



2024年高校图书馆发展论坛



山东大学
SHANDONG UNIVERSITY

图书馆
LIBRARY

基于大语言模型的 图书馆学科服务转型研究

报告人：吴帼帼

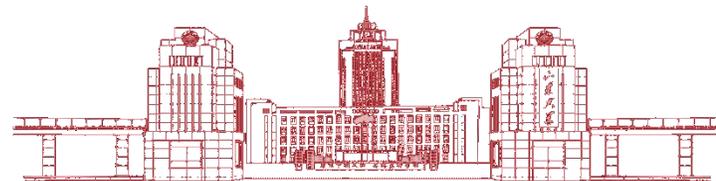
山东大学图书馆学科服务中心

2024年6月



目录

- 一、大语言模型起源与发展
- 二、高校学科服务现状分析
- 三、学科服务转型思路
- 四、基于开源LLMs的尝试





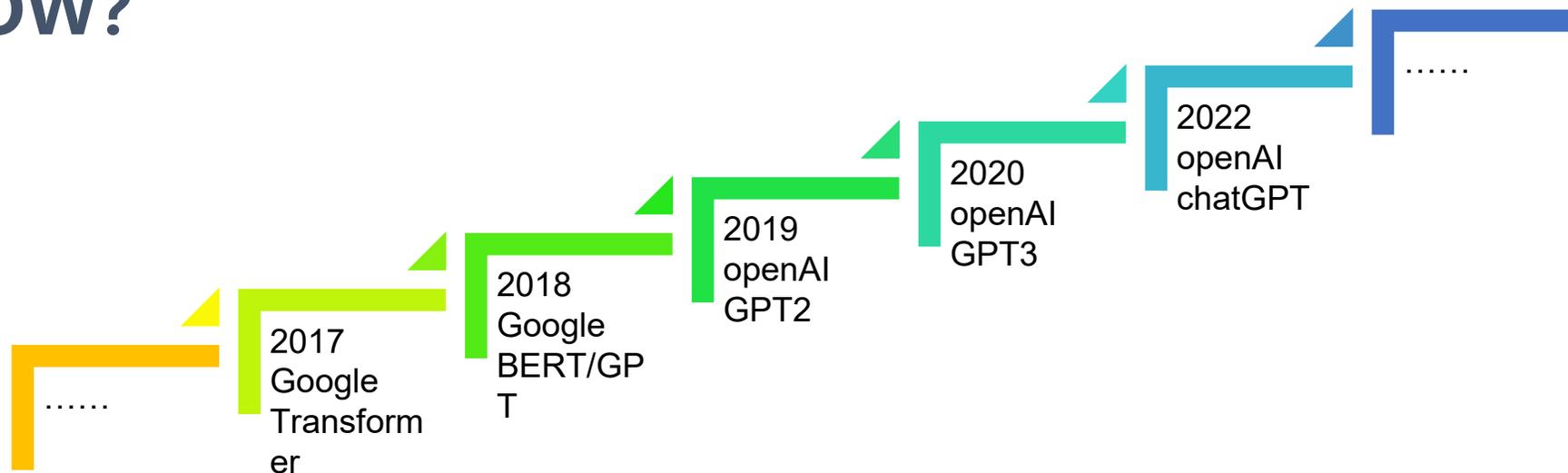
一、大语言模型

(1) WHAT ?

**GPT (Generative Pre-trained Transformer) 、
BERT (Bidirectional Encoder Representation from
Transformers) 、AIGC (AI-Generated Content)
RAG (Retrieval-Augmented Generation)**

大语言模型 (Large Language Models, 简称 LLMs)

(2) HOW?



一、大语

Attention Is All You Need

Ashish Vaswani*
Google Brain
avaswani@google.com

Noam Shazeer*
Google Brain
noam@google.com

Niki Parmar*
Google Research
nikip@google.com

Jakob Uszkoreit*
Google Research
uszko@google.com

Llion Jones*
Google Research
llion@google.com

Aidan N. Gomez* †
University of Toronto
aidan@cs.toronto.edu

Lukasz Kaiser*
Google Brain
lukaszkaizer@google.com

Illia Polosukhin* ‡
illia.polosukhin@gmail.com

Abstract

The dominant sequence transduction models are based on complex recurrent or convolutional neural networks that include an encoder and a decoder. The best performing models also connect the encoder and decoder through an attention mechanism. We propose a new simple network architecture, the Transformer, based solely on attention mechanisms, dispensing with recurrence and convolutions entirely. Experiments on two machine translation tasks show these models to be superior in quality while being more parallelizable and requiring significantly less time to train. Our model achieves 28.4 BLEU on the WMT 2014 English-to-German translation task, improving over the existing best results, including ensembles, by over 2 BLEU. On the WMT 2014 English-to-French translation task our model establishes a new single-model state-of-the-art BLEU score of 41.8 after training for 3.5 days on eight GPUs, a small fraction of the training costs of the best models from the literature. We show that the Transformer generalizes well to other tasks by applying it successfully to English constituency parsing both with large and limited training data.

