

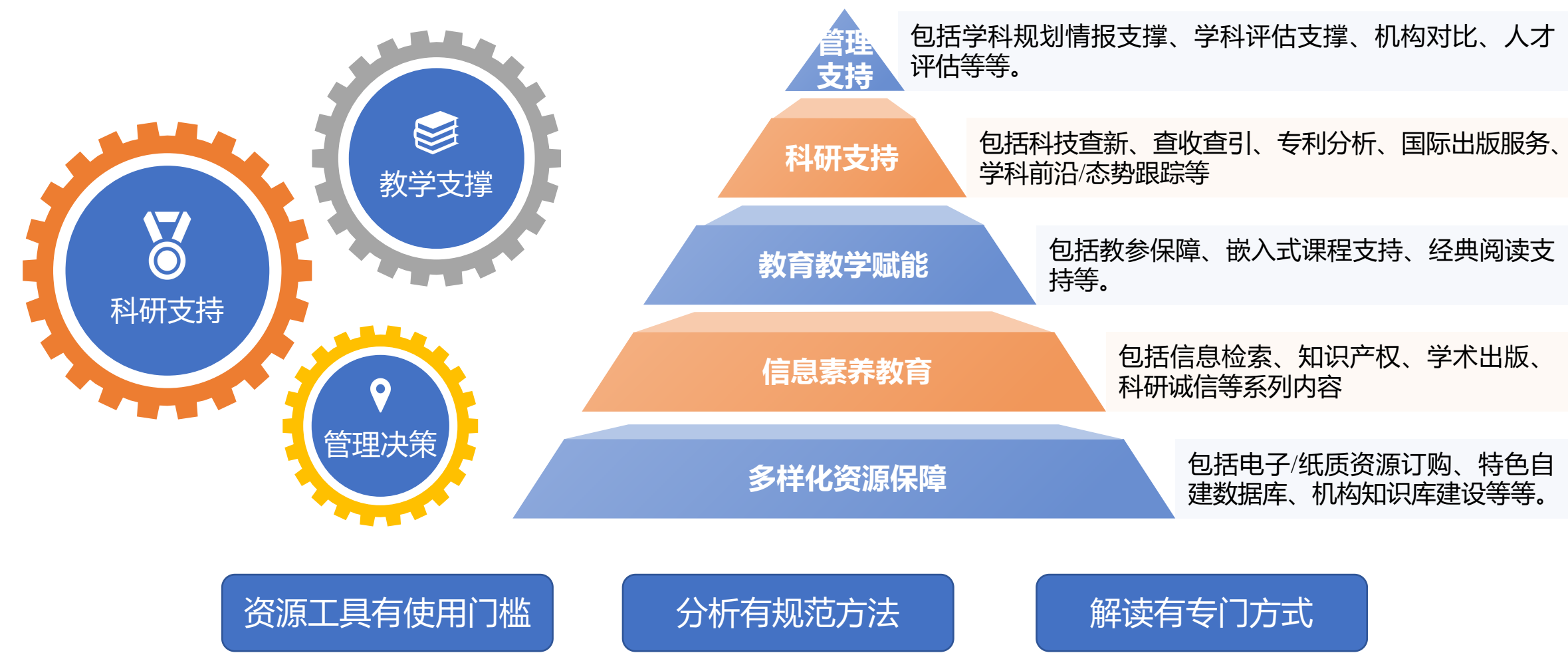
# AI与数据融合的学科服务支撑体系 探索与实践

万方数据知识产品开发部 刘建华

二零二五年六月二十七日

# 现有学科服务的主要构成

## □ 现有学科服务的主要构成



# 新环境下学科服务面对的一些问题

## 当前学科服务面临的挑战

逐渐降低对科技查新、简单的查收查引的强制要求；科技评价体系的改革对传统评价提出挑战。普遍经费紧缩。

### 客观的政策变动及经济原因

AI引发各类知识服务的生产关系革命性变革，包括但不限于知识学习长链条贯通、知识服务去中心化等。传统图书馆中介服务角色受到AI环境下端到端的服务冲击。

### 学术交流生态变化

正式学术交流与非正式学术交流逐渐趋向融合，对于资源组配、资源利用、科学评价等等都带来新的机遇与挑战。

### AI+学科、跨学科发展成为常态

如何能在AI融入学科发展、跨学科发展中，发挥图书馆的作用，开拓新的服务领地。

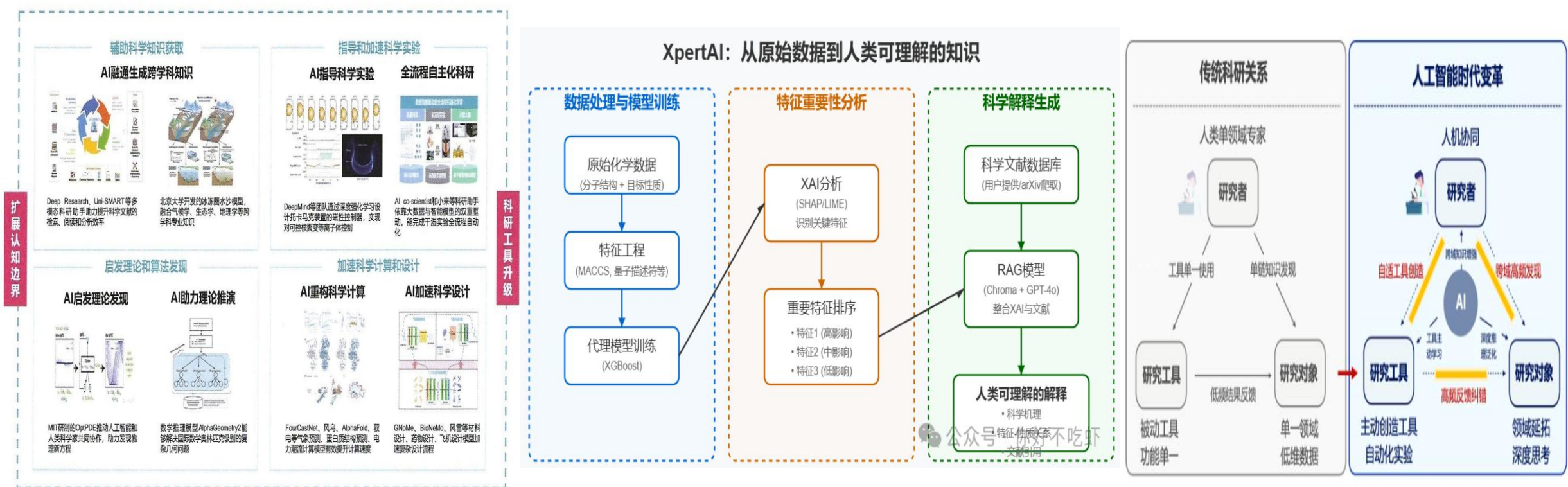
### AI推动各类关系变革

Ludo Waltman.Assessment reform and publishing reform need to go hand in hand.<https://www.leidenmadtrics.nl/articles/assessment-reform-and-publishing-reform-need-to-go-hand-in-hand>Ewen

张晓林:Library Inside:AI 赋能图书馆新质生产力的一种基础模型

# 人工智能与学科、行业的结合，推动AI4S、AI4E等的同时，也在推动着生产关系的变化

❑ AI背景下，用户需要资源、数据、知识与其本人/其所用AI实现无缝衔接，这既是效率优先的要求，也是AI4S的学科研究中打开AI黑箱的重要保障，端到端的服务服务需求加大，向深度的数据服务、新的服务方式与服务内容提出了需求。



中国科学技术信息研究所.AI for Science创新图谱，2025年3月；  
XpertAI如何用科学文献赋能机器学习解释化学结构-性质关系，打破黑盒困境.<https://mp.weixin.qq.com/s/HLKpCgrv0VtIsLeDutRA7A>  
Gottweis, J., Weng, W.-H., Daryin, A., Tu, T, et al. Towards an AI co-scientist. <https://arxiv.org/pdf/2502.18864>

# 学术图书馆在积极地应对需求的变化



Advancing learning  
Transforming scholarship

Association of College & Research Libraries  
A division of the American Library Association

美国大学与研究图书馆协会4月初发布  
《学术图书馆环境扫描》

2025 Environmental Scan

ACRL Research Planning and Review Committee



图书馆与人工智能”的切入点:  
旨在成为图书馆员评估人工智能  
伦理使用的辅助工具,列出了关  
键考量点及应用情境

IFLA

Entry point to  
Libraries and AI

Internet Manifesto Annex I

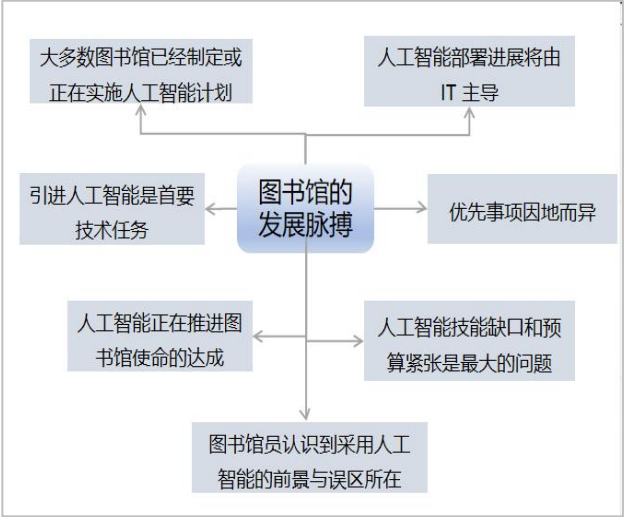
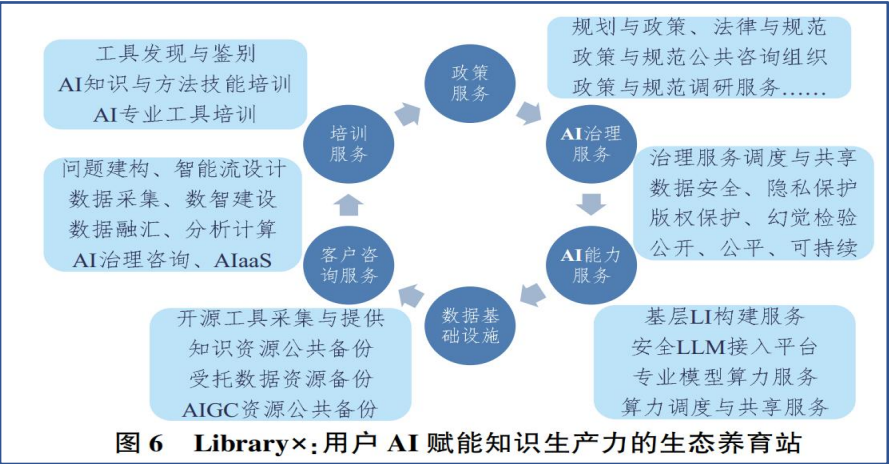


表 5 AI在高校图书馆的实践探索

| 类型       | 选项              | 人数  |
|----------|-----------------|-----|
| 自动化管理与服务 | 智能空间            | 662 |
|          | 机器人             | 352 |
|          | 自动化典藏与定位        | 344 |
|          | 自动化编目与分类        | 288 |
|          | 管理决策辅助          | 140 |
|          | 不了解             | 132 |
|          | 未开展相关应用         | 124 |
|          | 其他              | 17  |
| 教学与科研支持  | 智能搜索与学术助手       | 491 |
|          | 读者行为分析与预测       | 485 |
|          | 定制化知识资源推荐       | 410 |
|          | 人工智能素养培养与学术社群支持 | 340 |
|          | 不了解             | 232 |
|          | 未开展相关应用         | 137 |
|          | 其他              | 3   |
| 智能化辅助工具  | 数据分析与可视化工具      | 500 |
|          | 多媒体内容提取与分析工具    | 385 |
|          | 智能客服            | 381 |
|          | 跨媒体内容生成与创作工具    | 347 |
|          | 不了解             | 247 |
|          | 未开展相关应用         | 154 |
|          | 其他              | 4   |

2025 Environmental Scan.[https://www.ala.org/sites/default/files/2025-04/2025\\_ACRL\\_Environmental\\_Scan.pdf](https://www.ala.org/sites/default/files/2025-04/2025_ACRL_Environmental_Scan.pdf) ; 图书馆的发展脉搏.<https://img02.ma.scrmttech.com/18476/1812/resource/1726741375/图书馆的发展脉搏.pdf>; IFLA AI Entry Point for Libraries and AI.<https://repository.ifla.org/handle/20.500.14598/4034>; 张晓林:Library Inside:AI 赋能图书馆新质生产力的一种基础模型 赵兴胜,李小洁,程川生,等.认知与实践: AI技术在高校图书馆应用现状调研分析[J/OL].信息与管理研究,1-16[2025-01-07].<http://kns.cnki.net/kcms/detail/31.2118.G2.20250107.1047.002.html>.

