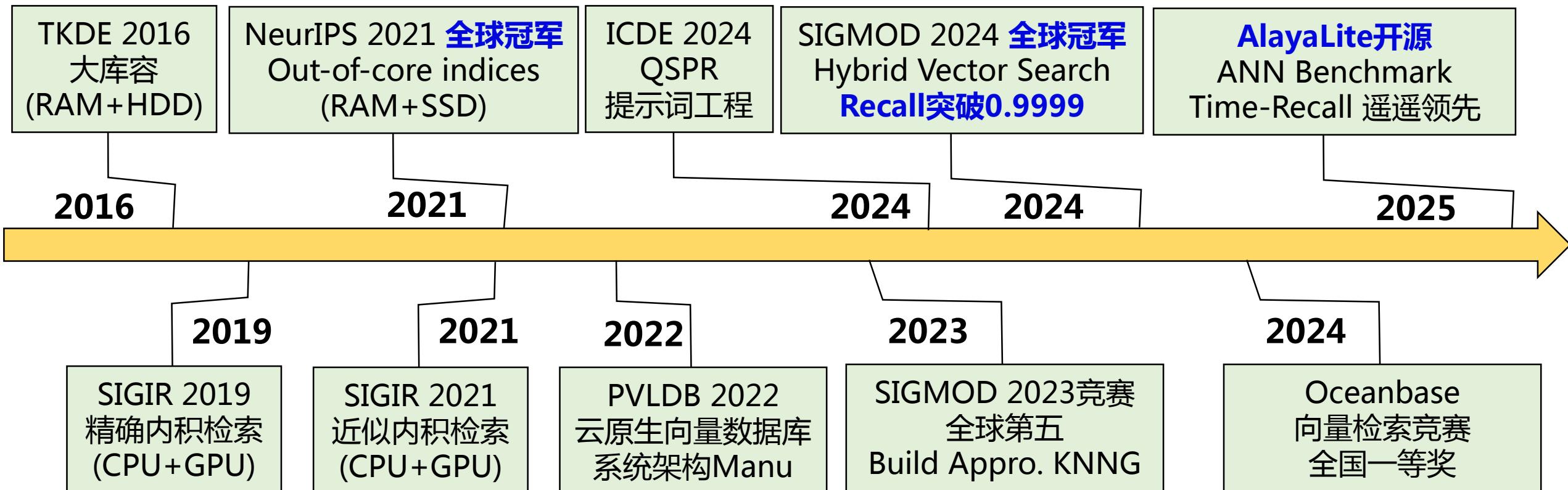


sift-128-euclidean

向量数据库系统中AlayaLite性能遥遥领先！！！



十年耕耘，厚积薄发



两个全球冠军，多篇CCF-A类论文，技术涵盖大库容到异构计算



内容提纲



- ✓ **RAG简介与应用**：记忆无痕，问道有源
- ✓ **RAG原理与流程**：横看成岭，侧见成峰
- ✓ **RAG向量数据库**：载舟亦水，活水为源
- ✓ **RAG上手与实战**：手中有剑，心中有数



