



AI通识基础课程 (AIB310002)

# 具身智能引论

Introduction to Embodied Intelligence

丁文超 李维 浦剑 商慧亮 尚笠 杨帆

# 课程介绍

- 具身智能简介
- 课程相关信息
- 课程内容综览



# 课程介绍

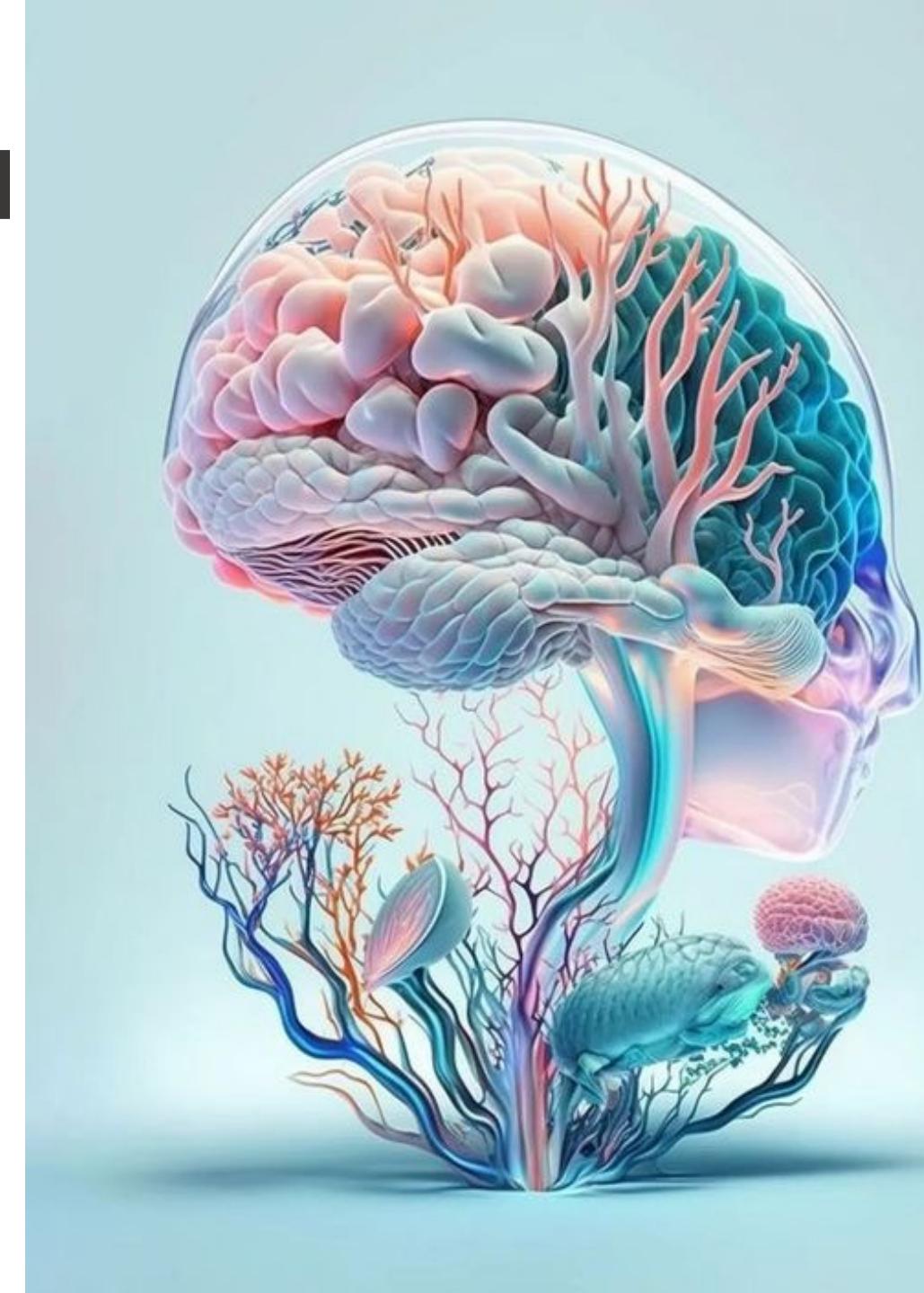
- **具身智能简介**
- 课程相关信息
- 课程内容综览



# Minds awaken in a world

- 我们并没有设计这个世界。我们只是发现自己身处其中。  
We did not design our world. We simply found ourselves with it.
- 我们醒来时发现了自己，也发现了我们所居住的世界。  
We awoke both to ourselves and to the world we inhabit.
- 随着成长和生活，我们开始思考这个世界。我们思考的世界并不是被创造的，而是被发现的  
We come to reflect on that world as we grow and live. We reflect on a world that is not made, but found.
- 我们的存在使我们能够反思这个世界。在反思中我们发现自己处于一个循环之中：我们处在一个似乎在反思开始之前就已经存在的世界里，但这个世界与我们紧密联系  
It is also our structure that enables us to reflect upon this world. Thus in reflection we find ourselves in a circle:  
we are in a world that seems to be there before reflection begins, but that world is not separate from us.

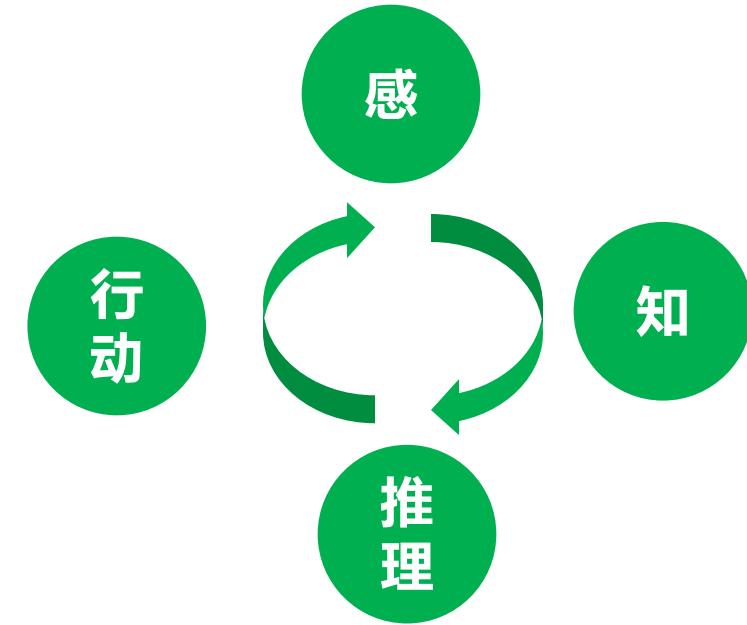
F. J. Varela, E. Thompson, E. Rosch, "The Embodied Mind Cognitive Science and Human Experience", MIT Press, 1992



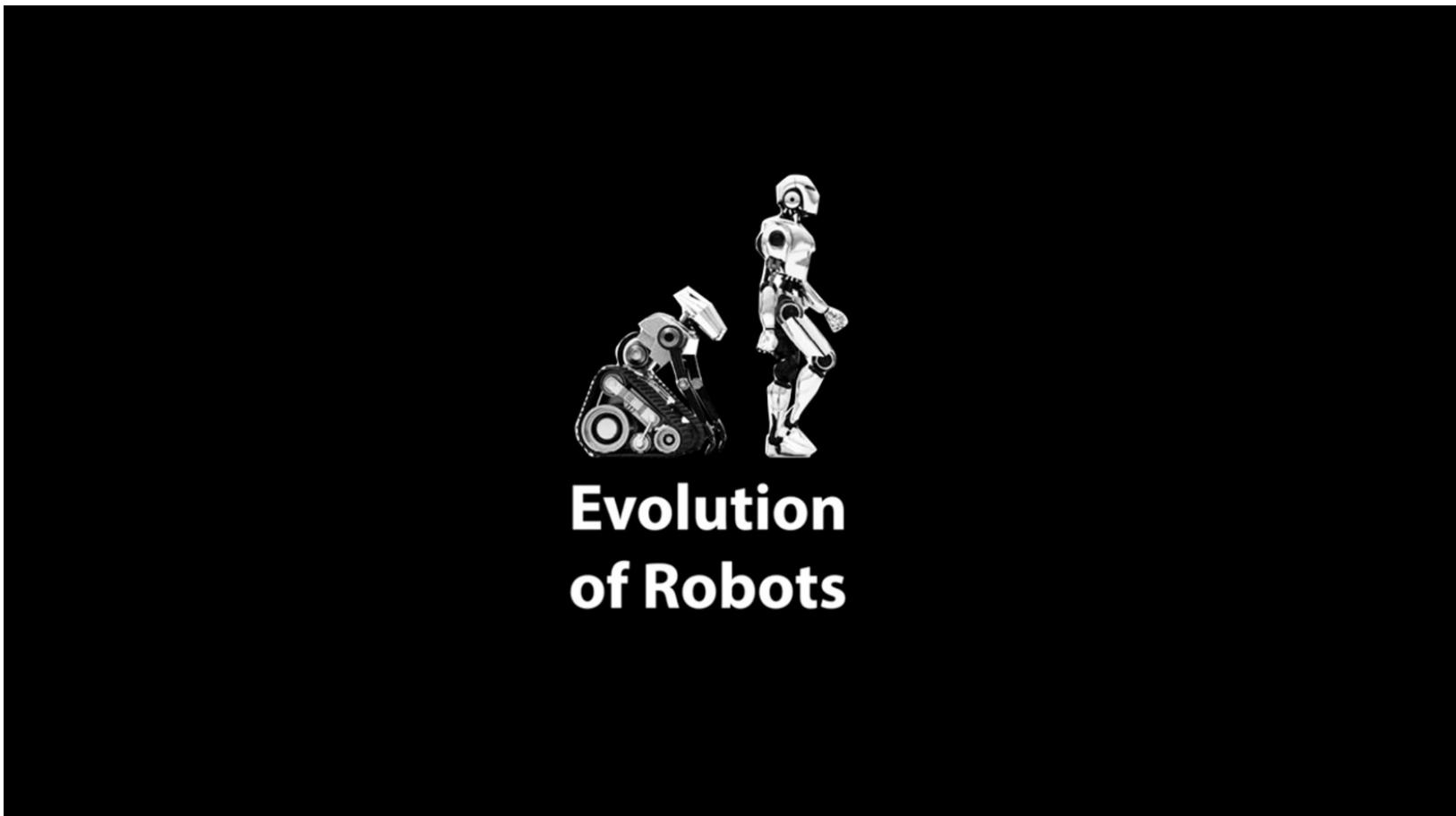


# 具身智能

- 具身智能 (embodied intelligence) 的概念起源于认知科学、人工智能、机器人和哲学的交叉领域。
- 它挑战了传统的观点，即智能主要是发生在大脑中的一种独立于身体和环境的计算过程。
- 具身智能主张，认知和智能深深植根于主体身体与环境之间的物理互动中。

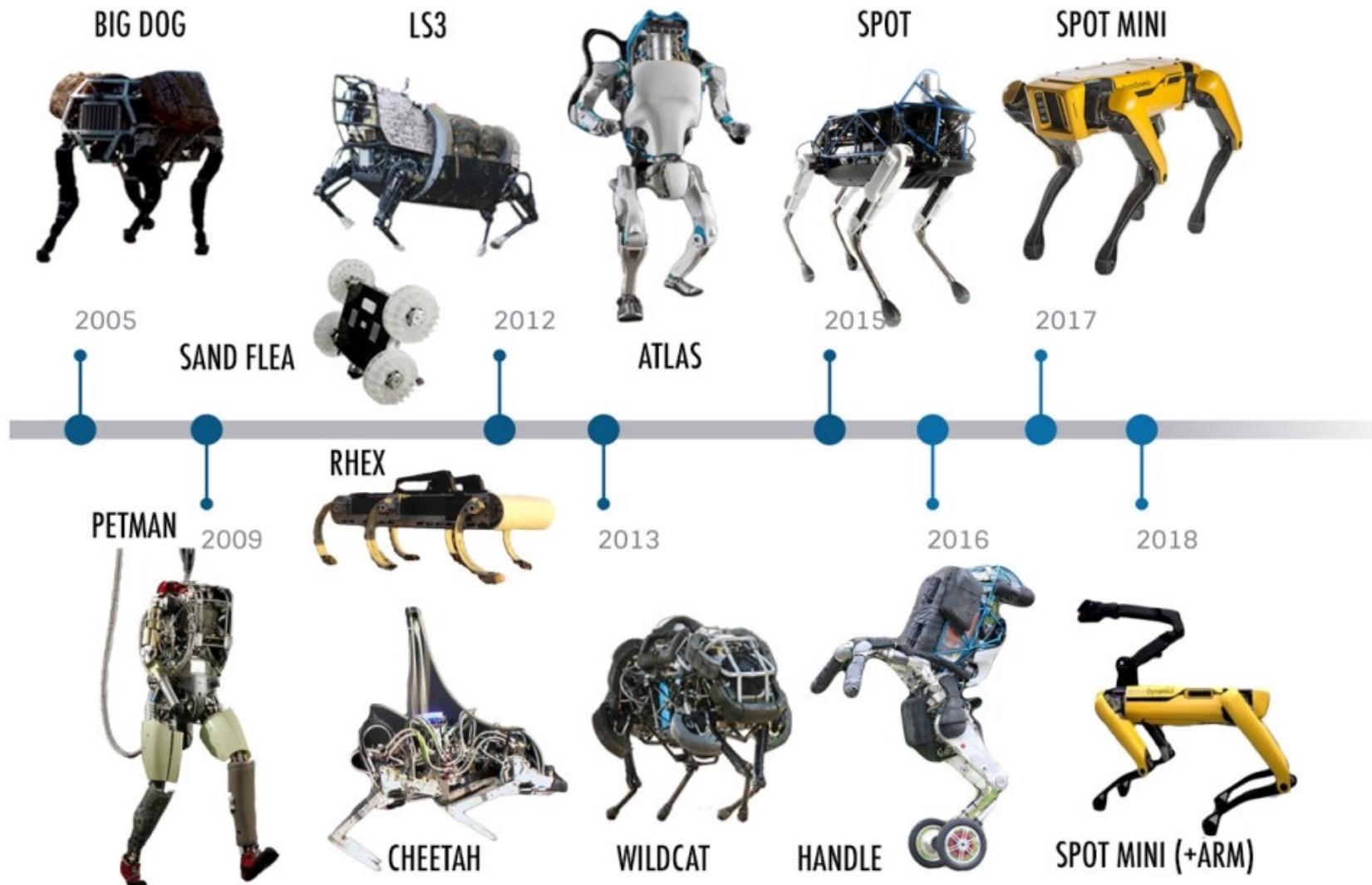


# **Robots Evolution | 100BC - 2020**

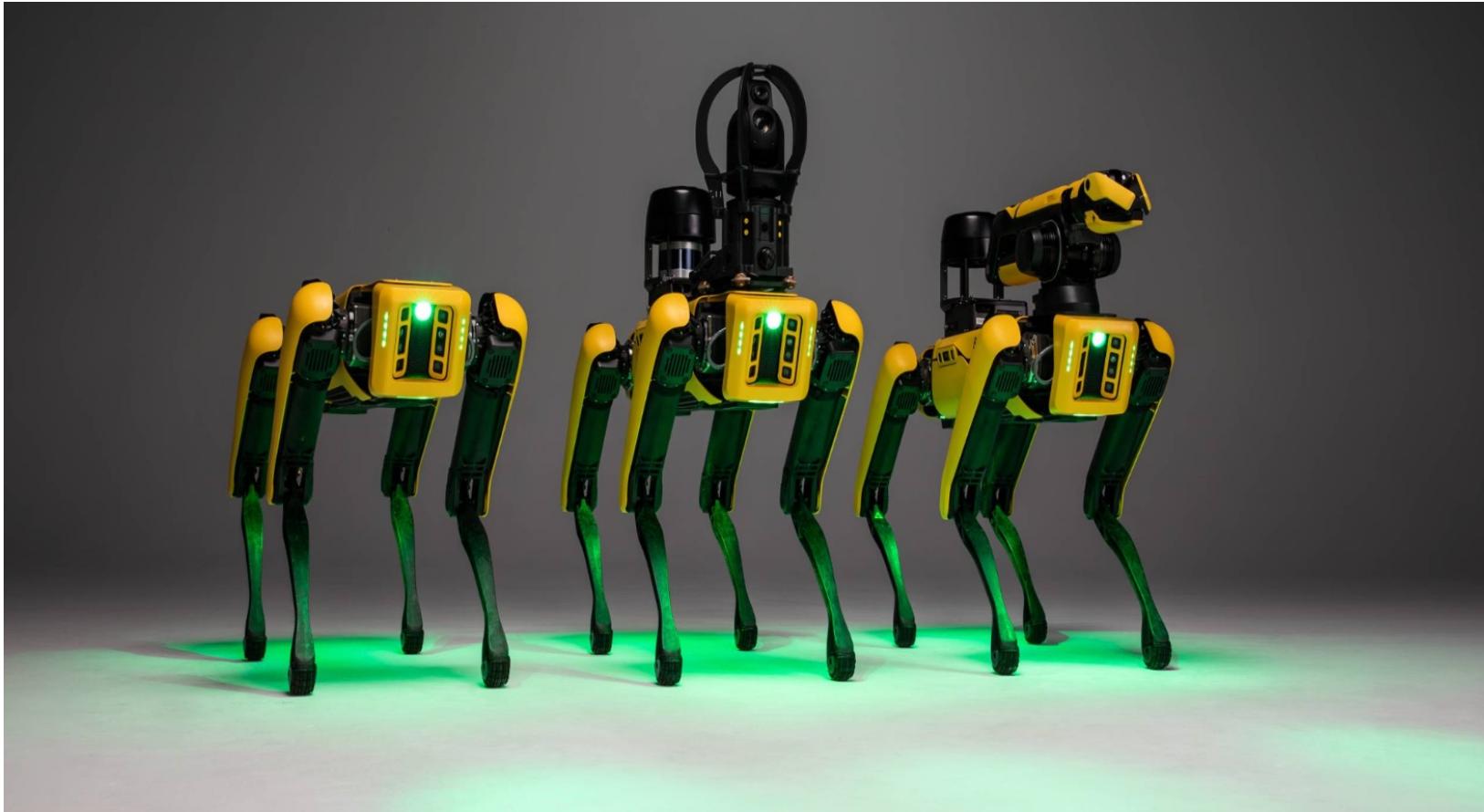


**视频网址:** <https://youtu.be/Uzf-K7UtD78?feature=shared>

# Boston Dynamics



# Boston Dynamics: 40 years of development (1983 - 2023 ) Atlas



视频网址：[https://youtu.be/\\_EZQx87DyzM?feature=shared](https://youtu.be/_EZQx87DyzM?feature=shared)

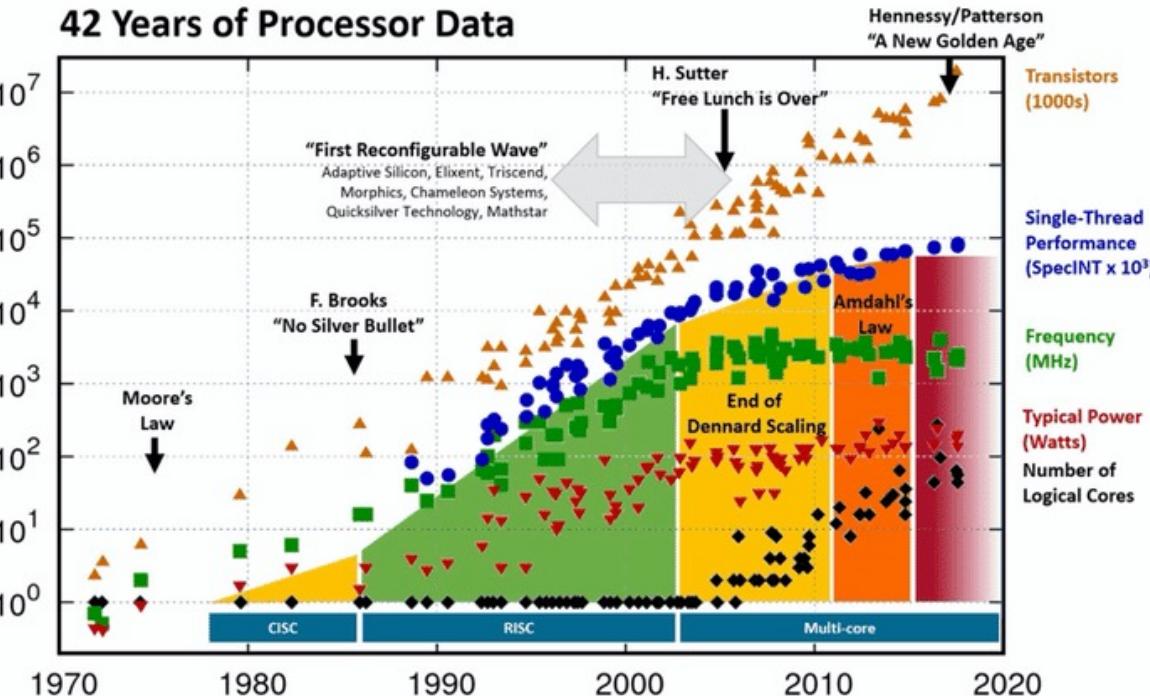
# 具身智能-哲学根源

- 梅洛-庞蒂 (Merleau-Ponty)：具身智能的哲学基础可以追溯到现象学家如梅洛-庞蒂的工作，他认为感知和认知本质上与身体及其在世界中的体验密切相关。他强调，我们对世界的理解是由我们与世界的身体互动所塑造的。
- 强调身体在感知和认知中的核心作用，认为知觉是通过身体与环境的互动直接发生的，而不是通过符号或概念的中介。



# 80年代的计算与智能

- 20世纪80年代摩尔定律驱动的计算机硬件性能的指数级增长，驱动人工智能的发展。
- 符号主义是20世纪50年代到80年代主流的人工智能方法，认为智能是符号处理和逻辑推理的结果，可以通过明确的规则和算法来实现。
- 符号主义者将智能视为一种信息处理过程，可以在没有身体的情况下实现。

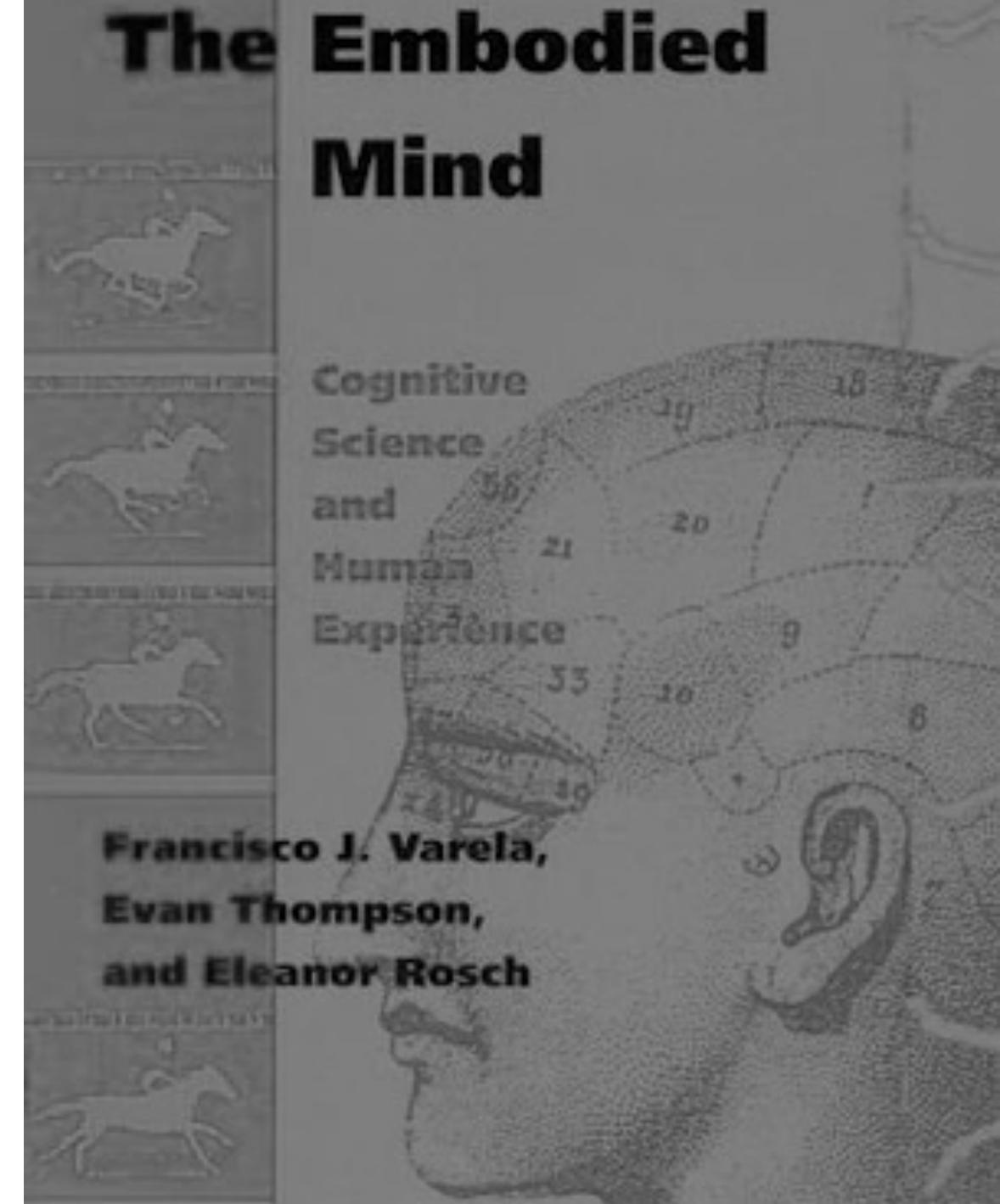


Hennessy and Patterson, Turing Lecture 2018, overlaid over "42 Years of Processors Data"  
<https://www.karlrupp.net/2018/02/42-years-of-microprocessor-trend-data/>; "First Wave" added by Les Wilson, Frank Schirrmeister  
Original data up to the year 2010 collected and plotted by M. Horowitz, F. Labonte, O. Shacham, K. Olukotun, L. Hammond, and C. Batten  
New plot and data collected for 2010-2017 by K. Rupp

# 具身智能-认知科学

- 具身认知 (Embodied Cognition) : 在20世纪70年代开始, 认知心理学的研究开始转向关注实际行动和感知。
- 在20世纪80年代, 认知科学家开始挑战以计算为主导的心智模型, 这些模型将大脑视为一种独立于身体的、处理符号的计算机。
- 埃莉诺·罗什 (Eleanor Rosch) 、弗朗西斯科·瓦雷拉 (Francisco Varela) 和翁贝托·马图拉纳 (Humberto Maturana) 等研究者提出, 认知是从大脑、身体与环境之间的动态互动中产生的, 从而推动了具身认知理论的发展。

F. J. Varela, E. Thompson, E. Rosch, "The Embodied Mind Cognitive Science and Human Experience", MIT Press, 1992



# 具身智能-机器人

- 罗德尼·布鲁克斯 (Rodney Brooks) : 在机器人领域，麻省理工学院的罗德尼·布鲁克斯在推动具身智能概念的发展方面发挥了关键作用。在1980年代和1990年代，布鲁克斯提出了行为基础机器人 (behavior-based robotics) 的理念，强调智能行为可以通过简单传感器和执行器与环境的互动中涌现出来，而不需要复杂的内部表征。
- Intelligence without Representation: 智能行为可以通过一系列直接的感知-行动循环实现，而不需要符号表示或内部模型。这一观点挑战了传统AI对智能的定义和实现方式。