# 新的学科服务岗位不断涌现

UC SANTA BARBARA

Recruit

[Home](#) / [Open Recruitment](#) / [Open Access Collection Strategist \(JPF02255\)](#)

Open Access Collection Strategist

Job #JPF02255

Library / Office of the Executive Vice Chancellor / UC Santa Barbara

We are no longer accepting applications for this recruitment. [Browse open recruitments](#)

POSITION OVERVIEW

Position title: Open Access Collection Strategist

Salary range: \$60,166-\$144,822

Percent time: 100%

Anticipated start: December 2022 or later

Position duration: Indefinite

FLORIDA STATE UNIVERSITY

News & Events

at College of Communication & Information

[Home](#) / [News](#) / [Academics](#) / [Events](#)

Reproducibility Librarian @ UF Libraries

Posted by Julia Orsola on September 18, 2020

The George A. Smathers Libraries at the University of Florida seeks a Reproducibility Librarian to develop an institutional strategy for education and support of transdisciplinary research reproducibility and open science. This position will be located in the Health Science Center Library (HSCU) in Gainesville. The Reproducibility Librarian is a year-round tenure-track library faculty position. The person who fills this position leads in designing and implementing a multifaceted program to enhance campus-wide efforts to promote and improve research reproducibility from design to dissemination. The Reproducibility Librarian participates as an active member of the library- and campus-wide teams to develop programming for and support information retrieval, storage, data science, and research. The incumbent will provide interdisciplinary information consultation services in a variety of modes, design workshops to promote research reproducibility, perform course-integrated instruction, and participate in the Health Science Center Library's teaching program. The position is responsible for special projects as assigned, such as new service development and evaluation, and development of web-based resources. The librarian works collaboratively in group efforts and maintains professional relationships with faculty, students and colleagues.

Artificial Intelligence Librarian

The University of Chicago Library - 4.0

Chicago, IL

Apply Now

Job Details

\$ Estimated: \$66.2K - \$79.8K a year

Qualifications

Senior level AI

Full Job Description

The University of Chicago seeks a new innovative team member to join the library as its first AI Librarian. This position will report to the Director of Digital Scholarship and play a vital role in advancing the University's engagement with artificial intelligence technologies in research and scholarship, as outlined in the Library's strategy. The AI Librarian will join a team of library experts committed to supporting open scholarship and will provide leadership in developing AI-focused services and resources. The successful candidate will collaborate with faculty, researchers, and students to explore the potential of AI in research, enhance data-driven scholarship, and contribute to the University's digital transformation initiatives.

The AI Librarian will lead the development of AI-based research support

- 加州大学圣巴巴拉分校“开放获取馆藏策划师”：
  - 评估机会，使学术交流系统向研究开放传播转变参与实施创新的馆藏模式和方法，管理UCSB开放获取（OA）出版基金。
  - 通过eScholarship、ORCID和Dryad促进开放出版选择；监测出版商和供应商提供的信息和数据，以确定开放获取相关的趋势和模式
  - 基于证据的馆藏和内容获取、管理和评估方面进行合作
  - .....

- （围绕科学数据，图书馆新职位嵌入科研全流程，参与数据管理活动）
  - 数据图书馆员（基于ALA JobLIST 2006-2018年4月发布的美国学术图书馆相关岗位）：
    - 包括数据素养馆员、研究数据馆员、数据可视化馆员、数据馆员、数据仓储馆员、数据服务馆员、数据与政府信息馆员等各类图书馆岗位名称。
  - 斯坦福大学图书馆：提供跨学科数字研究服务、数据服务，并形成人工智能计划。
  - 上海交通大学图书馆设置数据服务中心：汇聚研究数据、资源数据和管理数据，支撑科研范式变革；

- 芝加哥大学图书馆等发布“人工智能馆员”：
  - 在图书馆开展基于人工智能的研究支持服务开发；
  - 参与全校范围的人工智能与数据科学相关项目。
  - 参与制定促进人工智能合乎道德且负责任使用的战略及资源开发工作，并致力于提升师生数字素养

深圳大学图书馆

SHENZHEN UNIVERSITY LIBRARY

学习支持

新生指南

生成式人工智能专题

武汉大学图书馆

WUHAN UNIVERSITY LIBRARY

人工智能资源导航

四川大学图书馆 生成式人工智能专题

SICHUAN UNIVERSITY LIBRARY GENERATIVE ARTIFICIAL INTELLIGENCE

GenAI 工具与资源专题导航

Generative AI Tools and Resources

清华大学图书馆

Tsinghua University Library

借阅 资源 空间 学习支持 科研支持 概况

Library / LibGuides / GenAI-清华大学图书馆GenAI专题资源导航 / 首页 (Home)

GenAI-清华大学图书馆GenAI专题资源导航: 首页 (Home)

Search this Guide

[首页 \(Home\)](#) / [GenAI简介](#) / [GenAI工具](#) / [GenAI使用技巧](#) / [GenAI进阶使用](#) / [如何引用GenAI文字](#) / [GenAI出版伦理](#) / [信息素养微课堂: AI专题](#) / [图书馆AI助手](#)

# 科技信息服务行业中传统信息服务商、科技新贵纷纷涌入，渐成规范



## Generative AI in Higher Education

The Product Landscape

基本价值主张相似，以加速或丰富研究生命周期中发现、理解和创造三个关键阶段为核心目标

本质都是在商业大模型基础上的应用程序

产品特性与功能趋同性严重

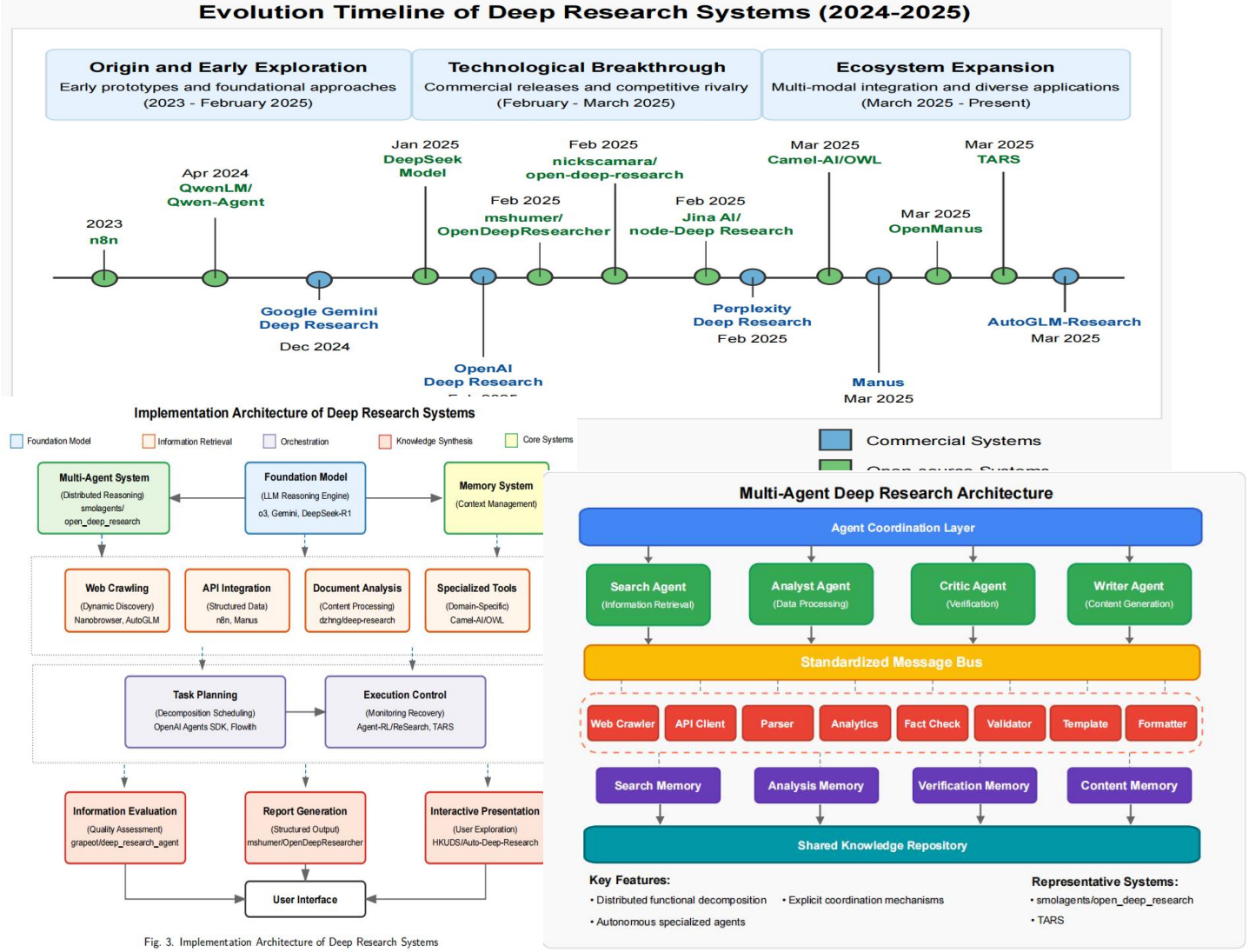
产品的学科纵深度与业务纵深度日趋加强

生成式内容的合理使用挑战日趋严重

与更大平台的融合正在成为优选路径

March 7, 2024

ITHAKA SR



# 在猛烈拥抱AI的同时需警惕潜藏的风险

## 1 内容准确性与严谨性

AIGC属于模糊相似中寻找共性，对于同一问题的结果并不完全一致，也加大了严谨性论证的难度。AIGC内容其实是对各种学习资源的重组，在模糊相似中找共性，并不判断对于生成内容本身的合理性。AIGC生成的内容是用户输入内容语义上关联性最大的，其内容未必正确，该特质就决定了可能生成大量虚假信息。而且生成式 AI 并不会提供信息的出处，带来的误导性更大。

## 3 隐私与伦理道德

使用AIGC时自身隐私信息无差别被人工智能捕获并学习，导致隐私泄露，被侵犯知识产权。滥用AIGC生成的虚假信息，不经鉴别地传播和出版，导致侵犯他人隐私或知识产权，进而也带来信息欺诈风险。

