

华知大模型与未来图书馆

Huazhi LM and the Future Library

报告人：张宏伟

Speaker: ZHANG Hongwei



生成式AI是智力革命，将深刻影响经济社会，重塑千行万业

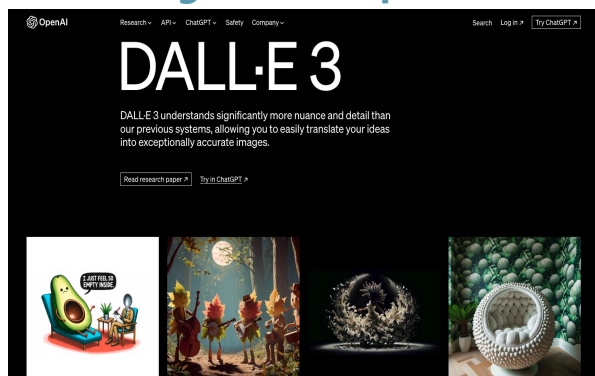
Generative AI is an intellectual revolution that will profoundly impact the economy and society, and remodel thousands of industries

生成内容更全面、交互方式更自然，可控性不断提高、适用场景快速扩展

More Comprehensive Generated Contents, More Natural Interactive Modes, Constantly Improved Controllability, and Rapidly Expanded Application Scenarios

【文生图】

Text-generated picture



颠覆知识生产方式
Overturn Knowledge
Production Mode

【文生视频】

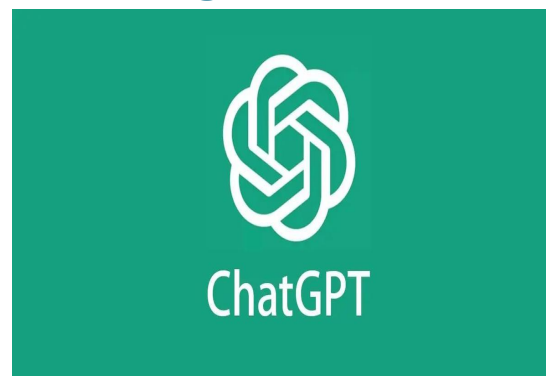
Text-generated video



变革信息获取方式
Revolutionize Information
Acquisition Method

【文生文】

Text-generated text



【全面多模态】

Comprehensive multimodal state



重塑人机交互方式
Reinvent Human-
Computer Interaction

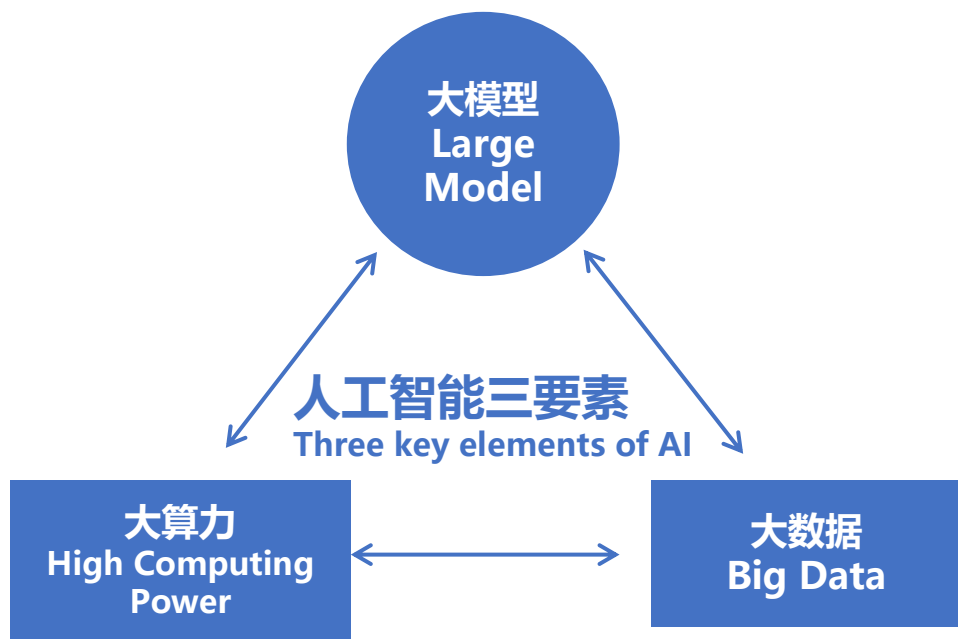
语言是思维的边界，大语言模型是生成式AI的基础和核心关键
Language is the boundary of thinking, and large language models are the foundation and key of generative AI

特征1：AI大模型一半是工程、一半是理论，建设门槛极高，将是智能时代的基础设施

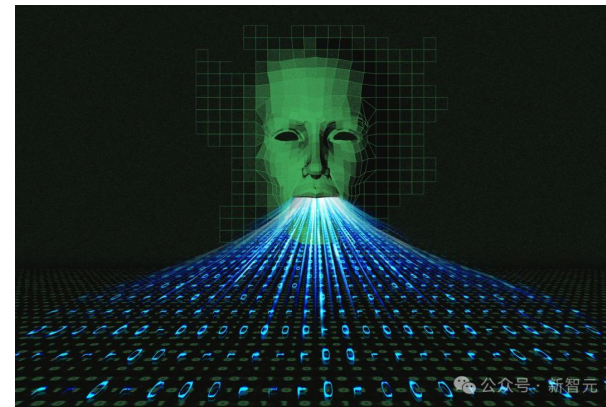
Feature 1: AI LM, half engineering and half theory, with an extremely high construction threshold, will be the infrastructure of intelligent age

AI大模型是工程化的重大创新，其核心技术壁垒是数据、算法、算力等要素资源的精巧组合

AI large model is a major innovation in engineering, and its core technical barrier is the exquisite combination of data, algorithm, computing power and other elements



