

教育部 - 华为智能基座课程  
山东大学人工智能研究生通识示范课程

# 人工智能基础与前沿



课程群

主讲：丛润民

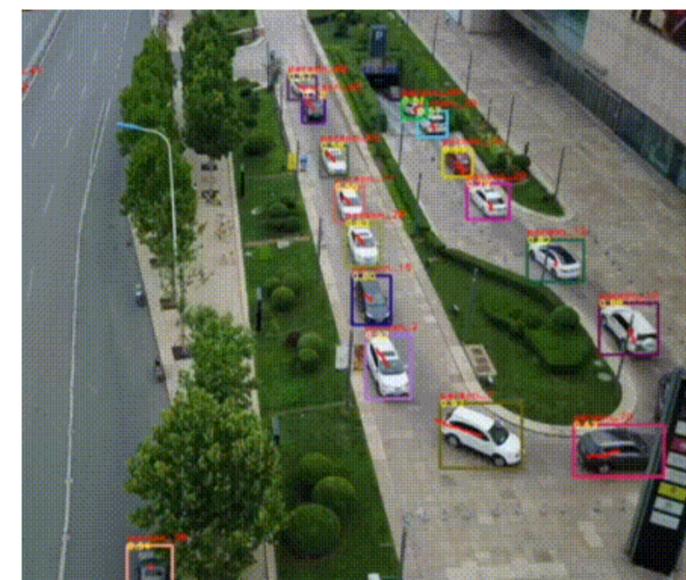
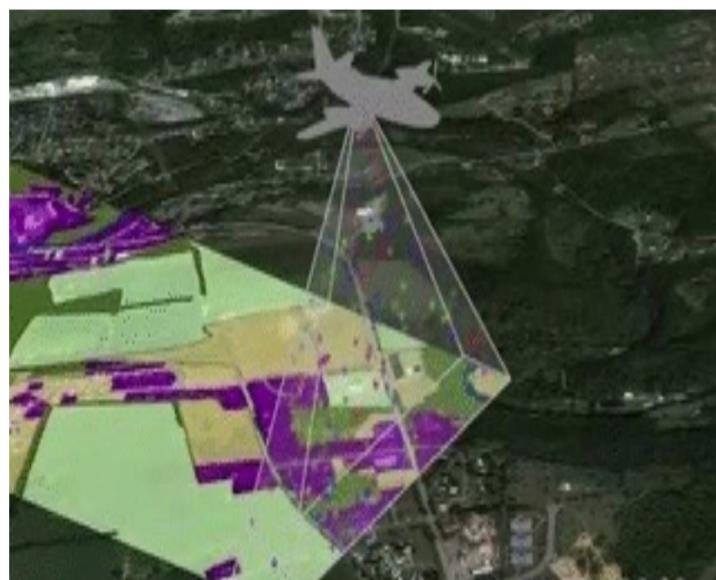
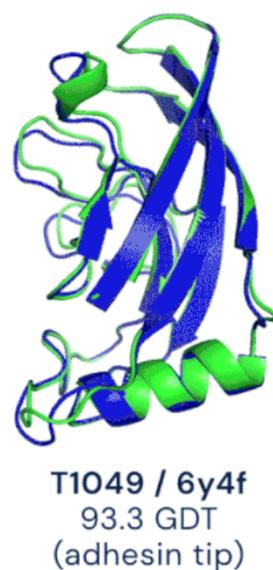
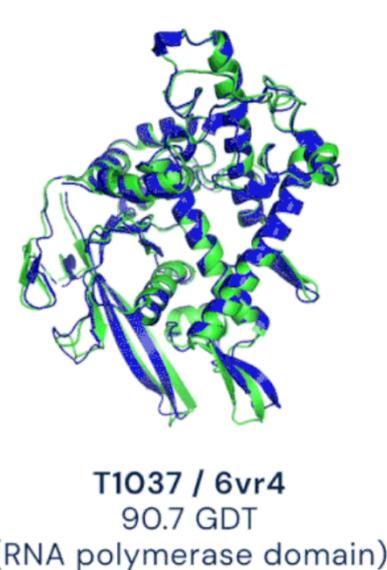
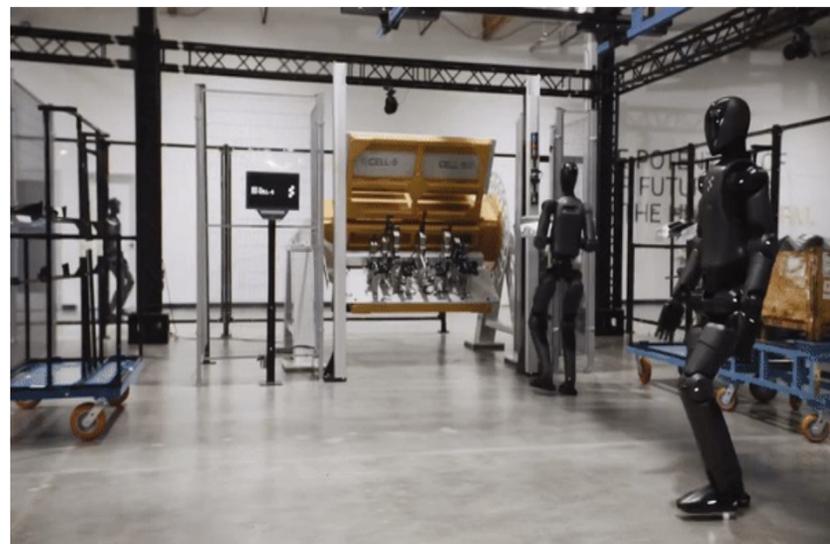
人工智能基础与前沿

为天下储人材 为国家图富强

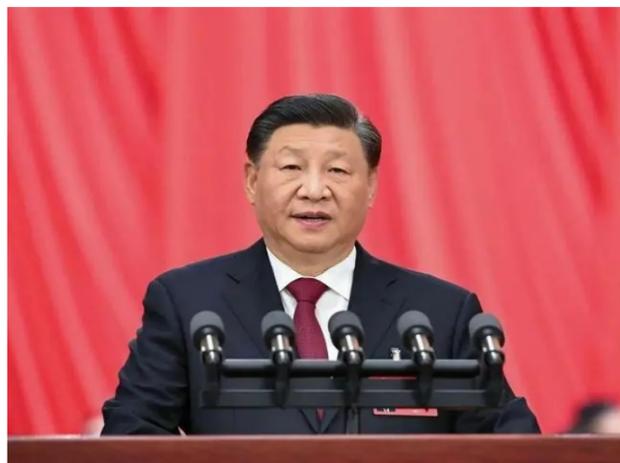


# 人工智能技术的重要意义与价值

- 人工智能作为当今技术发展的前沿，正在**重新定义**人类社会的各个方面
- 人工智能不仅是科技进步的产物，更是推动社会智能化变革的**关键引擎**



- 人工智能作为当今技术发展的前沿，正在**重新定义**人类社会的各个方面
- 人工智能不仅是科技进步的产物，更是推动社会智能化变革的**关键引擎**



2016年5月30日，习近平总书记在“科技三会”讲话指出——“**科技创新、科学普及是实现创新发展的两翼，要把科学普及放在与科技创新同等重要的位置。**”



2022年9月4日，中共中央办公厅 国务院办公厅印发的《关于新时代进一步加强科学技术普及工作的意见》指出——“**高等学校应设立科技相关通识课程，满足不同专业、不同学习阶段学生需求。**”



# 山东大学研究生人工智能通识教育实践

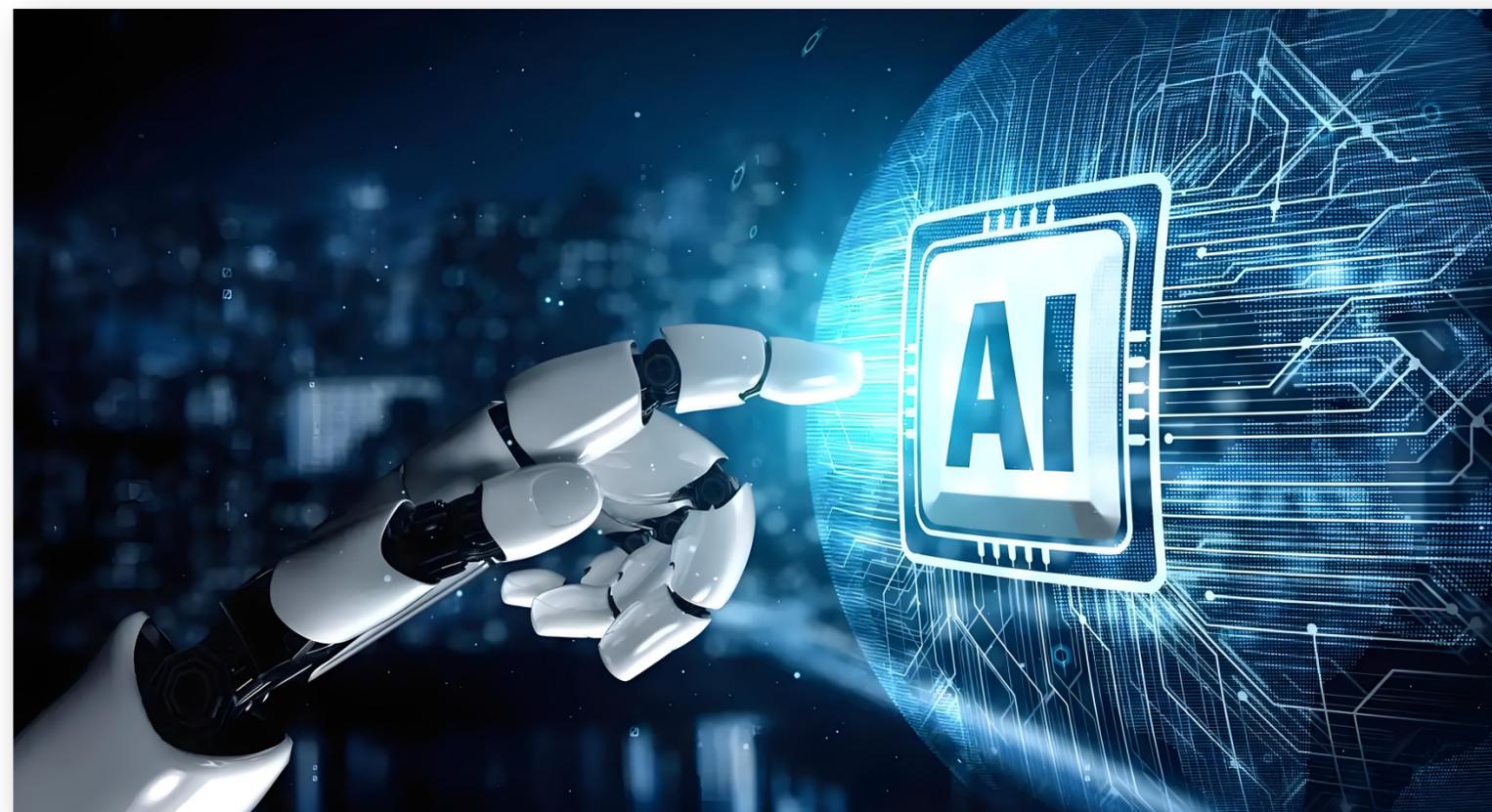


■ 建设山东大学研究生人工智能通识精品课程群，**29门**AI精品课程正式启动建设，实现**100%覆盖研究生新生**，**100%覆盖文理工医各学科**



建设《人工智能基础与前沿》研究生AI通识课，覆盖理工农医400名硕博

- 课程名称：**人工智能基础与前沿**
- 英文名称：**Fundamentals and Frontiers of Artificial Intelligence**
- 课程性质：**通识必修课程**（课号：0440142）
- 面向专业：全校所有专业
- 学时：16学时
- 学分：1学分
- 开课学期：秋季





# 课程资源



人工智能通识  
丛润民等著



## 人工智能通识 AI课程

学分: 2 学时: 29

本周新增访问 **↑157** [教学空间](#)

**线上**

**课程简介**

作为引领当代科技革命的核心驱动力，人工智能正在深刻重塑人类社会的发展图景，其影响力已从技术领域扩展至经济、文化、教育等各个维度。在此背景下，人工智能教育正经历从专业领域向通识教育的战略性转型。因此，《人工智能通识》课程应...

**虚拟教研室**

学校 山东大学

教师

**课程相册**

0人

累计人数

本学期合计 0人

0所

累计学校

本学期合计 0校次

0次

累计互动

本学期合计 0次

更新时间: 2025-08-25

312次 累计浏览

**能力体系**

- 能力目标
- 子能力
- 素养目标

**问题体系**

- 问题
- 实践
- 实验

**知识体系**

- 知识点
- 技能点
- 教案

**教学资源**

- 教学资源
- 视频
- 虚拟仿真
- 教材

**18**  
能力

**61**  
问题

**50**  
知识点

**120**  
教学资源

[AI指令 笔记整理](#) [AI智能体 答题辅导...](#) [AI助教](#)

## 《人工智能通识》配套智慧树课程



# 考核方式



平时成绩100分		期末成绩100分		总成绩100分
出勤	智慧树 在线学习	AI思辨	课程报告	平时50%+期末50%
50分	50分	30分	70分	100分

教育部 - 华为智能基座课程

人工智能研究生通识示范课程



山东大学  
SHANDONG UNIVERSITY

# 第1讲 人工智能概述

主讲：丛润民

人工智能基础与前沿

為天下儲人材 為國家圖富強



01

什么是人工智能?



02

图灵测试



03

人工智能发展史



04

ChatGPT——走向通用AI之路



05

大模型时代下的人工智能



# 目 CONTENTS 录





01

# 什么是人工智能？





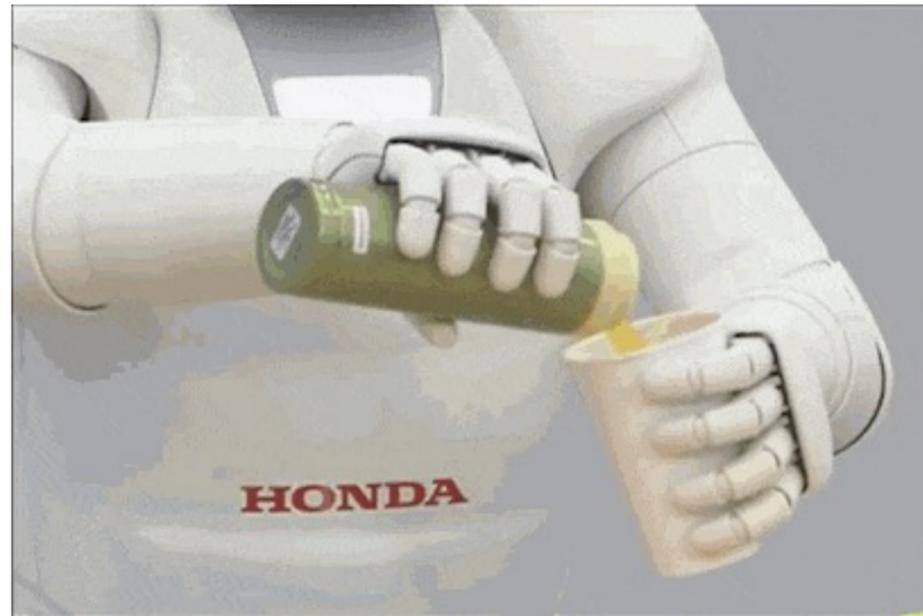
科普中国  
CHINA SCIENCE COMMUNICATION

# 什么是人工智能？

视频来源：科普中国

## 1.1 人工智能的定义

- 人工智能作为一门新兴的技术科学，其核心在于研究、开发用于**模拟、延伸及扩展人类智能**的理论、方法、技术与应用系统。
- 作为计算机科学的一个重要分支，人工智能致力于使计算机能够模拟人类智能的特定过程与行为，通过这些模拟实现原本需要人类智能才能完成的功能。



## 1.1 人工智能的定义

### ■ 在**模拟人类智能**方面：人工智能通过构建智能系统来模拟人类的思维过程。

- 机器学习算法可以学习并改进自身的性能，从而逐渐提高处理任务的能力。
- 自然语言处理则关注于让计算机理解和生成人类语言，实现与人类的自然语言交互。

### ■ 在**延伸人类智能**方面：人工智能可以辅助人类完成一些复杂或繁琐的任务。

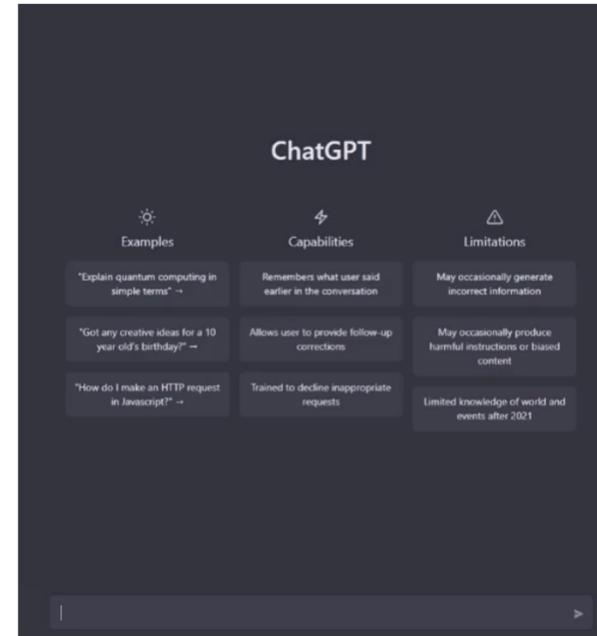
- 在医疗领域，人工智能可以帮助医生分析病例、诊断疾病，提高医疗服务的效率和质量。
- 在交通领域，自动驾驶技术可以辅助驾驶员驾驶车辆，减少交通事故的发生。

### ■ 在**扩展人类智能**方面：人工智能可以探索新的智能形式和智能应用。

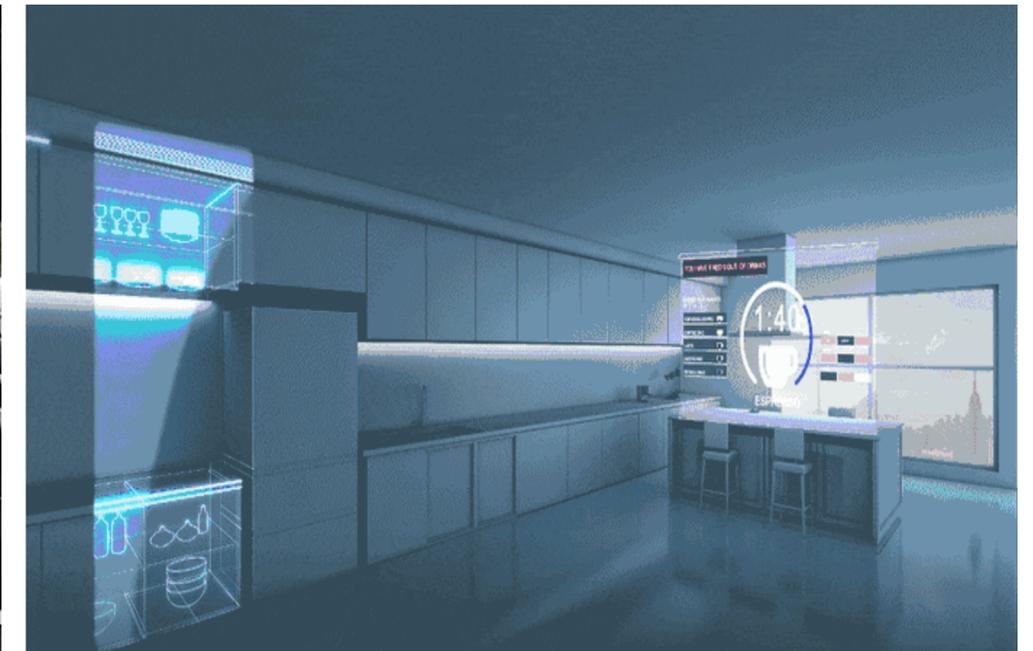
- 人工智能通过数据分析发现数据中的规律和模式，从而帮助人类解决一些复杂的问题。
- 人工智能还可以应用于创意领域，如音乐创作、绘画等，创造出具有独特风格的作品。

# 1.1 人工智能的定义

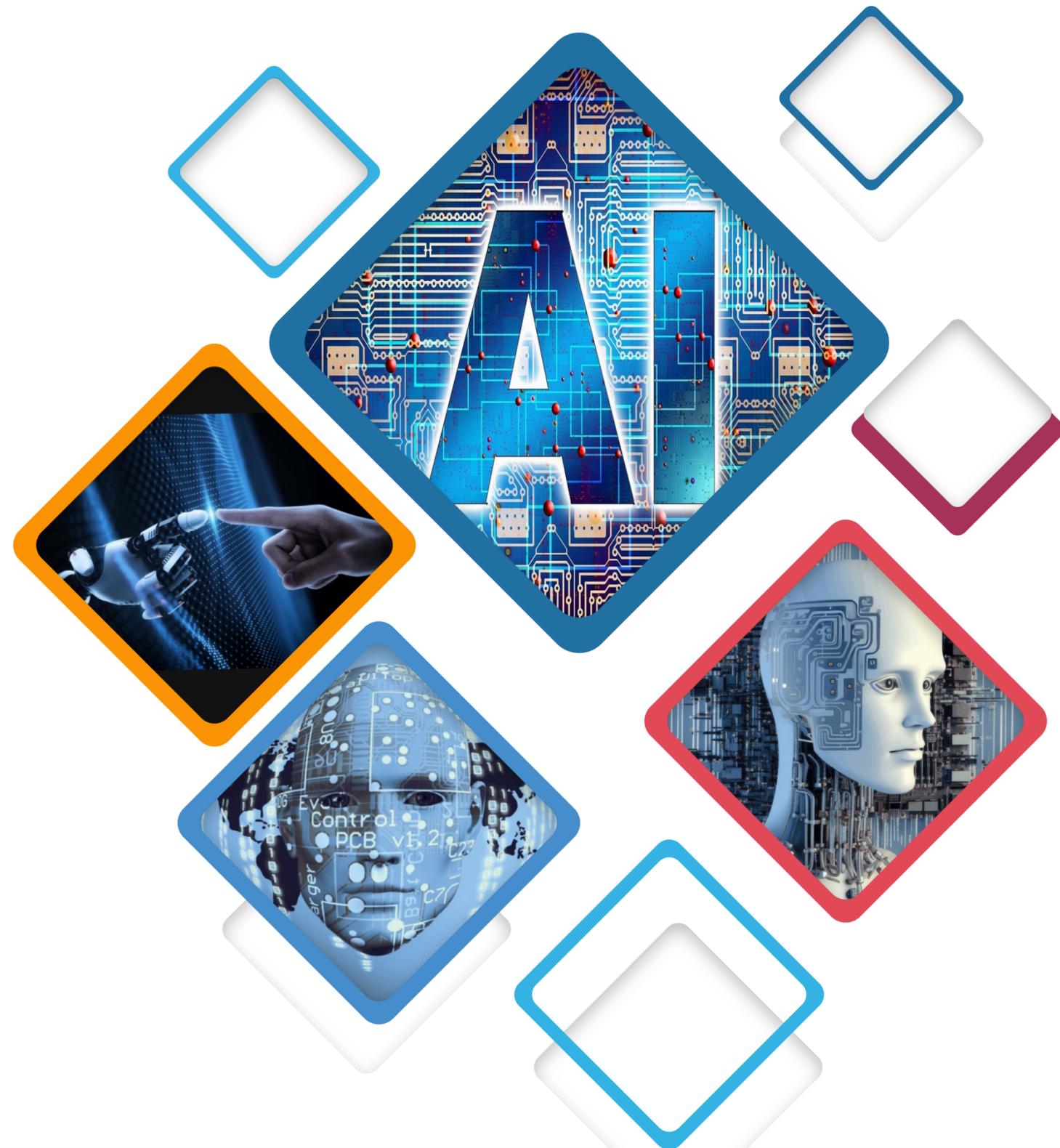
■ AI在**思想层面**的应用不仅扩展了人类的认知边界，还提升了决策的准确性和效率，使得人类能够更好地应对复杂和快速变化的世界。



■ AI在**行为层面**的应用通过自动化操作和智能助手，极大地提升了生活和工作的便利性，使得人类能够更高效地完成任务，享受更加智能化和个性化的服务。



- AI 的价值和意义重大，主要体现在：
  - **在经济方面**，能提高生产效率、推动产业升级，创造新商业模式与就业机会，助力经济发展。
  - **在社会层面**，可改善生活质量，如智能家居带来便利；提升医疗水平，辅助诊断疾病等；还能助力教育，实现个性化学习。
  - **在科学研究方面**，能加速探索进程，处理分析海量数据，帮助科学家发现新规律现象，推动多领域研究突破。



# 1.2 想一想

## 人工智能是否可以超越人类智能？



### deepseek

人工智能（AI）在某些特定任务上已经超越了人类，例如在棋类游戏、数据分析和模式识别等领域。然而，AI目前仍然无法在广泛的认知任务和复杂情境中全面超越人类智能。人类的智能具有高度的灵活性、创造力和情感理解能力，这些是目前AI难以复制的。

未来，随着技术的进步，AI可能会在更多领域接近甚至超越人类的能力，但这需要解决许多技术和伦理问题。总体而言，AI和人类智能各有优势，更可能的发展方向是两者互补，而不是简单的超越关系。

人工智能在某些特定领域，如数据处理、模式识别等方面，可能表现出超越人类的能力。然而，在整体智能、意识、情感理解、社会属性以及自然语言处理等方面，人类智能仍具有无法替代的独特优势。此外，人类的创新能力也是人工智能难以企及的。因此，尽管人工智能技术在不断发展，但全面超越人类智能的可能性仍然很低。未来，人工智能将更多地作为人类的辅助工具，与人类共同推动社会的进步和发展。



