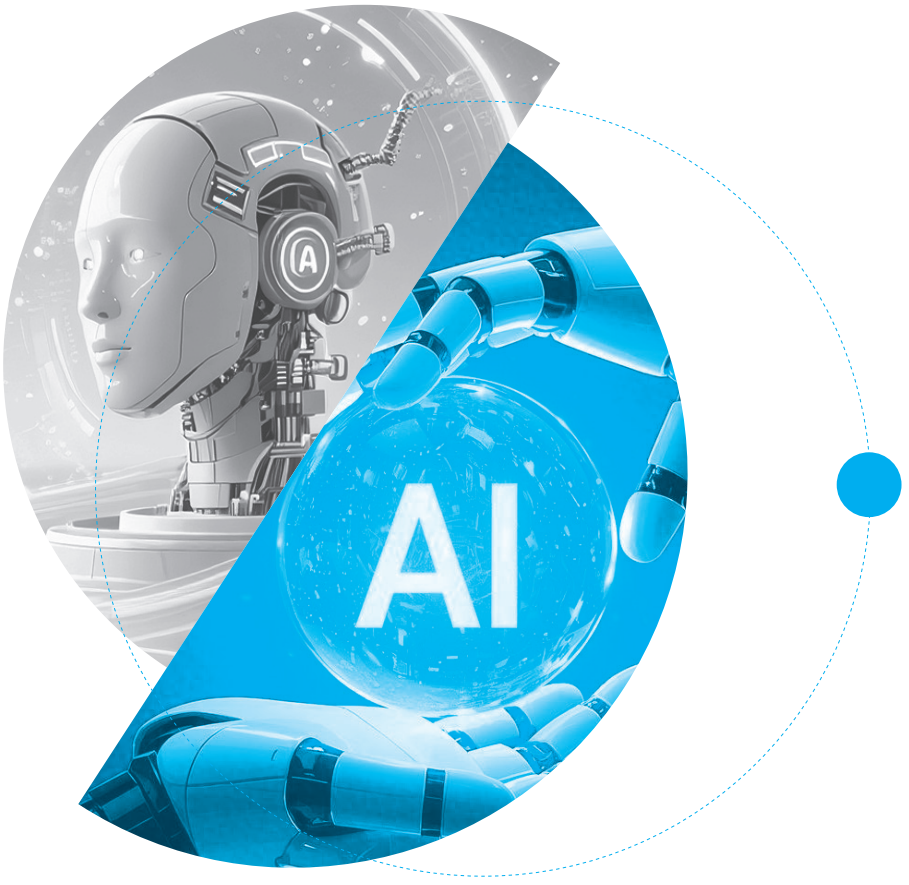
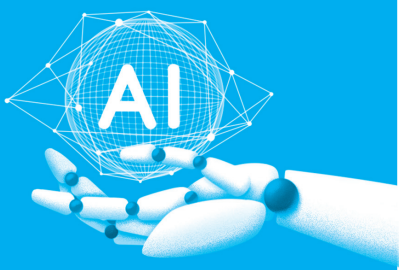


第 1 部分

走进智能时代
理解国家战略



第 1 章 认识人工智能，把握时代浪潮

知识图谱

学习目标

知识目标

1. 理解人工智能基本概念、研究目标，能区分其与传统程序系统的核心差异。
2. 了解人工智能发展历程及阶段特征，认识算力、数据对其发展的关键作用。
3. 了解国家“人工智能 +”行动背景与内涵，把握人工智能与产业融合趋势。

能力目标

1. 能从系统视角分析人工智能应用场景，识别人工智能在系统中的关键作用环节。
2. 能结合案例，对人工智能应用的价值与局限进行初步分析判断。
3. 能围绕场景提出“人工智能 +”应用设想，并说明基本技术逻辑。

素养目标

1. 理性认识人工智能发展，不神化也不过度否定。
2. 树立人工智能应用的伦理与安全意识，关注数据隐私、公平性等问题。
3. 培养严谨求实、尊重技术规律的科学态度，明确科研与工程实践的责任担当。

场景启思

使用手机刷脸支付时，摄像头瞬时扫描面部便能完成身份核验；早晚高峰的城市道路上，无人出租车能够自动识别路况、规划路线并安全行驶；校园中，AI 门禁系统能够快速识别师生身份并自动放行。这些日常场景正在不断提醒我们：人工智能已经从实验室中的研究课题，逐渐走进现实生活，成为推动社会发展的重要技术力量。

为了更直观地理解人工智能的作用，我们通过两个贴近生活的案例，了解人工智能系统是如何在实际场景中发挥作用的。

案例 1：人脸识别智慧校园系统

在许多高校中，传统的刷卡门禁逐渐被人脸识别系统取代。师生只需在门禁设备前短暂停留，系统便能够自动识别身份并完成通行验证。

这一过程依赖人工智能的人脸识别技术：系统首先建立师生的面部特征数据库，当摄像头捕捉到面部图像时，算法会对图像中的关键特征进行分析，并与数据库中的信息进行快速比对，从而确认身份。识别成功后，系统自动打开门禁，同时记录出入时间和人员信息。

在实际应用中，这一技术不仅用于校园门禁，还被应用到宿舍管理、图书馆借阅以及校园支付等场景，使校园管理更加高效，也为师生提供了更加便捷的服务体验。

案例 2：城市无人出租车

近年来，无人出租车逐渐出现在一些城市的道路上。用户只需通过手机下单，车辆便可以自动前往指定地点接送乘客，并在没有司机的情况下完成整个行程。

无人出租车的运行依赖自动驾驶技术。车辆通过摄像头、激光雷达等传感设备实时获取道路环境信息，例如交通信号灯、道路标线以及周围车辆和行人的位置。人工智能算法对这些信息进行分析和判断，从而规划行驶路线并控制车辆的加速、转向和制动，实现安全驾驶。