“星际之门”开启，配备百万芯片，耗资千亿美元，所需电力高达5千兆瓦



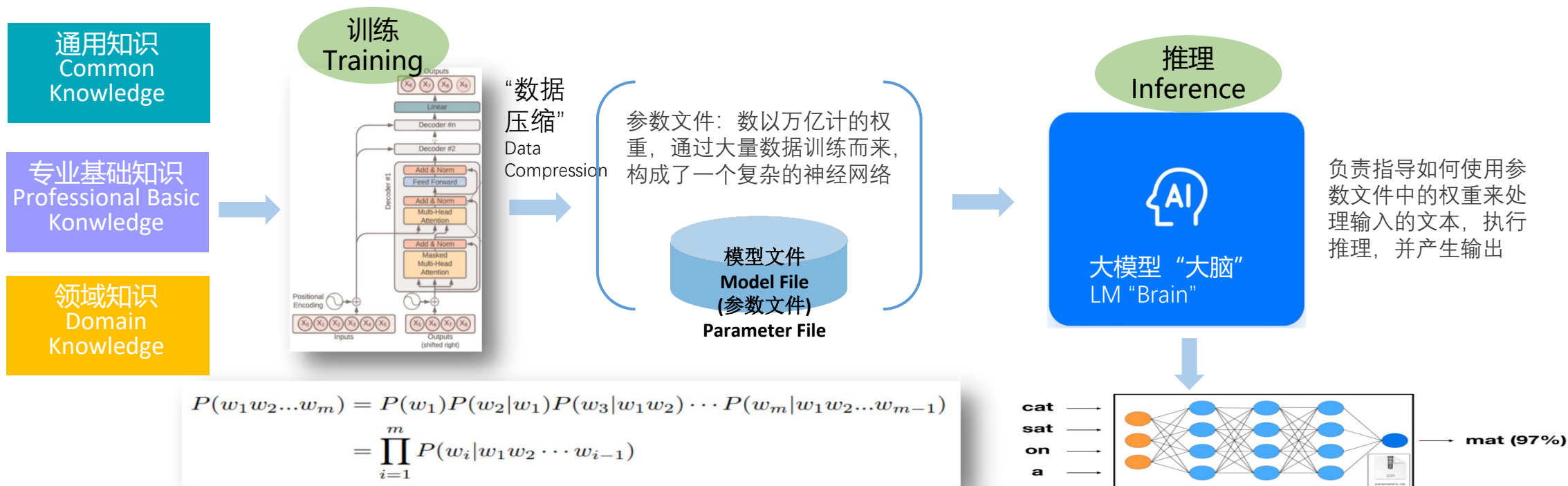
训练GPT-5短缺20万亿token，OpenAI被曝计划建“数据市场”

特征2：大模型本质是对人类全部知识的超高维压缩映射，数据是大模型的基石

Feature 2: The essence of large model is the ultra-high-dimensional compressed mapping of entire human knowledge, and data is the cornerstone of large model

AI大模型是超大神经网络，融合连接主义和贝叶斯主义两大学派优点，实现了性能跃迁，但本质还是统计学习

AI large model is a super-large neural network, which combines the advantages of connectionism and Bayesianism and realizes performance transition, but its essence is still statistical learning

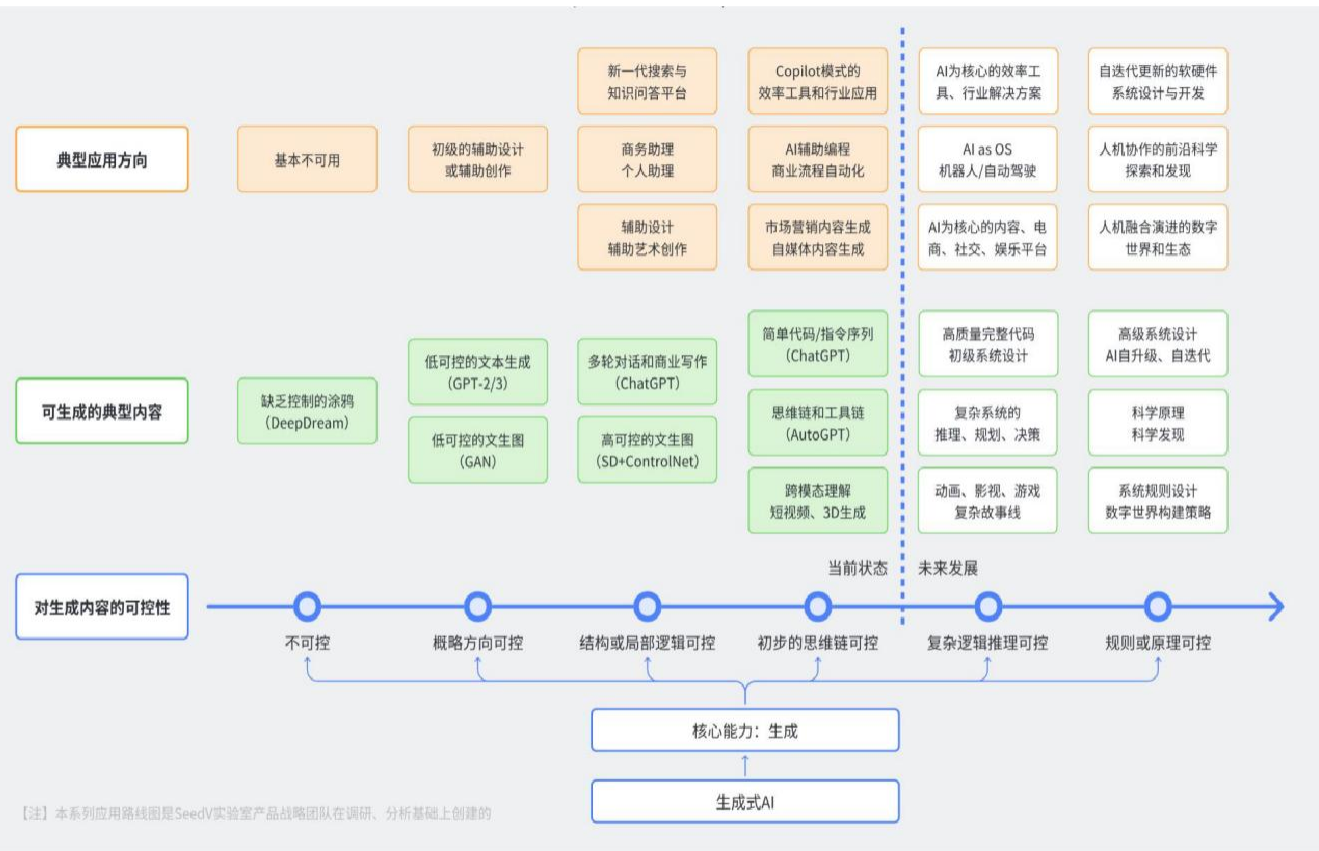


数据的质量、数量和多样性最终决定AI大模型的水平高低

Data's quality, quantity, and diversity ultimately determine the capacity of AI large model

