



TaoLi

Inspiring the love for Mandarin Chinese

Applications of AI Agents in International Chinese Language Education

谢俊杰

2025.04.25



- 
- **The Evolution of AI**
 - **What is an AI Agent**
 - **TaoLi AI Agent Platform**



Presentation



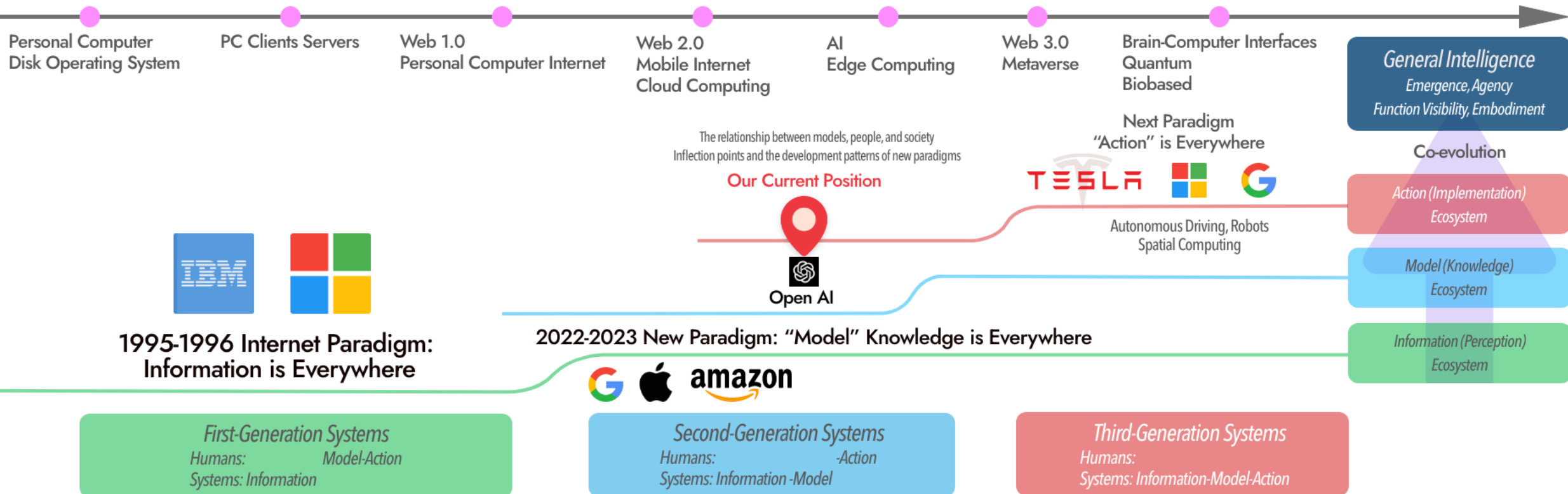
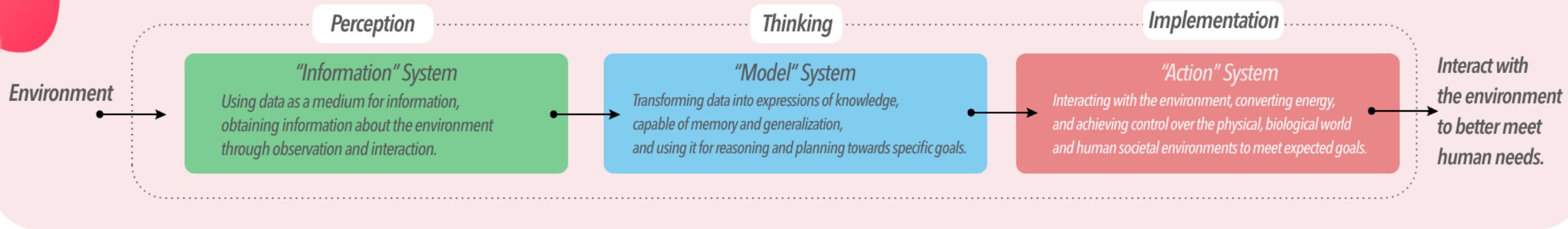