RAG上手与实践

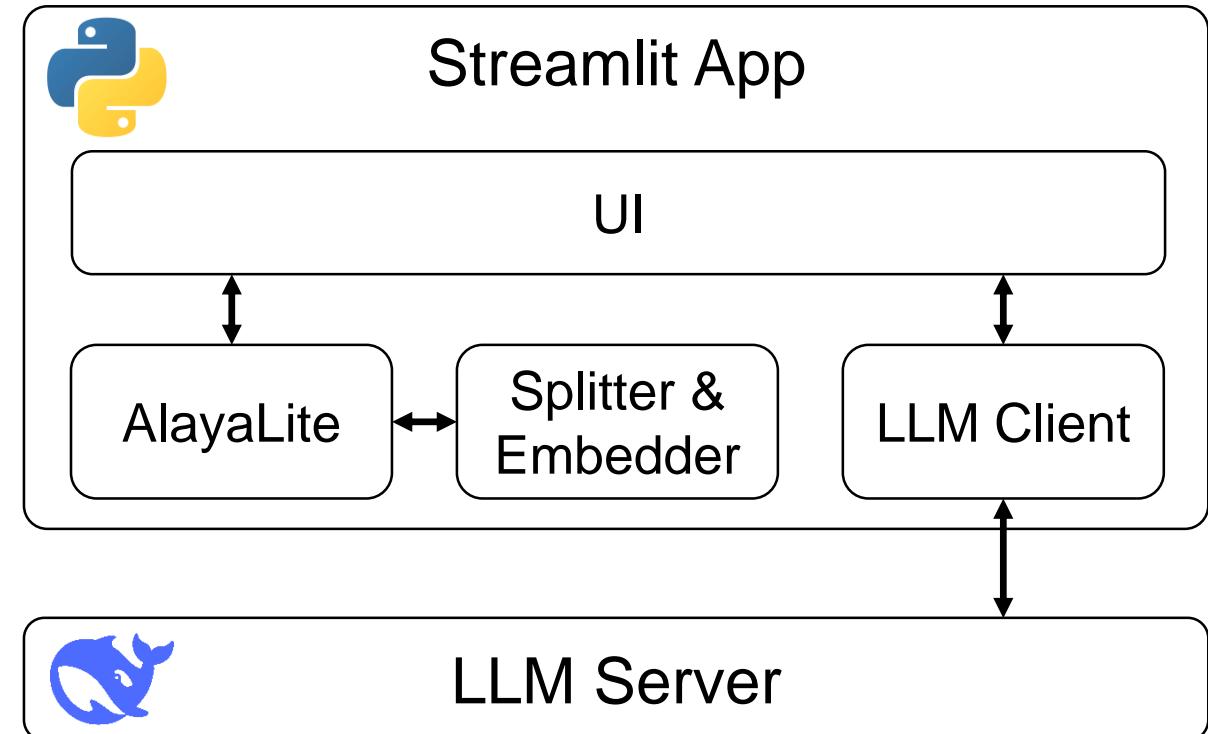


你想拥有为你定制服务的DeepSeek吗？

现在是时候了！

Tutorial: 部署自己的 RAG 应用

- ① 获取代码仓库
- ② 创建Codespace
- ③ 配置环境
- ④ 运行应用





Step 1: 获取代码仓库



- 在Github上Star并且Fork代码仓库

<https://github.com/AlayaDB-AI/AlayaLite>

Star 然后 Fork 仓库到自己的帐号上

The screenshot shows the GitHub repository page for 'AlayaLite'. At the top right, there are buttons for 'Edit Pins', 'Unwatch', 'Fork', and 'Starred'. Red arrows point from the text 'Star 然后 Fork 仓库到自己的帐号上' to the 'Star' and 'Fork' buttons. The repository has 2 stars and 0 forks. The main area displays a list of commits by 'haotian-l' with messages like 'chore: initial commit'. On the right side, there are sections for 'About', 'Activity', 'Custom properties', 'Watching', 'Forks', and 'Report repository'. Below that are sections for 'Releases' and 'Packages'.