随着技术的发展，无人驾驶系统正逐渐从实验测试走向实际应用，成为未来智能交通的重要发展方向。

从智慧校园到无人出租车，这些真实场景展示了人工智能在不同领域的应用潜力。人工智能不仅能够提升效率，还正在改变人们的生活方式和产业运行模式。接下来，我们将从人工智能的定义、发展历程以及技术体系等方面，系统了解这一正在深刻影响社会发展的重要技术。

人工智能的定义、内涵与里程碑

1.1.1 人工智能的定义与内涵

人工智能（Artificial Intelligence，AI）的诞生，离不开 1956 年美国达特茅斯学院那场具有里程碑意义的研讨会。正是在这次会议上，美国计算机科学家约翰·麦卡锡（John McCarthy）联合多位跨学科学者，首次正式提出“人工智能”这一概念，为这一新兴领域划定了探索边

界。在他们看来，人工智能的核心是打造具备高级模拟能力的机器——这类机器无须复刻人类的意识本身，却能精准复制、模拟那些可被清晰描述、明确界定的人类学习特质与智能表现，让机器在特定场景中展现出类人智能的核心功能。

作为人工智能领域的重要开创者，美国斯坦福大学人工智能研究中心的尼尔斯·约翰·尼尔森（Nils John Nilsson）教授，从学科本质的角度深化了对人工智能的认知。他指出，人工智能是关于知识的学科——怎样表示知识以及怎样获得知识并使用知识的学科。

美国麻省理工学院的帕特里克·温斯顿（Patrick Winston）教授则从应用价值角度对人工智能作出了更通俗的界定。在他看来，人工智能的核心目标十分明确，就是通过技术创新，让计算机具备胜任以往唯有人类依靠智力才能完成的工作的能力。无论是复杂的逻辑推理、精准的模式识别，还是创造性的任务处理，只要是过去需要人类智力介入的场景，都属于人工智能的探索与应用范畴。

这些表述虽视角各异，却共同勾勒出人工智能的核心轮廓：它并非单一技术或工具，而是一门融合计算机科学、数学、心理学、神经科学、语言学等多学科的交叉前沿领域，以模拟人类智能为核心、以拓展机器能力边界为目标，最终通过技术赋能，让机器更好地辅助甚至替代人类完成复杂任务。而这一目标的实现，离不开人工智能具备的四大核心内涵（见图 1–1），它们是人工智能落地应用的核心支撑。

图 1–1　人工智能内涵图

（1）感知和交互能力：如同人类的“五官”，是 AI 与外部世界建立连接的基础。AI 通过摄像头、麦克风、传感器等设备，能够精准捕捉图像、语音、环境数据（如人脸识别系统捕捉面部特征、无人出租车通过雷达感知路况）；同时能够以人类易懂的方式反馈结果，例如语音助手应答指令、智能设备执行操作，实现“输入需求—接收反馈”的双向互动。

（2）学习与适应能力：这是 AI 突破固定程序、实现自我进化的关键。AI 无须人工逐一预设规则，而是通过分析海量数据自主发现规律、优化性能。例如，推荐算法通过学习用户浏览习惯持续优化推送内容，医疗 AI 通过学习数百万份病历不断提升诊断准确率，机器在实践中不断适配新场景、响应新需求。

（3）推理和决策能力：相当于 AI 的“大脑”，是其处理复杂问题的核心。基于感知到的数据和学习到的规律，AI 能进行逻辑分析、判断与选择。例如，智能导航结合实时路况、天气、车流量等多维度数据，推理出最优出行路线；金融 AI 通过分析用户信用信息与交易记录，决

策是否批准贷款申请，高效应对复杂场景中的决策需求。

（4）自主行动能力：将推理与决策转化为具体动作的执行能力，是 AI 落地应用价值的最终环节。从工业机器人根据决策指令完成精密零件组装，到无人出租车根据决策指令实现转向、避让、平稳行驶，再到智能设备根据语音指令同步日程、控制家电，都体现了 AI“把想法变成行动”的核心价值，让智能从理论走向实践。

伦理思辨

随着人工智能技术的快速发展，越来越多原本需要人类判断与决策的任务开始由智能系统参与完成。人脸识别技术已被广泛应用于校园门禁与公共安全管理，自动驾驶技术也正在逐步进入城市交通系统。这些技术在提升效率和便利性的同时，也引发了一系列新的社会问题。

例如，人脸识别系统需要采集并存储大量个人生物特征数据，如果管理不当，可能带来隐私泄露风险；自动驾驶系统一旦发生事故，责任应如何界定也成为社会关注的焦点。此外，越来越多工作被自动化系统替代，也可能对就业结构产生影响。

思考与讨论：

（1）在校园或公共场所使用人脸识别技术时，应如何在安全管理与隐私保护之间取得平衡？

（2）如果自动驾驶车辆发生交通事故，责任应由技术开发者、运营方还是使用者承担？

（3）当人工智能不断提高自动化程度时，社会应如何应对由此带来的就业结构变化？

微课视频

1.1.2 人工智能的发展历程

从人工智能概念的提出，到如今广泛应用于生活方方面面的各类智能技术，人工智能的发展并非一帆风顺，而是历经了“萌芽—起伏—爆发”的多阶段迭代，每一次突破与停滞，都推动着技术边界的拓展。