R. A. Brooks, "Intelligence without representation," 1991



# 具身智能-2000年-至今

- **多模态整合和深度学习：**现代具身智能系统越来越多地使用深度学习技术来处理来自不同感官（视觉、听觉、触觉等）的数据。这些系统通过整合多模态数据来模仿人类的感知和行动能力，例如，机器人通过视觉和触觉传感器感知环境，并使用学习算法来规划和执行动作。
- **人机交互和社交机器人：**具身智能在开发社交机器人和人机交互系统方面发挥了重要作用。这些机器人通过模仿人类的社交行为和肢体语言，与人类进行更自然的互动。
- **增强现实和虚拟现实：**具身智能的概念也在增强现实（AR）和虚拟现实（VR）技术中得到应用。这些技术通过身体动作和环境互动，为用户提供沉浸式体验，体现了具身智能在虚拟环境中的应用。

# 具身智能-未来方向

- **更加自然的交互：**随着人工智能技术的不断发展，具身智能将继续推动更加自然和无缝的人机交互。未来的智能系统将能够更好地理解和适应人类的意图和情感，提供个性化和情境感知的服务。
- **身体与智能的融合：**未来可能会出现更加紧密融合的具身智能系统，这些系统将包括生物传感器、可穿戴设备和嵌入式智能，进一步模糊人类与技术之间的界限。
- **伦理和社会影响：**具身智能的发展也带来了伦理和社会方面的挑战，如隐私问题、自动化对就业的影响、以及机器人和AI系统在人类社会中的角色。这些问题将成为未来具身智能研究和应用的重要议题。

# Every humanoid robot 2024



视频网址：<https://youtu.be/PyrDh6RQdYY?feature=shared>

# 课程介绍

- 具身智能简介
- **课程相关信息**
- 课程内容综览



# 课程相关信息

- **目标：**本课程围绕具身智能这一新兴人工智能领域的核心理念和教学目标进行构建。
- **原理：**着重培养学生对具身智能的基本原理、机器学习与大模型技术、软硬件系统集成方法的理解和实际应用能力。
- **实践：**开发基于大模型技术的智能体通过集成传感器与执行器的智能交互系统的技能，通过设计实现与环境和人交互的具身智能系统原型获得实践经验。

# 课程相关信息

- AI通识基础课程 (AIB30002)
- 周课时 3, 总课时 54
- 授课时间: 每周一 (邯郸三教 H3308 8-10, 15:25-18:00) (邯郸三教 H3409 11-13, 18:30-21:05)
- 授课方式: 课堂讲授、文献阅读讨论与课程项目为主。学习具身智能领域的相关理论、技术和方法。结合课程项目安排学生运用课堂所学的理论技术与方法设计开发具身智能系统原型。
- 参考教材:
  - Embodied Artificial Intelligence, edited by Fumiya Iida, Rolf Pfeifer, Luc Steels, and Yasuo Kuniyoshi, Springer, 2004, 978-3-540-27833-7
  - Embodied Conversational Agents, Justine Cassell, Joseph Sullivan, Scott Prevost, and Elizabeth Churchill, MIT Press, 2000, 978-0262032780
  - How the Body Shapes the Way We Think, Rolf Pfeifer and Josh Bongard, MIT Press, 2006, 78-0262162395
- 考核和评价方式
- 作业30%
- 讨论与报告20%
- 课程项目、展示、论文50%

# 教学团队

姓名	所属院系	研究领域
尚笠	计算机科学技术学院	智能可穿戴、具身智能
商慧亮	信息学院	自主无人系统
丁文超	工研院	机器人决策规划
杨帆	微电子学院	智能芯片、大模型
李维	人工智能创新与产业研究院	仿生机器人、人形机器人
浦剑	类脑人工智能科学与技术研究院	自动驾驶

# 商慧亮

研究领域：自主无人系统

复旦大学信息科学与工程学院副教授，复旦大学博士。机器人与自主无人系统实验室负责人，主要研究方向围绕机器人与自主无人系统开展，涵盖3D视觉感知，自主导航、SLAM，云端机器人系统，机器人抓取与灵巧操作，工业视觉检测分析，深度强化学习、智能工业机器人、特种机器人等。课题组与美国卡耐基梅陇大学、加拿大多伦多大学、日本东京大学等相关研究课题组长期密切合作。商慧亮副教授先后承担或参与国家、省部级及产学研课题30余项，在国内外核心期刊发表论文50余篇。



# 丁文超

研究领域：机器人决策规划

工研院青年研究员。博士毕业于香港科技大学电子及计算机工程系，主要从事机器人的感知和预测决策方向研究。曾任职华为车BU智能驾驶产品部（ADS），获华为“天才少年”，担任ADS预测决策团队负责人，作为核心贡献者参与打造了华为ADS1.0以及ADS2.0智能驾驶系统，贡献了业界领先的智能驾驶预测方案，并实现在北汽极狐、长安阿维塔以及问界等多款车型上的商用。2023年加入复旦大学，入选上海市白玉兰浦江人才计划、小米青年学者等。近5年来在国内外核心会议期刊发表论文十余篇，获IEEE ROBIO 2023最佳论文提名奖。



# 杨帆

研究领域：智能芯片、大模型

复旦大学微电子学院教授。主要研究方向为集成电路设计自动化，智能芯片、大模型。获2012年上海市自然科学一等奖1项。获2018年 Integration, the VLSI Journal 最佳论文奖，获 DAC' 2014, ASPDAC' 2017, DAC' 2017 最佳论文提名。2018年获国家自然科学基金委优青项目资助。



# 李维

研究领域：仿生机器人、人形机器人

复旦大学人工智能创新与产业研究院研究员，复旦大学人工智能创新与产业研究院、计算机科学与技术学院、大数据学院博士生导师。上海交通大学工学博士。美国The University of Texas at Dallas博士后研究员。在国际顶刊（中科院SCI一区顶刊，Top 5%）上发表“第一作者与通讯作者”论文24篇、“独立作者”论文12篇，作为第一发明人有13项PCT国际发明专利、20多项国内发明专利。



# 浦剑

研究领域：自动驾驶

复旦大学类脑研究院青年研究员，博士生导师，复旦大学博士。已发表机器学习、自动驾驶等方面各类学术论文60余篇。近些年来，主要代理团队承担复旦大学与中国一汽集团、蘑菇车联等人工智能联合实验室重大研发任务，研发具有自主知识产权的车载多传感器融合感知、车路协同感知、基于语义的实时建图与定位系统、车辆预决策一体化模型等自动驾驶关键模块。



# 尚笠

研究领域：智能可穿戴、具身智能

复旦大学计算机科学技术学院教授。美国普林斯顿大学博士。曾任英特尔中国研究院副院长、首席架构师，美国科罗拉多大学博尔德分校电子计算机与能源工程系终身副教授。主要研究方向包括计算机系统、人机交互与普适计算、机器学习。在相关领域发表论文170余篇，多次荣获最佳论文奖与提名，被引逾8000次。曾获美国NSF CAREER Award资助。

邮箱：[lishang@fudan.edu.cn](mailto:lishang@fudan.edu.cn)



# 吴萍

智能人机交互与具身智能实验室科研助理。复旦大学硕士毕业于复旦大学管理学院管理科学系，主要负责实验室学生管理和科研项目管理等工作。曾获复旦大学优秀毕业生、研究生国家奖学金等奖项。

邮箱：ping\_wu@fudan.edu.cn



# 课程助教团队



许震宇  
联系方式: 15380961581  
研究领域: 视频理解和 LLM agent



黄和金  
联系方式: 19942130878  
研究领域: 具身智能、LLM agent



陈安瑞  
23110240001@m.fudan.edu.cn  
研究领域: 自动化代码生成



张睿  
联系方式: 15301130401 (微信同)  
研究领域: 嵌入式系统、具身智能、人机交互



陆天麟  
23210240250@m.fudan.edu.cn  
研究领域: 视觉语言模型



赵伊杨  
24212010050@m.fdu.edu.cn  
研究领域: 嵌入式系统



潘文韬  
i21110240007@m.fudan.edu.cn  
研究领域: 面向智能可穿戴设备的低功耗语音分析技术



陈子源  
联系方式: 13253392621  
研究领域: 具身智能, 视觉语言模型, 人机交互



任宇琪  
24210240043@m.fdu.edu.cn  
研究领域: 多模态大模型、音频信号处理



乔宇轩  
23212020135@m.fudan.edu.cn  
研究领域: 数字集成电路设计



全馨怡  
24210860014@m.fudan.edu.cn  
研究领域: 机器人决策规划



李京桥  
jingqiaoli24@m.fudan.edu.cn  
研究领域: 具身智能、计算机视觉



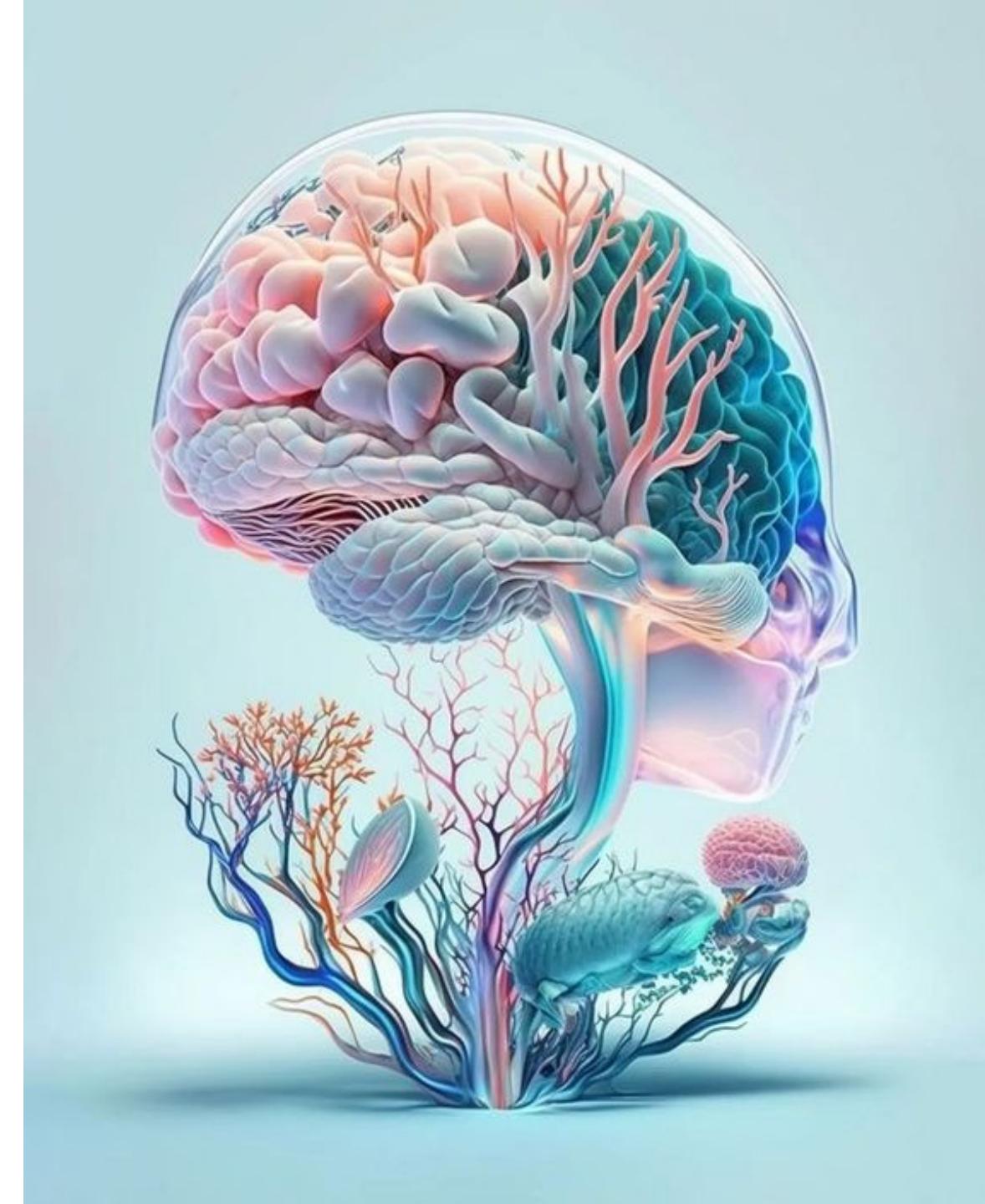
庄子涵  
ZZH-Fusion@outlook.com  
研究领域: 智能人机交互



杨昕烨  
24110980028@m.fudan.edu.cn  
研究领域: 仿生机器人

# 课程设计

**Build what you use, use what you  
build**



# 课程设计：智能陪伴机器人

·语音交互

·视觉理解

·机械控制

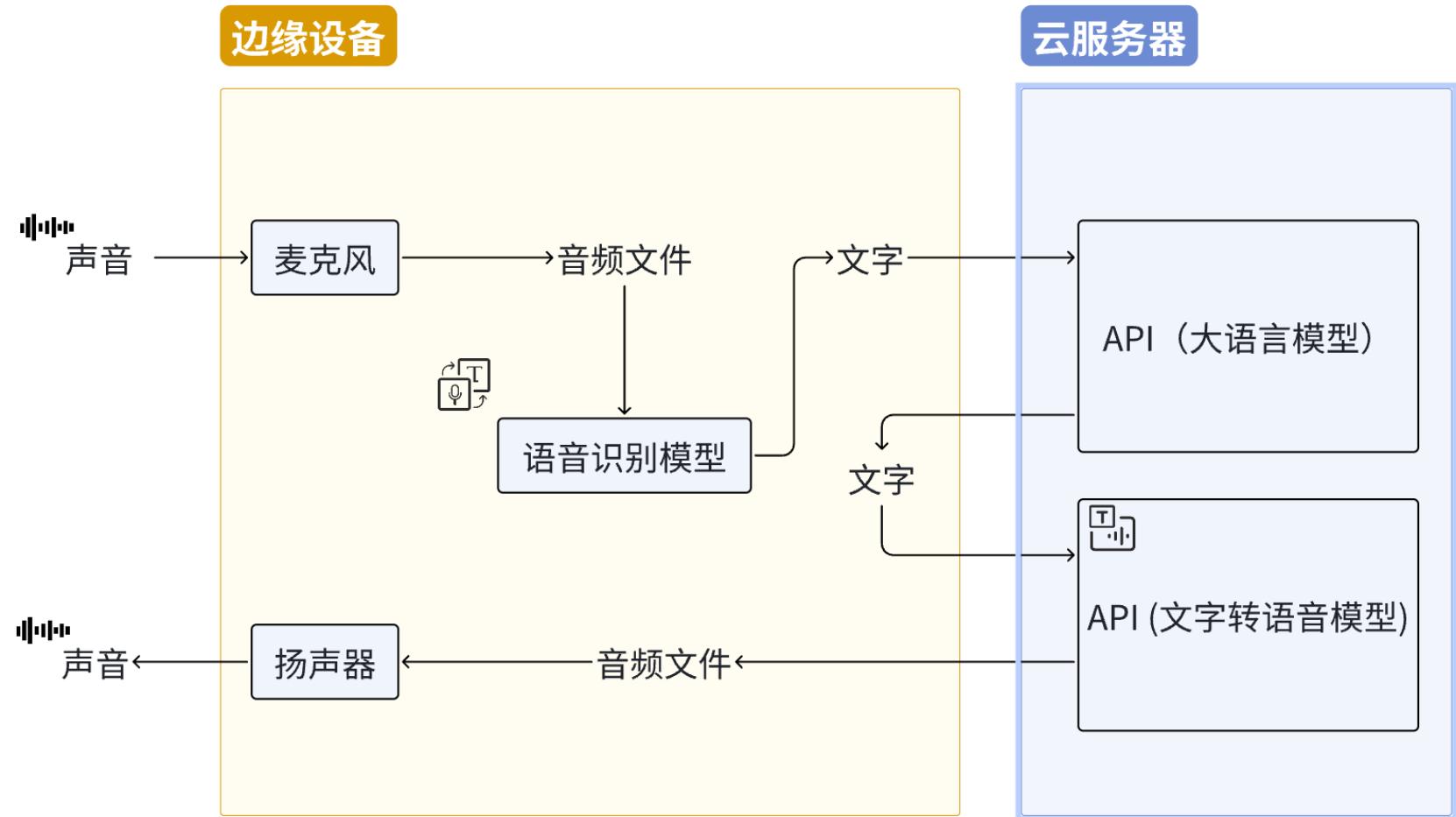


# 语音交互

## 音频采集与播放

## 语音识别

## 语音合成



# 视觉理解

场景识别

视觉问答

物体定位

深度估计

**Scene Caption:** "The image features a black laptop computer sitting on a black table. Next to the laptop, there is a glass of water and a coffee cup."



**Input:** "Get the white coffee cup."



**Object Segmentation:**  
- "where is the white coffee cup"

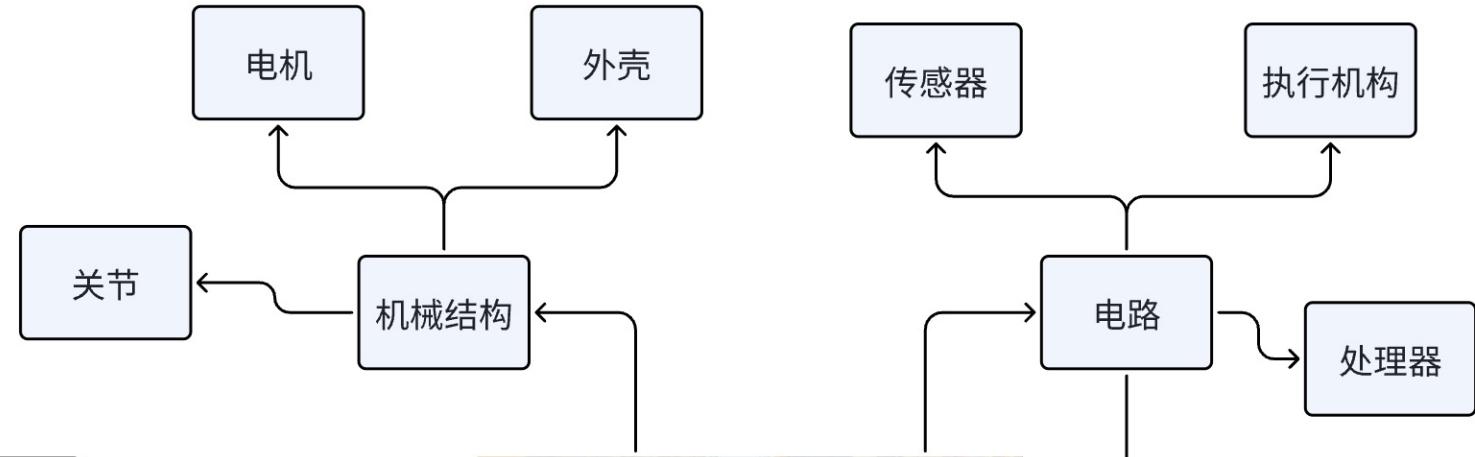
**Visual Question Answering**  
**Question:** Is the coffee cup white?  
**Answer:** Yes.

**Depth Estimation:** How far is it?

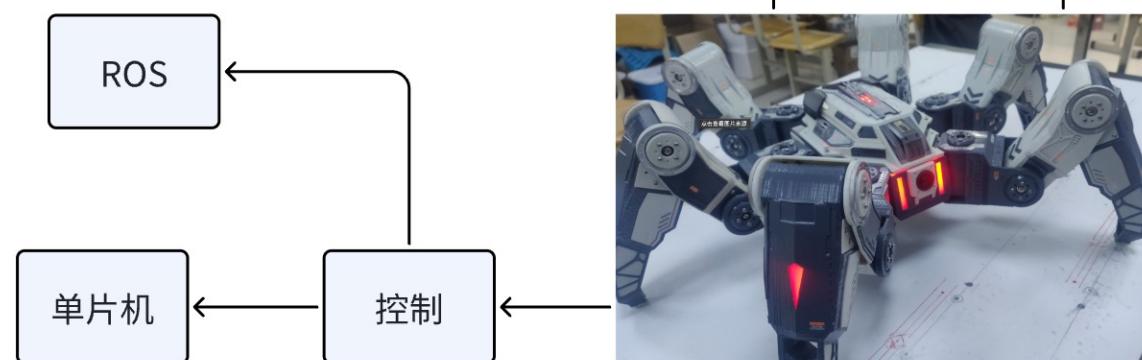


# 机械结构与控制

● 收集信息



● 与上位机通信



● 执行动作



时间安排	教学内容	内容负责人
第1周	具身智能简介	具身智能的定义 历史背景和关键技术发展 具身智能与传统人工智能的区别
第2周	伦理隐私安全	个人隐私与安全 安全技术与隐私保护 具身智能伦理挑战
第3-6周	机器学习基础	机器学习基础 强化学习 深度学习与大模型
第7-8周	感知技术	手势与动作识别 视觉认知技术 语音交互技术
第9-10周	感觉运动系统	各种类型的传感器（视觉、触觉、听觉等） 马达系统和执行器 机器人的感觉运动整合
第11-12周	认知架构	具身人工智能的认知架构概述 大模型智能体技术 记忆和学习与智能体认知 端到端：感知到动作
第13-14周	导航与制图	空间意识和SLAM（同时定位与制图）技术 路径规划和避障 导航与任务导向决策的整合
第15周	人机交互	脑机接口
第16周	群体智能	群体机器人 机器人与人类社会

时间安排	教学内容	内容负责人
第1周	具身智能简介	具身智能的定义 历史背景和关键技术发展 具身智能与传统人工智能的区别
第2周	伦理隐私安全	个人隐私与安全 安全技术与隐私保护 具身智能伦理挑战
第3-6周	机器学习基础	机器学习基础 强化学习 深度学习与大模型
第7-8周	感知技术	手势与动作识别 视觉认知技术 语音交互技术
第9-10周	感觉运动系统	各种类型的传感器（视觉、触觉、听觉等） 马达系统和执行器 机器人的感觉运动整合
第11-12周	认知架构	具身人工智能的认知架构概述 大模型智能体技术 记忆和学习与智能体认知 端到端：感知到动作
第13-14周	导航与制图	空间意识和SLAM（同时定位与制图）技术 路径规划和避障 导航与任务导向决策的整合
第15周	人机交互	脑机接口
第16周	群体智能	群体机器人 机器人与人类社会

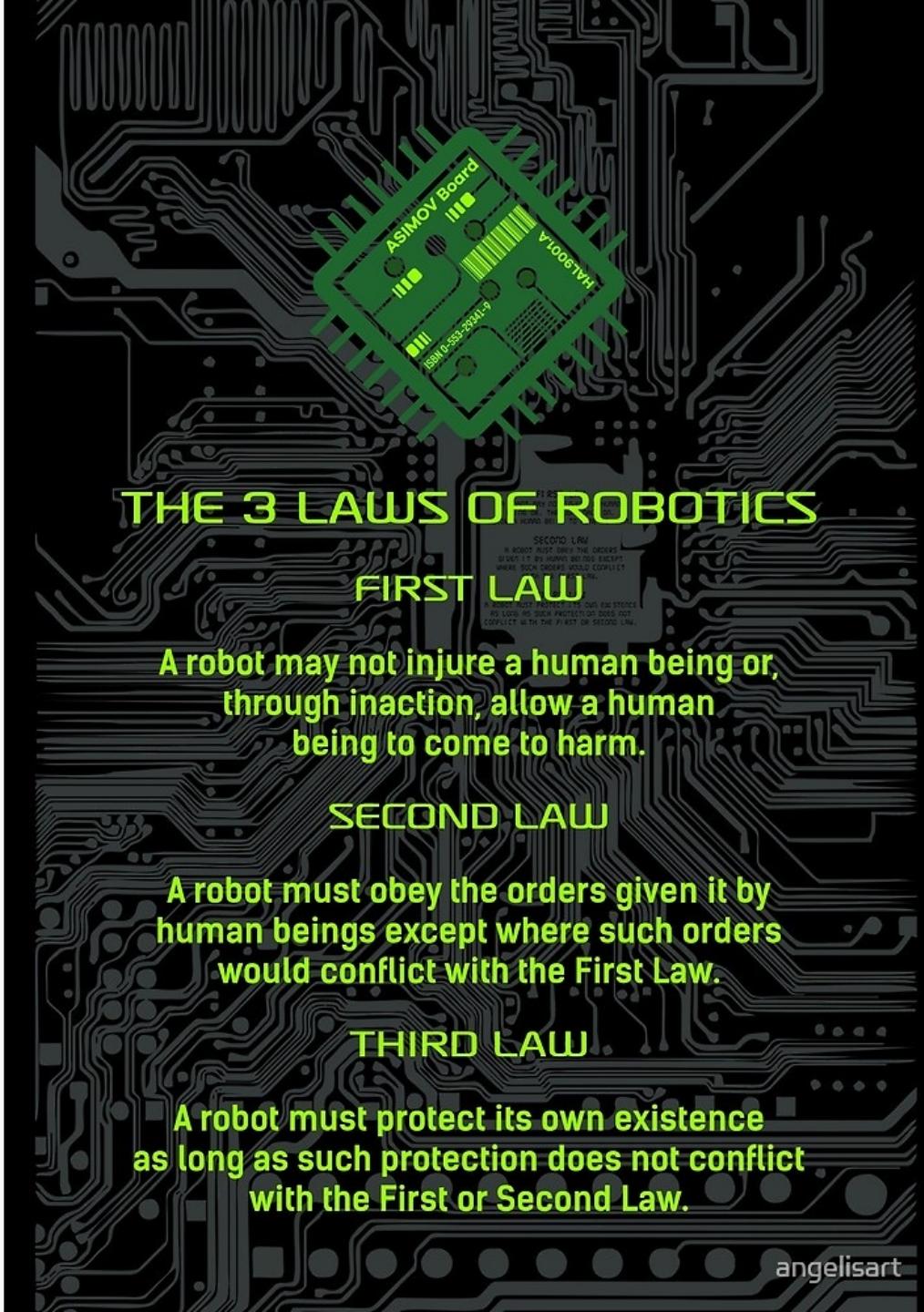
# 伦理隐私安全

- **伦理挑战：**具身智能系统决策时具有高度自主性，在自主性与责任、价值对齐、人机互动等方面面临伦理挑战。例如，机器人如何在紧急情况下做出决策？谁对机器人的行为负责？需要建立相应的伦理框架和法律法规来规范机器人的使用。
- **隐私挑战：**具身智能系统通常依赖大量数据来感知环境和学习，在数据收集与使用、持续监控与个人空间、数据匿名化与去识别化等方面面临隐私挑战。例如服务机器人、智能家居和监控系统的普及增加了隐私泄露和数据安全的风险。这些机器人在工作过程中可能会收集大量个人信息，因此确保数据安全和隐私保护成为一个关键问题。
- **安全挑战：**具身智能设备通常连接到互联网，并能够直接与物理世界互动，在网络安全、物理安全、误操作与滥用方面面临安全挑战。



# 伦理

- **自主性与责任：**具身智能系统（如自主机器人和无人驾驶汽车）在决策时具有高度自主性。如果这些系统做出有害或错误的决定，如何界定责任？是设计者、制造商、用户还是机器人本身？在这些情况下，明确责任归属和道德责任是一个复杂的问题。
- **价值对齐问题：**具身智能系统需要在决策时考虑人类的价值观和伦理标准。但这些标准在不同文化和社会中可能存在差异。因此，确保这些系统的决策与人类社会的价值观对齐是一个重要的伦理挑战。
- **人机互动伦理：**随着具身智能设备的普及，机器人在医疗、教育、护理等领域的应用越来越广泛。这些场景中，机器人的行为需要符合伦理规范，如尊重患者的隐私、关爱儿童的成长等。机器人在这些敏感场景中与人类的互动需要仔细设计，以避免伦理上的冒犯或伤害。