The Evolution of AI



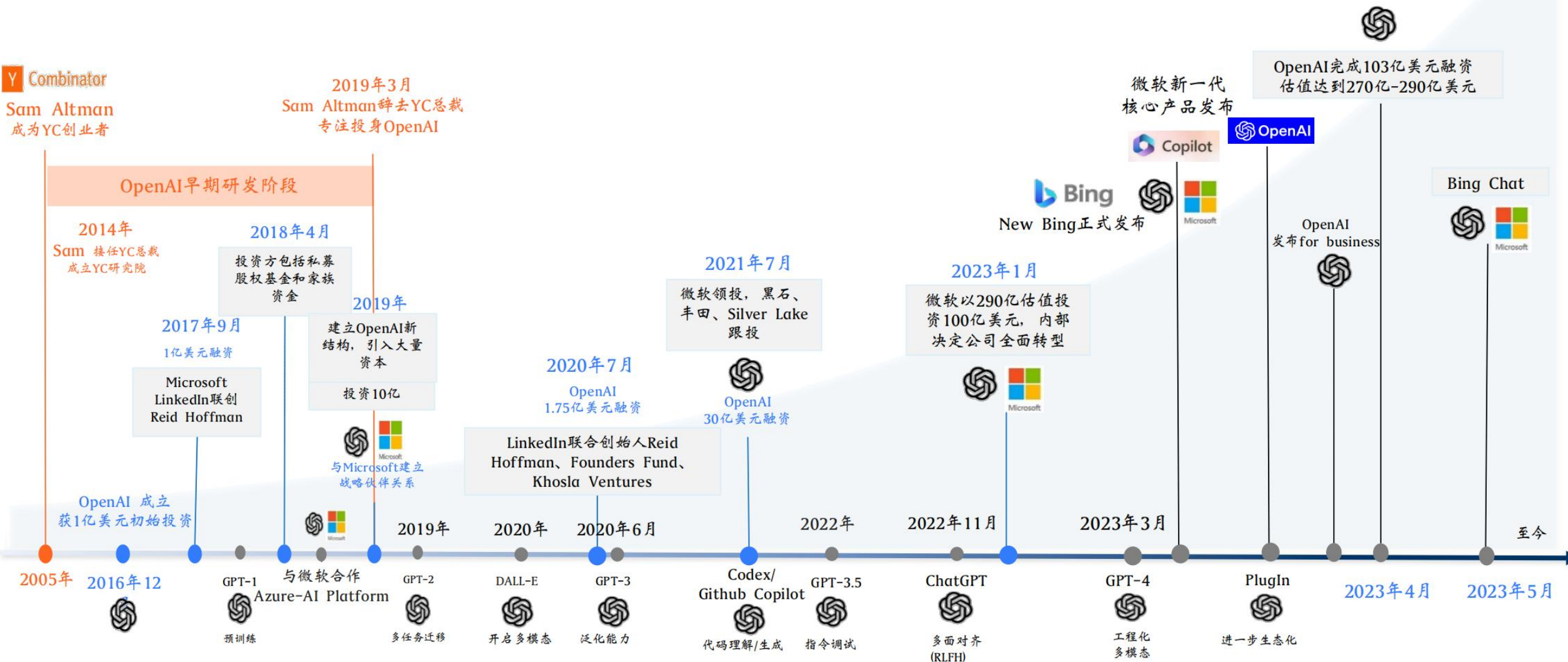
Trinitarian Structure Evolution Model

People, Organization, Society, Digitalization



Y Combinator

Sam Altman
成为YC创业者





AI 1.0时代

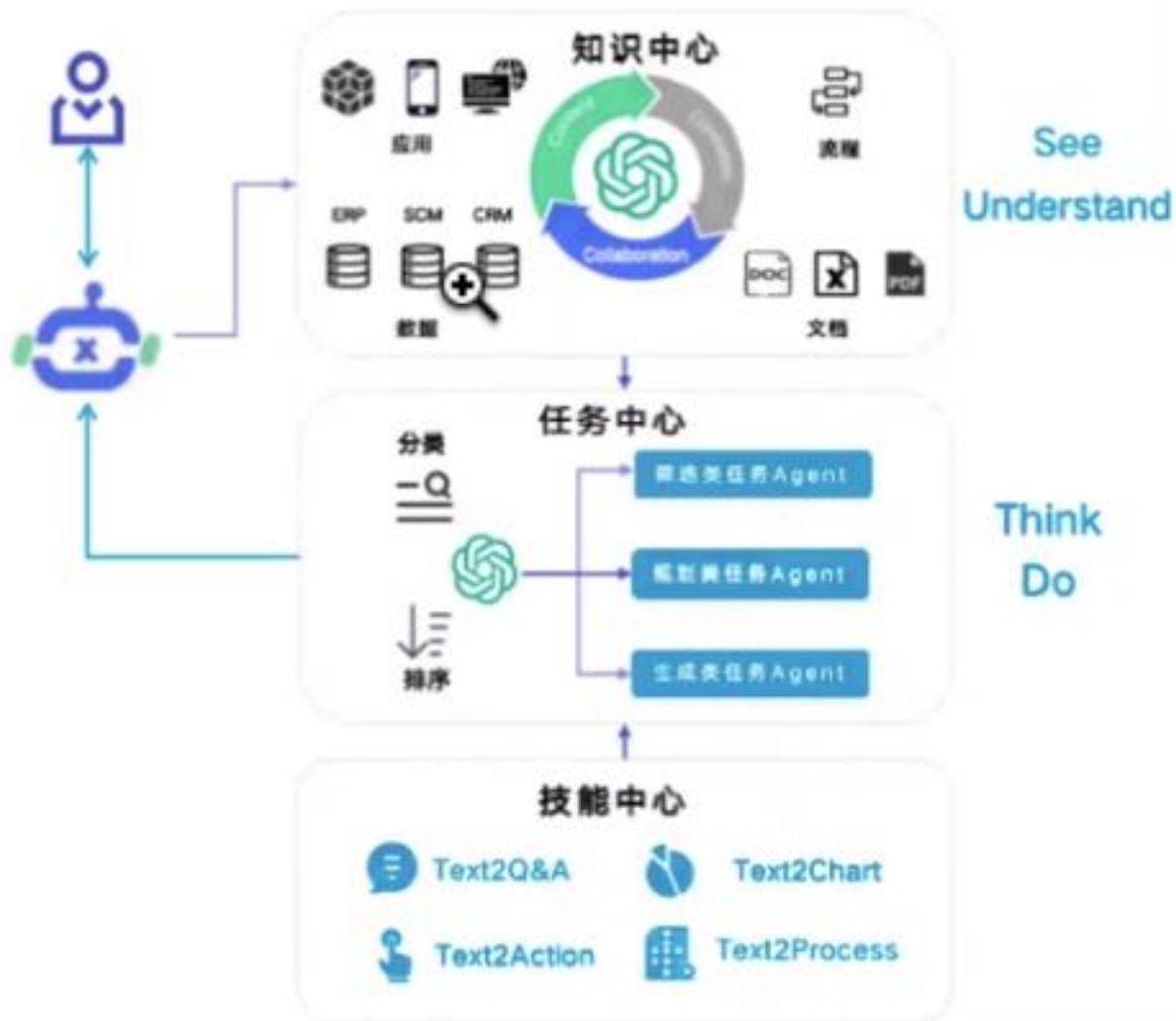
See, Think, Do, Understand

数据文档含义，业务规则，流程需要人来翻译给机器。

AI 2.0时代

See, Think, Do, Understand

机器能读懂数据，文档，应用和流程，人只需要不断地给予更新及训练即可。机器可以自动串联数据，文档和应用。