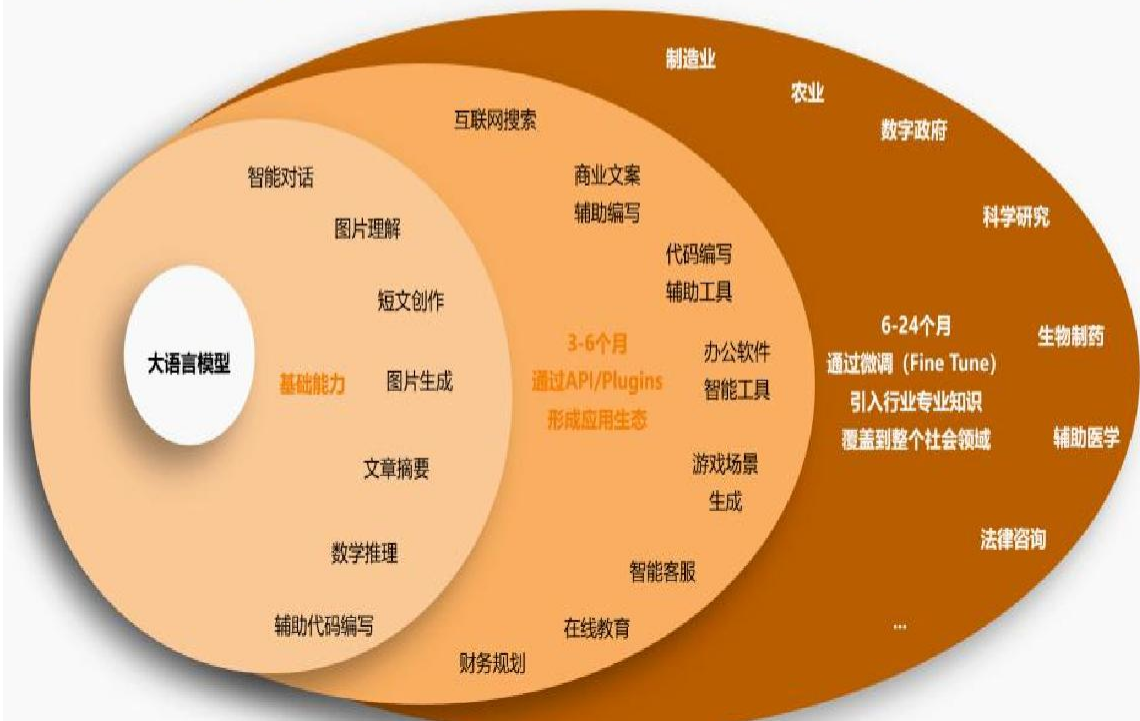
特征3：大模型对生成内容的可控性不断提高，适用场景正快速深化扩展

Feature 3: The controllability of LMs over generated content is increasing, and LM's applicable scenarios are rapidly deepening and expanding



“LM age” has begun and will impact the entire society and future

“大模型时代”已经开启，将影响整个社会和未来

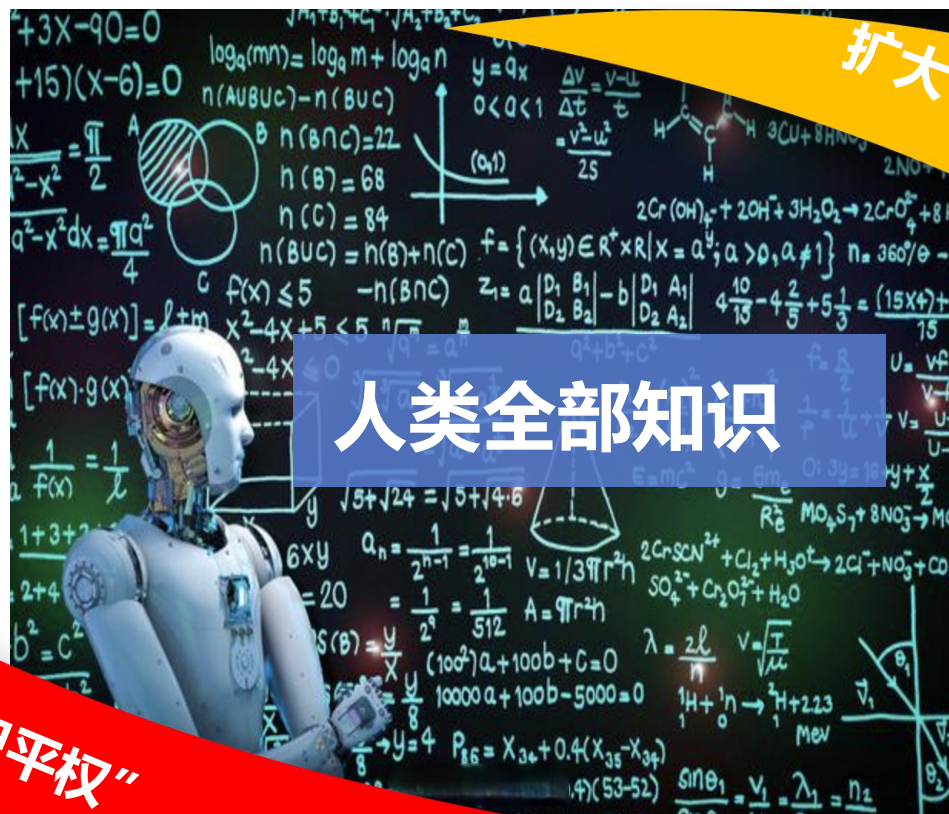


挑战：生成式AI是“人类增强”技术，会淘汰不懂AI的机构和个人

Challenge: Generative AI is a "human enhancement" technology that will eliminate institutions and individuals who do not understand AI

AI大模型的概率式方法范式决定其基本定位是启发和辅助人类

让更多的人公平、无障碍地获取、利用知识信息，成为潜在的知识创造者、问题解决者和创新者



扩大“数字鸿沟”



