基于大模型数智化底座的CNKI系列AI应用介绍

华知：中华知识大模型

大模型时代的图书馆智能化服务

2023·上海

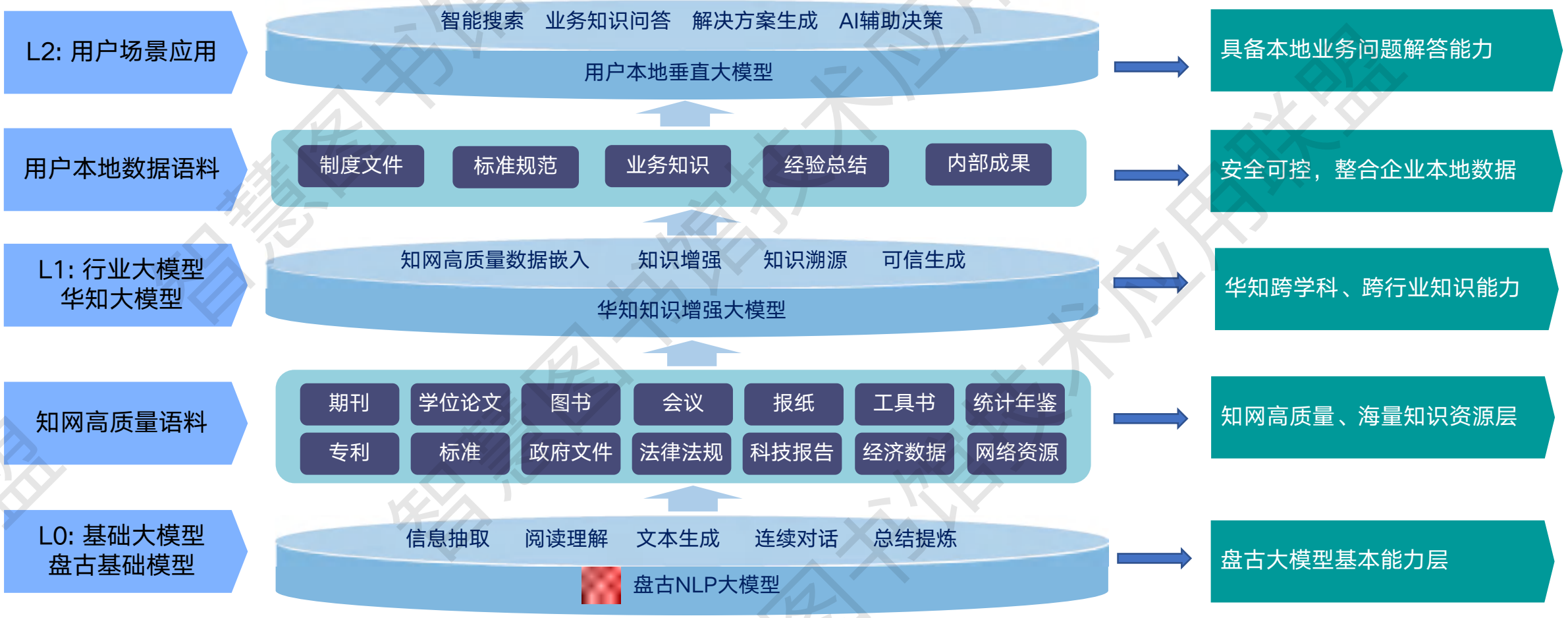
智慧图书馆技术应用联盟

知网数据+华为盘古打造服务科技创新和学术研究的中华知识大模型

与华为携手成立人工智能联合创新实验室，发挥华为核心技术优势和知网高质量数据优势，面向知网成熟的应用场景，共同打造自主可控、可信增强的**中华知识大模型（华知大模型）**，在具备通用大模型基本能力同时形成华知大模型8大特色能力，包括文献理解、文章伴读、知识推荐、知识溯源、技术分析、论文综述、文档写作、图表解读。

从L0到L1再到L2的三层架构，华知融合图书馆本地数据可信增强

华知大模型分层架构体系，从L0（盘古基础模型）到L1（华知行业应用模型）再到L2（企业垂直模型）的三层架构体系，有效解决通用通用大模型存在的专业性差、可信度低等问题，整合知网高质量行业专业数据和用户本地数据，为企业用户构建具有行业属性和个性化特色的垂直大模型。



