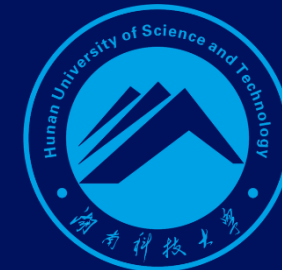




湖南科技大学
Hunan University of Science and Technology



人工智能基础

主讲人：康国胜

电 话：18202105624

Email：guoshengkang@gmail.com

个人主页：<https://guoshengkang.github.io>

实验室主页：<https://intelligentservicelab.github.io>

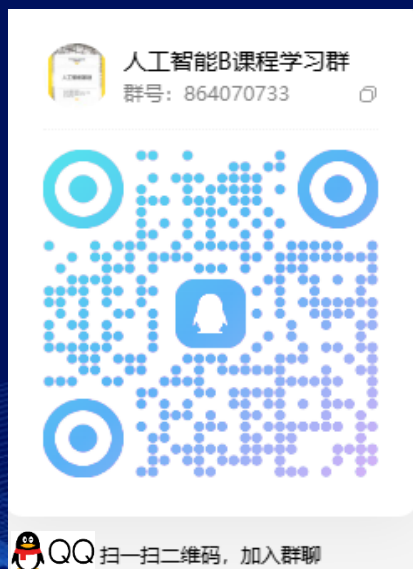
办公室：逸夫楼659





湖南科技大学
Hunan University of Science and Technology

第1章 人工智能概述





学完本章，应能用一张地图解释人工智能：它是什么、怎样发展、由哪些技术构成、正在改变哪些场景。

概念

先搭框架

历程

看清演进

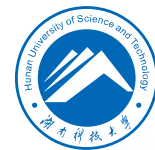
体系

理解方法

应用

判断边界





湖南科技大学
Hunan University of Science and Technology

1.1 人工智能的发展历程





发展历程：四次范式跃迁



人工智能发展历程与关键技术支撑

不是线性上升

AI 历史包含突破、低谷和再爆发，每次浪潮都依赖新的技术条件。

范式不断迁移

从逻辑推理到专家知识，再到统计学习、深度学习和大模型。

技术彼此叠加

新方法不是完全替代旧方法，而是在复杂任务中重新组合。



早期探索：智能问题被形式化（上世纪50到60年代）





符号主义：早期 AI 的核心思路

核心假设

智能可以通过符号、逻辑和规则进行表达；只要知识和推理规则足够完整，机器就能求解复杂问题。

典型局限

现实问题充满模糊、例外和不确定性，手写规则很难穷尽；这也为后续学习方法登场留下空间。

优势是可解释，难点是难以覆盖开放世界。



知识工程：把专家经验装进系统（上世纪70年代）

知识工程的价值

DENDRAL、MYCIN、XCON 等系统说明：当领域知识被结构化表达后，机器可以在专业任务中接近专家水平。

Ref. P3

知识工程的瓶颈

规则获取成本高、维护难，面对开放环境时组合爆炸和泛化不足会迅速暴露。

这一阶段让人们看到“知识是智能的基础”。



机器学习：数据开始成为核心燃料（上世纪90年代）

数据增长

互联网普及和数据规模扩大，为统计学习提供训练材料。

算法成熟

SVM、随机森林、AdaBoost 等方法完善并广泛应用。

算力可得

云计算降低训练成本，使复杂模型训练更可行。

机器学习把“人工写规则”转向“从数据中找规律”。



从规则到学习：AI 范式的变化

	规则驱动	数据驱动	模型驱动
知识来源	专家显式总结	样本统计规律	大规模参数学习
适用任务	边界清晰、规则稳定	可标注、可验证任务	复杂、多模态、开放任务
主要挑战	知识获取困难	数据质量和泛化	算力、对齐与可靠性