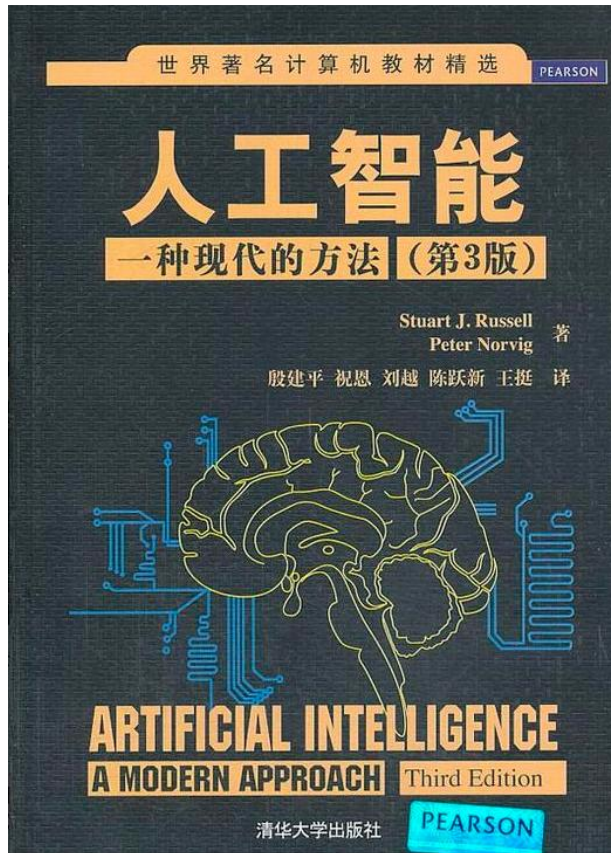


What is an AI Agent



Origins of Agents

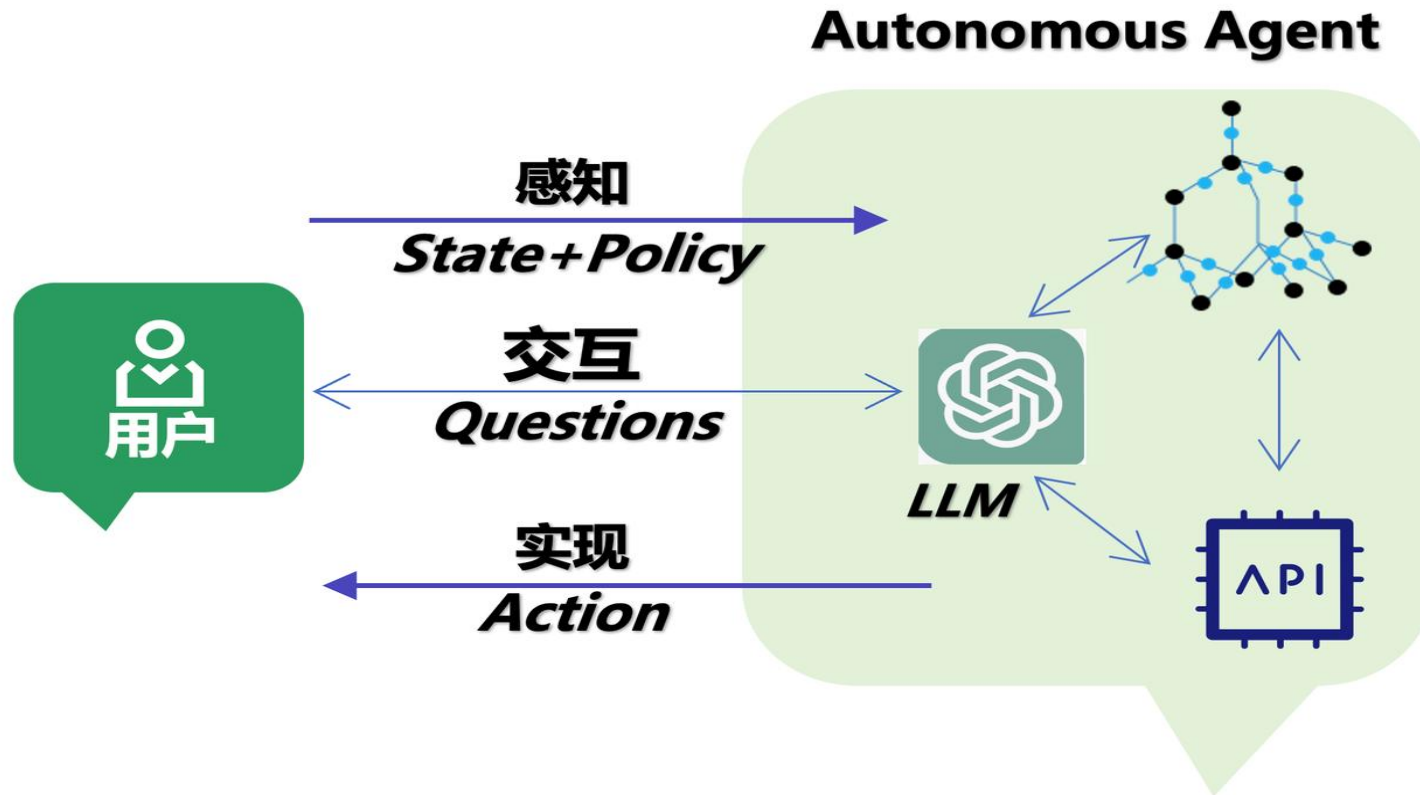
- 1995年, Russell 和 Norvig发表《Artificial Intelligence: A modern approach》(人工智能: 一种现代的方法)。



The book defines AI as the study of agents capable of perceiving environmental information and taking actions, with intelligent agents serving as the unifying theme.