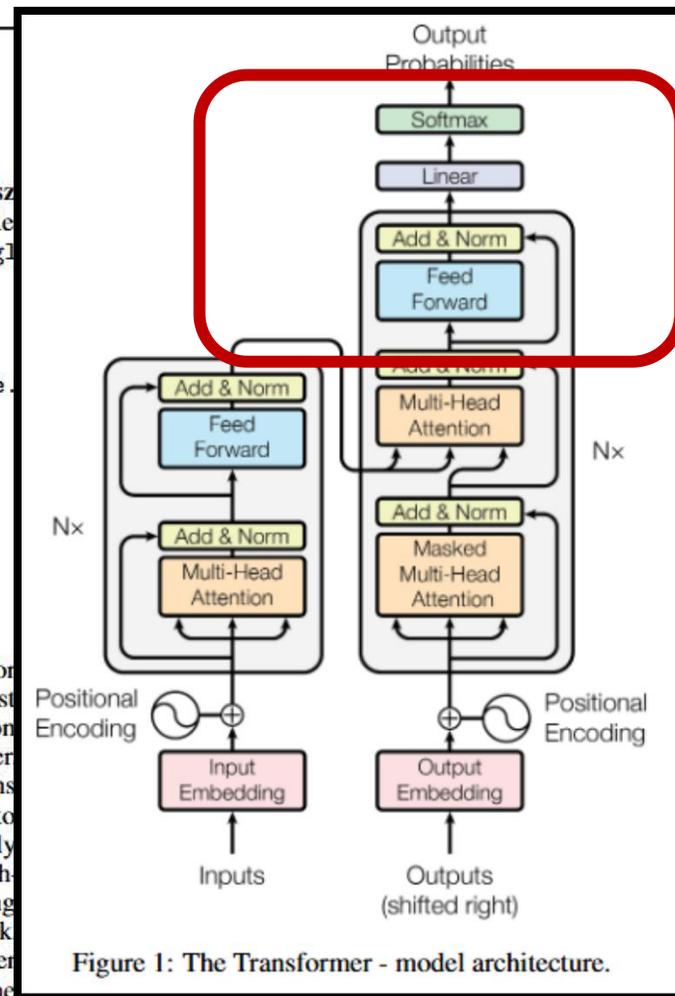


Figure 1: The Transformer - model architecture.

arXiv:1706.03762v7 [cs.CL] 2 Aug 2023

2017
Google
Transformer



一、大语言模型 (3) WHERE ?





二、高校学科服务现状分析

“双一流”高校学科服务现状调查

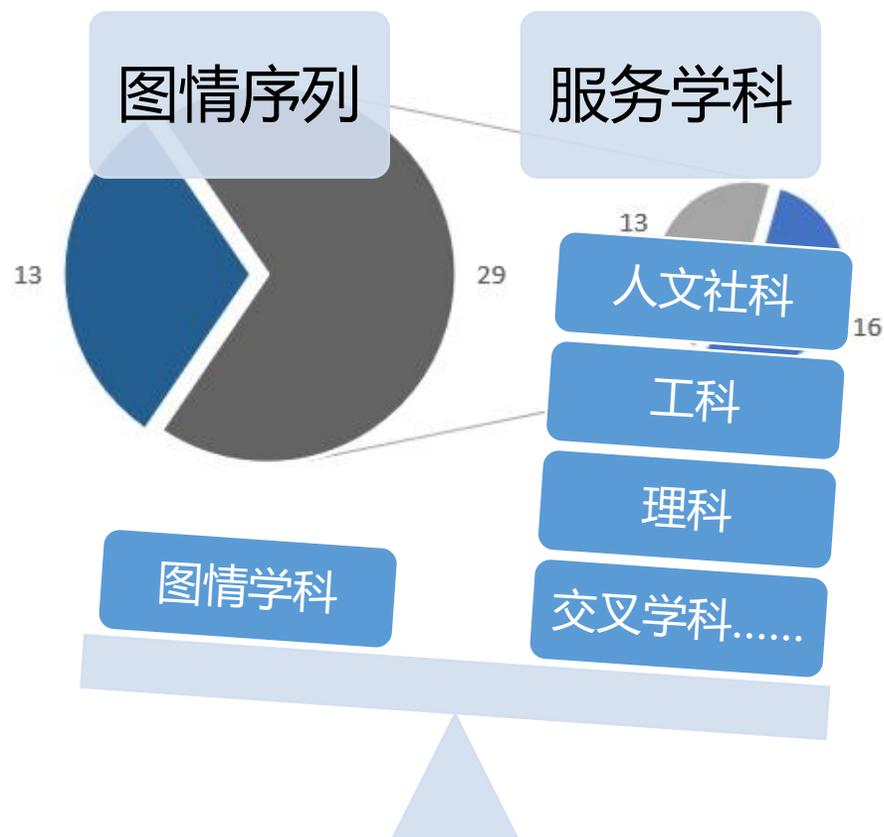
调研对象：“双一流”高校图书馆

调查时间：2024年1月—4月

调研项	调查项目详情
学科服务专栏	图书馆主页是否含有学科服务专栏
学科服务范畴	学科服务具体内容、与信息素养教学关系
服务学院情况	学科服务面向的院系、专业情况
学科馆员配置情况	学科馆员人数、是否有兼职馆员
学科服务报告	服务报告的内容、公开程度
学科服务平台	学科服务平台的使用率、主流平台