## 2 学术失信与公平失衡

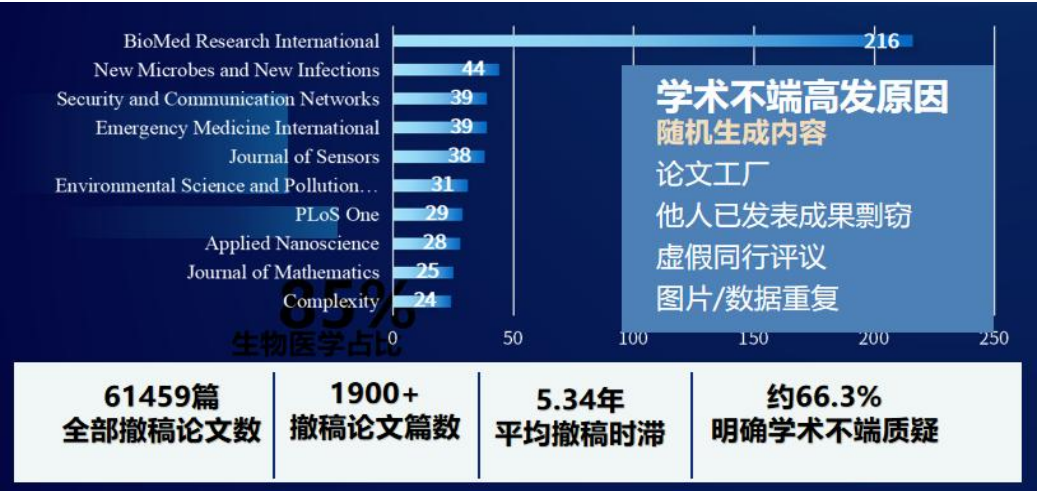
1.并非每个人都能使用到AIGC工具，使用权并不平等。2.使用AIGC的抄袭监测难度更高，学术不公现象会只增不减，由此开展的学术不端调查耗费时间与精力也更巨大。

## 4 算法歧视与偏见

1.训练数据受限和数据偏差会影响输出模型  
2.无法区分事实和虚构的缺陷，甚至生成的文本中能提取到充斥着人性弱点的信息。

## 5 技术成瘾产生创新依赖

AIGC全面的能力和对现实的补偿效应带有强烈的成瘾风险，科研工作者一旦离开AI就不会思考，或者被AI牵着鼻子走。



nature

Explore content ▾ About the journal ▾ Publish with us ▾ Subscribe

[nature](#) > [news](#) > article

NEWS | 24 April 2025

## Science sleuths flag hundreds of papers that use AI without disclosing it

Telltale signs of chatbot use are scattered through the scholarly literature — and, in some cases, have disappeared without a trace.

□ 数据建设规范化、丰富化，无论对本身产品及服务的增强，还是试水新的数据要素服务，都是新机遇

表 2 出版行业内容许可交易记录<sup>[5][6]</sup>

| 出版商                              | 购买者                 | 交易内容类型                                        | 内容用途                                            |
|----------------------------------|---------------------|-----------------------------------------------|-------------------------------------------------|
| Wiley                            | 大型科技公司              | 学术与专业图书                                       | 训练 LLM                                          |
|                                  | Potato              | 专注于生物类、经同行评议的文献内容                             | 用于构建 Potato 工具，包括自动论文评审以及实验室指南生成工具              |
|                                  | Perplexity          | 护理、商业和工程领域的教育资源，包括电子书                         | 订阅 Perplexity 专业版机构用户能访问购买的合作 Wiley 教育资源集。      |
|                                  | Amazon web Services | 科技期刊全文                                        | 用于针对医疗保健和生命科学的开源工具包。2025 年合作推出科学文献搜索的生成式人工智能代理。 |
| Taylor & Francis/ <u>Informa</u> | Microsoft           | 图书（可能未包含期刊）                                   | 用于改进 AI 系统的相关性与性能                               |
|                                  | 大型科技公司              |                                               |                                                 |
| OUP                              | 未知                  | 未知                                            | 合作开发 LLM                                        |
| CUP                              | 未知                  | 未知                                            | 计划用于训练 LLM                                      |
| SAGE                             | 未知                  | 出版的高质量科技内容                                    | 用于训练 LLM                                        |
| <u>De Gruyter Brill</u>          | 未知                  | 未明确，但公布信息特别聚焦图书                               | 训练和测试基础 LLM，将开发工具和应用，并开发授权内容数据库验证 AI 内容准确性      |
| NEJM Group                       | <u>OpenEvidence</u> | 从 1990 年以来的新英格兰医学杂志及相关出版物                     | RAG 模型支持 <u>OpenEvidence</u> 平台，提升面向医生的服务       |
| AAAS                             | <u>ProRata</u>      | 先期包括 Science Advances 期刊同行评议的 OA 论文及科学杂志的新闻故事 | 用于 Gist.ai 搜索引擎提高透明度与可信度，聚焦于少量高质量的内容            |

高质量数据支撑LLMs

版税分配模式重新计量

AI许可解决方案

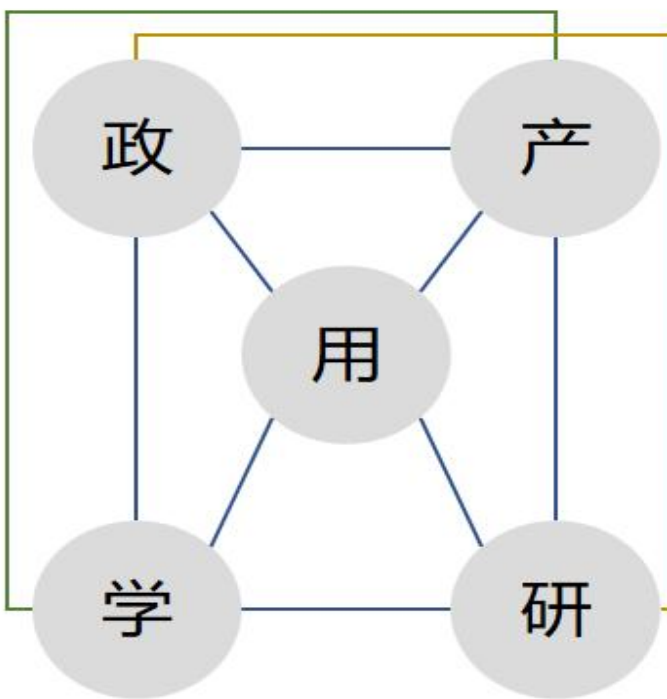
针对特定行业的集体版权许可

垂直领域模型的联盟生态共建

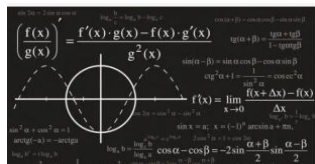
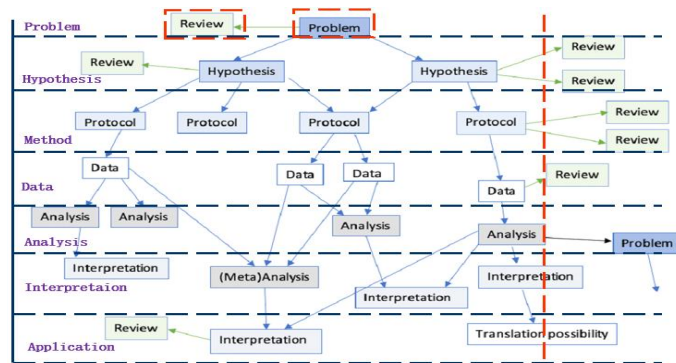
# 始于文献，超越文献

## 覆盖原生资源->解构数据->知识数据的链式数据体系

## AI与科技信息服务纵深化需求的结合



### 持续推动科技创新全谱段基础资源建设（渠道、类型、模态）

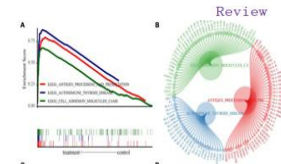


**Transforming Research Paradigms Through AI: A Study on Artificial Intelligence Literacy and Education Strategies from an Interdisciplinary Perspective**  
Cai Yingchun<sup>1,2</sup>, Yu Chenlin<sup>1,2</sup> (1 Shanghai International Studies University Library; 2 Digital Scholarship Center of Shanghai International Studies University)

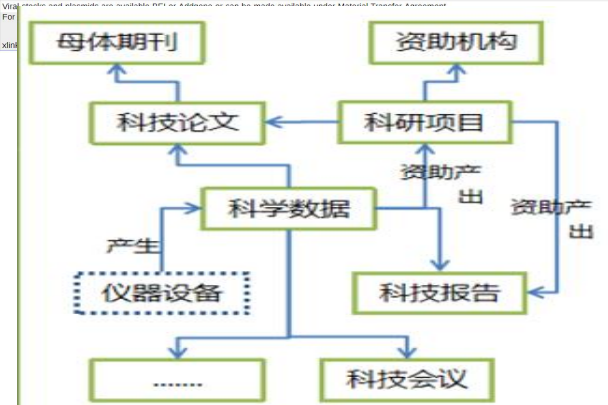
**Abstract** Building an AI literacy education and training system for artificial intelligence (AI) for research is crucial in cultivating more scientific researchers equipped with technical knowledge. In addition, AI literacy education is also an important part of the national science and technology education policies, courses, and training programs in major countries; the proposed system is believed to be able to help cultivate more scientific researchers with technical knowledge, scientific research responsibility, interdisciplinary thinking, and emotional connection. AI literacy strategies are crucial for the development of scientific research, and the proposed system is believed to be able to help cultivate more scientific researchers in educational institutions, enterprises, and public welfare organizations like libraries and communities. They are also believed to be able to help cultivate more scientific researchers in the scientific field, and help cultivate more scientific researchers in the scientific field, and help cultivate more scientific researchers in the scientific field. They not only promote the application of AI technology in the scientific field, but also help cultivate more scientific researchers enabling them to have the ability to use AI technology for interdisciplinary research.

**Keywords** AI for Research (AI4R), Interdisciplinary, Artificial Intelligence Technology, Core elements, AI literacy education, AI literacy training system, AI literacy education and training system

**Keywords** AI for Research (AI4R), Interdisciplinary, Artificial intelligence literacy, Core elements, Education and training