1. 萌芽期：概念诞生与初步探索（20 世纪 50—60 年代）

人工智能的思想源头可追溯至 1950 年艾伦·图灵发表的《计算机器与智能》一文。文中提出的“图灵测试”，首次为“机器是否具备智能”提供了可验证的标准——若人类无法通过对话区分交流对象是机器还是人类，则可认为该机器具备智能，这一理念为后续研究奠定了思想基础。

1956 年，达特茅斯会议的召开标志着人工智能正式成为一门独立学科。会议上，约翰·麦卡锡等学者明确了人工智能的研究目标：“让机器模拟人类的学习、推理等智能行为。”此后，这一领域迎来早期探索的小高潮：1957 年，赫伯特·西蒙与艾伦·纽厄尔研发出“逻辑理论家”程序，首次实现机器自动证明数学定理，验证了机器模拟人类推理的可行性；1966

年，聊天机器人 ELIZA 问世，通过模仿心理治疗师的对话方式与人类互动，让大众首次直观感受到机器具备类人交流能力。

这一阶段的人工智能虽仅能完成简单任务，却为领域发展指明了核心方向：让机器具备类人的智能行为。

智领未来

2. 起伏期：技术瓶颈与行业寒冬（20 世纪 70—90 年代）

早期 AI 的快速探索，让学界与产业界对其抱有过高期待，但技术与现实的落差，让这一领域进入了两次“寒冬期”。

1）第一次人工智能寒冬（20 世纪 70—80 年代初）

受限于当时的算力水平（仅能支撑简单运算）与数据规模（缺乏海量训练数据），早期人工智能系统（如基于规则的专家系统）仅能处理特定领域的单一问题，难以应对复杂、动态的现实场景。例如，某医疗专家系统仅能诊断特定类型的肺炎，却无法适应患者的个体差异。

受技术局限性与高昂研发成本的双重制约，政府与企业纷纷削减投入，AI 领域由此步入第一个寒冬。

2）短暂复苏与第二次人工智能寒冬（20 世纪 80 年代中期—90 年代）

20 世纪 80 年代中期，“神经网络”算法的优化与算力的小幅提升，让人工智能迎来短暂复苏：日本提出“第五代计算机计划”，试图打造具备推理能力的智能计算机；美国研发出可进行语音识别的人工智能系统。

但此次复苏仍未突破核心瓶颈——人工智能系统的泛化能力极弱（仅能适配训练场景），商业化价值未达预期。20 世纪 90 年代后，“第五代计算机计划”失败，企业研发项目终止，人工智能再次陷入寒冬，领域发展放缓。

这一阶段的起伏，让学界意识到：AI 的发展不能仅依赖理论推演，更需要算力、数据等基础条件的支撑。

榜样力量

图灵——探索机器智能的先驱

艾伦·图灵（Alan Turing）是英国著名数学家和计算机科学先驱。从学生时代起，他就对“机器是否能够像人一样思考”这一问题产生了浓厚兴趣，并长期坚持研究计算与智能的关系。在那个计算机尚未普及的年代，图灵大胆提出了许多富有前瞻性的思想。

第二次世界大战期间，图灵参与密码破译工作，在极其紧张和复杂的环境中带领团队攻克技术难题，为战争胜利作出了重要贡献。战后，他继续投身计算机和人工智能研究，并提出了著名的“图灵测试”，为后来人工智能的发展奠定了重要思想基础。图灵勇于探索未知领域、坚持科学研究的精神，激励着一代又一代科研工作者不断挑战技术边界、推动人类科技进步。

3. 爆发期：技术突破与全面普及（2000 年至今）

2000 年后，算力的指数级提升（GPU、TPU 等专用芯片普及）、互联网带来的海量数据，以及深度学习算法的成熟，共同推动人工智能进入爆发期。

1）技术突破：深度学习开启新赛道（21 世纪 10 年代初）

2012 年，深度学习模型 AlexNet 在图像识别竞赛中，将错误率从 26% 降至 15%，远超传统算法。这一突破证明：基于神经网络的深度学习，能让机器从海量数据中自主学习复杂规律。此后，深度学习成为 AI 的核心技术支撑，推动图像识别、语音识别等领域算法准确率快速逼近甚至超越人类水平。

2）大众认知：标志性事件破圈（21 世纪 10 年代中期）

2016 年，AlphaGo（基于深度学习的围棋 AI）击败世界围棋冠军李世石，让全球意识到 AI 已具备复杂策略决策能力——围棋的可能性走法远超宇宙原子数量，此前被认为是“人类智能的最后堡垒”。这一事件让 AI 从专业领域走进大众视野，成为社会关注的焦点。

3）全面普及：通用智能与场景渗透（2020 年至今）

2020 年后，以 GPT 系列、文心一言为代表的大语言模型爆发，结合多模态技术，让 AI 从“专用人工智能”向“通用人工智能”迈进：生活中，生成式人工智能（AIGC）工具可生成图文、音视频内容；产业中，人工智能已渗透医疗、制造、交通等领域；社会层面上，“人工智能 +”成为国家战略，推动技术与各行业深度融合。

如今的 AI，已从实验室中的技术概念，转变为重塑生活、产业与社会的核心力量，其发展仍在加速——未来，通用人工智能的探索、AI 伦理的规范，将成为领域发展的新方向。随着人工智能逐渐进入社会运行的关键领域，人们不仅关注技术能力本身，也开始思考其可能带来的社会影响与伦理问题。