二、高校学科服务现状分析

困境一：学科服务队伍松散，学科馆员发展受限



“双一流”高校图书馆中，在图书馆主页上设有“学科服务”栏目的29所，服务内容中学科服务同时承担着信息素养教育培训的有16所。

人员配置上：Max37 Min9 专职 VS 兼职（南开、山大）

服务院系上：Max52 Min15

人均对口服务院系数：Max 5 Min1.162





二、高校学科服务现状分析

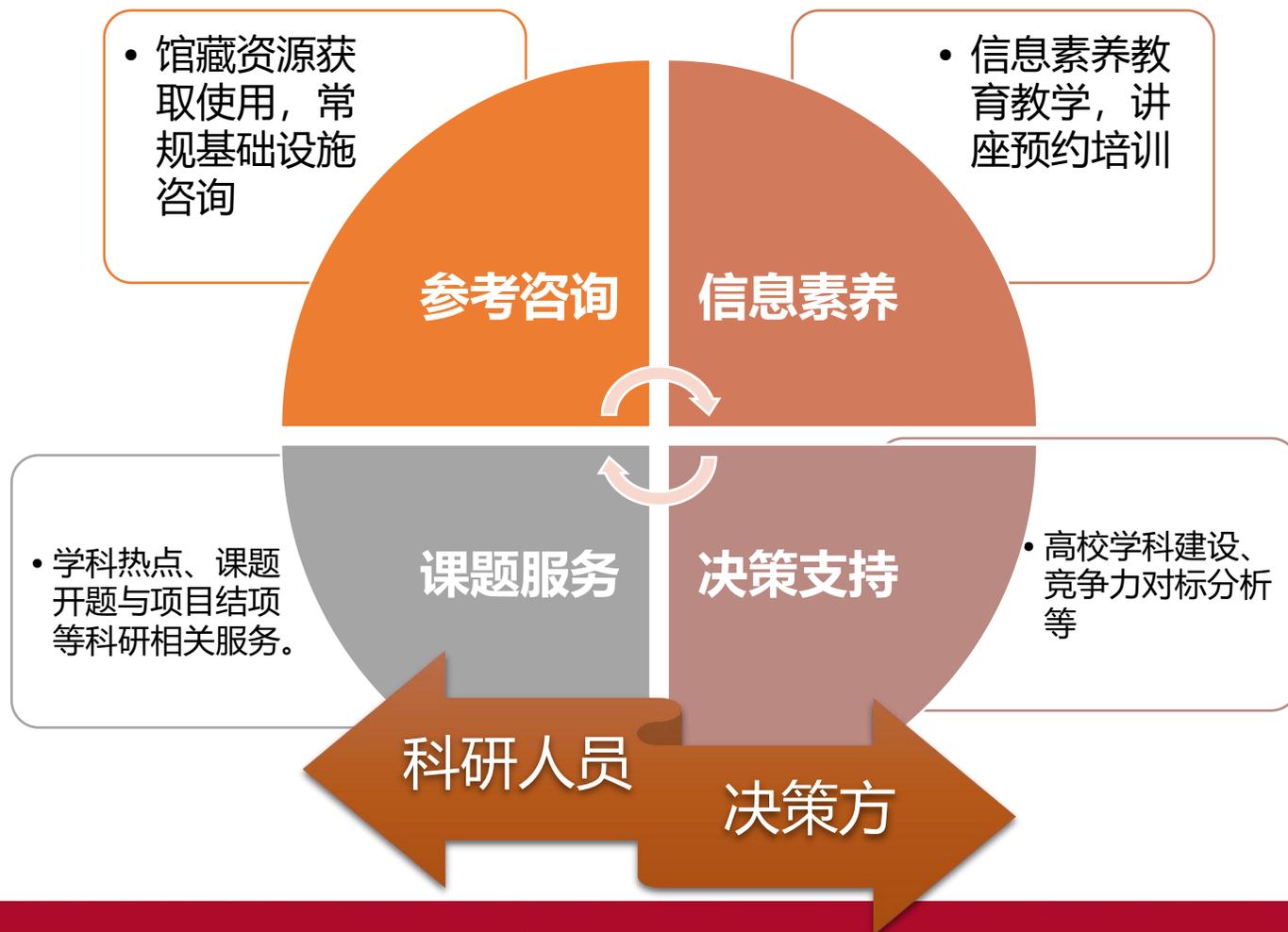
困境二：传统学科服务平台，交互性差利用率低





二、高校学科服务现状分析

困境三：服务内容基础边界模糊，重点难把握



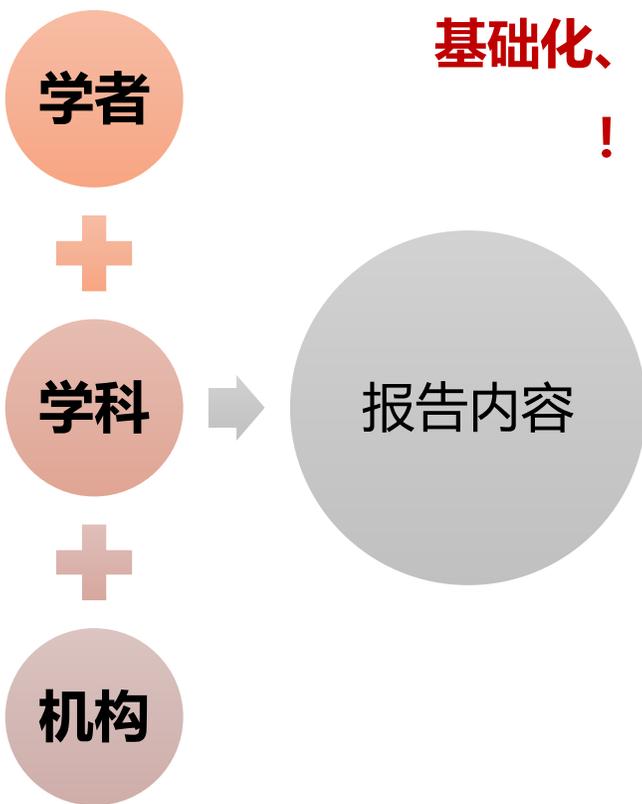