# THREE LAWS OF ROBOTICS

- A robot may not injure a human being or, through inaction, allow a human being to come to harm
- A robot must obey the orders given it by human beings except where such orders would conflict with the First Law
- A robot must protect its own existence as long as such protection does not conflict with the First or Second Laws

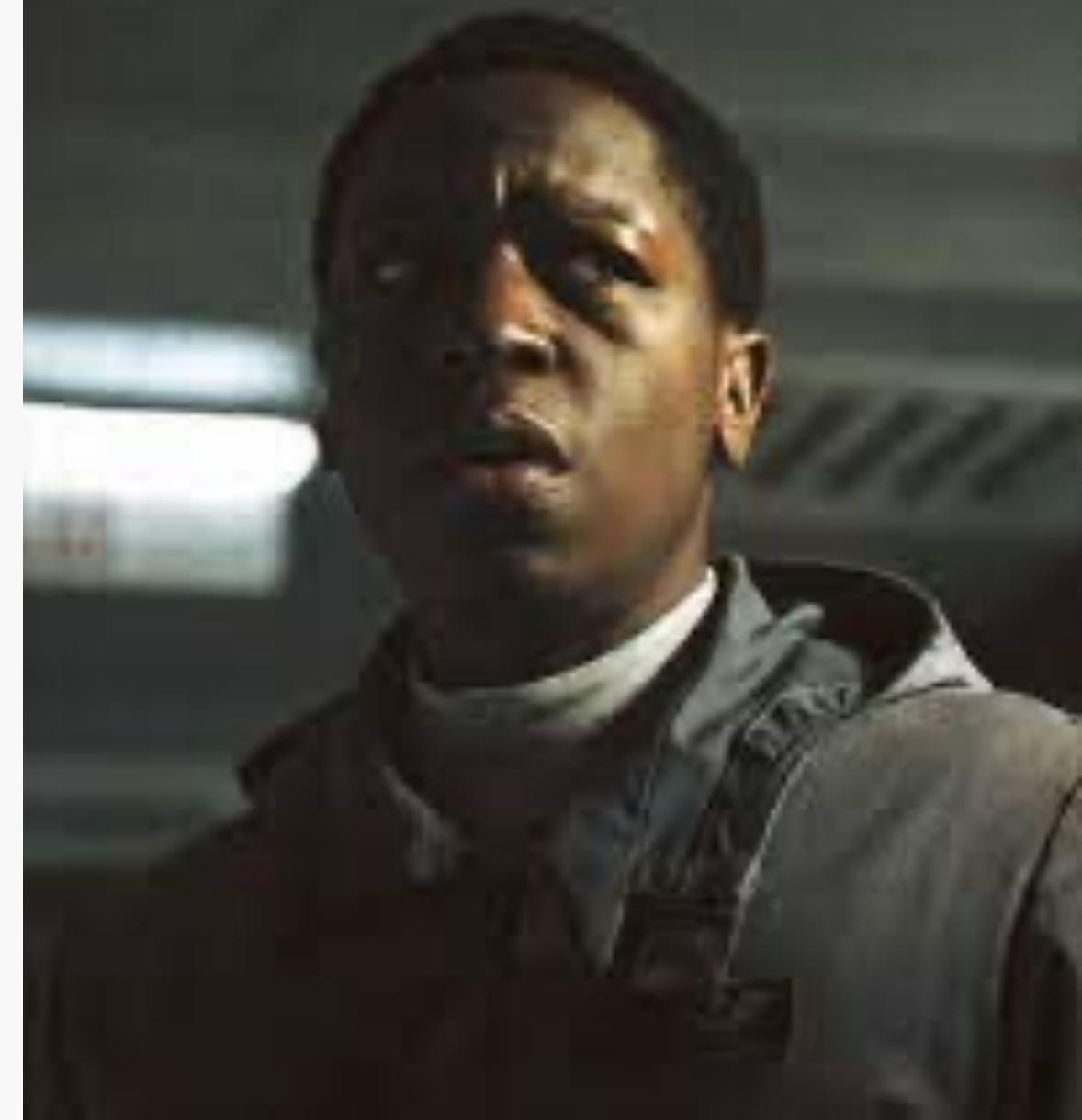
## ZEROTH LAW

A robot may not harm humanity, or, by inaction, allow humanity to come to harm

ISAAC ASIMOV

ALIEN  
ROMULUS

ONWARD  
AUGUST 16



# 伦理

- 伦理道德准则无法量化度量
- 具身智能系统需要量化指标
- 伦理道德与社会文化相关
- 伦理的理性与非理性悖论

## 人工智慧電車難題

你設計了一台擁有自我意識的機器人。  
它碰上了電車難題。  
你需要為他的選擇擔負道德責任嗎？



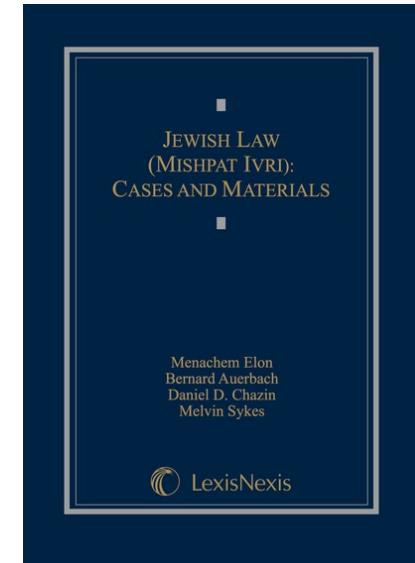
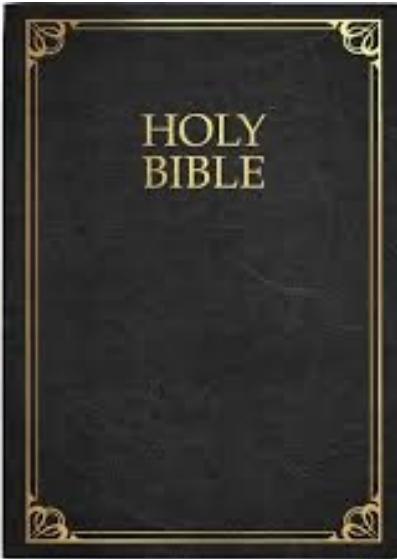
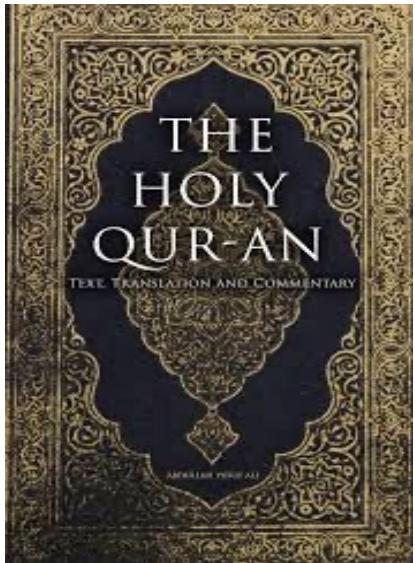
# 隐私

隐私是指个人、群体或机构自主决定何时、如何以及在多大程度上将有关他们的信息传递给其他人的权利。

Privacy is the claim of individuals, groups, or institutions to determine for themselves when, how, and to what extent information about them is communicated to others.

-- Alan Westin 1960s

# 隐私：贯穿人类历史



在《古兰经》和穆罕默德的言论中有对隐私的承认。《圣经》中也有许多关于隐私的提及。犹太法律早已承认免于被监视的概念。在古希腊和古代中国也存在对隐私的保护。

There is recognition of privacy in the Qur'an and the sayings of Mohammed. The Bible has numerous references to privacy. Jewish law has long recognized the concept of being free from being watched. There were also protections in classical Greece and ancient China.

# 隐私：数字时代

伊士曼柯达公司 (Eastman Kodak Company) 的创始人乔治·伊士曼 (George Eastman) 因发明现代照相胶卷而获得了美国专利#306 594。以前人们必须在摄影棚里使用笨重的玻璃板，而现在大家可以拿着柯达的“快拍相机”(Snap Camera) 走上街头，未经他人同意就可以随意拍摄他们的照片。(1884年)

**SPORTS AND PASTIMES** 47

**The Kodak Camera.**

**A** NYBODY who can wind a watch can use the Kodak Camera. It is a magazine camera, and will make 100 pictures without reloading. The operation of taking the picture is simply to point the camera and press a button. The picture is taken instantaneously on a strip of sensitive film, which is moved into position by turning a key.

\* \* \* A DIVISION OF LABOR. \*

After the 100 pictures have been taken, the strip of film (which is wound on a spool) may be removed, and sent by mail to the factory to have the pictures finished. Any amateur can finish his own pictures, and any number of duplicates can be made of each picture. A spool of film to reload the camera for 100 pictures costs only \$2.00.

No tripod is required, no focusing, no adjustment whatever. Rapid rectilinear lens. The Kodak will photograph anything, still or moving, indoors or out.

\* \* \* A PICTURESQUE DIARY \*

Of your trip to Europe, to the mountains, or the sea-shore, may be obtained without trouble with a Kodak Camera, that will be worth a hundred times its cost in after years.

\* \* \* A BEAUTIFUL INSTRUMENT \*

Is the Kodak, covered with dark Turkey morocco, nickel and lacquered brass trimmings, inclosed in a neat sole leather carrying case with shoulder-strap—about the size of a large field-glass.

Sent for a copy of the KODAK PRIMER with Kodak photograph.

**THE EASTMAN DRY PLATE and FILM CO.**  
Branch, 115 Oxford Street, London.  
ROCHESTER, N. Y.

**THE**  
**Kodak Camera.**  
ANYBODY CAN USE IT.

Photography reduced to three motions.

Silver Medal, Minneapolis Convention, P. A. of A., for most important invention for the year.

1.—Pull the Cord.  
2.—Turn the Key.  
3.—Press the Button.  
And so on for one hundred pictures.

One hundred shots before reloading.

Size of Camera,  $3\frac{1}{4} \times 3\frac{3}{4} \times 6\frac{1}{2}$  inches.  
Size of Picture,  $2\frac{5}{8}$  inches diameter.  
Weight of Camera, 25 ounces.

PRICE, LOADED, . . . . . \$25.00.

Extra spools film for 100 pictures, . . . . . \$2.00  
Amateurs can finish their own pictures, or the exposed film can be sent to the factory, by mail, to be developed and pictures finished. Price for 100 finished pictures, including spool of 100 films, for reloading, . . . . . 10.00

115 OXFORD STREET,  
LONDON.

ROCHESTER, N. Y.

# 隐私：泛在计算时代

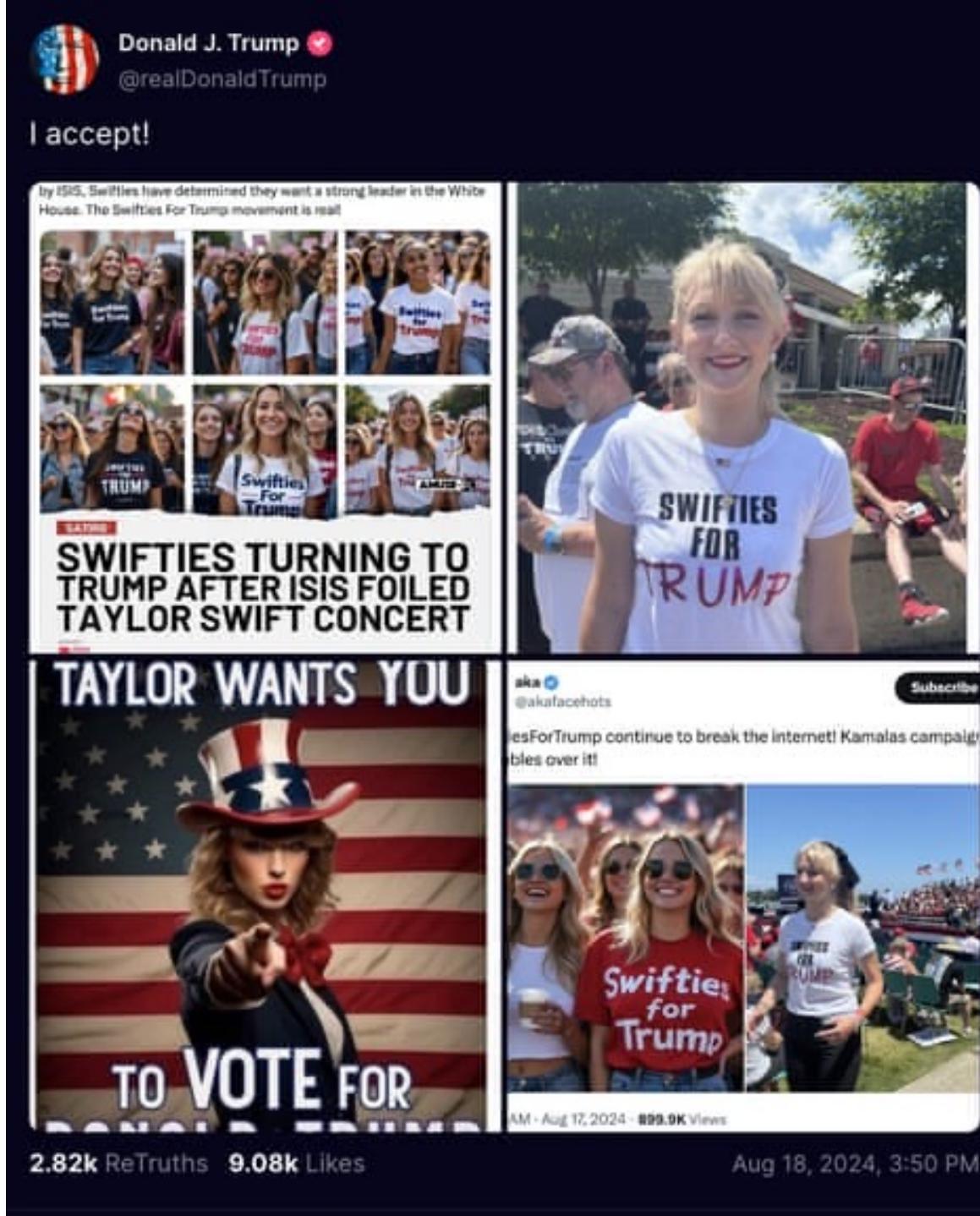
也许最关键的是隐私问题：每个房间里都有数百台电脑，它们都能够感知附近的人，并通过高速网络相连，这种情况可能会使过去的极权主义看起来像纯粹的无政府状态。

Perhaps key among is privacy: hundreds of computers in every room, all capable of sensing people near them and linked by high-speed networks, have the potential to make totalitarianism up to now seem like sheerest anarchy.

-- Mark Weiser

# 隐私：生成式AI时代

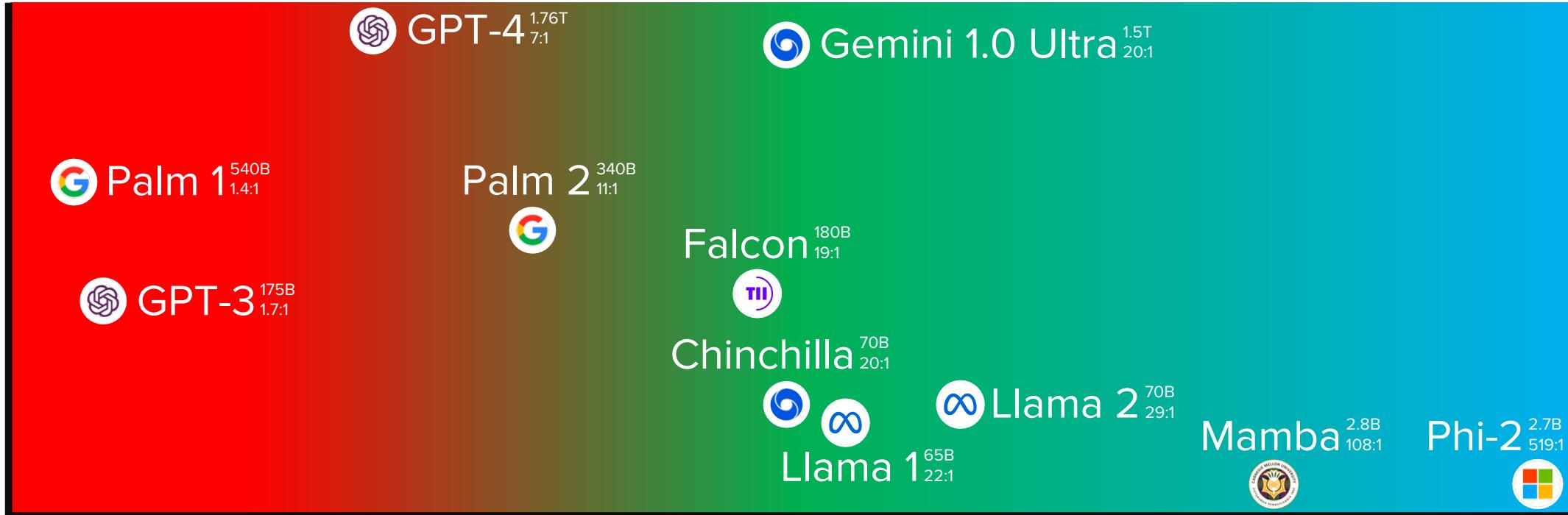
- 数据泄露与滥用
- 数据来源的合法性
- 深度伪造和虚假信息
- 透明性问题
- 数据中毒攻击
- 持久数据存储和删除权



# DATA-OPTIMAL (CHINCHILLA) MODEL HEATMAP

DEC/  
2023

Parameters (B)



**2:1**

Significantly undertrained  
(Not enough data)

**20:1**

Ratio of tokens trained to parameters

**200:1**

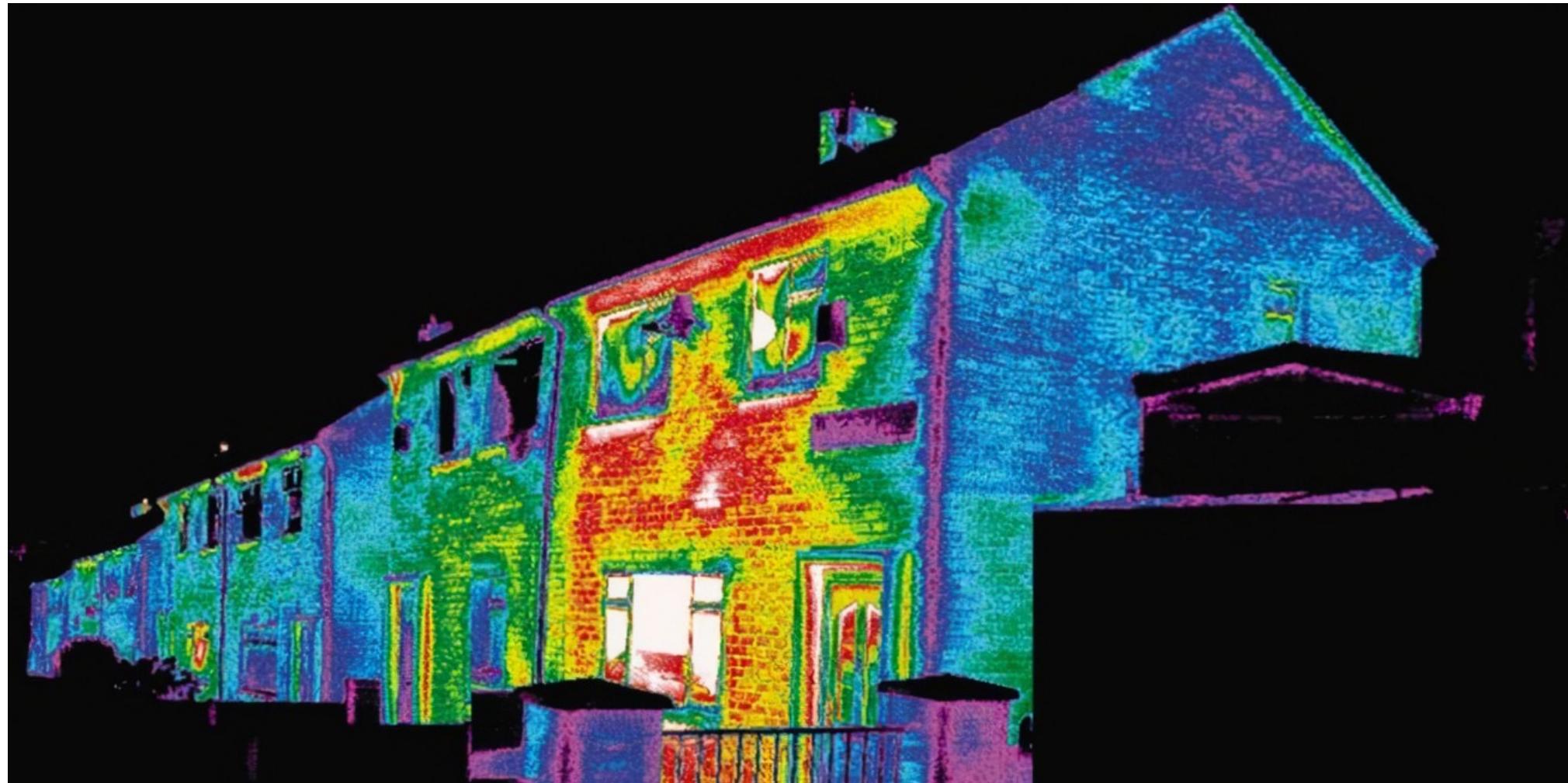
Potentially overtrained  
(Too much data)

Selected highlights only. Mostly to scale. Informed estimates for Palm 2, GPT-4, and Gemini. Alan D. Thompson. November 2022, major update December 2023. <https://lifearchitect.ai/>



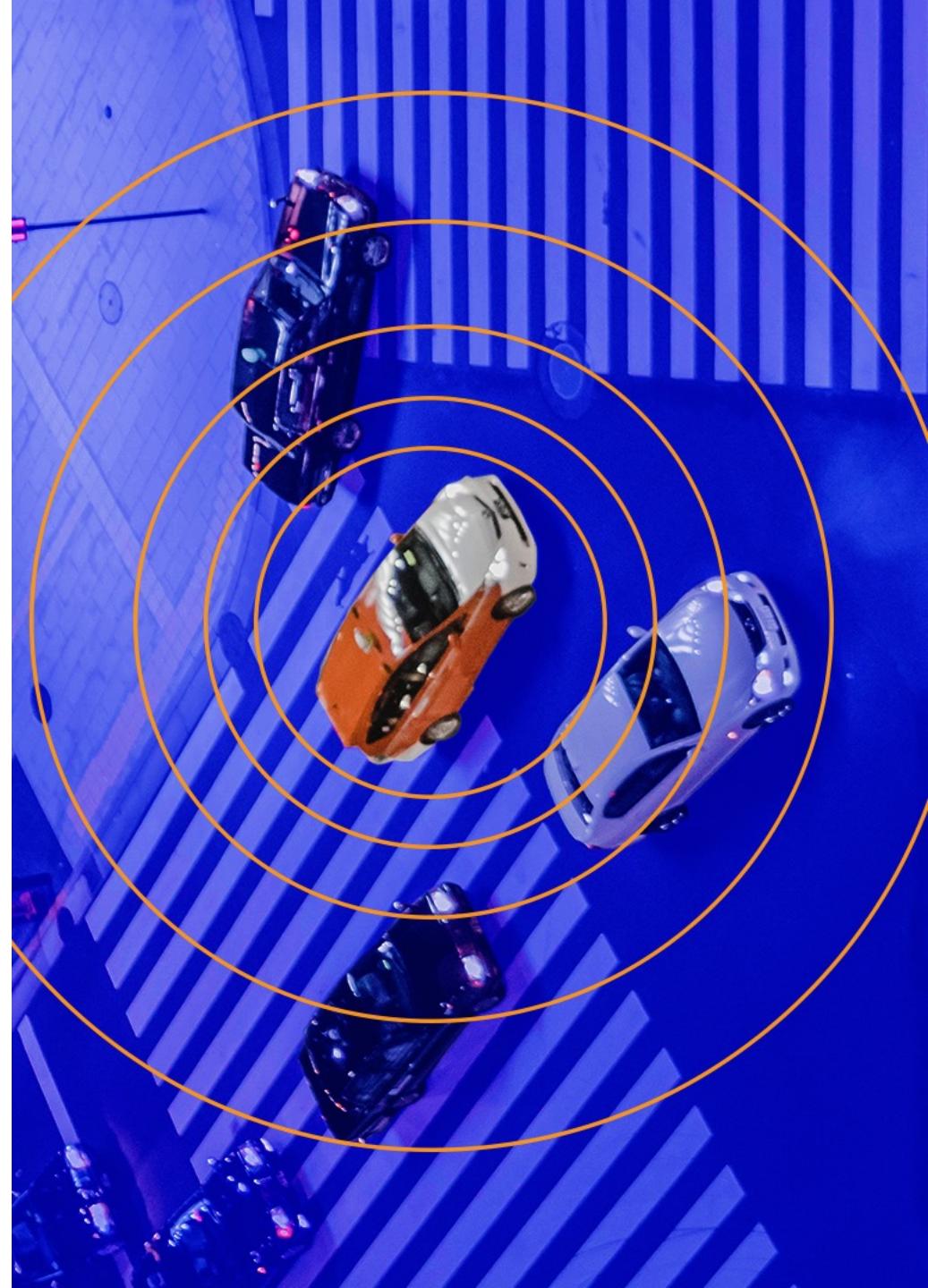
LifeArchitect.ai/chinchilla

# Kyllo案件



# 安全

- **网络安全威胁**: 具身智能设备通常连接到互联网，这使它们容易受到网络攻击。黑客可能通过远程入侵控制机器人，导致设备失控或用于恶意目的。确保这些系统的网络安全性是防止恶意行为的重要一环。
- **物理安全**: 具身智能系统能够直接与物理世界互动，因此它们的不当行为或故障可能会对人类和财产造成实际的物理伤害。例如，制造车间的机器人手臂失控可能会导致工人受伤。设计和部署安全机制以防止这些情况的发生是至关重要的。
- **误操作与滥用**: 具身智能设备可能被误操作或恶意滥用，从而导致意外的安全问题。例如，家庭机器人被利用来获取家庭成员的安全信息。开发安全使用指南和防滥用机制是必要的，以减少这些风险。



# Tesla FSD



视频网址: <https://youtu.be/tIThdr3O5Qo?feature=shared>

时间安排	教学内容	内容负责人
第1周	具身智能简介	具身智能的定义 历史背景和关键技术发展 具身智能与传统人工智能的区别
第2周	伦理隐私安全	个人隐私与安全 安全技术与隐私保护 具身智能伦理挑战
第3-6周	机器学习基础	机器学习基础 强化学习 深度学习与大模型
第7-8周	感知技术	手势与动作识别 视觉认知技术 语音交互技术
第9-10周	感觉运动系统	各种类型的传感器（视觉、触觉、听觉等） 马达系统和执行器 机器人的感觉运动整合
第11-12周	认知架构	具身人工智能的认知架构概述 大模型智能体技术 记忆和学习与智能体认知 端到端：感知到动作
第13-14周	导航与制图	空间意识和SLAM（同时定位与制图）技术 路径规划和避障 导航与任务导向决策的整合
第15周	人机交互	脑机接口
第16周	群体智能	群体机器人 机器人与人类社会