科技中国

我国人工智能发展的重要进展

近年来，中国在人工智能技术研究与产业应用方面取得了一系列重要进展。在语音智能领域，科大讯飞的语音识别和语音合成技术已经广泛应用于教育、医疗和智能设备等场景；在计算机视觉领域，人脸识别和图像识别技术在智慧城市、移动支付和公共安全等领域得到广泛应用。

在智能计算基础设施方面，国内企业持续推进人工智能计算平台的发展，例如华为昇腾人工智能处理器和相关计算平台，为大规模模型训练和智能应用提供了重要支撑。与此同时，中国在自动驾驶、智能机器人以及生成式人工智能等领域也不断取得新的进展。以大语言模型为代表的新一代人工智能技术，正在推动智能问答、内容生成和智能助手等应用快速发展。

这些技术突破不仅展示了我国在人工智能领域的创新能力，也为人工智能在各行业的广泛应用奠定了重要基础。

国家“人工智能 +”行动与产业融合趋势

随着人工智能技术不断发展，其角色已经从局部工具应用，逐渐扩展为具有广泛渗透力的通用技术。人工智能不再只存在于实验室或个别行业的试点项目中，而是越来越多地参与产业运行、公共服务和社会治理等核心环节。在这一背景下，如何系统性地推动人工智能与经济社会发展深度融合，成为国家层面必须作出的重要部署。

在 2024 年《政府工作报告》中，我国首次明确提出要开展“人工智能 +”行动，强调以人工智能为关键技术支点，推动其与制造业、服务业以及社会治理等领域深度融合，加快培育新质生产力。这一表述标志着人工智能已从“前沿技术布局”阶段，进入“全面融入经济社会发展”的新阶段。“人工智能 +”由此不再只是一个技术趋势概念，而成为国家推动高质量发展的一项现实行动路径。

需要指出的是，“人工智能 +”并不意味着在各行各业简单叠加一项智能技术。其核心并不在于是否应用了 AI，而在于人工智能是否真正融入业务流程和决策体系之中。与早期以单点功能为主的应用不同，当前的“人工智能 +”更强调通过数据、模型和算力的综合应用，对原有依赖经验和人工判断的运行方式进行重构。这种重构既体现在效率提升上，也体现在决策方式、服务模式和产业形态的变化之中。

在产业实践中，这种融合趋势正在不断显现。人工智能正从局部辅助工具，逐步演变为系统能力的一部分。在制造领域，人工智能不再仅用于某道工序或某类检测任务，而是逐步嵌入生产调度、质量控制和设备管理等整体流程；在医疗和教育等领域，人工智能更多承担分析和支持决策的角色，提升服务效率与公平性；在城市治理和公共服务中，人工智能通过对多源数据的综合分析，为精细化管理和科学决策提供技术支撑。可以看到，“人工智能 +”推动的并非某个行业的局部升级，而是一种以智能化为特征的产业融合趋势。

随着大模型和平台化能力的发展，人工智能的融合方式也在发生变化。通用模型为各类应用提供基础能力，而行业知识和业务规则的引入，使人工智能能够更好地适配具体场景。这种“通用能力与行业深度融合”的模式，使人工智能既具备跨领域迁移能力，又能够在具体应用中形成稳定价值，为“人工智能 +”的持续推进提供了现实路径。

然而，将人工智能从技术能力转化为可长期运行、可规模化复制的系统，并非一蹴而就。这一过程不仅依赖技术本身，更离不开科研与工程技术人员在复杂现实场景中的持续探索与实践。

从更宏观的角度看，“人工智能 +”行动不仅推动了产业形态的变革，也对人才结构和能力需求提出了新的要求。未来社会不仅需要少数从事算法研究的专业人才，更需要大量能够在各专业领域理解并合理使用人工智能技术的复合型人才。无论是工程、管理、医学，还是人文

艺术领域，人工智能都将成为从业者必备的能力支撑。

因此，理解“人工智能 +”行动，对于高校学生而言，其意义不仅在于了解一项国家政策，更是认识人工智能时代发展方向的重要途径。通过这一视角，学生可以更清晰地思考自身专业在智能时代的定位与发展路径，为未来的学习与职业选择奠定认知基础。

智领未来

从专用人工智能到通用人工智能

过去几十年，人工智能的发展主要集中在“专用人工智能”（Narrow AI）领域，即针对特定任务设计的智能系统。例如，人脸识别用于身份验证，推荐算法用于内容推荐，语音识别用于语音助手。这类系统在特定任务上可以达到很高的准确率，但通常只能完成单一类型的工作。

近年来，随着深度学习、大模型和算力技术的突破，人工智能正逐渐向“通用人工智能”（General AI）的方向演进。以大语言模型为代表的新一代人工智能系统，能够在同一平台上完成文本生成、知识问答、代码编写、图像理解等多种任务，并通过持续学习不断提升能力。这种跨任务、多能力的特征，使人工智能逐渐从“工具型系统”向“智能助手型系统”转变。

未来，人工智能还将进一步与机器人技术、物联网、大数据等领域深度融合，在医疗健康、智慧城市、智能制造等场景中发挥更大的作用。同时，随着技术能力的不断增强，如何在推动创新的同时加强技术治理、保障数据安全和促进社会公平，也将成为人工智能发展过程中需要持续关注的重要议题。


