

上海图书馆
SHANGHAI LIBRARY

开放数据与人工智能协同创新的探索研究 ——以上海图书馆开放数据竞赛为例

上海图书馆副馆长 徐强

第七届上海图书馆阅读季（2025-2026）重点项目



国际大赛 历史人文 海量数据

智阅千年 古韵新生

第十届 上海图书馆开放数据竞赛

全球征集优秀应用、论文和AIGC作品

<https://opendata.library.sh.cn/>

指导单位

上海市文化和旅游局
上海市数据局

主办单位

上海图书馆（上海科学技术情报研究所）
国际科学技术信息理事会（ICSTI）
上海市中心图书馆

协办单位

智慧图书馆技术应用联盟
厦门大学图书馆
CADAL项目管理中心
浙江大学信息技术中心
《图书馆杂志》期刊
《信息与管理研究》期刊

合作单位

公共图书馆

南京图书馆
浙江图书馆
安徽省图书馆
深圳图书馆
苏州图书馆
嘉兴市图书馆
南通市图书馆
上海市静安区图书馆
上海市闵行区图书馆
上海市嘉定区图书馆
上海市金山区图书馆
上海市奉贤区图书馆
上海市崇明区图书馆

其他文化机构（按首字母排序）

《全国报刊索引》编辑部
CADAL项目管理中心
东方社区信息苑
复旦大学图书馆
广州搜韵文化发展有限公司
杭州弘意科技有限公司
华东师范大学图书馆
上海宋庆龄研究会
上海韬奋纪念馆
中国科学院上海营养与健康研究所n指数语义技术平台
中国历代人物传记资料库项目（CBDB）
中国人民大学图书馆