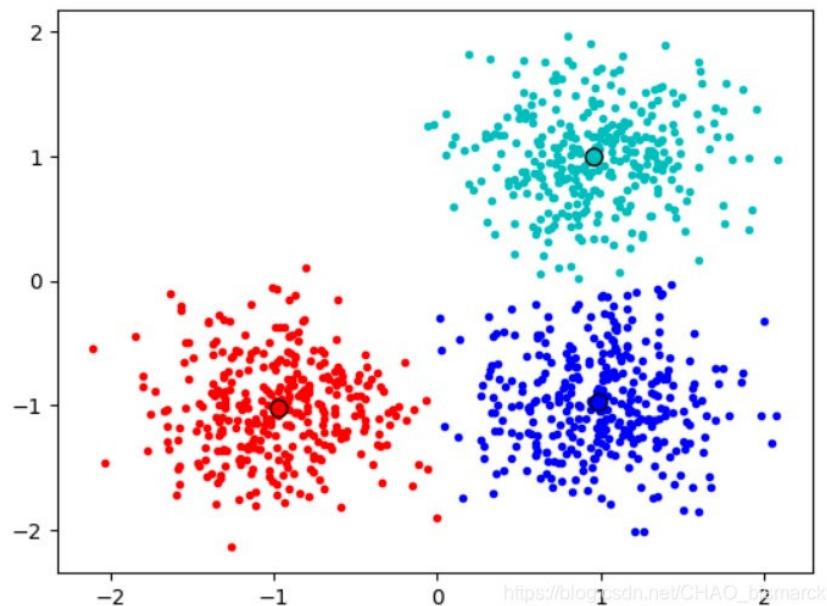
# 机器学习

机器学习（Machine Learning）是一种通过算法和统计模型，让计算机系统从数据中自动学习和改进的方法，而无需依赖于明确的编程指令。它是人工智能（AI）的一个子领域，旨在使机器能够从经验中学习并做出决策。

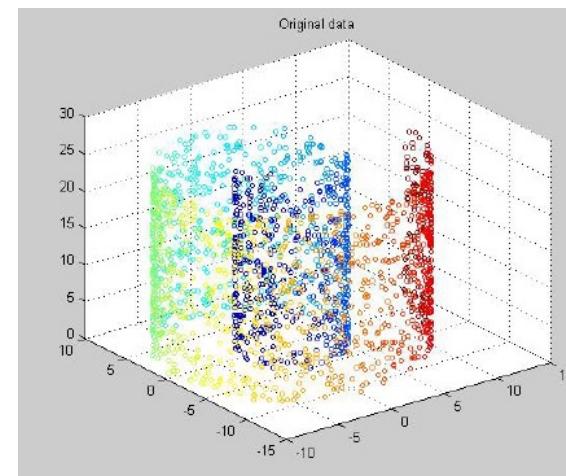
- 无监督学习：聚类、降维等
- 监督学习：分类、回归等
- 深度学习：卷积神经网络、循环神经网络
- 强化学习：通过与环境交互来获得最大收益

# 无监督学习：聚类与降维

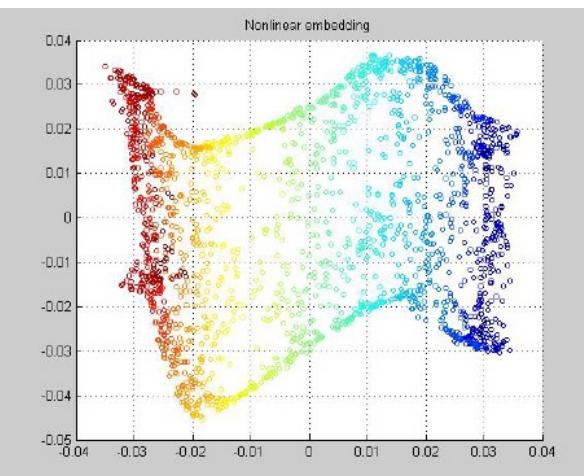
无监督学习无需对数据进行标记，借助数据特征在空间中的分布，实现对数据的特征分析和提取



数据的聚类



数据的降维



# 监督学习—回归与分类

- 监督学习给定了数据的标记，利用标记来训练模型，使得模型的输出与标记趋向一致
- 线性回归：给定数据 $\{x_1, x_2 \dots, x_n\}$ 及回归值 $\{y_1, y_2 \dots, y_n\}$ ，其中 $x_i \in R^m$ ，求得系数向量 $\alpha \in R^m$ ，使得回归误差的二范数最小

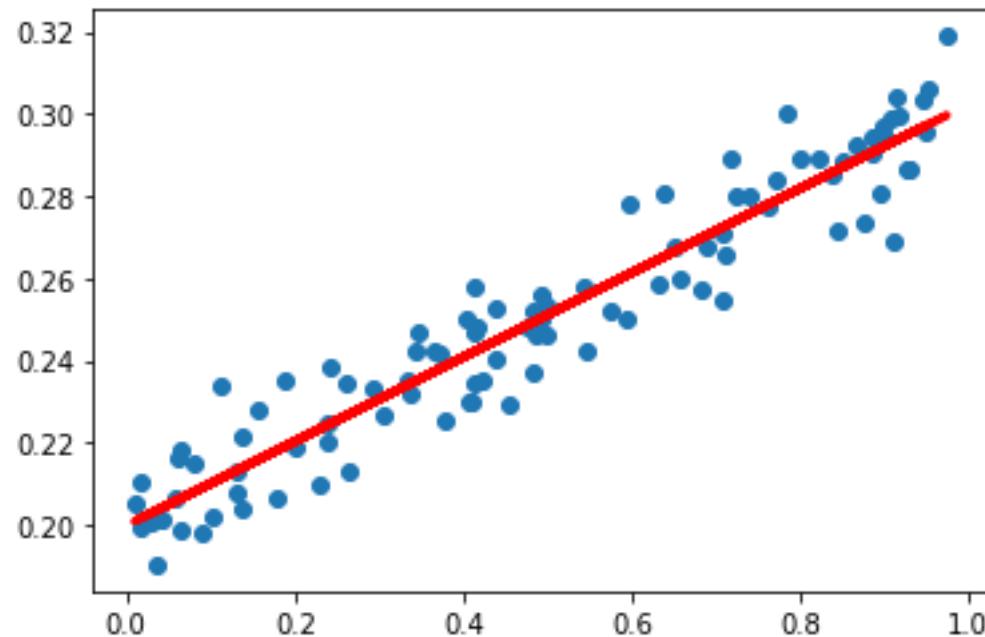
$$\|y - X\alpha\|_2^2$$

- 其中 $X = [x_1^T, x_2^T \dots, x_n^T]^T$ ,  $y = [y_1, y_2 \dots, y_n]^T$ .
- SVM线性二分类：给定数据 $\{x_1, x_2 \dots, x_n\}$ 及二分类标签 $\{y_1, y_2 \dots, y_n\}$ ，其中 $y_i \in \{-1, 1\}$ ，寻找超平面 $w^T x + b = 0$ ，使得该超平面与所有样本点距离的最小值最大化，假定 $\|w\| = 1$ .

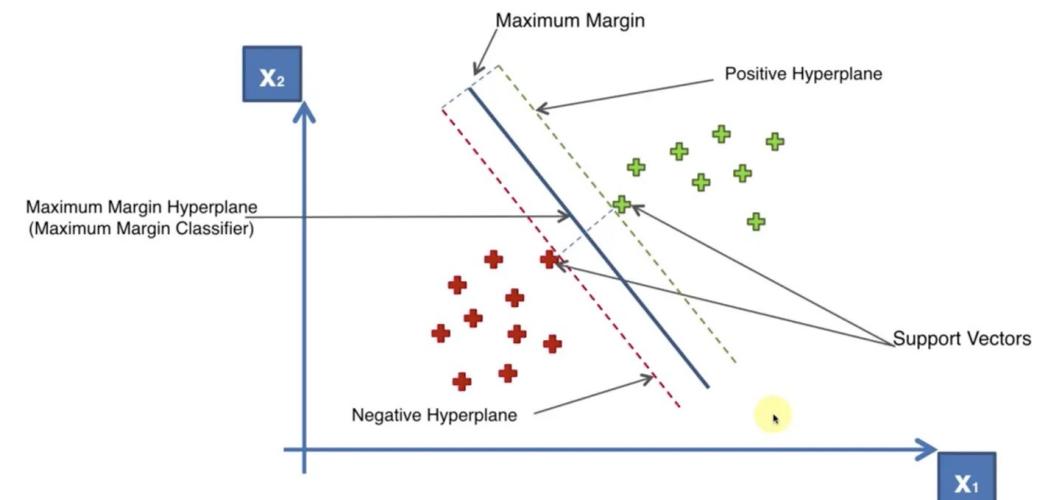
$$\begin{aligned} & \max \min \gamma_i \quad i = 1, \dots, n \\ \text{s.t.} \quad & y_i(w^T x_i + b) \geq \gamma_i, i = 1, \dots, n \end{aligned}$$

第*i*个数据点到超平面的最小距离

# 监督学习—回归与分类



线性回归



支持向量机分类

# 深度学习

深度学习（Deep Learning）是机器学习的一个子领域，它通过模拟人脑的神经网络结构来自动学习和提取数据中的特征和模式。深度学习特别擅长处理大规模和复杂的数据集，如图像、视频、文本和语音等，在具体智能应用中得到了广泛的应用。

## ● 深度学习的挑战和应对方法

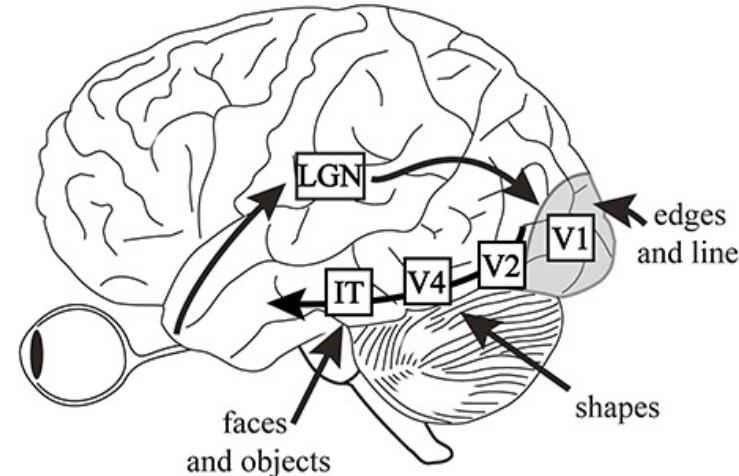
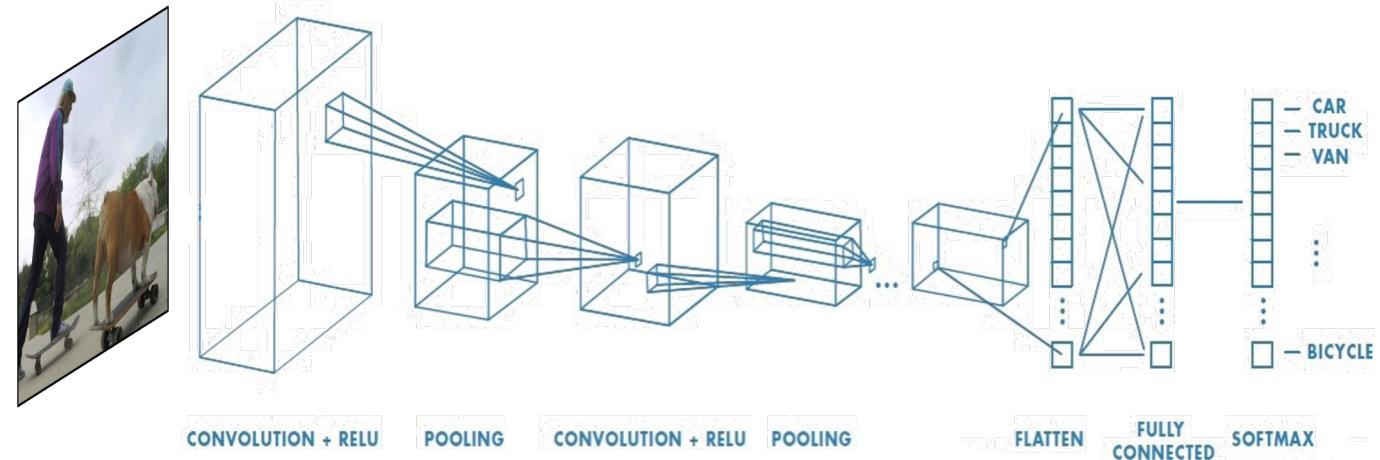
- 大量标记数据：众包标记、预训练模型
- 训练巨大计算量：GPU/NPU/TPU等SIMD处理器
- 梯度消失/爆炸问题：新的激活函数/残差网络结构

## ● 不同的深度网络：

- 卷积神经网络：图像结构化数据的固定相关性提取
- 循环神经网络：对序列化数据进行表征建模

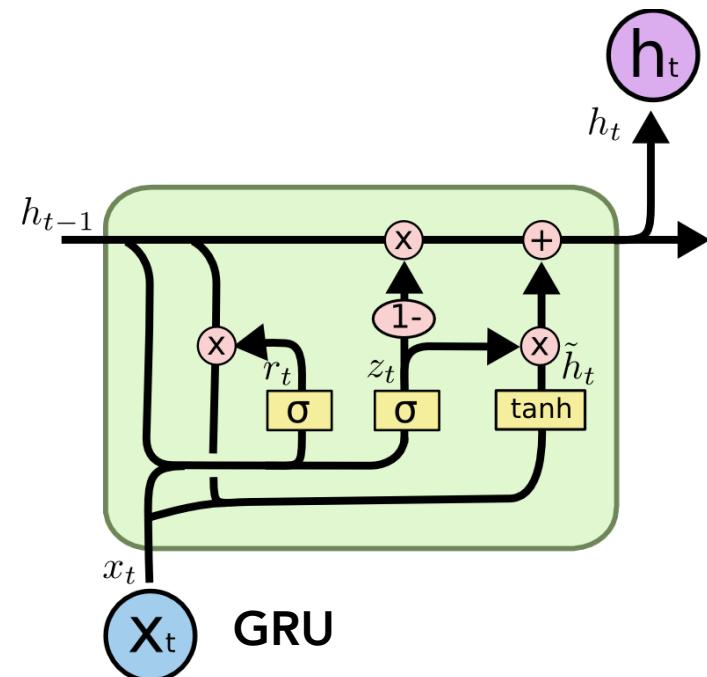
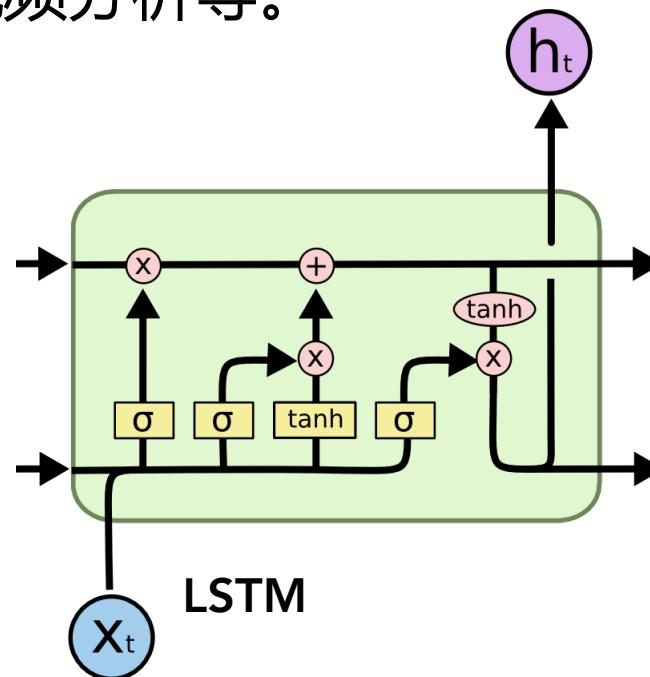
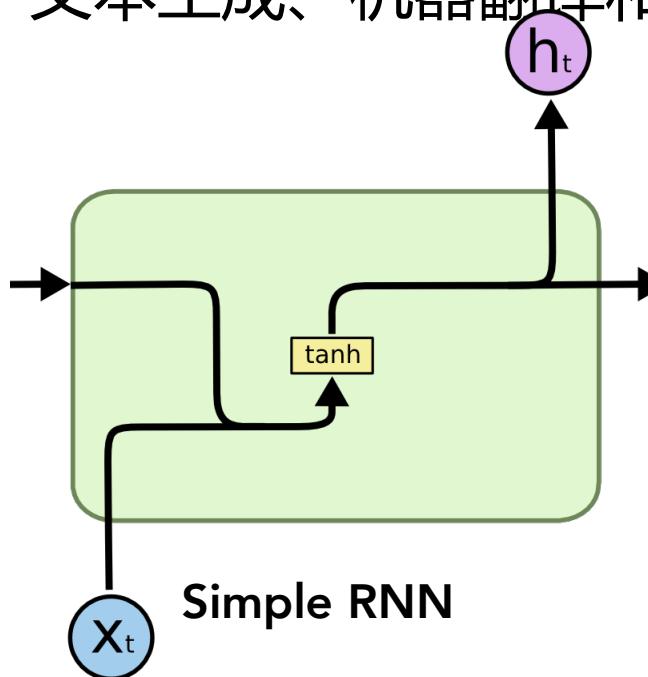
# 卷积神经网络

逐层利用卷积提取相邻格点特征，并通过激活和池化过滤后进入下一层，获得更高层次特征，最终根据特征通过全连接层输出分类结果



# 循环神经网络

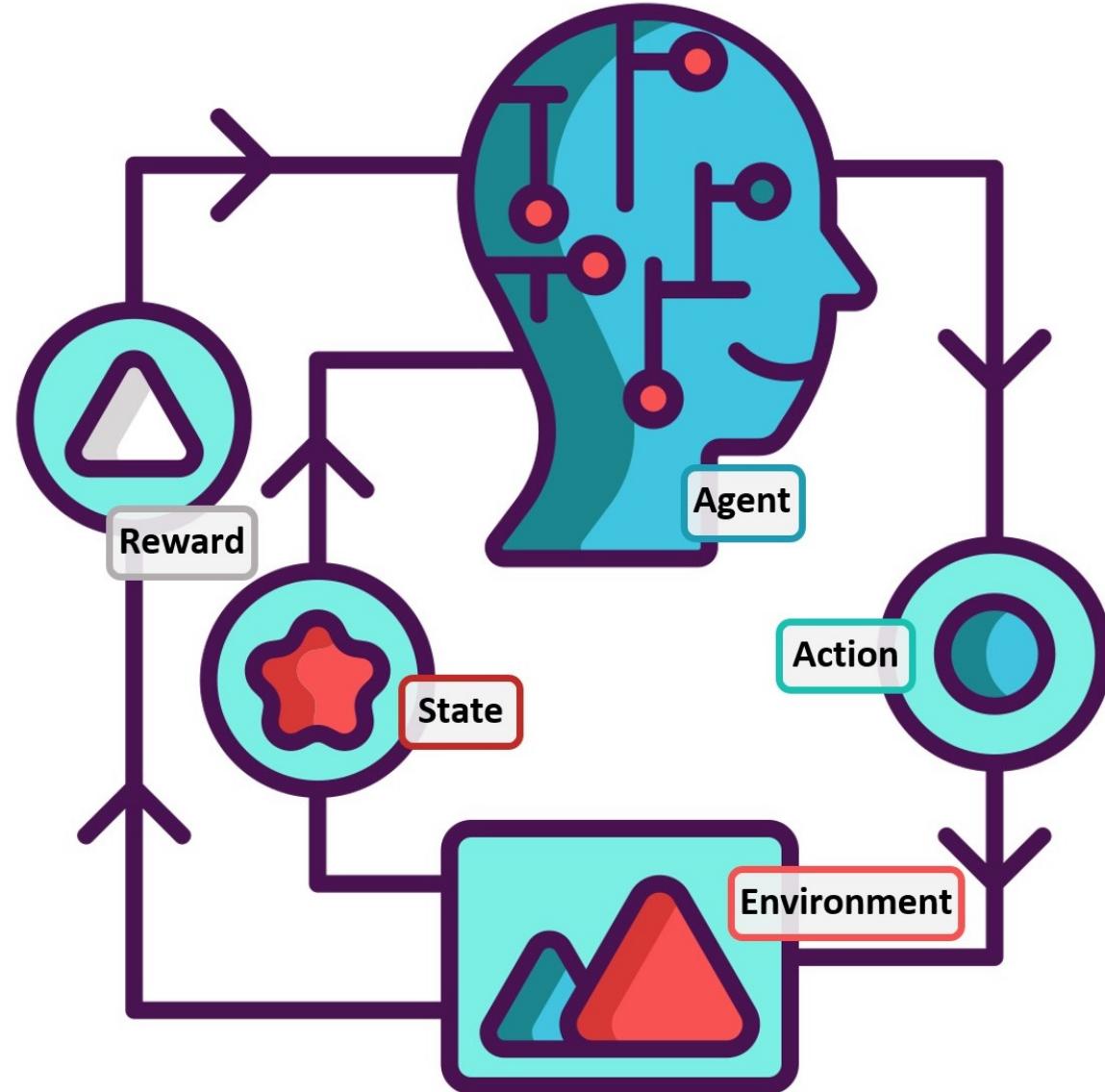
**循环神经网络 (Recurrent Neural Networks, RNN)** 是一种擅长处理序列数据的人工神经网络模型。与传统的前馈神经网络不同，循环神经网络具有循环连接，能够记住先前的信息，并利用这些信息对当前的输入进行处理。这使得RNN特别适合于处理涉及时间序列的任务，如语音识别、语言建模、文本生成、机器翻译和视频分析等。



# 强化学习

强化学习 (Reinforcement Learning, RL) 是一种机器学习方法，通过与环境的交互来学习如何采取行动，以最大化累积奖励。

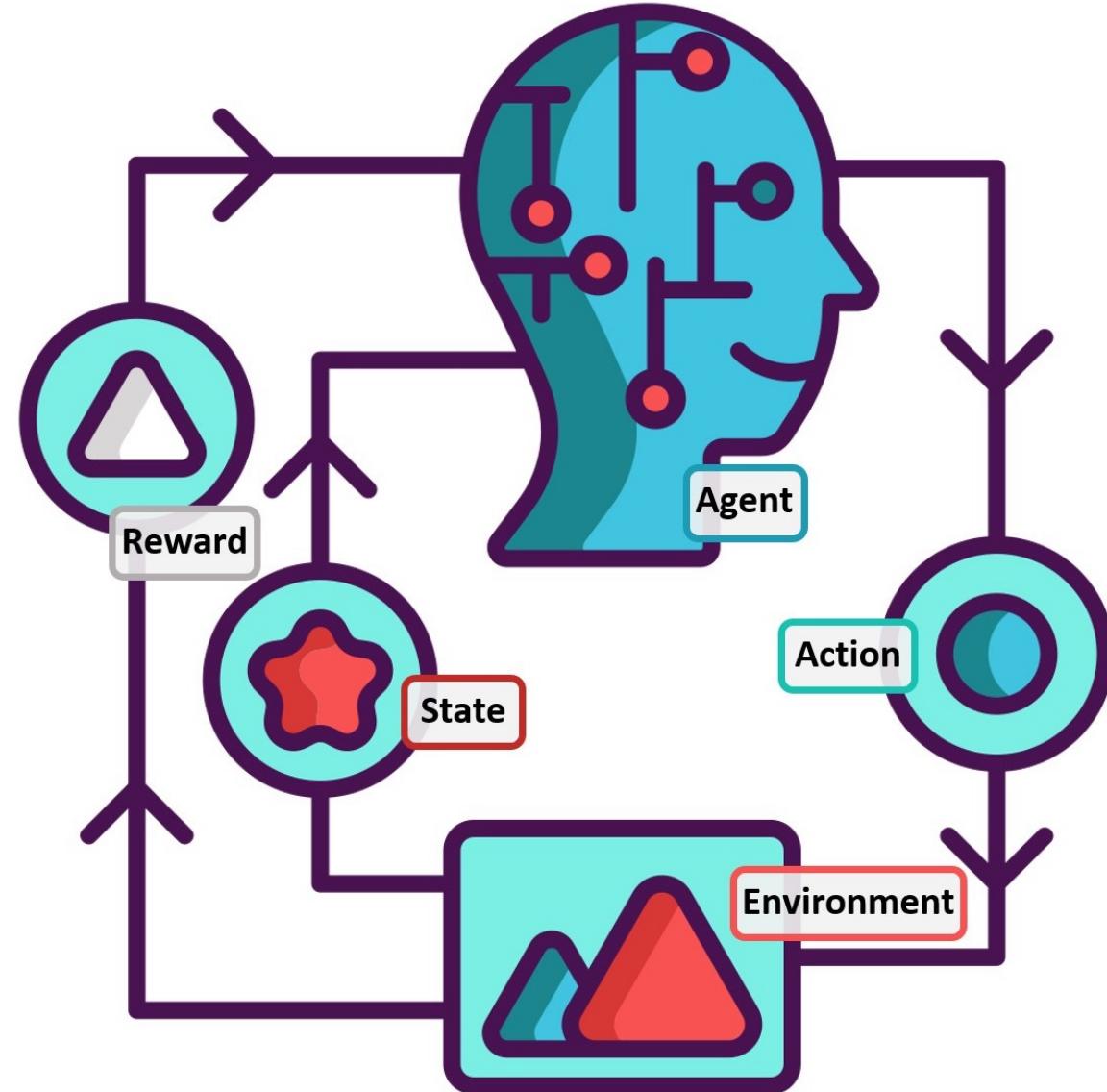
- **智能体 (Agent)**：智能体是执行动作的学习者，在给定环境中作出决策。智能体的目标是学习一个策略，以便在特定情境下选择最佳行动。
- **环境 (Environment)**：智能体所处的外部世界。环境根据智能体的动作产生反馈，提供新的状态和奖励。环境可以是任何场景，如棋盘游戏、机器人操作或自动驾驶。
- **状态 (State)**：环境的特定配置或描述，代表智能体在给定时间点的观察或信息。状态决定了智能体可以采取哪些动作。
- **动作 (Action)**：智能体在给定状态下可以执行的行为或操作。每个动作会导致环境状态的变化。
- **奖励 (Reward)**：智能体在采取某个动作后从环境中获得的反馈信号。奖励用于评估该动作的好坏，指导智能体的学习过程。目标是最大化累积奖励。
- **策略 (Policy)**：策略是智能体在每个状态下选择动作的规则或函数。策略可以是确定性的（每个状态对应一个动作）或随机的（在概率意义上选择动作）。



# 强化学习

强化学习 (Reinforcement Learning, RL) 是一种机器学习方法，通过与环境的交互来学习如何采取行动，以最大化累积奖励。

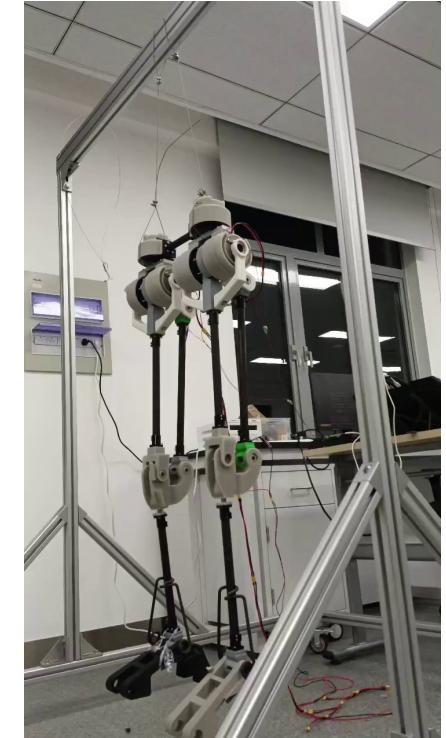
- **价值迭代 (Value Iteration) 和策略迭代 (Policy Iteration)**：基于动态规划的方法，用于确定最优策略。
- **Q学习 (Q-learning)**：一种无模型的强化学习算法，通过更新状态-动作值函数 (Q函数) 来学习最优策略。
- **深度Q网络 (DQN)**：将深度学习与Q学习相结合，利用神经网络来近似Q函数，处理高维状态空间的问题。
- **策略梯度方法 (Policy Gradient Methods)**：直接优化策略，通过梯度上升法找到最优策略，适用于连续动作空间。