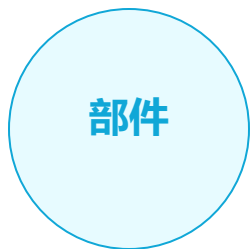


深度学习突破了人工特征瓶颈（2012年以后）

深度学习的关键，不只是“层数更深”，而是让模型自动学习层级化特征。



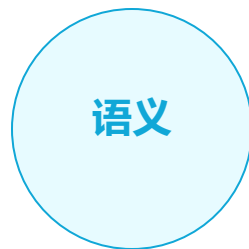
低层信号



结构组合



目标识别



高层理解



生成式 AI：智能开始承担创作任务

文本生成

对话、写作、
摘要、翻译、
代码辅助。

图像生成

海报、概念图
、风格迁移与
图像修复。

视频生成

动态内容生成
拓展多模态创
作边界。

智能体

从回答问题走向
规划、调用工具
和执行任务。

生成式 AI 让 AI 从“**识别与预测**”进一步走向“**生成与协作**”。



历史主线：每次突破都回答一个旧难题

规则不足

知识工程尝试用专家知识补足早期系统的推理能力。

知识难获

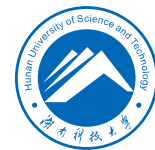
机器学习通过样本数据缓解手工建模成本。

特征难设

深度学习用多层网络自动抽取复杂表示。

任务更开放

大模型和生成式 AI 通过预训练、多模态和工具调用拓展任务范围。



湖南科技大学
Hunan University of Science and Technology

1.2 人工智能的定义与分类





人工智能的基本定义

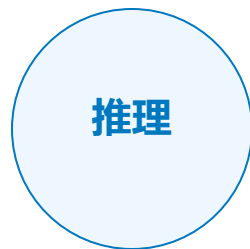
人工智能研究如何利用机器模拟、延伸和扩展人类智能，使其在特定任务中表现出类似人类的能力。