支持单位

上海阿法迪智能数字科技股份有限公司



扫码进入竞赛网站



扫码加入参赛群聊



目录

第十届上海图书馆开放数据竞赛

Open Data+AI

01

02

03

04

05

引言

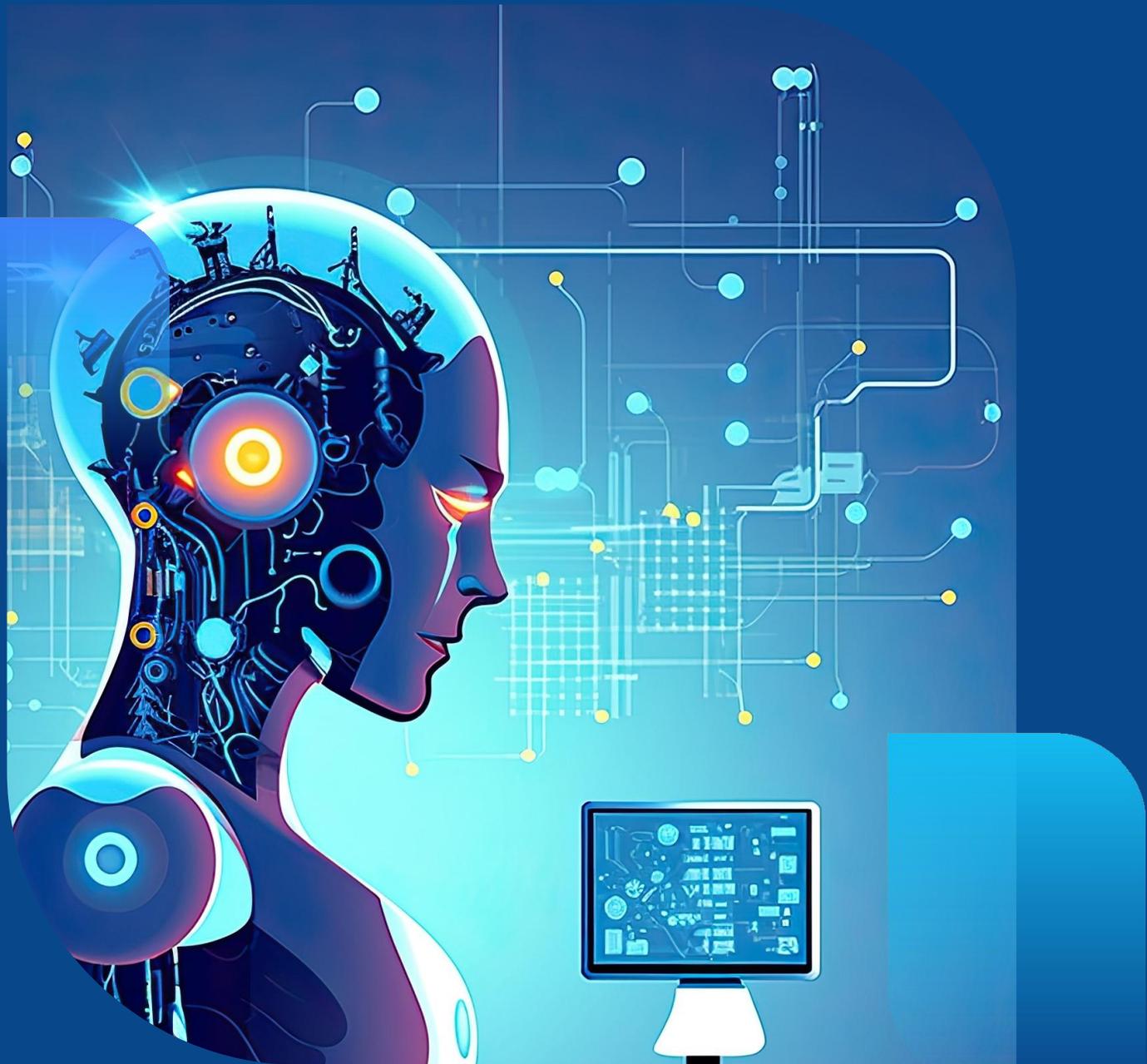
AI与开放数据基础

应用场景

竞赛详情

参与方式

智阅千年 古韵新生



引言

理解竞赛背景与目标



开放数据与AI的重要性



竞赛目标

推动AI与开放数据结合的实际应用

创新应用开发

鼓励参赛者设计并实现基于AI的创新应用，提升开放数据的利用率。

提升社会影响力

通过竞赛作品展示，提高公众对开放数据和AI技术的认知和应用意识。

提供实践平台

为参赛者提供动手实践的机会，通过竞赛积累项目经验与技术能力。



促进跨领域合作

汇聚技术专家、数据科学家等多领域人才，共同探索智能体应用的潜力。

发掘前沿技术应用

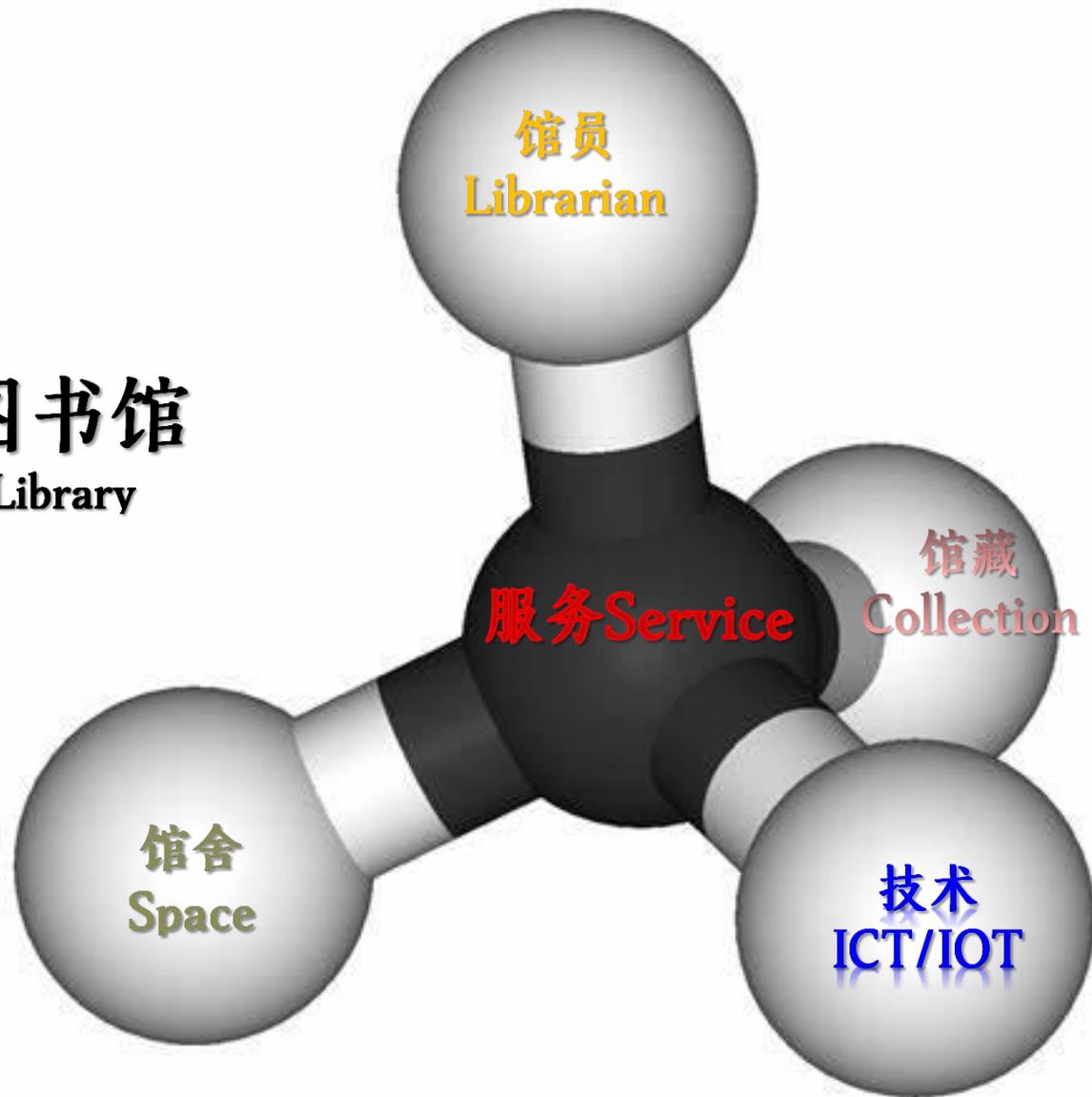
探索AI与开放数据结合的新兴技术，推动智能体应用的前沿发展。

推动可持续发展

借助AI与开放数据的力量，助力社会、经济、环境的可持续发展目标实现。

数字图书馆

Digital Library







Knowledge Base of Stele Rubbings
碑帖知识库

高级搜索

输入关键字

王羲之 化度寺邕禅师舍利塔铭 楷书 化度寺邕禅师 王 欧阳询楷书 魏征撰文

<https://beitie.library.sh.cn/home>



中国家谱知识服务平台
Chinese Genealogy Knowledge Service Platform

首页 家谱中心 先祖名人 登录 注册

Genealogy Digital Collection

家谱检索-踏上寻根问祖的文化之旅

高级检索 | 地图检索

请输入谱名/姓氏/谱籍地/堂号

家谱文化 家规家训 迁徙图 数字展览

<https://jiapu.library.sh.cn/>

<https://ldy.library.sh.cn/index>

770TB



上海新华 华语老电影知识库 Old Chinese Film Knowledge Base

中文 | English 帮助手册

电影名 / 演员 / 导演 / 编剧 / 年份.....

+高级搜索

电影 影人 影戏院

Historical Photo Gallery



上海文化总库 年谱 老照片 老电影 老唱片 红色文献 武康路

历史图片库

“图片上海”发掘和整理上海图书馆馆藏的地方文献中的数字化图片资源，以各种关联实现对图片多角度、多层次的解读。用以图说史的形式，揭示晚清、民国期间上海城市文化、城市空间的多彩多姿，和百年上海的各种永恒瞬间，以及“历史影像”的强大生命力和学术价值。

请输入关键词搜索

笔墨长城：宣传画里的中国抗战 抗战老照片 上海与世博 话剧图片 辛亥革命在上海

<https://scc.library.sh.cn/#/photo/home>



上海图书馆 馆藏唱片知识库

首页 唱片 曲目 乐人 机构团体

搜索唱片、曲目、乐人、机构团体等关键词 高级检索

路德维希·凡·贝多芬

路德维希·凡·贝多芬 (Ludwig van Beethoven, 1770年12月16日或17日—1827年3月26日)，出生于神圣罗马帝国-科隆选侯国的波恩，维也纳古典乐派代表人物之一，欧洲古典主义时期作曲家。贝多芬一生创作题材广泛，重要作品包括9部交响曲、1部歌剧、32首钢琴奏鸣曲、5首钢琴协奏曲、多首管弦乐序曲及小提琴、大提琴奏鸣曲等。因其对古典...