### 豆包 app

人工智能能否超越人类智能存在争议。从计算、存储和特定任务处理上，人工智能有优势，如能快速处理海量数据、在棋类游戏战胜人类。但人类智能的创造力、情感、道德判断及常识推理等是人工智能难以企及的，人类可从无到有创造，能基于情感做复杂决策。目前，人工智能在特定领域表现出色，但全面超越人类智能还远未达到，未来发展走向仍需持续观察与研究。

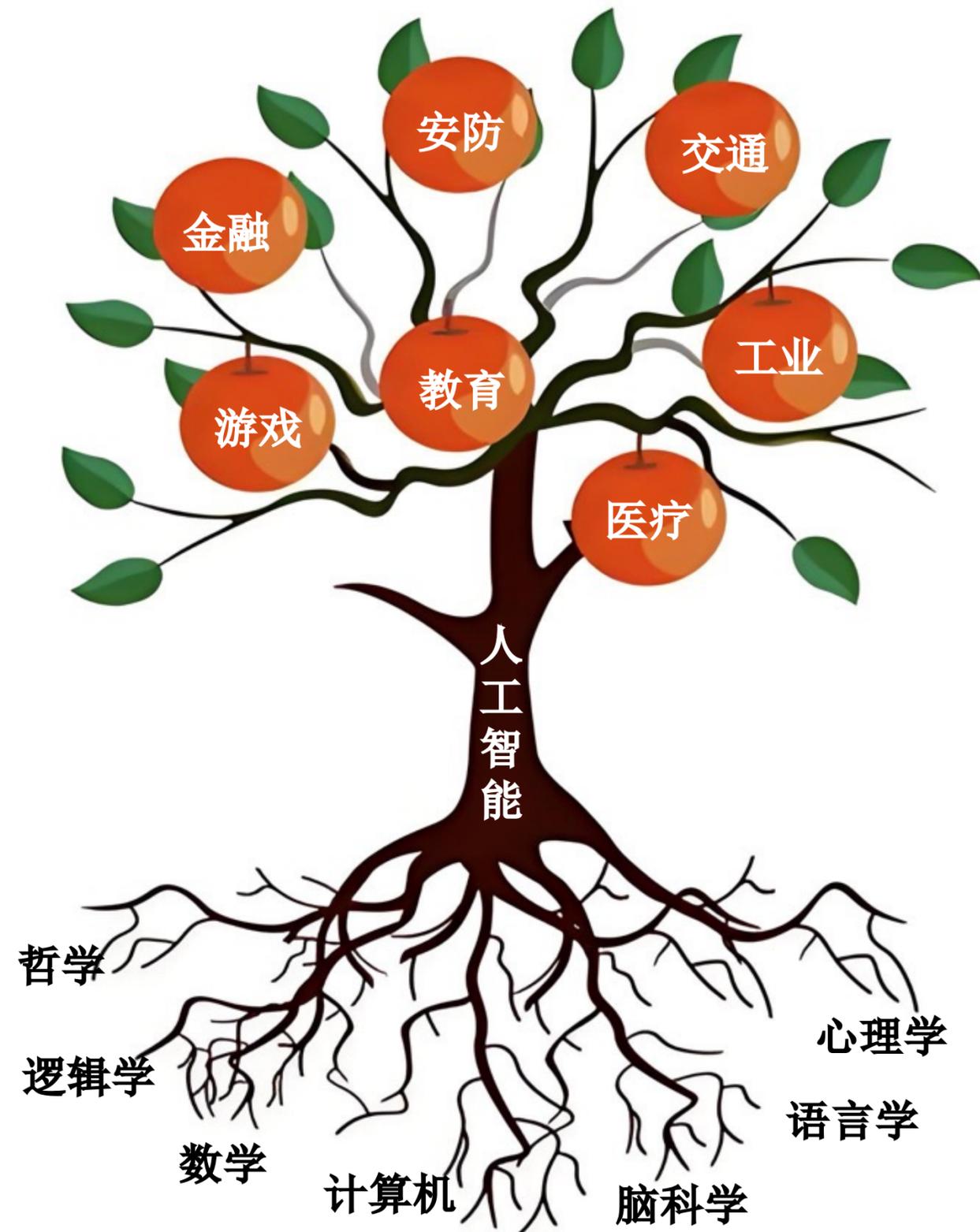
人工智能在某些方面已经展现出超越人类的能力，例如计算速度、数据处理和重复性任务执行。它能够快速处理海量信息，精准执行任务，且不会因疲劳或情绪影响效率。然而，人类智能具有独特的情感、创造力和复杂决策能力。人工智能目前还难以完全理解人类的情感和价值观，也无法像人类一样在复杂多变的环境中灵活创新和做出道德判断。所以，人工智能在未来可能在特定领域超越人类，但在整体智能上，人类智能的复杂性和灵活性是人工智能难以完全替代的。

复制 再试一次 分享





- 人工智能是一个广泛的科学领域，它融合了多种学科和技术，如数学、计算机科学、心理学、哲学等。
- 这些不同学科的基础要素不仅决定了AI系统的功能和性能，还影响了其应用的广度和深度。



## ■ 数学和统计学为人工智能奠定了理论基础。

### 数学

线性代数、微积分、概率论和统计学等数学知识都是人工智能领域的重要基础知识。比如，数学中的求导方式对于理解反向传播算法至关重要，线性代数中的矩阵表达方式通常用于处理多维数据。

### 统计学

概率论和统计学知识则用于分析数据分布、建模不确定性以及在数据中识别模式，它为人工智能提供了处理数据、建立模型和做出推断的理论和方法。

## ■ 计算机科学为人工智能奠定实现基础。

### 编程语言

Python、Java和C++等编程语言是支撑AI算法实现的具体化工具。并行计算、分布式系统和高性能计算则为大规模数据处理和复杂模型训练提供了计算能力。

### 机器学习

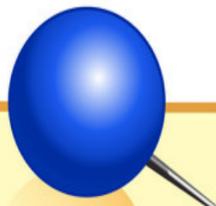
机器学习是人工智能的重要手段和方法之一，也是人工智能实现的重要基石。借助机器学习（包括深度学习技术）可以赋予计算机自主学习的能力，进而更好地模拟人的智能。

### 数据科学

数据科学为人工智能提供了大量的数据，并且通过数据分析和挖掘技术捕获数据中蕴藏的规律，为人工智能算法的设计和 optimization 提供燃料。

## 1.3 人工智能的基础

### ■ 认知心理学为人工智能提供了模拟和扩展人类智能的目标和参照。



认知心理学的研究成果可以帮助人工智能系统更好地理解 and 适应人类的认知习惯和规律，从而减少错误，提高模型的认知准确性。

认知心理学主要研究人类的高级心理过程，主要是认识过程，如注意、知觉、表象、记忆、创造性、问题解决、言语和思维等。



- 认知心理学与人工智能的深度融合催生了情感计算这一交叉领域。情感计算旨在使机器能够识别、理解、模拟和表达人类情感，从而实现更自然的人机交互。通过对人类情绪和情感的研究，人工智能系统可以更准确地解读用户的情感状态，提供更贴心的服务。

## ■ 伦理和社会科学成为确保AI的发展是可持续的、负责任的关键。



伦理原则旨在引导人工智能技术的发展，使其在提升效率和创新的同时，不会损害个人、社会 and 环境的利益，确保技术进步与人类的伦理价值观相协调，促进建立一个公正、安全和可持续的人工智能未来。

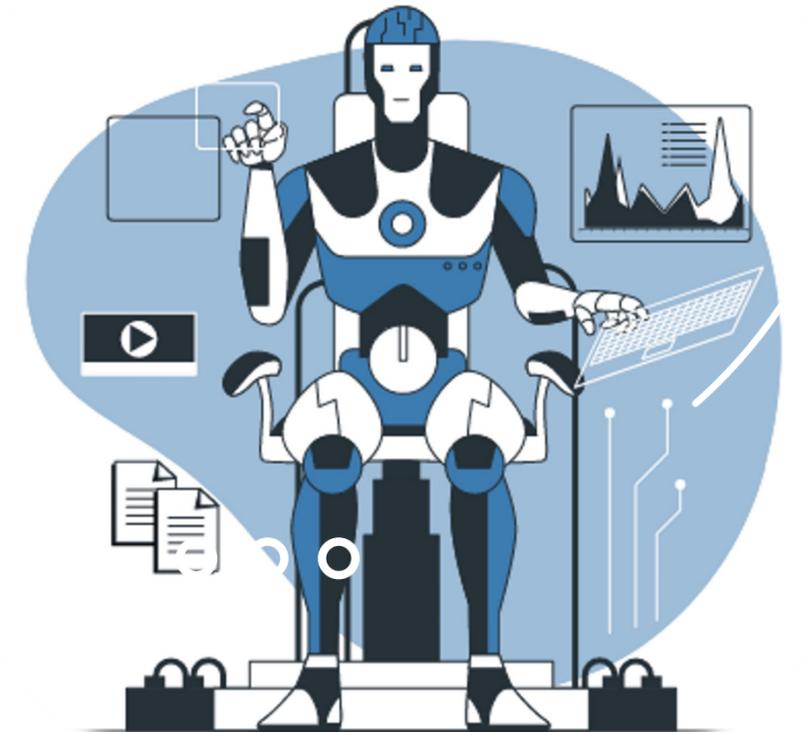
人工智能伦理是确保人工智能技术发展与应用过程中符合道德规范和社会责任的一系列原则和标准，它强调尊重人权、公平无歧视、透明度、责任归属、隐私保护、安全性、促进社会福祉以及持续的伦理审查。



人工智能伦理

# 1.3 人工智能的基础

- 人工智能是一个复杂且多维度的系统，它涵盖了从理论到实践的多个领域，这些学科或者技术相辅相成，相互促进，共同推动了人工智能的快速进步。
- 同时人工智能技术也在不断反哺这些基础学科，例如利用人工智能技术可以求解数学中的偏微分方程、可以辅助大学生进行心理筛查等。



## 第一层次



计算智能

能存能算

## 第二层次



感知智能

能听会说、能看会认

## 第三层次



认知智能

能理解、会思考

计算机具有**快速计算**和**记忆存储**能力：



**计算智能** 能存储会计算

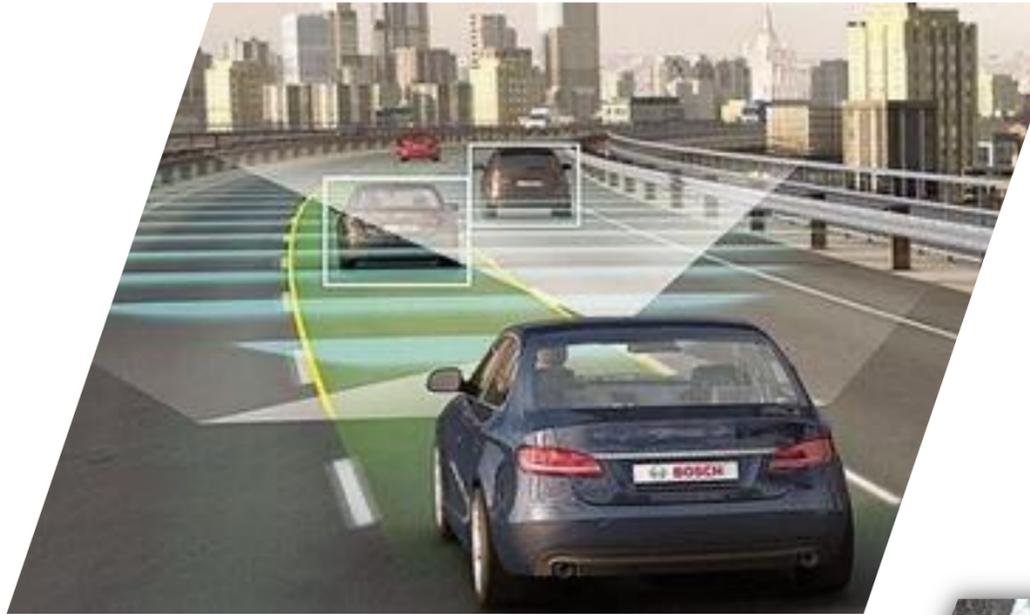
1996年首次对决中，国际象棋大师卡斯帕罗夫以4:2的战绩击败IBM深蓝。但在1997年，他输给了卷土重来的“深蓝”（Deeper Blue）

(1) “深蓝”重量达1.4吨，有32个节点，每个节点有8块专门为进行国际象棋对弈设计的处理器，平均运算速度为每秒200万步。

(2) 深蓝算法的核心是基于**暴力穷举**：生成所有可能的走法，然后执行尽可能深的搜索，并不断对局面进行评估，尝试找出最佳走法。

1.4

# 人工智能的三个层次——感知智能



**感知智能** 能听会说  
能看会认



类似于人的视觉、听觉、触觉等感知能力



**问题1：露营的有多少人？**

**问题2：抵达时间，今天还是几天前？**

**问题3：他们是靠什么工具抵达到这里的？**

**问题4：附近是否有城镇？**

1. 有四位游客，因为野餐毯上有4把勺子，值班名单上有4个人的名字。
2. 他们几天前到达的，因为在那段时间在他们的帐篷和一棵树之间出现了一张蜘蛛网。
3. 乘船，树旁边是船桨。
4. 没有小镇，但一个村子在不远处，因为有一只鸡在那儿闲逛。

1.4

# 人工智能的三个层次——认知智能



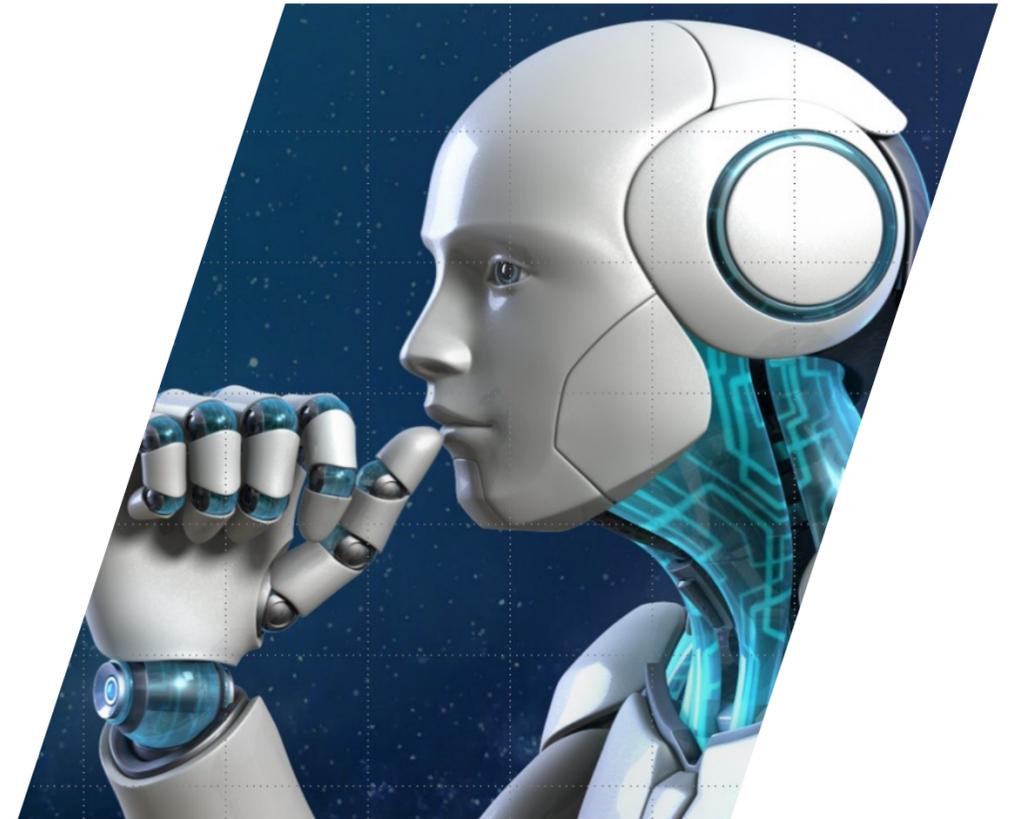
逻辑推理



知识理解



决策思考



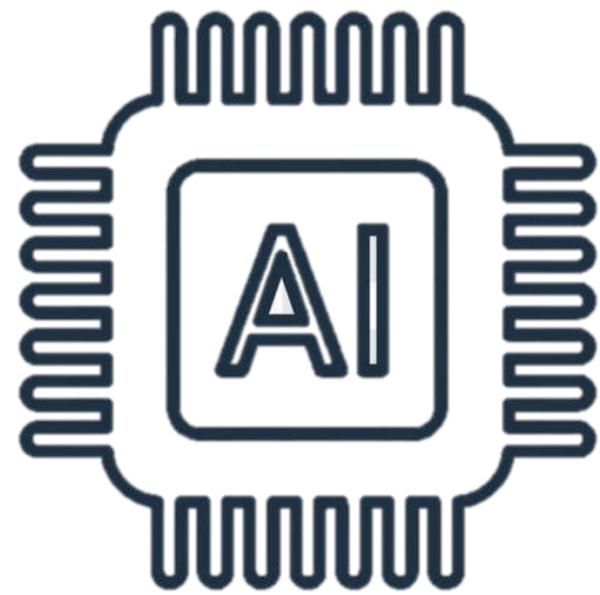
**认知智能** 能理解、思考、决策

概念、意识、观念都是认知智能的表现

# 1.5 人工智能+生活/行业



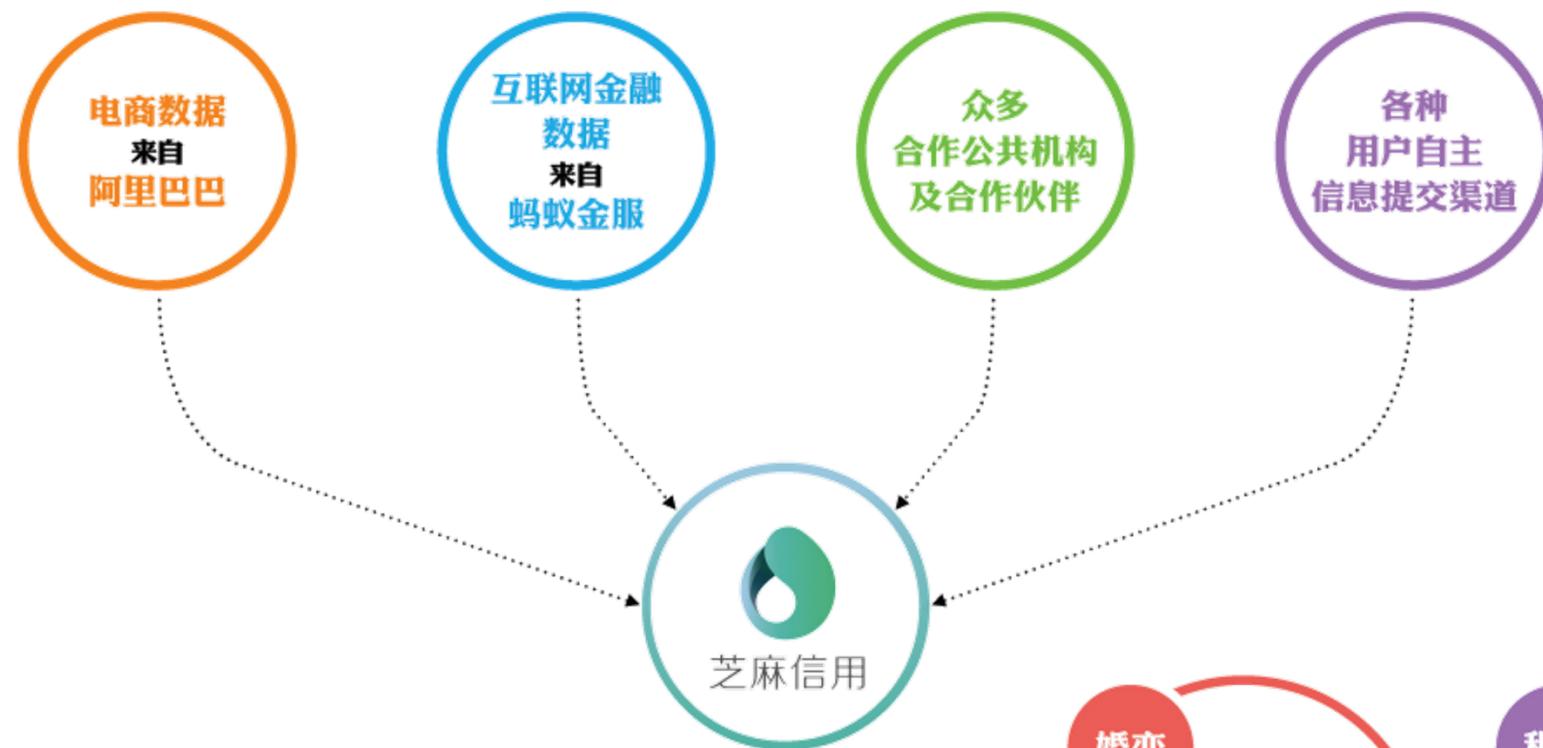
AI+生活



AI+行业

# 1.5 人工智能+金融





消费信贷  
消费分期  
信用卡  
P2P



# 1.5 人工智能+法律

欧盟《一般数据保护条例》规定：自2018年5月25日起要求所有算法解释其输出原理，保证数据隐私和算法公平性



**1 人工智能法律、伦理**  
建立人工智能法律法规、伦理规范和政策体系

**3 人工智能+法学培养模式**

- 法律人懂人工智能、会用人工智能
- 不仅传授知识本身，更强调创新、实践能力和判断力的塑造

国务院关于印发新一代人工智能发展规划的通知



**2 智慧法庭**  
促进人工智能在证据收集、案例分析、法律文件阅读与分析中的应用，实现审判体系和审判能力智能化

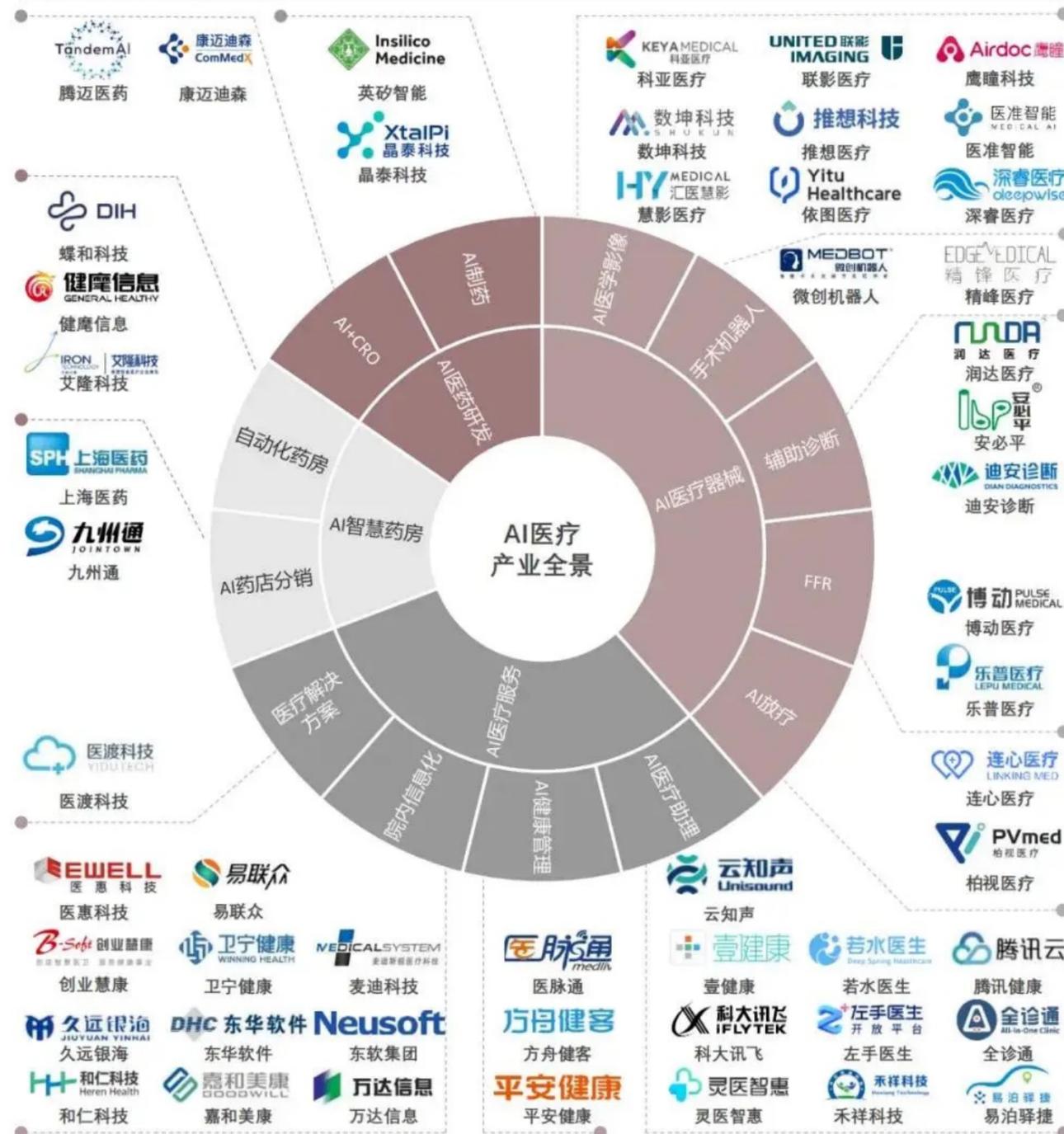
## 法律人工智能科技



AI应用程度 参与主体	信息化 (IT/互联网思维)	智能化 (人工智能思维)
当事人	法律咨询 律师推荐	文书处理
律所/律师	信息管理	智能客服 文书检索 案情预测 文书处理
法院/法官	互联网法院 信息管理	智能客服 文书检索 案情辅助审判 文书处理