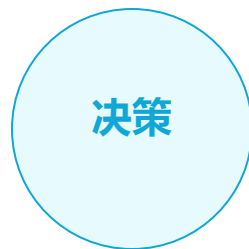
获取信息



形成模型



解决问题



采取行动



“人工”与“智能”：定义中的两个关键词

人工

强调机器、程序、算法和工程系统，是人类设计与制造出来的智能载体。

智能

强调感知、记忆、学习、推理、创造和适应环境的综合能力，不等同于单一计算速度。





AI 能力可以拆成一个协作系统



人工智能技术框架示意

外圈是总体目标

人工智能统摄机器学习、知识表示、自然语言处理等技术方向。

内圈是典型应用

人脸识别、语音助手、智能推荐、自动驾驶等都是由多项技术组合实现。

学习要看连接

不要孤立记概念，要理解能力之间如何相互支撑。



图灵测试与中文屋：一场关于“理解”的辩论

图灵测试

如果测试者通过对话无法分辨对方是人还是机器，就可认为机器表现出智能。它关注外在行为。



艾伦·麦席森·图灵 ---- 1950

中文屋思想实验

机器可能只是按照规则处理符号，即使外部看来像理解，也未必真正理解语义。



约翰·罗杰斯·塞尔 ---- 1980

Ref. P7

一个看表现，一个追问本质。



理解智能的三种理论视角

思维理论

智能源于思维活动，强调逻辑推理和抽象思考。

知识阈值理论