融合数据、知识和AI，打造新基座能力体系支撑智慧图书馆建设

为智慧图书馆建设的打造全新底座，嵌入图书馆全业务提供智能服务

知识底座

人工智能底座

大数据底座

数据中台\知识中台\AI中台

智能化能力

应用AI技术，综合数据和知识的能力体系，面向上层业务应用形成综合智能能力体系

数据智能 | 机器智能 | 认识智能 | 感知智能 | 创作智能

知识服务能力

面向用户、面向业务应用提供一站式知识发现和获取，知识资源全生命周期的服务

知识导航 | 知识检索 | 智能推送 | 知识关联 | 知识地图 | 知识重组

AI基本能力

大模型基于数字化、资产化的数据和知识，进一步提升智能化能力，发掘数据价值

智能标引 | 智能挖掘 | 语义理解 | 信息抽取 | 语义问答 | 文本生成

数据服务能力

数据体系化盘点、资产化服务，面向用户（人）、系统（业务）、机器（AI大模型）

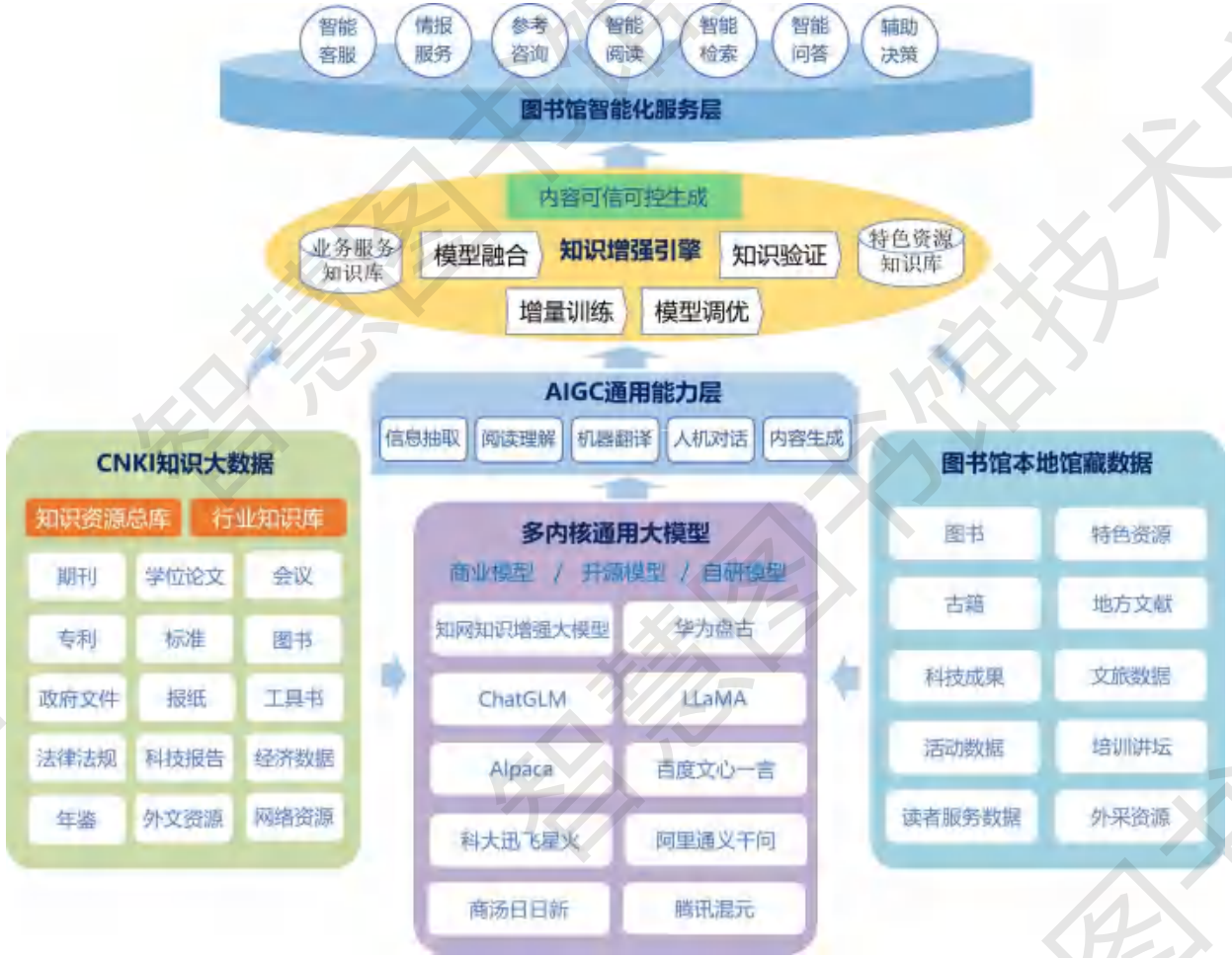
数据目录 | 数据地图 | 数据检索 | 数据分析 | 数据计算 | 指标检索

综合能力输出

- 文献服务
- 参考咨询
- 馆藏服务
- 借阅服务

融合数据资源搭建图书馆自己的人工智能大模型

基于多内核通用大模型，融入**知网高质量知识资源**和**图书馆内部馆藏数据**，采用增量训练、局部微调、知识嵌入、外挂知识库等增强机制和手段，解决大模型生成内容不精准、专业性差、可信度低问题，实现大模型及其人工智能应用在图书馆本地低成本、轻量化部署，帮助图书馆建设具有自身特点的大模型底座。