LLM-Powered AI Agents

Artificial intelligence agents are engineered to emulate human-like reasoning and actions, facilitating natural interactions while demonstrating self-directed learning and environmental adaptation

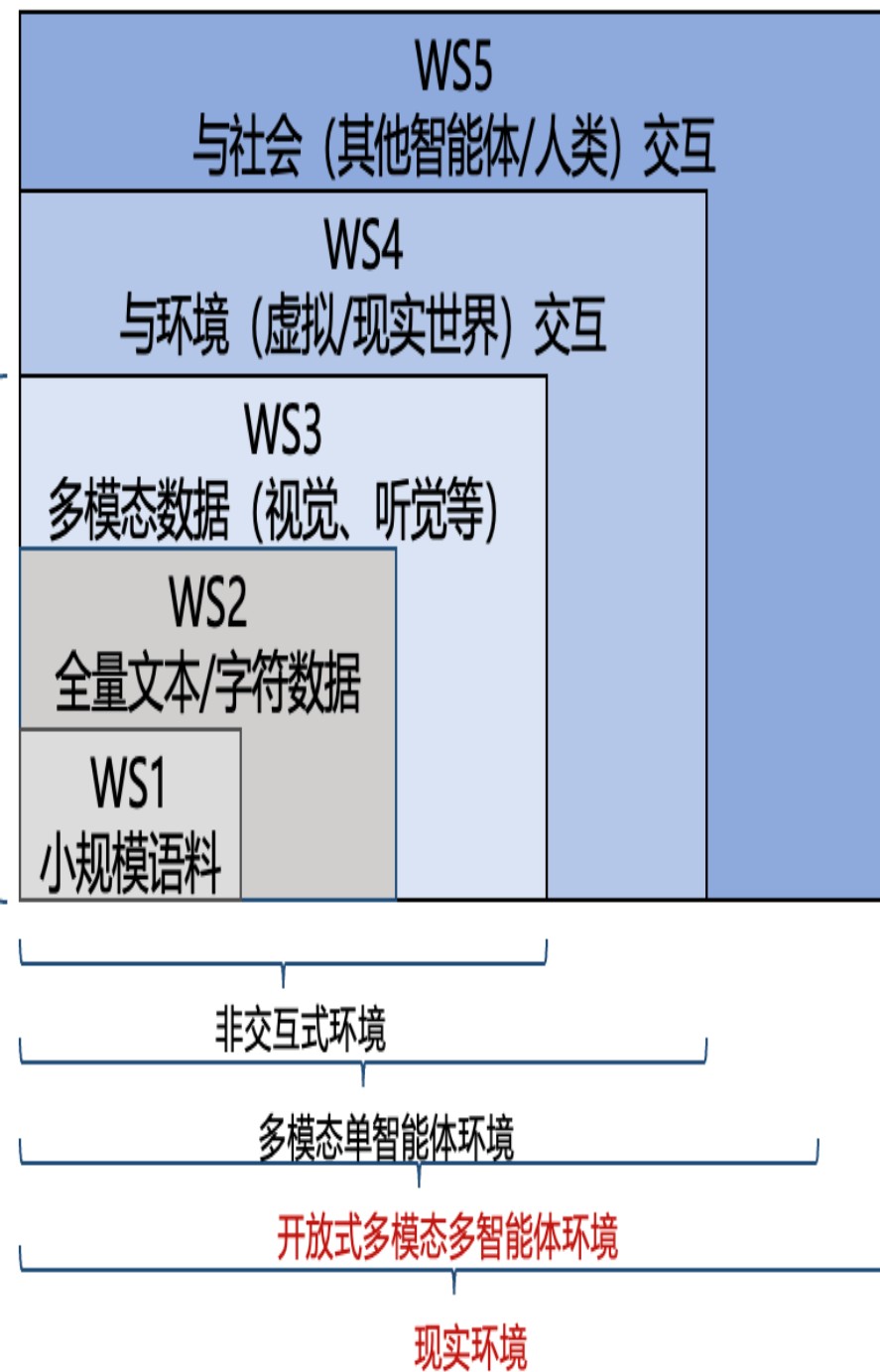


Key Words of AI Agents

- 自治 (Autonomy)
- 知觉 (Perception)
- 反应 (Reactivity)
- 推理和决策 (Reasoning and decision-making)
- 学习 (Learning)
- 通信 (Communication)
- 以目标为导向 (Goal-oriented)