智能是在巨大搜索空间中快速找到满意解的能力。

进化理论

智能可由系统与环境持续互动而涌现。

这些视角帮助我们理解：智能既有规则推理，也有学习适应。



基于**实现维度**的人工智能层级划分

弱人工智能

面向特定任务，依赖明确目标、数据和模型；当前绝大多数应用属于弱人工智能。

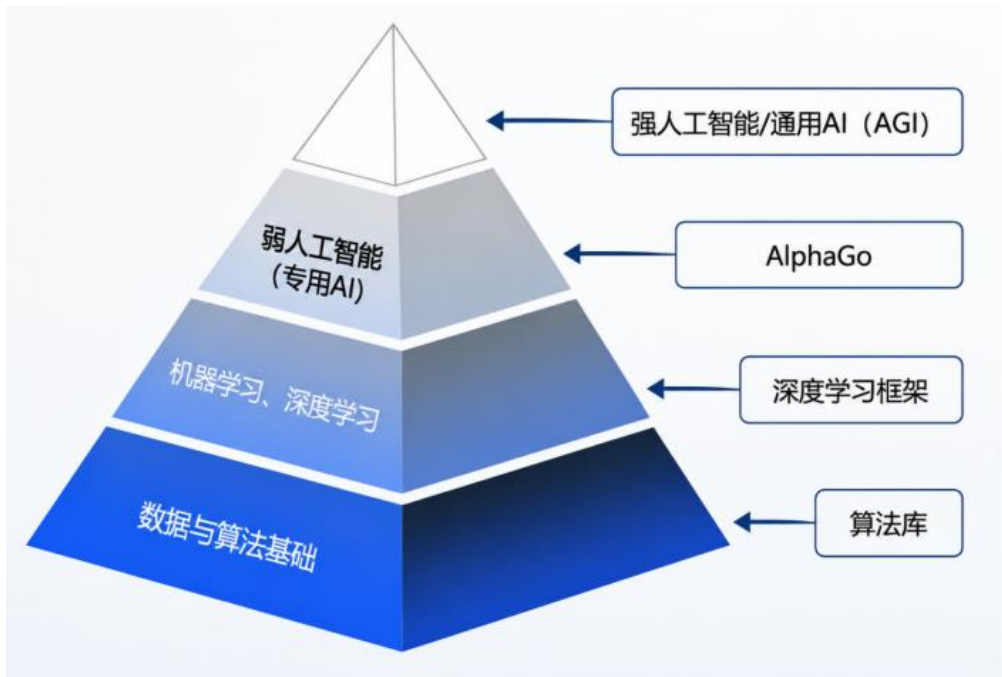
强人工智能

期望像人类一样具有通用理解、迁移、推理和自我适应能力，目前仍是长期研究目标。

现实主流是专用能力的持续增强。



人工智能的层级结构与技术架构



底层支撑

数据、算法库和学习框架为上层智能提供基础。

当前主体

弱人工智能和专用 AI 是现实应用主流。

长期目标

强人工智能/通用 AI 仍是理论和工程共同探索的方向。

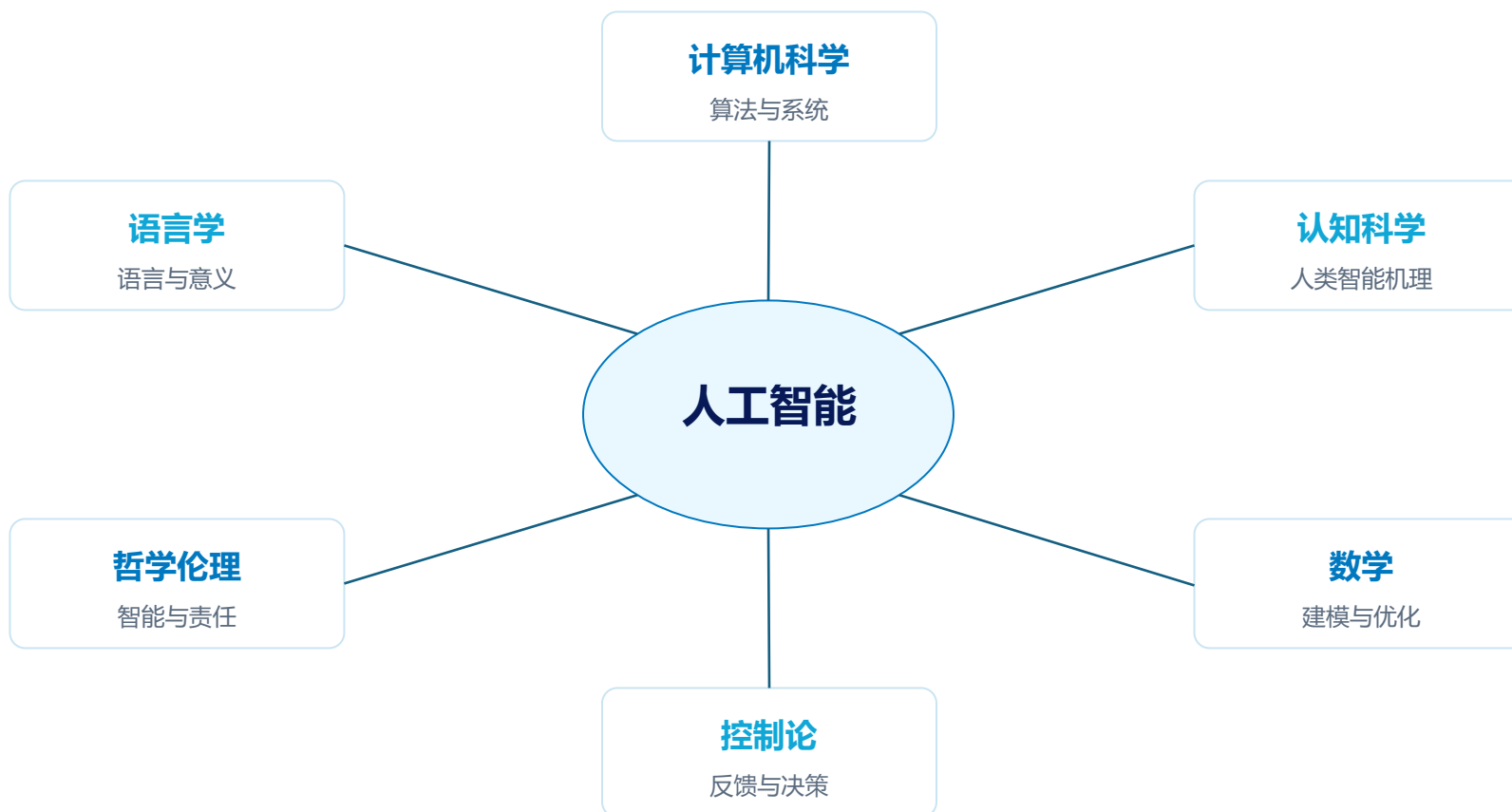


基于应用能力的人工智能分类：专用 AI 与通用 AI

	专用AI	通用AI
任务范围	聚焦单一或少数任务	面向广泛任务与开放环境
迁移能力	依赖任务边界和训练数据	希望跨领域迁移与组合推理
评价方式	用准确率、效率、稳定性 衡量	还需理解、规划、适应和安 全性评价
代表状态	已经大规模落地	仍处于探索阶段