学科服务内容词频云图

二、高校学科服务现状分析

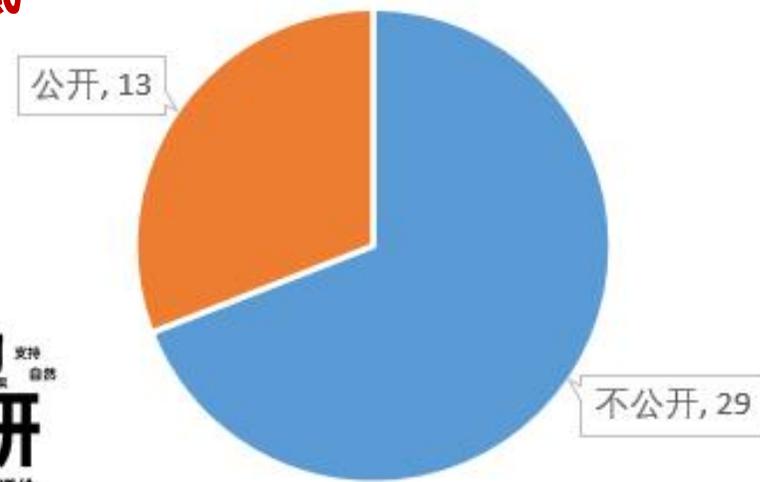
困境四：服务报告开放程度低，内容范式化

基础化、范式化、可替代性强、认可度低

！！闭门造车、大环境封闭



公开, 13



不公开, 29

服务报告内容词频云图

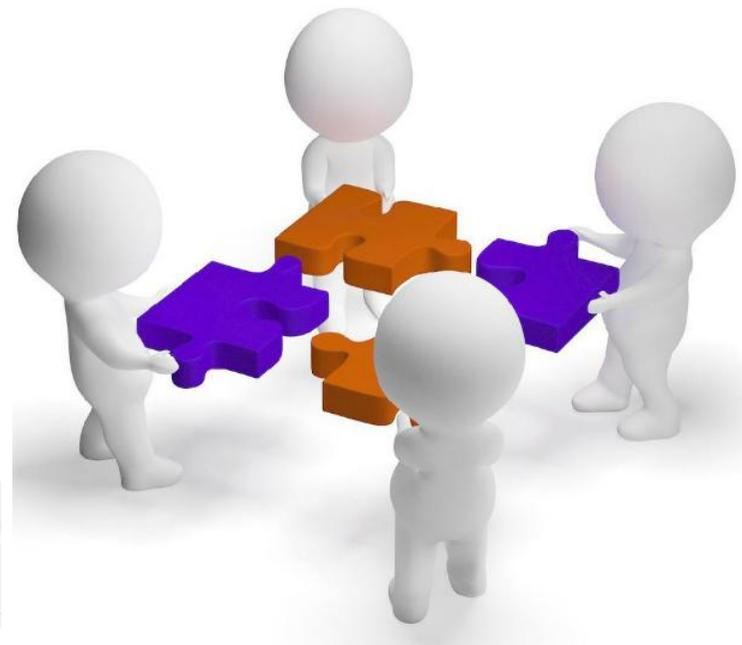


三、学科服务转型思路

3.1 政策规范与体系建设

国内外高校都在开展学科知识服务，高度重视学科发展建设

1. 高校图书馆、校科研院、学科发展中心都在关注学科动态，不同高校、同一高校的不同职能部门都存在一定程度上的重复交叉劳动
2. 关于学科服务报告目前业界尚无规范，业内开放交流环境较差。