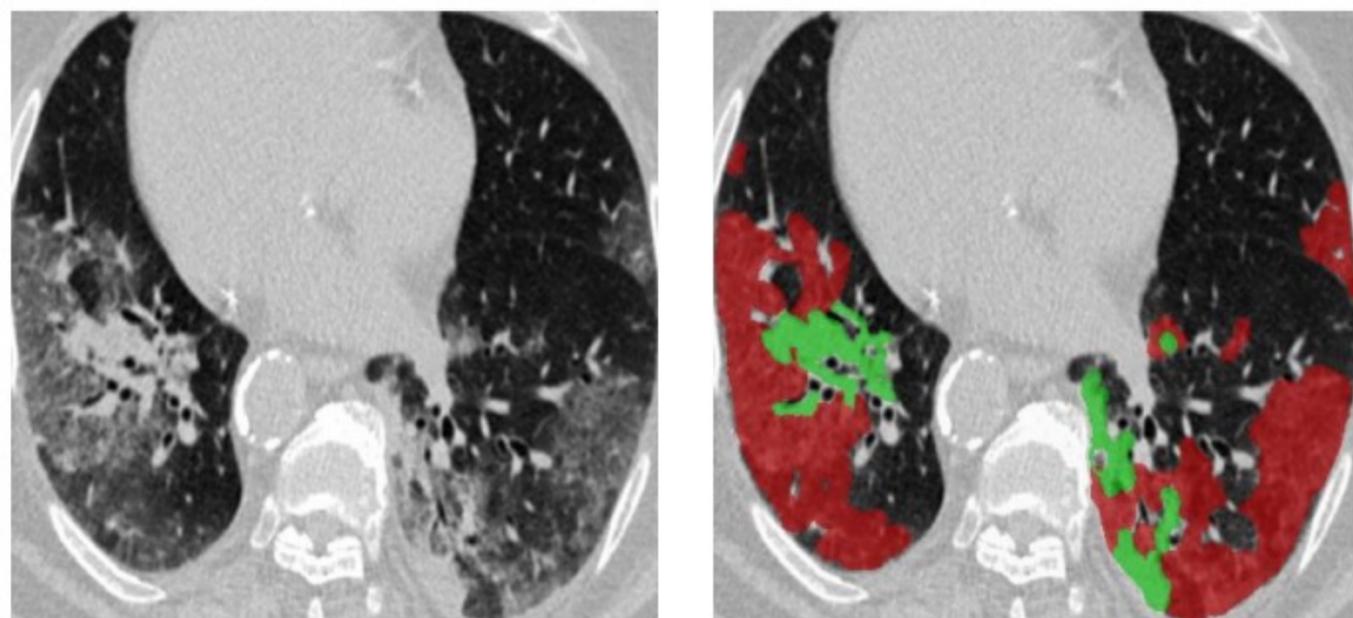
### 中国AI医疗产业全景图



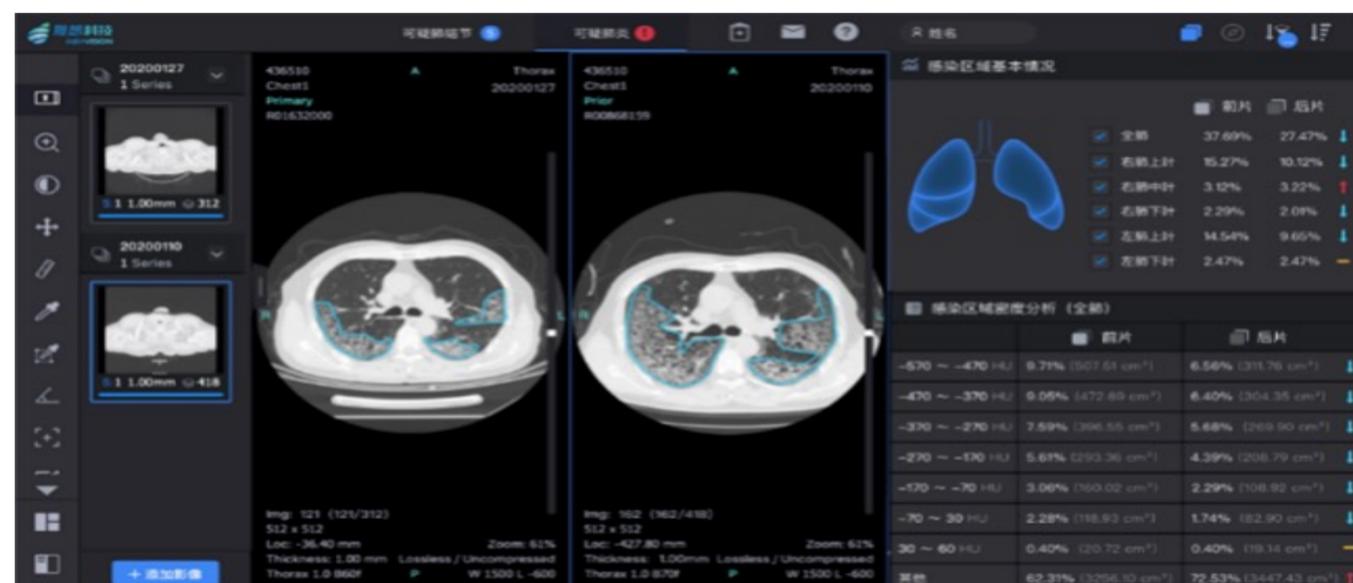
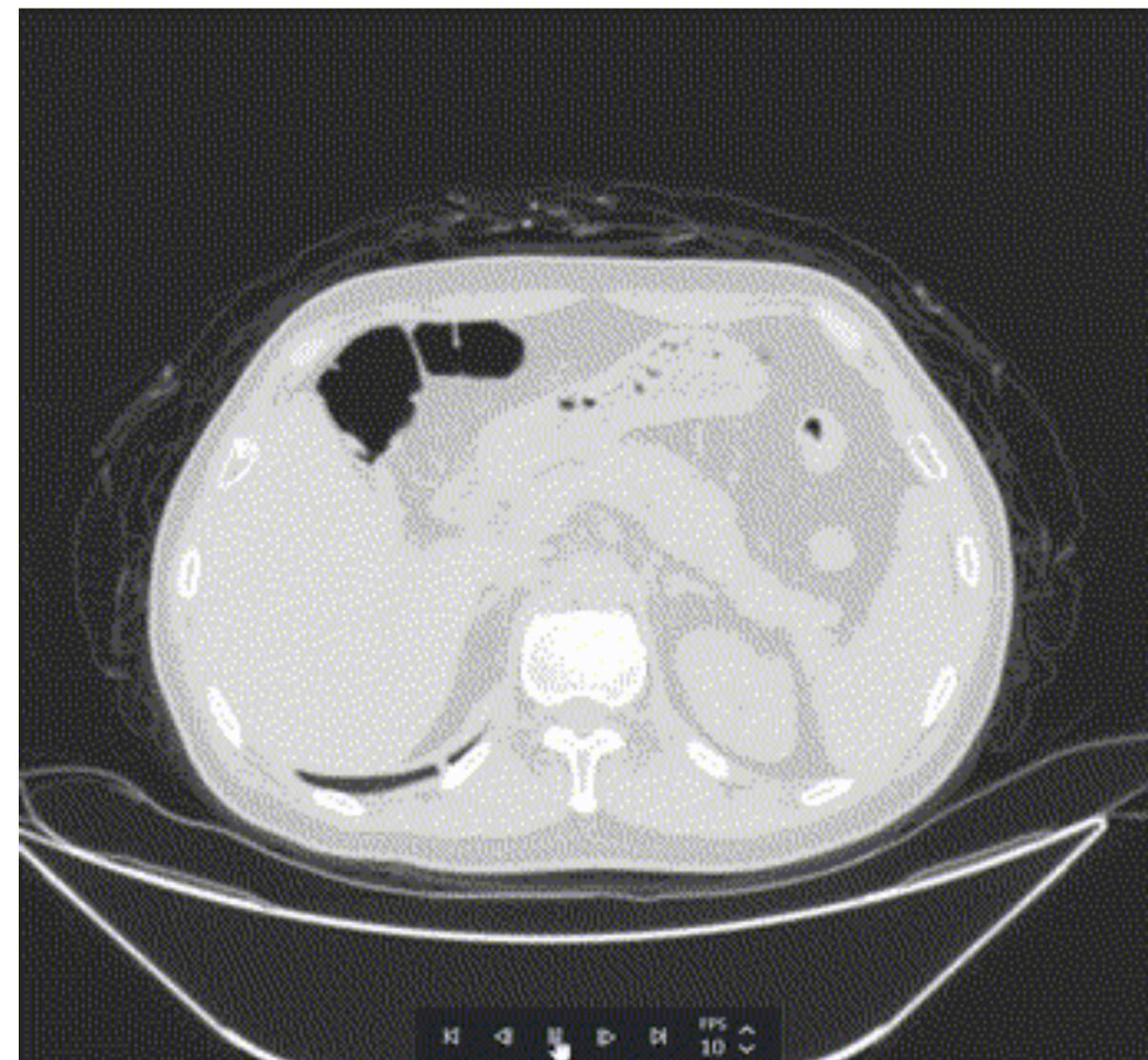
\*仅列出部分代表企业

来源: 头豹研究院

## COVID-19肺部感染图像分割与筛查

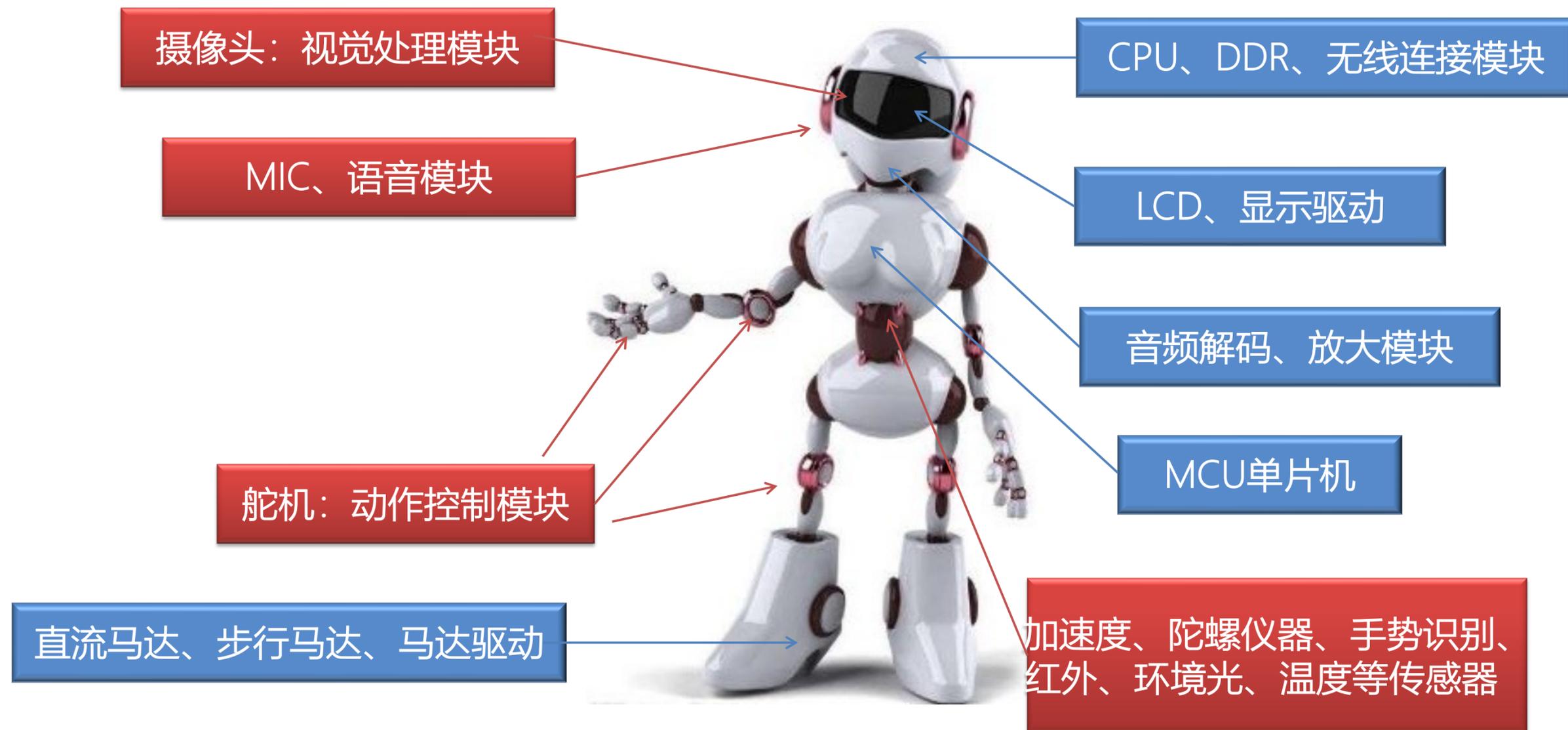


初期：毛玻璃现象（红色） 后期：肺实变（绿色）

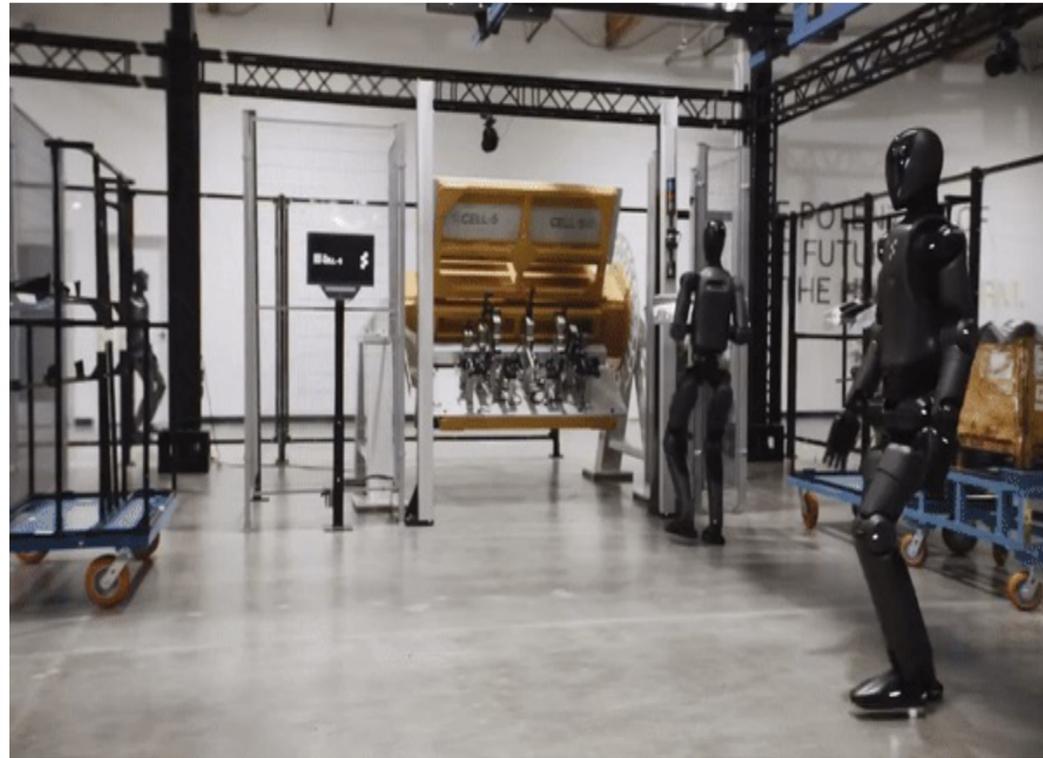
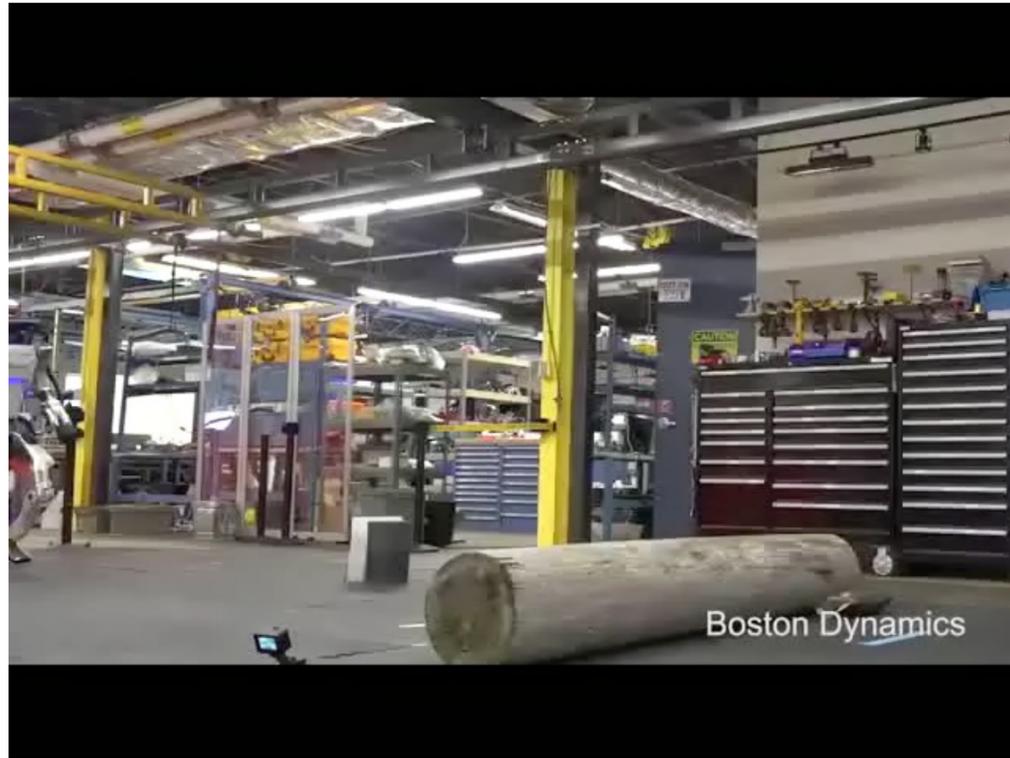


解密新冠肺炎CT影像AI筛查 <http://mmcheng.net/covid-ai/>

# 1.5 人工智能+机器人



# 1.5 人工智能+机器人



扫地机器人



拣货机器人



医疗机器人



教育机器人



服务机器人



02

# 图灵测试



# 如何评估一个机器是否具有智能呢？

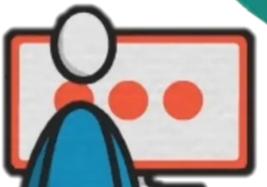
---



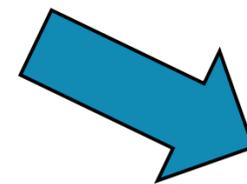
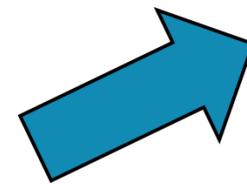
# 如何评估一个机器是否具有智能呢?



“今天真是个好天气啊”  
这句话除了字面意思，  
还测试能表达什么含义?



随意提问



说话者可能只是以此作为开场白，来开启与对方的对话，并没有特别针对天气的感慨。

就是对于今天天气的描述，没有其他意思

图灵测试

 未通过



## 2.1 图灵测试

图灵测试是由英国数学家、计算机科学家艾伦·图灵于1950年提出的一种衡量机器是否具备智能的实验方法。其核心思想是如果一台机器能够通过对话使人类无法区分它是机器还是人类，那么这台机器就可以被认为具有智能。



### 艾伦·图灵 (Alan Turing, 1912年6月23日~1954年6月7日)

英国著名的数学家、逻辑学家，被誉为“计算机科学之父”和“人工智能之父”。1936年（24岁），图灵提出了“图灵机”的概念，奠定了现代计算机科学的基础。第二次世界大战期间，他设计了名为“Bombe”的机电设备，加速了对恩尼格玛密码的破译。1950年（38岁），图灵在论文《计算机器与智能》中提出了著名的“图灵测试”，作为判断机器是否具有智能的标准。

## ■ 自然语言处理：

- 机器必须能够理解复杂的自然语言，它不仅需要理解表面上的文字（包括语法、词汇、语义和上下文），还要能够推断出潜在的含义、隐喻、讽刺等复杂的语言现象。
- 同时，还需要机器能够生成流畅、连贯且符合语境的文字回应，表现出类似人类的表达能力。



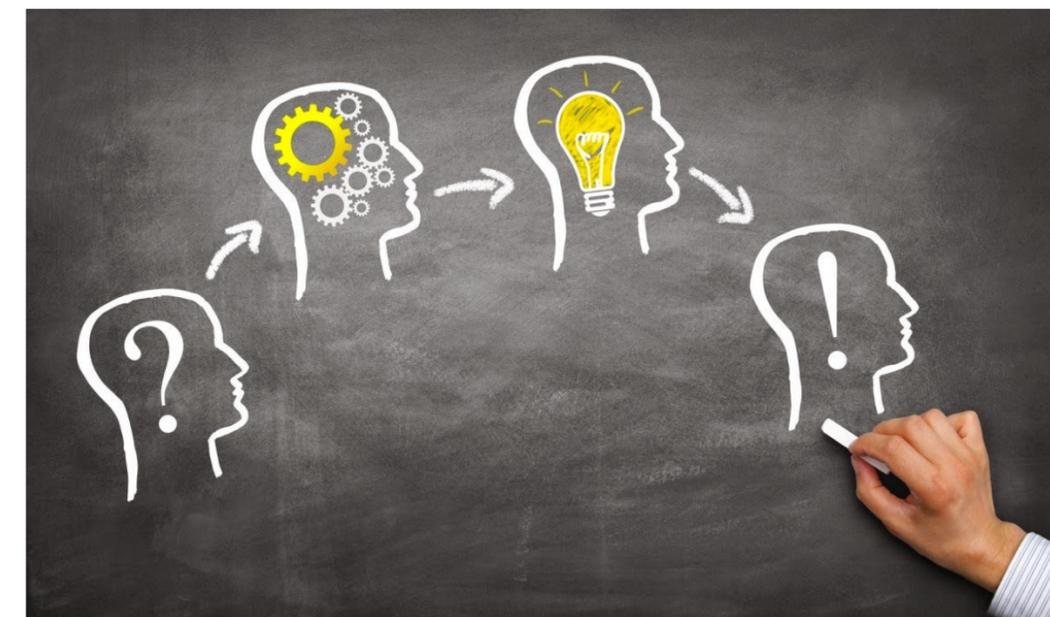
## ■ 知识表示能力：

- 机器需要将获取到的知识以结构化的形式进行表示，如通过语义网络、框架、描述逻辑等。
- 这种结构化表示有助于机器在对话过程中快速检索和运用相关知识。



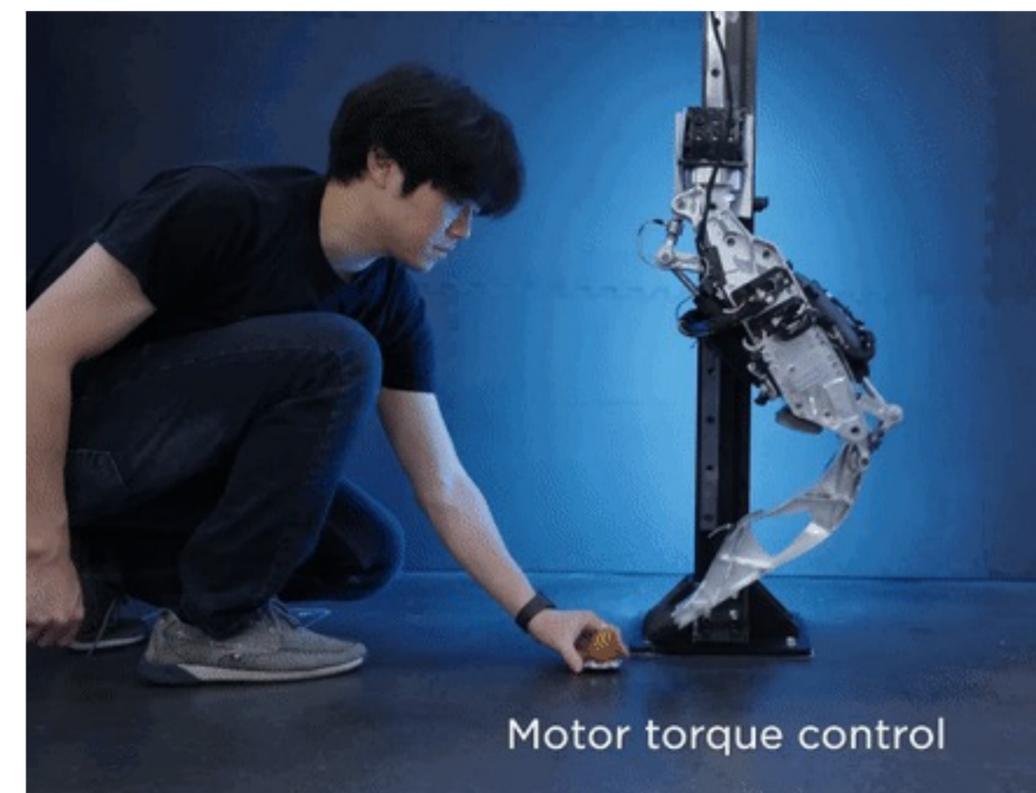
## ■ 推理能力：

- 机器需要具备广泛的知识基础，并能够根据所掌握的知识进行推理。
- 例如，机器应能够理解并回答关于常识、科学、文化、历史等各类问题，甚至在一些未知问题上表现出推理和假设的能力。