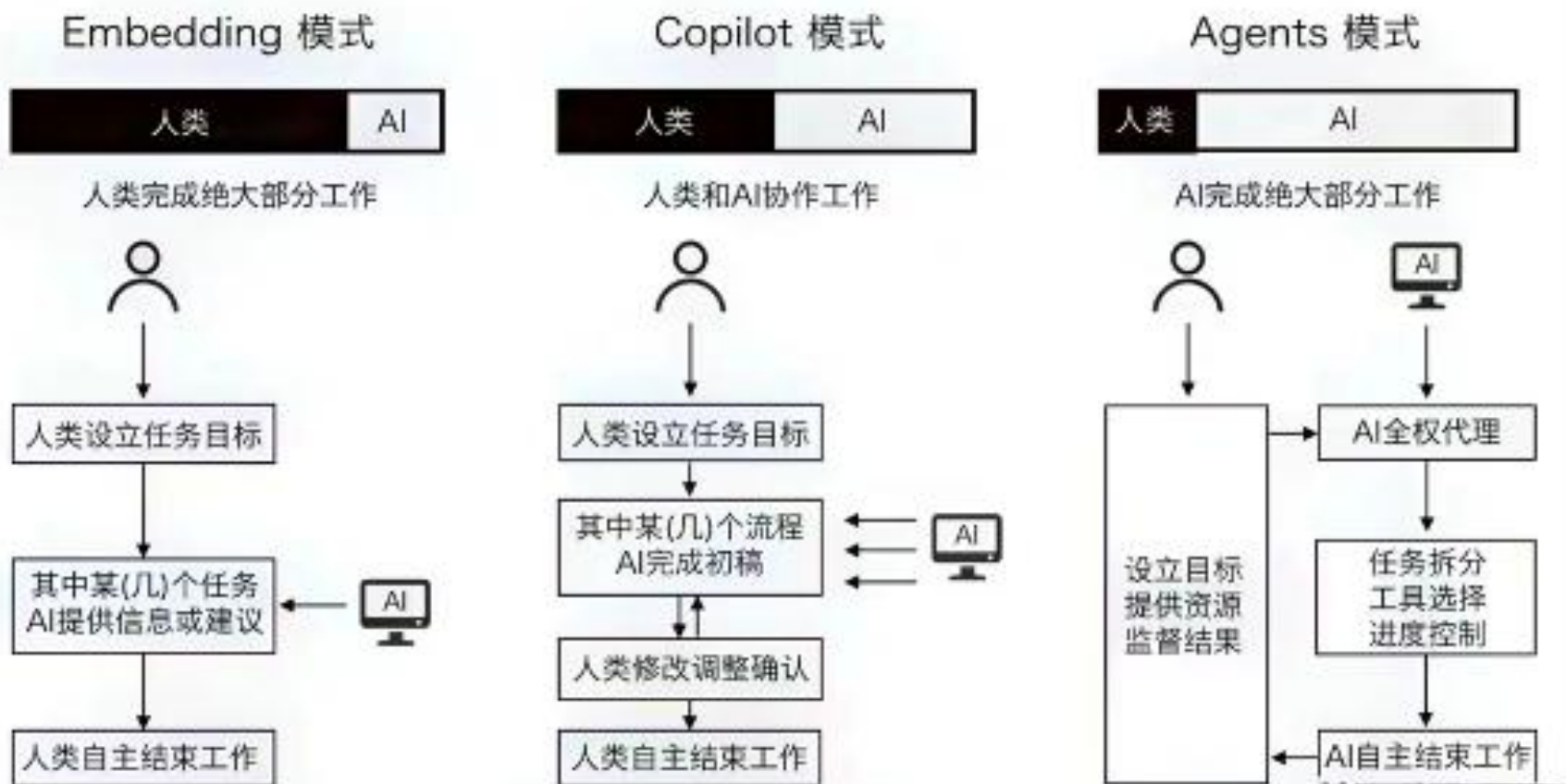
面向决策任务的 (多模态)
预训练大模型

文本/多模态
预训练大模型
(ChatGPT/GPT-4)