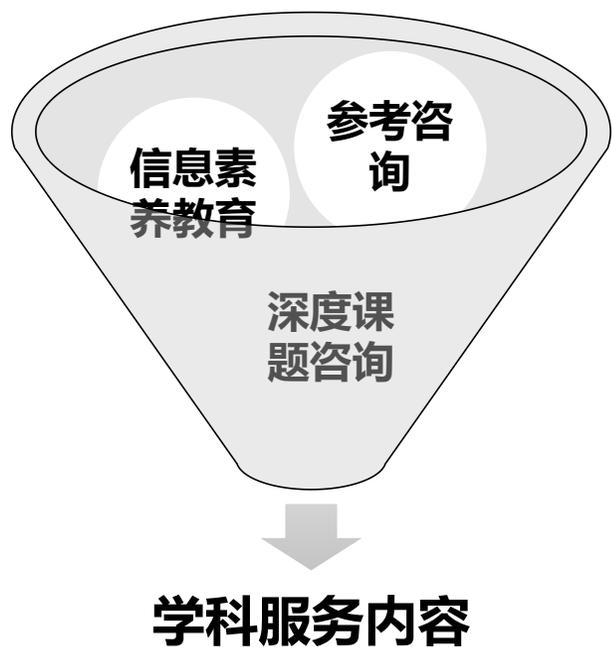


积极打破壁垒，摒弃“闭关锁门”，有意识的加强业界开放交流。推进服务体系与行业规范建设，共建共享，构建内部开放的服务联盟



三、学科服务转型思路

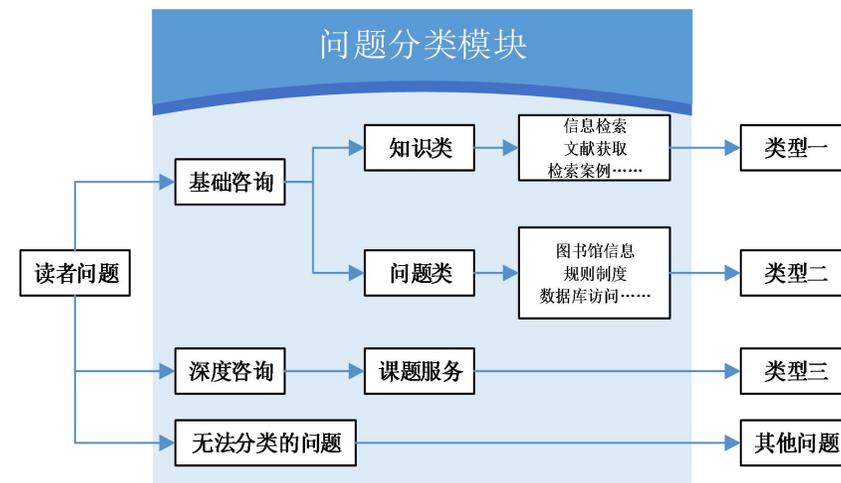
3.2 问题分类，建立规范化的分层服务体系



参考咨询——问题类

信息素养教育——知识类

深度课题咨询——定题情报类