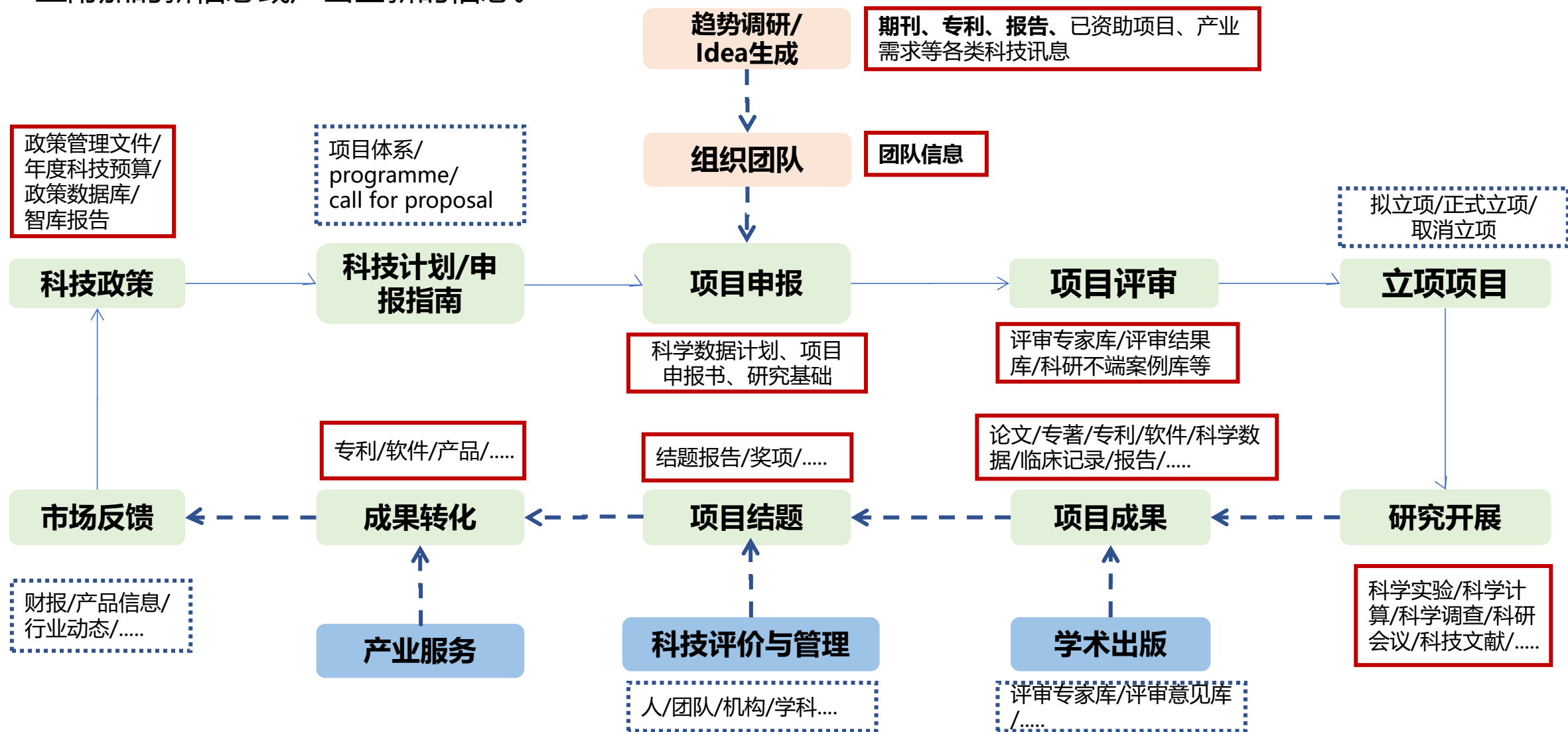
## 细化多模态原始数据解构重组与知识向量化



## 纵深数据的知识化挖掘与链式关联

# 立足文献，向左向右持续扩展

- **原生数据**：从完整的科技创新价值链看原生资源的扩展，并关注各价值环节在利用信息的同时会针对旧信息衍生附加的新信息或产出全新的信息。



# 立足文献，向下深探

□ **原生数据**：文献是AI4S的学科服务中打开AI黑箱的重要保障，需要注意科技文献不仅只有TOC信息，图像、公式、软件、数据、基金信息等等，解析文献的过程实际上是在重现一个研究的过程

所在期刊

标题

作者信息

机构信息

出版变更

www.aging-us.com

AGING 2020, Vol. 12, No. 4

Identification of biomarkers related to CD8<sup>+</sup> T cell infiltration with gene co-expression network in clear cell renal cell carcinoma

Jiaxing Lin<sup>1,2</sup>, Meng Yu<sup>2,3</sup>, Xiao Xu<sup>3</sup>, Yutao Wang<sup>1</sup>, Haotian Xing<sup>1</sup>, Jun An<sup>1</sup>, Jieping Yang<sup>1</sup>, Chaozhi Tang<sup>1</sup>, Dan Sun<sup>1</sup>, Yuyan Zhu<sup>1</sup>

<sup>1</sup>Department of Urology, The First Hospital of China Medical University, Shenyang 110001, Liaoning, China  
<sup>2</sup>Department of Reproductive Biology and Transgenic Animal, China Medical University, Shenyang 110001, Liaoning, China  
<sup>3</sup>Department of Pediatric Intensive Care Unit, The Shengjing Hospital of China Medical University, Shenyang 110001, Liaoning, China

\*Equal contribution

Correspondence to: Yuyan Zhu; email: yzyzhu@cmu.edu.cn  
Keywords: clear cell renal cell carcinoma (ccRCC), CIBERSORT, weighted gene co-expression network analysis (WGCNA), CD8<sup>+</sup> T cells, CCL5  
Received: October 1, 2019 Accepted: February 4, 2020 Published: February 20, 2020

Copyright: Lin et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY 3.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**ABSTRACT**

Clear cell renal cell carcinoma (ccRCC) is an extremely common kind of kidney cancer in adults. Immunotherapy and targeted therapy are particularly effective at treating ccRCC. In this study, weighted gene co-expression network analysis and a deconvolution algorithm that quantifies the cellular composition of immune cells were used to analyze ccRCC expression data from the Gene Expression Omnibus database, and identify modules related to CD8<sup>+</sup> T cells. Ten hub genes (*LCK*, *CD2*, *CD3D*, *CD3G*, *IRF1*, *IFNG*, *CCR5*, *CD8A*, *CCL5*, and *CXCL9*) were identified by co-expression network and protein-protein interactions network analysis. Datasets obtained from The Cancer Genome Atlas were analyzed and the data revealed that the hub genes were meaningfully up-regulated in tumor tissues and correlated with promotion of tumor progression. After Kaplan-Meier analysis and OncoPrint meta-analysis, *CCL5* was selected as a prognostic biomarker. Finally, the experimental results show that reduced expression of *CCL5* decreased cell proliferation and invasion in the ccRCC cell line. Various analyses were performed and verified, *CCL5* is a potential biomarker and therapeutic target which related to CD8<sup>+</sup> T cell infiltration in ccRCC.

**INTRODUCTION**

Renal cell carcinoma (RCC) is one of the most common malignant tumors, ranking seventh among male malignant tumors and tenth among female malignant tumors [1]. RCC accounts for 80% of all kidney cancer, and clear cell renal cell carcinoma is the most common subtype of renal cell carcinoma [2]. Smoking, obesity, and high blood pressure increase the risk of kidney cancer [3]. In recent years, immune checkpoint inhibitors have become the standard for first-line treatment of renal cell carcinoma [4–6]. However, no

specific molecular markers have been for immunotherapy of renal cell carcinoma [2]. Therefore, the exploration of immune-related molecular markers is an important focus of renal cell carcinoma research.

RCC are prone to immune infiltration, and the characteristics of tumor microenvironment strongly alter the response to immunotherapy [7]. CD8<sup>+</sup> T cells contribute to tumor adaptive immunity. Among the immune cells of ccRCC, CD8<sup>+</sup> T cells account for the largest proportion [8]. In most solid tumors, highly infiltrating CD8<sup>+</sup> T cells are beneficial to tumor

www.aging-us.com 3694 AGING

Significant progress has been made in the study of the molecular basis of ccRCC by using animal tumor models, *in-vitro* cell lines, and clinical samples. However, the complexity of the ccRCC micro-environment requires further analysis and larger datasets. We used gene expression matrix to construct the co-expression network and calculate the infiltration

level of T cells, and correlations were determined to identify the genes most related to CD8<sup>+</sup> T cells. The gene enrichment analysis of the selected module shows that it is a highly immune-related module. The most highly connected genes in the co-expression network and the protein-protein network were considered hub genes (*LCK*, *CD2*, *CD3D*, *CD3G*, *IRF1*, *IFNG*, *CCR5*,

**Figure 9. GSEA and experiment of CCL5.** (A) The above picture shows the enrichment fractional broken line of the three pathways, and the lines in the lower figure correspond to the genes of each pathway. (B) The circle diagram shows three enrichment pathways and the core genes that play a role in the enrichment process. The larger the circle corresponding to each gene, the larger the rank metric score value. (C) The relative mRNA levels of normal control and si-CCL5 of 769-P cells. (D) The results of CCK-8 assay showed decreased proliferation ability of 769-P cells when treated with si-CCL5. (E) The invasion assay results showed decreased invasion ability of 769-P cells treated with si-CCL5.

www.aging-us.com 3703 AGING

**Cell culture and siRNA-PTEN construction**

The 769-P cells were cultivated in DMEM with 10% of fetal calf serum, 100 U/ml penicillin and 100 U/ml streptomycin. The cells were cultured at 37 °C, with 5% CO<sub>2</sub>. The CCL5-310 siRNA sequences were as follows: 5'-GCUGAACAAGGGAAGCUUTT-3' 5'-AAGCUU GCCCUUGUUCAGGTTT-3'. The CCL5-286 siRNA sequences were 5'-GACAGGAUUCUUGUUAUG ACTT-3' and 5'-GUCAUACAGAAUUCUGGTTT-3'. The CCL5-240 siRNA sequences were: 5'-UC GUCCACAGGUAAGGATT-3' 5'-AUCCUUGAC CUGUGGAGGATT-3'.

**RNA extraction and quantitative RT-PCR**

RNA samples were extracted from cells with the FastPure Cell/Tissue Total RNA Isolation kit (Vazyme Biotech) and they were reverse-transcribed by HiScript III RT SuperMix for qPCR (Vazyme Biotech). To determine the relative transcript level, PCR was quantified in real-time using a LifeECO PCR machine (BIOER Technology Co, Ltd). SYBR Green was used as the fluorophore. The CCL5 primers sequences were as follows: 5'-CTC ATTGCTACTACTGCGCTCTCGGCTCCGTCG-3' and 5'-GCTCATCTCCTAAAGAGTTGATGATCTC -3'. The PCR parameters were: 95 °C 5min, followed by 50 cycles of 30 s and 1 min at 60 °C. Each sample was measured in three independent reactions. The threshold values of each sample/primer were determined, and the average error and standard error were calculated. Melting curve analysis was performed, and the mRNA expression levels were normalized against that of β-actin.

**Cell proliferation analysis and invasion experiment**

769-P cells were cultivated in five 96-well cell culture plates (1500 cells/well) respectively for 5 days. A volume of 10 μl of CCK-8 solution was added to each well in a plate at an interval of 24 h. The cells were cultured for 3 h in 37 °C, and then a photometry test was performed at wavelength of 450 nm. The 769-P cells were then transferred into the top layer holes in a basement membrane plate with wells containing the medium without serum. The lower cavity was filled with 12% fetal calf serum medium to allow chemical attraction. The medium was removed after 12-hour incubation and the implant in the upper cavity was discarded. Cells invading the lower cavity were fixed with 700 ml of 4% PFA and stained with crystal for 1 h to visualize. The cells in the lower cavity were counted and normalized to the number in control conditions to measure the relative invasion capacity.

**Abbreviations**

ccRCC: Clear cell renal cell carcinoma; WGCNA: Weighted gene co-expression network analysis; CIBERSORT: Cell-type identification by estimating relative subsets of RNA transcripts; TILs: Tumor-infiltrating immune cells; PPI: Protein-protein interactions; GEO: Gene Expression Omnibus; TCGA: The Cancer Genome Atlas; GSEA: Gene set enrichment analysis.

**ACKNOWLEDGMENTS**

We want to thank TCGA, GEO, TIMER, and TISIDB for free use.

**CONFLICTS OF INTEREST**

The authors declare that there are no conflicts of interest.

**FUNDING**

This work was supported in part by the National Natural Science Foundation of China (grants 81672523, 81472403, 81272834, and 31005072), the 2018 Support Plan for innovative talents in Colleges and Universities of Liaoning Province, the 2018 "million talents Project" funded Project of Liaoning Province, and the 2019 Key R & D projects of Shenyang.