- 唱片 22种
- 机构团体 6个
- 乐人 36人

路德维希·凡·贝多芬

Record Collection Knowledge Base

<https://cp.library.sh.cn/>

数字人文 (Digital Humanities)



 武康路
  街道全景
  分享空间
  知识库

 登录
  注册

建筑照片



武康路113号
巴金故居



武康路391弄1-5号
周璇旧居



武康路393号
黄兴故居



武康路40弄1号
唐绍仪旧居

<https://wkl.library.sh.cn/>

简介
人物
事件
资料

1916年建造，为砖木结构的假层四西班牙式小楼。1943年初夏，中国最早的歌影两栖明星、流行乐坛的先驱者——“金嗓子”周璇和养母一起搬进武康路391弄1号，居住在三楼的一套房子。此前，由于时局关系，周璇已歇影一年多。与此同时，周璇加入张善琨主持的中华电影联合股份有限公司，复出拍片。主演了《渔家女》、《鸾凤和鸣》、《红楼梦》、《凤凰于飞》四部影片，并为百代唱片公司录制了这些影片的全部插曲。1946年春，周璇应香港大中华影业公司之邀赴香港拍片，离开了居住三年多的武康路寓所。

语音简介 

街道全景



Map showing historical sites along Wukang Road:

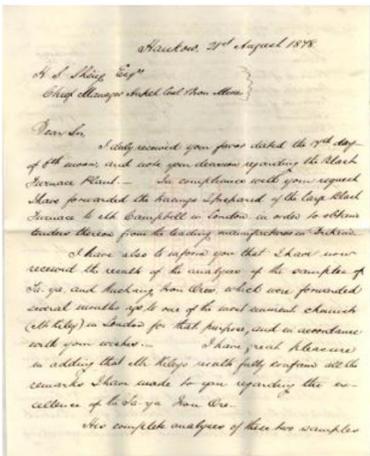
- 孔令侃故居
- 颜福庆旧居
- 王元化故居
- 正广和老屋
- 丁香花园
- 顾祝同旧居
- 湖南别墅
- 巴金故居
- 密丹公寓
- 开善堂公寓
- 周作人旧居
- 贝祖诒、蒋士云夫妇旧居 (E)

Wukang Road Knowledge Base

开放数据平台 (Open Data)

首页 / 项目 / 【上海图书馆】盛宣怀档案之郭师敦盛宣怀往来信函

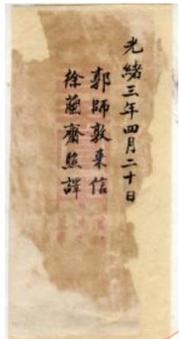
全部 未认领 抄录中 待审核 已完成 被退回



【上海图书馆】盛宣怀档案之郭师敦盛宣怀往来信函

截止时间: 2022-12-31

14个任务 58张图片



郭师敦复盛宣怀函

作者: 郭师敦

关键词: 矿务; 兴国州; 勘探; 工程

页数: 6页

抄录者: test03



郭师敦复盛宣怀函

作者: 郭师敦

关键词: 矿务; 英国; 唐廷枢; 机器; 勘探; 天津府...

页数: 10页

抄录者: zmrsxg

审核者:



盛宣怀复郭师敦函

作者: 盛宣怀

关键词: 运输; 李鸿章; 轮船; 湖北; 荆州府; 天津...

页数: 2页

抄录者: heinz

审核者: 91149554

Crowdsource Historical Archives

<https://zb.library.sh.cn/>



首页 基础词汇库 文献词汇库 本体词表 开发者指南 工具 论文/课件 中文

本网站作为上海图书馆数字人文项目的开放数据平台, 将陆续以关联数据 (Linked Data) 的方式向互联网公开发布上图数字人文项目所用的基础知识库 (人、地、时、事、物), 文献词汇库 (家谱、手稿档案、古籍等), 本体词表, 和数字人文项目建设过程中所用到的各种数据清洗和转换工具, 以及项目组发表的相关论文、课件等研究资料。并以REST API, Sparql Endpoint, 内容协商 (Content Negotiation) 等方式提供各种数据消费接口供开发人员调用, 以促进数据的开放获取、共享和重用。调用接口前, 需注册API Key, 并阅读开发者指南。

人名规范库

More...

华人姓氏列表

More...

中国历史纪年表

More...

地理名词表

More...

收藏机构名录

More...

上海年华知识图谱
(地名志、建筑、历史文化事件...)

More...

印章知识库

More...

避讳字知识库

More...

古籍刻工名录

More...

<https://data.library.sh.cn/index>

馆藏唱片知识库

乐人名家



梅兰芳

京剧表演艺术大师,戏剧活动...



马连良

北京人。回族。字温如,号...



黎锦光

黎锦光 (1907 - 1993年)...



吕文成

广东音乐大师。生于广东。...



言菊朋

北京人。蒙古族。姓马拉特氏...



金少山

原名仲义。北京人。满族。...



程砚秋

程砚秋 (1904年1月1日-195...



谭富英

谱名豫升,小名升楼。原籍...



白驹荣

广东顺德人。原名陈荣。早...



唐诃

作曲家。原名张化愚。河北...



陈歌辛

陈歌辛,原名陈昌寿,出生...



周璇

江苏常熟人。乳名小红,又...



江河水
1979



Symphony No. 101...
Haydn, Joseph, composer.



芭蕾舞剧《天鹅湖》...
(英)国家爱乐交响乐队演奏
1981



Sonata No. 5 in F...
Beethoven, Ludwig van, compo...



Symfonia D-dur nr ...
Haydn, Joseph, composer.



孙家馨独唱歌曲
孙家馨独唱
1982



Symphony no. 41, i...
Mozart, Wolfgang Amadeus, co...



黄金台
吉胜奎,刘永春,小子红[合唱]
1906



Symphony No. 6 in...
Beethoven, Ludwig van, compo...



我的家乡沂蒙山
王音璇演唱
1982?



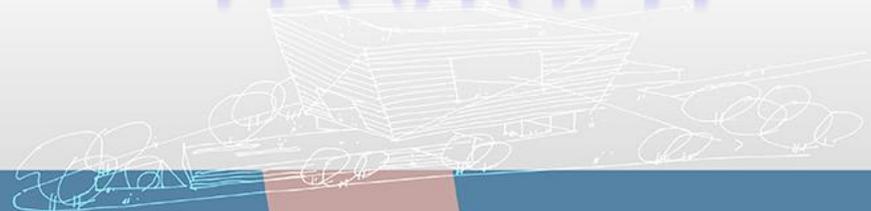
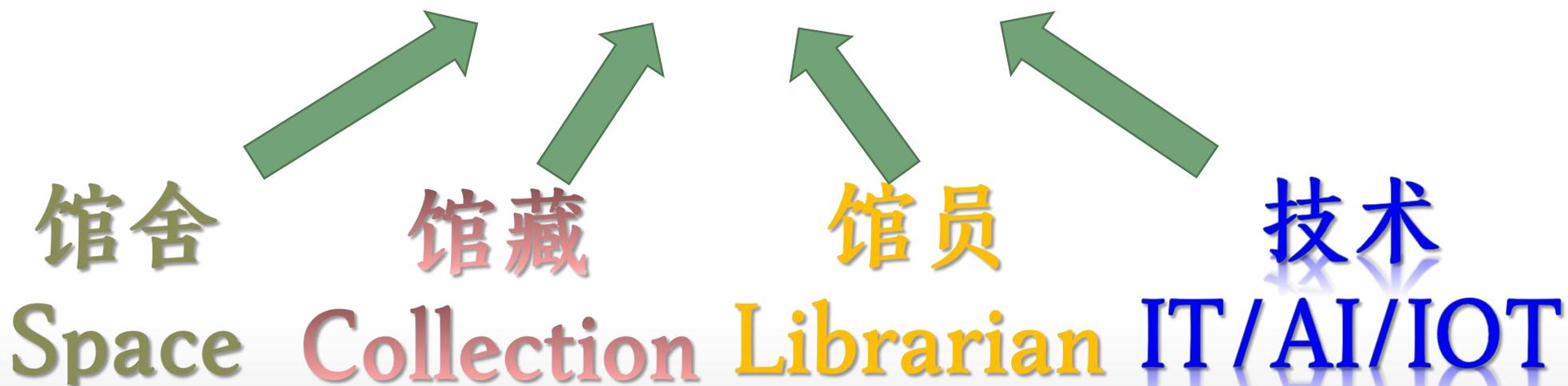
Franz Joseph Haydn