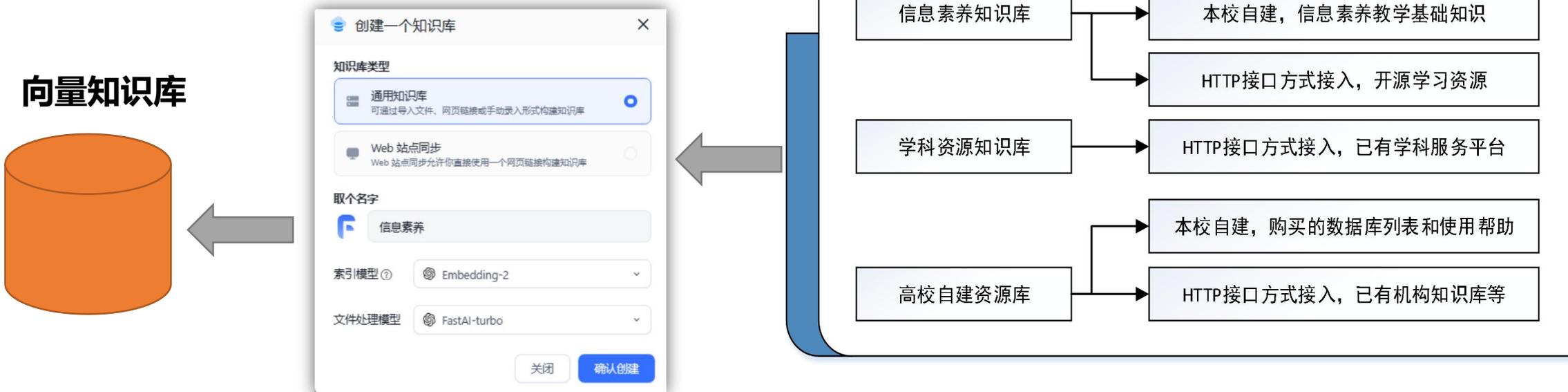
三、学科服务转型思路

3.3 三位一体，打造学科知识自主学习平台

国内高校学科服务平台的功能可分为联络属性、宣传服务、学科资源推荐等。

基于LLMs建立学科知识学习平台包含三个子库：

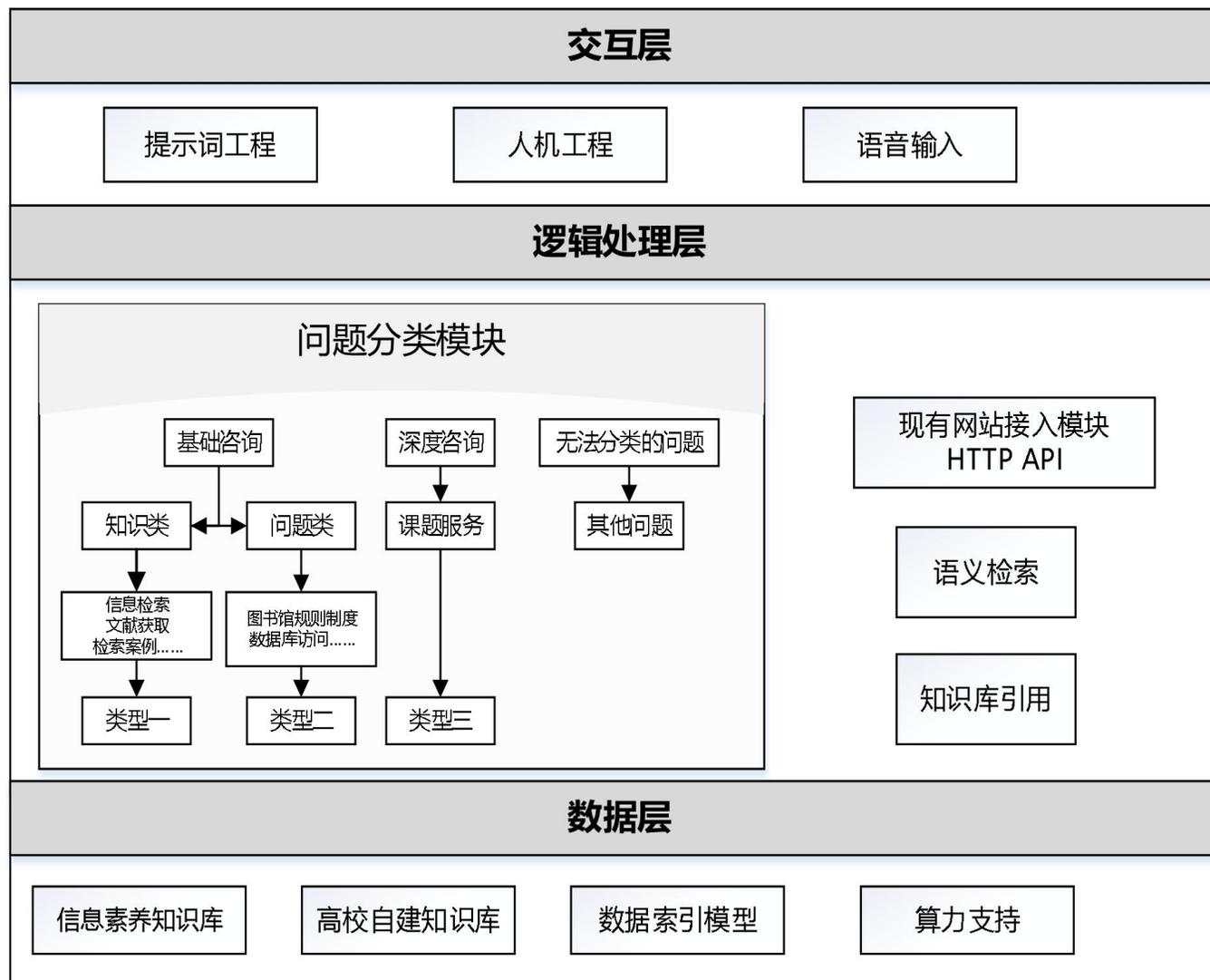
- 问题类——图书馆基础设施服务库，用于处理基咨询
- 知识类——信息素养知识库，用于解决教学任务重
- 定题情报服务类——学科资源知识库