**REFERENCES**

1. Siegel RL, Miller KD, Jemal A. Cancer statistics, 2016. *CA Cancer J Clin.* 2016; 66:7–30. <https://doi.org/10.3322/caac.21332> PMID: 26742998
2. Escudier B, Porta C, Schmidinger M, Rioux-Leducq N, Bex A, Khoo V, Grünwald V, Gillesen S, Horwich A, and ESMO Guidelines Committee. Electronic address: [clinicalguidelines@esmo.org](mailto:clinicalguidelines@esmo.org). Renal cell carcinoma: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. *Ann Oncol.* 2019; 30:706–20. <https://doi.org/10.1093/annonc/mdz056> PMID: 31987338
3. Tahbaz R, Schmid M, Merseburger AS. Prevention of kidney cancer incidence and recurrence: Lifestyle, medication and nutrition. *Curr Opin Urol.* 2018; 28:62–79. <https://doi.org/10.1097/MOU.0000000000000454> PMID: 29059103
4. Conskline B, Hurwitz ME. Current Status and Future Directions of Immunotherapy in Renal Cell Carcinoma. *Curr Oncol Rep.* 2019; 21:34.

www.aging-us.com 3706 AGING

基金资助

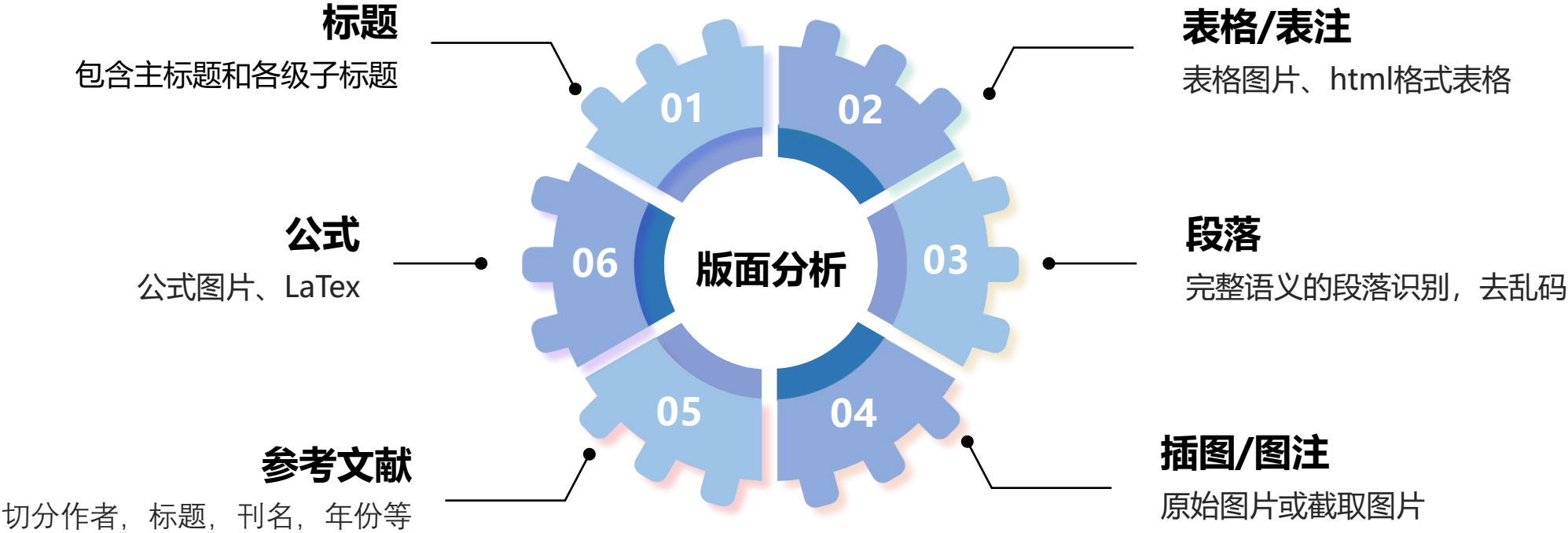
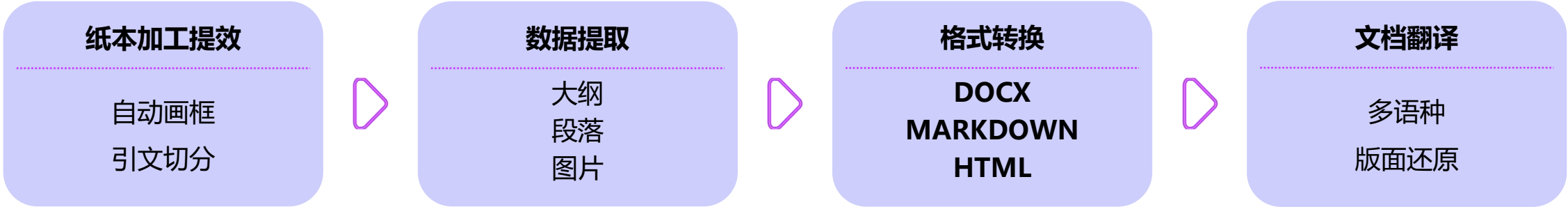
引文

文本内容、文本结构

图像、应用的数据、使用的仪器设备/研究软件

# 立足文献，向下深探-利用AI全面升级技术体系

解构数据：AI加持的智能版式识别全面提升原生数据的解构粒度与性能，从描述元数据加工生产到知识要素解构



从单一文本到多元知识关联，支撑知识推理。

Design of an energetic particle radiation diagnostic spectroscopy system based on national core chips and Qt on Linux in EAST

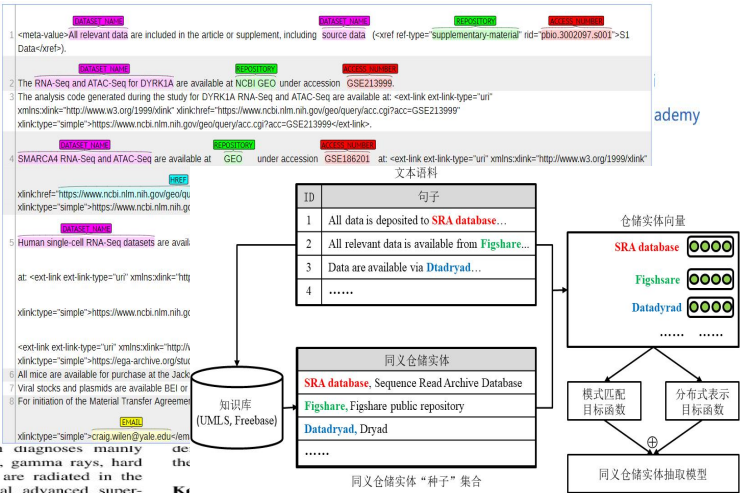
Ying-Ying Zheng<sup>1</sup> Zi-Han Zhang<sup>2</sup> Qiang Li<sup>2</sup> Hong-Rui Cao<sup>3</sup> Yong-Qiang Zhang<sup>1</sup> Jin-Long Zhao<sup>3</sup> Qi-Ping Yuan<sup>3</sup> Bing-Jia Xiao<sup>3</sup>  
Ling-Ling Yan<sup>1</sup> Jian-Qiu Zhu<sup>1</sup>

1.Institute of Plasma Physics,Hefei Institutes of Technology of China,Hefei 230026,China; 2.Hefei University of Technology,Hefei 230031,China;Hefei University of Technology of Sciences,Hefei 230031,China

Received: 20 February 2021 / Revised: 19 May 2021  
© China Science Publishing & Media Ltd. (CSPMP)  
Nuclear Society 2021

**Abstract** Energetic particle radiation diagnoses mainly detect the particles (such as neutrons, gamma rays, hard X-rays, and escaping electrons) that are radiated in the discharge process of the experimental advanced superconducting tokamak device to characterize the operating state of the plasma in real time. The upgrading of these diagnoses requires new instruments based on national (here, “national” means developed and produced by a Chinese company) core chips and open-source software with advanced digitization, a high sampling rate, and a high time resolution. The new spectroscopy system designed in this study adopts the national field-programmable gate array (FPGA) and an analog-to-digital converter as the core chip, and it is developed using Qt on Linux. The communication between the FPGA and embedded controller occurs via a high-speed peripheral component interconnect eXtension for instrument express protocol with a direct memory access mode. On this basis, the time resolution of the system is improved from 2 to 1 ms, the maximum channel address is increased to 4096, and the sampling rate is increased from 10 to 80 Msps. Calibration experiments of the spectroscopy system with <sup>152</sup>Eu and <sup>137</sup>Cs sources

This work was supported by the National MCF Energy Research and Development Program of China (No. 2018YFB030100) and the National Natural Science Foundation of China (No. 12075285).



1 Introduction

The experimental advanced superconducting tokamak (EAST) is the most advanced experimental superconducting tokamak platform in China [1, 2]. Several new operating parameter records have been achieved by EAST since the first discharge operation in 2006; for example, 100-s pulse H-mode steady-state operation was realized in 2017 [3, 4], and the plasma center electron temperature reached one million degrees ( $T_{e0} = \sim 10$  keV) in 2018 [5]. Notably, certain components of the device aged as the service lifetime increased [6].

To enable upgrading of the plasma diagnostic system and plasma control system (PCS), the accuracy and stability of these diagnoses should be improved, and the PCS feedback control speed should be increased to realize real-time and stable control [7]. In these diagnoses, the energetic particle radiation diagnostic spectroscopy (EPRDS) system is important for measuring the energy spectrum of

关联数据集

中国科学院数据索引  
CHINA RESEARCH DATA INDEX

首页 数据浏览 统计分析 动态信息 注册与出版 关于我们

登录

基于国产核心芯片和Linux下QT软件的高能粒子辐射诊断谱仪系统设计

数据集

Design of an energetic particle radiation diagnostic spectroscopy system based on national core chips and Qt on Linux in EAST

Ying-ying Zheng , Zi-han Zhang , Qiang Li , Hong-rui Cao , Yong-qiang Zhang , Jin-long Zhao , Qi-ping Yuan , Bing-jia Xiao , Ling-ling Yan , Jian-qiu Zhu

点击下载数据集 收藏 分享

创建日期: 2022-10-23 发布日期: 发布日期: 最近更新日期: 2022-11-24

数据集基本信息

摘要: 提供已发表论文中的图表类原图供开放获取, 包括: 高能粒子辐射诊断谱仪系统结构图, FPGA设计结构逻辑框图, 输入信号幅值与能谱地址的线性拟合图, 实验测量和标定后的能谱图。

数据仓储于 ScienceDB

引用量 0 (来自所有数据库)

下载量 5 (来自所有数据库)

访问量 44 收藏量 0

版本 V1

数据集关联信息

关联论文

科慧 Sci-Fund

首页 申报指南 free 浏览导航 统计分析 服务资源 new 欢迎 Nationa...