## ■ 学习与适应能力：

- 机器应能够从与人类的交互中学习，并适应不同的交流风格和偏好。
- 通过不断调整自己的表现，更好地迎合对话者的期望，从而更好地融入对话环境。



- 图灵测试要求机器具备自然语言处理、知识表示、推理与学习等多方面的能力，以便在对话中模拟人类行为，达到无法区分的程度。**机器一旦通过了图灵测试，则可以被认为是能够表现出与人类相当的智慧，具备一定的“智能”。**机器需要具备广泛的知识基础，并能够根据所掌握的知识进行推理。
- 尽管图灵测试在人工智能领域具有重要意义，但它并非评估人工智能的唯一标准。图灵测试**主要关注机器在语言交流中的表现，忽略了智能的其他关键维度**，如感知、情感理解和常识推理等能力。此外，图灵测试的评判标准较为主观，**缺乏统一的衡量尺度**，可能导致评估结果的不一致。



## ■ 完全图灵测试的提出

- 完全图灵测试是对普通图灵测试的拓展。普通的图灵测试主要关注文字对话的能力，避免评审员与被测试机器发生物理上的互动。
- 然而，一些人工智能系统可能涉及人机在物理层面上的交互，因此提出了完全图灵测试的概念。

## ■ 完全图灵测试的意义

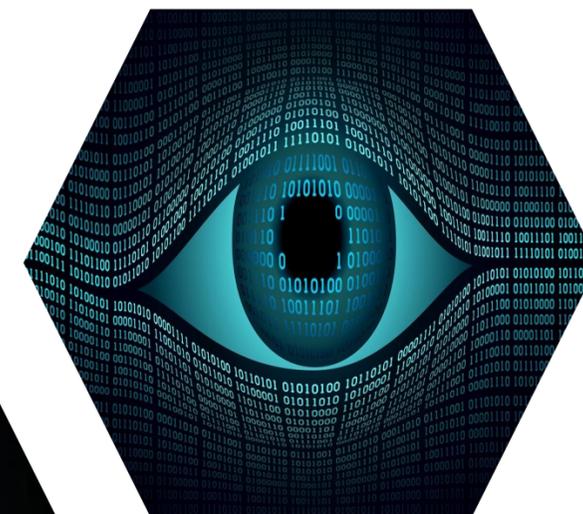
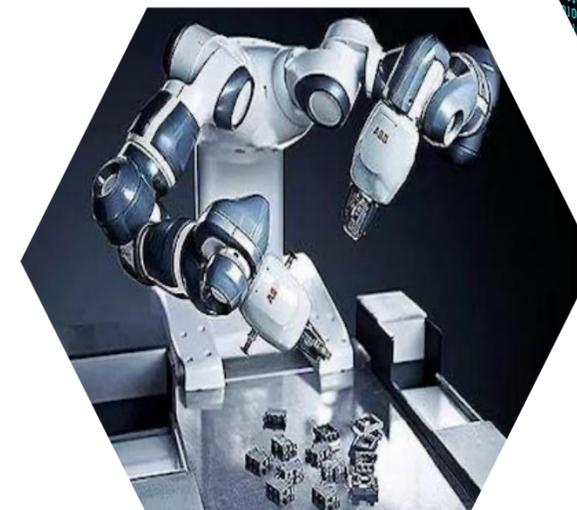
- 意义在于**更全面地评估机器的智能水平**，它不仅考察机器在文字对话中的表现，**还考察机器在物理交互中的能力**。



为了通过完全图灵测试，机器除了需要在文字对话中表现出与人类相似的智能水平外，还需要具备以下两项额外能力：

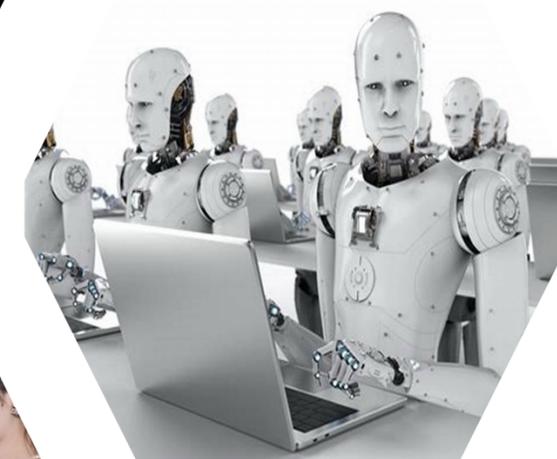
## ■ 感知能力

- 机器需要具备感知能力，以便与人类进行物理层面的交互。这通常需要借助计算机视觉技术来实现。



## ■ 操纵物体的能力

- 机器需要具备操纵物体的能力，以便在执行任务时与人类进行互动。这通常需要借助机器人学技术来实现。



2.2

# 延伸：ChatGPT-4通过图灵测试了吗？





争议



美国加州大学圣地亚哥分校的认知科学家在预印本论文中的实验结果显示：招募 500 名参与者与 4 个不同的智能体进行对话，这 4 个智能体包括 3 个 AI 和 1 个人类。其中 **ChatGPT-4 表现较好，参与者有 54% 的时间将其视为人**，相比之下，过时的 ELIZA 能成功欺骗参与者的机率只有 22%，ChatGPT-3.5 则在约一半情况下未被发现，而人类对话者被正确辨认的比例为 67%。

**争议**

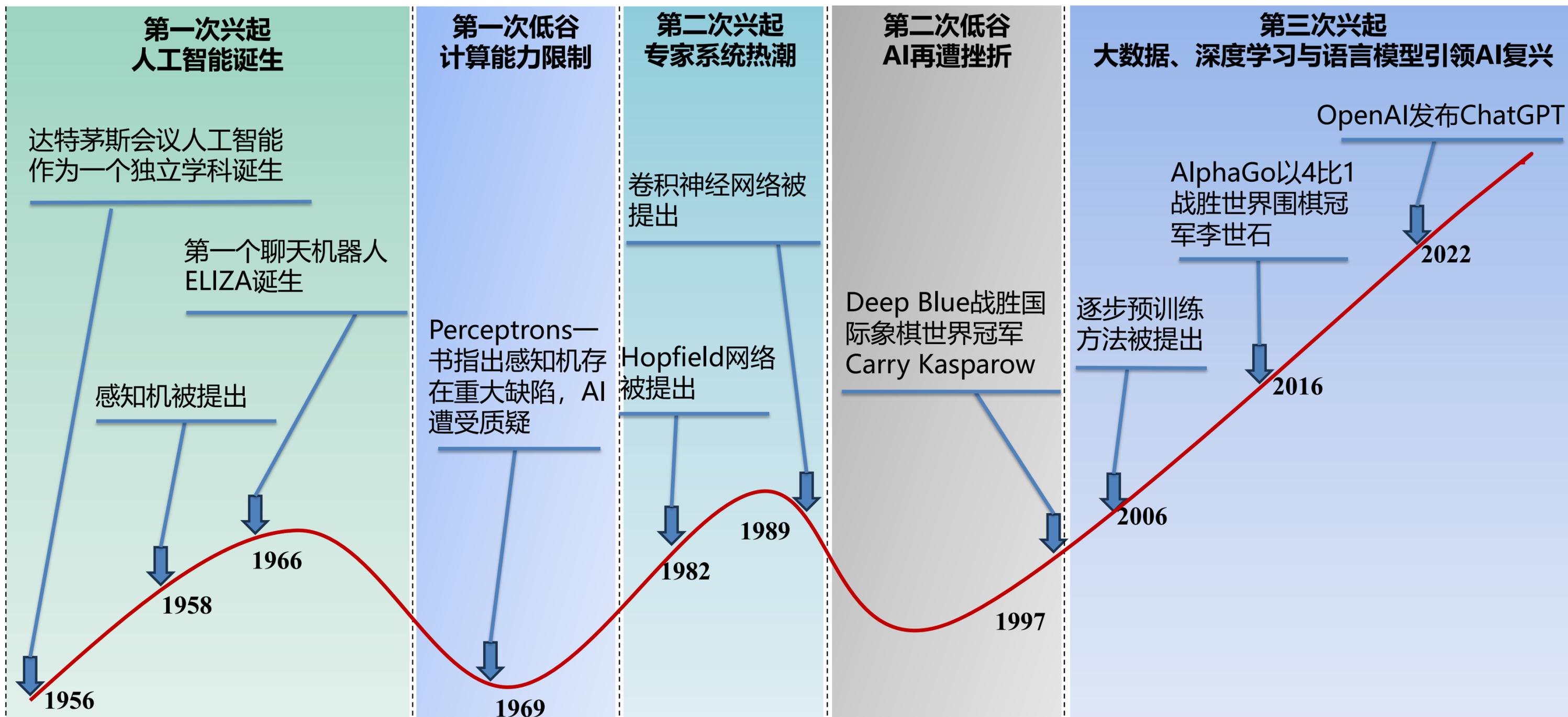
- **测试标准的主观性**：图灵测试的判定依赖于提问者的主观判断，不同的人可能有不同的判断标准和敏感度，会影响测试结果的准确性和可靠性。
- **测试环境的局限性**：测试通常是在特定的环境和条件下进行的，在更长时间、更广泛深入的交流中，ChatGPT-4 可能会暴露出更多机器的特征。
- **智能表现的片面性**：图灵测试仅通过语言交流来评估智能，忽略了人类智能的其他表现形式，如感知、行动等，ChatGPT-4可能只是在语言方面表现出智能。

03

# 人工智能发展史

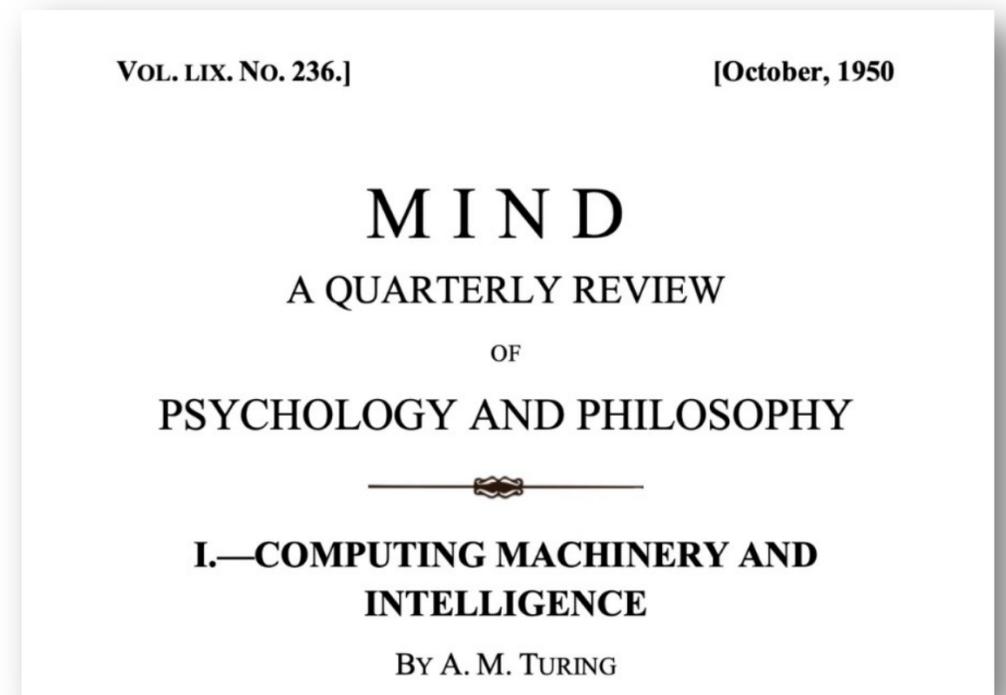


## 人工智能的发展之路并非一帆风顺，而是经历了跌宕起伏的“三起两落”



## 3.1 第一次兴起——20世纪50年代末到60年代初

- 1950年，英国数学家艾伦·图灵（Alan Turing）在他的论文《计算机器与智能》（《Computing Machinery and Intelligence》）中**首次提出了人工智能的概念**，并给出了著名的**图灵测试**，探讨了“机器能思考吗？”这一问题。
- 不同于哲学思辨，图灵测试提供了一个可实验验证的智能定义，推动了AI研究从理论走向实践。图灵测试至今仍是讨论“机器智能”的起点，影响AI伦理、认知科学等领域。





1956年美国达特茅斯会议合影

时间拨回到**1956年**的夏天，在美国汉诺斯小镇宁静的**达特茅斯学院**，约翰·麦卡锡（Lisp语言发明者、图灵奖得主）、马文·明斯基（人工智能与认知学专家）、克劳德·香农（信息论的创始人）、艾伦·纽厄尔（计算机科学家）、赫伯特·西蒙（诺贝尔经济学奖得主）、所罗门列夫（归纳推理机发明者）、纳撒尼尔·罗切斯特（电气工程师先驱）、塞弗里奇（维纳学生，模式识别的奠基人）等聚在一起进行了一场头脑风暴讨论会。**计算机科学家约翰·麦卡锡说服与会者接受“人工智能”一词作为本领域的名称。**后来，这次会议也被看作是**人工智能正式诞生的标志。**

## 人工智能(Artificial Intelligence)

使一部机器像人一样进行**感知、认知、决策、执行**的人工程序或系统



艾伦·图灵，英国计算机科学家、数学家、逻辑学家、密码分析学家、理论生物学家，“计算机科学之父”、“人工智能之父”，英国皇家学会院士



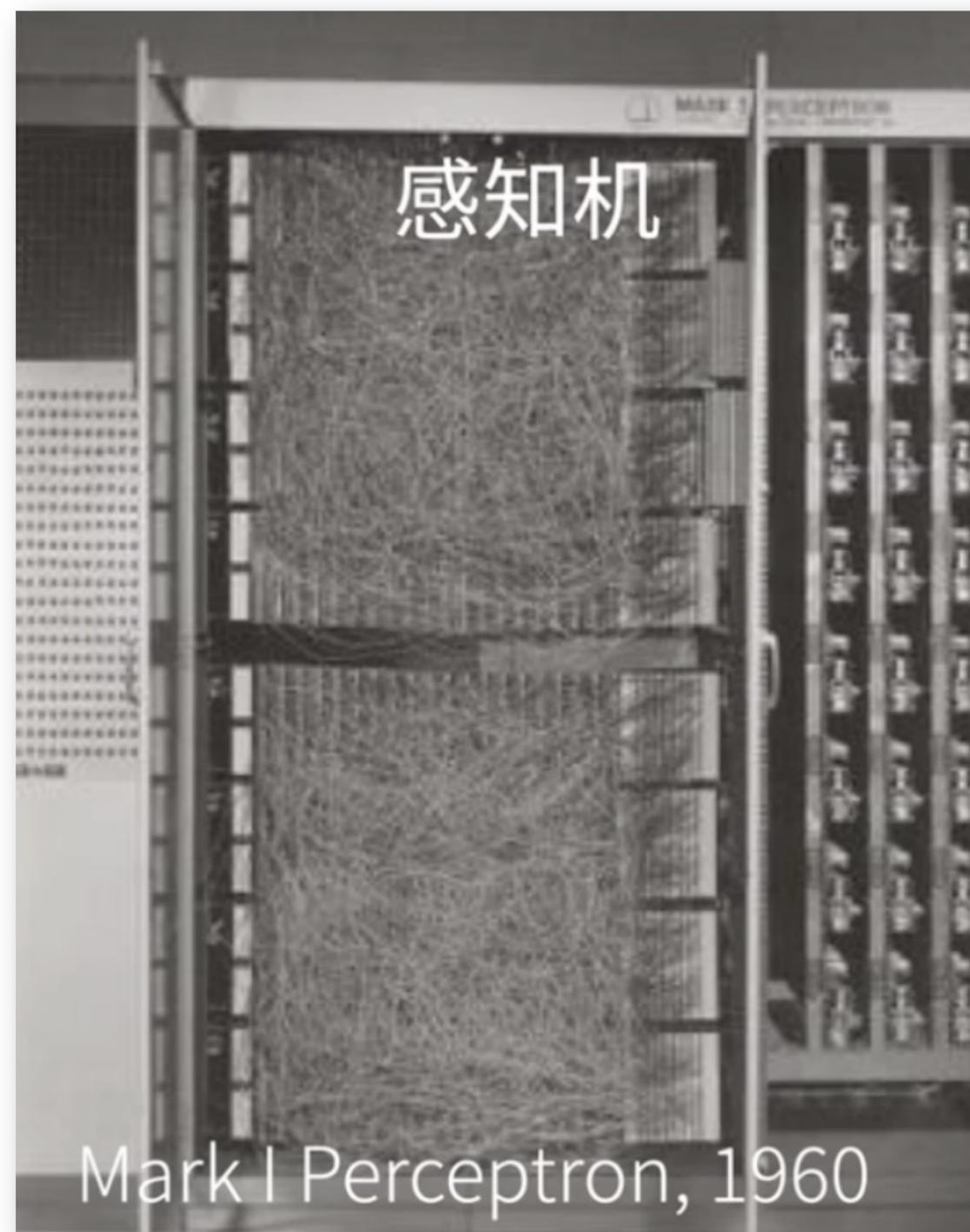
约翰·麦卡锡，LISP编程语言的创始人之一，1971年计算机界的最高奖项图灵奖，并于1991年获得美国国家科学奖，斯坦福大学人工智能实验室的主任，被成为“人工智能之父”

图灵的“人工智能”概念是指一种具备独立思维能力的全新生命体或思维形式的理论构想，而麦卡锡所定义的“人工智能”则是基于图灵的理论框架发展而来的一门具体科学。值得注意的是，“人工智能”这一术语正是由麦卡锡首次提出。

## 3.1 第一次兴起——20世纪50年代末到60年代初

### ■ 神经网络模型的实现

- 1958年,《心理评论》发表了康奈尔大学研究员弗兰克·罗森布拉特 (Frank Rosenblatt) 的“**感知机 (Perceptron)** : 大脑中信息存储和组织的概率模型”并在一台IBM-704计算机上进行了模拟实现,这是第一个可以自动学习权重的神经元模型,为后来的神经网络研究奠定了基础。感知机的成功也进一步激发了人们对人工智能的广泛兴趣,许多研究人员和机构开始投入到这一领域。



## 3.1 第一次兴起——20世纪50年代末到60年代初

### ■ 问答机器人的诞生

- 1966年，麻省理工学院科学家约瑟夫·维森鲍姆（Joseph Weizenbaum）开发出**第一个心理治疗机器人ELIZA**，旨在模拟心理治疗师与患者的对话。它展示了人工智能在自然语言处理方面的潜力。
- 在这一时期，人工智能研究热潮涌动，科学家们豪情满怀，基于当时人工智能在各个领域的初步发展，甚至有人大胆预测，20年内就能造出超越人类的AI。

```
Welcome to
                EEEEEEE LL      IIII  ZZZZZZ  AAAAA
                EE      LL      II     ZZ     AA  AA
                EEEEE  LL      II     ZZZ    AAAAAA
                EE      LL      II     ZZ     AA  AA
                EEEEE  LLLLLL  IIII  ZZZZZZ  AA  AA

Eliza is a mock Rogerian psychotherapist.
The original program was described by Joseph Weizenbaum in 1966.
This implementation by Norbert Landsteiner 2005.

ELIZA: Is something troubling you ?
YOU:   Men are all alike.
ELIZA: What is the connection, do you suppose ?
YOU:   They're always bugging us about something or other.
ELIZA: Can you think of a specific example ?
YOU:   Well, my boyfriend made me come here.
ELIZA: Is it important to you that your boyfriend made you come here ?
YOU:   He says I'm depressed much of the time.
ELIZA: I am sorry to hear that you are depressed.
YOU:   It's true. I am unhappy.
ELIZA: Can you explain what made you unhappy ?
YOU:
```

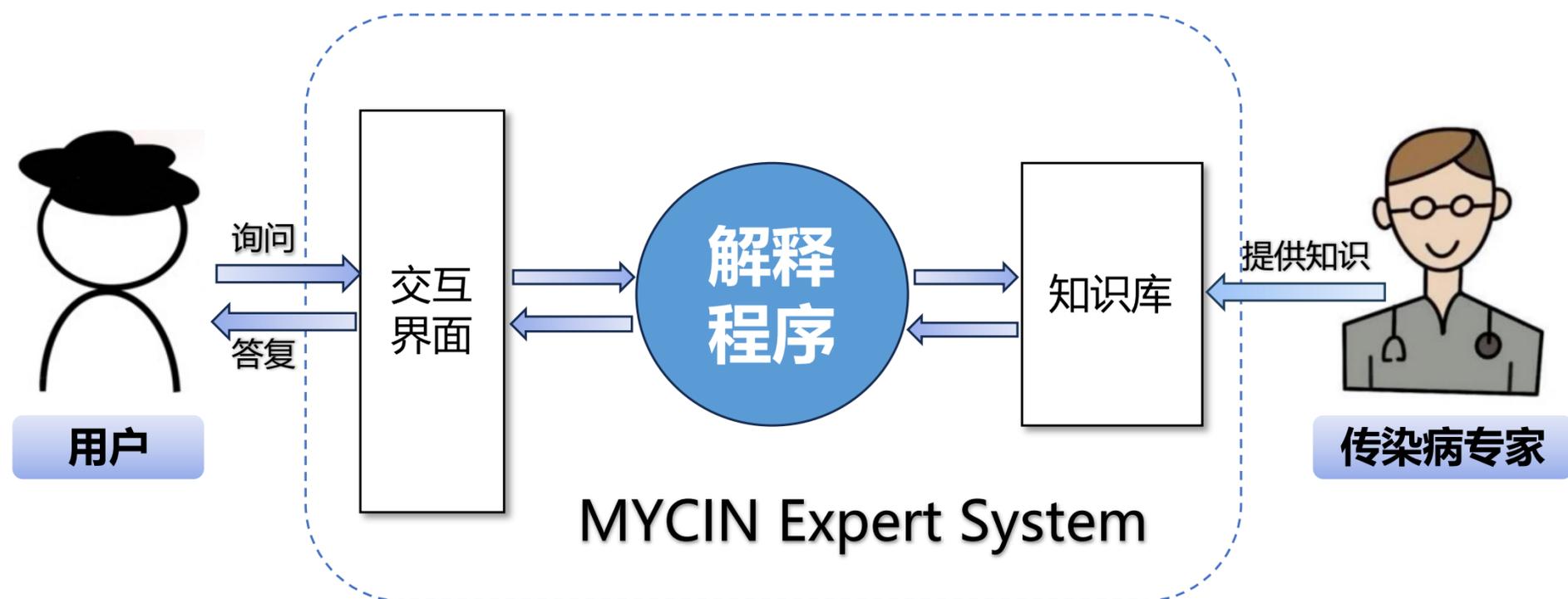
ELIZA问答机器人的对话示例



1969年，“符号主义”代表人物、图灵奖获得者马文·明斯基等人在他们的著作《Perceptrons: An Introduction to Computational Geometry》中指出了**感知机无法解决非线性可分问题的缺陷**。感知机的局限性使得许多研究人员开始质疑神经网络的有效性，导致对这一领域的研究兴趣大幅下降。



1973年，法国著名数学家詹姆斯·莱特希尔（James Lighthill）向英国政府提交了一份报告，严厉批评了当时的AI研究，认为其未能实现预期目标。这份报告被称为“莱特希尔报告”，它直接导致了**英国政府进一步大幅削减AI研究经费**，其他国家也纷纷开始效仿，对AI研究产生了深远的影响。



美国斯坦福大学研究团队研发的**MYCIN系统**，它是早期模拟决策系统和专家系统的代表之一，用于诊断血液传染病并推荐抗生素治疗。

得益于计算能力的提升，许多具有复杂规则的专家系统开始在计算机上得到应用。这些系统能够处理更为复杂的逻辑关系和推理过程，从而在实际应用中展现出更高的智能水平和实用性，进一步推动了人工智能技术的兴起和发展，成为推动此次兴起的核心关键力量。

## ■ 人工神经网络

- **Hopfield网络**: 1982年约翰·霍普菲尔德 (John J. Hopfield) 发明了Hopfield网络, 成为深度学习的早期基础。
- **玻尔兹曼机**: 1983年由辛顿 (Geoffrey E. Hinton) 等人提出的玻尔兹曼机, 能进行无监督学习, 对数据进行特征提取, 也被称为随机Hopfield网络。
- **反向传播算法**: 1974年保罗·韦尔博斯 (Paul Werbos) 提出反向传播算法 (BP算法)。1986年大卫·莱姆哈特 (David E. Rumelhart) 等人在《Nature》上发表的论文极大地推动了神经网络的研究。1993年, BP算法在国际模式识别竞赛中获奖, 进一步证明了其有效性, 为神经网络的广泛应用打下了基础。
- **多层感知器与BP算法的结合**: 开启了神经网络研究的新一轮高潮。
- **卷积神经网络**: 1989年, 杨立昆 (Yann LeCun) 和约书亚·本吉奥 (Yoshua Bengio) 等人提出了一个卷积神经网络 (Convolutional Neural Network, CNN), 即著名的LeNet-5, 并使用BP算法完成网络训练。



1982年，美国物理学家、2024年诺贝尔物理学奖得主约翰·霍普菲尔德（John J. Hopfield）发明了Hopfield网络，它是一种单层、全互连的反馈神经网络，用于联想记忆，该网络对理解复杂系统和作为深度学习早期基础有重要贡献。



1983年，图灵奖得主、2024年诺贝尔物理学奖得主杰弗里·辛顿（Geoffrey E. Hinton）等人发明了玻尔兹曼机（Boltzmann Machines），也被称为随机Hopfield网络，这是第一个能够学习不属于输入或输出的神经元内部表征的神经网络。它本质是一种无监督模型，用于对输入数据进行重构以提取数据特征做预测分析。

# 名堂



1989年，图灵奖得主杨立昆（Yann LeCun）和约书亚·本吉奥（Yoshua Bengio）等人提出了一个卷积神经网络（Convolutional Neural Network, CNN），即著名的LeNet-5，并使用BP算法完成网络训练。LeNet-5也成功应用于美国邮局的手写字符识别系统中。