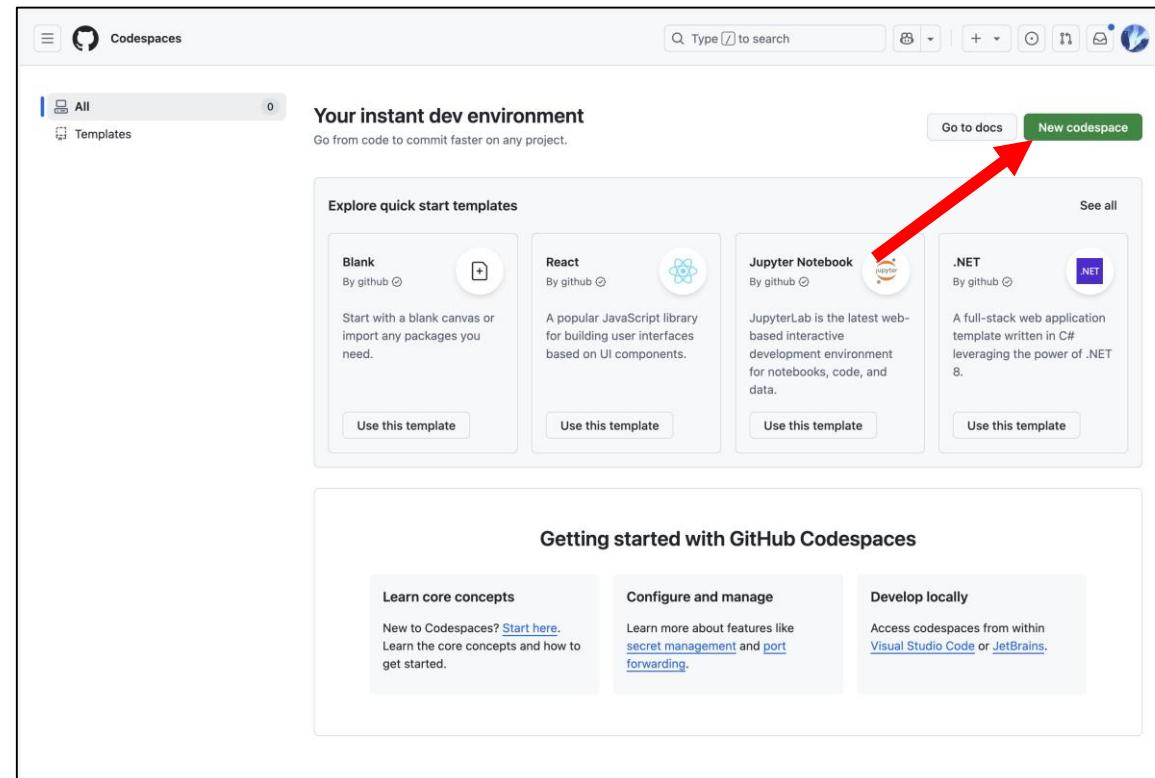


Step 2: 创建 Codespace



- 在Github Codespaces上创建开发环境

<https://github.com/codespaces/>





Step 2: 创建Codespace



□ 在Github Codespaces上创建开发环境

Create a new codespace

Repository
To be cloned into your codespace
DBGroup-SUSTech/AlayaLite

Branch
This branch will be checked out on creation
master

Region
Your codespace will run in the selected region
Southeast Asia

Machine type
Resources for your codespace
4-core

Create codespace

选择AlayaLite

选择master分支

可选2-core或4-core

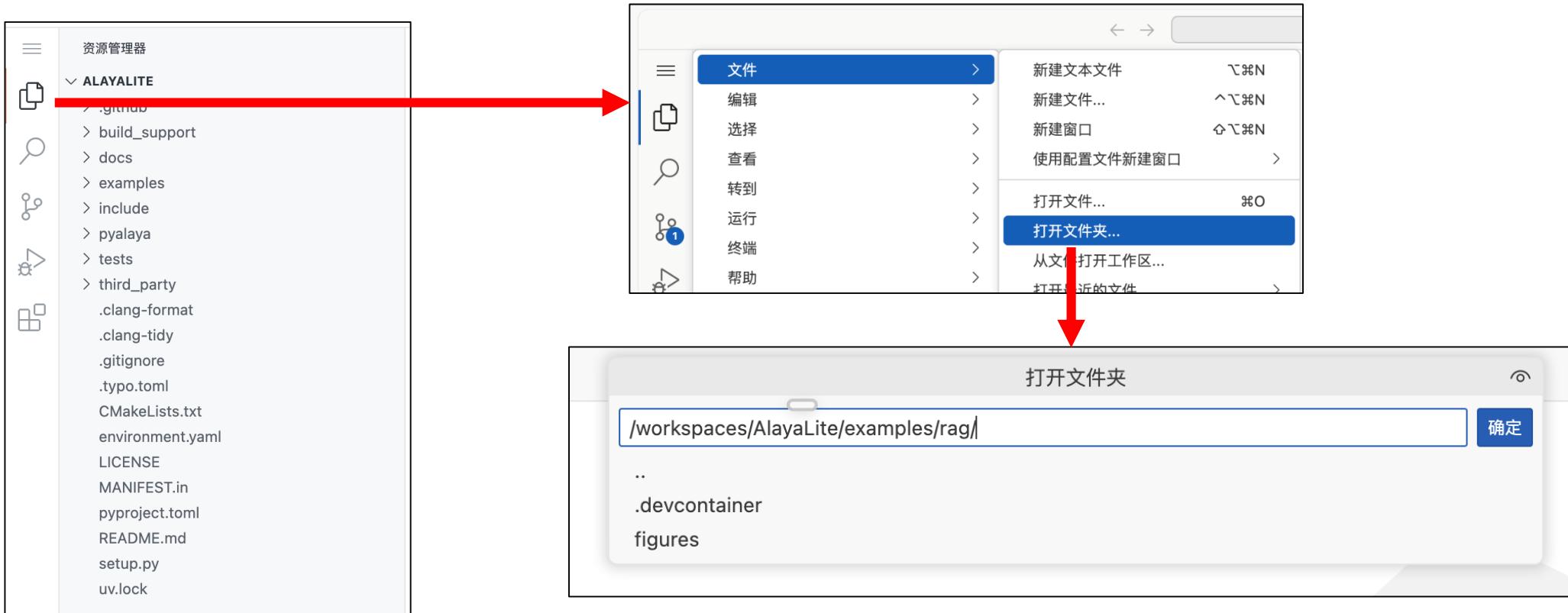
最后点击创建



Step 3: 配置环境

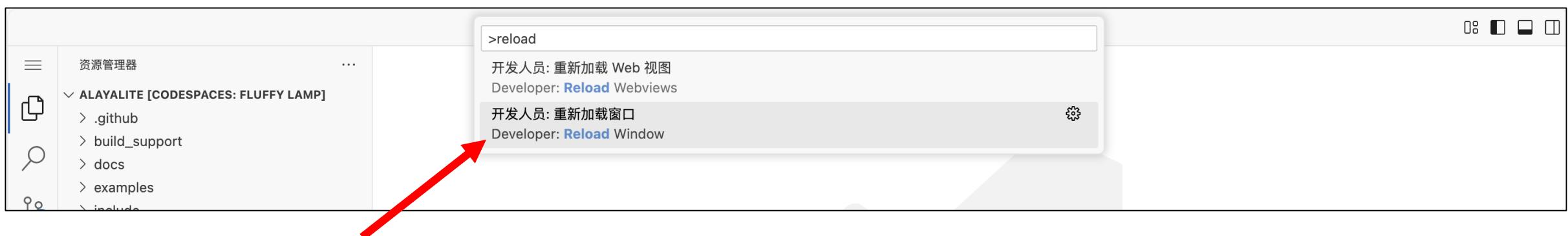


□ 进入RAG示例文件夹



Step 3: 配置环境

- 参考README文档步骤1，完成环境配置
- 没有端口转发，可尝试重新加载工作区，以激活配置





Step 4: 运行应用



1. 配置API Key

2. 上传文档

可使用项目文件夹中的示例文档test_docs.txt

3. 处理文档

RAG QA

What are higher-order chunking techniques?

Higher-order chunking techniques include small-to-bigger and sliding windows, which are advanced methods for segmenting content into meaningful chunks.

References

4. 发起提问。示例问题：
What are advanced chunking techniques?