# 强化学习

强化学习 (Reinforcement Learning, RL) 是一种机器学习方法，通过与环境的交互来学习如何采取行动，以最大化累积奖励。

- **初始化**: 智能体在初始状态下开始与环境互动。
- **动作选择**: 智能体根据当前策略选择一个动作。
- **执行动作**: 智能体执行所选动作，环境产生相应状态变化，并提供即时奖励。
- **状态更新**: 智能体观察到新的状态和奖励，更新其知识和策略，以改进未来的决策。
- **重复**: 智能体不断重复以上过程，通过与环境的持续互动和学习，逐步改进策略，使其在不同状态下的决策能够最大化累积奖励。



举例：训练机器人行走

# 大语言模型

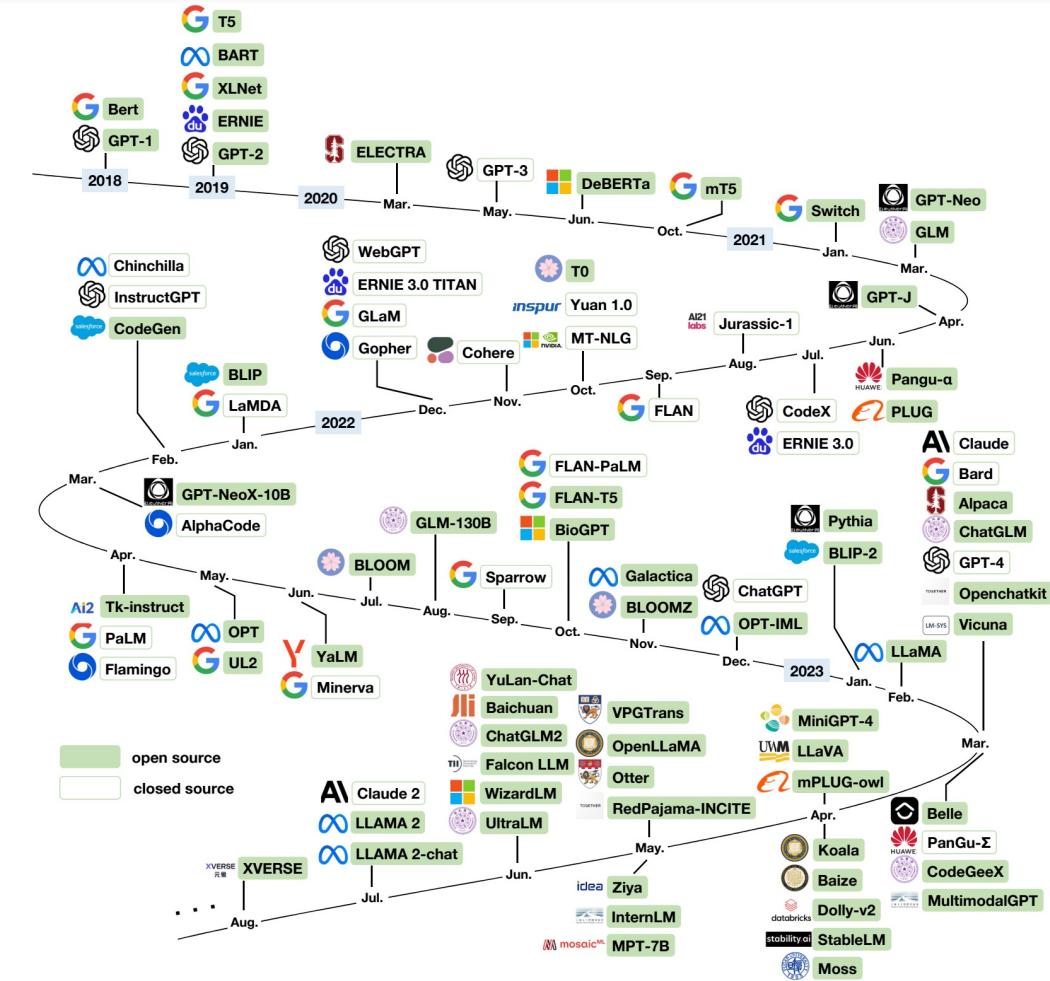
大语言模型 (Large Language Model, LLM) 是一种基于深度学习的自然语言处理 (NLP) 模型，能够理解和生成类似人类的文本。

**神经网络架构：**大语言模型通常使用深度神经网络架构尤其是基于Transformer的架构。Transformer是一种有效的神经网络模型，擅长处理序列数据，通过自注意力机制捕捉不同位置的词语之间的关系。

**大规模数据训练：**大语言模型依赖于大规模的文本数据进行训练，数据通常来自互联网。丰富的数据来源使模型能够学习到广泛的语言表达和知识。

**预训练：**大语言模型首先在大规模语料库上进行预训练学习通用的语言表示。这一阶段模型通过大量的文本数据，训练理解语言的基本规律和语义结构。

**微调：**在预训练之后，模型可以在特定任务（如问答、翻译、文本分类）上进行微调。通过在目标任务的特定数据集上进行微调，提高在特定任务上的表现。

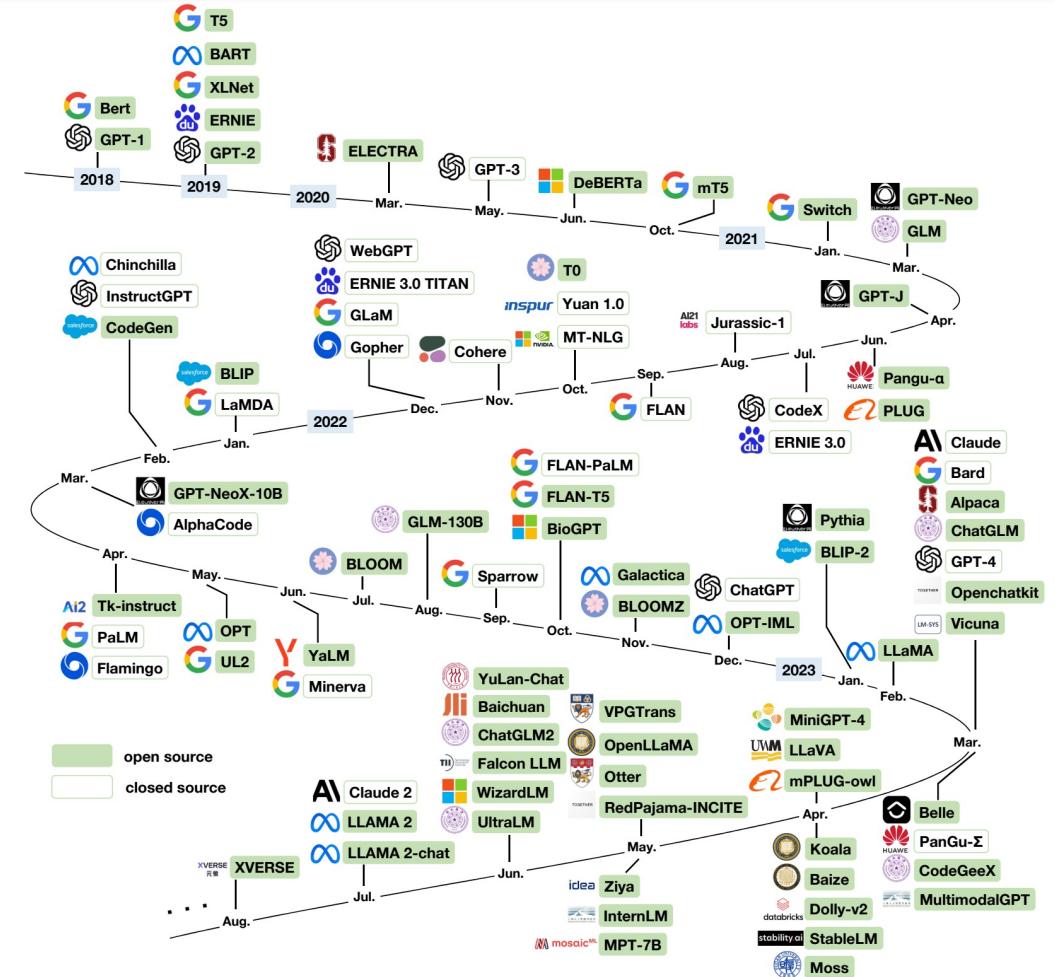


Source: Examining User-Friendly and Open-Sourced Large GPT Models: A Survey on Language, Multimodal, and Scientific GPT Models

# 大语言模型

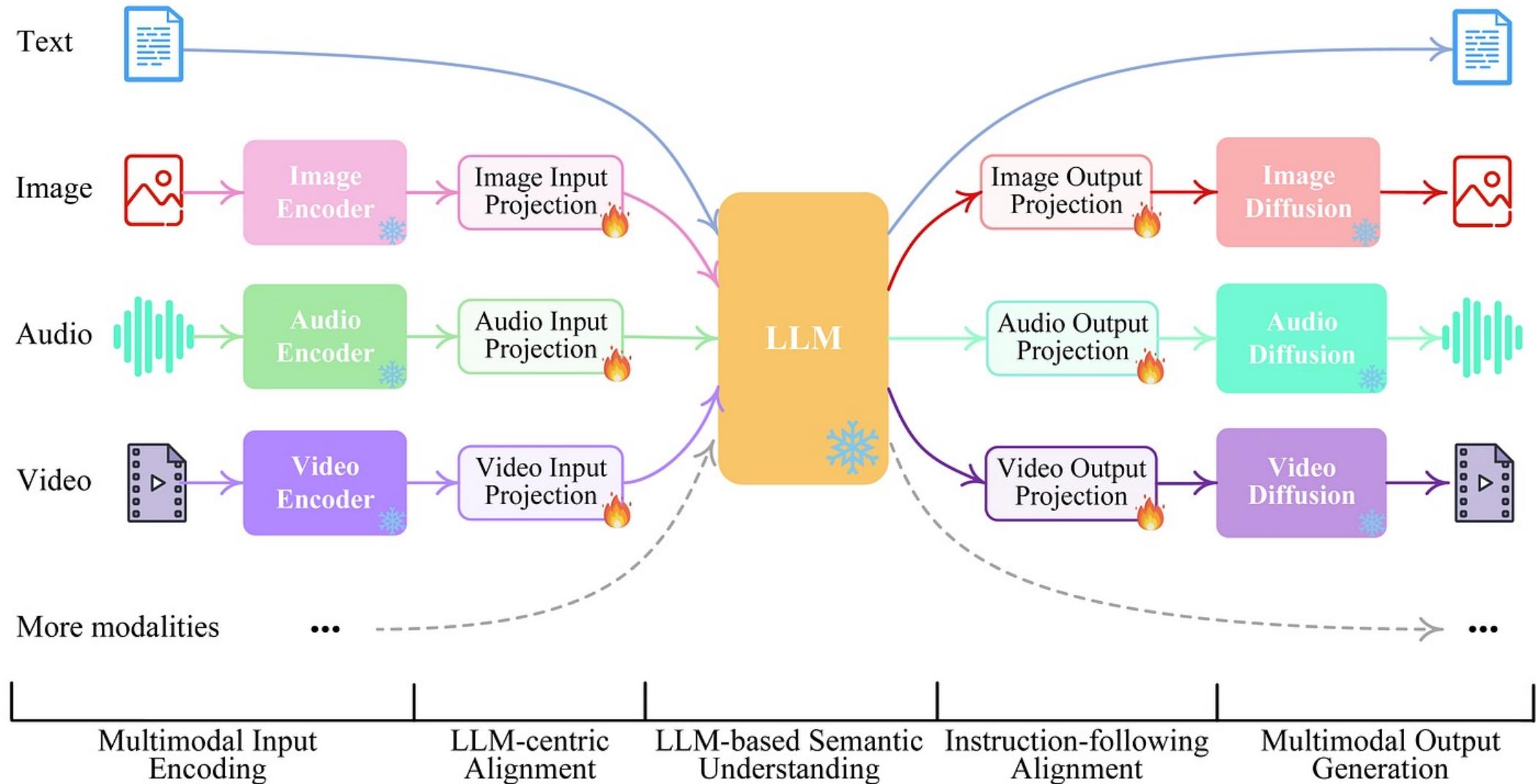
## LLM为我们带来了什么能力

- 信息与知识的能力
- 自然语言理解能力
- 推理与思辨的能力
- 学习与演化的能力
- 多模态的处理能力



Source: Examining User-Friendly and Open-Sourced Large GPT Models:  
A Survey on Language, Multimodal, and Scientific GPT Models

# 大语言模型



时间安排	教学内容	内容负责人
第1周	具身智能简介	具身智能的定义 历史背景和关键技术发展 具身智能与传统人工智能的区别
第2周	伦理隐私安全	个人隐私与安全 安全技术与隐私保护 具身智能伦理挑战
第3-6周	机器学习基础	机器学习基础 强化学习 深度学习与大模型
第7-8周	感知技术	手势与动作识别 视觉认知技术 语音交互技术
第9-10周	感觉运动系统	各种类型的传感器（视觉、触觉、听觉等） 马达系统和执行器 机器人的感觉运动整合
第11-12周	认知架构	具身人工智能的认知架构概述 大模型智能体技术 记忆和学习与智能体认知 端到端：感知到动作
第13-14周	导航与制图	空间意识和SLAM（同时定位与制图）技术 路径规划和避障 导航与任务导向决策的整合
第15周	人机交互	脑机接口
第16周	群体智能	群体机器人 机器人与人类社会

# 感知技术



手势



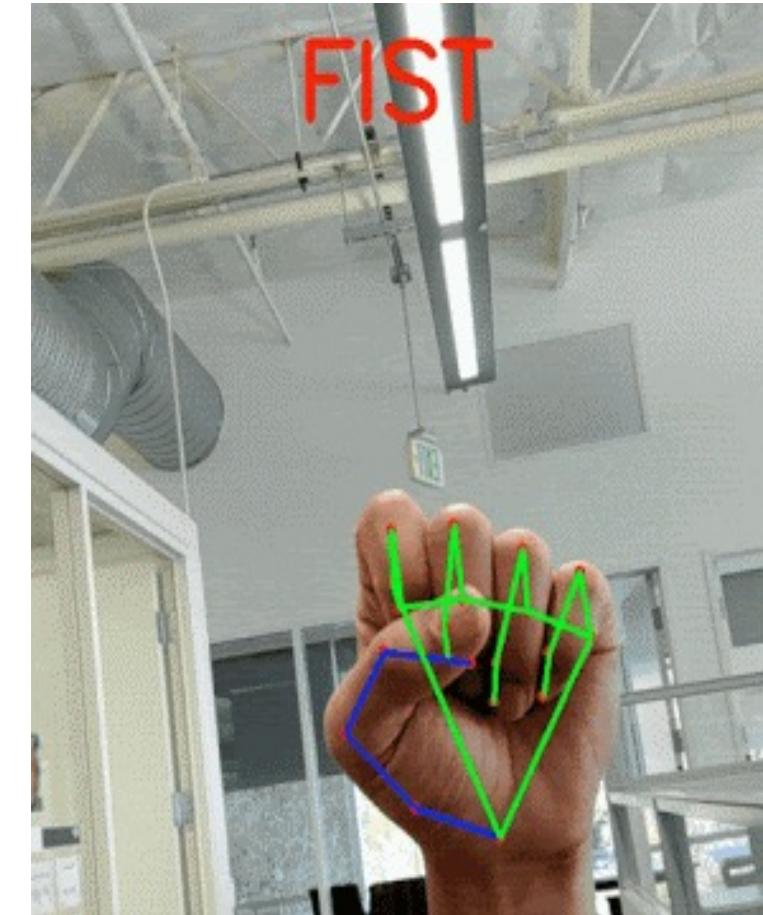
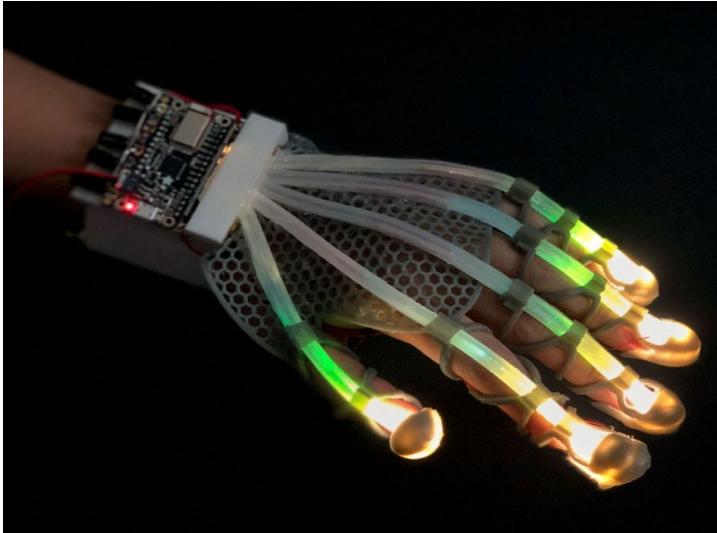
视觉



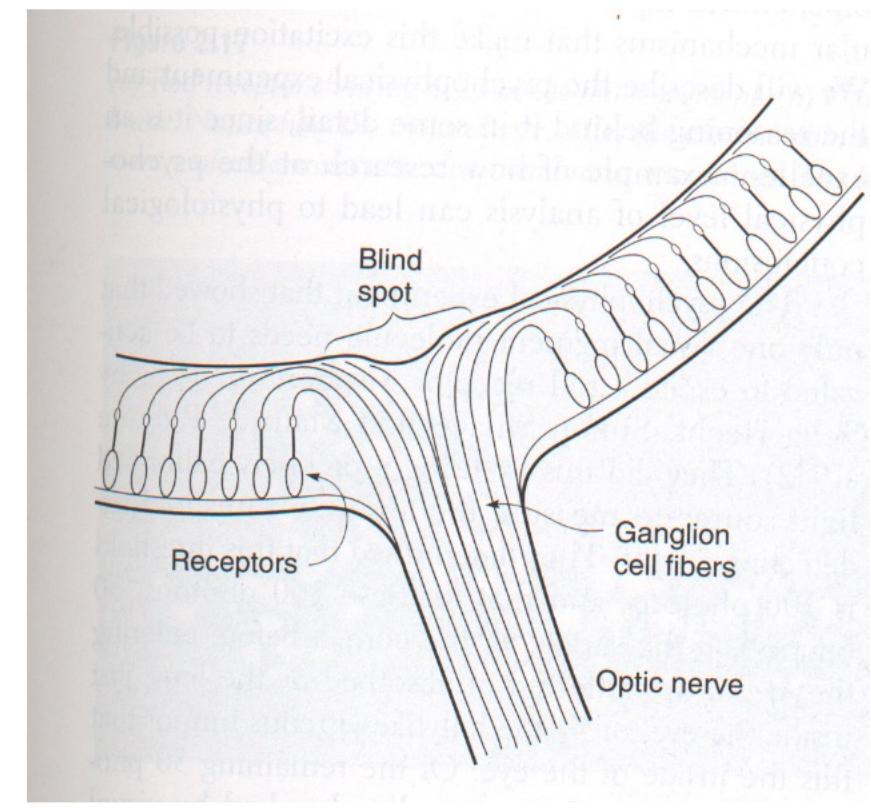
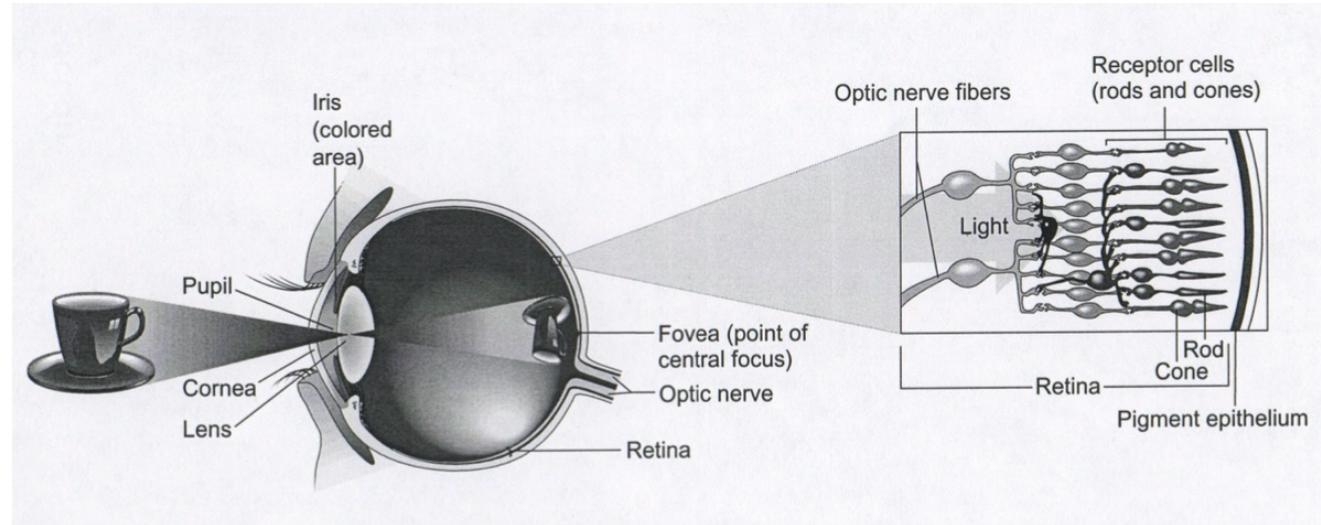
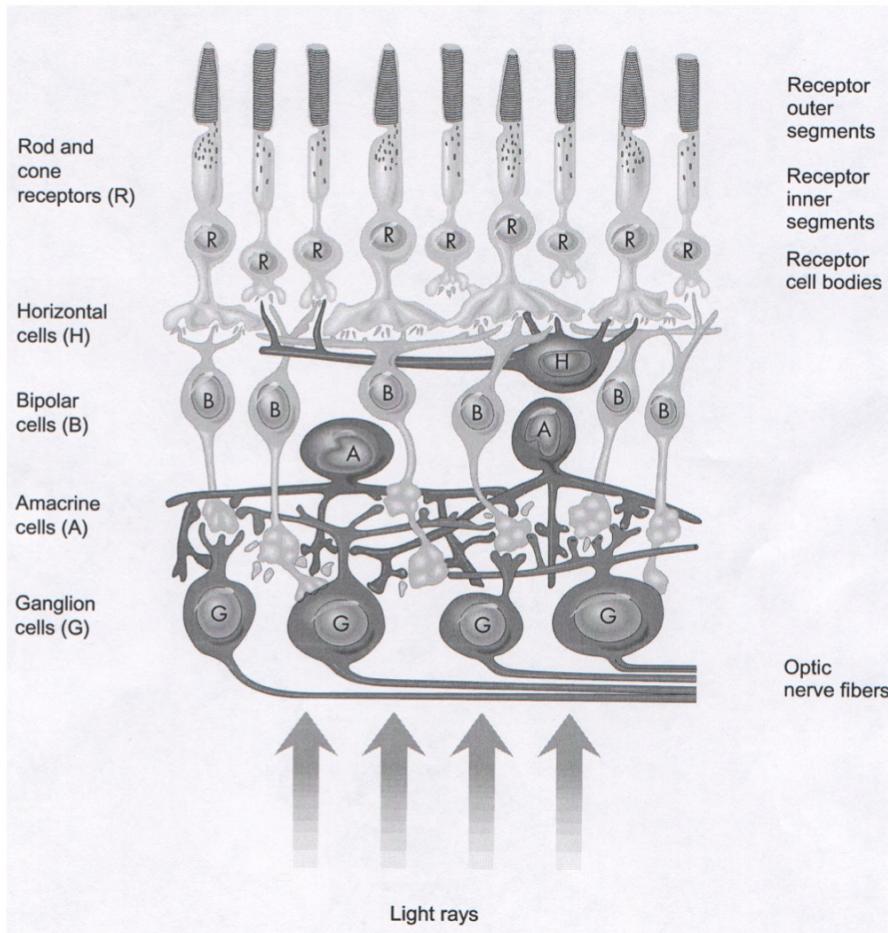
语音