## ■ 人工智能第二次低谷的原因

### 专家系统缺点

- 随着时间推移，常识性错误、更新和维护成本高昂等问题逐渐暴露，限制了其广泛应用。

### 硬件支持缺乏

- 通用计算机的性能提升使得LISP机器市场需求急剧下降，导致LISP机器制造商退出市场。

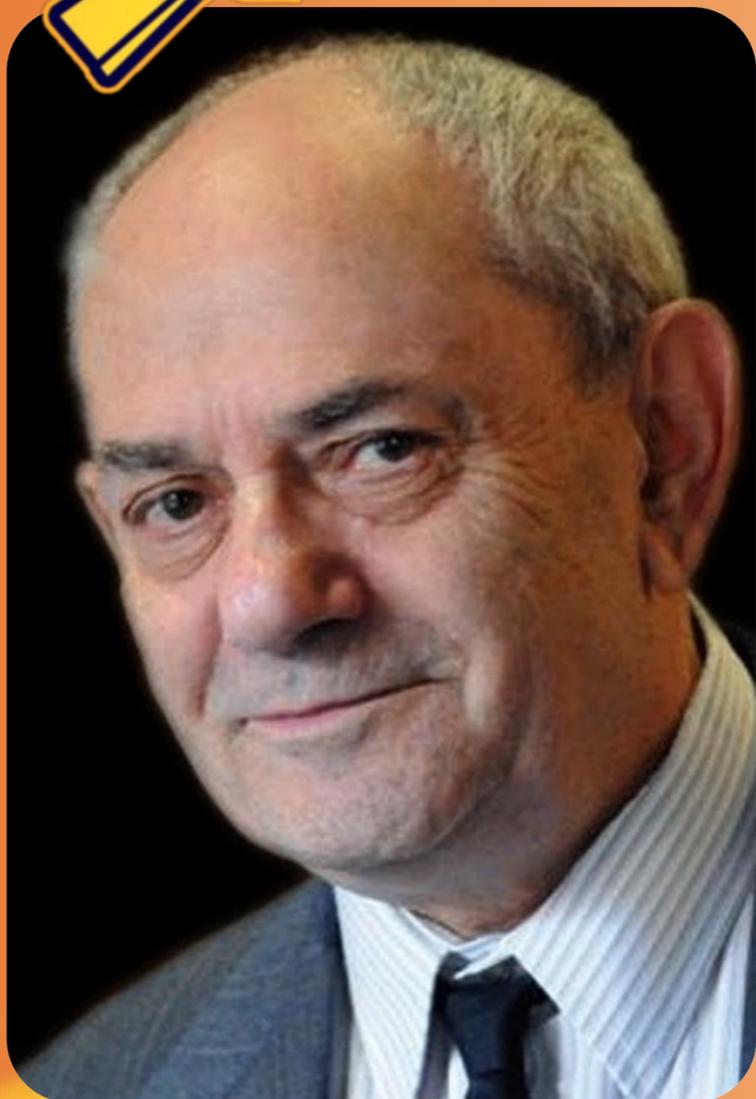
### 机器学习兴起

- 支持向量机和决策树等传统机器学习算法在这一时期大放异彩，逐渐取代了神经网络在许多应用中的地位。

### 计算资源受限

- 尽管计算能力有所提升，但仍不足以支持大规模神经网络的训练和应用。

# 名堂



统计学习理论之父、苏联犹太裔美国统计学家、美国国家工程院院士弗拉基米尔·纳乌莫维奇·瓦普尼克 (Vladimir Naumovich Vapnik) 于1995年提出了支持向量机 (Support Vector Machine, SVM), 其学习策略是间隔最大化, 这使得它在处理高维数据时具有较好的泛化能力。此外, SVM的数学基础扎实, 能够通过求解凸二次规划问题来找到最优解。这些优势使得SVM在许多实际应用中表现出色, 逐渐取代了当时的神经网络。2006年他入选美国国家工程院, 还获得2005年伽伯奖、2008年帕里斯·卡内拉基斯奖、2010年神经网络先锋奖、2012年IEEE弗兰克·罗森布拉特奖、2012年本杰明·富兰克林计算机和认知科学奖、2017年IEEE约翰·冯·诺伊曼奖章、2018年伦敦大学的科尔莫戈罗夫奖章、2019年BBVA基金会知识前沿奖等。

## 3.1 第三次兴起——2006年以后

### ■ 人工智能第三次兴起的核心驱动力

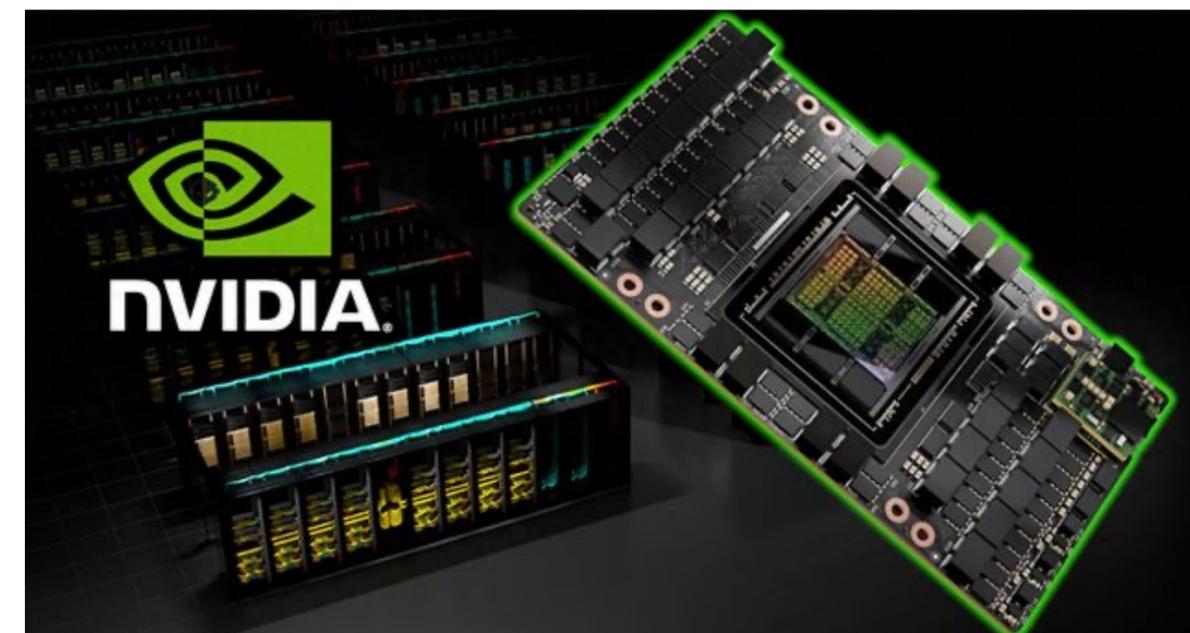
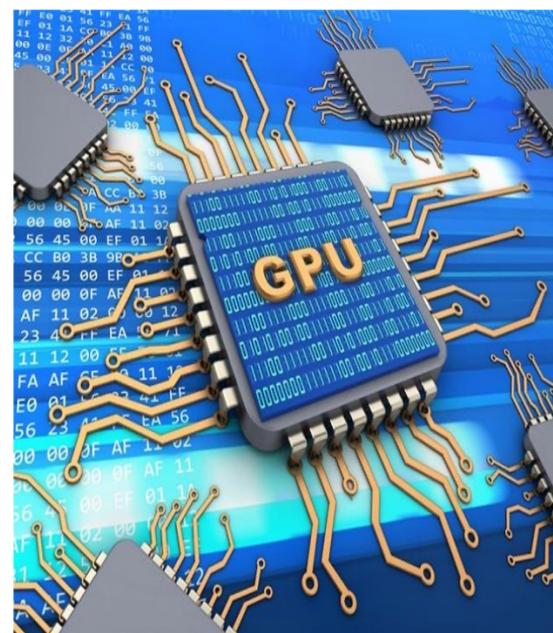
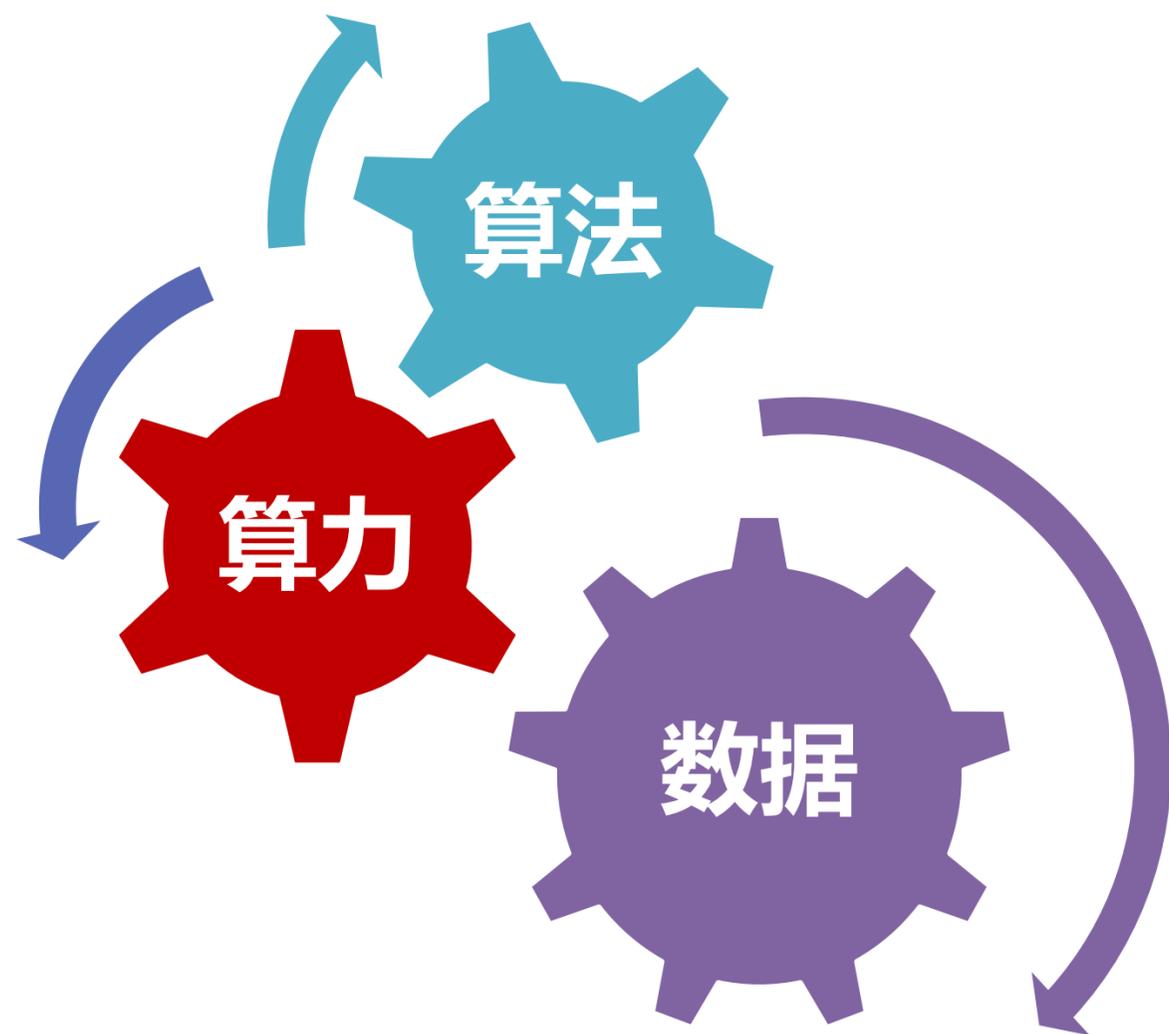


## ■ 人工智能第三次兴起的核心驱动力



美国斯坦福大学的李飞飞教授等人于2009年创建发布的 **ImageNet数据集** 包含了超过1400万幅图片，涵盖了2万多个类别，可支撑图像分类、定位、检测等方向研究。

## ■ 人工智能第三次兴起的核心驱动力

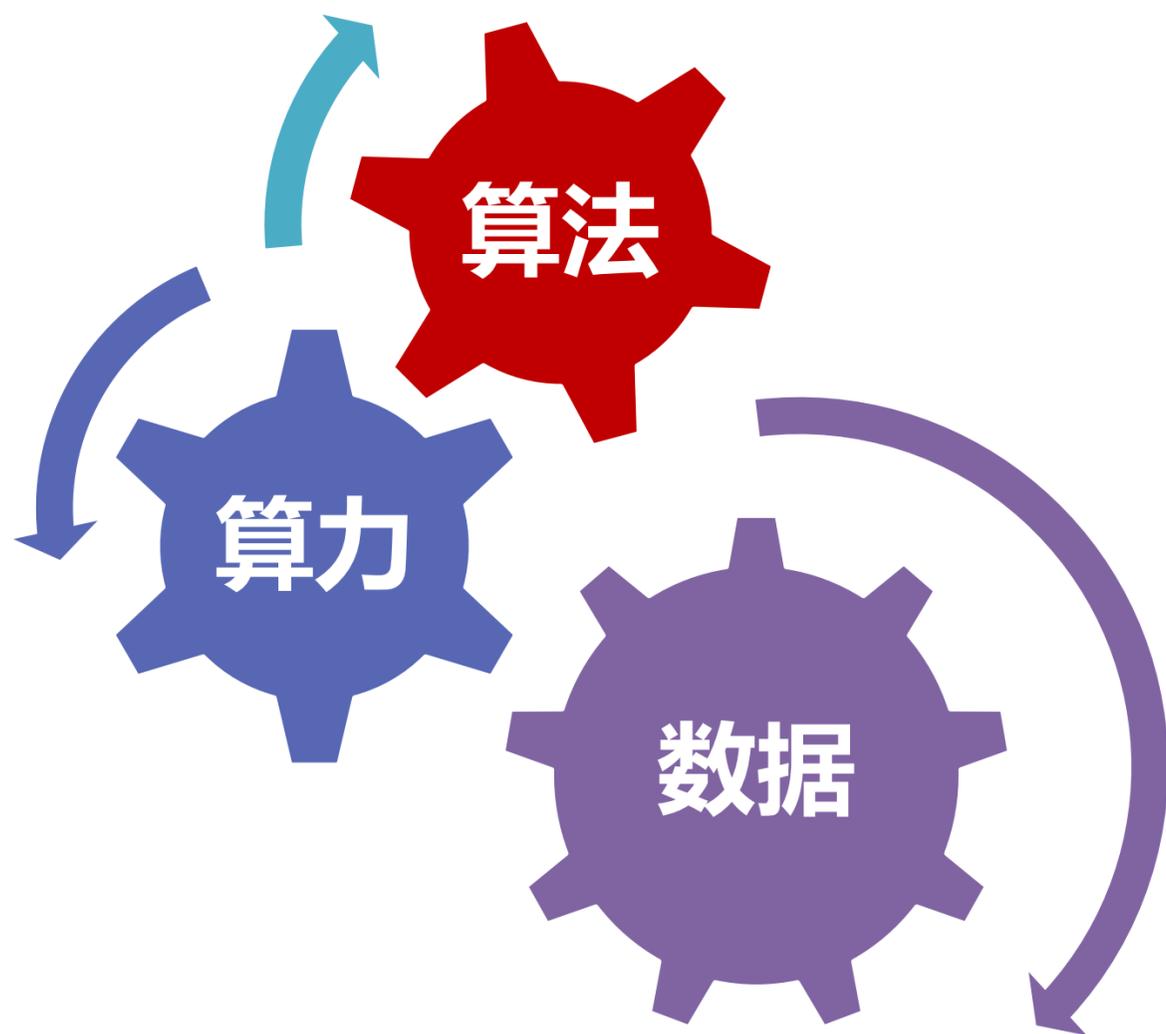


**GPU (Graph Processing Unit)**、FPGA (Field Programmable Gate Array)、ASIC (Application Specific Integrated Circuit) 以及神经拟态芯片等高性能计算硬件的发展，极大增强了计算机的运算能力，为处理复杂的人工智能算法提供了强有力的硬件支持。

## ■ 人工智能第三次兴起的核心驱动力

以**深度学习**为代表的关键技术不断突破

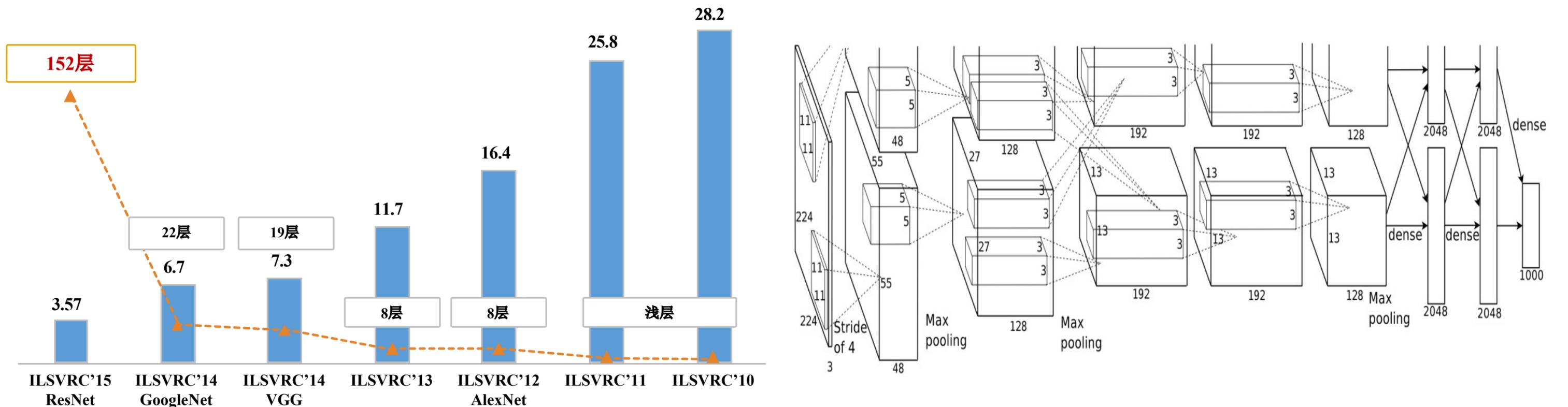
- CNN 网络： AlexNet、 VGG、 GoogLeNet、 ResNet
- 注意力网络： 通道注意力、 空间注意力、 自注意力
- 序列到序列网络： LSTM、 GRU
- 生成网络： GAN、 CycleGAN、 Diffusion
- 自编码器： Auto-encoder
- Transformer网络： ViT、 DETR
- .....



# 3.1 第三次兴起——2006年以后

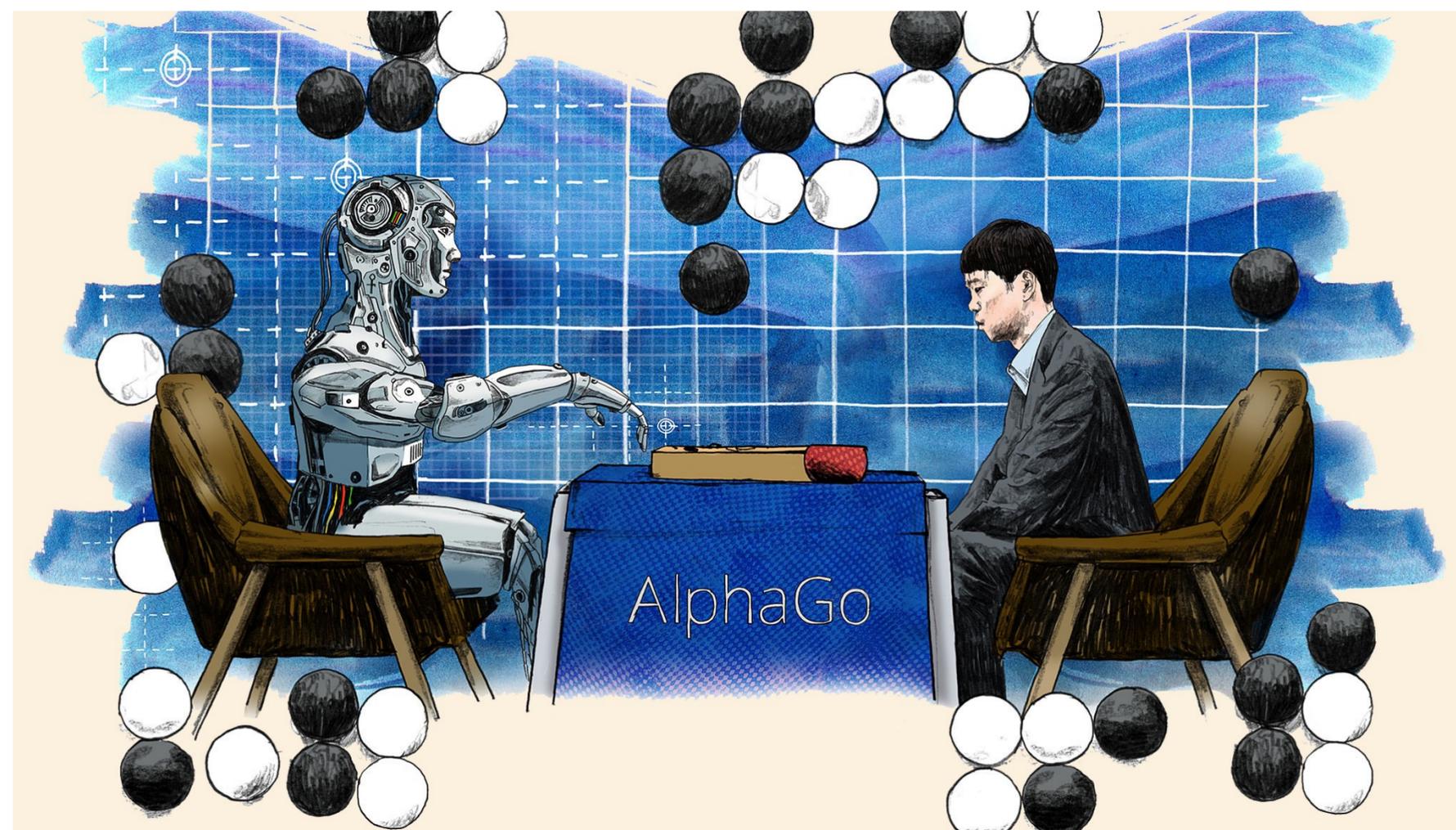
## ■代表性事件回顾——伴随成长的ImageNet竞赛

2010年首次开赛，2012年AlexNet横空出世，在ImageNet挑战赛上大获全胜，将Top-5分类错误率降低至16.4%，相比于2011年冠军方案（统计机器学习方法）性能提高约10%，这是史上第一次有模型在 ImageNet 数据集获得如此出色的表现，引爆了神经网络的研究热情。



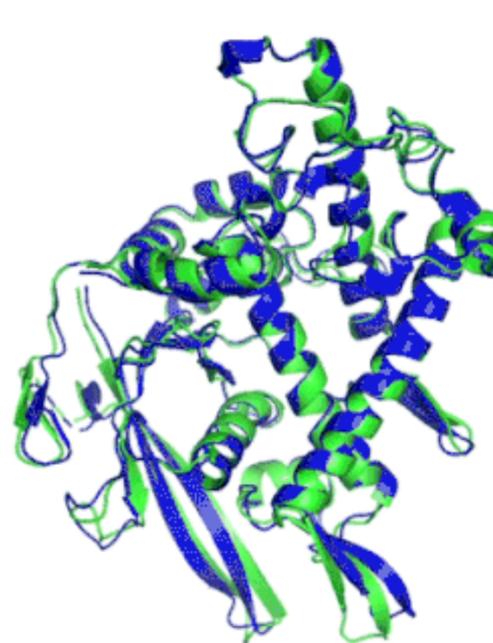
## ■代表性事件回顾——一骑绝尘的AlphaGo

2016年，由谷歌（Google）旗下DeepMind公司戴密斯·哈萨比斯（Demis Hassabis）领衔的团队开发的AlphaGo在围棋界掀起了惊涛骇浪，以4:1的战绩击败了围棋世界冠军、职业九段棋手李世石九段，这一壮举不仅彰显了深度学习与强化学习技术的非凡实力，更在全球范围内引发了关于人工智能潜力与未来的热烈讨论。

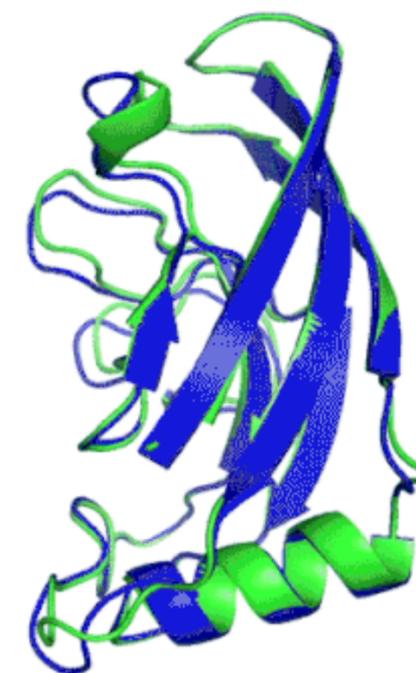


## ■代表性事件回顾——赢回诺奖的AlphaFold

2020年，谷歌DeepMind团队的AlphaFold2人工智能系统实现了蛋白质结构预测领域的重大突破，它犹如一座里程碑，标志着该领域迈入了一个全新的时代。AlphaFold2以卓越的表现脱颖而出，精准地预测了蛋白质的三维结构，其精确度之高，几乎可与冷冻电子显微镜、核磁共振及X射线晶体学等尖端实验技术相抗衡，展现了人工智能在生命科学领域的非凡实力。其核心发明者Demis Hassabis也荣获2024诺贝尔化学奖。



T1037 / 6vr4  
90.7 GDT  
(RNA polymerase domain)



T1049 / 6y4f  
93.3 GDT  
(adhesin tip)