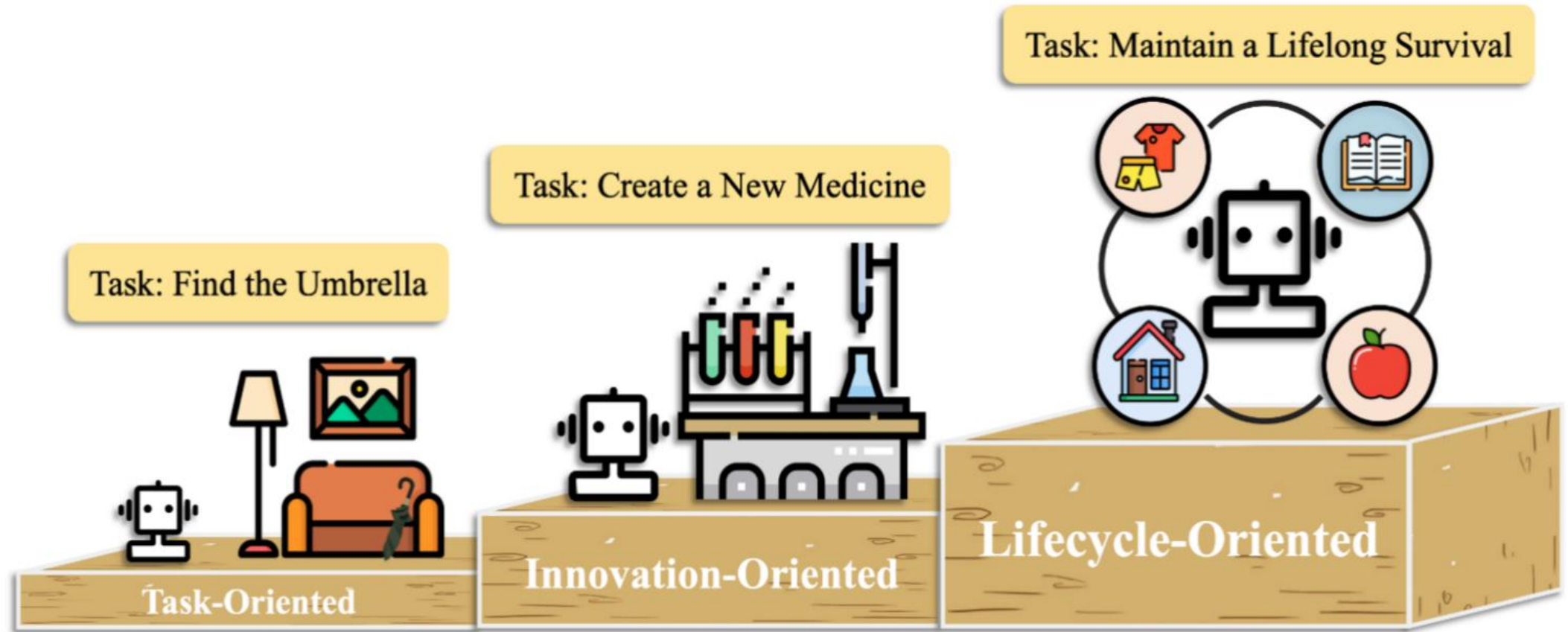


Categorization of AI Agents

人类与AI协同的三种模式



Categorization of AI Agents





TaoLi AI Agent Platform

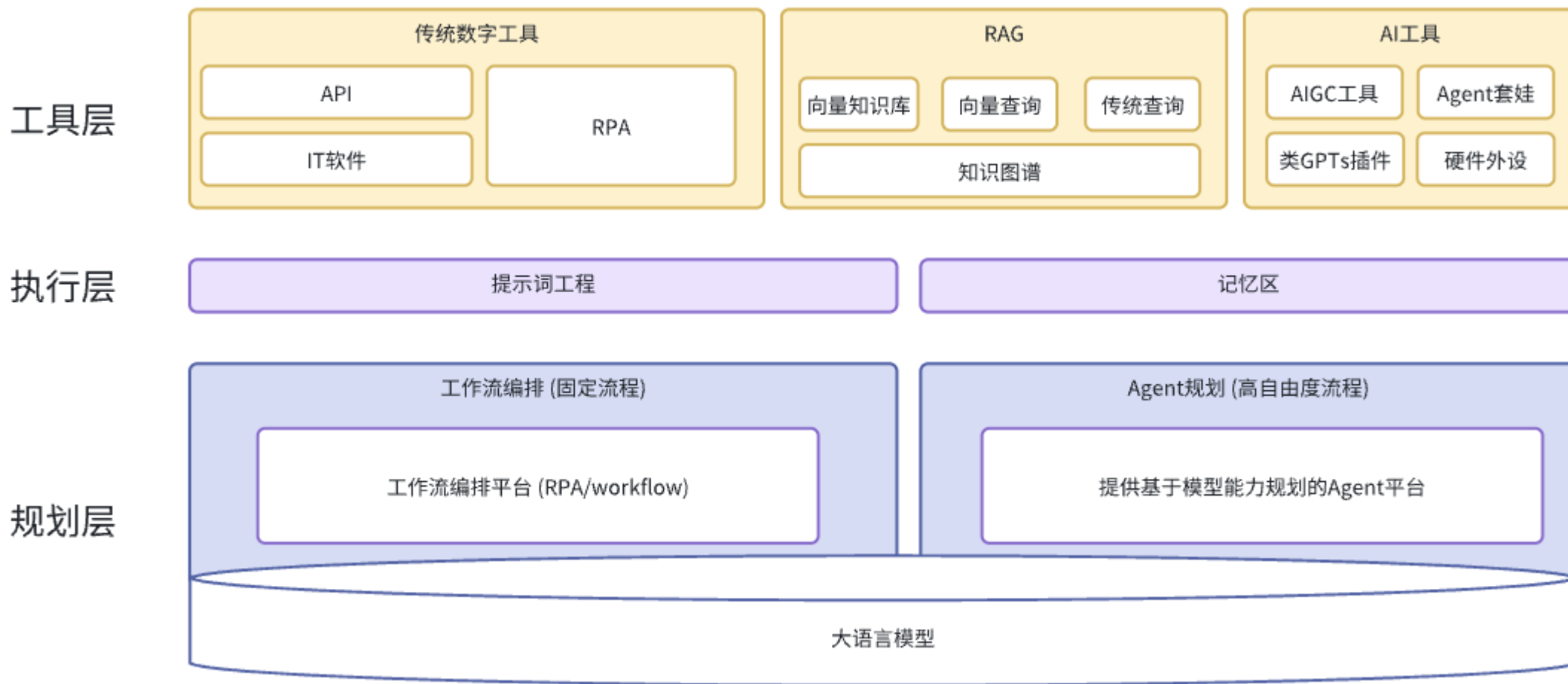


Generative AI Application Level

	描述	示例
L1 Tool	人类完成所有工作，没有任何明显的AI辅助	绝大部分应用
L2 Chatbot	人类直接完成绝大部分工作。人类向AI询问意见，了解信息 AI提供信息和建议但不直接处理工作	初代ChatGPT
L3 Copilot	人类和AI进行写作，工作量相当。AI根据人类要求完成工作 初稿，人类进行目标设定，修改调整，最后确认	Github Copilot Midjourney ChatGPT with Plugin
L4 Agent	AI完成绝大部分工作，人类负责设定目标、提供资源和监督 结果。AI完成任务拆分，工具选择，进度控制，实现目标后 自助结束工作	AutoGPT
L5 Intelligence	完全无需人类监督，AI自主拆解目标，寻找资源，选择并使 用工具，完成全部工作，人类只需给出初始目标	类冯·诺依曼机器人 或者.....人？



TaoLi AI Framework



The Educational Scenario

等级	按应用场景推荐	按学科推荐	其他方式推荐
L1	词汇出题	语文（单词）、英语（单词）	文本
L2	知识点出题	历史、文学、语文（阅读理解）	文本（包含推理）
L3	错题集个性化出题、判卷	理论物理、理论化学、英语（语法）	图片
L4	实时教学助手	数学、化学、物理、政治	数字