- 支持插拔式多内核模型引擎
- 高质量专业知识资源注入
- 实现用户本地轻量化部署
- 生成内容的可信可控保障

场景一：CNKI AI 学术研究助手，探索向“生成式”知识服务转型

对大模型生成内容进行可信增强，提供内容出处，实现可信溯源

AI智能推荐相关学者研究领域问题

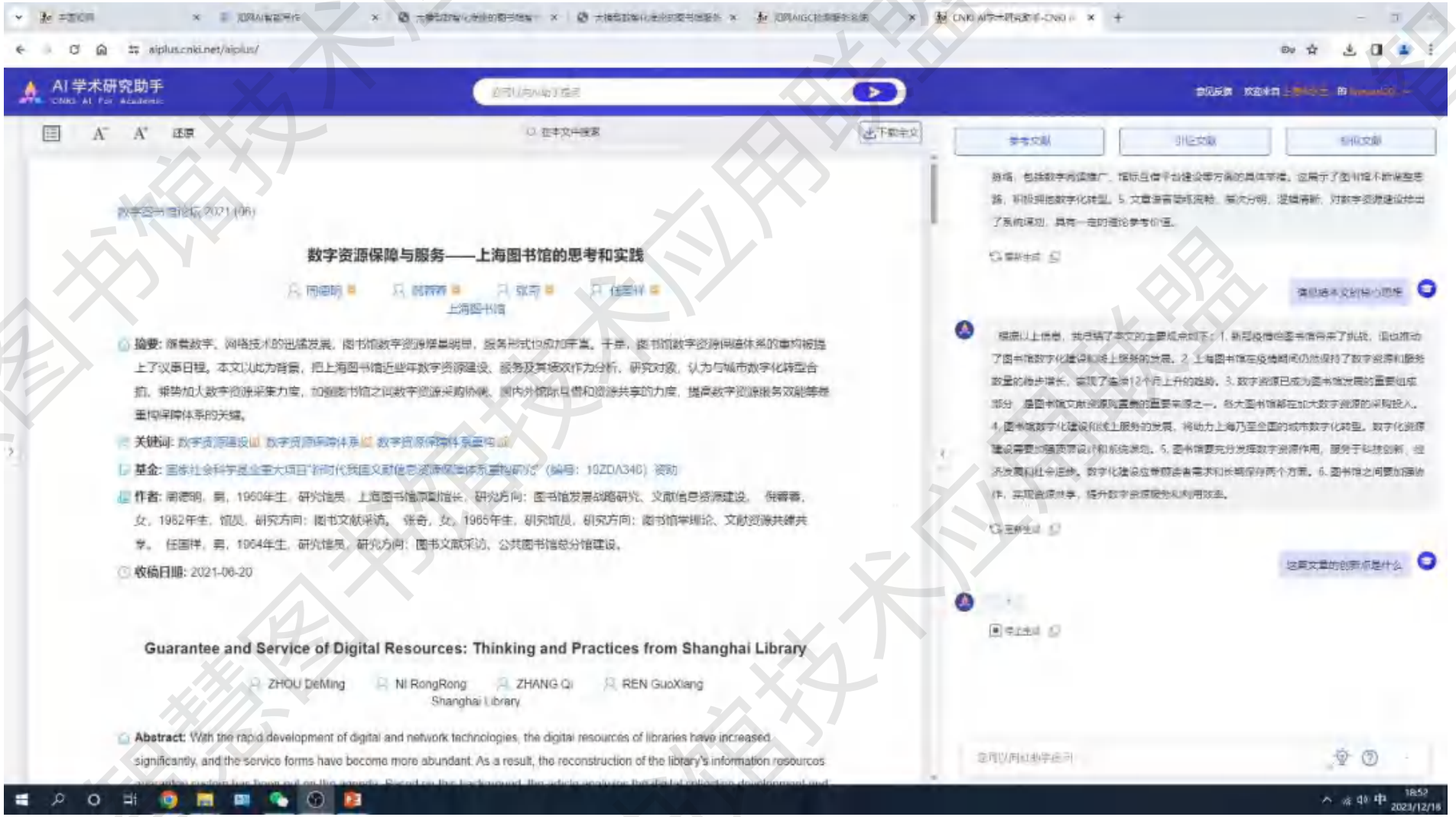