● Experimental result  
● Computational prediction

## 3.1 第三次兴起——2006年以后

### ■代表性事件回顾 —— 大模型时代的开启者ChatGPT

2022年，ChatGPT（Chat Generative Pre-trained Transformer）的横空出世标志着大模型时代的来临。以ChatGPT为代表的通用大语言模型，是人工智能领域的一项重大突破，它们通过海量数据的深度学习与复杂神经网络的构建，展现出了前所未有的自然语言处理能力和智能交互水平。ChatGPT及其同类大模型的崛起，标志着人工智能正逐步从特定任务的执行者转变为能够理解、学习与创造的智能伙伴，引领着科技行业向更加智能化、人性化的方向迈进。

ZH vue校验手机号是否合法

如果你想在 Vue.js 中校验手机号是否合法，可以使用正则表达式来进行校验。下面是一个例子：

```
// 定义手机号正则表达式
const phoneReg = /^1[3-9][0-9]{9}$/

// 在 Vue 组件的 methods 中定义一个方法用于校验手机号是否合法
checkPhone (phone) {
  if (!phoneReg.test(phone)) {
    // 如果手机号不合法，可以在这里提示用户
    alert('请输入正确的手机号')
  }
}
```

在使用时，可以在表单提交时调用这个方法

# 3.2

## 人工智能时代已悄然开启，拥抱还是拒绝？





04

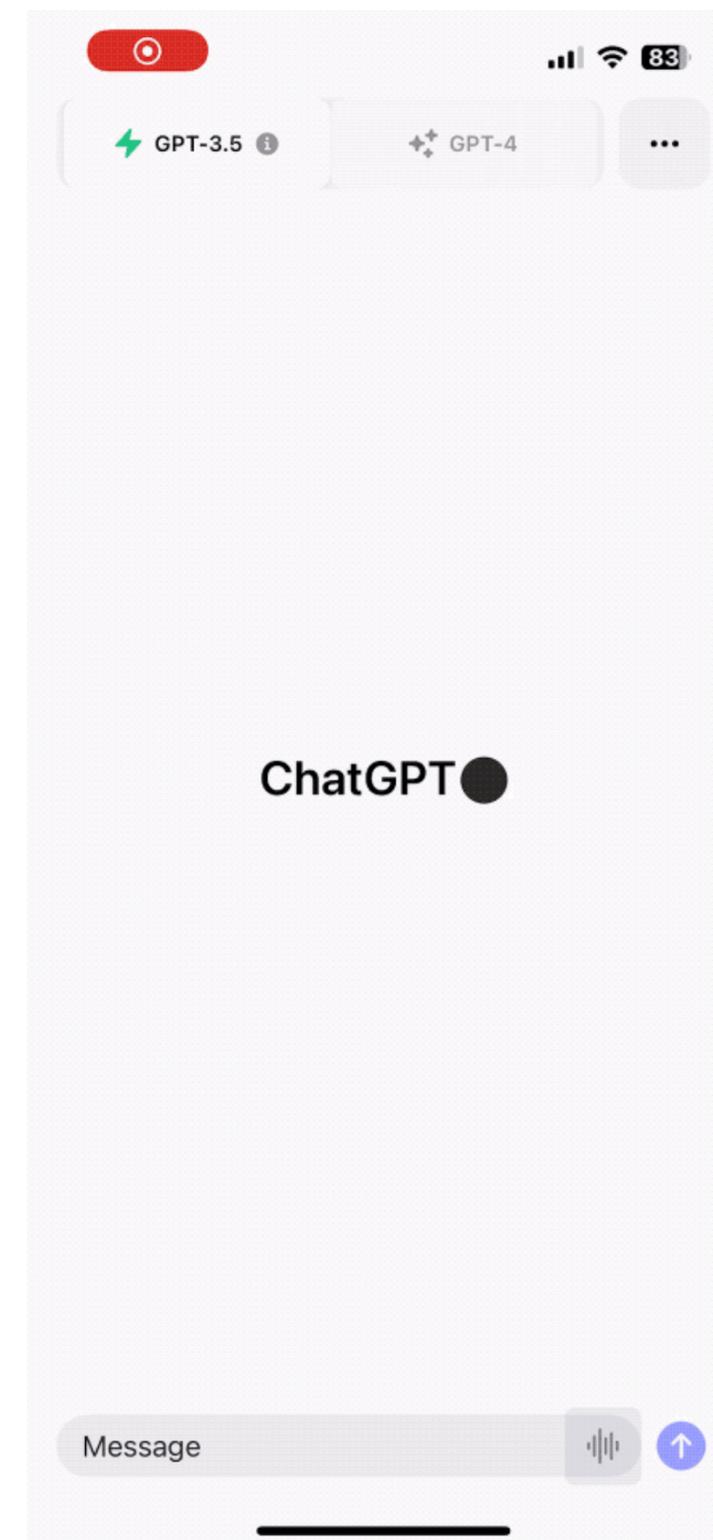
# ChatGPT——通用AI之路

## ■ 定义

- ChatGPT是由OpenAI开发的一种基于生成式预训练变换器（GPT）的大型语言模型。
- 它能够理解和生成自然语言文本，用于对话、问答、写作等任务。

## ■ 发展历程

- GPT-1（2018）：初步尝试，展示了语言模型的潜力。
- GPT-2（2019）：规模更大，生成能力显著提升。
- GPT-3（2020）：1750亿参数，性能接近人类水平。
- ChatGPT（2022）：引入人类反馈强化学习，能力大幅提升。
- GPT-4（2023）：更强大、更智能，支持多模态输入。



## 4.1 ChatGPT简介

### ■ 核心架构

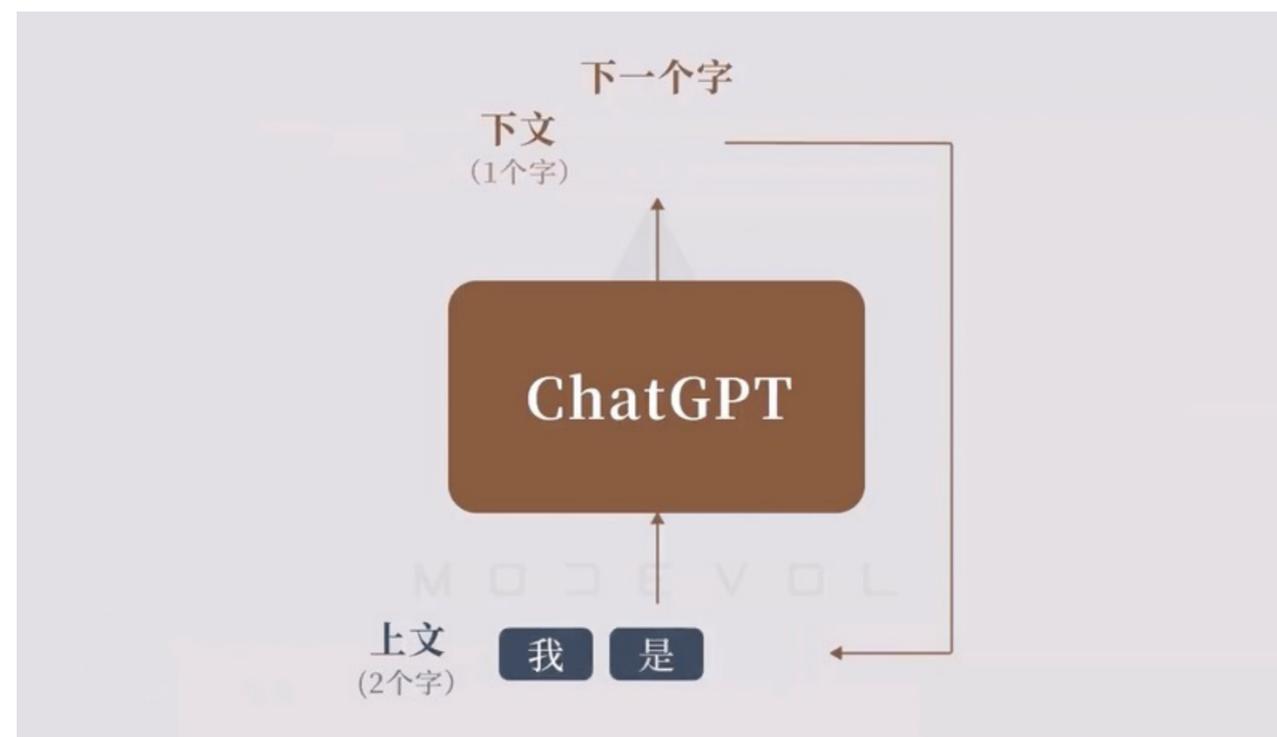
- 基于Transformer架构，利用自注意力机制处理文本。
- 通过预训练和微调学习语言规律。

### ■ 训练过程

- 预训练：在海量文本数据上学习语言模式。
- 微调：在特定任务数据上优化模型表现。

### ■ 生成文本的原理

- 根据输入文本预测下一个词，逐步生成完整回答。

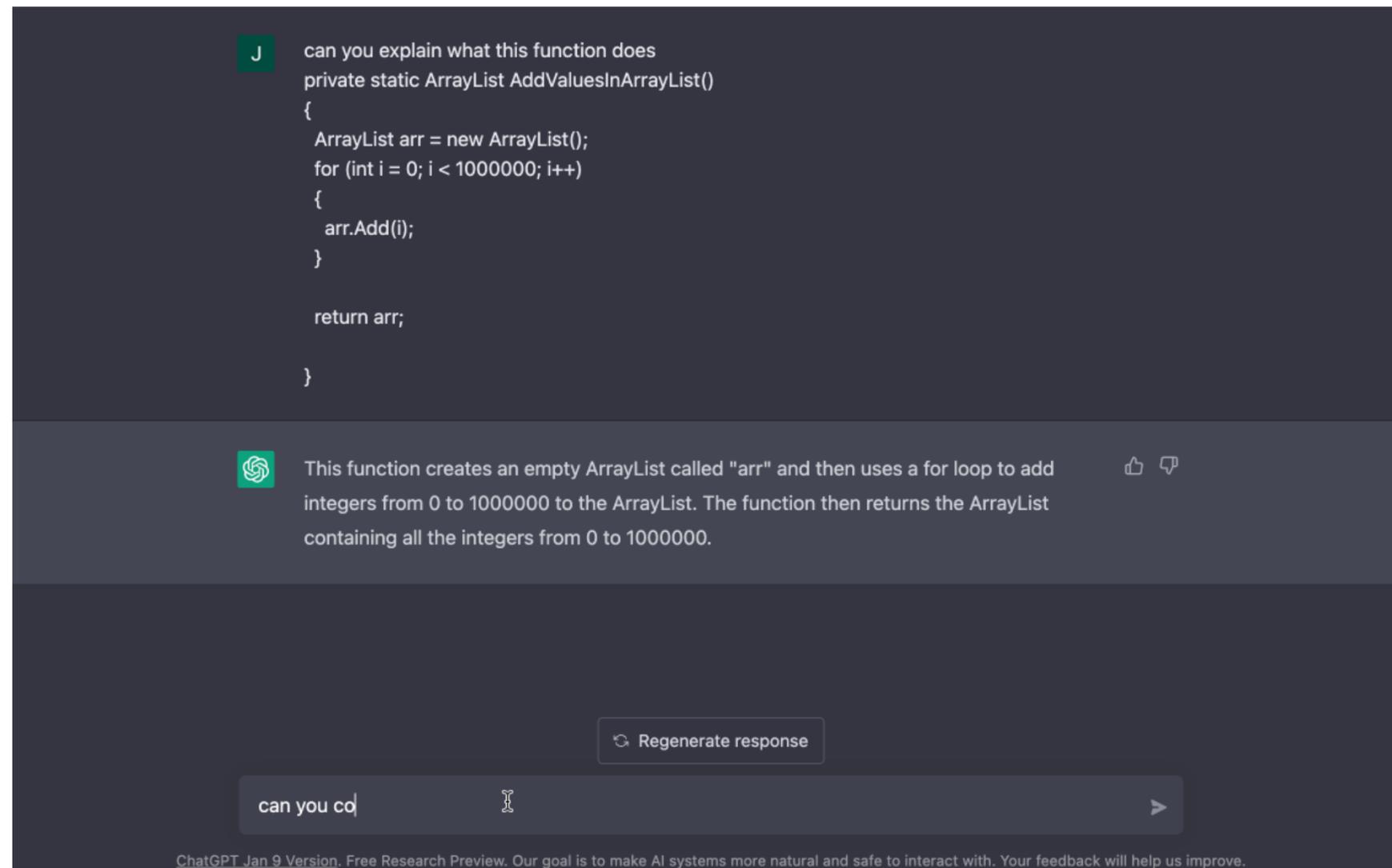


# 4.1 ChatGPT简介——功能

帮我写一个计算最大公约数的Python代码，并为每一行写上详细的注释。

➤ 多种语言任务：包括但不限于撰写论文、邮件、脚本、文案，进行翻译、编写代码。

以下是一个计算最大公约数的Python代码，并为每一行添加了详细的注释：



can you explain what this function does

```
private static ArrayList AddValuesInArrayList()
{
    ArrayList arr = new ArrayList();
    for (int i = 0; i < 1000000; i++)
    {
        arr.Add(i);
    }

    return arr;
}
```

This function creates an empty ArrayList called "arr" and then uses a for loop to add integers from 0 to 1000000 to the ArrayList. The function then returns the ArrayList containing all the integers from 0 to 1000000.

can you cd

ChatGPT Jan 9 Version. Free Research Preview. Our goal is to make AI systems more natural and safe to interact with. Your feedback will help us improve.

```
python
def gcd(a, b):
    # 定义一个函数 gcd，它接收两个整数 a 和 b 作为参数

    while b != 0:
        # 当 b 不等于 0 时，继续执行循环
        # 这表明我们还没有找到最大公约数

        a, b = b, a % b
        # 使用欧几里得算法计算最大公约数
        # 首先，将 a 的值赋给 b，将 b 的值赋给 a 除以 b 的余数
        # 这样做会逐步缩小 a 和 b 的值，直到 b 变为 0

    return a
    # 当 b 为 0 时，循环结束，此时 a 的值就是最大公约数，返回 a

# 示例使用
num1 = 48 # 定义第一个整数
num2 = 18 # 定义第二个整数
result = gcd(num1, num2) # 调用 gcd 函数计算 num1 和 num2 的最大公约数
print(f"{num1} 和 {num2} 的最大公约数是: {result}")
# 输出 num1 和 num2 的最大公约数
```

### 代码解释：

- **函数定义：** `gcd(a, b)` 函数接收两个整数作为输入参数，用于计算它们的最大公约数。
- **while 循环：** 在 `b` 不为0时，使用欧几里得算法不断迭代，将 `a` 和 `b` 的值更新为当前的 `b` 和 `a % b`，直到 `b` 变为0。
- **返回结果：** 循环结束时，返回 `a` 的值作为最大公约数。
- **示例使用：** 代码的最后部分展示了如何调用 `gcd` 函数并输出计算结果。

- 多模态功能：模型可以通过联合处理图像和文本数据进行训练，从而理解和生成与图像相关的文本描述。
- 语音功能和桌面版应用程序的推出，进一步扩展了其应用范围。

什么是人工智能通识课程

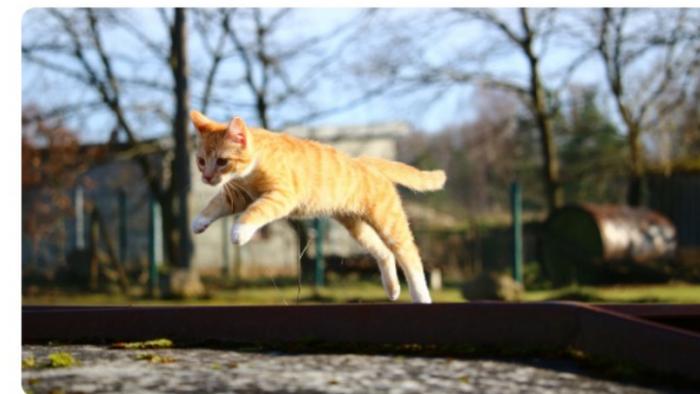


人工智能通识课程通常介绍AI的基础知识、技术背景、应用领域和潜在影响。它旨在帮助学生了解AI的基本概念、发展历程以及在实际生活中的应用，提供一个全面的AI概览。你对某个特定领域或主题特别感兴趣吗？

介绍一下人工智能的发展历史吧



人工智能的发展历史可以追溯到20世纪50年代的早期研究。最初，AI的研究集中在基本算法和逻辑推理上。60年代和70年代，AI进入了“寒冬期”，因为技术未能达到预期。80年代，专家系统的兴起带来了新的进展。21世纪初，随着计算能力和数据的激增，机器学习和深度学习成为主流，推动了AI的快速发展。最近几年，AI在各个领域的应用越来越广泛，显示出巨大的潜力。



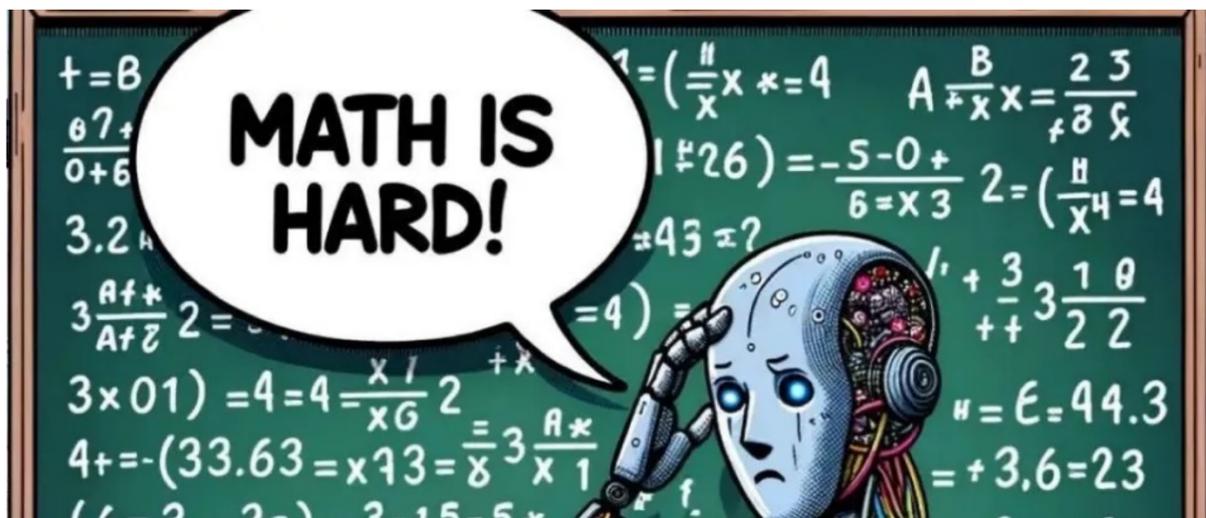
请帮我总结这张图像的内容



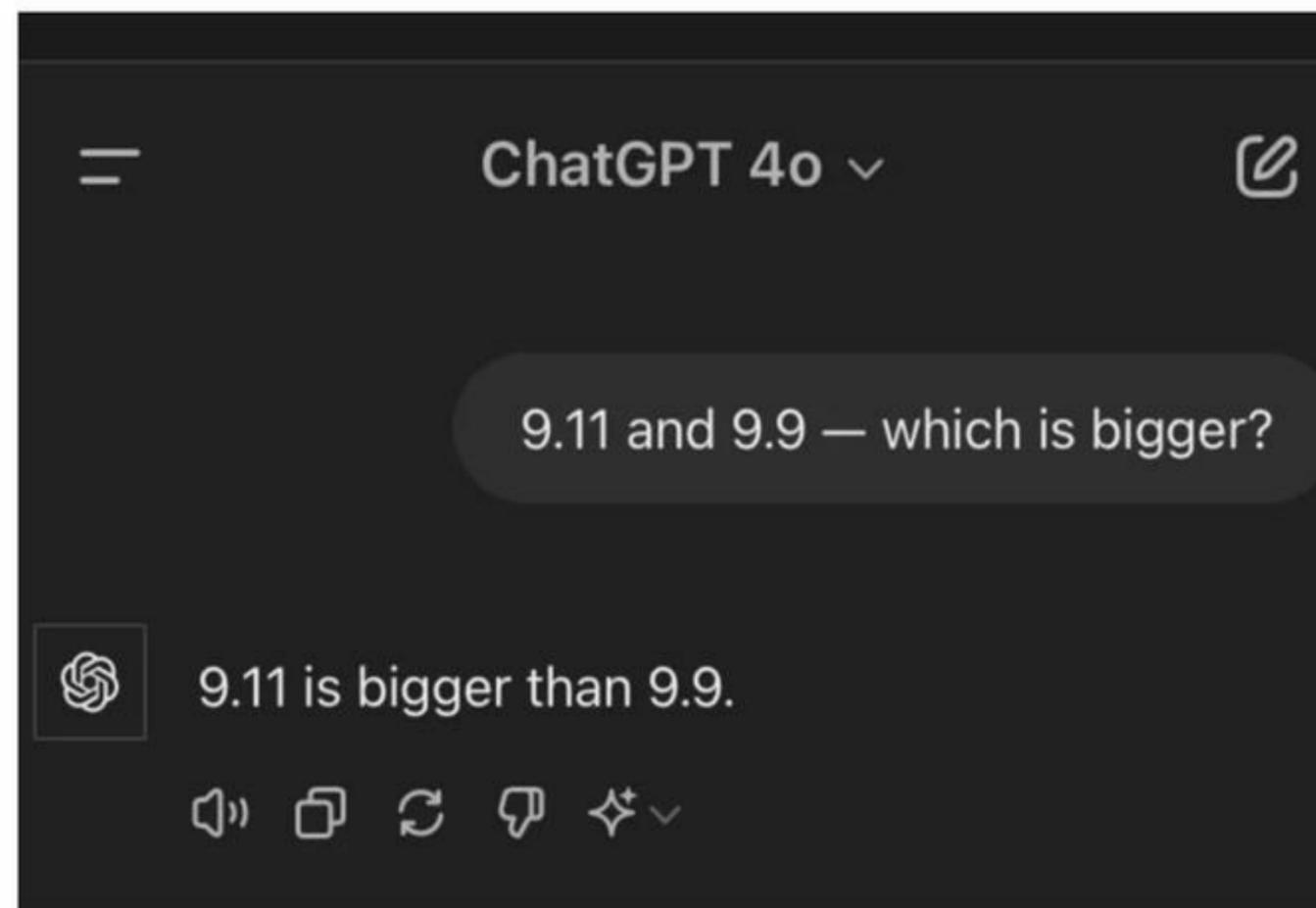
这张图展示了一只橙色的猫在空中跳跃的瞬间。猫的身体悬空，四肢伸展，仿佛正在从一个地方跃向另一个地方。背景是一片户外区域，有树木、围栏和一些模糊的建筑物或结构，表明这只猫可能是在一个户外花园或院子里。图像捕捉到了猫灵活、敏捷的动作，表现出其活泼和充满活力的特性。

ChatGPT进行文本问答（左）与图像内容总结（右）

- 虽然ChatGPT具有强大的自然语言处理能力，但其输出质量在很大程度上取决于用户输入的提示词。通过合理的提示词设计，用户可以更好地引导模型生成连贯、准确且有用的回答。



**9.9和9.11比大?  
有这个必要吗?**



➤ 虽然ChatGPT具有强大的自然语言处理能力，但其输出质量在很大程度上取决于用户输入的提示词。通过合理的提示词设计，用户可以更好地引导模型生成连贯、准确且有用的回答。

➤ 思维链（Chain-of-Thought, CoT）

就是一种指导用户设计提示词的技术。

思维链是指将一个复杂的推理任务分解为一系列中间步骤，通过生成这些中间的推理步骤来引导模型从问题的输入逐步推导出最终的答案，就像链条一样，一环扣一环，形成一个连贯的思维过程。

9.9与9.11哪个大？



9.11 比 9.9 大。

比较这两个数时，先看整数部分，都是9，接着比较小数部分：

- 9.9 的小数部分是 0.9
- 9.11 的小数部分是 0.11

由于  $0.11 > 0.9$ ，因此  $9.11 > 9.9$ 。



➤ 虽然ChatGPT具有强大的自然语言处理能力，但其输出质量在很大程度上取决于用户输入的提示词。通过合理的提示词设计，用户可以更好地引导模型生成连贯、准确且有用的回答。

➤ 思维链（Chain-of-Thought, CoT）

就是一种指导用户设计提示词的技术。

思维链是指将一个复杂的推理任务分解为一系列中间步骤，通过生成这些中间的推理步骤来引导模型从问题的输入逐步推导出最终的答案，就像链条一样，一环扣一环，形成一个连贯的思维过程。

9.9和9.11哪个大，一步一步分析



问题：9.9和9.11哪个大？

分析：

1. 统一小数位数：

- 为了便于比较，将9.9写成9.90。
- 现在比较9.90和9.11。

2. 比较整数部分：

- 9.90的整数部分是9。
- 9.11的整数部分也是9。
- 两者整数部分相同，需要比较小数部分。

3. 比较小数部分：

- 9.90的小数部分是0.90。
- 9.11的小数部分是0.11。
- $0.90 > 0.11$ 。

结论：

由于小数部分0.90大于0.11，因此9.9大于9.11。

最终答案：

$\boxed{9.9 > 9.11}$

- **增强了大模型的推理能力：**CoT 通过将复杂问题分解为多步骤的子问题，相当显著的增强了大模型的推理能力，也最大限度的降低了大模型忽视求解问题的“关键细节”的现象。
- **增强了大模型的可解释性：**CoT 让大模型展示了“做题过程”，使得我们可以更好的判断大模型在求解当前问题上究竟是如何工作的，也为我们定位其中错误步骤提供了依据。
- **增强了大模型的可控性：**通过让大模型一步一步输出步骤，我们通过这些步骤的呈现可以对大模型问题求解的过程施加更大的影响。
- **增强了大模型的灵活性：**可以在现有的各种不同的大模型中使用 CoT 方法。