人工智能是一门综合性交叉学科





相关学科如何支撑 AI

实现基础

计算机科学提供算法、系统、工程和部署能力。

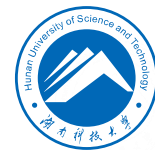
理论支撑

数学、控制论、信息论支撑形式化建模与反馈机制。

价值边界

哲学、伦理学、语言学帮助理解智能、意义与责任。

理解交叉性，才能避免把 AI 简化成单一算法。



湖南科技大学
Hunan University of Science and Technology

1.3 人工智能的核心知识体系





核心知识体系：从基础到应用

应用层

视觉、语言、语音、推荐、自动驾驶与行业智能

核心层

机器学习、深度学习、强化学习、神经网络与 Transformer

基础层

知识表示、问题求解、搜索算法、数据与算法基础



知识表示：机器理解世界的格式

知识的类型

陈述性知识描述事实，过程性知识描述操作过程，元知识控制知识如何被使用。

Ref. §2.1-P41

知识的边界

知识具有相对正确性和不确定性，必须放在具体场景和条件中理解。

机器先要“表示”，才能“推理”。



三种典型知识表示方法

产生式表示法

用 IF-THEN 规则表达因果关系，直观、可解释，常见于专家系统。

框架表示法

用“槽-侧面”描述对象属性，适合表达层级和继承关系。

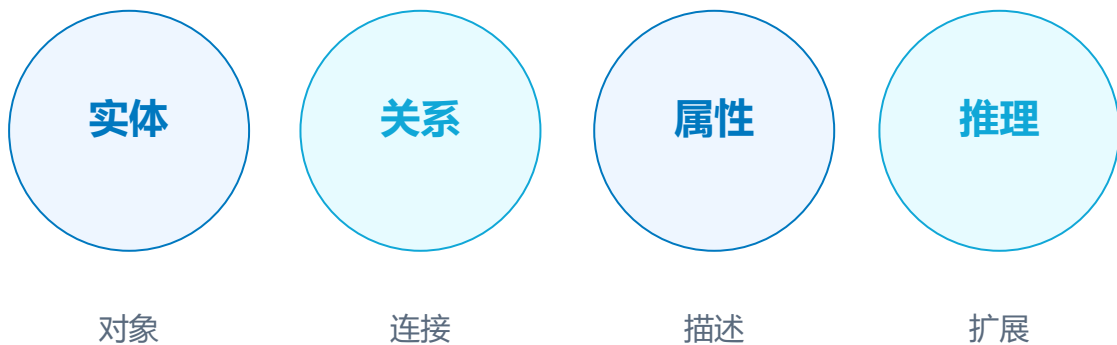
状态空间表示法

把问题转化为状态、操作、目标和代价构成的搜索空间。



知识图谱：从规则表到关系网络

知识图谱把知识组织为**实体**、**关系**和**属性**构成的网络，让机器能够在复杂连接中进行查询与推理。





问题求解：机器寻找答案的过程

初始状态

可选操作

目标测试

路径代价

1

2

3

4

明确任务从哪里开始。

定义每一步能做什么。

判断是否已经达到目标。

评估方案是否足够优。





搜索方法分类：盲目搜索、启发式搜索

盲目搜索

深度优先、广度优先和一致代价搜索不依赖额外领域知识，策略简单但效率可能较低。

启发式搜索

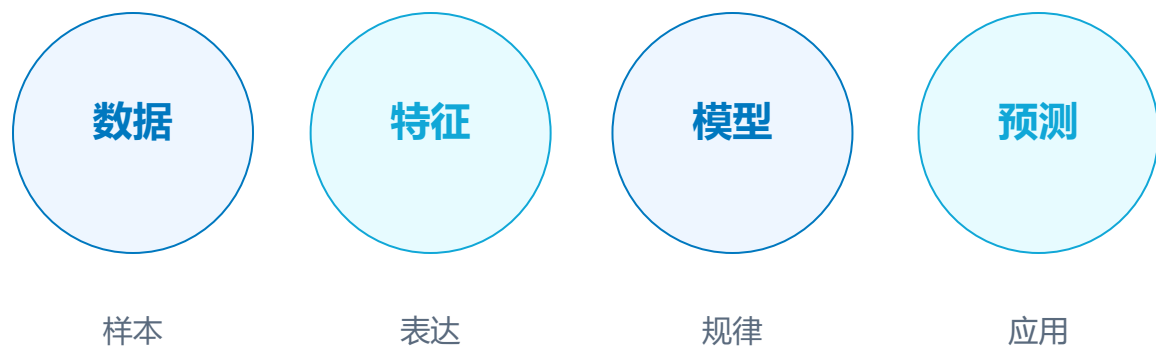
贪心最佳优先和 A* 搜索引入领域知识，用评价函数提升搜索效率并兼顾最优性。

启发式搜索更像“带着经验找路”。



机器学习：从数据中学习规律

机器学习把传统编程的“写规则”转变为“用数据训练模型”。





三类机器学习范式

监督学习

使用带标签数据，学习输入到输出的映射，适合分类和回归。

无监督学习

在无标签数据中发现潜在结构，适合聚类、降维和异常发现。

强化学习

智能体通过与环境交互获得奖励，优化长期决策策略。



机器学习的一般流程





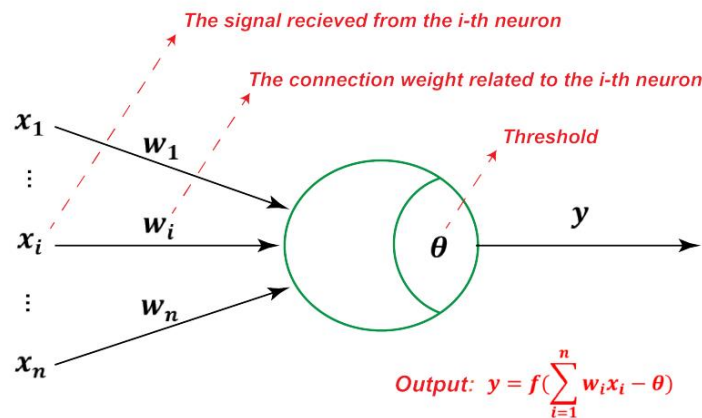
神经网络：从生物启发到数学模型

单个神经元

M-P 模型用输入、权重、阈值和输出模拟神经元的基本响应过程。