实训项目1-1

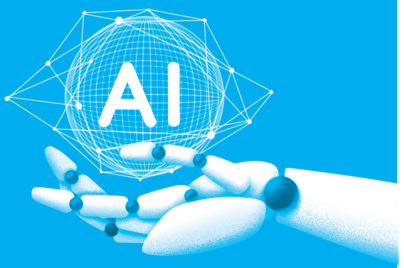

实训项目1-2


实训项目1-3


学以致用


综合评价表

第 2 章



人工智能思维构建，实现思维跃迁

知识图谱

学习目标

知识目标

1. 理解计算思维的内涵及分解、模式识别、抽象、算法设计“四要素”。
2. 明确计算思维的适用边界，理解“规则可写”与“规则难穷举”的差异。
3. 了解“人工智能＋专业”的融合路径及合规、可控、责任等关键边界。

能力目标

1. 能用计算思维对真实问题进行结构化拆解与流程化表达。
2. 能判断一个任务更适合数据驱动还是模型驱动，并说明理由。
3. 能结合具体专业场景，分析人工智能在系统中的技术落点与应用逻辑。

素养目标

1. 建立从“规则驱动”到“学习驱动”的人工智能思维框架。
2. 形成“技术＋场景＋约束”的系统观，强化对AI应用边界的认知。

场景启思

在校园里买一瓶水，将钱塞进自动售货机，机器会准确吐出商品并完成找零；回到宿舍，打开手机相册，系统会自动将照片分类成“人物”“美食”“风景”，甚至还能精准识别出

“猫”“狗”“夕阳”“海边”。这两件事从表面上看都体现了机器的“智能”，但它们背后的原理却完全不同。

一种能力来自规则与步骤：只要把流程写清楚，机器就能稳定执行，输入相同，输出必然相同；另一种能力来自数据与学习：我们很难把规则写出来，只能让系统从大量样本中学会判断。

为了更直观地理解这两种能力来源的差异，我们通过两个极具反差的日常案例，拆解它们的工作逻辑，看看“智能”究竟是如何产生的。

案例 1：自动售货机找零——规则驱动的“可计算系统”

“我投了 10 元，买 3 元的水，机器为什么一定会找我 7 元?”这件事非常典型地体现了计算思维。售货机的任务目标非常清晰。

输入：投入金额、商品价格。

输出：商品 + 找零金额。

约束：找零必须准确，不能多也不能少。

流程：按固定步骤计算差额，再根据硬币 / 纸币库存组合出找零方案。

在这里，机器既不需要“学习经验”，也不需要“理解人类”。只要把规则写清楚，它就能稳定执行：同样的输入永远得到同样的结果，并且每一步都可以追踪与验证。这类系统的智能本质是：把现实问题转化为结构化表达，再转化为可执行步骤。这正是计算思维最擅长的领域。

案例 2：手机相册自动分类——学习驱动的“感知智能”

“为什么手机能自动识别这张照片里是猫?”

如果让我们用规则写出来，通常会凭直觉这样描述：猫有耳朵、有尾巴、有胡须、有四条腿；猫的眼睛比较圆；猫的脸比较小。

但很快就会发现，这些规则根本不够用。因为现实中照片的情况太复杂：光照不同、角度不同、遮挡不同；猫可能趴着、蜷着、背对镜头；有些狗长得像猫，有些玩偶也像猫；同一只猫在不同姿态下外观差异巨大。

也就是说，这类问题真正的困难并不是“算不出来”，而是写不出来。当规则难以穷举时，系统只能换一种方式获得能力：收集大量图片样本，让模型从数据中学习猫的特征模式，并在学习过程中不断调整参数，逐步形成识别能力。因此，相册分类的结果往往不是绝对正确的，而是一种带置信度的判断：“这张图更像猫”“这张图更像狗”。它的输出既具有概率性，也允许存在一定误差。这类系统的智能本质是：用数据替代规则，让系统从经验中学习规律。这正是人工智能思维的核心。

计算思维与人工智能思维

2.1.1 计算思维的内涵与四要素

在信息化时代，计算机之所以能够深刻改变社会，并不是因为它“像人一样聪明”，而是因为它能够以极高的速度与稳定性执行规则。只要人类能够把规则讲清楚，把步骤写明确，机器就可以不知疲倦地重复执行，并且保证每次执行结果一致。这种把问题转化为可被机器执行的步骤的思维方式，就是计算思维。

很多同学对计算思维的第一印象是编程。但严格来说，计算思维并不等于写代码。写代码只是把思维落到机器上的一种方式，而计算思维本身是一种更普遍的能力：它要求我们在面对问题时，能够先把复杂情境“理清楚”，再把解决过程“写清楚”，最终让流程能够被自动化执行，如图 2–1 所示。

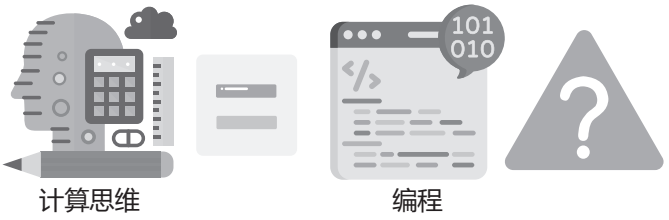


图 2–1　计算思维内涵图

为了让这种思维更直观，我们先从一个校园中典型的系统任务出发：排课。学校教务系统每学期都会生成课表，要同时考虑教师时间、教室容量、班级冲突、必修课优先级、实验课设备要求等大量约束因素。如果完全依赖人工安排，不仅效率低，而且容易出错、难以复核。而当这些约束被整理成清晰规则后，排课问题就可以交给系统自动求解。这里的关键并不是系统“理解教学”，而是人类把排课过程抽象成一个可计算的结构：哪些是输入（课程、教师、教室、时间段），哪些是约束（冲突、容量、优先级），哪些是输出（课表），以及如何在规则下寻找可行解。换句话说，排课系统并不依赖“智能”，它依靠的是规则化、结构化与流程化。这正是计算思维的典型体现。