关联当前问答参考和引用的原始内容素材

场景二：通过AI为用户提供辅助阅读工具，自动阅读总结图书和文献内容

将大模型内容理解和信息抽取应用到数字阅读环节，读者可以机器对话，了解相关图书的主要内容以及不同图书之间的内容差异，直接将结果反馈给读者，从而节省读者时间。

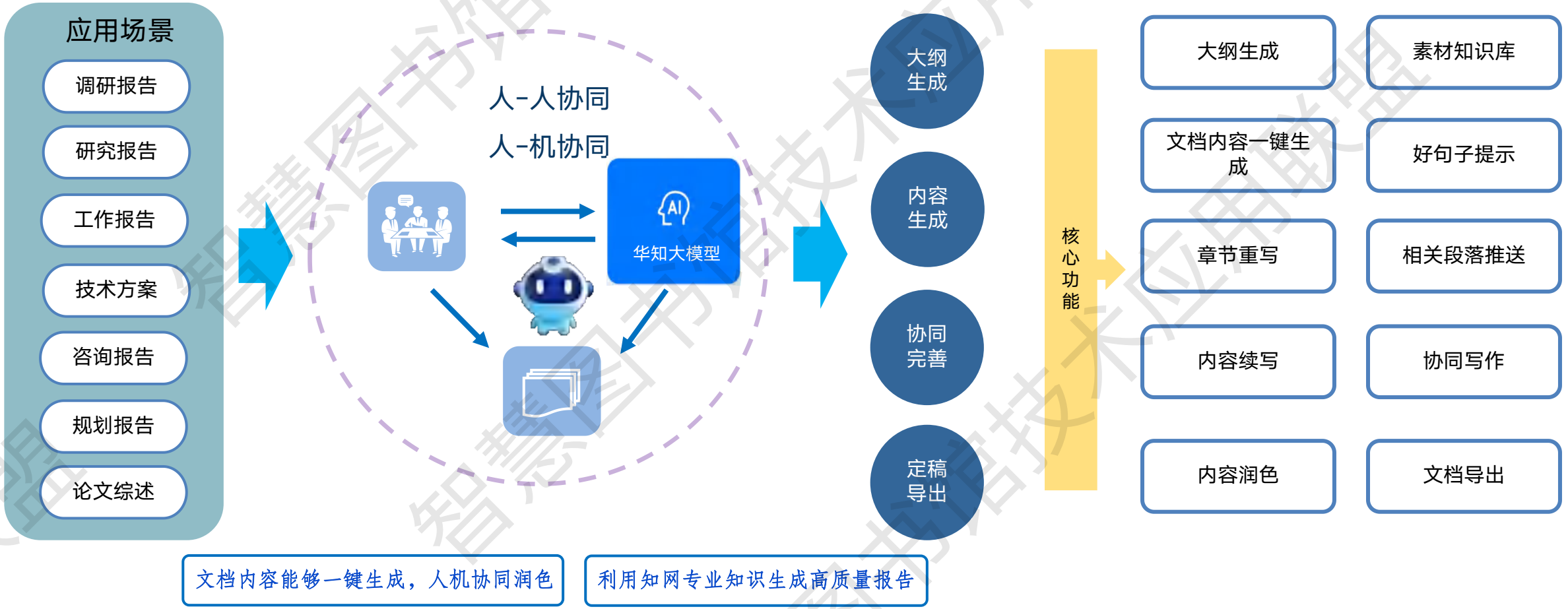
机器阅读理解