四、基于开源LLMs的尝试

框架设计



四、基于开源LLMs

libLLMs

中国法分类数学属于什么类 6条记录 FastAI-3.5, FastAI-3.5, FastAI-3.5

你好, 我是lib助手, 请不要忘记选择知识库-

- 开放资源
- 知识学习
- 课程资源

文献是指记录知识的载体, 文献可以包括书籍、期刊、报纸、论文、手册等各种形式, 用

引用

手动录入

1条引用 2条上下文 1.02s 查看详情

课程资源咨询

吴老师 理科学科服务 ssc001@sdu.edu.cn
刘老师 工科学科服务 ssc002@sdu.edu.cn
李老师 社科学科服务 ssc003@sdu.edu.cn
冯老师 人文学科服务 ssc004@sdu.edu.cn

1.3s 查看详情

根据中国法(中国图书馆分类法)的分类体系, 数学属于类别“O”的自然科学和数学类,

引用

手动录入

1条引用 6条上下文 0.91s 查看详情

完整响应 ?

问题分类

AI积分消耗

0.283

运行时长: 1.03s

模型: FastAI-3.5

总Tokens: 283

问题/检索: 文献的概念

上下文总数: 4

分类结果: 关于信息

问题分类

- 关于信
- 关于图
- 课题深

完整响应 ?

问题分类 知识库搜索 AI 对话

1045

问题/检索词:

文献的概念
文献的定义是什么?
文献的含义
文献的概念解释

搜索模式: 语义检索

相似度: 0.4

单次搜索上限: 1500

结果重排: false

问题优化模型: FastAI-3.5

问题优化结果: 文献的定义是什么?
文献的含义
文献的概念解释

引用内容: #1 | 综合排名 #1 语义检索: 0.9222

文献是指记录有知识的载体

12 手动录入 前往知识库 >

系统实现 1.0

文献的概念

关于大语言模型的适配资源

中国法分类数学属于什么类



四、基于开源LLMs的尝试

优化与展望

1) 基础调研自动化:

对于课题服务中的学术专业名词定义调研、课题起源发展、参考标引等基础性调研性工作借助大语言模型实现，让学科馆员能集中精力于深度服务；

真实性、学术性、准确性

2) 服务报告规范化:

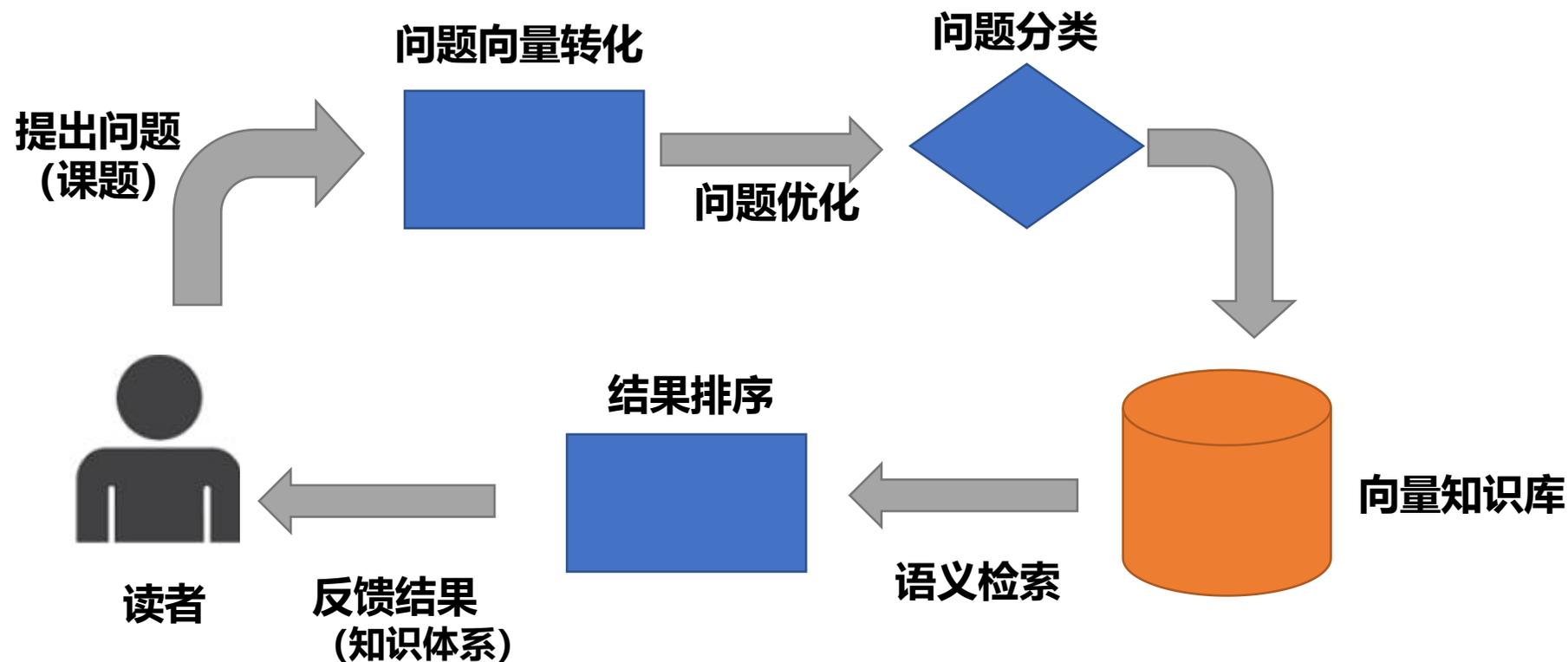
针对不同学科领域，基于已积累的本学科领域相关的服务报告，自学习建立学科服务报告模版，可参考目前较成熟的自动报表生成实现。