从更本质的角度来看，计算思维的核心任务可以用一句话概括：把现实问题转化为结构化表达，再转化为可执行步骤。

在这一过程中，计算思维最常见、也最关键的思考方式通常包含四个环节：分解、模式识别、抽象与算法设计。它们并不是彼此割裂的步骤，更像是一套可以反复往返的工具箱：当你

觉得问题太复杂时，就先把任务拆开；当你发现不同子任务结构相似时，就寻找其中的共性模式；当你被大量细节淹没时，就提炼出真正影响结果的关键变量；当关键要素与关系逐渐清晰后，就可以把解决过程写成一套可执行、可重复的步骤，也就是算法。

（1）分解：把复杂问题分成若干可处理的子任务。

（2）模式识别：在不同子任务之间发现重复结构，提升复用与效率。

（3）抽象：忽略无关细节，提取关键变量与关系。

（4）算法设计：把解决方案写成清晰、可重复执行的步骤。

这四个环节并不是一次性线性执行的流程，而是一种可反复迭代的思维循环。如图 2–2 所示，在实际问题解决中，人们往往会在“分解—抽象—算法设计”的过程中不断修正，直到形成稳定、可执行的方案。

图 2–2　计算思维四要素循环图

这种思维方式并不局限于计算机系统。在学习与生活中，许多任务同样可以用类似的方法组织起来。例如，制订考试周复习计划时，可以先把任务拆分成课程清单、时间安排、重点难点、练习与复盘等模块；再观察不同课程复习过程中的共性结构；再抓住掌握程度、题型结构、考试权重等核心要素；最后把每天要做的事情写成清晰的执行步骤，并根据完成情况动态调整。你会发现这个过程并不要求你会写程序，但它已经具备了可执行、可复现、可优化的典型特征。

计算思维之所以成为信息化时代的重要能力，还因为它天然强调一种“可验证性”。在计算思维框架下，一个方案是否可靠，往往可以通过逻辑推导、测试用例、边界条件分析等方式进行检验。这种特性使得计算思维特别适合那些对可靠性要求极高的场景，例如金融结算、考试评分、库存管理、航班调度等系统。在这些系统中，“差不多”是不允许的，必须做到“每一步都可追踪、每一次都可复现”。

当然，计算思维并不是万能的。它的强大来自规则的清晰，但它的边界也同样来自规则。一旦规则难以穷举，流程就很难继续推进。为了更准确地把握计算思维的适用范围，表 2–1 对其典型优势与局限做了一个归纳。

表 2–1　计算思维的能力画像：适用条件、优势与局限

对比维度	计算思维的典型特征	典型例子（校园 / 社会）	对后续学习的启示
适用问题类型	结构清晰、边界明确、规则可描述、目标可验证	排课系统、成绩计算、库存管理、银行转账	规则与流程仍是智能系统的“骨架”
解决方式	规则驱动：先写清楚规则，再让系统执行	“如果时间冲突则不可选”“如果容量不足则排除”	先学会把问题表达清楚，再谈智能
输出特征	确定性输出：相同输入必得相同结果	课表生成、考试分数计算	可靠性强，便于审计与追责

续表

对比维度	计算思维的典型特征	典型例子（校园 / 社会）	对后续学习的启示
核心优势	可解释、可追踪、可复现、易测试	系统出错时可定位到某一步规则	这也是很多关键系统必须采用的方式
典型局限	依赖规则完备性；规则难穷举时效果下降	图像识别、语音识别、语言理解难靠规则覆盖	需要“学习驱动”

当规则可以被写清楚时，计算思维足以支撑一个稳定可靠的系统。但在许多现实任务中，真正的困难并不在计算量，而在于我们根本写不出完整规则。面对这类问题，人类又是如何获得能力的？这就需要引入另一种思维方式——以“学习”为核心的人工智能思维。

2.1.2 计算思维的边界

在 2.1.1 小节中，我们看到计算思维之所以强大，是因为它能够把问题转化为结构化表达，再进一步转化为可执行步骤。只要输入、输出与约束足够清晰，计算机就能稳定、快速、可复现地完成任务。排课系统、成绩计算、库存管理等应用之所以可靠，本质上都是因为规则能够被明确写出，并且在较长时间内保持有效。

但当我们把目光从结构化系统转向真实世界时，会很快发现一个现象：很多我们认为“很智能”的能力，并不是通过写规则获得的。人类能够识别人脸、理解语言、感知情绪、判断风险、进行创作，这些能力往往难以被写成完整的规则流程。换句话说，计算思维并非失效，而是遇到了它的天然边界。

为了更直观地理解这种边界，不妨先从一个校园里常见的任务对比开始。

在成绩计算中，规则通常是清晰明确的。例如“平时成绩占 30%，期末考试占 70%”“缺勤超过三次不得参加考试”等规定。规则一旦确定，系统就可以自动执行，并给出确定结果。即使结果出现争议，也可以回溯到输入数据与规则条款，解释每一步的来源。这是一类典型的“可写规则”问题。

但当任务变成作文评分时，情况立刻变得复杂起来。我们当然可以写出一些规则：字数是否达标、是否偏题、是否有明显错别字、结构是否完整。但这些规则只能覆盖一小部分。真正决定一篇文章质量的因素，往往包括立意深度、论证力度、语言表达、逻辑严密性、创新性等，这些指标具有明显的模糊性与主观性，很难用有限规则穷举描述。即使勉强写出规则，也会发现不同教师、不同语境、不同写作目的下的评价标准并不完全一致。于是，“写规则”这条路很快就走到了尽头。