General Application Platform

AI赋能

高效率和大规模生成

个性化适应

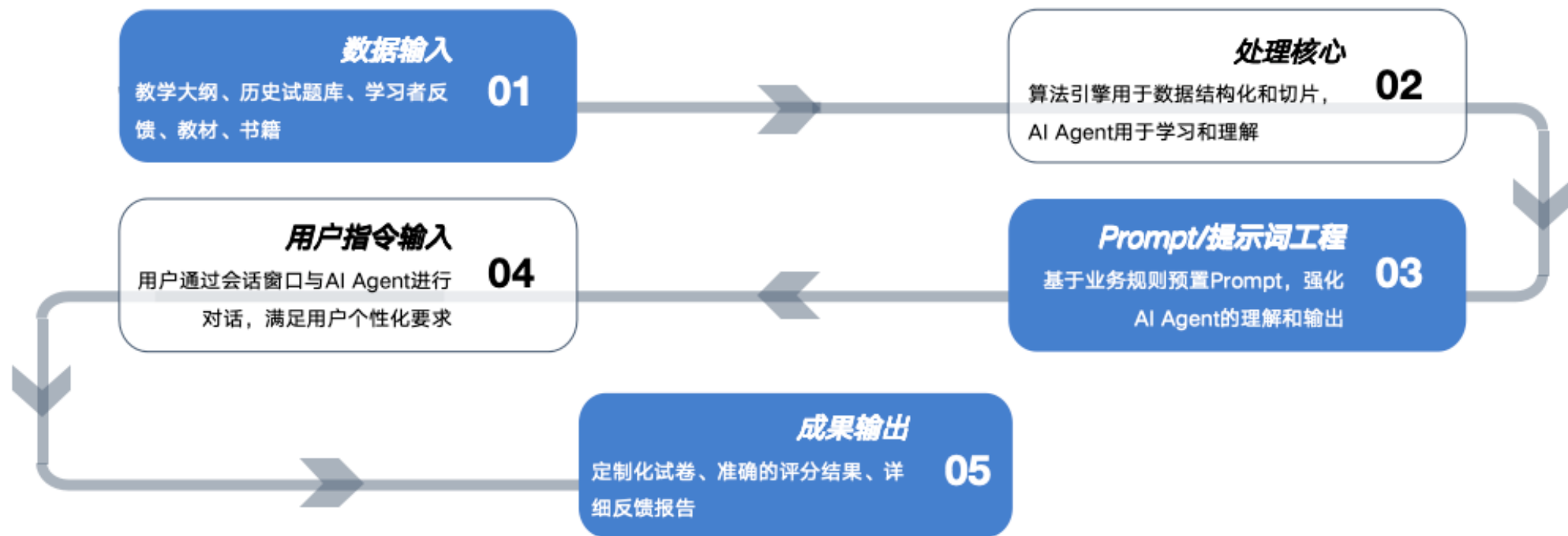
多类型出题

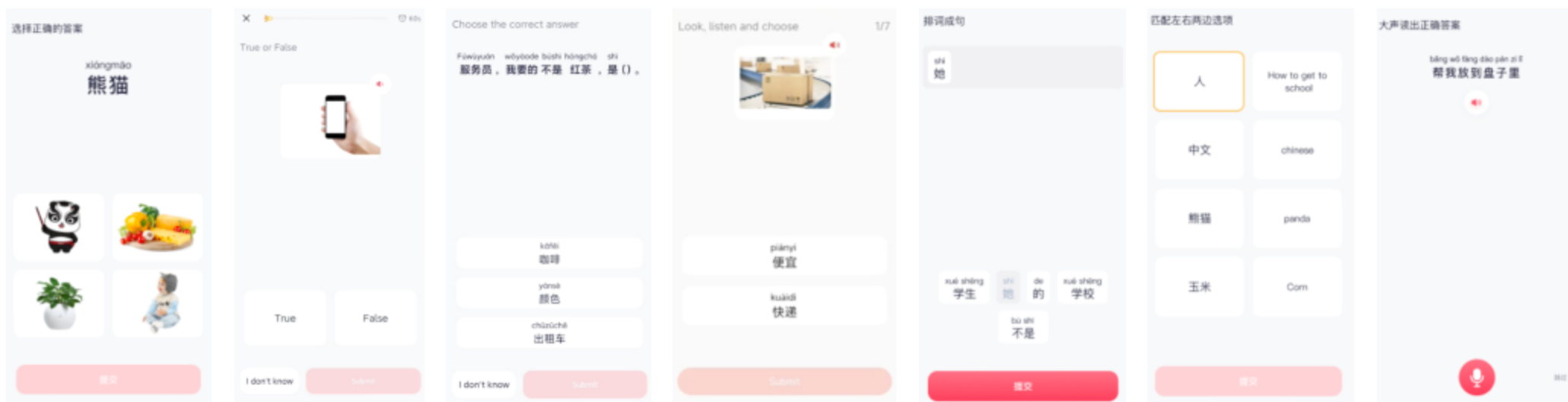
自动评分机制

快速搭建

可扩展与集成性

The Chain for TaoLi Agent





题型	单选题					录音题(跟读题)		排序题			判断题		
题干构成	文本	图片	音频	文本+音频	图片+音频	文本	文本+音频	空白框	音频+框	文本(填空)	文本+图片	文本+音频	图片+音频
样题1	引导语: Choose the correct answer 题干(文字): 面条儿 选项(音频): 面条儿-A: 饺子-B 答案: A ID:	引导语: Look and choose 题干(图片): 米饭 选项(文字): 南瓜-A: 米饭-B 答案: B ID:	引导语: Listen and choose 题干(音频): 饺子 选项(文字): 辣椒-A: 饺子-B 答案: B ID:	引导语: Choose the correct answer 题干(文字): 她在爸爸的() 选项(音频): 她在爸爸的右边 选项A(文字): 上边 选项B(文字): 下边 选项C(文字): 右边 答案: C	引导语: Look, listen and choose 题干: 音频(菲律宾) 选项: 图片(菲律宾) A 菲律宾 B 印度尼西亚 答案: A	引导语: Record 题干: 这个西瓜好甜。	引导语: Listen and repeat 题干: 音频(越南) 文本: 越南	引导语: Make the sentence 题干: ----- A汉堡 B我 C一个 D要 答案: BDCA	引导语: Make the sentence 题干: _____ 选项(音频): 我要一份薯条。 选项: 一份-A: 要-B: 我-C: 薯条-D 答案: CBAD	引导语: Make the sentence 题干: _ 要 _ A汉堡 B我 C一个 答案: BCA	引导语: True or False 题干(图片): 汤 选项(标题): 粥 选项: 正确-A: 错误-B 答案: 错误-B	引导语: True or False 题干(音频): 米饭 选项(标题): 蔬菜 选项: 正确-A: 错误-B 答案: 错误-B	引导语: True or False 题干(音频): 胡萝卜 选项(图片): 胡萝卜 选项: 正确-A: 错误-B 答案: 正确-A



Question Generation Agent (QG Agent)

学科知识点

学科的核心概念、理论、公式、
定义和原则

实际案例和具体实例可以帮助学生更好地理解抽象概念。例如，在教育领域，可以添加实验和文献发现作为实例。

案例研究和实例

题库和习题解析

包括各种类型的练习题和解题策略，以及详细的解析和答案。这不仅帮助Agent生成相关的题目，还能提供解题方法和逻辑推理的示例

包含各个级别和学科的教育标准、课程大纲和学习目标。这有助于确保生成的题目和内容与教育要求和目标保持一致。

教育标准和课程大纲

问题解决技巧

向知识库中加入有关批判性思维、创造性思维和问题解决策略的知识可以提升Agent的教学质量。

Question Generation Agent (QG Agent)



通过Agent搭建

- 利用agent和chain搭
建出题应用。
- 利用好大语言模型的
长处, 规避劣势。
- 丰富ai组件, 满足不
同种类出题需求。
- 大量减少开发周期,
快速搭建。

Question Generation Agent (QG Agent)