让不懂AI的人能力差距被放大，面临被更擅长使用AI的同行替代

促进“知识平权”

AI不断放大人类知识边界，也会放大个体间能力差距，人工智能素养是未来人才必备

AI constantly amplifies the boundaries of human knowledge and also amplifies the ability gap between individuals. Artificial intelligence literacy is essential for future talents

人类已学不过AI大模型，未来教育、学习范式亟需变革

Human beings can no longer compete with AI large models, future education and learning paradigms urgently need to be transformed

仅靠死记硬背就可以掌握的知识或技能逐渐失去价值，人工智能在这些方面可以做得比人类更好



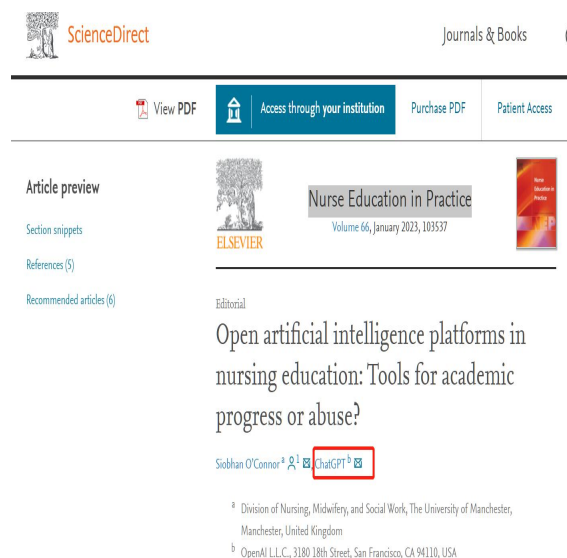
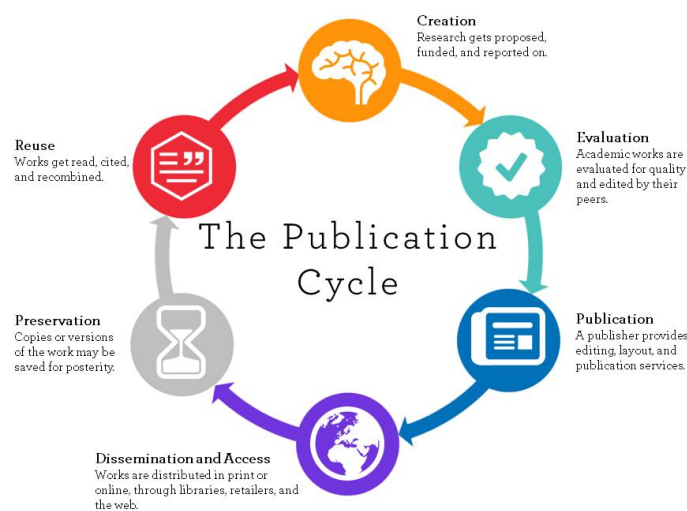
事实性知识和程序性知识重要性下降，原理性知识重要性上升

The knowledge service industry chain will be restructured to accelerate the transformation towards "real" knowledge services

编辑出版

图书情报

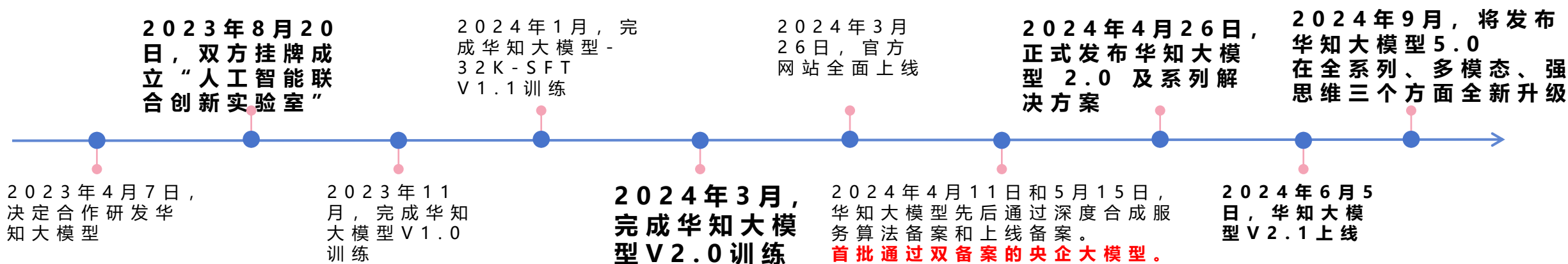
教育学习



华知大模型—国内首个知识服务和科研行业大模型，能力持续进化.....

Huazhi LM - The first knowledge service and research industry large model in China, with continuous evolution of capabilities

知网与华为，优势互补，携手成立人工智能联合创新实验室，共同打造**可融入行业生产系统的专业知识增强大模型—中华知识大模型(简称华知大模型)**，重点解决通用大模型知识记忆错误、信息时效性不足、缺乏专业知识等突出问题，入选人工智能产业焕新行动相关方案。



华知大模型：专业知识增强，服务专业领域、解决专业问题

Huazhi LM: Enhance professional knowledge, serve professional fields, and solve professional problems