EAST等离子体稳态实时控制集成系统研究

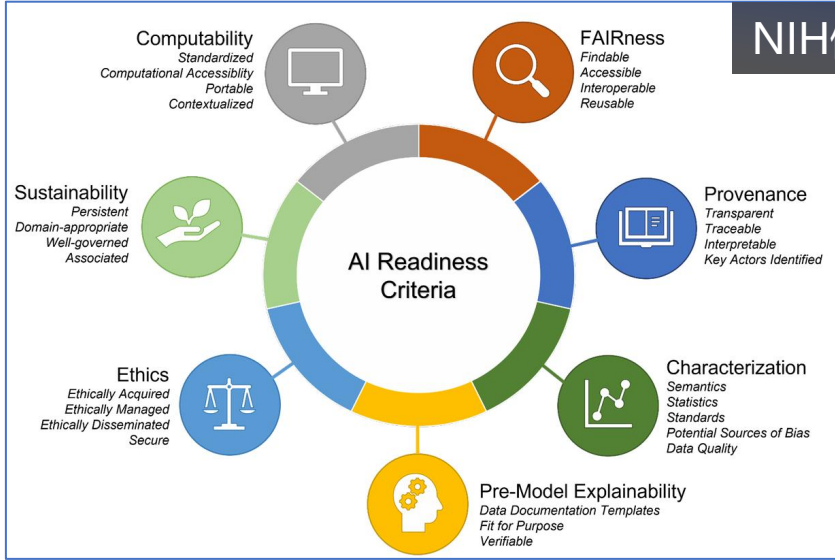
|      |                              |         |                         |
|------|------------------------------|---------|-------------------------|
| 项目编号 | 12075285                     | 负责人     | 袁旗平                     |
| 承担机构 | 中国科学院合肥物质科学研究院               | 受资国家/地区 | 中国 (安徽)                 |
| 资助机构 | 中国自然科学基金委数理科学部               | 项目类型    | 研究项目 (面上项目)             |
| 项目级别 | 国家级                          | 项目状态    | 进行中                     |
| 批准经费 | 90,304.00USD (640,000.00CNY) | 立项年份    | 2020                    |
| 开始年份 | 2021                         | 项目起止    | 2021-01-01 - 2024-12-31 |
| 原始学科 | 等离子体物理-磁约束等离子体               |         |                         |

展开更多信息

项目产出

期刊论文

# 标准规范与AI技术同行，推动AI for Data<->Data for AI双循环



## NIH侧重数据质量与功能性

对照FAIR标准，AI-Readiness 是指数据具有**机器可读性、可靠性、准确性、可解释性、预测性**，并且对将来的人工智能应用是**可访问的**。并提供了一系列对照工具。

## Gartner强调场景适用

AI-Ready是指数据能够充分表征使用场景中的多样化模式、错误、异常值以及特定用途所需的意外情况，用于训练或运行AI模型。这是一种基于元数据的系统性流程与实践，通过数据对齐、验证及治理来实现



Source: Gartner  
© 2024 Gartner, Inc. and/or its affiliates.  
All rights reserved. GTS\_3255850

EBOOK

## The Big Book of AI-Ready Data



## data.world侧重数据治理与验证

AI-Ready数据需经历严格的验证过程，如：格式检查、范围检查、引用完整性、名称验证、唯一性检查、一致性测试、问题跟踪系统、安全测试。

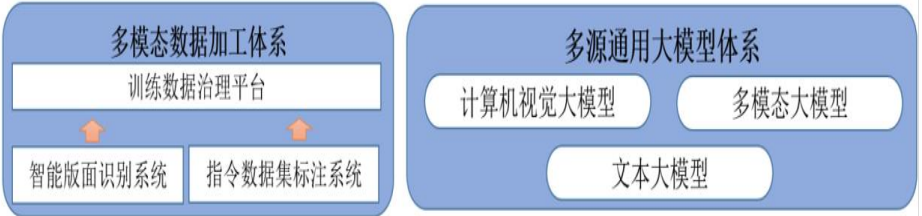
## NOAA四级成熟度模型

从**数据分级与评估**的视角，对照**数据质量、访问、文档**三个维度要素形成标准

|                      | 数据一致性角度                           | 数据访问角度                                     | 元数据角度                            |
|----------------------|-----------------------------------|--------------------------------------------|----------------------------------|
| 级别: 0 (Not AI-Ready) | 未进行内部一致性的检查                       | 仅通过请求或订单系统对公众开放使用                          | 没有数据字典可用，或数据字典不是机器可读格式（例如PDF）    |
| 级别: 1 (Minimal)      | 手动一致性检查                           | 只有一个非程序化的访问选项，例如文件下载                       | 数据字典是机器可读格式（例如CSV、XML、JSON）      |
| 级别: 2 (Intermediate) | 一致性检查部分自动化，部分结果有文档记录              | 多种访问方式，且至少包括一个可编程访问方法（API）                 | 数据字典使用机器可读的元数据标准                 |
| 级别: 3 (Optimal)      | 完全自动化的内部一致性检查和报告；且考虑了与社区数据集的外部一致性 | 除多种交付选项（下载、API）外，还有基于云和高性能计算提供的数据即服务（DaaS） | 机器可读的元数据标准；参数与行业标准或其他机构数据集标准协调一致 |

# 标准规范与AI技术同行，推动AI for Data<->Data for AI双循环

❑ **解构数据与知识数据：** 依托AI技术全面升级数据治理与分析工具，按照工具链思路进行体系化建设与组配，构建数据治理与分析AI智能体。



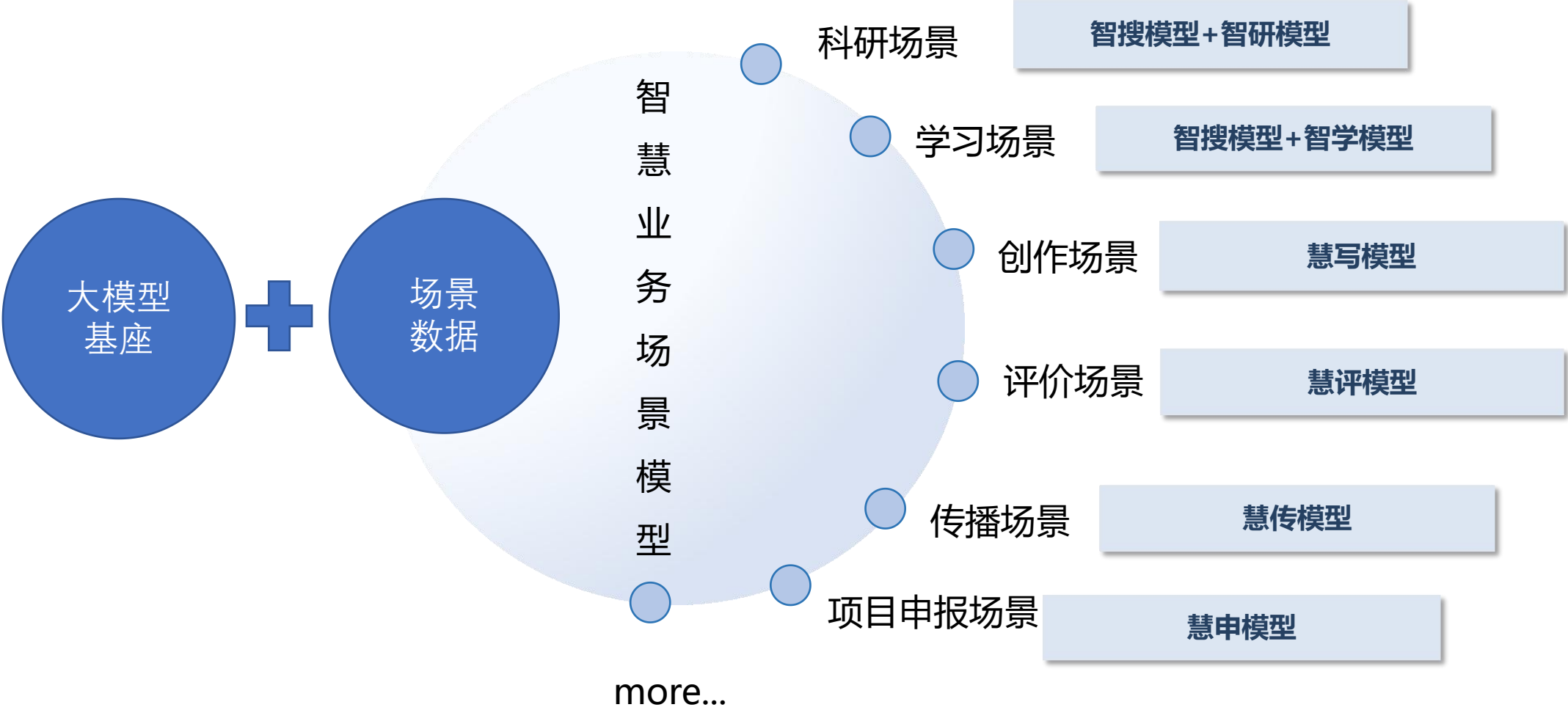
# 标准规范与AI技术同行，推动AI for Data<->Data for AI双循环

❑ **解构数据与知识数据：**基于“通用大模型+面向场景任务的小模型”策略，形成系列工具，利用向量数据库存储并索引高维空间数据，支撑智能计算和复杂关系计算，提升性能与效果



# 超越文献保障，迈向智能服务

□ 基座大模型与场景数据的碰撞，推动文献资源服务转向工具服务、组织服务



# 场景与业务驱动的智慧服务矩阵重组与创新

面向学生、教师、研究者、科研管理人员等不同创新主体角色，以学习、研究、管理、决策等关键环节的实际业务场景为驱动，打造系列工具集，沉淀不同层次的一体化智能服务能力。

|       |          |         |           |         |      |          |       |         |         |
|-------|----------|---------|-----------|---------|------|----------|-------|---------|---------|
| 服务组件  | 基础计算模型   |         | 工具模型      |         | 实体画像 |          |       | 数据集模型   |         |
|       | 文献多维共引聚类 | 新兴词监测   | 发展趋势分析    | 热点演化分析  | 学者   | 机构       | 期刊    | 学者产出    | 机构产出    |
|       | 引文分析     | 合作网络    | 文献导出工具    | 热点生成工具  | 主题   | 地区       | 行业    | 行业数据集   | 领域数据    |
| 工具/服务 | 文献保障服务   |         | 科研全流程支持服务 |         |      | 科技决策支持服务 |       |         |         |
|       | 知识服务平台   |         | 万方选题      | 慧写      | 刊寻   | 排版       | 科慧    | 灵析      | 学者知识脉络  |
|       | 知识检索     | 多粒度资源揭示 | 回溯学术脉络    | 生成标题    | 智能匹配 | 期刊论文排版   | 基金检索  | 主题分析    | 学者知识脉络  |
|       | 便捷获取保障   | AI文献速读  | 追踪研究重点    | 撰写大纲    | 学科速览 | 参考文献校对   | 多维聚类  | 学者分析    | 关键词知识脉络 |
| 技术    | AI增强价检索  | 章节脉络    | 拓展研究边界    | 推荐段落    | 精准查刊 | 同屏直观比对   | 可视化分析 | 机构分析    | 机构知识脉络  |
|       | 智能翻译     | 可视化关联分析 | 发掘新兴主题    | 文本润色    | 多维分析 | 学位论文排版   | 案例分析  | 学科分析    | 标准管理平台  |
|       | 跨语言检索    | 全文对话    | 新颖性评测     | 智能翻译    | 征稿启事 | 公文排版     | 订阅管理  | 地区分析    | 检索/导航   |
|       | 全新学术评价   | 智能推荐    | 优质文献推荐    | 引文推荐    | 分析报告 | 多格式导出    | 便捷导出  | 期刊分析    | 有效性查询   |
| 数据    | 分布式搜索索引  |         | 知识组织与关联   | 大数据计算系统 |      | 知识图谱     | 语义模型  | RAG增强检索 | 语义标引    |
|       | 期刊       | 学位      | 会议        | 专利      | 科技成果 | 科技报告     | 地方志   | 视频      | 引文库     |
|       |          | 学者      | 机构        | 基金      | 科学数据 |          |       |         |         |