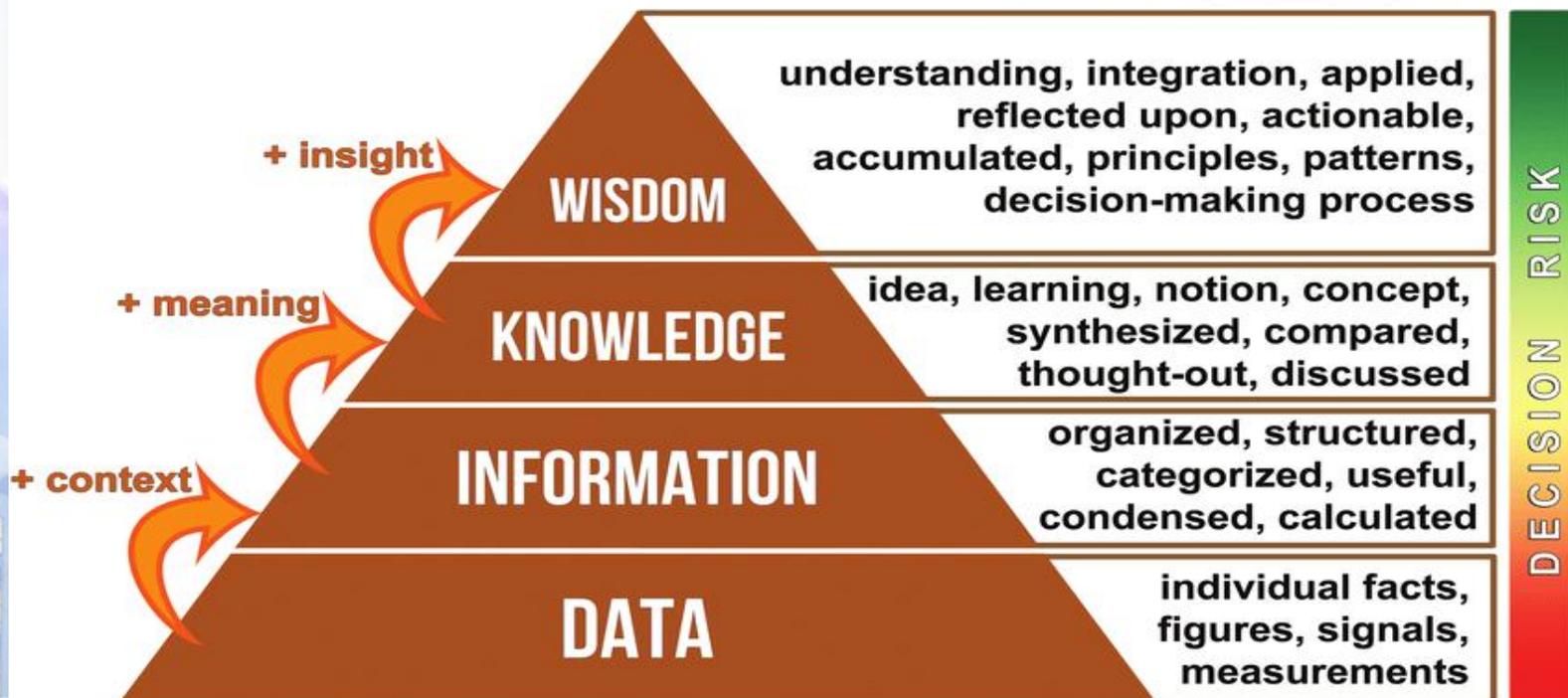
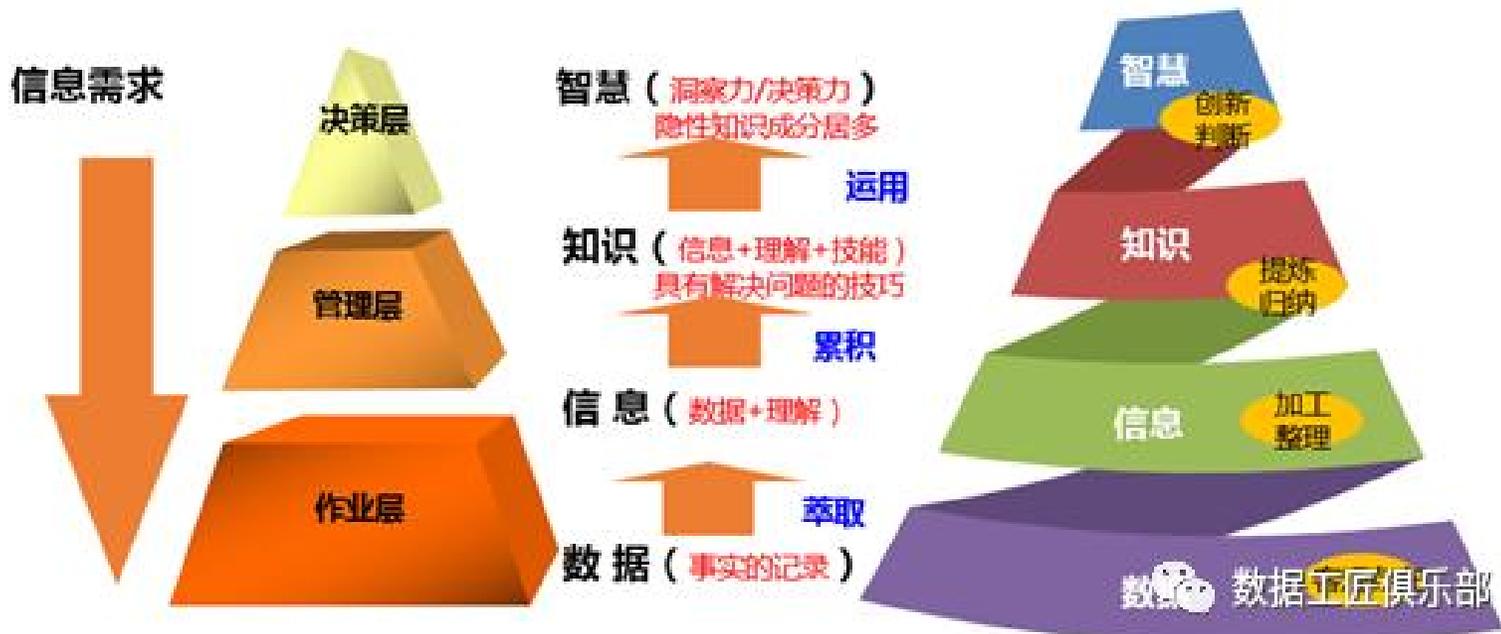
- 唱片 18张
- 乐人 28人
- 机构团体 1个