内容智能总结



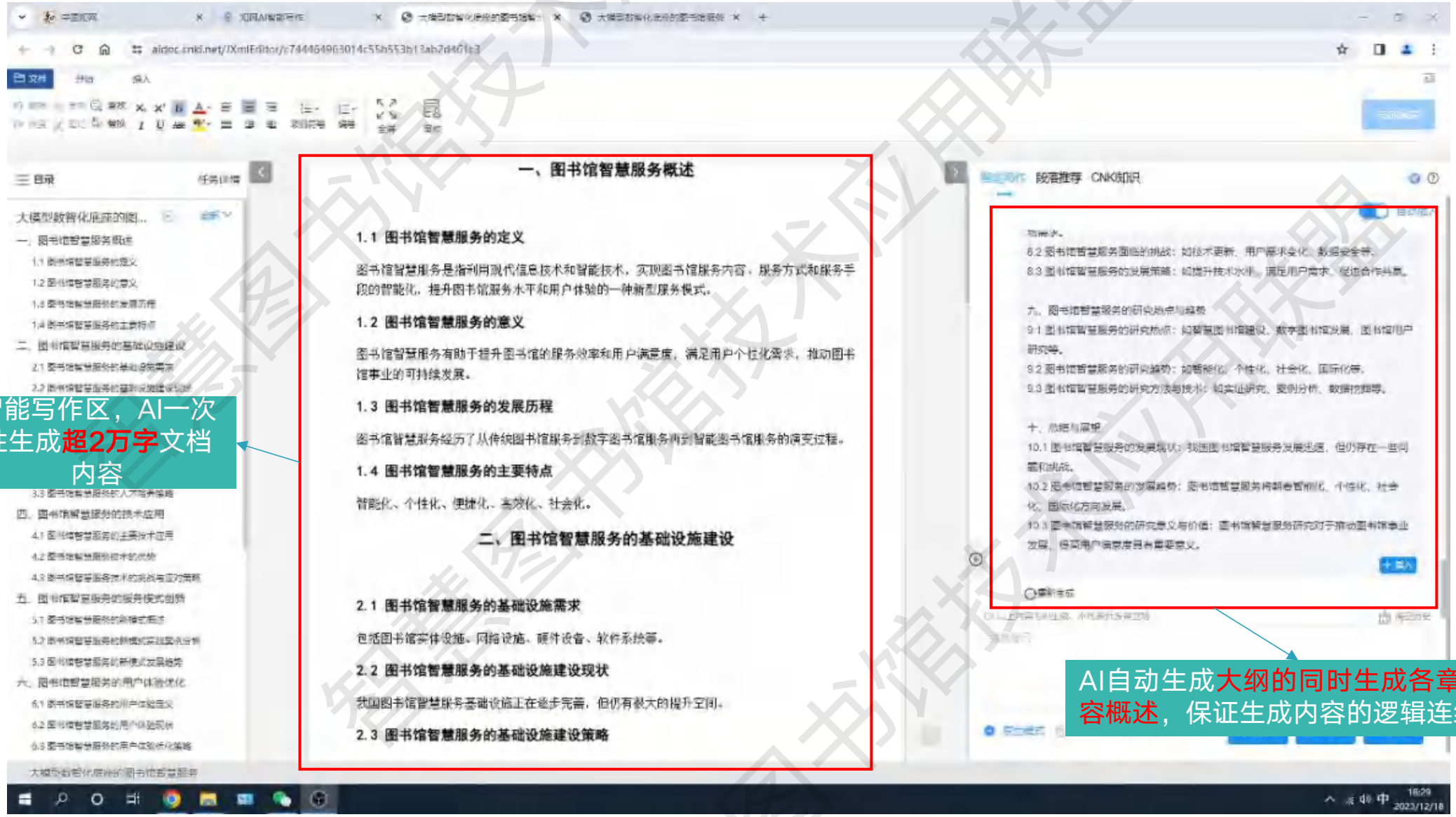
场景三：智能写作，AI智能生成高质量研究报告，助力科研创新

智能写作将大模型生成能力与协同文档、知识库三者进行融合，嵌入技术方案、调研报告、规划报告、研究报告、工作总结等各类文档编写场景，开发类ChatGPT的人工智能助手，辅助用户快速、高效地编写文档。