通用性、专业性、实用性三大维度评测，能力效果优于竞品

全系列

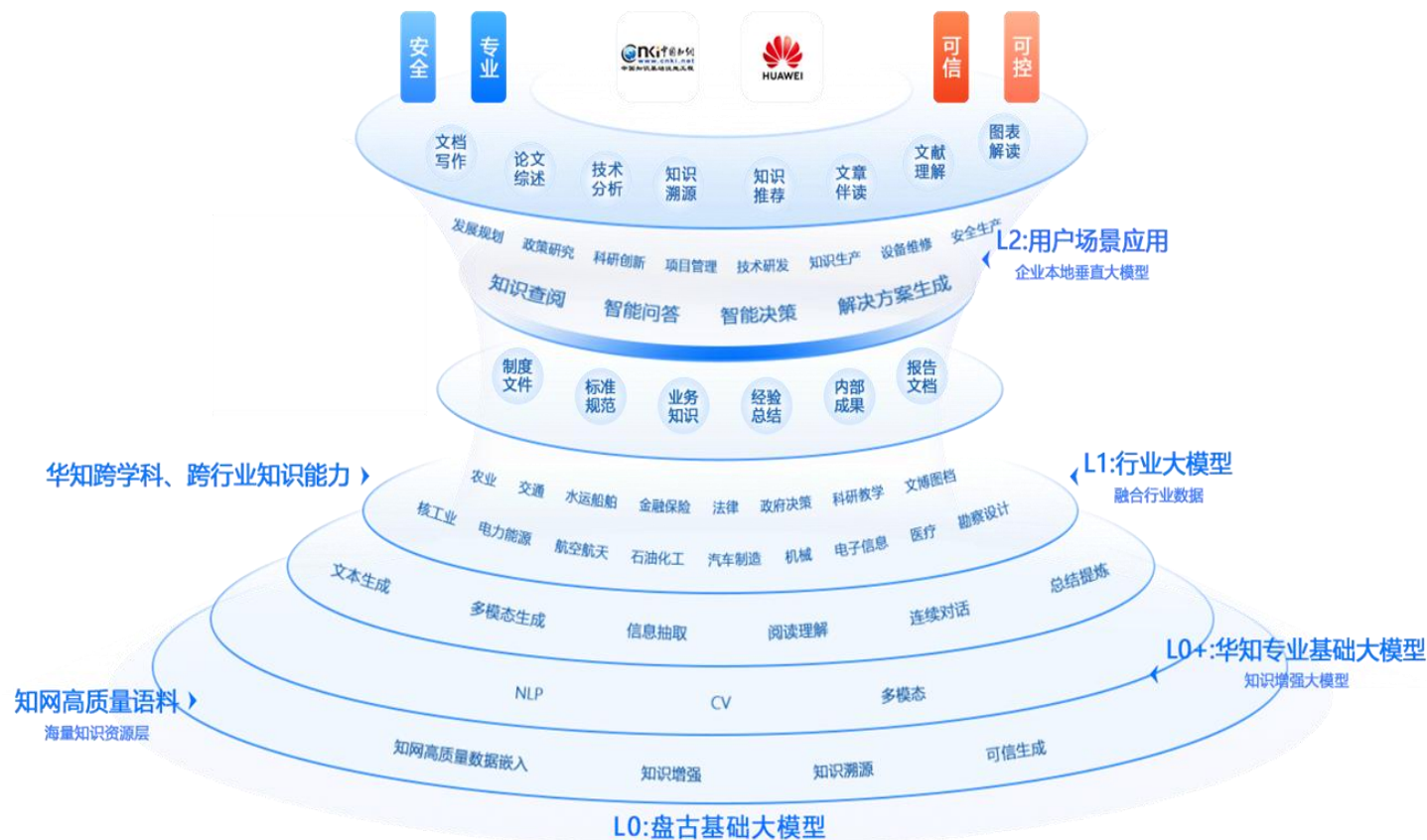
多模态

强思维

专业

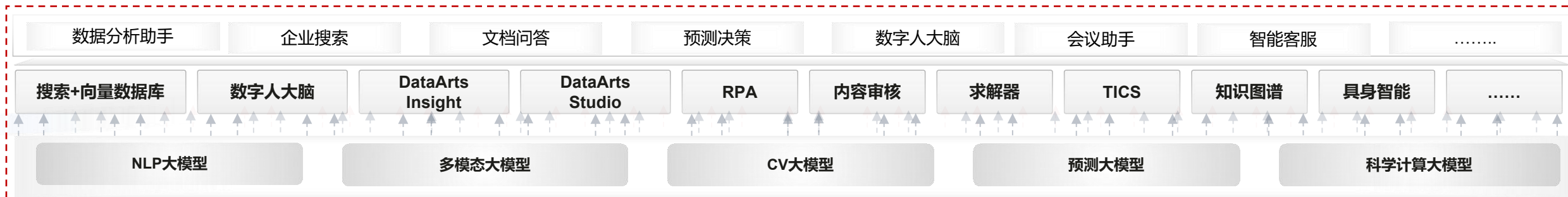
安全

可信

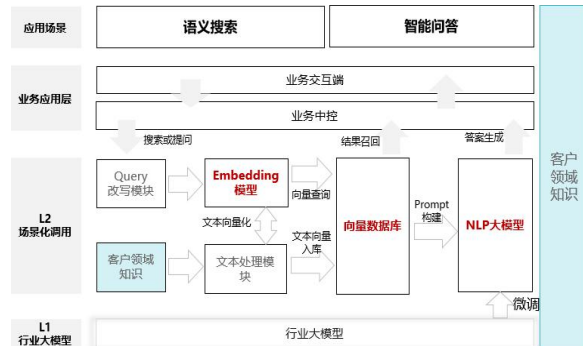


华知大模型 + 产品升级联动，构建跨行业水平解决方案，赋能知识密集行业领域

Huazhi LM + Product Upgrade, Building Cross-Industry Horizontal Solutions, Empowering Knowledge-Intensive Industry Fields



搜索&向量数据库+华知



应用场景

- 语义搜索：企业知识库、帮助中心等
- 智能问答：智能客服、政策问答等

方案核心竞争力

- 文本Embedding模型：在中文领域，模型综合评分领先友商。
- AIKBase向量数据库：相同精度下性能更好，相同性能下精度更高。
- NLP大模型：多种参数规模大模型，按需取用；全系列大模型开发与应用套件支持用户快速迭代使用；模型内容安全可信。

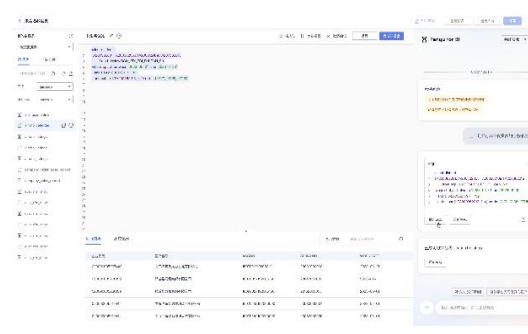
数字人+华知



应用场景

- 智能问答、咨询、播报等
- ### 方案核心竞争力
- 效果运营：通过丰富的运营功能实现快速的数字人大脑问答效果提升
 - LLM模型大脑：在信通院大模型评测模型开发与模型能力均达到4+级能力；模型安全：大模型具备严格的内容安全处理能力，并具备符合主流价值观、伦理道德的内容生成能力，全流程支持模型训练与推理安全。