<https://cp.library.sh.cn/>

智慧图书馆 Smart Library

服务 Service





人工智能 Artificial intelligence

GenAI生成式人工智能：

GenAI是指能够创造新内容的人工智能技术。它通过深度学习和神经网络等技术，从大量数据中学习，理解特定的风格或模式，并据此生成新的内容。

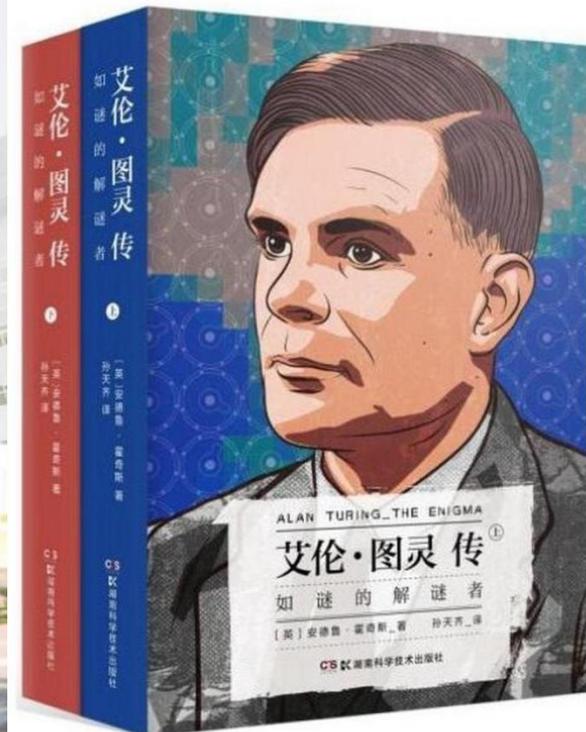


AIGC人工智能生成内容：

AIGC是指利用AI技术生成的各种形式的内容，如文本、图像、音频或视频等。它代表了继专业生产内容（PGC）和用户生产内容（UGC）之后的新型内容创作方式。

AGI通用人工智能：

AGI追求的是让机器在广泛的认知任务上表现得与人类一样好，甚至更好。它的目标是打造一个能够像人类一样，甚至在更多领域超越人类的智能系统，具有广泛的适应性和学习能力。



人工智能 Artificial intelligence

LLMs大语言模型 (Large Language Models) :

大语言模型 (Large Language Model , 简称LLM) 是指在大规模文本语料上训练、包含百亿级别 (或更多) 参数的语言模型。这些模型采用与小模型类似的Transformer架构和预训练目标, 如语言建模 (Language Modeling) , 但主要区别在于模型的大小、训练数据量和计算资源的大幅增加。

LVMs大视觉模型 (Large Vision Models) :

是指在大规模图像或视频数据集上训练的深度学习模型, 这些模型通常包含数亿甚至更多参数。它们能够处理和理解视觉信息, 并在各种视觉任务上展现出卓越的性能。大视觉模型应用在自动驾驶、医疗影像分析、安防监控、内容创作等。2D

SIMs大空间模型 (Spatial Intelligence Models)

大空间模型专注于模拟和理解三维空间, 包括空间关系、导航、物体交互等, 使AI系统能够在现实世界的三维空间中进行理解和操作。通常用于机器人、自动驾驶汽车、增强现实 (AR) 、虚拟现实 (VR) 、地理信息系统 (GIS) 、三维建模等领域。大空间模型专注于处理与三维空间中对象的位置、形状、运动和相互作用有关的信息。3D