临时Key: sk-vyeqeODAk4AhUXMQbD0zaY7EA8PCcc5zL6p2O2tTOYxTmtfc , 你也可以用你自己的



Step 5: 专属知识库问答



南科大2025年寒假图书馆开放情况？

南科大2025年寒假图书馆开放情况？

已搜索到 46 个网页 >

根据搜索结果，目前没有找到关于南方科技大学（南科大）2025年寒假图书馆开放情况的具体信息。不过，可以参考其他高校图书馆的寒假开放安排，通常寒假期间图书馆会采取部分开放或限时开放的模式，同时电子资源一般保持24小时开放 1 4 6。

如果您是南科大的师生，建议通过以下方式获取准确信息：

1. 访问南科大图书馆官网：通常图书馆会在官网发布最新的开放通知。
2. 关注图书馆官方微信公众号：许多高校图书馆会通过微信公众号推送重要通知。
3. 联系图书馆服务台：直接拨打图书馆的服务电话，获取寒假开放的具体安排。

如果您需要其他高校图书馆的寒假开放安排作为参考，可以查看搜索结果中的相关内容 1 2 6。

取消 取消 喜欢 分享



Service Setting

LLM Base URL: https://api.lkeep.cloud.tencent.cor

LLM API Key

LLM Model: deepseek-v3

Embedding Model: BAII/bge-small-zh-v1.5

Knowledge Base Management

Upload documents

Drag and drop files here
Limit 200MB per file • TXT, PDF, DOCX, MD

Browse files

图书馆2025年... x 20.3KB

Start processing

RAG QA

南科大2025年寒假图书馆开放情况？

南科大2025年寒假期间，琳恩图书馆和涵泳图书馆1月20日至2月12日闭馆。一丹图书馆1月13日至2月14日开放，其中1月16日至19日闭馆清洁，春节假期（1月28日至2月4日）仅开放一楼学习区，时间为8:00-22:00。图书和京东阅读本可延期至2月21日归还。

References ▾

Please enter your question... >





南方科技大学

SOUTHERN UNIVERSITY OF SCIENCE AND TECHNOLOGY



谢谢 !

DBGroup @ SUSTech

Dr. Bo Tang (唐博)

tangb3@sustech.edu.cn