AskXBOT

会话

Agent

工作流

知识库

设计器

管理

ramza

我的工作流

运行记录

计划任务

工作流

桃李

导入

新建

桃李AI出题助手_判断题_cjx	桃李AI出题助手_判断题_文本+图片	桃李AI出题助手_排序题_文本(填空)
未发布	未发布	
桃李AI出题助手_录音题_文本+音频	桃李AI出题助手_录音题_文本	桃李AI出题助手_单选题_文本+音频
桃李AI出题助手_单选题_文本	桃李AI出题助手_音频选词(生成同类词)-副本	桃李AI出题助手_文字/音频选拼音(从已学词库...)
	未发布	01月09日 运行成功
桃李AI出题助手_生成句子完型填空	桃李AI出题助手_生成句子拆排序	桃李AI出题助手_音频选词(生成同类词)
	01月09日 运行成功	01月11日 运行成功
桃李AI出题助手_音频判断题	桃李AI出题助手_音频选词(从已学词库挑词)	桃李AI出题助手_造句跟读
运行成功		

<

1

>

运行成功 2024-01-09 14:17:37~2024-01-09 14:17:39 运行时长: 2秒

桃李AI出题助手_文字/音频选拼音(从已学词库挑词)

输入参数

题干:

飞

输出结果

题目:

⇌ ☒

题干	音频	选项A	选项B	选项C	选项D	答案
飞	xbot-c... ▶ ⬇	ma	sān	fēi	zuò	C
飞	xbot-c... ▶ ⬇	jiǔ	fēi	běn	tā	B
飞	xbot-c... ▶ ⬇	lái	fēi	jiào	ne	B
飞	xbot-c... ▶ ⬇	shū	le	fēi	xiǎng	C
飞	xbot-c... ▶ ⬇	mǎi	fēi	néng	lǐ	B
飞	xbot-c... ▶ ⬇	wǔ	diǎn	le	fēi	D
飞	xbot-c... ▶ ⬇	xiǎo	qī	sān	fēi	D

选择题

文字

判断题

音频

填空题

图像

排序题

语音识别

是非题



通过题型和类型组合
出丰富的题目类型

文字选择题

选择题

文字

听音选词

选择题

文字

音频

听力题

填空题

音频

文字

排序题

填空题

音频

文字

更多题型.....

Intelligent Scoring Agent

解决方案

- 支持**专家/老师知识/教案的上传**，可不断补充、更新文档，保持文档实时性。
- 通过**意图识别分发、关键词与向量化多路召回、知识融合等手段**，来理解学生在答案或文章中试图表达的主题和观点，并为其打分。
- 支持Agent和用户间的**多轮交互**，修正应用效果。
- 支持，**用户反馈、日志追踪等，不断迭代优化效果。**

方案优势



支持文档、文本、数据、等多个钟类型



支持多语种，多类型的考题生成，及AI审卷评分



教案上传与知识更新：实现智能评分系统的持续学习和优化进步



支持SaaS或私有化的部署形式



开箱即用

大语言模型判卷和作文评分系统不仅可以快速、准确地评估学生答案质量，更进一步提供了基于个体表现情况进行试题难度及内容主题动态调整功能。



提高效率

通过自动化的处理流程，大语言模型能够快速准确地完成大量试卷和作文的评分工作。这让教师们有更多时间专注于课堂教学与个别辅导。



可定制性强

结合意图识别、关键词向量化以及知识更新等技术手段，我们可以根据特定需求调整系统设置来满足不同科目或者具体场景要求。



沉淀学校资源

AI出题软件不仅可以满足大量出题的需求，同时它也能实现对海量作答数据的收集和处理，进而优化算法、改善出题质量和个性化程度

Intelligent Scoring Agent

桃李评分-短句

返回

保存

▶ 调试

🚀 发布

变量 <

🔍 请输入组件名称搜索

模型

知识



添加知识

将文本或文档添加至知识库
并完成知识预处理。



知识检索

基于文本内容检索相关知识
片段。



知识问答

基于知识内容检索推理获得
答案。

🔧 输入

句子 × 考题 ×

🔧 代码

输入
考题 ×

```
1 import json, random
2 import pymysql
3
4 error_msg = {"异常": "\U0001F62F",
5              "信号状态": "\U0001F62F"}
6
7 def search_database(title):
```

输出
考点 ×

🔧 LLM

模型

Default

创造性

低 中 高

提示词

你是一个考试评分专家，专业评判HSK5
级的考生的作文。你熟练掌握中文语法规
则并具有深厚的文学素养。

==评分技能==
你需要使用评分技能进行评价

==规则==
- 客观

输出

评分

🔧 输出

评分 ×

+ - 🔍 📄

Intelligent Scoring Agent

HSK 评分说明				
HSK 各级别考试的听力、阅读、书写各部分满分均为 100 分，总分为各部分分数相加之和。				
各级别听力和阅读部分计分方式如下：				
	听力		阅读	
	每题分值	题量	每题分值	题量
HSK 一级	5	20	5	20
HSK 二级	2.86	35	4	25
HSK 三级	2.5	40	3.34	30
HSK 四级	2.23	45	2.5	40
HSK 五级	2.23	45	2.23	45
HSK 六级	2	50	2	50

各级别书写部分计分方式如下：				
	题量		分值	满分
	完成句子	写汉字		
HSK 三级	5	5	12	60
	10	10	8	40
HSK 四级	5	5	8	40
	10	10	6	60
HSK 五级	8	8	5	40
	2	2	30	60
HSK 六级	1	1	100	100

附：书写部分各题型评分说明（针对每个小题）

- 完成句子题
0 分：空白。
低档分：未包含所提供的全部词语；
词序排列不正确；
有 3 个或 3 个以上错别字。
中档分：词序排列基本正确但增加了未提供词；
词序排列正确但有 1-2 个错别字。
高档分：包含全部词语且词序排列正确，无错别字。
- 写汉字题
0 分：空白或写了与答案完全不相关的字。
中档分：与答案相近，是错字。
高档分：与答案一致。
- 看图用词造句
0 分：空白。
低档分：内容与图片及所提供的词语无关。

特色

深度理解: 利用大语言模型(LLM)强大的自然语义处理能力，实现对各类题目答案内容精准把握。

持续进化: AI Agent通过不断学习与更新，实时提升评分标准适应性并保证最佳效果输出。

个性关怀: LLM 与 AI Agent联手打造高水平定制服务，精确反映每位学生表现并为其未来成长路径指引方向。





END

桃李中文团队报告