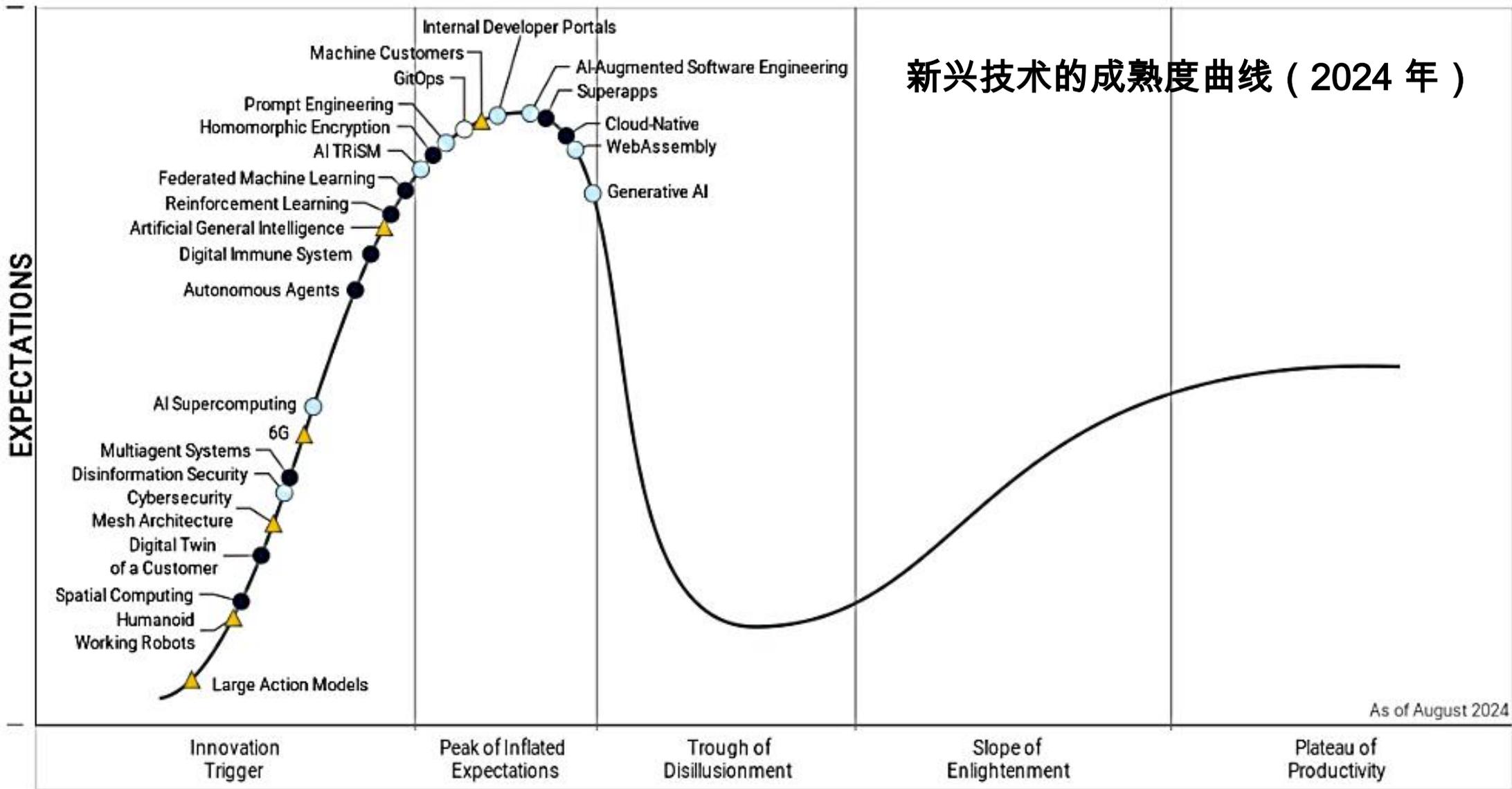




AI与开放数据基础

了解关键技术及其结合

新兴技术的成熟度曲线 (2024年)



As of August 2024

Plateau will be reached: ○ <2 yrs. ● 2-5 yrs. ● 5-10 yrs. ▲ >10 yrs. ⊗ Obsolete before plateau