场景三：智能写作，AI智能生成高质量研究报告，助力科研创新

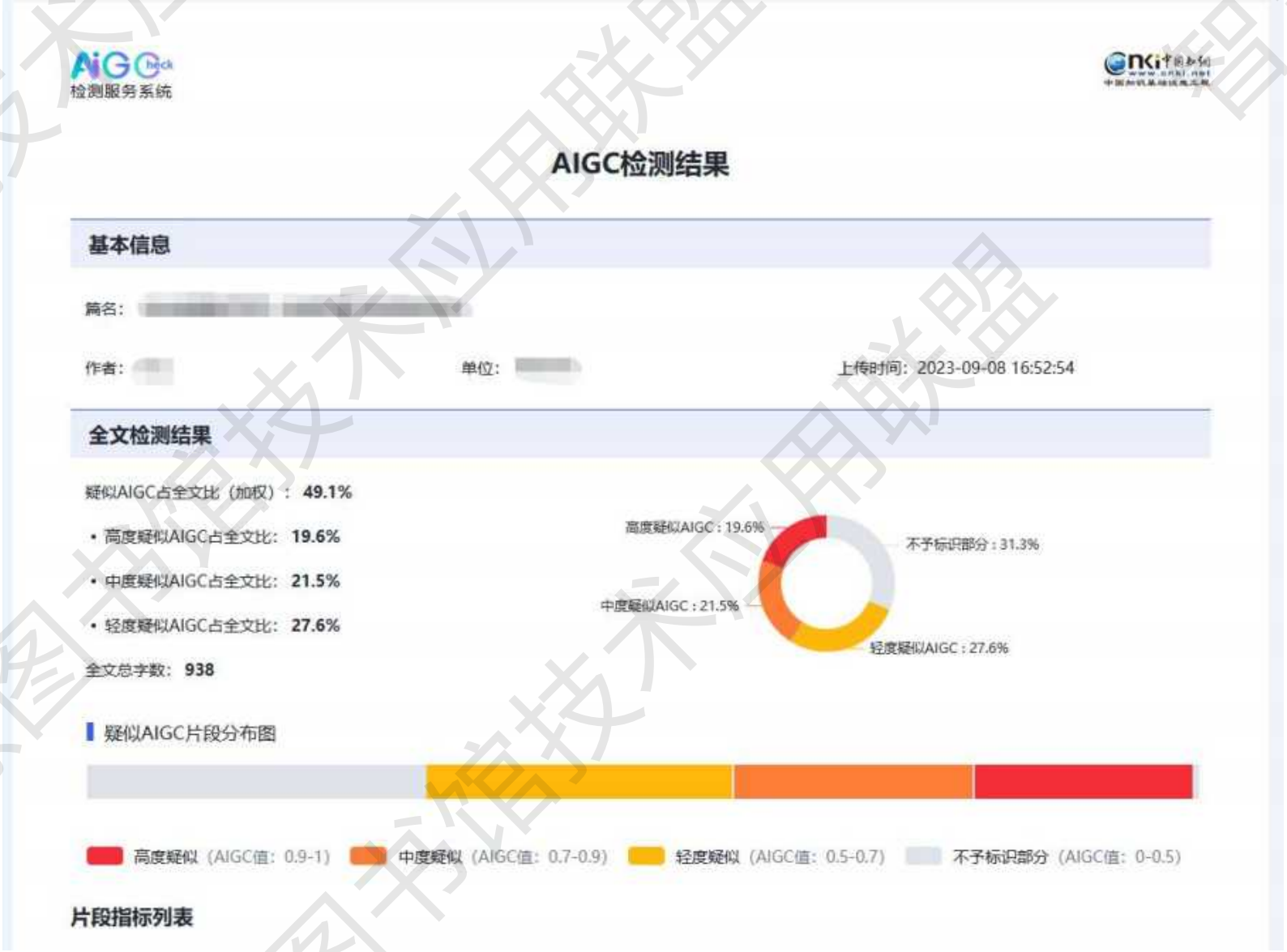
只需要三步（生成大纲、概述、全文）即可完成上万字报告的自动生成



场景四：AIGC检测服务，助力学术诚信与学术规范建设

系统以知网结构化、碎片化和知识元化的高质量文献大数据资源为基础，基于预训练大语言模型算法逻辑，结合“知识增强AIGC检测技术”和若干检测算法，从语言模式和语义逻辑两条链路，用AI检测AIGC，可快速、准确识别包括期刊论文、学位论文、图书专著和科研成果等在内的学术文本中不同程度疑似AI生成的内容。

- 引导和监管AIGC技术的科学合理使用
- 维护学术诚信和保护知识创新
- 加强学术生态建设、健全学术治理体系
- 提高人才培养质量、加快科技自立自强



诚邀体验

https://huazhi.cnki.net/

同方知网 · 上海

2023.12

智慧图书馆技术应用联盟