这一对比揭示了一个关键事实：计算思维虽擅长处理规则可穷举的问题，但在面对规则难穷举的问题时会显得力不从心。现实世界中大量困难的任务，并不是“算不出来”，而是“写不出来”。从计算思维的角度看，“写不出规则”并不是因为人类懒惰，而是因为现实世界本身

具有多层复杂性。典型来说，这些复杂性主要来自以下四个方面。

1）输入往往是非结构化的

在排课系统中，输入是课程编号、教师名单、时间段、教室容量等结构化数据；但在很多任务中，输入可能是一段语音、一张图片、一段视频、一篇文章。它们不是天然的“表格数据”，而是包含大量连续信号与高维特征的复杂对象。面对这种输入，想用规则直接描述关键点在哪里，往往非常困难。

2）现实存在噪声与不确定性

即使是同一个对象，在不同光照、角度、遮挡条件下的表现也不完全相同。例如同一张脸，在戴口罩、戴眼镜、侧脸、低光照、表情变化时的视觉特征差异巨大。规则一旦写得过于具体，就会失去泛化能力；规则如果写得过于宽泛，又会导致误判。

3）目标往往带有模糊性与多样性

很多任务的目标并不是“对 / 错”，而是“更好 / 更差”。例如“生成一张更符合审美的海报”“写一段更自然的对话”“推荐更合适的课程”。这种目标很难被写成统一的硬性规则，因为不同人对“好”的理解本身就存在差异。

4）环境与需求会持续变化

排课规则可能多年不变，但在推荐系统、舆情分析、网络安全、金融风险等任务中，数据分布和行为模式会快速变化。规则一旦固定，就会很快过时。即使不断更新规则，也会带来维护成本爆炸。

因此，在这些场景中，计算思维遇到的问题并不是“计算能力不足”，而是“表达能力不足”——我们难以把问题完整表达成可执行规则。当规则难以表达时，一个常见的工程直觉是继续增加规则，试图把“漏网之鱼”补上。但这种做法很快会遇到现实困境：规则越多，系统越复杂；系统越复杂，维护越困难。人工智能发展早期的专家系统就是典型例子。专家系统的思路是把专家经验总结成大量“如果……那么……”的规则，让系统通过规则链进行推理。在特定领域，它确实能发挥作用，例如设备故障排查、部分医疗诊断辅助等。但随着规则数量增长，系统会出现明显的知识工程瓶颈：一方面，专家知识难以完全显性化，很多经验是隐性的；另一方面，新增规则可能与旧规则冲突，导致不可预期的连锁反应，系统反而更难稳定运行。工程上通常将这种现象称为“规则爆炸”。

为了更直观地呈现计算思维的边界，可以将不同任务放在一个简单的二维图中进行理解：当输入越非结构化、规则越难描述时，计算思维的优势就越难发挥。如图 2–3 所示，排

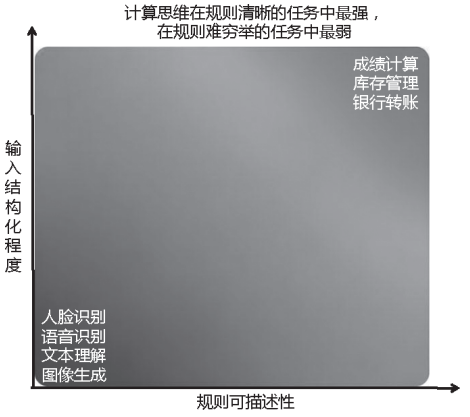


图 2–3　计算思维的边界

课、成绩计算等任务位于“规则容易描述、输入结构化”的区域；而人脸识别、语音识别、语言理解等任务则位于“规则难以穷举、输入高度非结构化”的区域。两类任务的差异，并不是难度高低的简单对比，而是问题性质的根本不同。

需要强调的是，计算思维的边界并不意味着计算思维被淘汰。恰恰相反，在现代智能系统中，计算思维仍然发挥着不可替代的作用：它负责系统流程、权限控制、安全策略、日志记录与可追溯性；在高风险场景中，它负责规则兜底与约束；在系统工程中，它负责模块化拆解与接口设计。也就是说，计算思维依然是智能系统的“骨架”，只是它无法独立解决所有问题。

为了帮助读者更清晰地区分两类问题的典型特征，表 2–2 给出了一个对比总结。通过这张表可以看到：当问题从左侧逐渐走向右侧时，单纯依赖规则往往难以取得理想效果，能力的来源需要从“人工编写规则”转向“从数据中学习规律”。

表 2–2　计算思维的适用边界

对比维度	可写规则的问题	难写规则的问题	典型例子
输入类型	结构化、符号化、可穷举	非结构化、高维、连续信号	课表数据 vs 图像 / 语音 / 文本
规则表达	规则清晰、可穷举	规则模糊、难穷举	税费计算 vs 情绪识别
输出特征	确定性强，可验证	概率性强，需容错	成绩计算 vs 推荐系统
环境变化	相对稳定，规则长期有效	快速变化，规则易过时	教务流程 vs 舆情分析
工程维护	维护成本可控	规则越多越难维护	小型业务系统 vs 大规模智能系统
常见风险	主要是逻辑漏洞与边界条件遗漏	主要是偏差、误判与不可解释	程序漏洞 vs 模型偏差