AI技术概述

理解AI技术的基础

语音识别技术

语音识别技术将人类的语音信号转换为文本，提高人机交互的自然性和便捷性。

强化学习

强化学习通过智能体与环境的交互，逐步优化决策过程，实现更高效的学习。

自然语言处理

自然语言处理技术使计算机能够理解、解析并生成人类语言，提升人机交互体验。

机器推理

机器推理技术使机器能够根据已有数据进行逻辑推理，辅助复杂问题的解决。

计算机视觉

计算机视觉技术赋予机器识别和分析图像、视频的能力，广泛应用于自动驾驶等领域。

深度学习

深度学习通过构建多层神经网络，实现对复杂数据的高效处理与模式识别。

开放数据简介

开放数据资源与应用

开放数据定义

开放数据指无需授权即可自由获取和利用的数据资源，促进信息共享。

开放数据来源

政府、企业和科研机构是主要的开放数据提供者，确保数据广泛获取。

数据清洗与预处理

清洗和预处理步骤确保数据质量，去除噪音和不相关信息。

开放数据政策

各国政府制定开放数据政策，推动数据资源的透明化和共享。

数据标准化

标准化数据格式和协议有助于不同系统间的数据交互与整合。

数据隐私保护

在开放数据应用中，隐私保护至关重要，需采用匿名化等技术。

数据共享平台

开放数据共享平台促进数据的集中存储与便捷访问，提高利用率。



应用场景

探索AI与开放数据的实际应用

智能助手

AI在日常生活中的应用

语音交互

语音助手通过自然语言处理，实现智能对话，提高用户交互体验。

个性化推荐

利用用户数据，提供个性化的产品或服务推荐，增强用户满意度。

情感分析

分析用户情感倾向，帮助企业优化服务策略，提升客户关系管理。

实时翻译

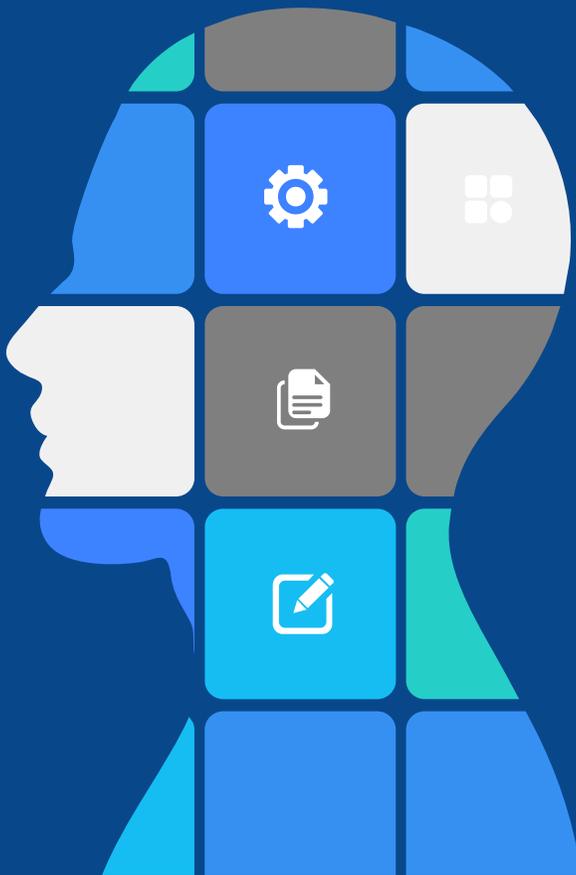
支持多语言实时翻译，打破语言障碍，促进跨文化交流与协作。

01

02

03

04



05

日程管理

智能助手帮助用户管理日常事务，安排会议和提醒重要事项。

06

健康监测

结合健康数据，提供个性化健康建议，实时监测用户健康状况。

07

智能客服

自动处理客户咨询与投诉，提高服务效率，降低人工成本。

08

虚拟助手

提供虚拟形象互动，为用户提供娱乐与信息查询的全新体验。

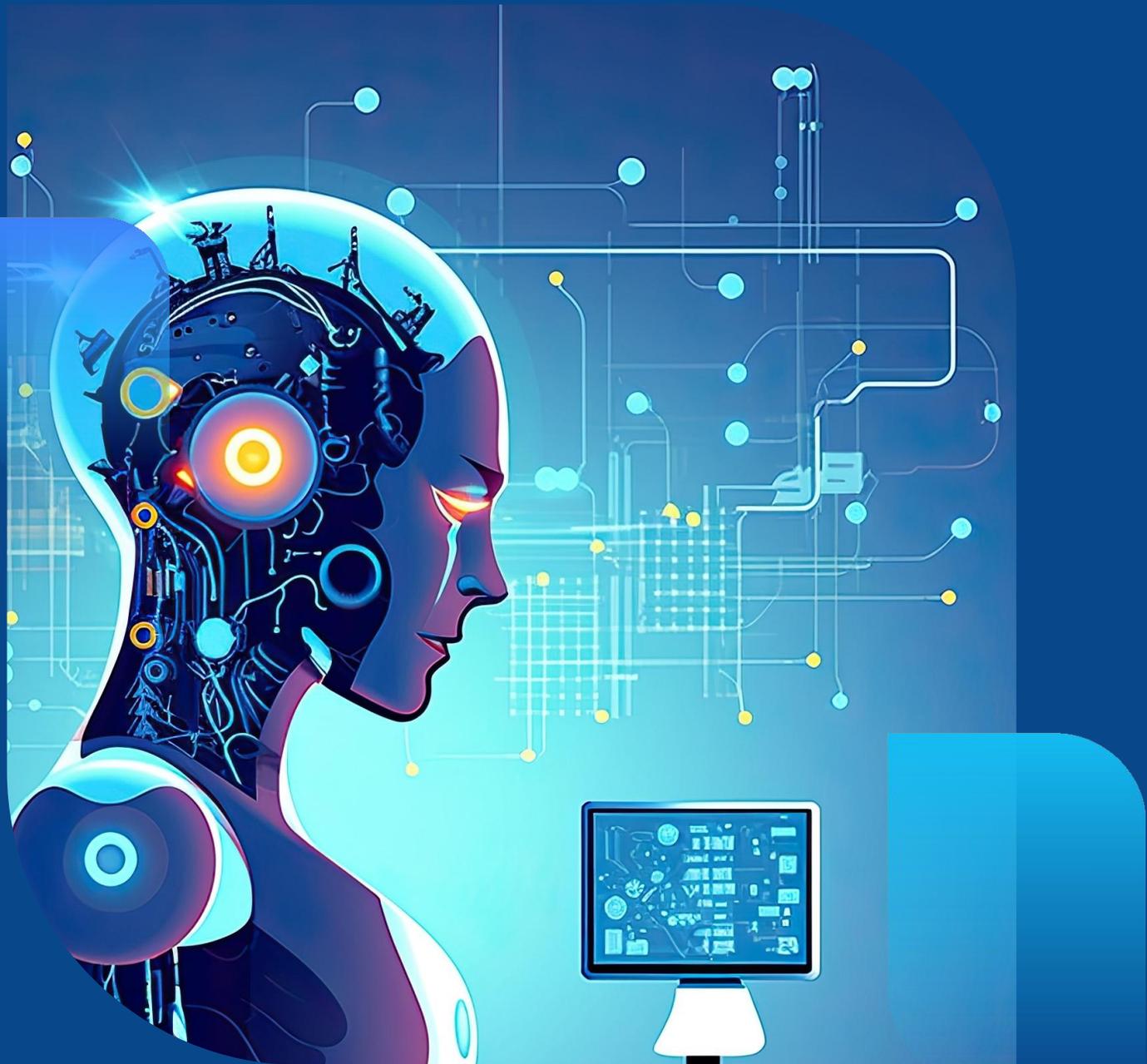
GenAI生成式人工智能

➤ **AI智能体 AI Agent**，即人工智能代理，是指能够执行任务或服务的智能系统。这些代理可以是软件或硬件，甚至两者的结合，它们能够模拟人类智能行为，以自主或半自主的方式执行任务。



AI Agent不同自动化程度所需要的能力简要分析

等级	感知能力	认知能力	执行能力	规划能力
L1 (部分自动化)	“所见即所得”的感知，处理单一模态下的相对简单的数据类型，应用于简单场景	利用大量人类监督信号获得的一定程度的理解语言、利用语言人机交互能力	少量的常见标准工具的调用，简单的工具调用逻辑	静态地执行特定的、预定义的任务。涉及少量的、简单串并联的流程节点
L2 (有条件自动化)	多模态感知能力，能处理更广泛的数据类型，应用于更多样、更长尾、更复杂的场景	全面的认知能力，包含记忆能力、决策能力、高度智能的对话能力、内容生成能力	可使用的工具数量、类型、实现的业务逻辑的复杂度得到极大提升	以业务规模达到端到端最大化自动化为目标，可以规划和编排大量流程节点和复杂逻辑
L3 (高度自动化)	综合利用认知能力，环境交互结果，在少量人类干预下获得超高精度的感知力	通过综合利用环境知识、人类少量的监督信号，达到高精度的认知水平	在人类少量干预下，可以实现绝大多数工具调用代码	能够主动洞察问题域和求解域的环境变化，实现业务流程的灵活适应和编排，环境适应能力强
L4 (完全自动化)	在无人工干预下智能体自主进化获得超高精度的感知能力	利用环境信号自主学习提升认知水平	能自动学习工具使用的方式、实现100%的自动化调用工具的能力	能利用过程反思、经验沉淀，难例挖掘等高度智能化的决策机制，自主提升规划和编排能力，自主进化