网络训练

前向传播得到预测，损失函数衡量误差，反向传播和梯度下降更新参数。



深度学习从这里向多层网络展开。



1.3 人工智能的核心知识体系——核心层

深度学习：逐层抽象复杂特征

深度模型像一层层筛选器，从底层信号中提取更高层、更接近语义的表示。





深度学习中的代表结构

CNN

适合图像局部特征提取和视觉任务。

RNN/LSTM

适合序列建模，曾广泛用于文本与语音。

Transformer

依靠注意力机制支撑大模型发展。

深度强化学习

把深度网络与试错决策结合。



计算机视觉：让机器看见并理解世界

基础表示

数字图像由像素、通道和颜色空间构成，机器通过矩阵形式处理视觉信息。

核心任务

图像分类判断是什么，目标检测定位在哪里，图像分割刻画像素级边界。



视觉任务从“看见”走向“看懂”。



自然语言处理：从识词到理解

复杂应用

机器翻译、智能问答、文本摘要、对话系统和文本生成

语义理解

词义消歧、指代消解、情感分析和语义相似度

基础解析

分词、词性标注、命名实体识别和句法分析



AI 大脑的语言中枢：自然语言处理

让机器理解文字、生成语言，并完成语言任务





1.3 人工智能的核心知识体系——应用层

智能语音处理：听懂、说出与识别身份

感知层

表达层

认知层

部署层

1

采集、数字化、预处理和特征提取。

2

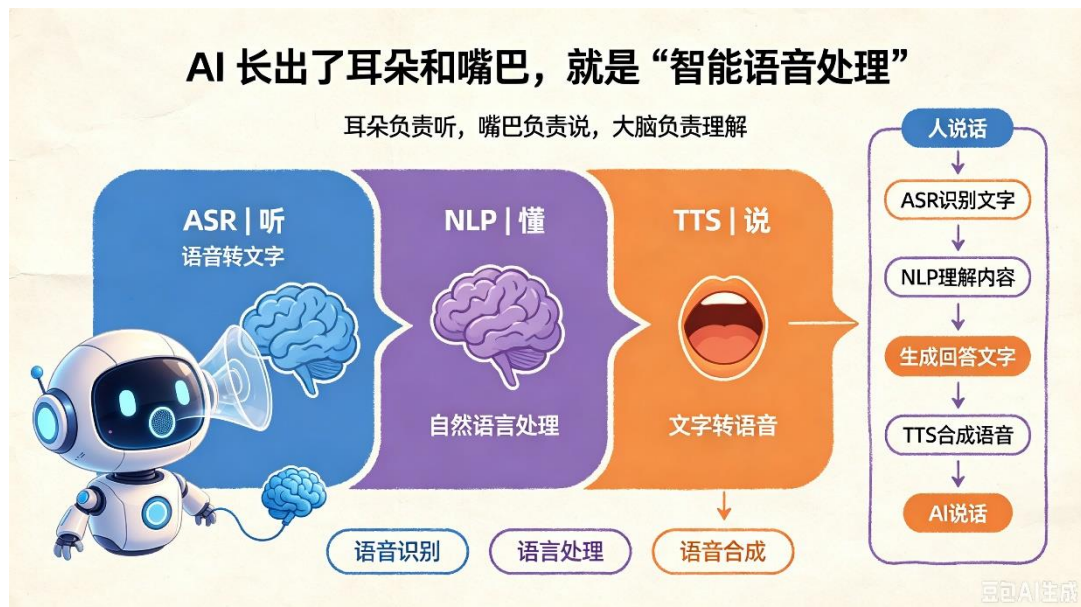
语音识别与语音合成完成声文转换。

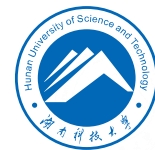
3

声纹识别、情感分析与语音唤醒。

4

云端与端侧协同兼顾效率和隐私。





湖南科技大学
Hunan University of Science and Technology

1.4 人工智能的应用领域





1.4 人工智能的应用领域

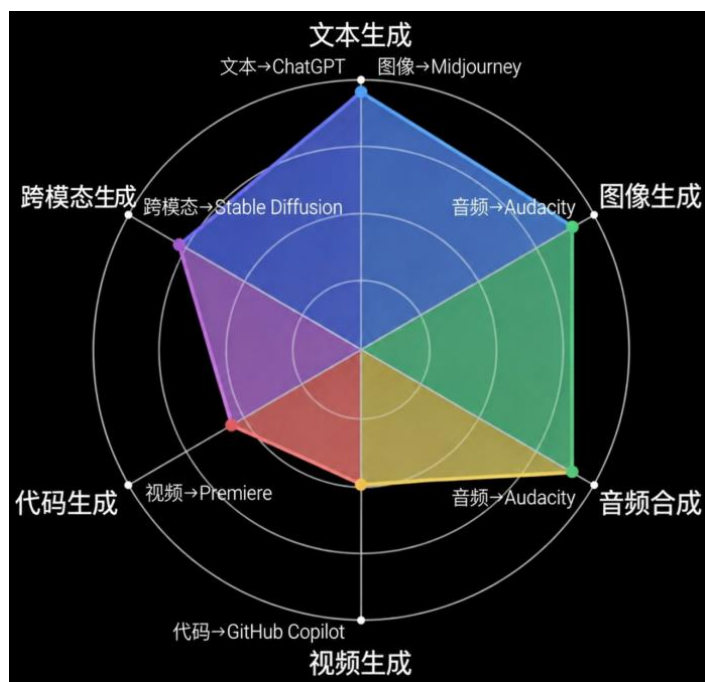
智能游戏与人机对弈：AI 的试验场





1.4 人工智能的应用领域

生成式人工智能：内容生产的新范式



AIGC 多模态生成能力示意

能力扩展

AIGC 覆盖文本、图像、音频、视频、代码和跨模态生成。

范式变化

AI 从分析器和分类器，进一步成为内容创造工具。

风险并存

真实性、版权、偏见和滥用问题需要治理机制配套。

文本生成(ChatGPT)、图像生成(Midjourney)、音频合成(Suno)、视频生成(Sora)、代码生成(GitHub Copilot)和跨模态生成(Stable Diffusion)



1.4 人工智能的应用领域

AIGC 的典型流程

从一句需求到一份作品，生成式 AI 往往经历“输入意图—模型生成—人工筛选—再编辑”的循环。



这是未来必须要掌握的“人机协同”工作模式。同时也带来了版权、隐私和“AI幻觉”等新问题。



1.4 人工智能的应用领域