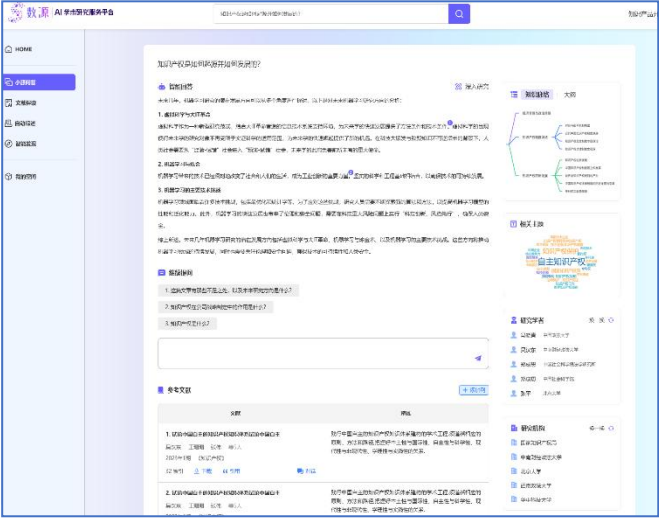
# 检索发现式的原文服务迈向认知引擎构建

在传统检索基础上，深度融合大语言模型，构建「语义理解-知识关联-动态优化」三位一体的认知智能引擎，实现从关键词匹配到语义向量认知的范式跃迁。

AI增强检索



循证式问题解答



智能文献伴读



AI增强检索

答案综述

参考文献快读

可视化知识脉络

文献伴读

全文精炼

术语解释

单篇问答

多篇对比

文献综述

阅读快报

智能翻译

大模型全文标引

语义检索模型

检索排序效果监测

科研实体、研究问题、创新点、实验方法、  
缺点和不足等

改进检索排序算法，实现真实需求理  
解与精准检索

首位命中率、前三条结果点击率、首屏  
效率值、翻页点击率

# 感知终端用户场景，提供链路式工具集自组配机制

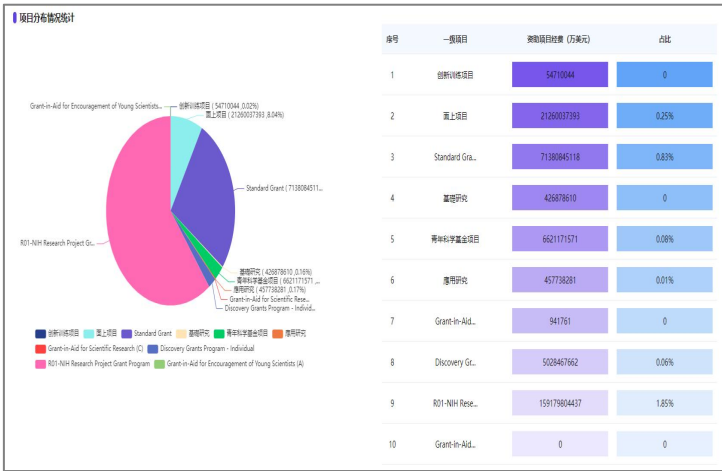
□ 智研：贯穿科研学习全流程，集“搜、读、写、研”为一体的智慧研究服务平台。

- **研学探知**：文献检索、文献阅读、阅读推荐、科技资讯推送。
- **知识管理**：个人知识库构建，文献收藏、标签管理、阅读笔记、订阅推荐、学术成果管理。
- **创作投稿**：论文选题、智能创作、论文投稿、智能排版、智能审校、论文检测。
- **研习手账**：科研计划个性化生成、日历提醒、科研任务进度提醒等。

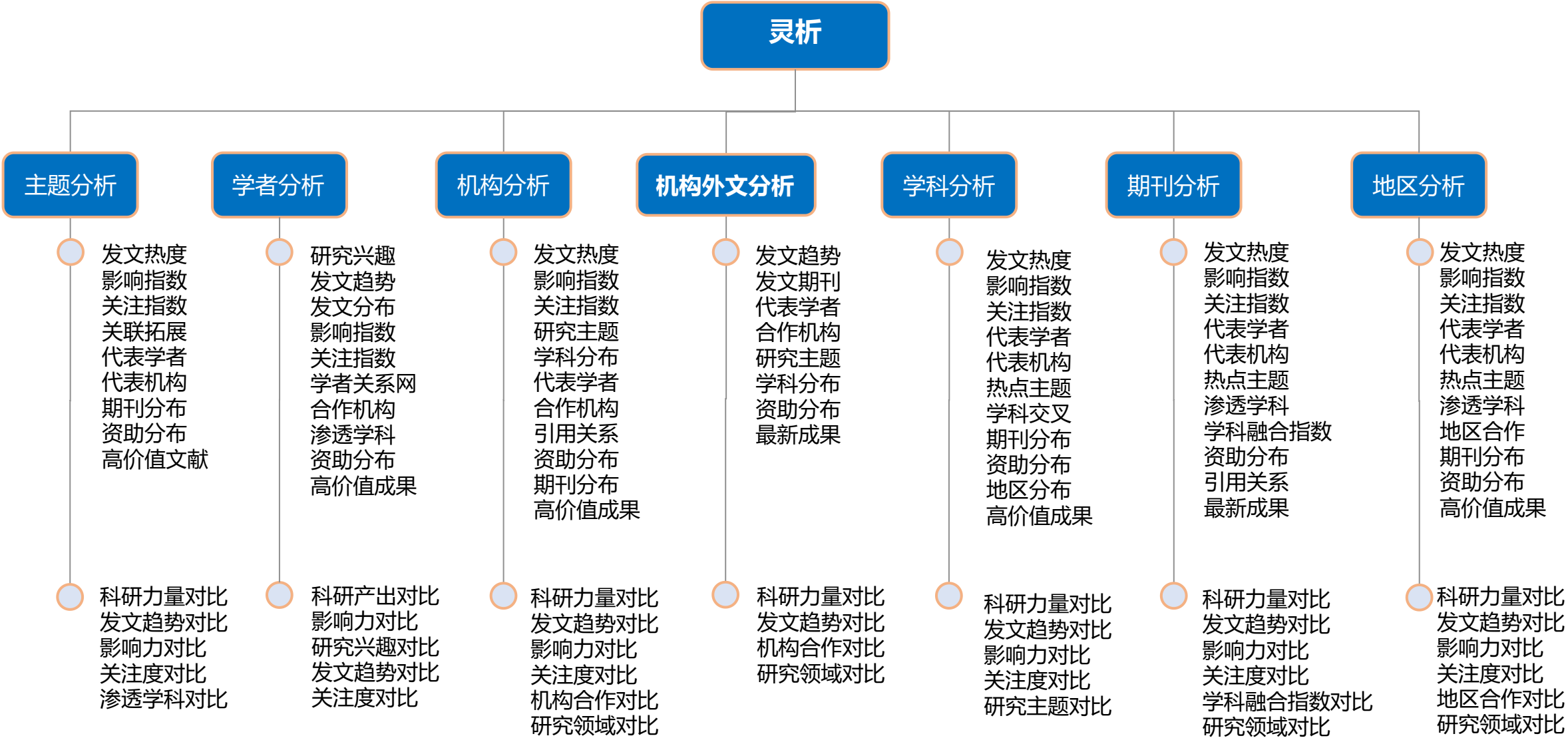


# 从文献发现服务迈向更多业务场景

❑ **科研项目全周期：**面向科研项目的申报与实施，串联基金申报提醒、选题推荐、资助态势分析、推荐申报代码、立项申请书辅助撰写（国内外研究现状、前期研究基础）、诚信检测、成果汇集、科学数据仓储等全流程服务。



# 面向创新主体的多维立体评估



# 面向学科跟踪与评估需求，分别构建灵析与学科发展支撑服务

找科研方向



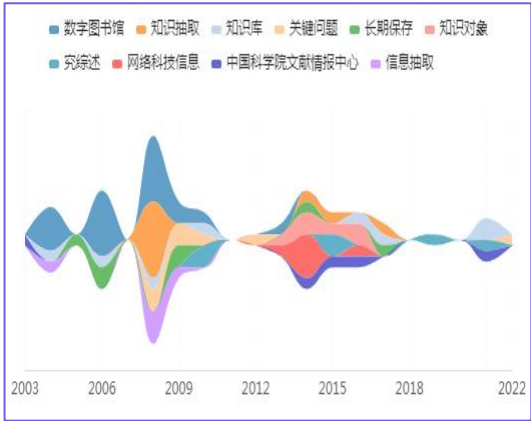
找科研合作



做学科建设



定位领域人才



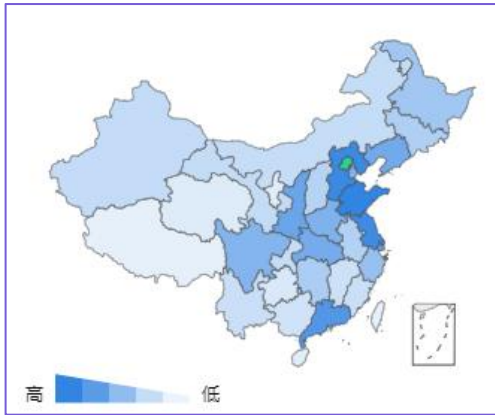
找发文期刊



查看基金分布



地区科研定位



# 打通“事前-事中-事后”全链条覆盖的科研诚信建设与学术不端治理体系

**Type 1 文献相似性比对资源库**

**学术文献** 期刊、学位、会议等学术论文

**专业文献** 标准文献、专利信息、法律法规等专业特色资源

**学术网文** 互联网学术网文、重要行业报纸等资源

**比对资源库** 按照资源类别等维度建设14个比对资源库

**Type 2 科研诚信风险大数据资源**

国内外微博信息、发表后同行质疑信息、学术预警期刊信息、图像及文本筛查数据、学术不端公开惩戒信息等

**Type 3 科技文献图像数据资源库**

**学术论文** 期刊论文、学位论文、会议论文等中英文、医学为主

**图像提取** 自动化图像识别、提取、切图，并元数据管理

**向量化** 对图像进行向量化处理，建立图像矢量化库

**向量化建库**

**提图、切图**

Figure 1 illustrates the Researcher's Academic Integrity Research Process. The process is a linear flow from 'Knowledge Acquisition' to 'Service System'. It includes 'Trust Classroom' and 'Monitoring and Early Warning' as intermediate steps. A red banner at the top reads 'Academic Integrity Research and Academic Integrity Research'. A red banner at the bottom reads 'Academic Integrity Research and Academic Integrity Research'.

[illegible]

## 教育-预警-查验-检测全流程覆盖

The diagram illustrates a comprehensive inspection service for scientific paper integrity risks. It features a central computer monitor displaying a document with a red lightning bolt icon, symbolizing a detected risk. Surrounding the monitor are several callout boxes and icons representing different types of risks and services:

- Text Similarity (文本相似性):** Indicated by a green box and a gear icon.
- Image Reproduction (图像复用):** Indicated by a blue box and a camera icon.
- Funding Source (基金资助):** Indicated by a purple box and a document icon.
- AIGC Risk (AIGC风险):** Indicated by an orange box and a shield icon.
- Citation (引文):** Indicated by a green box and a document icon.
- Author (作者):** Indicated by a blue box and a person icon.
- Publication Date (发表期刊):** Indicated by an orange box and a document icon.

Below the monitor, there are two buttons: **New Paper Submission (新论文上架)** and **Published Paper Inspection (已发表论文验照)**.