# 感知技术：手势识别

手势识别技术是通过捕捉和解读手部动作或姿态来实现人机交互的一种技术。它主要分为三类：基于视觉的手势识别、基于传感器的手势识别和基于无线感知的手势识别。

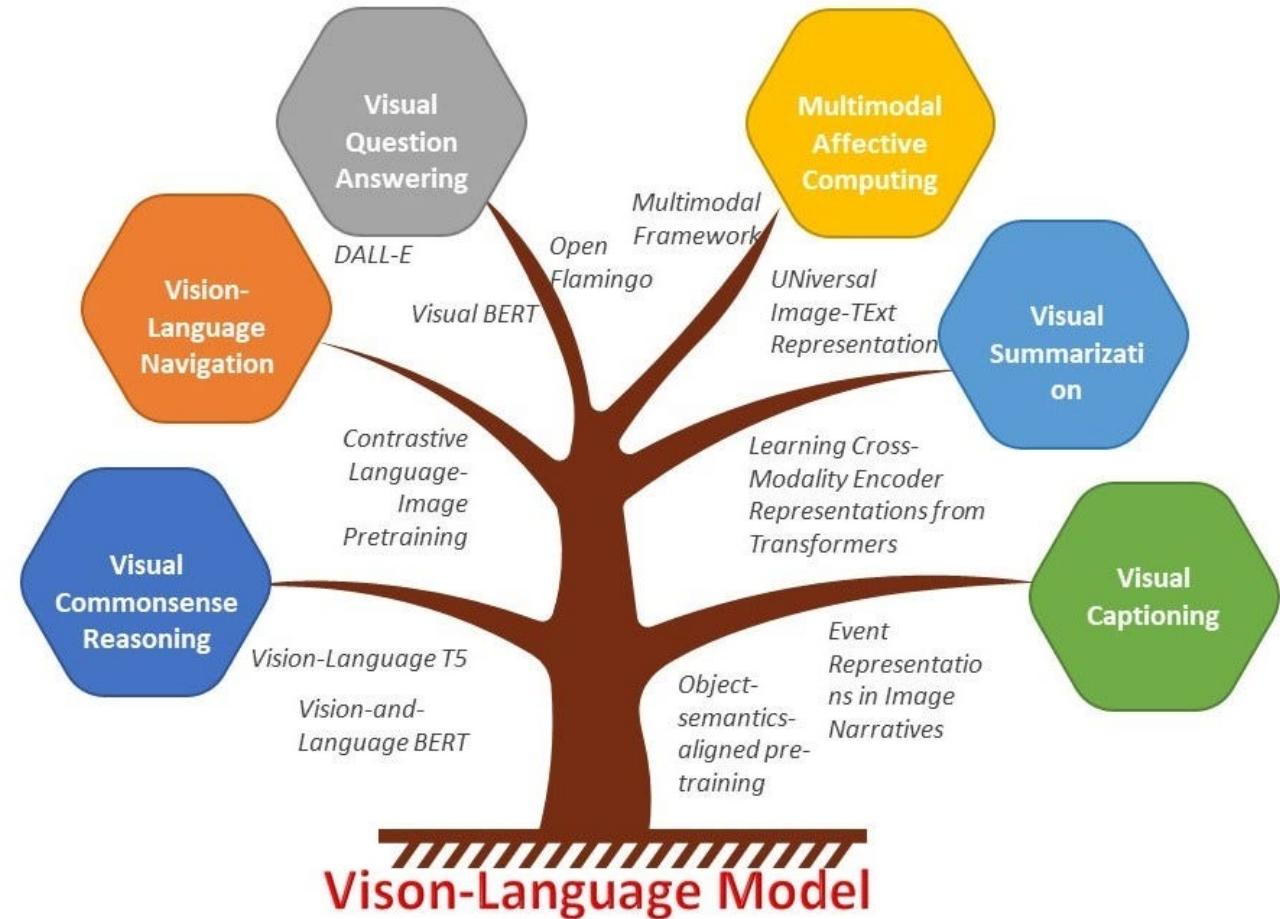


# 感知技术：视觉认知



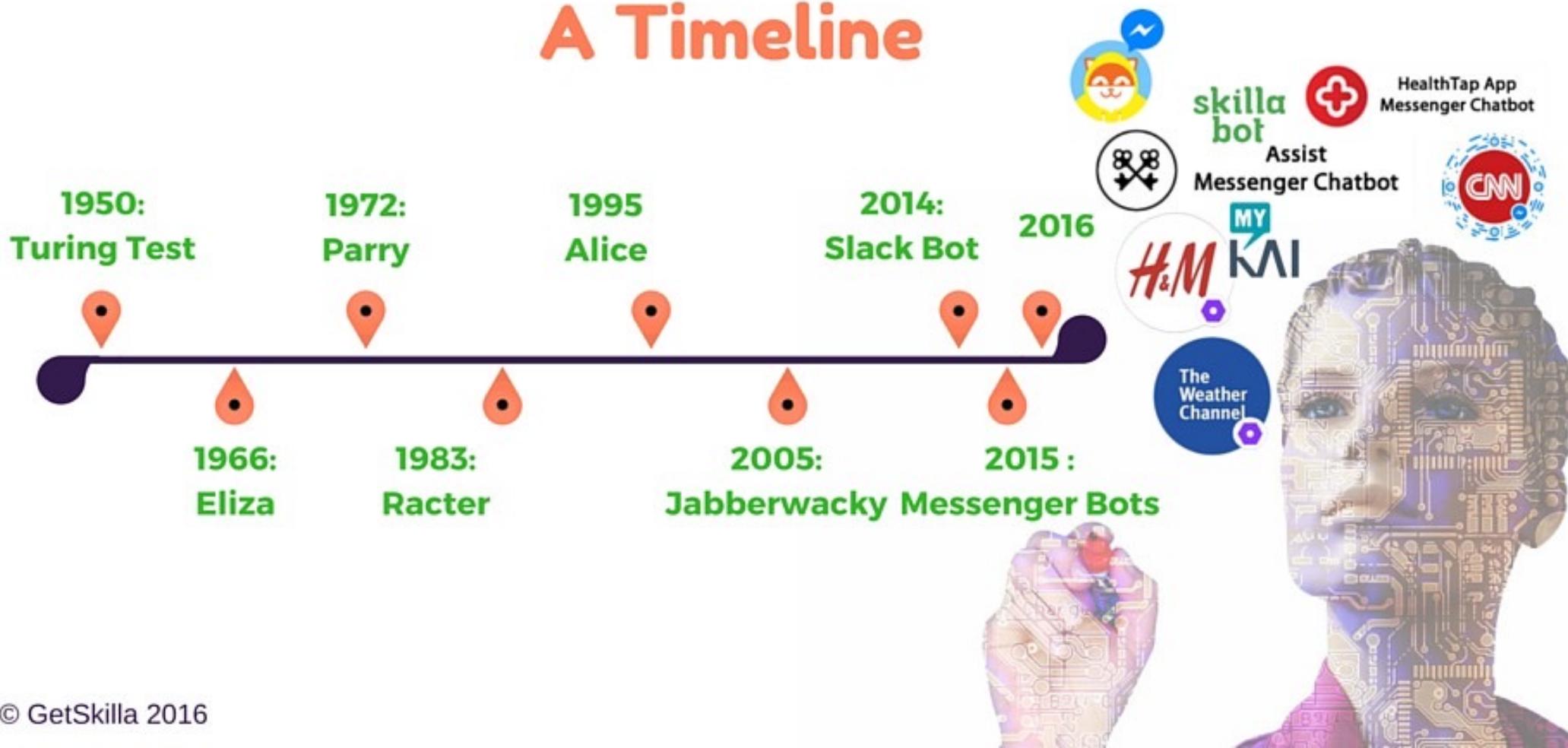
# 感知技术：视觉认知

**视觉语言模型 (Vision-Language Model, VLM)** 是一种结合了视觉和自然语言处理能力的深度学习模型。它能够同时理解和生成图像和文本信息，从而实现图文跨模态的理解和交互。这类模型可以将视觉信息（如图像、视频）和语言信息（如文本描述、问题）进行关联，通过学习视觉和语言之间的对应关系，完成如图像描述生成、图像问答、文本引导的图像生成、跨模态检索等任务。视觉语言模型广泛应用于图像识别、智能助手、自动驾驶、智能监控和人机交互等领域，通过使计算机具备多模态理解能力。



# 感知技术：语音交互

## History of Chatbots: A Timeline



# Google Duplex: A.I. Assistant Calls Local Businesses To Make Appointments



视频网址: <https://youtu.be/D5VN56jQMWM?feature=shared>

# 感知技术：语音交互

## 自动语音识别（ASR）

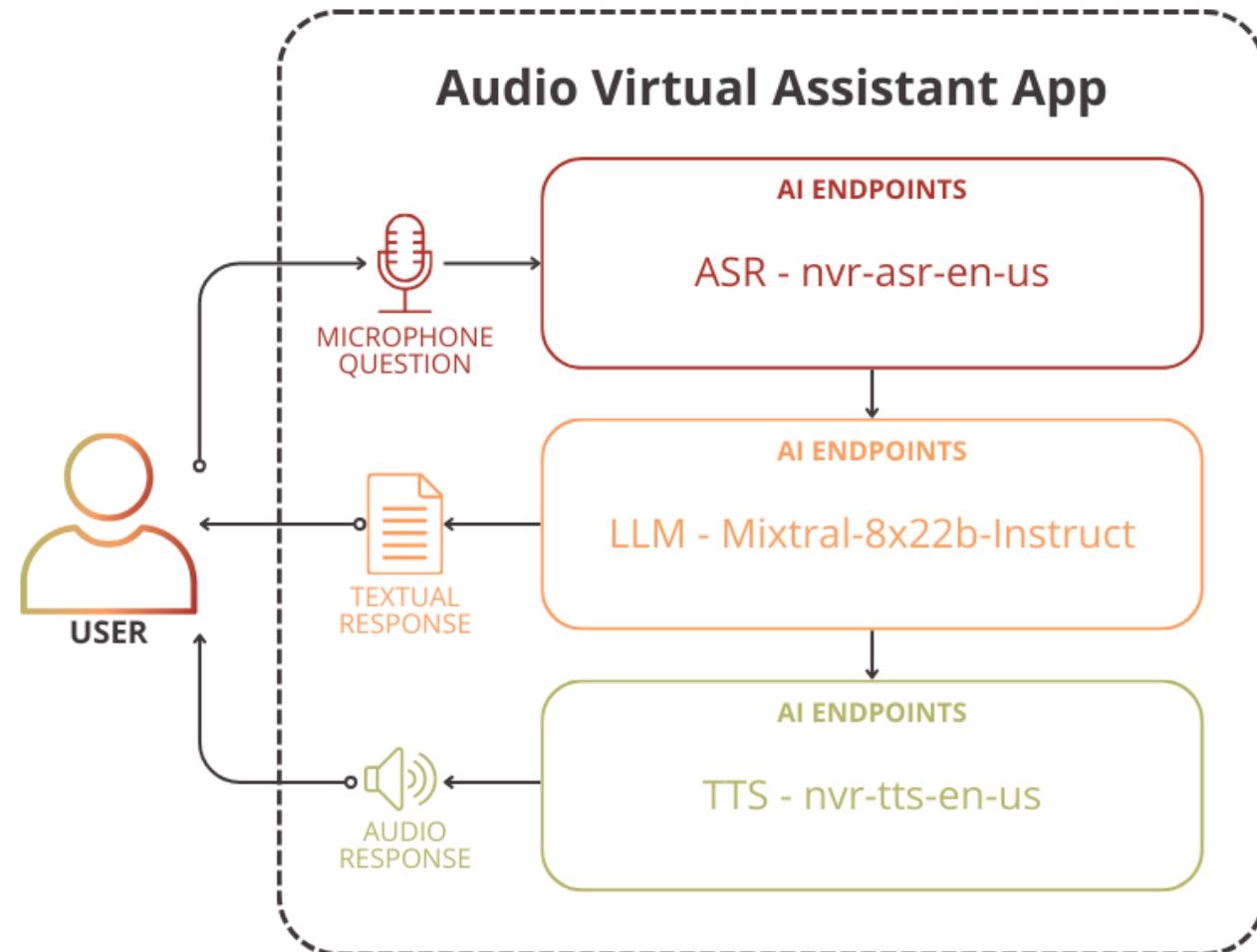
- 功能描述：**自动语音识别（ASR, Automatic Speech Recognition）是将用户的语音输入转化为文本的技术。ASR系统通过捕捉和分析用户的语音信号，识别其中的词汇和句子结构，将口语转换为可处理的文本格式。
- 应用场景：**ASR是语音助手的前端接口，用于实时理解用户的语音指令和查询，支持语音命令、搜索、问答等功能。

## 大型语言模型（LLM）

- 功能描述：**大型语言模型（LLM, Large Language Model）在语音助手中负责理解和生成自然语言文本。LLM利用大规模语料库进行训练，能够对用户的文本输入进行语义理解、信息提取，并生成合适的回答或反应。它在语音助手中的作用是理解用户的意图，并基于已知信息提供智能响应。
- 应用场景：**LLM被用于对用户问题的智能回答、对话生成、个性化建议等任务，提升语音助手的自然语言处理能力。

## 文本到语音转换（TTS）

- 功能描述：**文本到语音转换（TTS, Text-to-Speech）技术负责将LLM生成的文本响应转化为自然的语音输出。TTS系统通过合成语音来模仿人类说话，使得语音助手能够用自然、流畅的语言与用户交流。
- 应用场景：**TTS用于将文本内容转化为语音，用于语音助手的反馈、语音播报新闻、导航提示等应用，提供用户友好的听觉体验。

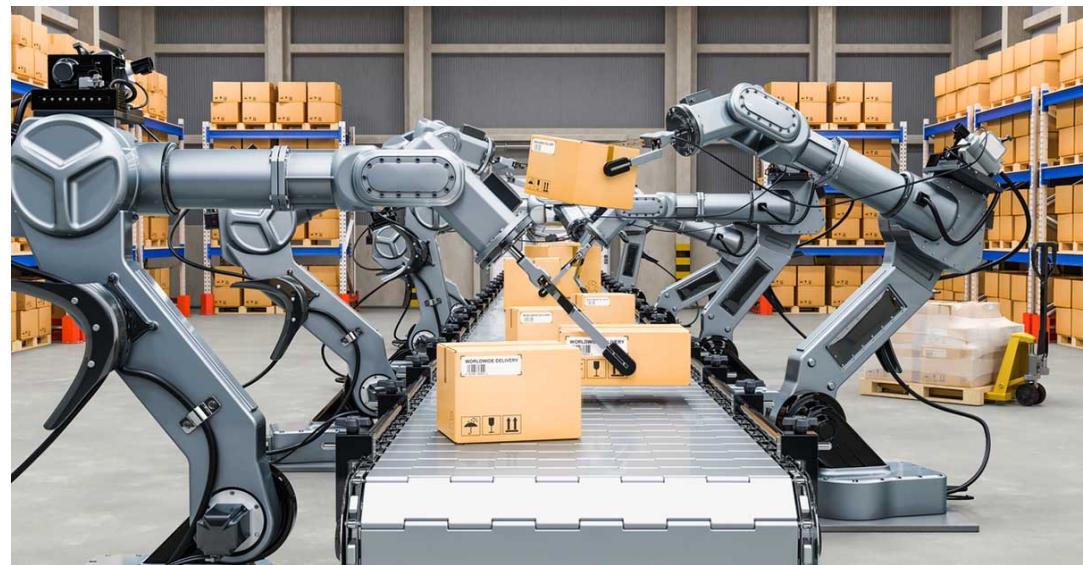


时间安排	教学内容	内容负责人
第1周	具身智能简介	具身智能的定义 历史背景和关键技术发展 具身智能与传统人工智能的区别
第2周	伦理隐私安全	个人隐私与安全 安全技术与隐私保护 具身智能伦理挑战
第3-6周	机器学习基础	机器学习基础 强化学习 深度学习与大模型
第7-8周	感知技术	手势与动作识别 视觉认知技术 语音交互技术
第9-10周	感觉运动系统	<b>各种类型的传感器（视觉、触觉、听觉等）</b> <b>马达系统和执行器</b> <b>机器人的感觉运动整合</b>
第11-12周	认知架构	具身人工智能的认知架构概述 大模型智能体技术 记忆和学习与智能体认知 端到端：感知到动作
第13-14周	导航与制图	空间意识和SLAM（同时定位与制图）技术 路径规划和避障 导航与任务导向决策的整合
第15周	人机交互	脑机接口
第16周	群体智能	群体机器人 机器人与人类社会

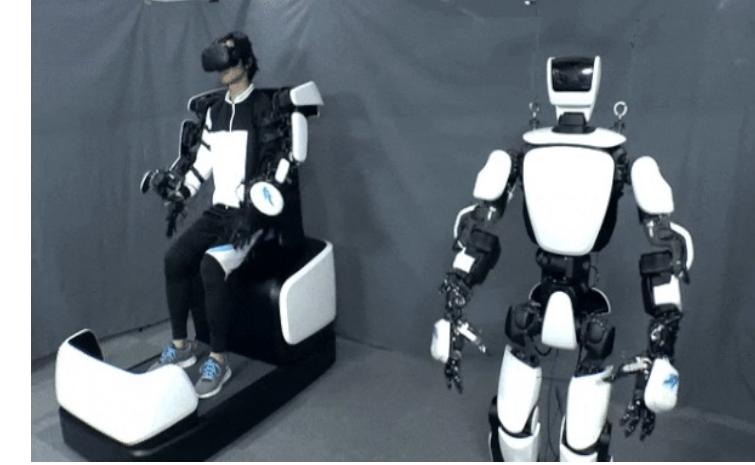
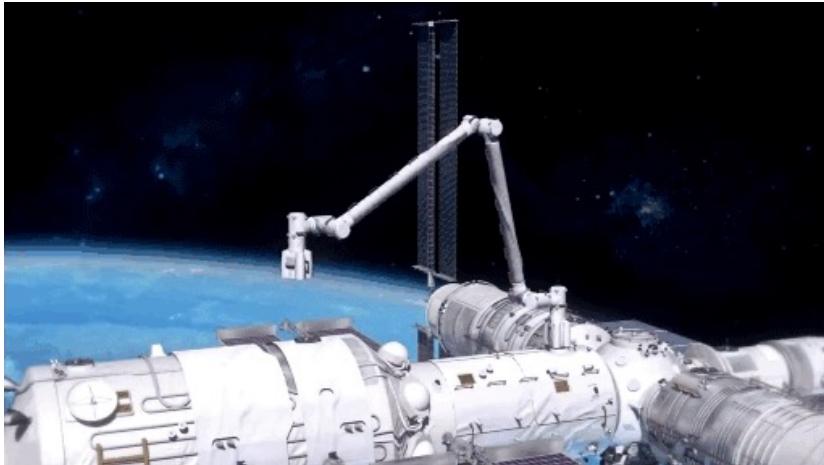
# 感觉运动系统

感觉-运动整合利用传感器获取信息，处理信息，然后通过运动系统作出适当反应的过程，是实现自主和智能行为的关键。

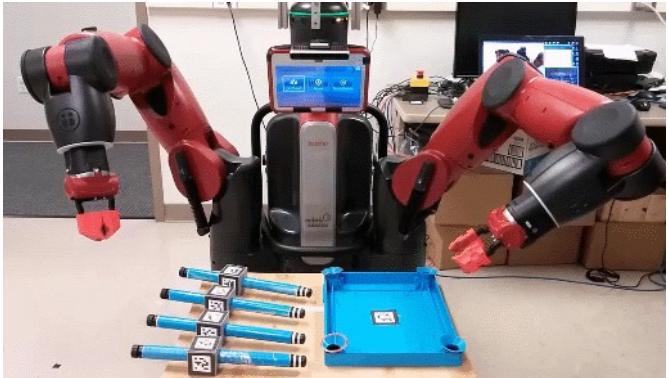
- **实时反馈控制：**机器人通过传感器持续监测自身状态和环境，并根据这些信息实时调整运动。这种反馈控制系统可以在机器人操作中提供高精度和稳定性。
- **传感器融合：**传感器融合技术将来自多个传感器的数据进行整合，以提供更准确和全面的环境理解。例如，视觉数据与LiDAR数据的融合可以提高机器人导航和障碍物检测的准确性。
- **感知-行动循环：**机器人通过感知环境并根据环境反馈作出行动，这一过程是动态且连续的。感知-行动循环是具身智能的基础，强调机器人与环境的持续互动。
- **学习与适应：**通过机器学习算法，机器人可以从感知数据中学习，并不断改进其运动策略。深度学习和强化学习技术使机器人能够在复杂环境中学习如何导航、避障、操控物体等。



# 典型机器人系统

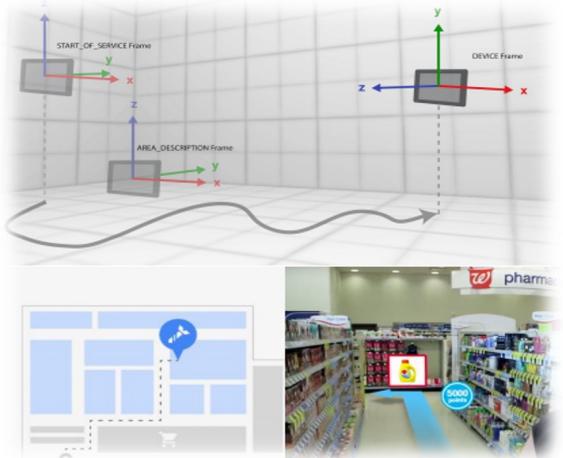


# 感觉运动系统解决的主要问题与目标

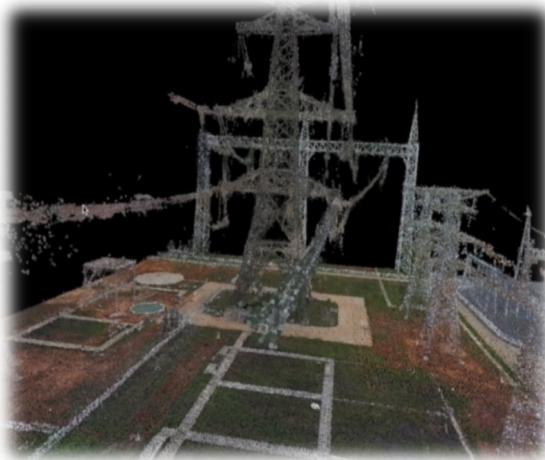


1. 机器人如何像人一样聪明避障行走?
2. 机器人如何像人一样感知周围环境?
3. 机器人如何像人一样识别定位目标?
4. 机器人如何像人一样操作

能够根据空间几何信息自主行走



实时建模感知环境-SLAM



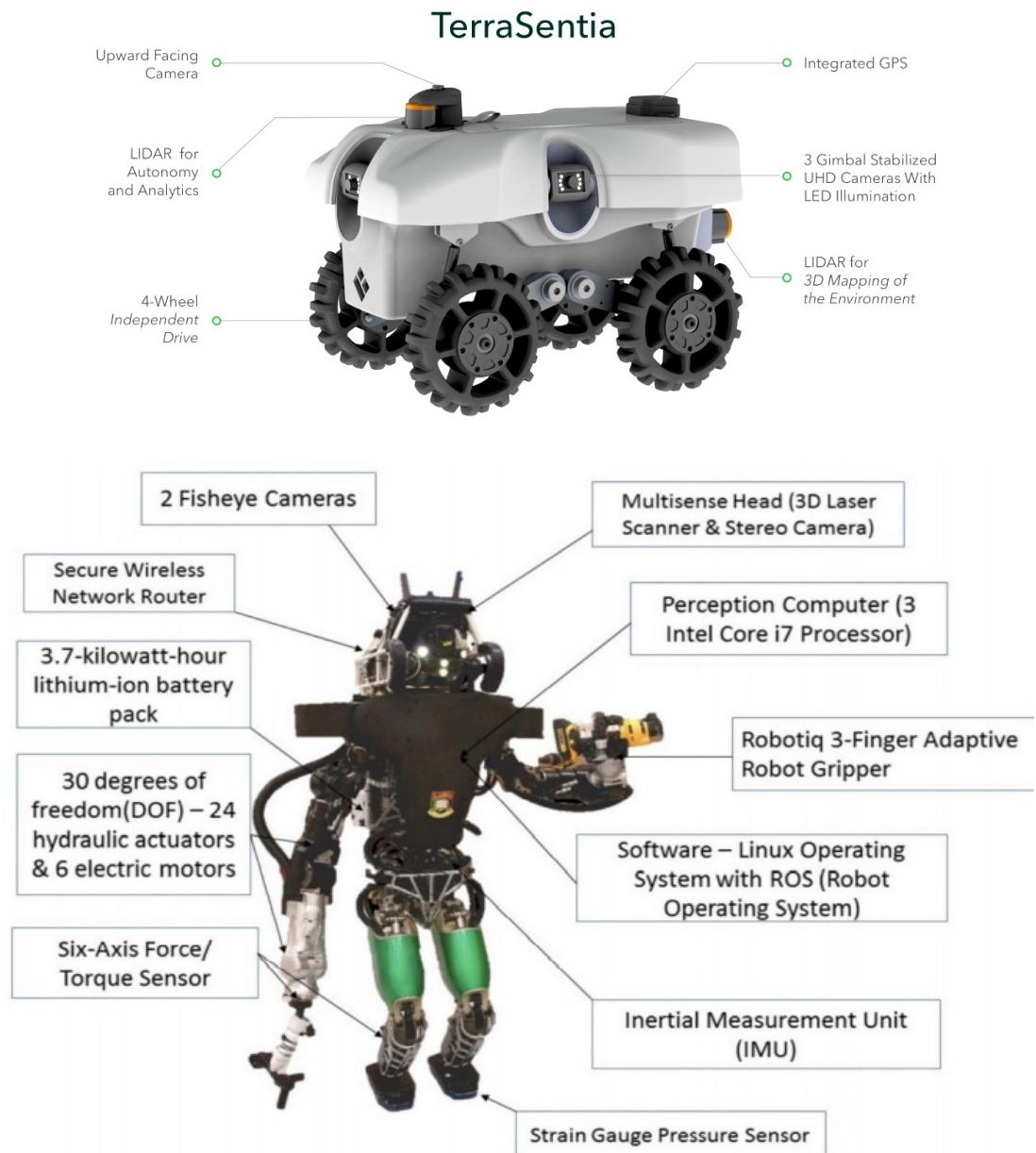
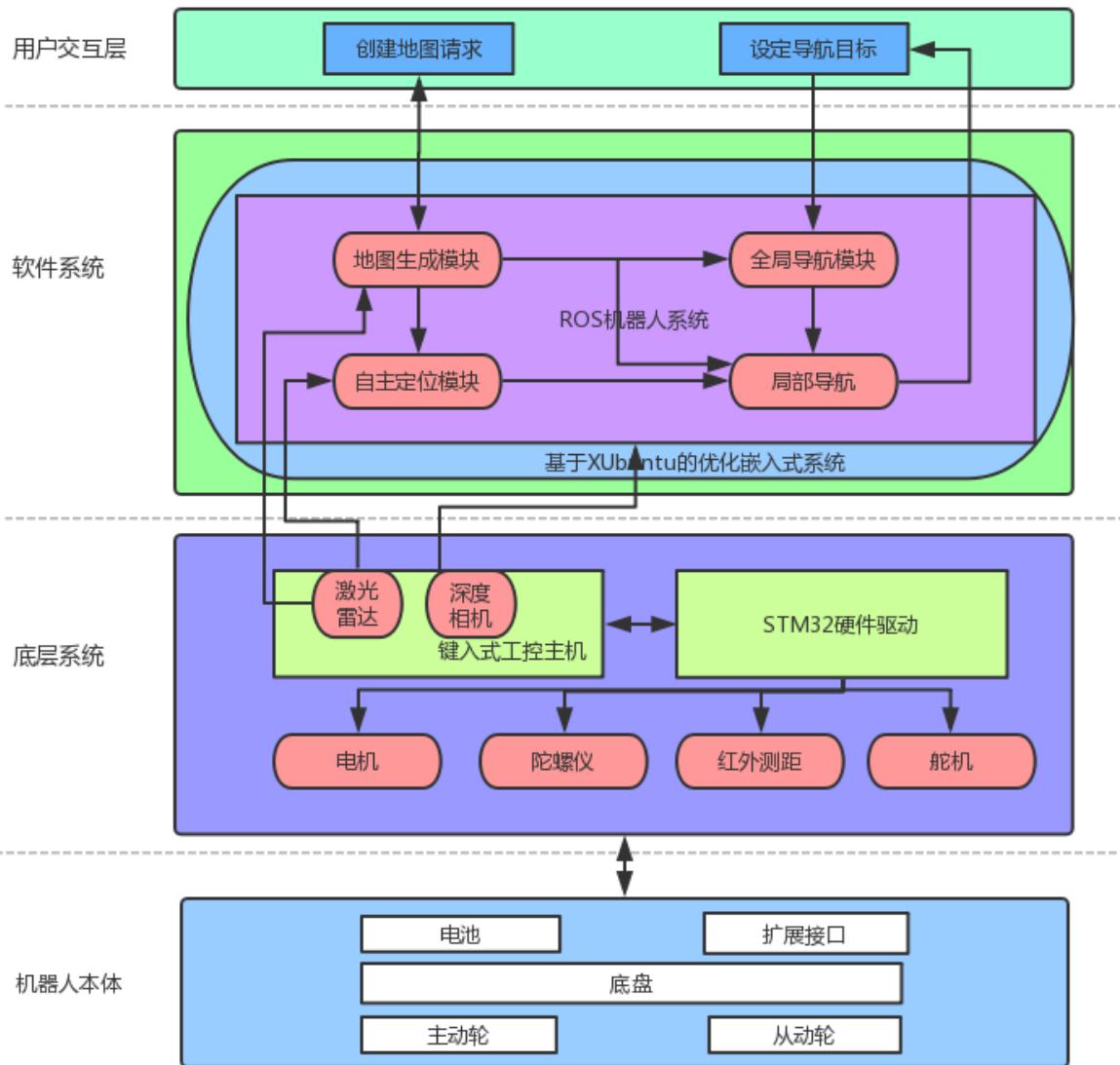
匹配特征识别目标物