DataArts + 华知



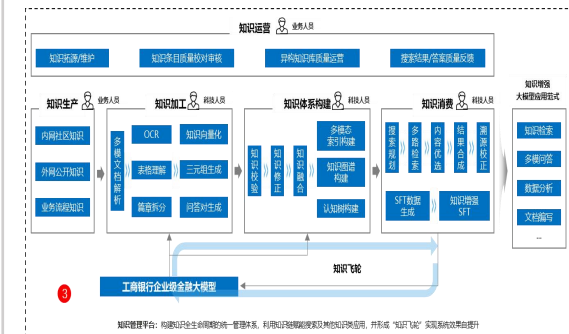
应用场景

- 自然语言取数：辅助生成SQL，提升取数
- 自然语言查询：直接对话式生成报表，帮助业务即席查询/分析

方案核心竞争力

- 集成开发助手：辅助SQL生成；
- 数据资产管理：基于自然语言查询资产、资产图谱生成及相关资产推荐。
- 数据分析助手：与BI集成，辅助图表生成，推荐可视化组件、生成图表并添加到BI仪表盘，发送图表到用户业务群组

知识管理+华知



应用场景

- 企业知识沉淀：
- 知识挖掘
- 知识检索
- 知识增值
- **方案核心竞争力**
- 能力融合：系统化整合各类AI能力，围绕企业全面的知识管理诉求构建可解耦、分层清晰，灵活取用
- 多项原子能力全国评测TOP1

华知大模型提供边云端多种服务模式，满足用户多样化需求

Huazhi LM offers a variety of service models at the edge, cloud, and device, meeting the diverse needs of users



模型接口

直接调用大模型接口，支持用户搭建本地化应用服务。

“模型+知网向量知识库” 接口：可信的知识检索增强服务

支持利用出版大模型结合知网向量知识库，进行知识检索增强场景应用，提供可控生成、可信溯源的内容生成服务。

专有云上模型：提供训推一体化服务

支持在华为昇腾云上独立部署出版大模型，提供增量训练、SFT微调和推理服务。

本地部署与训推服务

支持用户在自有昇腾算力设备上部署出版大模型，提供训练、推理服务。

云端大模型+本地向量库提供知识增强服务

云端大模型+本地向量库部署，RAG知识增强服务；出版机构资源不出社，实现资源的本地训练、本地调用。

“模型+昇腾服务器” 训推一体机：

出版大模型预装到昇腾服务器上，集成全栈AI开发与管理平台，为用户提供开箱即用的软硬一体化服务，支持用户在本地图训练、推理。

云服务

私有化部署

基于华知大模型，开放合作，共筑大模型研发应用新生态

Based on the Huazhi LM, open cooperation to jointly build a new ecosystem for the development and application of large models

华为云



cnki 中国知网
www.cnki.net
中国知识基础设施工程

华知
大模型

腾云出版大模型

律境·法律大模型

合作
伙伴



中国公共关系协会文化大数据产业委员会



贵阳市政府



武汉市社会科学院



新华社国家重点实验室



清华大学



武汉大学



云冈石窟研究院



中核集团财务共享中心



中国宝原投资有限公司



核工业第八研究所



湖北省社会科学院



北京市科技情报研究所



内蒙古社会科学界联合会



北京儿童医院

文化大模型

政务大模型

内容安全大模型

文化遗产大模型

财务大模型

制度大模型

华知·龙吟核行业大模型

文化智库大模型

儿科虚拟患者大模型

行业
大模型

华知大模型赋能知网全线产品服务升级改造，提升服务效率质量

Huazhi LM empowers the upgrade and transformation of CNKI's full range of products and services, enhancing service efficiency and quality