智能交通与车联网

交通系统智能化

智能交通把感知、决策、通信和控制技术用于城市出行，提高安全性与效率。

车联网协同

车辆、道路、云平台和终端共享信息；联邦学习可在隐私保护下支持协作建模。

交通场景是 AI 落地的重要考场。



1.4 人工智能的应用领域

具身智能与机器人

具身智能

通过身体、传感器和行动与物理世界交互，形成感知-行动回路。

机器人

在制造、服务、医疗等场景中承担感知、规划和执行任务。

关键挑战

真实世界具有动态、不确定和安全约束，需要多模态感知与稳健控制。



1.4 人工智能的应用领域

其他典型应用场景

智慧医疗

辅助影像诊断、药物发现与个性化健康管理。

金融科技

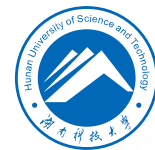
风控、反欺诈、智能投顾与客户服务。

智能制造

质检、预测维护、排产优化和柔性生产。

智慧城市

交通治理、公共安全、能源管理和城市服务。



湖南科技大学
Hunan University of Science and Technology

1.5 人工智能的技术生态与工具





Python 与数据工具降低了 AI 门槛



Python 生态与创新民主化示意

语言入口

Python 语法简洁、生态完备、社区活跃，适合 AI 入门和快速原型。

工具链协同

NumPy、Pandas、可视化、OpenCV、NLTK 等工具连接数据和模型。

学习重点

会调用工具，也要理解问题建模、数据质量和算法边界。



框架、平台与专业工具：本章收束



AI 能力超市与大模型服务平台示意

框架支撑

TensorFlow、PyTorch、MindSpore、PaddlePaddle 等框架降低模型开发门槛。

平台服务

大模型平台把模型能力封装成 API 或网页服务，支持文本、图像、代码和多模态任务。

本章主线：从概念、历程到核心技术、应用领域和工具生态，AI 学习要同时关注能力边界与责任。

人工智能全局地图

第四层：应用场景

具身智能与机器人

行业赋能

医疗金融制造等

智能交通与车联网

智能游戏与对弈

问题求解盲目与启发搜索

基础层

知识表示规则与图谱

智能语音处理

计算机视觉CV

能力层

机器学习监督无监督强化

核心层

深度学习CNN与Transformer

自然语言处理NLP

第三层：核心技术

第一层：概念与本质

智能层级

强人工智能通用AGI

弱人工智能专用AI

核心定义

感知、学习、推理、决策

哲学辩论

中文屋实验内在理解

图灵测试外在表现

第五层：生态与工具

数据工具

NumPy与Pandas

底层框架

PyTorch与TensorFlow等

语言入口

Python

大模型平台

API调用与能力超市

知识工程

专家系统知识驱动

机器学习

深度学习与大模型

统计规律数据驱动

自动抽象模型驱动

早期探索

符号主义规则驱动

THANKS

谢谢