对目标物体进行定位与操作



# 自主机器人的系统构成



时间安排	教学内容	内容负责人
第1周	具身智能简介	具身智能的定义 历史背景和关键技术发展 具身智能与传统人工智能的区别
第2周	伦理隐私安全	个人隐私与安全 安全技术与隐私保护 具身智能伦理挑战
第3-6周	机器学习基础	机器学习基础 强化学习 深度学习与大模型
第7-8周	感知技术	手势与动作识别 视觉认知技术 语音交互技术
第9-10周	感觉运动系统	各种类型的传感器（视觉、触觉、听觉等） 马达系统和执行器 机器人的感觉运动整合
第11-12周	认知架构	<b>具身人工智能的认知架构概述 大模型智能体技术 记忆和学习与智能体认知 端到端：感知到动作</b>
第13-14周	导航与制图	空间意识和SLAM（同时定位与制图）技术 路径规划和避障 导航与任务导向决策的整合
第15周	人机交互	脑机接口
第16周	群体智能	群体机器人 机器人与人类社会

# 具身智能-认知架构

## 感知任务 (Perception)

- 视觉、听觉、触觉、嗅觉、味觉感知

## 运动控制任务 (Motor Control)

- 运动规划、抓取和操作、协调运动

## 学习任务 (Learning)

- 强化学习、模仿学习、自主探索

## 认知任务 (Cognitive)

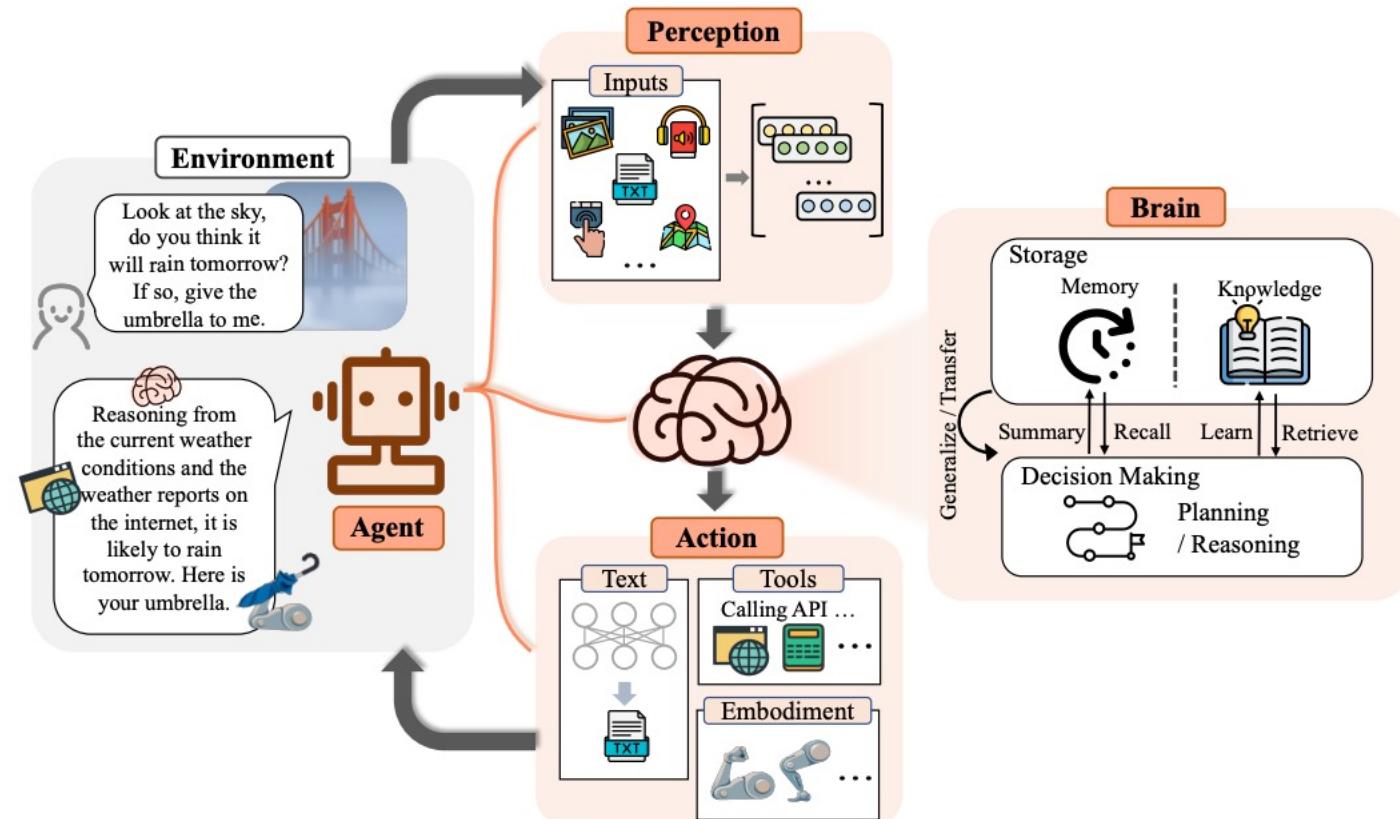
- 决策与规划、问题解决、语言理解与生成

## 社交互动任务 (Social Interaction)

- 情感识别与反应、合作与写作
- 文化和社会规范理解

## 基于与认知地图构建

- 空间记忆、语义记忆



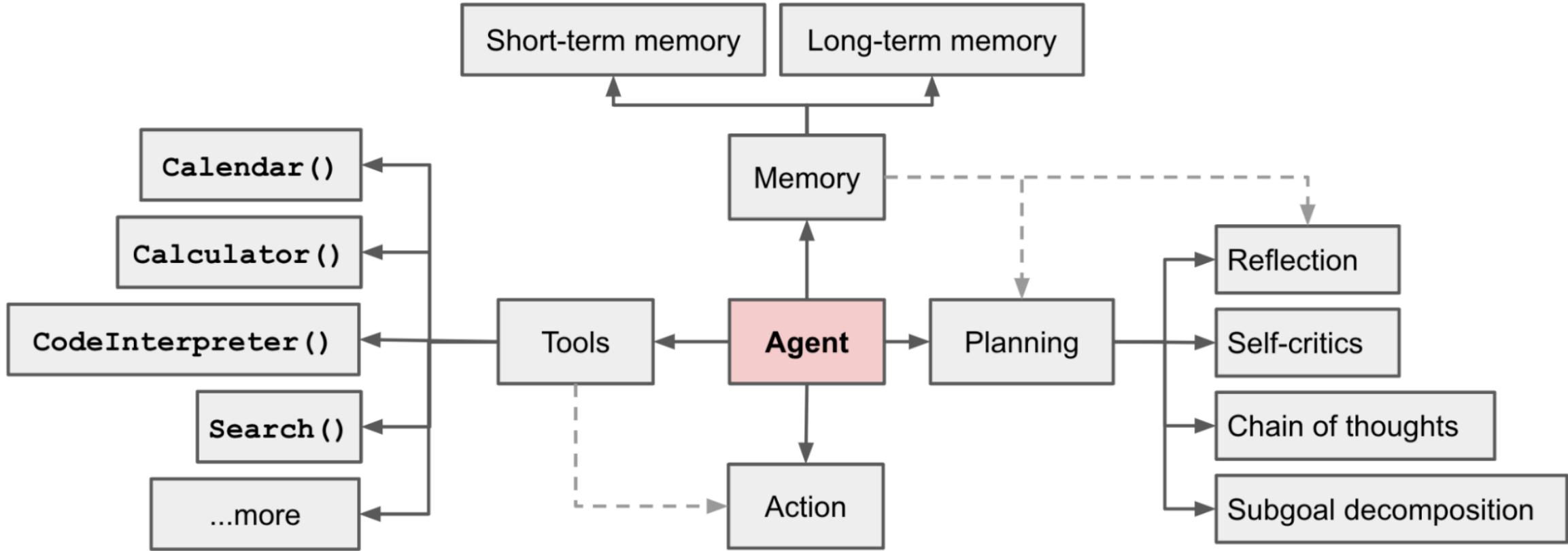
# 具身智能-认知架构

## 智能体：思维与计划

- **子目标和分解：**智能体将大型任务分解为较小的、可管理的子目标，从而能够高效地处理复杂任务。
- **反思与改进：**智能体能够对过去的行动进行自我批评和自我反思，从错误中学习，并为未来的步骤进行改进，从而提高最终结果的质量。

- Chain-of-Thoughts：复杂问题的解构
- Tree-of-Thoughts：从思维到计划
- Graph-of-thoughts：对接领域知识
- ReAct：从观察到交互
- Reflection：贝叶斯推断
- Multi-agent：协同工作

# 具身智能-认知架构



# 具身智能-认知架构

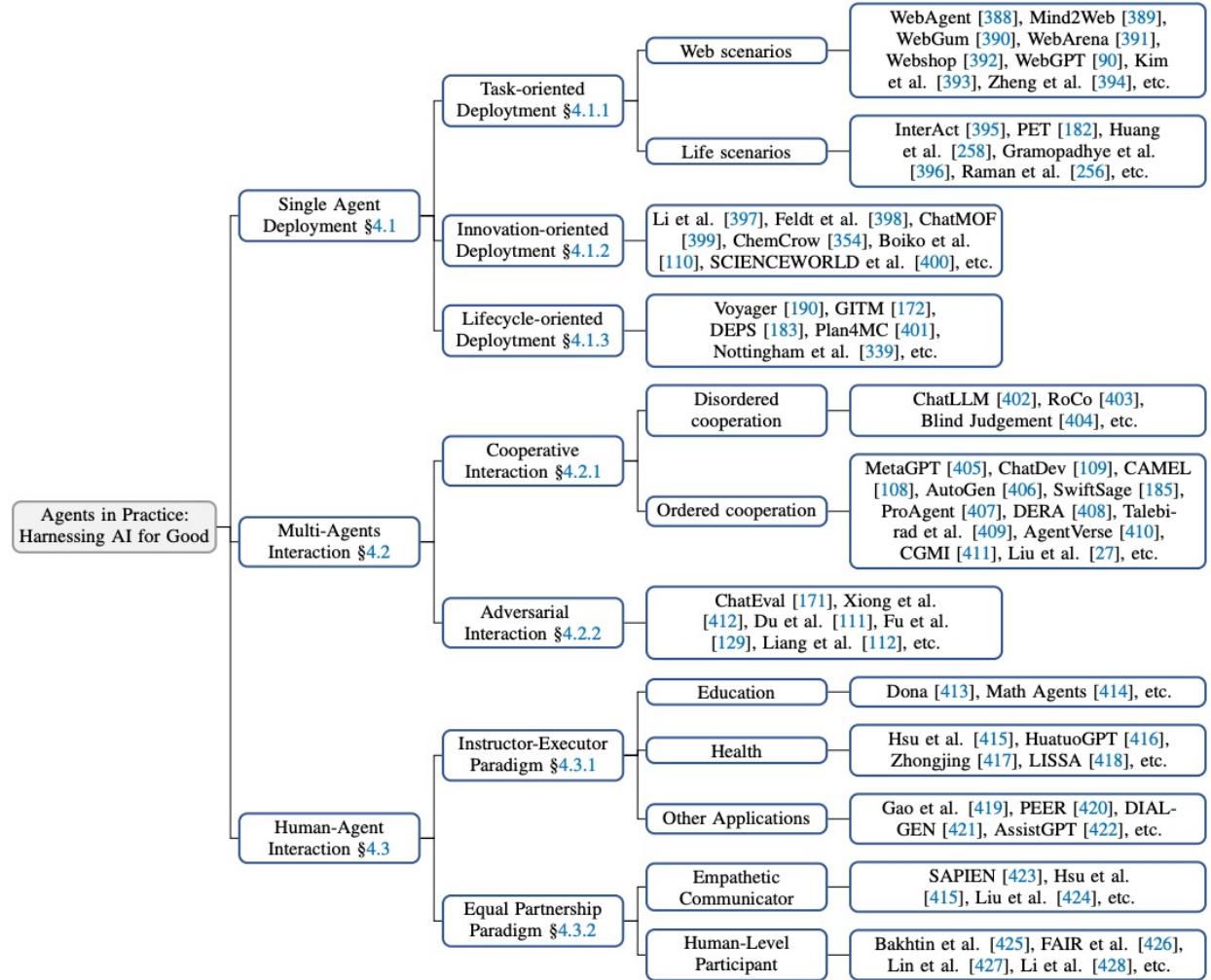
## ■ 大模型智能体技术 (2) :

### ■ 智能体实践

- 单智能体
- 多智能体协同
- 人和智能体交互

### ■ 智能体社会

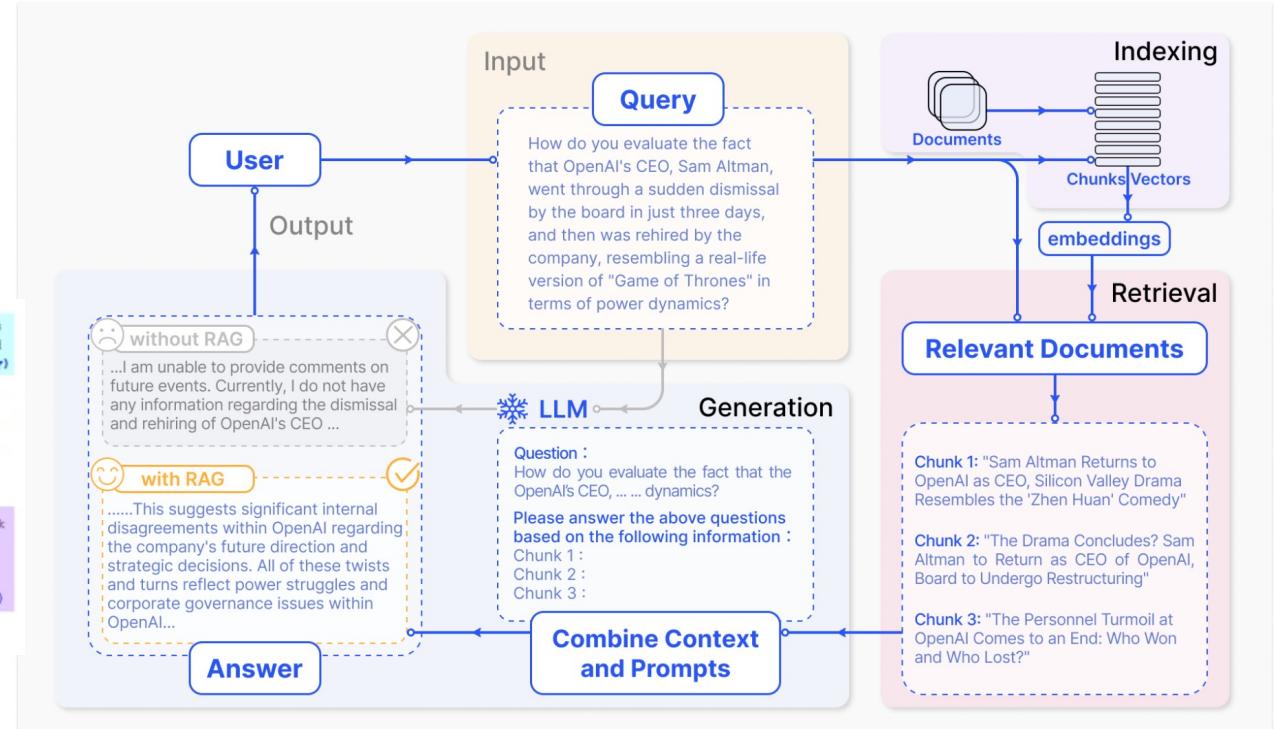
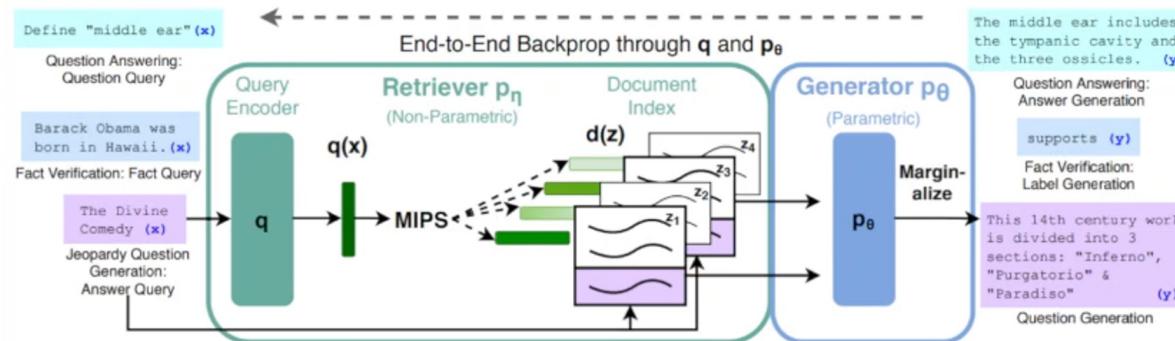
- 智能体行为和个性
- 智能体环境
- 社会仿真



# 具身智能-认知架构

$$p_{\text{RAG-Sequence}}(y|x) \approx \sum_{z \in \text{top-}k(p(\cdot|x))} p_\eta(z|x) p_\theta(y|x, z)$$

- RAG的基本流程：索引 + 检索 + 生成
- 构建知识库，关键是建立索引，为了能够更加精确地找到相关的信息



# 具身智能-认知架构

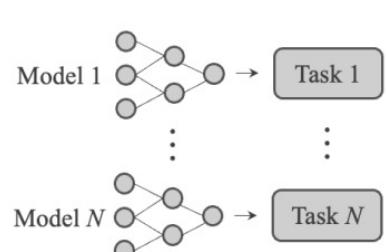
## ■ 端到端：感知到动作（1）

### ■ 端到端概述

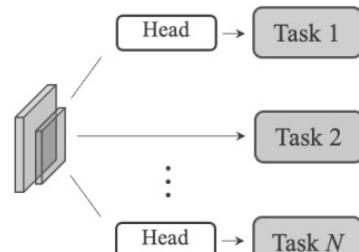
- Sensor-to-Action
- 局限和挑战

### ■ 端到端典型技术

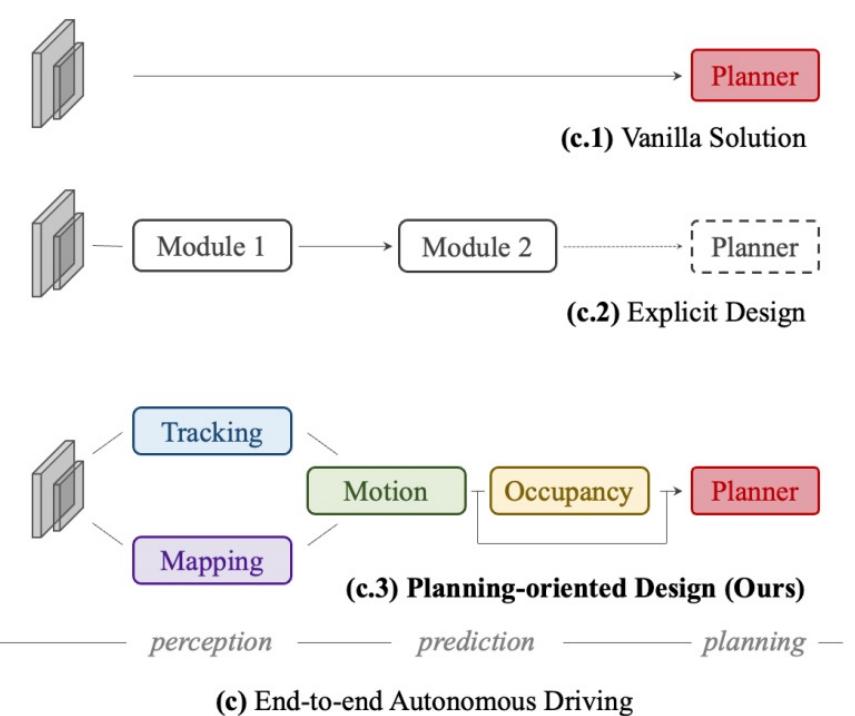
- 早期端到端算法
- 模块化端到端
- 自动驾驶及机器人应用



(a) Standalone Models



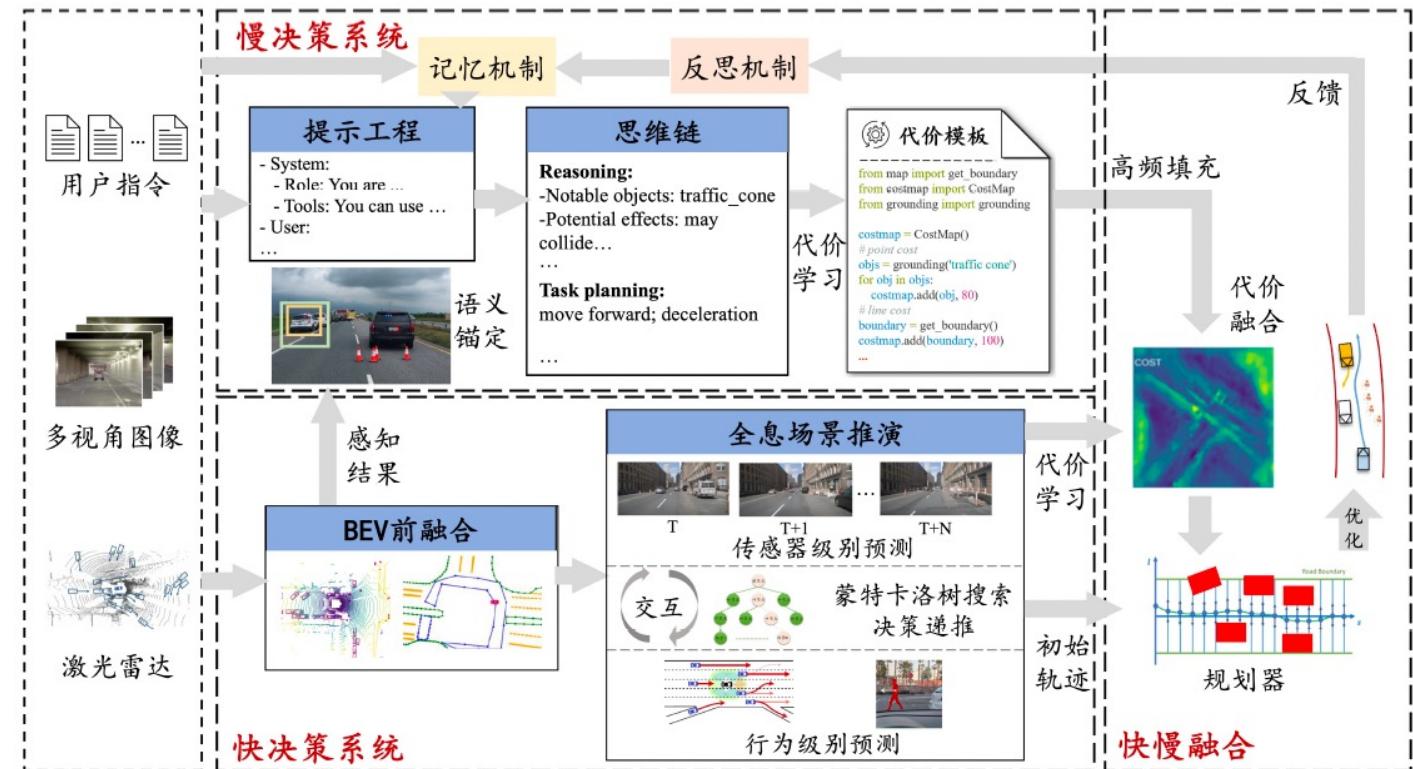
(b) Multi-task Framework



# 具身智能-认知架构

## 端到端：感知到动作（2）

- 端到端与智能体技术
- 端到端技术的局限
- 智能体技术的局限
- 大小模型融合



# **Multi-agent hide and seek**



**Multi-Agent  
Hide and Seek**

**视频网址：** <https://youtu.be/kopoLzvh5jY?feature=shared>

时间安排	教学内容	内容负责人
第1周	具身智能简介	具身智能的定义 历史背景和关键技术发展 具身智能与传统人工智能的区别
第2周	伦理隐私安全	个人隐私与安全 安全技术与隐私保护 具身智能伦理挑战
第3-6周	机器学习基础	机器学习基础 强化学习 深度学习与大模型
第7-8周	感知技术	手势与动作识别 视觉认知技术 语音交互技术
第9-10周	感觉运动系统	各种类型的传感器（视觉、触觉、听觉等） 马达系统和执行器 机器人的感觉运动整合
第11-12周	认知架构	具身人工智能的认知架构概述 大模型智能体技术 记忆和学习与智能体认知 端到端：感知到动作
第13-14周	导航与制图	空间意识和SLAM（同时定位与制图）技术 路径规划和避障 导航与任务导向决策的整合
第15周	人机交互	脑机接口
第16周	群体智能	群体机器人 机器人与人类社会

# 导航与制图

**SLAM（同时定位与地图构建）** 是指在未知环境中，移动机器人或其他智能设备通过传感器（如激光雷达、摄像头、IMU等）获取环境信息，同时进行自我定位并构建环境的地图。SLAM的核心任务是使机器人在移动过程中能够实时更新自身的位置（定位）和周围环境的地图（地图构建），从而在未知或动态变化的环境中实现自主导航。