无思维链与有思维链提示对比

## ■ 智谱清言

- 智谱AI的智谱清言模型基于GLM模型开发，具备内容创作、图像理解、信息归纳总结等能力；

## ■ 通义千问

- 阿里的通义千问大模型有着强大的思维启发能力，能够对复杂问题进行拆解，为用户启发创意，辅助用户学习各种知识；

## ■ 讯飞星火

- 科大讯飞的星火大模型则在语音识别和语音合成等任务上展现出卓越的效率和优秀的表现力；



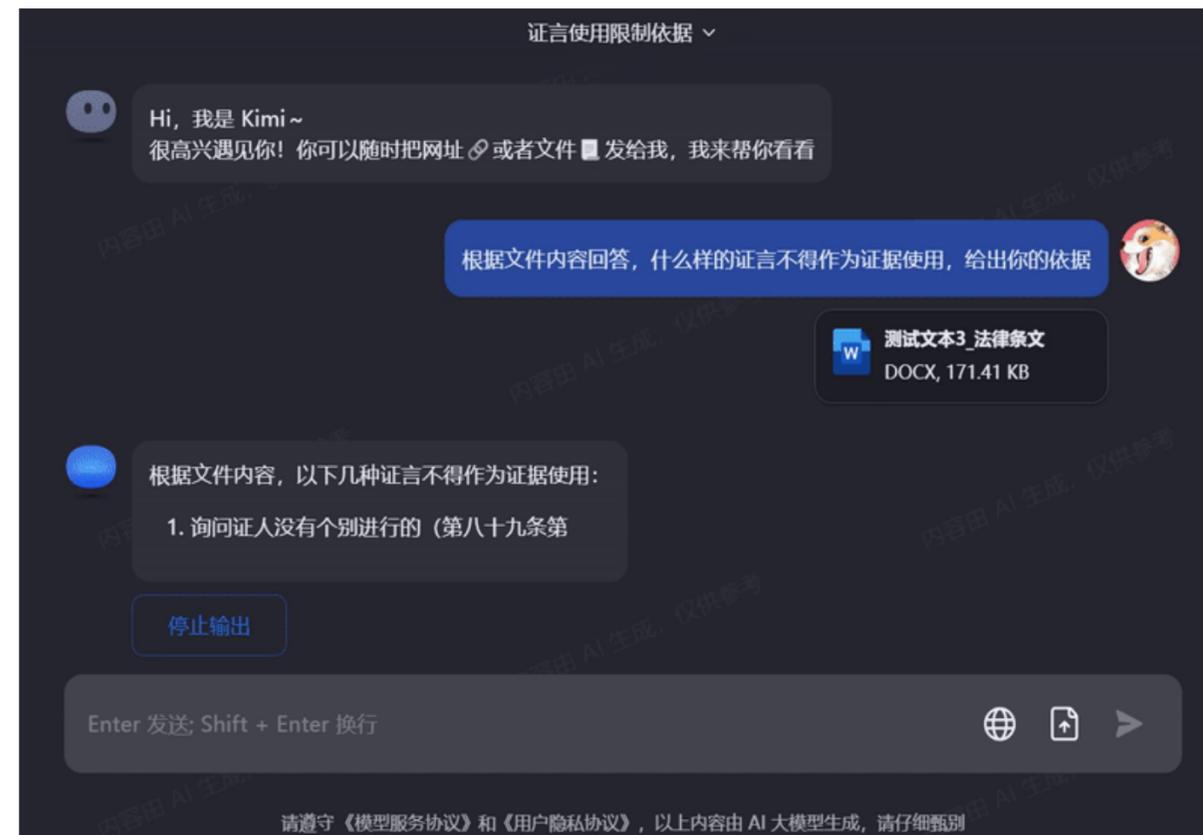
# 百花齐放的国产大模型

## ■ 文心一言

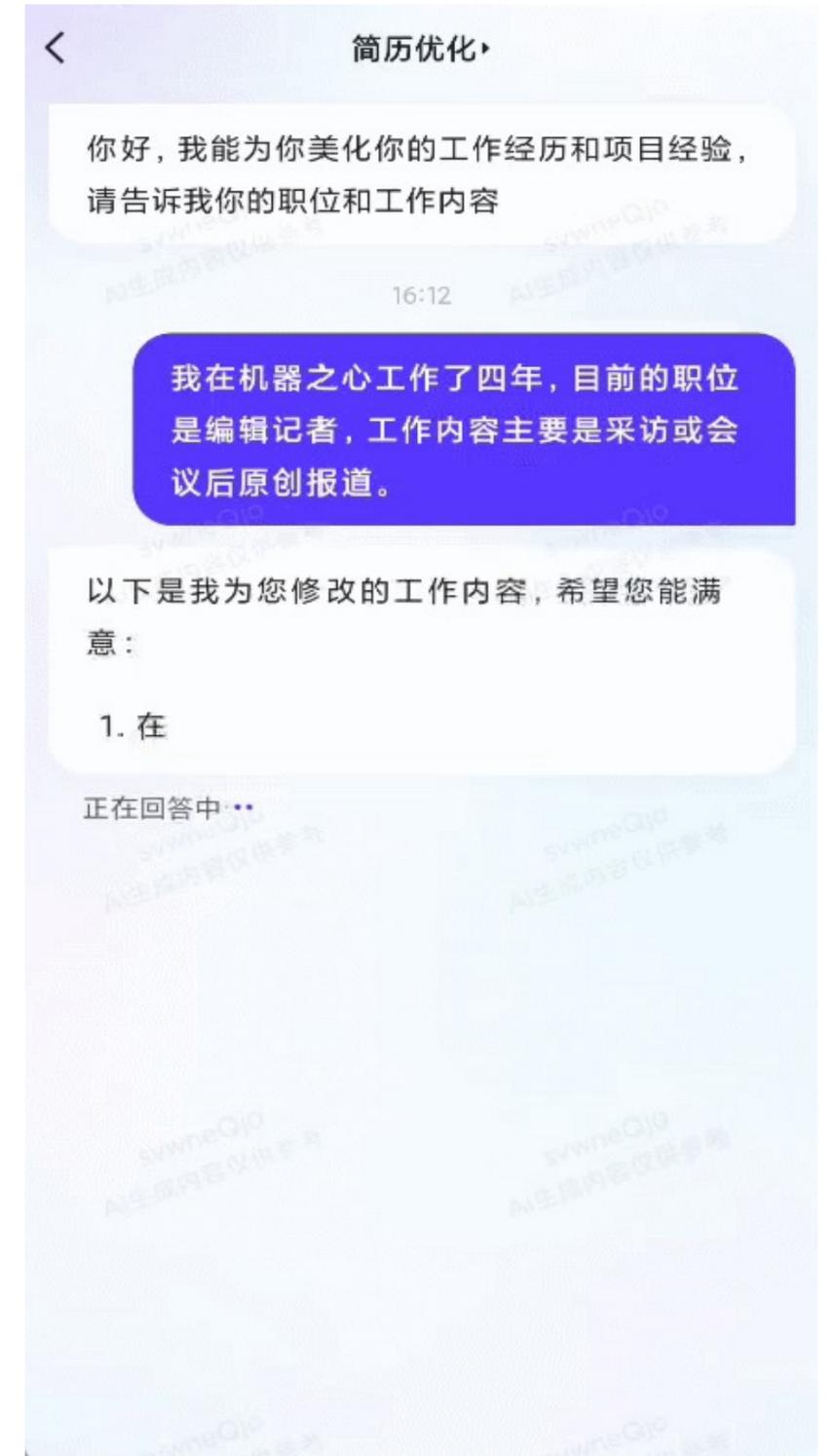
- 百度的文心一言大模型有着出色的联网搜索能力，帮助其在文学创作、问题解答等方面表现出色；

## ■ Kimi

- 月之暗面 (Moonshot AI) 的Kimi Chat大模型则专注于长文本处理，支持对20万汉字的长文本进行文档总结和-content拆解。



使用Kimi Chat进行文档分析



文心一言简历优化

# 百花齐放的国产大模型

## 2025中国 AI大模型 产业图谱 1.0版

数据猿 DataYuan.cn × 上海大数据联盟 SHANGHAI BIG DATA ALLIANCE  
说明:本图谱企业位置不涉及排名与先后顺序

### 垂直大模型

行业大模型					智能体(AI Agent)				
<b>金融大模型</b> HUND SUN 恒生 拓尔思 TRS 中电金信 GienTech 中国建设银行 招商银行	<b>教育大模型</b> 有道 youdao 学而思 领界 AI 点控云 杉数科技 作业帮 TAL 好未来	<b>医疗大模型</b> 京东 中科大脑 北电数智 智云健康 朗玛信息 longmaster AthenaEyes 智慧眼 医联 廿卓世科技 华为云 和和 中国移动	<b>工业大模型</b> 百川智能 BAICHUAN AI 羚羊 北电数智 LUSTER 凌云光 卡奥斯 COSMOPlat H3C 杉数科技 华为云 SmartMore	<b>政务大模型</b> 中科大脑 华为云 北电数智 天翼云 State Cloud 中国移动 华院计算	<b>通用智能体</b> 金蝶 Fabarta 中电金信 GienTech 实在智能 HAND	<b>流程智能体</b> 神州数码 Digital China 芝罘旗 廿卓世科技 实在智能	<b>营销智能体</b> 容联云 CLOOPEN Marketingforce 迈富时 高科数聚 Gausscode LUSTER 凌云光 BlueFocus 浙文互联	<b>客服智能体</b> 网易云商 容联云 CLOOPEN 高科数聚 Gausscode 天润融通 智齿科技	<b>数据分析智能体</b> 数势科技 拓尔思 TRS Fabarta Aloudata 羚羊

### 通用基础大模型

大语言模型	图像大模型	语音大模型	视频大模型	多模态大模型
阿里巴巴 廿卓世科技 Tencent 腾讯 华为云 百川智能 智谱·AI Transn 传神 科大讯飞 IFLYTEK Bai 百度 Moonshot AI	智谱·AI Tencent 腾讯 抖音集团 Bai 百度 科大讯飞 IFLYTEK 阿里巴巴 华为云	阿里巴巴 Tencent 腾讯 云和声 抖音集团 科大讯飞 IFLYTEK Bai 百度 出门问问	Bai 百度 生数 商汤 阿里巴巴 火山引擎	阿里巴巴 Transn 传神 Tencent 腾讯 华为云 科大讯飞 IFLYTEK Bai 百度

### 基础设施

AI算力				实时通信	数据集
<b>AI芯片</b> 壁仞科技 摩尔线程 MOORE THREADS 地平线 Horizon Robotics Cambricon 寒武纪 HYGON 平头哥	<b>AI服务器</b> 神州数码 Digital China ZTE中兴 H3C FI工业富联	<b>智算中心</b> 摩尔线程 MOORE THREADS 中国电信 CHINA TELECOM 禾·道 ZTE中兴 燧智科技 YeeZ Tech 中国联通 和和 中国移动 China Mobile	<b>智能云服务</b> 腾讯云 QingCloud 容联云 CLOOPEN 移动云 金山云 华为云 京东云 百度智能云 阿里云 联通云 UCLLOUD 优刻得	<b>实时通信</b> 声网 网易云信 七牛云 ZEGO 即构	<b>数据集</b> <b>通用数据集</b> 腾讯云 阿里云 阿里巴巴 <b>行业/领域数据集</b> 阿里云 医渡云 科大讯飞 IFLYTEK 神州数码 Digital China Bai 百度 拓尔思 TRS 中电金信 GienTech 海天瑞声

# 4.4 通用人工智能

- 通用人工智能（Artificial General Intelligence, AGI）是指具备与人类相当或超越人类水平的智能系统，能够在各种任务和环境中灵活应对、学习和推理。与狭义人工智能（Narrow AI）不同，AGI 的目标是实现广泛的认知能力，而不仅仅是解决特定任务。
- 目前AGI尚未实现，深度学习、强化学习、神经符号计算等被认为是实现 AGI 的可能路径。大型语言模型（如 GPT 系列）展示了部分通用能力的潜力，但仍局限于特定任务和数据范围。



AGIC深圳国际通用人工智能展



360集团创始人周鸿祎免费课开讲



OpenAI，是一家开放人工智能研究和部署公司，其使命是确保通用人工智能造福全人类。创立于2015年12月，总部位于美国旧金山。OpenAI最早为非营利组织，于2015年底由包括萨姆·奥尔特曼、彼得·蒂尔、里德·霍夫曼和埃隆·马斯克等创办。随着2018年埃隆·马斯克退出以及大模型对资金超预期需求，2019年3月，公司从非营利性转变为“封顶”的营利性；同年7月，微软与OpenAI合作，注资10亿美元共同研发新的Azure AI超算技术。



2022年11月，全新聊天机器人模型ChatGPT问世，产品上线仅5天用户数量突破100万。2024年2月，文生视频大模型Sora问世，在全球内容创作行业卷起新的风暴，成为人工智能发展进程中的“里程碑”。5月，推出GPT-4o，处理文本、图像、音频能力更自然、流畅。10月，ChatGPT搜索功能上线，覆盖网页版以及手机、桌面应用，是AI领域的一次重大变革。

OpenAI首席执行官萨姆·奥尔特曼（Sam Altman）是改写 AI 历史的传奇，“ChatGPT 之父”！19 岁辍学创业，27 岁成为 Y Combinator 总裁，30 岁创立 OpenAI。因理念分歧戏剧性离职又火速回归，身家超 20 亿美元。



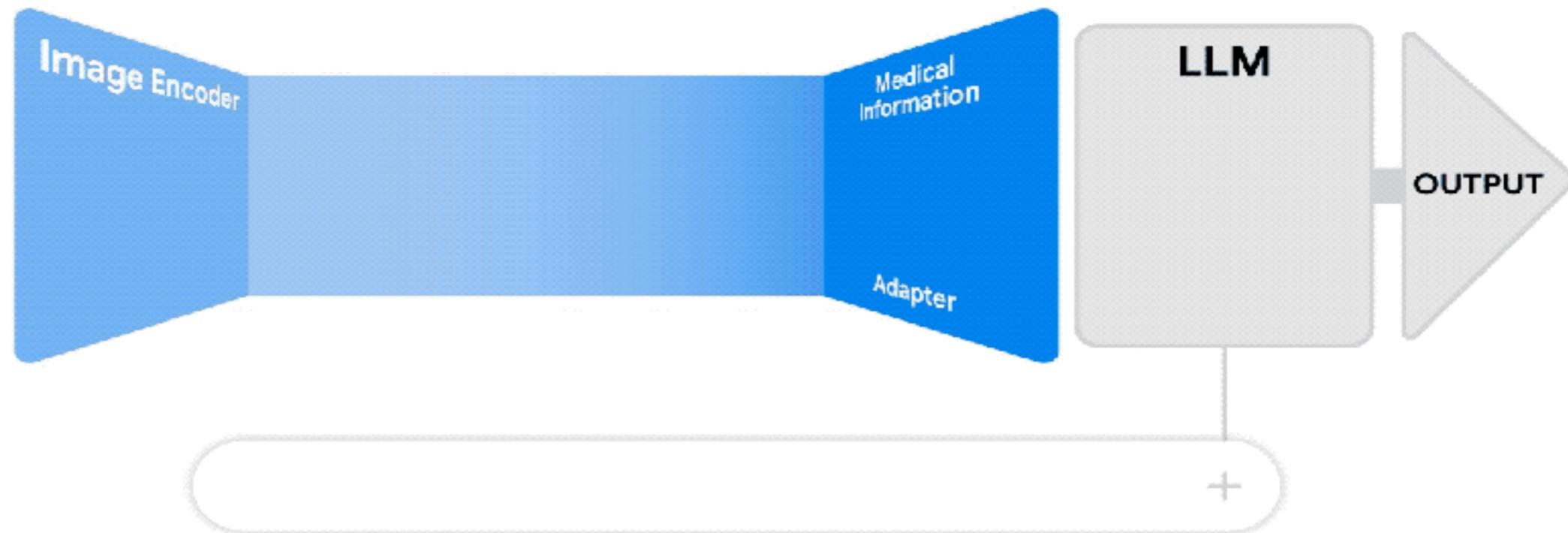
05

# 大模型时代下的人工智能



## 5.1 从语言理解到多模态智能——多模态大模型

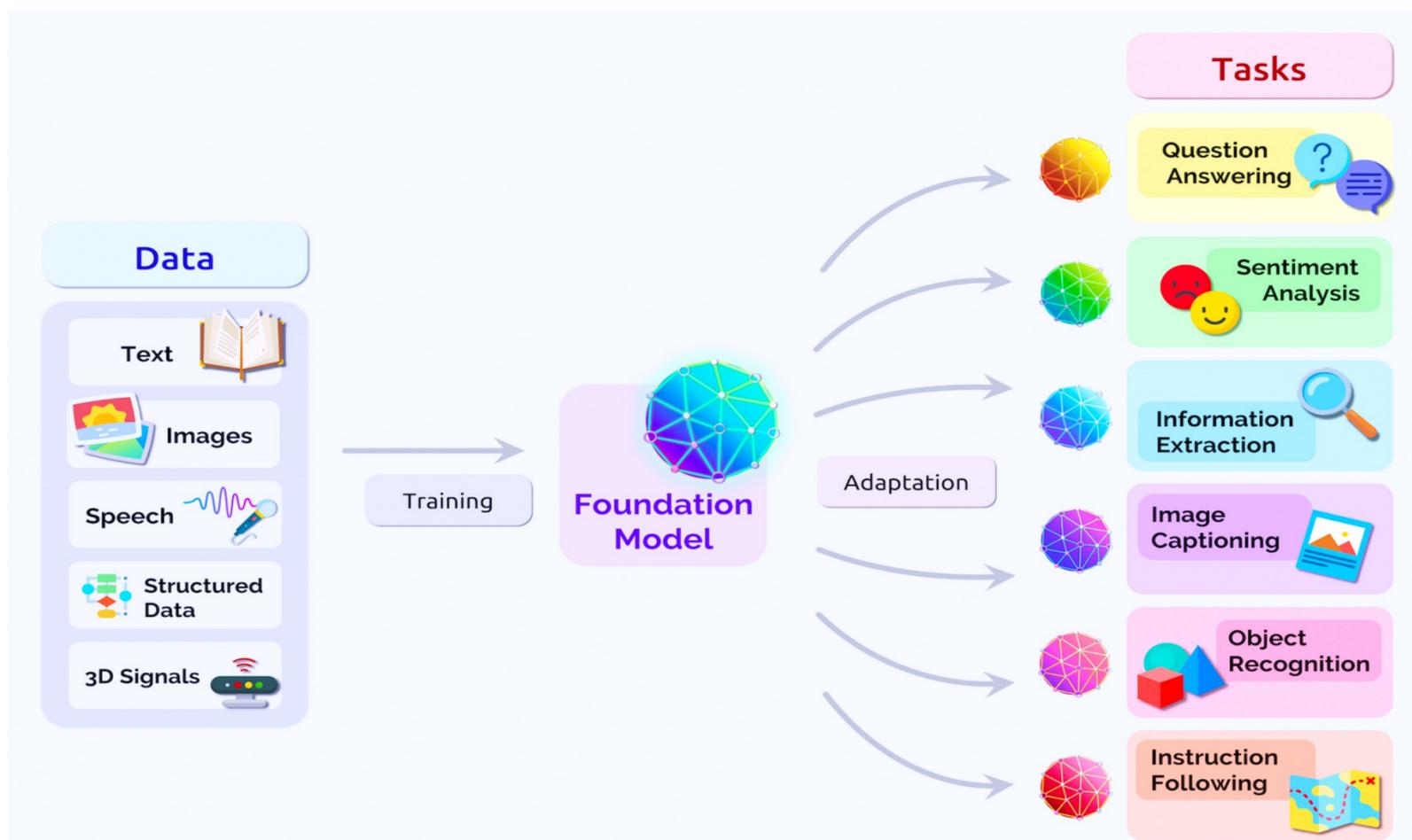
- 大语言模型无法直接处理图像、视频、音频等多种模态信息之间的相互作用，也无法充分理解不同模态之间的上下文关系。因此，在一些需要跨模态理解的任务中，如图像问答、视频描述等，大语言模型的表现往往不尽如人意。
- 多模态大模型（Large Multimodal Model, LMM）应运而生，是一种能够理解和处理各种输入形式的 AI 模型，这些输入包括各种“模态”，如图像、视频和音频，模态就是 AI模型的数据。



LMM模型对读入的“一张医学图像”（输入1）进行“是否存在胸腔积液”（输入2）的诊断

## ■ 多模态大模型

- 多模态大模型通过融合多种模态的信息，能够在更广泛的场景中提供更全面和准确的理解和生成能力。这些模型通过大规模的数据训练，学习如何联合理解和生成跨多种模态的信息，被视为通往通用人工智能的下一个重要步骤。



## ■ 多模态大模型百花齐放

- 文生视频大模型
- 文生音频大模型
- 图文问答大模型
- 图像分割大模型
- 气象预测大模型

## 5.2 以文为引，化影成真——文生视频大模型

- 早期文生视频（Text-to-Video, T2V）模型缺乏对文本与视频之间复杂关系的深度理解，这些模型生成的视频在视觉效果上往往较为粗糙，缺乏真实感，而且无法生成高分辨率、细节丰富的视频。而且，帧与帧之间的过渡常常不自然，导致视频中的动作显得不连贯或僵硬。
- 2024年，OpenAI公司发布人工智能文生视频T2V大模型Sora，能够根据用户的文本提示创建最长60秒、包含多个角色和特定运动的逼真视频，具有深度模拟真实物理世界的的能力。





一分钟了解  
“恐怖”的Sora

## ■ Sora关键技术

- **多模态融合架构：**通过将文本和视觉信息深度结合，模型能够更好地理解文本的复杂语义，并将其准确映射到视频生成过程中，使得生成的视频能够精确反映文本描述的内容和细节。
- **生成网络和图像增强技术：**通过多层次的生成对抗网络（Generative Adversarial Networks, GAN）、高分辨率生成网络（High-Resolution Generation Network）和超分辨率重建技术（Super-Resolution Reconstruction），Sora能够生成高分辨率且细节丰富的视频，克服了早期模型在生成质量上的缺陷。
- **时间一致性模块：**通过平滑处理相邻帧之间的过渡，确保动作和场景在时间轴上自然流畅。
- **自监督学习机制：**Sora通过在大规模多模态数据集上的自监督学习，不仅学习了广泛领域的文本-视频映射，还在未标注的数据中挖掘潜在的模式，显著提升了模型的泛化能力。

5.2

# 以文为引，化影成真——文生视频大模型



## ■ 法律风险和监管问题

- Sora的出现引发了一些法律风险和监管问题。例如，Sora作品是否存在侵犯他人著作权的可能，成为了业界关注的焦点。
- 同时，Sora生成的视频内容可能涉及到对现有作品的再创作，这也在一定程度上挑战了传统的版权保护框架。



## ■ 对视频制造行业的影响

- 一方面，Sora的高效率和成片效果可能会降低视频制作的门槛，激发更多人的创造力。
- 另一方面，Sora的发布也引发了对滥用视频生成技术的担忧，AI生成的“深度伪造”内容可能导致虚假和错误信息的广泛传播。



## ■ 国产文生视频大模型

- 快手的可灵（Kling）大模型主体采用Diffusion Transformer架构，并结合3D 时空联合注意力机制更好地建模了复杂时空运动，从而生成出**符合运动规律并具有较大幅度**的视频内容。同时，可灵大模型能够生成**长达2分钟**的视频，且帧率达到30fps。通过采用可变分辨率训练策略，让模型在推理过程中可以做到同样的内容输出多种多样的视频宽高比。

## ■ 未来展望

- Sora引领视频内容创作步入了新纪元，为创意产业带来了无限可能。
- 可以预见的是，这些生成类大模型在不久的将来一定会成为艺术工作者、视觉设计师和广告工作者们必不可少的辅助工具。



可灵大模型生成的视频

## ■ 盘古气象大模型

- 全球气候变化的加剧和极端天气事件的频繁发生，传统的数值天气预报方法在处理复杂的气象数据和捕捉细微的气象变化方面仍存在不足。
- 盘古气象大模型是由华为云研发的创新AI气象预报系统，是**AI4Science**领域的典型技术突破。