训练集数据量



四、基于开源LLMs的尝试

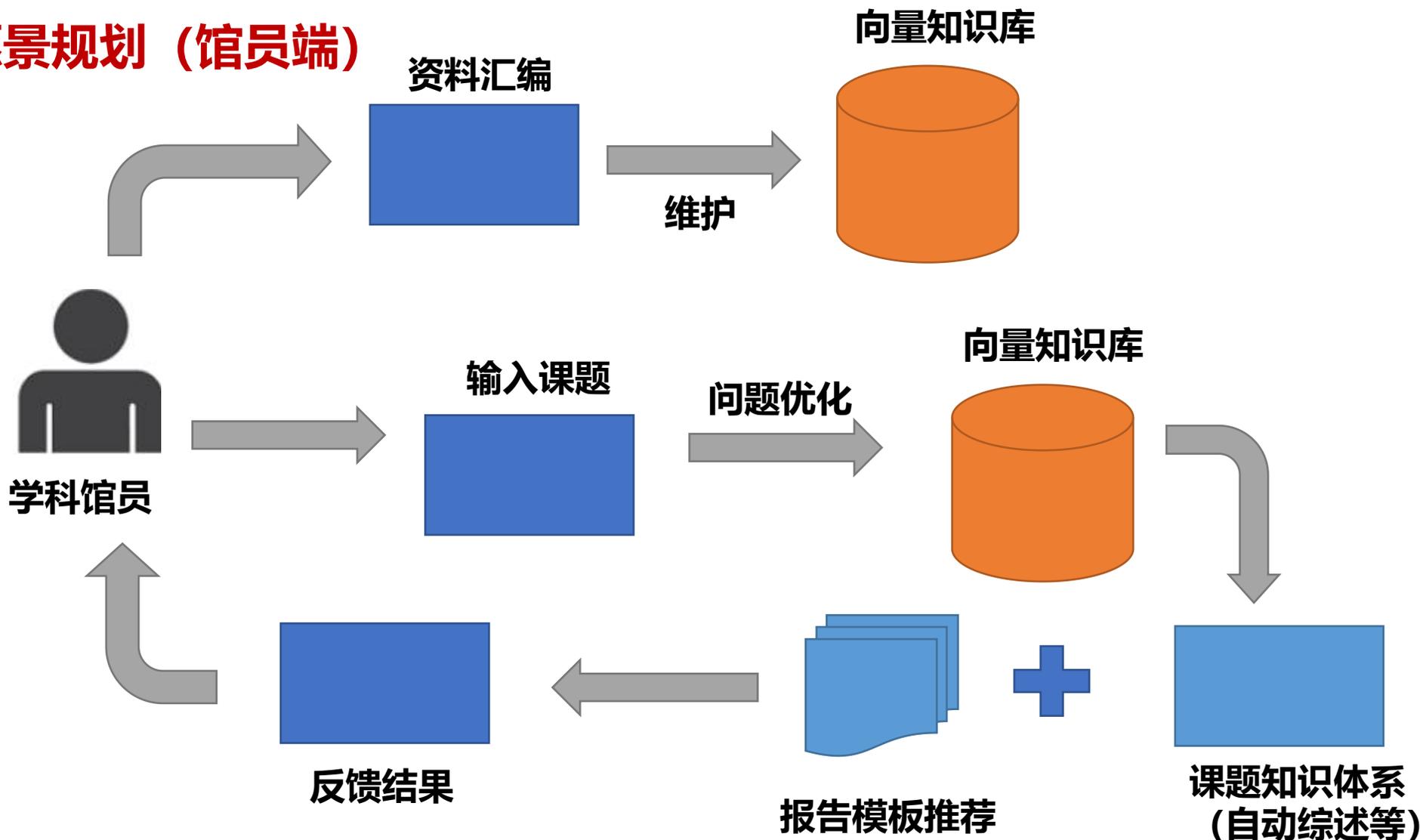
愿景规划（读者端）





四、基于开源LLMs的尝试

愿景规划 (馆员端)





山东大学
SHANDONG UNIVERSITY

图书馆
LIBRARY

敬请批评指正

山东大学图书馆 吴帼帼
guoguow@sdu.edu.cn