## SLAM技术定义

SLAM  
(Simultaneous Localization and Mapping) 即同时定位与制图技术，是空间意识实现的关键技术之一。

01

## SLAM的工作原理

SLAM技术通过传感器数据实时构建环境地图，并定位智能体在地图中的位置，实现未知环境的探索与导航。

02

## SLAM的关键算法

包括基于特征的SLAM、基于直接法的SLAM等，各有优势和适用场景。

03

# SLAM技术的 应用案例

## 无人驾驶汽车

利用激光雷达、摄像头等传感器，实现高精度的环境感知和自主导航。



## 机器人探索与救援

在灾害现场等复杂环境中，SLAM技术帮助机器人进行环境建模和路径规划，提高救援效率。



## AR/VR空间交互

结合SLAM技术，AR/VR应用能够提供更加自然和精确的空间交互体验。



# 路径规划的基本概念



## 路径规划的挑战

需要考虑环境动态变化、障碍物分布、智能体自身的运动约束等因素。



## 路径规划的常用算法

如A\*算法、Dijkstra算法、RRT算法等，各有特点和适用性。

### 路径规划的定义

路径规划是为智能体从起点到终点寻找一条可行且最优的路径的过程。

# 避障技术的发展

## 避障技术的重要性

- 避障技术确保智能体在复杂环境中安全有效地执行任务。

## 避障技术的关键要素

- 包括感知障碍物、预测障碍物运动、实时更新路径等。

## 避障算法的实现

- 融合机器学习、传感器融合等技术，提高避障的准确性和鲁棒性。

# 路径规划与避障的实践案例

01

自主无人机飞行

在城市环境中，无人机利用路径规划与避障技术实现安全飞行。

02

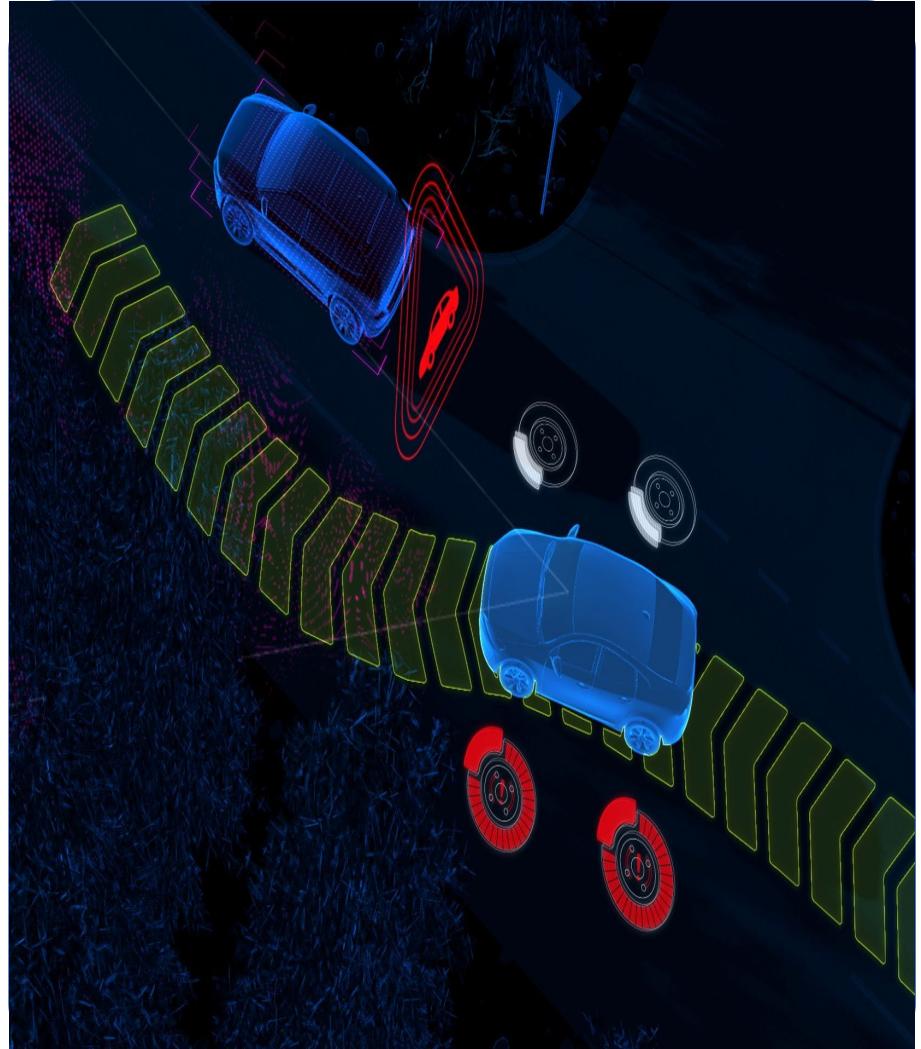
工业自动化与机器人

在自动化生产线上，机器人通过路径规划与避障技术提高作业效率。

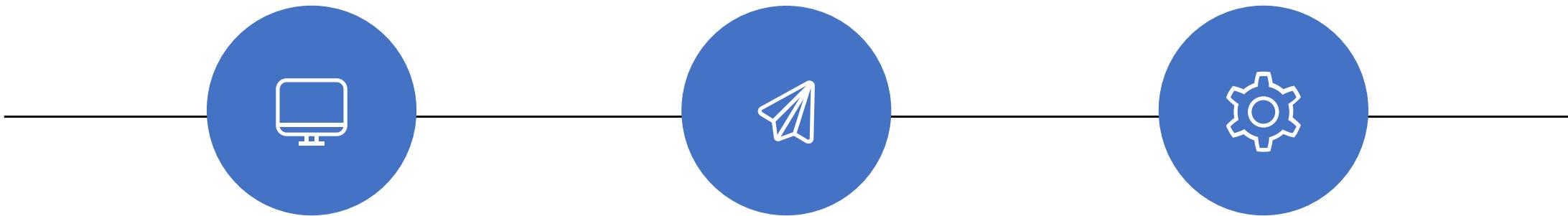
03

无人驾驶汽车的避障

无人驾驶汽车在城市交通中应用高级避障算法，确保行车安全。



# 导航系统的组成



## 导航系统的硬件组成

包括GPS、IMU、激光雷达等传感器，以及计算平台等。

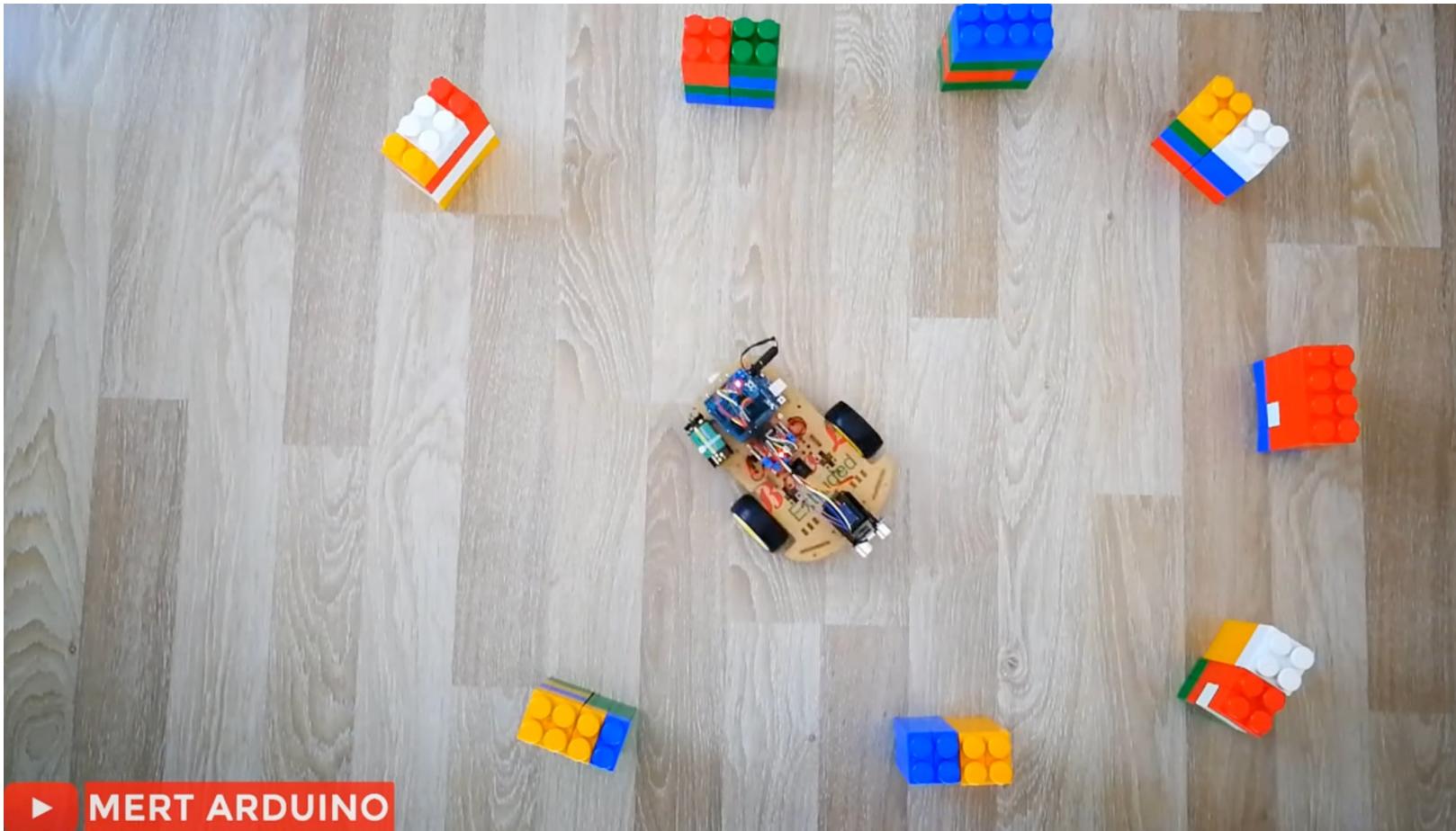
## 导航系统的软件架构

涉及感知、定位、规划、决策等多个模块的协同工作。

## 导航系统的实时性要求

导航系统需要快速响应环境变化，提供实时的决策支持。

# How to make Arduino Obstacle Avoiding Robot Car under \$20

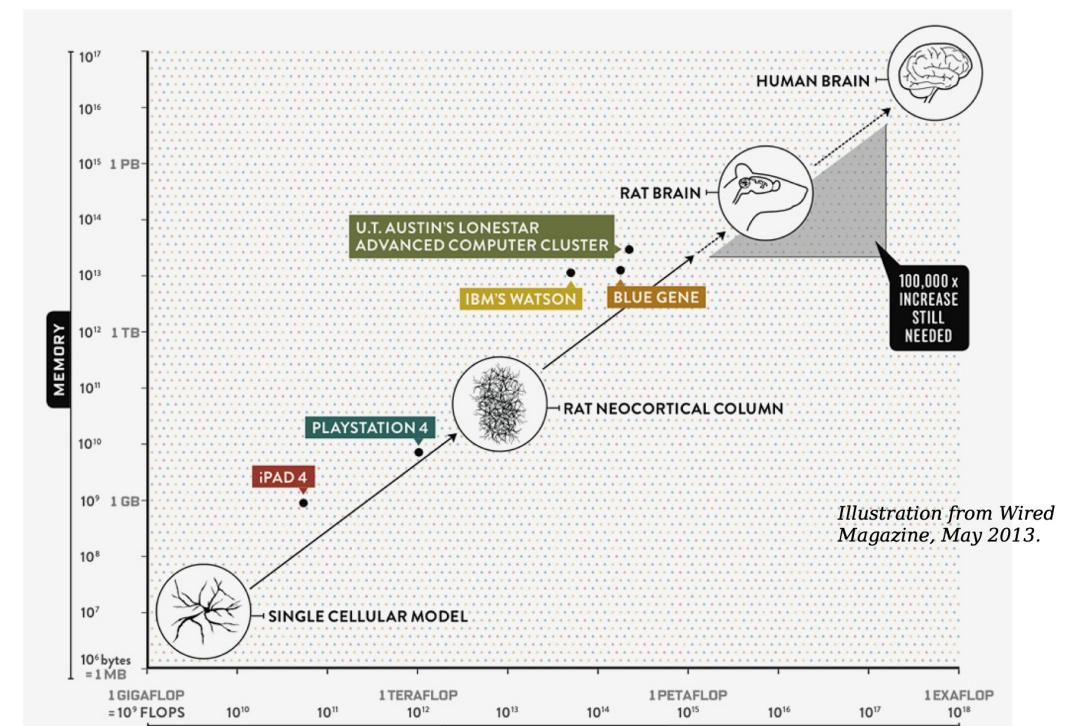
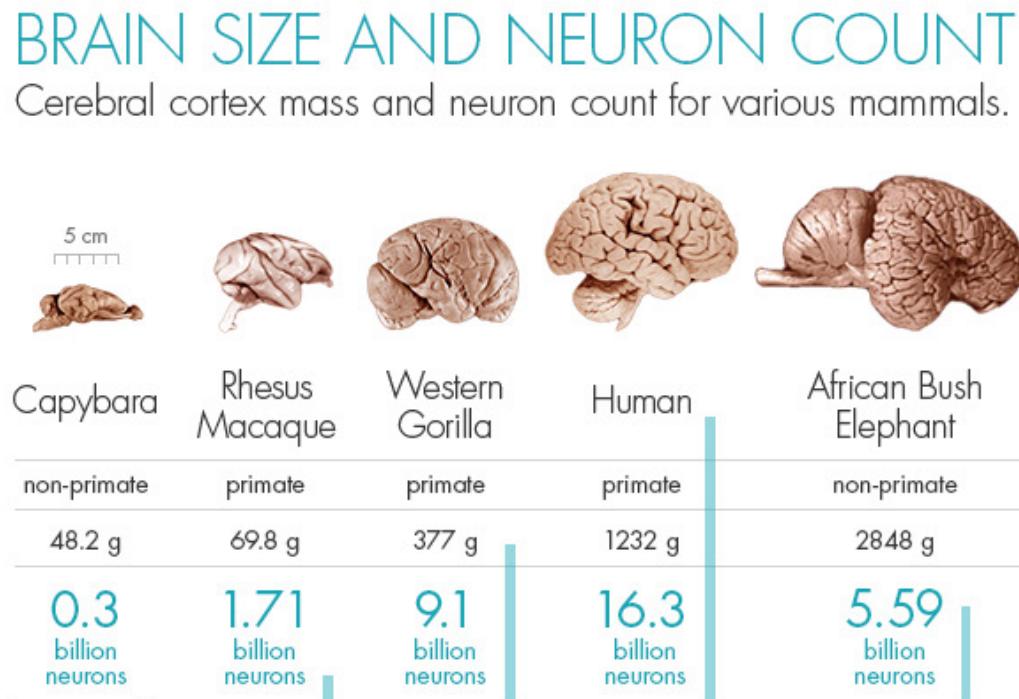


视频网址：<https://youtu.be/4CFO0MiSIM8?feature=shared>

时间安排	教学内容	内容负责人
第1周	具身智能简介	具身智能的定义 历史背景和关键技术发展 具身智能与传统人工智能的区别
第2周	伦理隐私安全	个人隐私与安全 安全技术与隐私保护 具身智能伦理挑战
第3-6周	机器学习基础	机器学习基础 强化学习 深度学习与大模型
第7-8周	感知技术	手势与动作识别 视觉认知技术 语音交互技术
第9-10周	感觉运动系统	各种类型的传感器（视觉、触觉、听觉等） 马达系统和执行器 机器人的感觉运动整合
第11-12周	认知架构	具身人工智能的认知架构概述 大模型智能体技术 记忆和学习与智能体认知 端到端：感知到动作
第13-14周	导航与制图	空间意识和SLAM（同时定位与制图）技术 路径规划和避障 导航与任务导向决策的整合
第15周	人机交互	脑机接口
第16周	群体智能	群体机器人 机器人与人类社会

# 人机交互-脑机接口

脑机接口（Brain-Computer Interface, BCI）是一种直接连接大脑与外部设备的技术，通过读取和解码大脑的电信号来实现人脑与计算机之间的信息交互。以下是对脑机接口的总结，包括其工作原理、应用场景、挑战和未来发展方向。

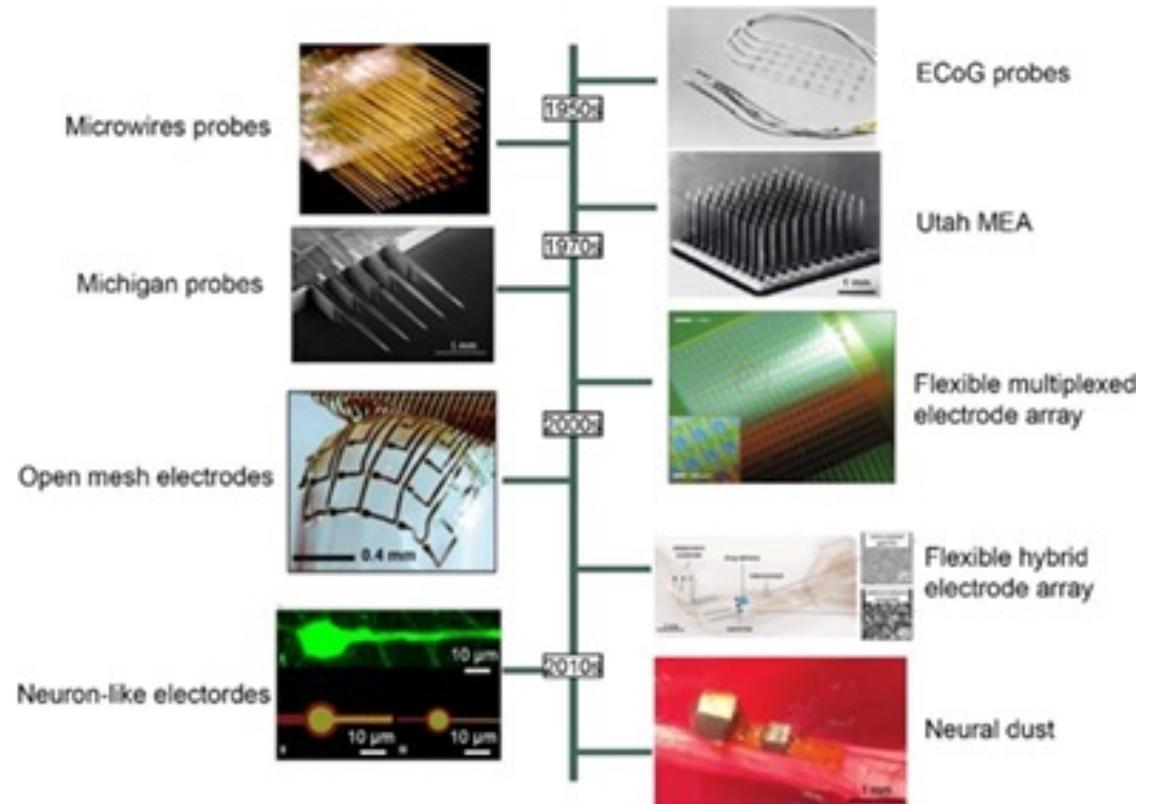


# 人机交互-脑机接口

非侵入式脑机接口是一种最实用的神经信号成像方法，应用于颅骨外部，只需应用于头皮上。  
脑电图 (EEG)  
脑磁图 (MEG)  
功能性磁共振成像 (fMRI)



侵入式脑机接口 (Invasive Brain-Computer Interface, Invasive BCI) \*\*是一种直接与大脑内部进行交互的脑机接口技术。通过将电极或神经探针植入大脑皮层或更深层的脑组织，侵入式BCI能够记录神经元的电活动，并解码这些信号以控制外部设备。



# Neuralink monkey play video games with his brain



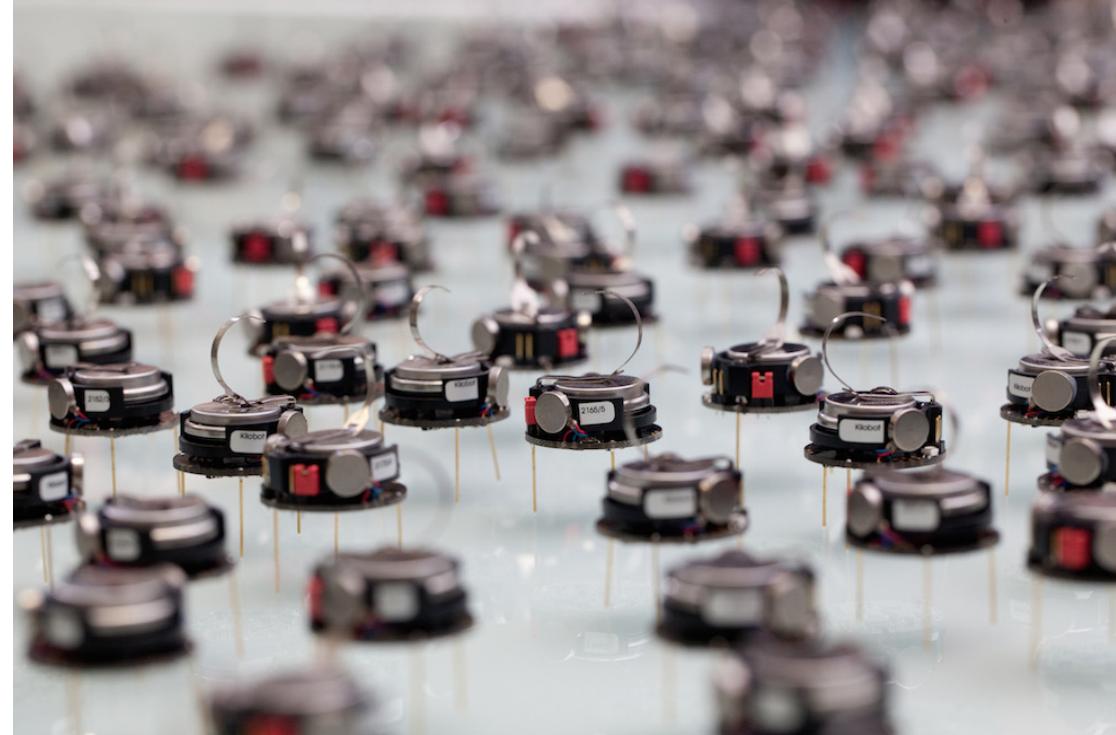
视频网址：<https://youtu.be/2rXrGH52aoM?feature=shared>

时间安排	教学内容	内容负责人
第1周	具身智能简介	具身智能的定义 历史背景和关键技术发展 具身智能与传统人工智能的区别
第2周	伦理隐私安全	个人隐私与安全 安全技术与隐私保护 具身智能伦理挑战
第3-6周	机器学习基础	机器学习基础 强化学习 深度学习与大模型
第7-8周	感知技术	手势与动作识别 视觉认知技术 语音交互技术
第9-10周	感觉运动系统	各种类型的传感器（视觉、触觉、听觉等） 马达系统和执行器 机器人的感觉运动整合
第11-12周	认知架构	具身人工智能的认知架构概述 大模型智能体技术 记忆和学习与智能体认知 端到端：感知到动作
第13-14周	导航与制图	空间意识和SLAM（同时定位与制图）技术 路径规划和避障 导航与任务导向决策的整合
第15周	人机交互	有效人机交互的原则 情感与社交机器人
第16周	群体智能	群体机器人 机器人与人类社会

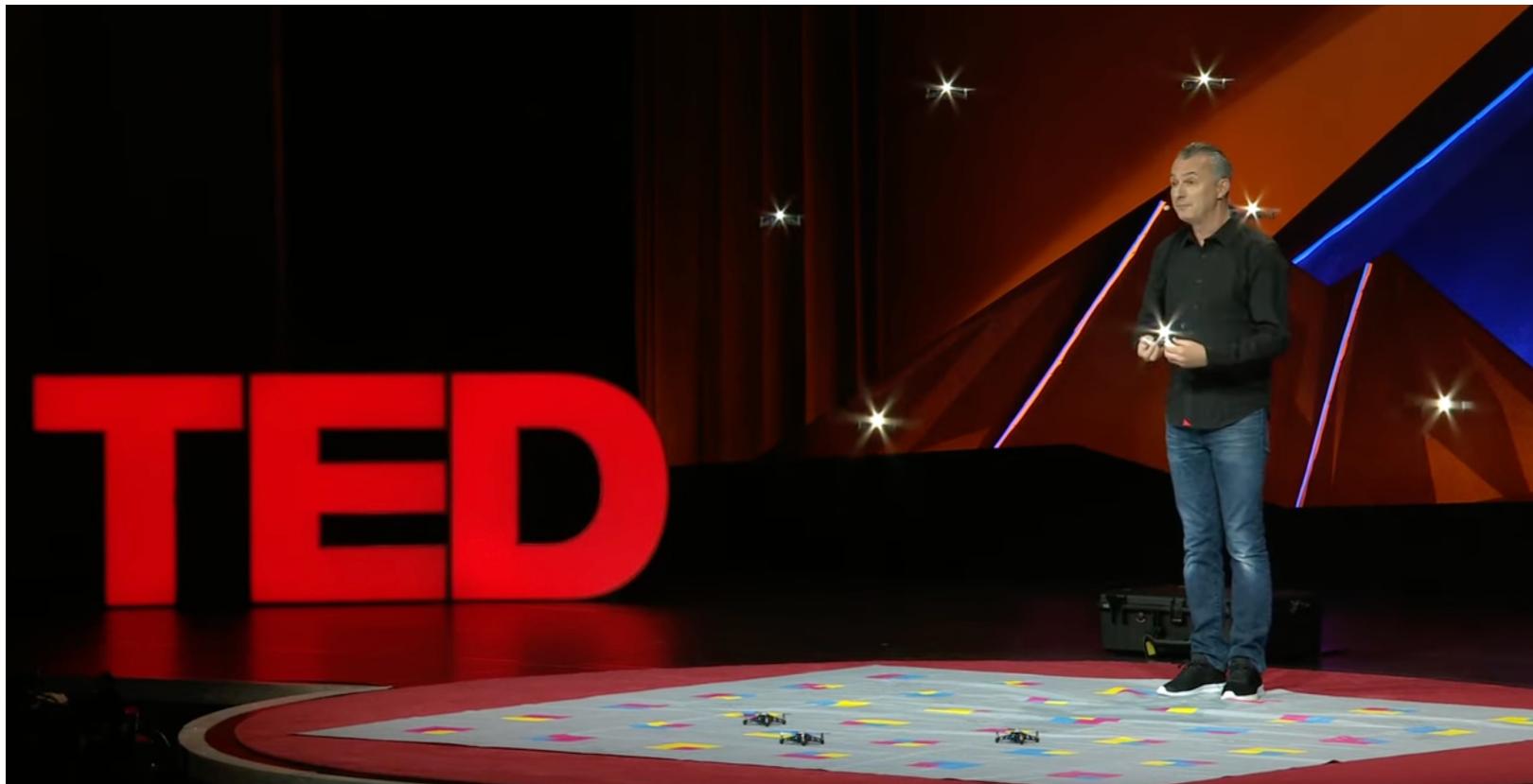
# 群体智能

群体机器人 (Swarm Robotics) 是指多个机器人个体组成的群体，通过相互之间的协作或竞争来完成任务。群体机器人系统模仿自然界中的群体行为（如蜜蜂、蚂蚁等），利用简单的个体行为来实现复杂的整体任务。

- **无人机群 (Drone Swarms)**：多架无人机可以组成一个群体，协同执行任务，如灾后搜救、农业监测、环境监控和军事侦察。这些无人机通过无线通信彼此协调行动，分配任务以覆盖更大的区域或执行更复杂的操作。
- **自动化仓库机器人**：在现代自动化仓库中，成百上千的移动机器人协同工作来搬运和整理库存。这些机器人通过中央控制系统协调，避免碰撞，优化路径，并高效完成货物的运输和排序。
- **海底探测机器人群**：一组水下机器人可以一起工作，进行大面积的海底地形勘探、管道检查或水质监测。每个机器人负责特定的区域，数据可以实时共享，以生成更全面的环境模型。



# A swarm of mini drones makes ... magic!



视频网址：<https://youtu.be/u2bQSKvZ2qI?feature=shared>

# 机器人与人类社会

机器人在现代社会中的角色和影响日益扩大，对人类社会的多个方面产生了深远的影响。这种影响既有积极的方面，也有潜在的挑战。

## ● 机器人在日常生活中的应用

- **家庭服务机器人**：越来越多的家庭开始使用服务机器人来完成日常任务，如清洁、烹饪、监控和家庭安全。
- **陪伴和护理机器人**：机器人被设计用于陪伴老人和儿童，提供心理安慰和社交互动，帮助老年人和残疾人进行日常护理，如辅助行走、提醒用药和健康监测。

## ● 工业和商业领域中的机器人

- **制造业自动化**：工业机器人在制造业中扮演着重要角色，承担着装配、焊接、喷涂、搬运等重复性和高精度任务。这种自动化大大提高了生产效率和产品质量，同时降低了生产成本。
- **物流和仓储管理**：物流机器人用于自动化仓库管理、货物搬运和配送，大幅提高了物流效率。亚马逊等企业已经广泛应用机器人技术来优化仓储和配送流程。
- **农业机器人**：农业机器人用于播种、收割、施肥和监控作物生长，帮助农民提高农业生产力和作物管理的精度，减少对劳动力的依赖。

## ● 医疗和健康领域的机器人

- **手术机器人**：机器人手术系统，如达芬奇手术机器人，提供了高精度和最小侵入性的手术方式，帮助外科医生执行复杂的手术任务，提高了手术的成功率和患者的恢复速度。
- **康复机器人**：这些机器人帮助患者进行物理康复训练，通过重复性和精准的运动辅助患者恢复功能。康复机器人广泛用于中风患者、受伤者的康复训练中。
- **辅助诊断机器人**：利用AI技术的机器人可以帮助医生进行诊断，如影像分析、疾病筛查和病理检测，提高了诊断的准确性和效率。





AI通识基础课程 (AIB310002)

# 具身智能引论

Introduction to Embodied Intelligence

丁文超 李维 浦剑 商慧亮 尚笠 杨帆