竞赛详情

了解竞赛规则与奖励

奖励设置

丰厚奖励与资源支持





参与方式

如何加入竞赛并获取支持

报名指南

报名流程与注意事项

在线报名

访问竞赛官网，填写报名表格并提交项目概述。

团队组建

建议跨领域组队，涵盖技术、设计等多方面人才。

提交材料

按要求提交项目计划书、技术方案等必要文件。

资格审核

竞赛组委会将对报名材料进行审核与确认。

常见问题

查阅官网FAQ，解决报名过程中的常见疑问。

截止日期

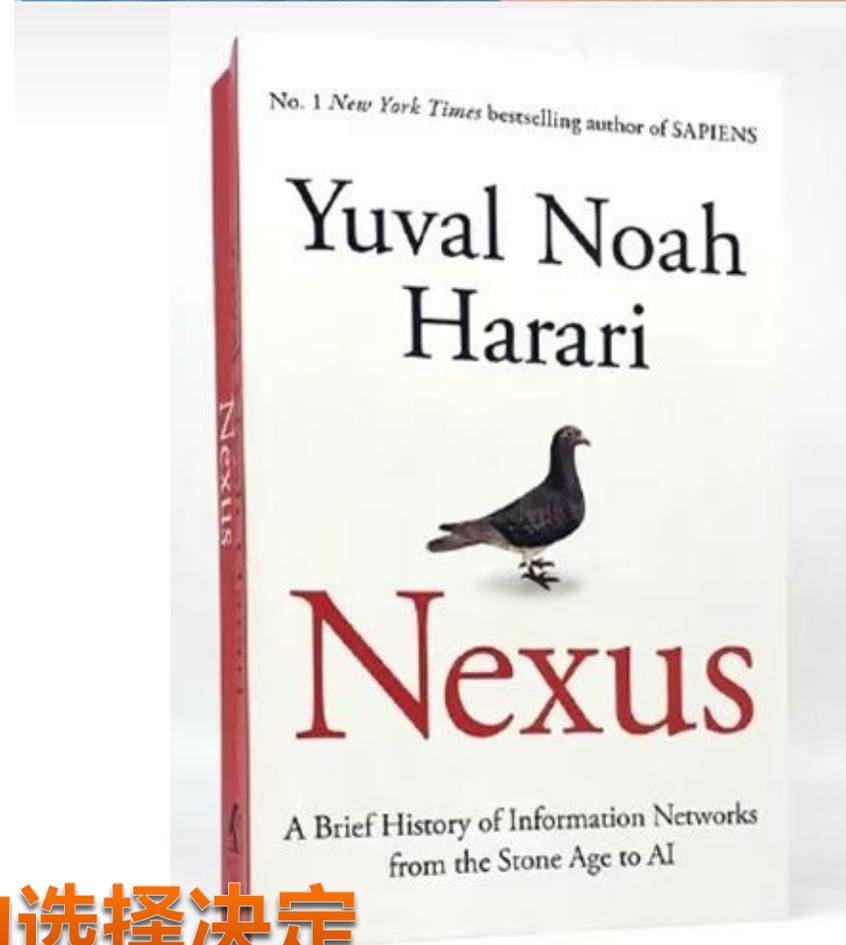
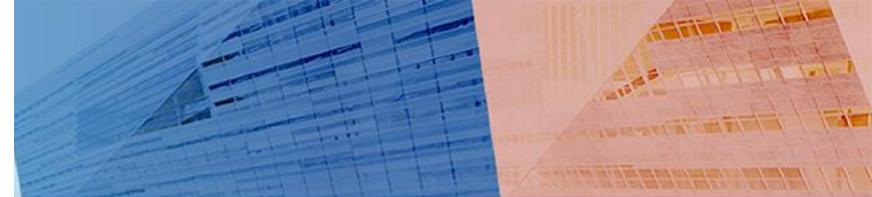
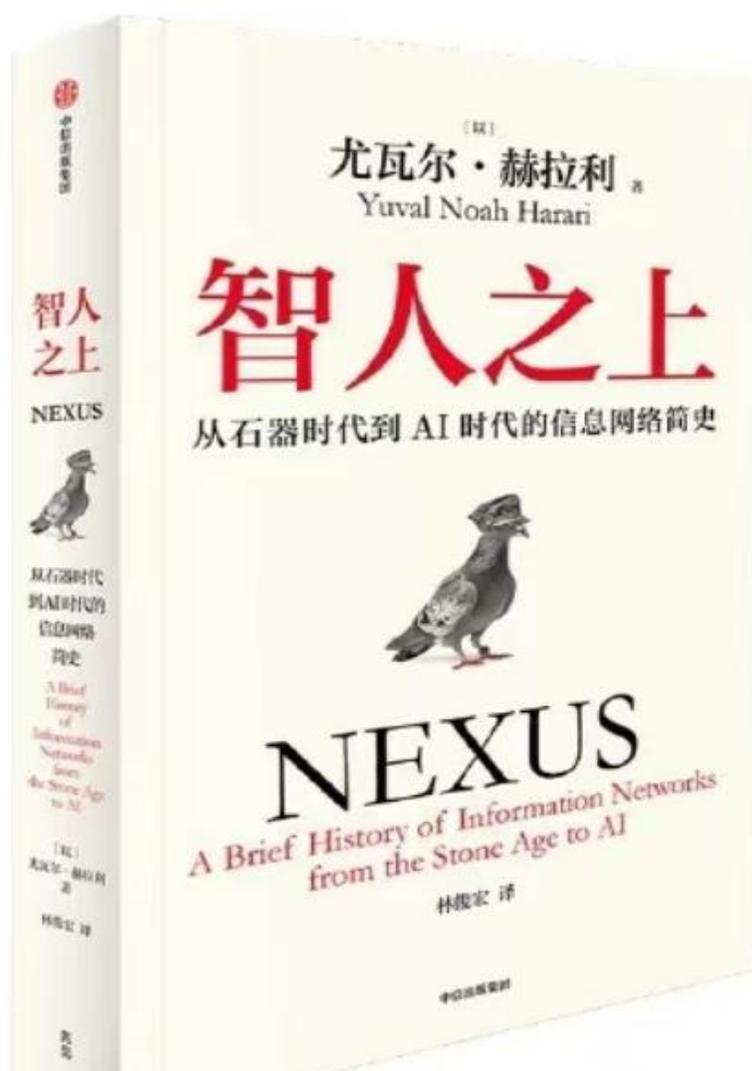
确保在截止日期前完成报名，避免错过参赛机会。

报名费用

竞赛报名免费，无需支付任何费用即可参与。

“我所能想象的天堂应该是
图书馆的模样。”

——豪尔赫·路易斯·博尔赫斯（1899—1986年）



未来的图书馆将由我们今天的选择决定

The future trajectory of libraries will be shaped by the decisions we make in the present.



第七届上海图书馆阅读季（2025-2026）重点项目

智阅千年 古韵新生

第十届 上海图书馆开放数据竞赛

全球征集优秀应用、论文和AIGC作品

<https://opendata.library.sh.cn/>

指导单位

上海市文化和旅游局
上海市数据局

主办单位

上海图书馆（上海科学技术情报研究所）
国际科学技术信息管理学会（ICSTI）
上海市中心图书馆

协办单位

智慧图书馆技术应用联盟
厦门大学图书馆
CADAL项目管理中心
浙江大学信息技术中心
《图书馆论坛》期刊
《信息与管理研究》期刊

合作单位

公共图书馆

南京图书馆
浙江图书馆
安徽图书馆
深圳图书馆
苏州图书馆
嘉兴市图书馆
南通市图书馆
上海市静安区图书馆
上海市闵行区图书馆
上海市嘉定区图书馆
上海市金山区图书馆
上海市南汇区图书馆
上海市崇明区图书馆

其他文化机构（按首字母排序）

《全国报刊索引》编辑部
CADAL项目管理中心
东方社区信息源
复旦大学图书馆
广州捷豹文化发展有限公司
杭州弘惠科技有限公司
华东师范大学图书馆
上海宋庆龄研究会
上海韬奋纪念馆
中国科学院上海营养与健康研究所i指数语义技术平台
中医历代人物传记资料库项目（CBDB）
中国人民大学图书馆

支持单位

上海珂途迪智能数字科技股份有限公司



扫码进入竞赛网站



扫码加入参赛群聊



国际大赛 历史人文 海量数据



谢谢

上海图书馆 徐强
qxu@libnet.sh.cn