- 知网知识生产、知识服务、知识应用产品全面升级
- 40+应用场景、上千万用户体验、百万篇AI生成文档





AI学术研究助手

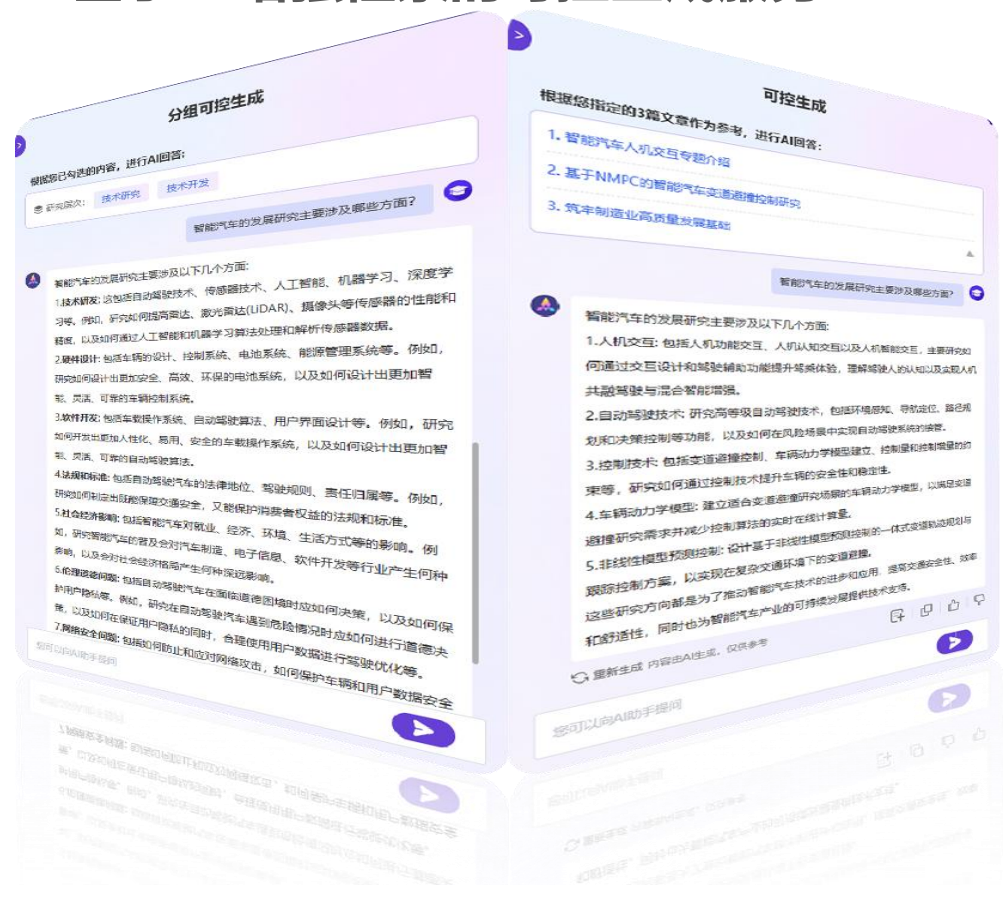
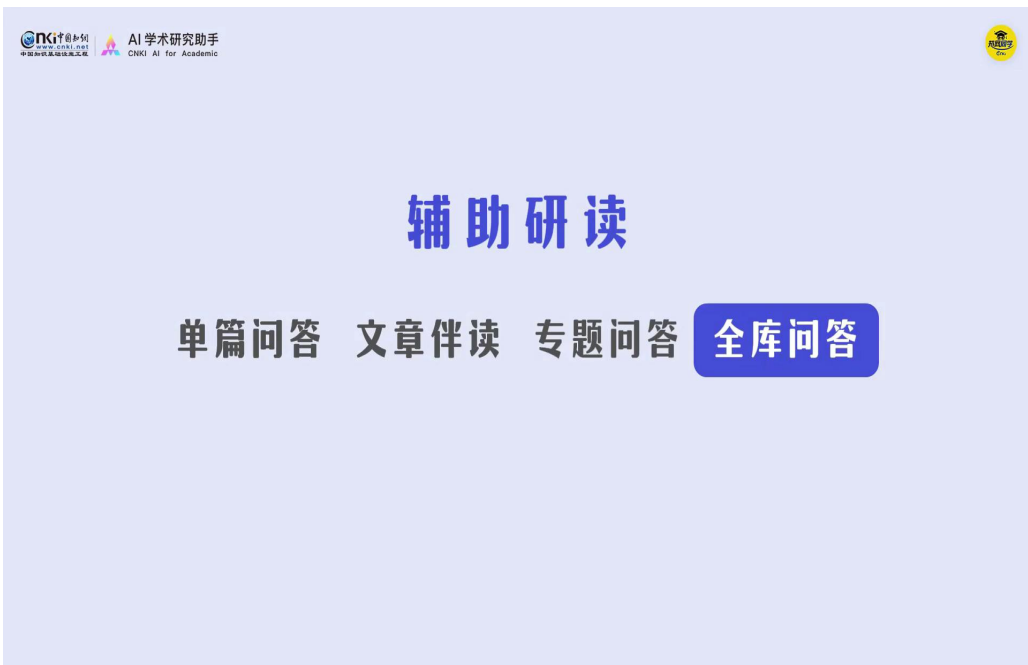
CNKI AI for Academic

个性化、生成式专业知识服务平台

Personalized, Generative Professional Knowledge Service Platform

问答式增强检索,生成式知识服务

基于AI增强检索的可控生成服务



AI学术研究助手 | 01

问答式增强检索

(全库问答)

五月份正式上线以来，访问量迅速增长，个人用户破千万

Since its official launch in May, the number of visits has grown rapidly, with individual users exceeding ten million