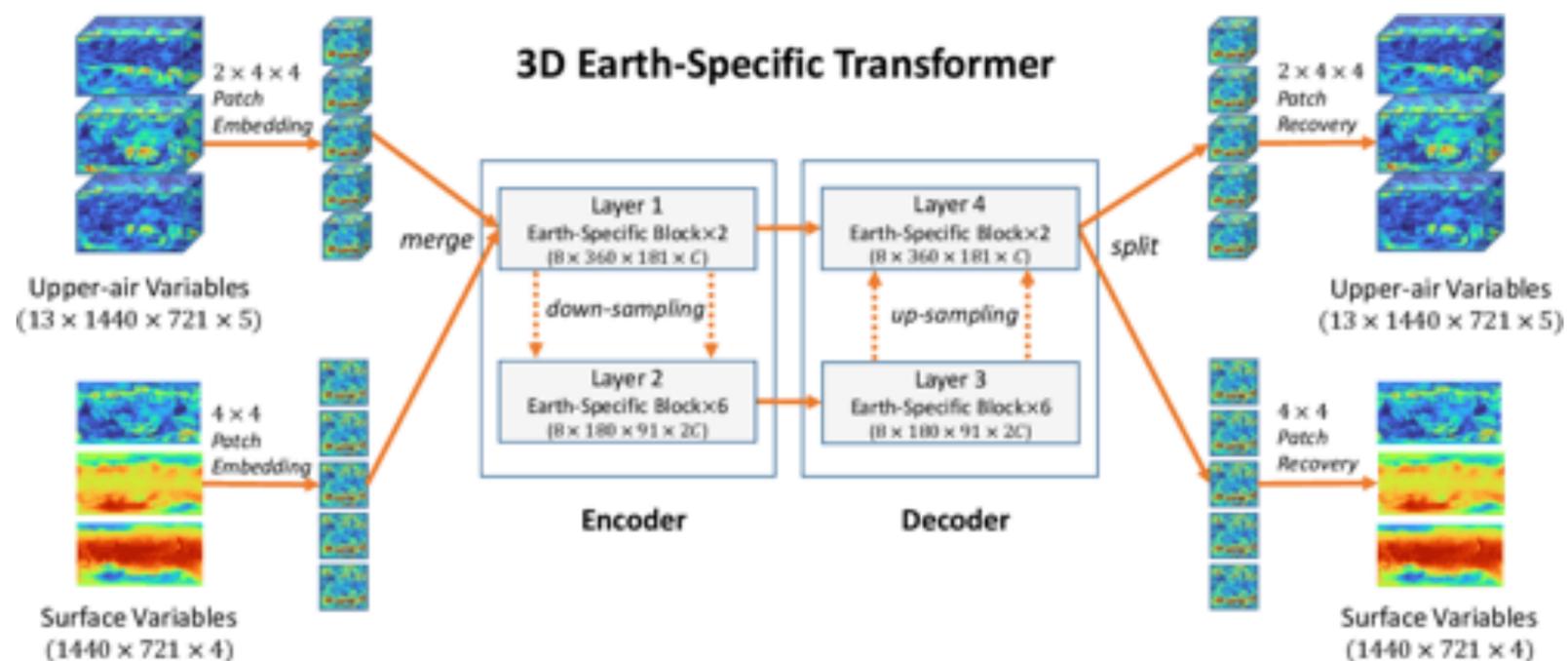
盘古气象大模型发布会现场

## ■ 盘古气象大模型的成就

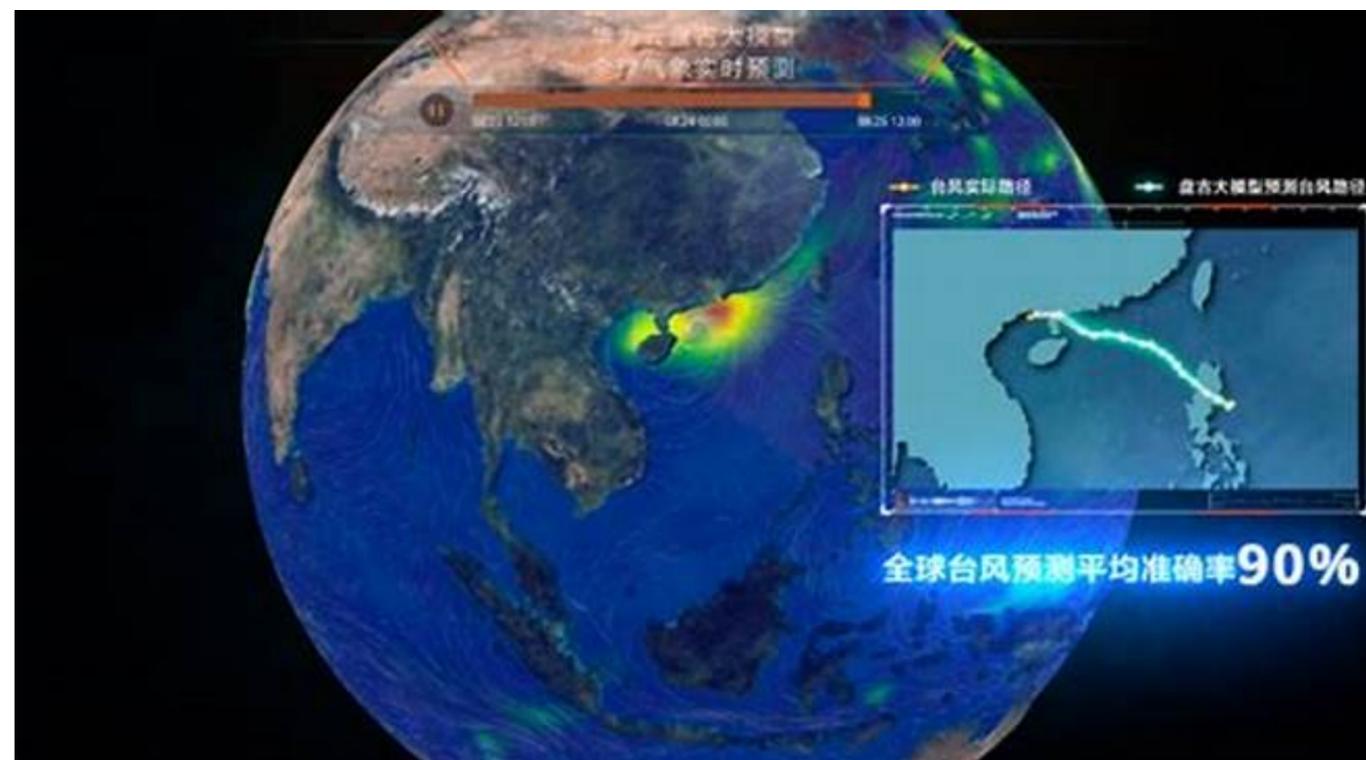
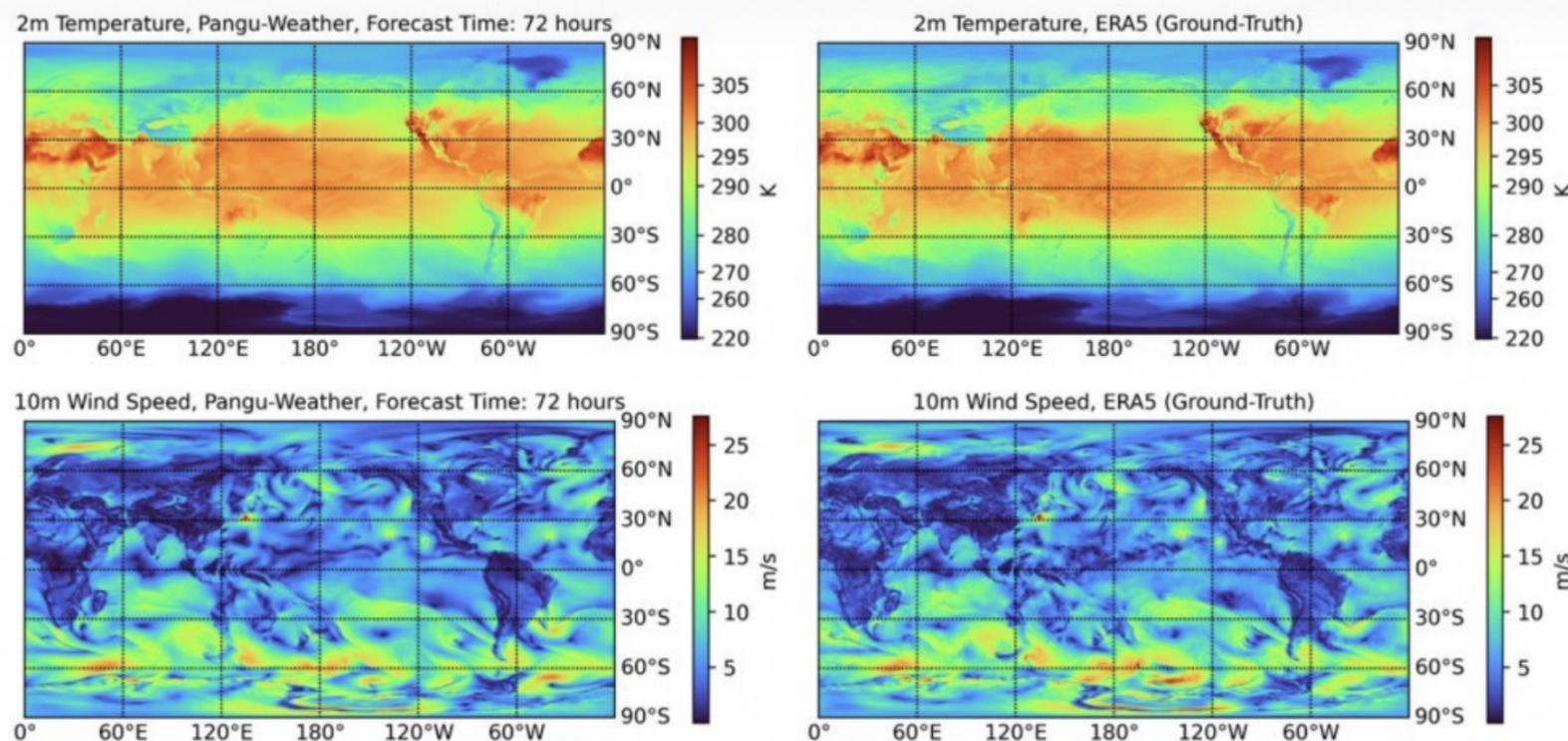
- 更为精准且及时的天气预报服务，显著增强了人类应对自然灾害的预见能力。
- 预测速度上实现了革命性的提升——**提速高达10000倍以上**。
- **首个**精度超过传统数值预报方法的AI模型。

- 由于气象数据与图像数据之间存在诸多的相似之处，那我们能否利用当前计算机视觉（Computer Vision, CV）领域的大模型，对于气象数据进行分析与预测？答案是肯定的。虽然如此，但是直接将现有的CV大模型架构应用于气象预测上还有诸多不足之处。
- 为此，盘古气象大模型设计了3D Earth-Specific Transformer (3DEST) 模块来处理复杂的不均匀3D气象数据，并且使用层次化时域聚合策略来减少预报迭代次数，大幅提升了模型预测的准确率和效率。

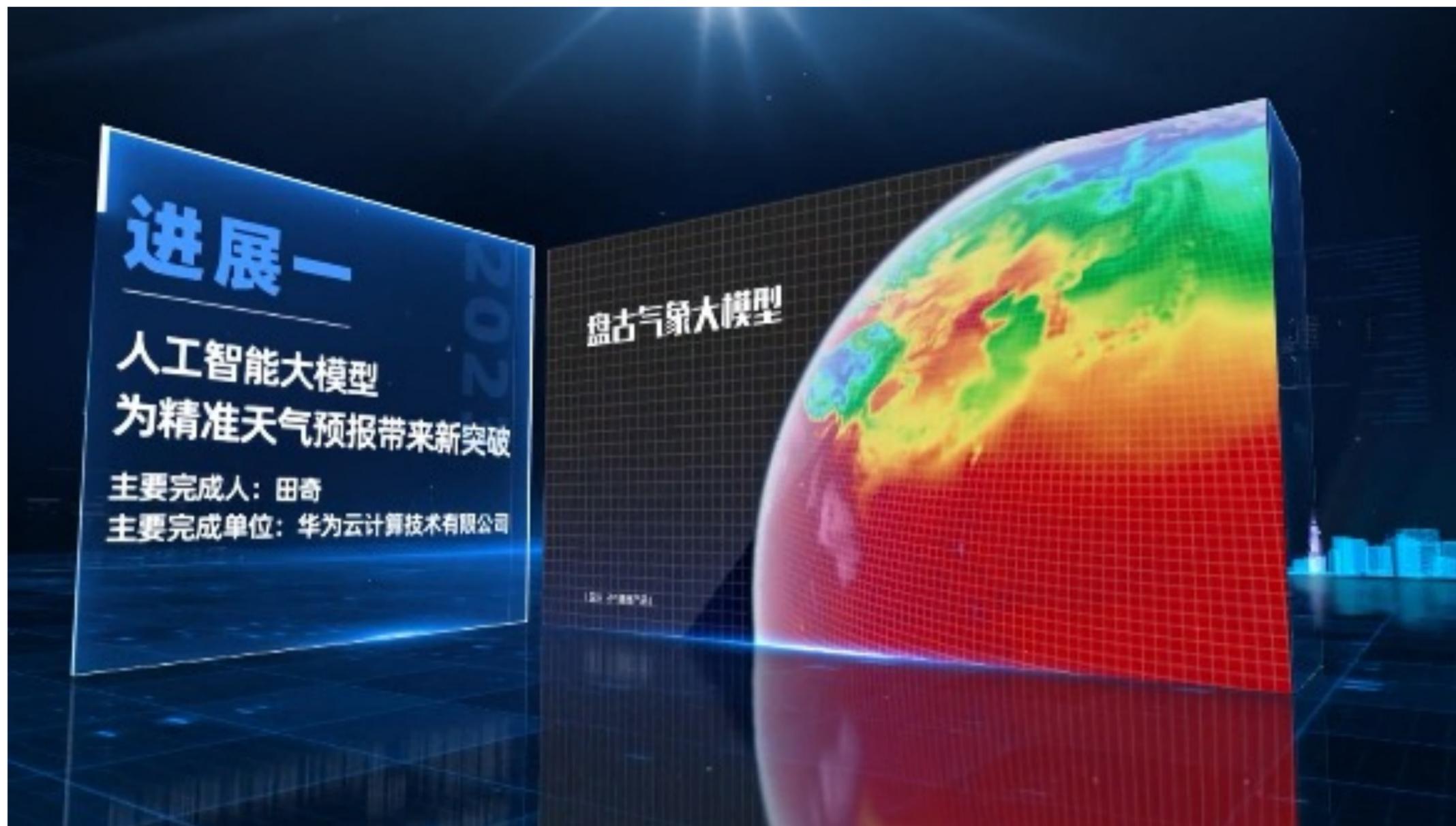
- 该大模型仅需在一张V100显卡上运行1.4秒，就能完成24小时全球气象的预报，位势、湿度、风速、温度、海平面气压等复杂信息可以在一块屏幕上一目了然。

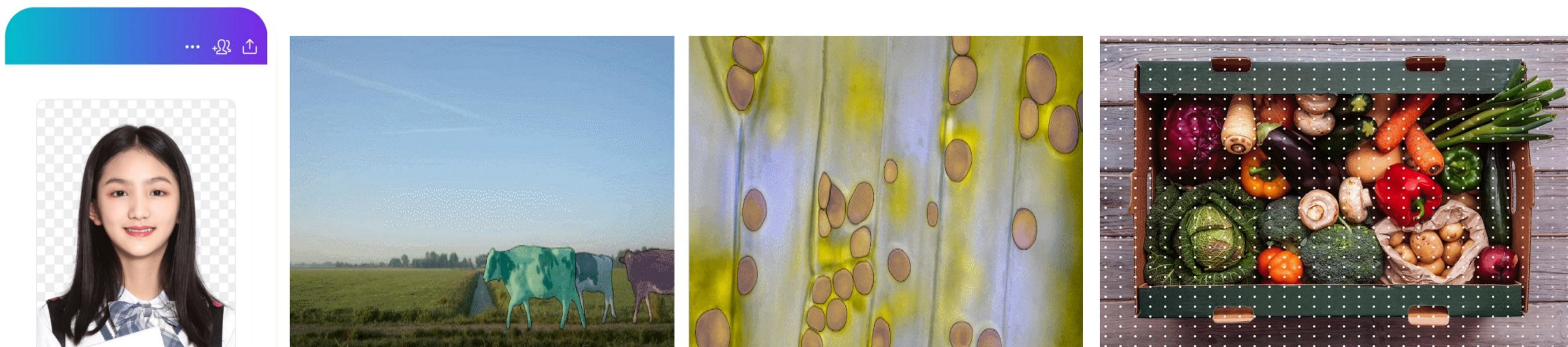


- 2023年10月，飓风“奥蒂斯”预计影响墨西哥地区，欧洲气象局预测该飓风会提前消散，而盘古气象大模型则给出了不同的预测结果并成功预测了飓风“奥蒂斯”的实际运动路径，其预测曲线与飓风实际运动路径几乎一致，精度达到了气象预报的新高度。这一预测结果展示了盘古气象大模型在复杂气象条件下的卓越预测性能。



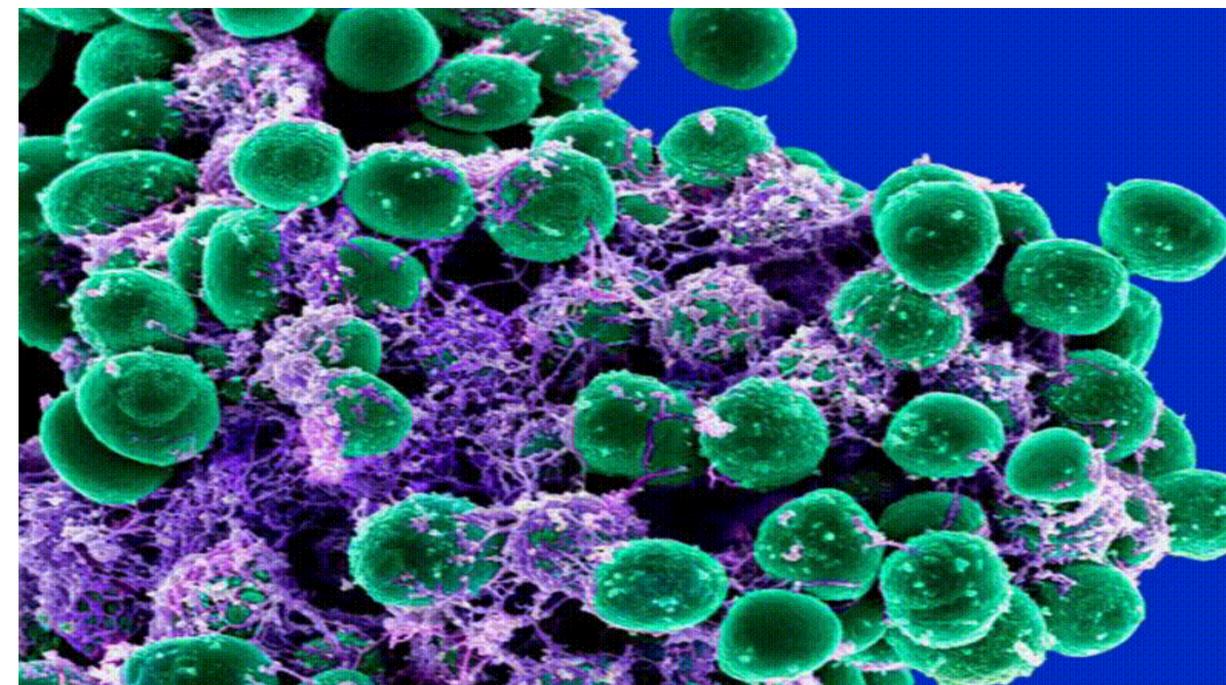
- 成果成为近年来中国科技公司首篇作为唯一署名单位发表的《Nature》正刊论文，入选2023年度中国科学十大进展。





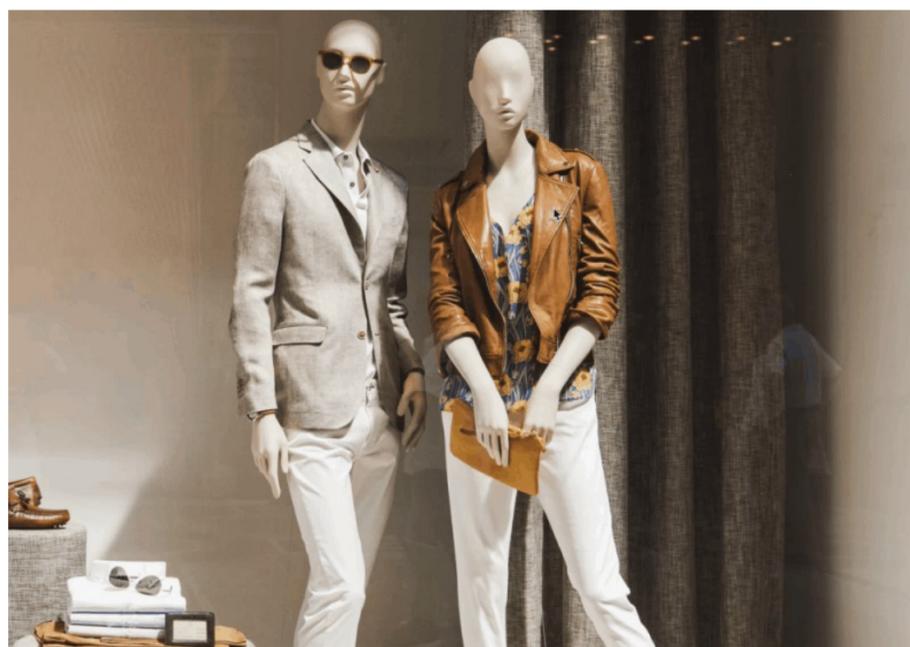
- SAM (Segment Anything Model) , 作为Meta公司的一项里程碑式创新成果, 在计算机视觉领域内开创了前所未有的图像分割新纪元。
- 用户**只需指尖轻点或随意勾勒几笔**, 计算机便能即时洞悉图像奥秘, 精准分离并识别出画面中的任意目标物体。而且SAM具有令人称叹的**零样本学习能力**, 即便面对全新、未见过的物体类别, 也能仅凭用户提供的简单线索, 迅速勾勒出物体的精确轮廓。

- 以前，针对特定图像任务构建高精度的目标分割模型，依赖于技术专家手工标注，耗时费力，极大地限制了模型开发与迭代的效率。SAM的问世则彻底革新了这一领域的运作模式。
- SAM内置的数据标注引擎，以其强大的自动化能力，自动生成精确的分割标签，众多数据标注平台将SAM作为图像对象分割标注的首选工具，节省了数百万小时的人工注释时间。
- SAM在科学探索、医学诊断、乃至众多工业领域激发出新的活力。



## ■ SAM2的升级与突破

- 为了更全面地捕捉动态场景的本质，Meta推出了SAM2。
- SAM2的核心优势在于其即时响应与高效传播机制：一旦接收到用户提示，它便能迅速在当前帧上生成初始掩码，并沿时间轴扩展至整个视频序列，精准描绘出目标对象的动态轮廓。
- 用户还能在视频播放过程中随时添加额外提示，对初始掩码进行迭代优化，增强了交互的灵活性与结果的准确性。



## ■ SAM2的应用

- 许多重要的工程应用也都需要在视频数据和3D数据中进行准确的对象分割，例如在混合现实、机器人、医疗手术机器人、自动驾驶汽车和视频编辑中。在工业领域，如自动驾驶汽车中使用的系统，它可以为视觉数据提供更快的注释工具，以训练下一代计算机视觉系统。
- 对于内容创作者，SAM2可以在视频编辑中实现创意应用，并为生成视频模型增加可控性。



## ■ 挑战

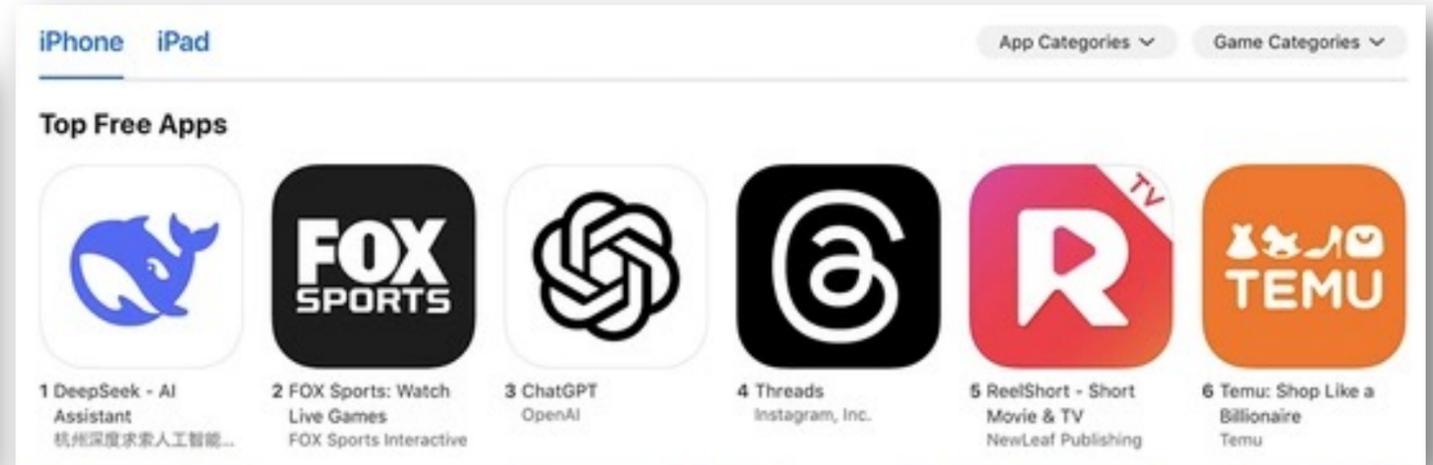
- 大模型技术的发展也面临着诸多挑战。
- 排在首位的便是计算资源消耗问题，训练一个通用基础大模型便要消耗数百万度电。
- 此外，模型的偏见、幻觉 等问题，以及可能被滥用于生成虚假信息等伦理问题，都是亟待解决的难题。

## ■ 未来

- AI的发展历程告诉我们，科技的进步往往不是一帆风顺的，它需要我们保持热情，同时也要脚踏实地，不断探索和创新！



**特朗普称DeepSeek崛起  
应当为美企敲响“警钟”  
也表示“这可能是非常积极的发展”**



**MONEYWATCH**  
**What is DeepSeek, and why is it causing Nvidia and other stocks to slump?**



By Aimee Picchi  
Edited By Alain Shertor  
Updated on: January 27, 2025 / 4:58 PM EST



**NEWS**  
Home | InDepth | Israel-Gaza war | War in Ukraine | Climate | UK | World | Business | Politics | Culture

World | Africa | Asia | Australia | Europe | Latin America | Middle East | US & Canada

**DeepSeek shows AI's centre of power could shift away from US**

教育部 - 华为智能基座课程

人工智能研究生通识示范课程



山东大学  
SHANDONG UNIVERSITY

# 第1讲 人工智能概述