**万方检测**  
for laboratory evaluation

**文献相似性检测服务**  
查重 · 查引 · 查不刊

[国际科研诚信服务平台](#)

[数字图书馆事业部](#)
[登录 / 注册](#)

# 课程作业管理系统 正式上线！

学科专业管理   作业全过程管理   多维度统计分析

[点击查看 >>>](#)



**：** 万方检测个人定制版已上线AGCC检测功能，欢迎使用 [查看详情](#)

个人用户专属



个人用户版  
PERSONAL USER SERVICE

机构用户专属



学术预审稿  
ACADEMIC PREVIEW SERVICE

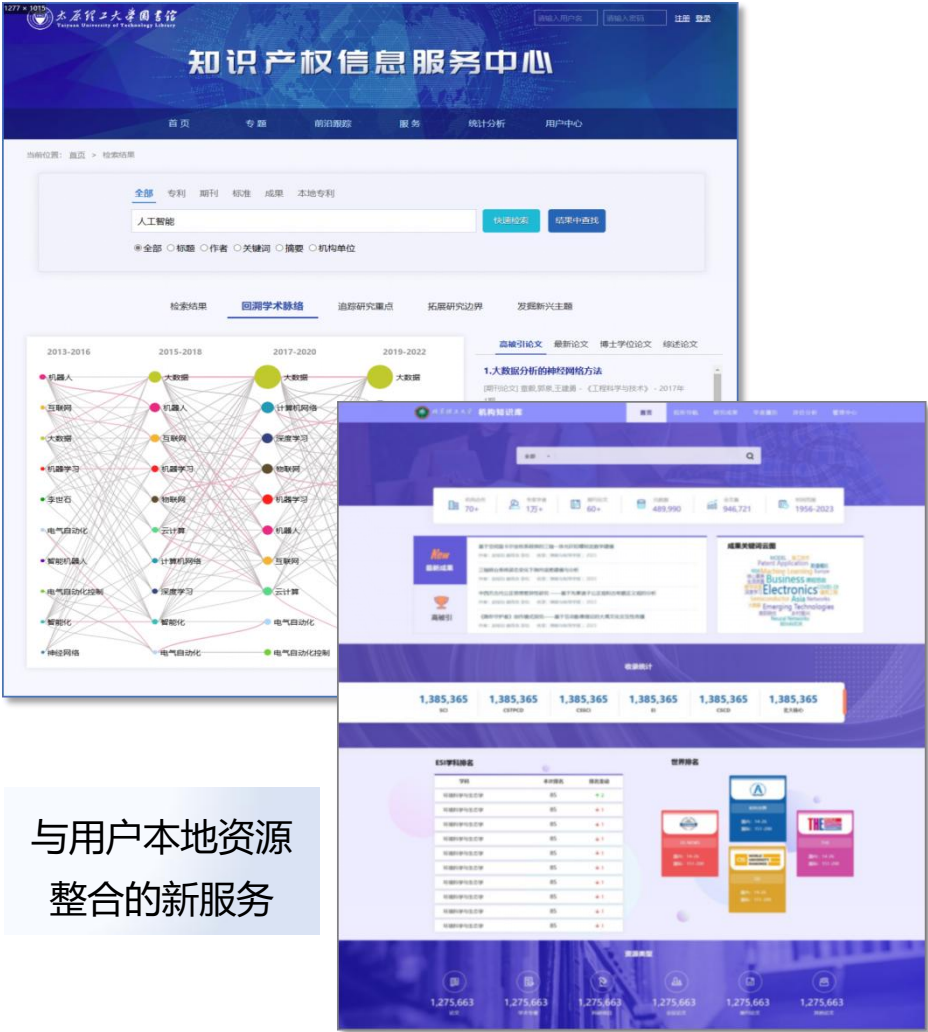
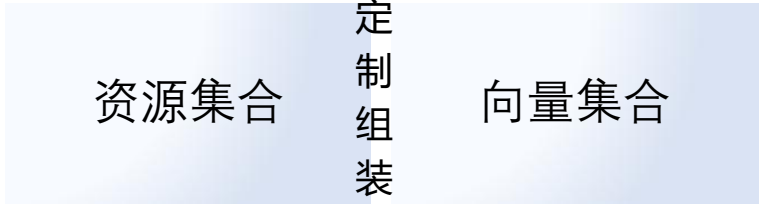
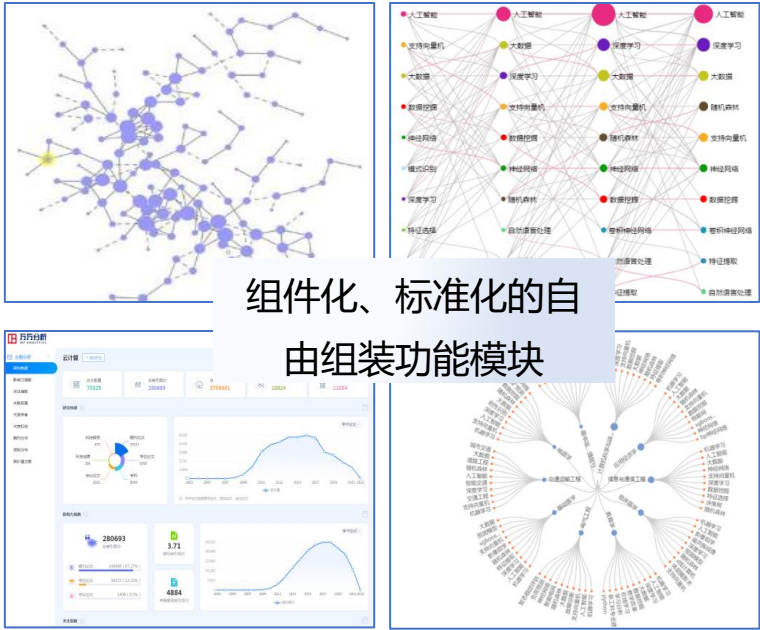
合作用户专属



期刊论文合作版  
JOURNAL PAPER COOPERATION SERVICE

# 贯彻通用与本地融合的一体化协同建设

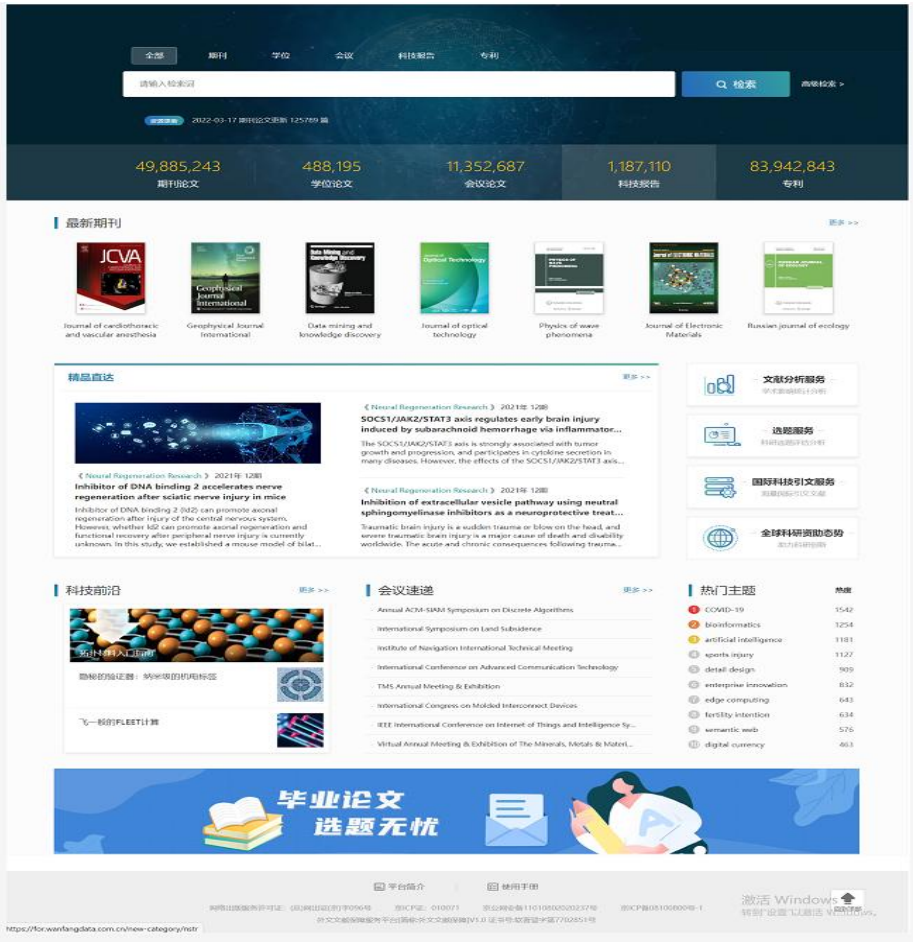
构建产品矩阵时，充分考虑产品的模块化、标准化及可自由组装性，一方面支持敏捷组配，另一方面支撑共建应用生态



与用户本地资源整合的新服务

更多.....

□ 应对汇率上涨+经费缩减压力，竭力利用国际合作渠道数据+OA资源+馆际互借渠道资源，构建AI支撑的西文文献保障服务，并在探索基于AI的垂直学科领域科技资源服务体系自动构建。



| 资源订购优化策略咨询服务--方案一 |              |                     |
|-------------------|--------------|---------------------|
| 分析维度              |              | 数据来源                |
| 订购目录              | 实际订购刊列表      | 机构提供（需规范提供信息）       |
| 订购成本              | 出版商公开的价目表    | 需在构建母体刊表过程中，补充单刊订购价 |
| 订购效益              | 实际订购价        | 机构提供                |
|                   | 订购资源的下载量     | 机构提供，counter数据      |
|                   | 订购资源中的机构发文   | 自主计算                |
| 文献保障成本            | 订购资源中的机构引用   |                     |
|                   | 订购资源中的机构被引   | 自主计算                |
|                   | OA直达满足率      | 自主计算                |
|                   | 可使用的历史订阅备份比例 | 机构提供                |
|                   | 原文传递成本       | 自主计算                |

| 资源订购优化策略咨询服务--方案二 |              |                             |
|-------------------|--------------|-----------------------------|
| 分析维度              |              | 数据来源                        |
| 订购目录              | 拟订购刊列表       | 机构高发文、高引用、高被引刊，补充机构重点学科领域顶刊 |
| 订购成本              | 出版商公开的价目表    | 需在构建母体刊表过程中，补充单刊订购价         |
| 文献保障              | OA直达满足率      | 自主计算                        |
|                   | 原文传递成本       |                             |
| 订阅永久备份比例          | 可使用的历史订阅备份比例 | 需根据出版商订阅政策                  |

# 持之以恒的价值理念



**1、AI-Ready:** 构建面向科技创新场景，应用驱动的资源体系，形成适配AI利用的从“原生数据”到“知识数据的”全链路数据建设体系（AI Ready Data）。

**2、AI In Technical:** 攻关文本挖掘、知识计算、认知智能等关键技术，加强**智能技术的研究和应用能力**。

**3、AI For S&I:** 聚焦科技创新典型需求场景，加深对用户发展战略的理解，场景驱动产品升级与重构，实现**人工智能+场景服务的深度拓展能力**。

|                                                                                                                                                                                                                                                          |                                                                                   |                                                                                   |                                                                                   |                                                                                   |                                                                                    |                                                                                     |
|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------|-----------------------------------------------------------------------------------|-----------------------------------------------------------------------------------|-----------------------------------------------------------------------------------|------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------|
| 产业协同                                                                                                                                                                                                                                                     |  | 标准管理服务系统                                                                          | 内部知识构建系统                                                                          |  |  |  |
| 科研管理                                                                                                                                                                                                                                                     |  |  |  |  |  |  |
| 学术出版                                                                                                                                                                                                                                                     |  | 学者知识脉络                                                                            | 关键词知识脉络                                                                           | 学科发展评估平台                                                                          |  |  |
| 研学支撑                                                                                                                                                                                                                                                     |  |  |  |  |  |  |
|    |                                                                                   |                                                                                   |                                                                                   |                                                                                   |                                                                                    |                                                                                     |

# 感谢大家!

万方数据将以更开放的心态，与各位精诚合作，建立更密切、多元的合作机制，共同实现高质量可持续发展。

万方数据 刘建华

liujh@wanfangdata.com.cn