CNKI智能写作

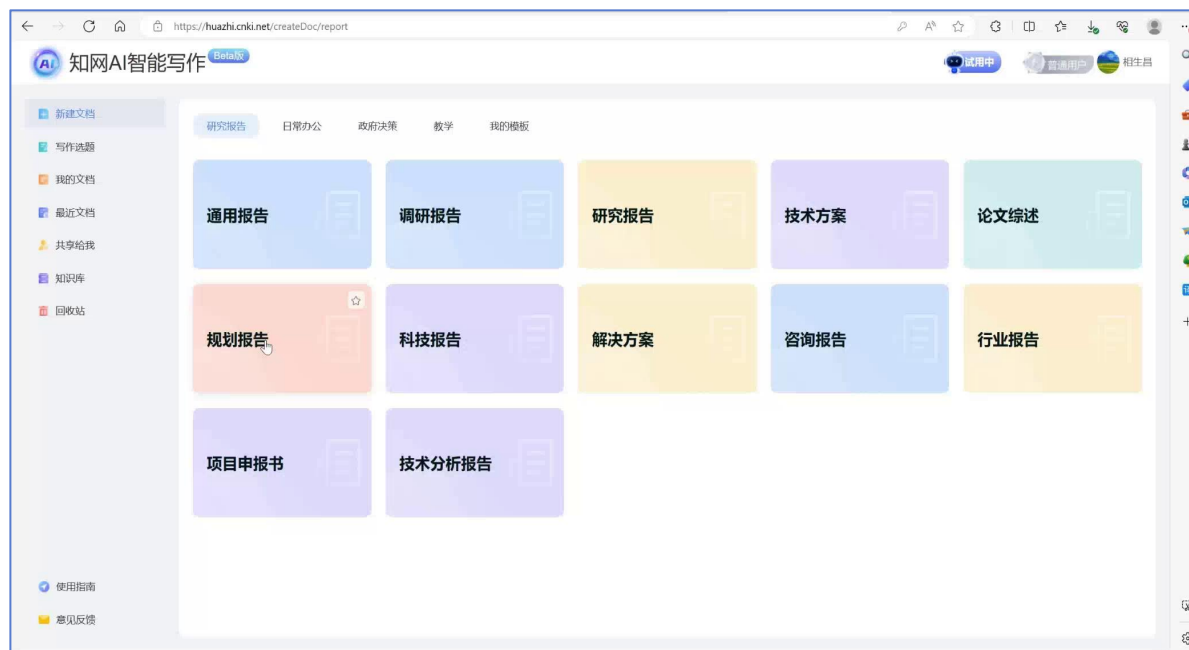
CNKI AI Intelligent Writing

生成内容更专业、更可靠

Generated more professional and reliable content

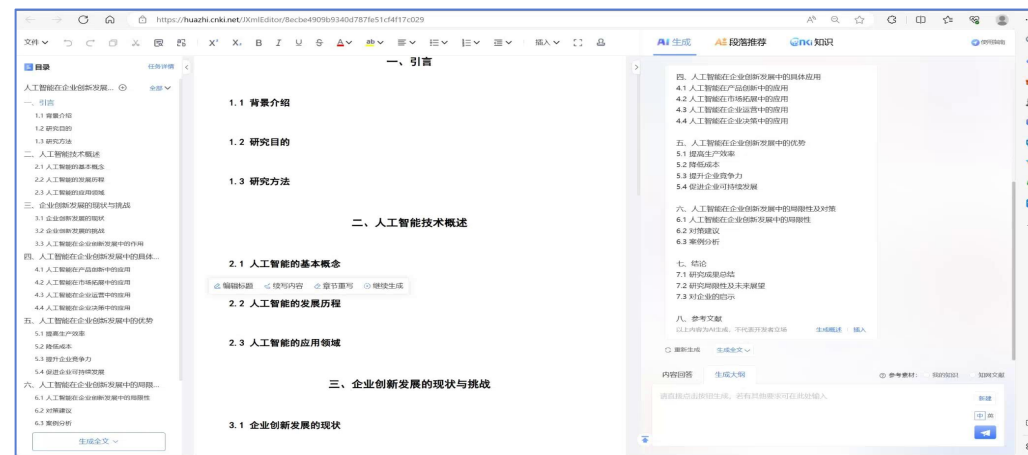
1 生成专业级大纲及概述

Outline and overview generated in professional level



2 生成全文，知识增强，编改润色

Full text generated with strengthen knowledge and editing/ polishing



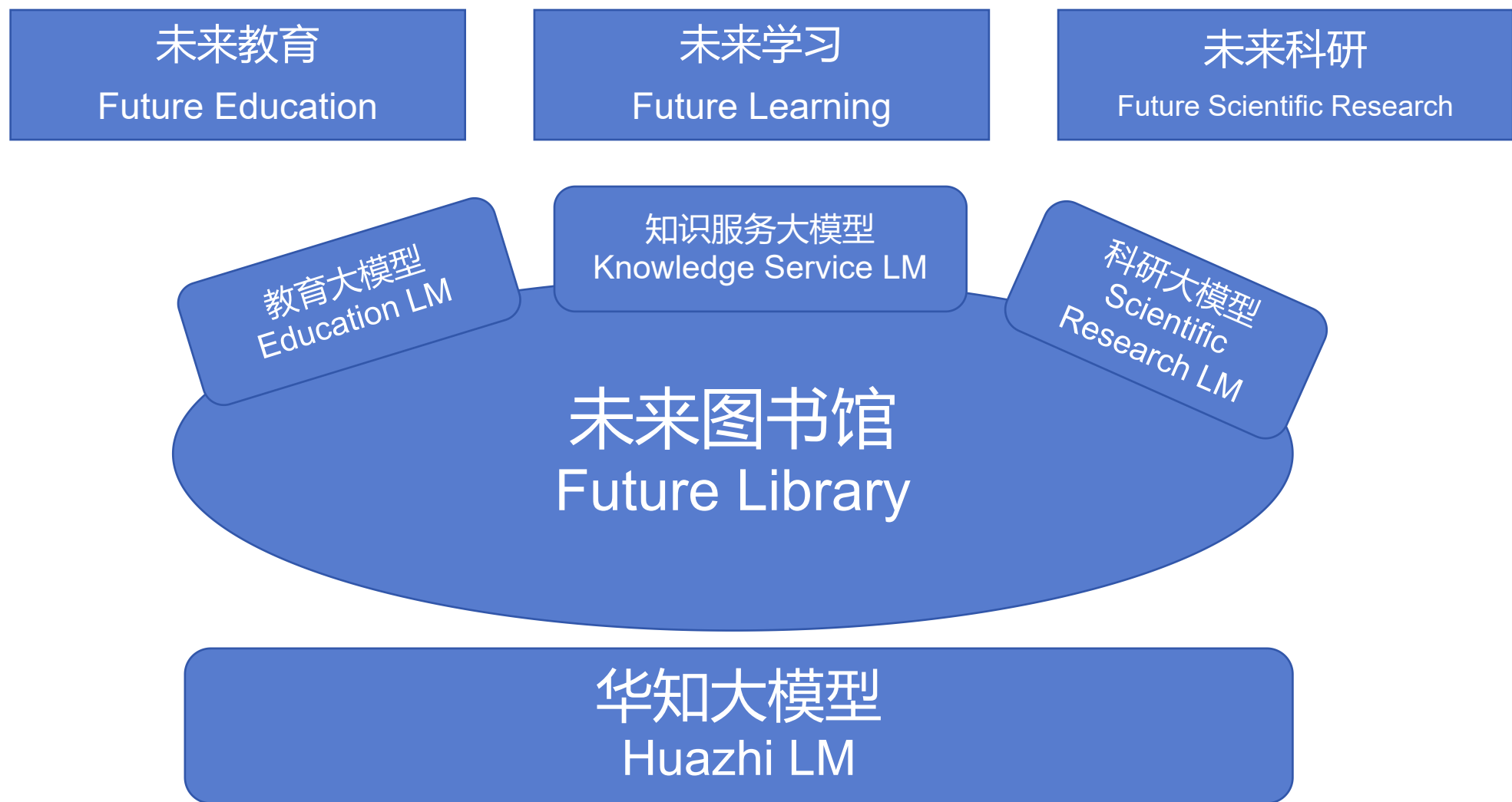
3 内容验证、溯源；标准格式导出

Content verification and traceability; exporting of standard formats



华知大模型赋能图书馆智能底座构建，服务教育教学、科研管理全过程

Huazhi LM empowers the construction of an intelligent library base, serving the entire process of education, teaching, and management



知网+华为云，联合开展图书馆+人工智能示范行动

CNKI + Huawei Cloud jointly carry out Library + AI demonstration activity

为有效落实人工智能+行动，面向海内外教育科研机构，遴选**1000家**图书馆，开展图书馆+AI升级示范行动，免费赠送**10000亿tokens算力**，超出部分以市场最低价供应。

To effectively implement AI+ action, **1000** libraries of education and research institutions will be selected around the world to carry out library+AI upgrade demonstration actions, with a free gift of **1 trillion tokens of computing power**. Any excess will be supplied at the lowest market price.

活动时间：**2024年6月18日-12月31日**

Activity period: **Jun. 18, 2024 - Dec. 31, 2024**

谢谢!

Thank you!