当问题落在“规则可描述、边界清晰”的区域时，计算思维足以支撑一个稳定可靠的系统。但当问题进入“非结构化输入、目标模糊、动态变化”的区域时，继续增加规则往往只会带来维护成本的急剧上升，却难以获得真正可用的效果。此时，能力的来源就必须发生变化：与其试图穷举规则，不如让系统从大量数据中学习规律。人类识别、理解、判断与创作等能力的形成，本质上也正是一种从经验中学习的过程。人工智能思维的关键，就是把这种能力获得方式转化为可被计算机实现的机制。

2.1.3　人工智能思维

在计算思维中，系统能力主要来自人类写出的规则。我们把任务拆解成明确步骤，把条件判断写清楚，把流程设计完整，再让计算机稳定执行。只要规则能够被描述，系统就可以做到快速、可靠、可追溯。

但当面对图像、语音、文本等复杂信息处理场景时，规则往往难以穷举。此时，人工智

能思维逐渐形成：它不再依赖人工把规则写完，而是让系统从数据中学习规律。换句话说，人工智能思维并不是更复杂的编程，而是一种新的问题解决方式——通过学习获得能力。

从人工智能思维的角度来看，一个系统要获得某种能力，通常需要回答三个基本问题：经验从哪里来、能力存放在哪里、能力如何形成。对应到人工智能系统中，这三个问题分别落在三个关键词上：数据、模型与学习。

◎ 数据是经验来源。它可以是一批带有标签的样本（例如“猫 / 狗”图片），也可以是大量未标注的真实记录（例如用户浏览行为、语料文本、设备日志）。数据是否真实、是否覆盖实际场景，往往决定了模型泛化能力的上限。

◎ 模型是能力载体。它可以理解为一种可训练的函数结构：给定输入，输出预测或生成结果。模型内部包含大量参数，这些参数决定了模型对输入信息的处理方式。

◎ 学习是能力获得机制。它的本质是让模型在大量样本上不断调整参数，使输出越来越符合目标。系统能力并不是一次性写出来的，而是在学习过程中逐步形成的。

这三个要素构成了人工智能思维的核心框架：数据提供经验，模型承载能力，学习完成能力的形成。当我们用这种方式重新理解人工智能时，就会发现它与传统软件的根本区别：传统软件主要依赖规则覆盖，人工智能系统主要依赖数据学习。

伦理思辨

在人工智能思维中，系统能力往往来自数据学习而不是人工编写规则。这意味着系统的判断并不是绝对确定的，而是一种基于概率的预测。例如，人脸识别、情感分析、风险评估等系统都会给出“更可能”的判断，而不是完全确定的结论。

这种能力在很多场景中能够显著提升效率，但同时也可能带来新的问题。例如，如果训练数据本身存在偏差，系统可能对某些群体产生误判；如果人们过度依赖系统判断，也可能忽视人工复核的重要性。

思考与讨论：

1. 当人工智能系统参与评分、推荐或风险评估时，人类是否应完全依赖系统判断？
2. 如果算法判断出现偏差，对某些群体产生不公平结果，应如何进行纠正？
3. 在重要决策中，应如何设计人工复核机制，来避免技术误判带来的风险？

通过这些问题可以看到，人工智能系统不仅需要技术能力，也需要合理的制度设计与伦理约束。

这种差异也会带来一个非常直观的现象：人工智能系统的输出往往不是“非黑即白”的确定性结果，而更像是一种带有置信度的判断。例如，“这张图片更像猫”“这段文本更可能表达消极情绪”“该用户更可能喜欢某门课程”。这种输出形式并不是人工智能系统的缺点，而是对

现实不确定性的合理表达。更重要的是，它为系统设计提供了空间：当置信度较高时可以自动处理，当置信度较低时可以交给人工复核，从而形成更可靠的人机协同机制。

这里需要明确一点，人工智能思维并没有取代计算思维。两者面向的是不同类型的问题：计算思维擅长构建稳定可靠的系统骨架，人工智能思维擅长在复杂现实任务中获得近似人类的感知与理解能力。现代智能系统往往同时依赖两者，在流程与约束层面采用计算思维保证可靠，在感知与理解层面引入人工智能思维获得能力。两种思维方式相互补充，才构成了今天我们所熟悉的智能应用形态。

智领未来

为什么人工智能的输出是概率性的？——置信度与容错机制

很多同学会发现，人工智能系统的输出往往不是“对 / 错”的确定答案，而是“更可能”的判断。例如相册识别会显示“这张图片更像猫”，语音识别会出现同音误差，推荐系统也只是给出“更可能喜欢”的排序。

这种现象并不是人工智能“不够聪明”，而是因为现实世界本身充满噪声与不确定性：光照变化、角度变化、遮挡、口音、表达歧义、数据分布变化……都可能让同一个输入在不同环境下呈现出不同特征。面对这种复杂性，学习型模型往往只能给出概率意义上的判断。

因此，工程上通常不会把模型输出当作“最终裁决”，而是会引入一套容错机制：用置信度衡量模型是否“足够确定”；设置阈值，决定自动处理还是交由人工复核；对高风险任务加入多模型校验与规则兜底；通过日志记录与反馈回流等方式持续改进模型。

换句话说，人工智能系统的关键不是“永远不犯错”，而是在允许一定误差的前提下，让系统整体更高效，同时在关键环节做到可控、可追溯、可负责。

2.2 数据驱动与模型驱动的核心范式

2.2.1 数据驱动与模型驱动

在 2.1 节中，我们已经知道：当规则难以穷举时，人工智能系统往往通过“学习”获得能力。但是一个更具体的问题自然会出现：学习到底靠什么？为什么有些系统看起来主要靠“喂

数据”，而有些系统看起来主要靠“设计规则和结构”？

为了把这个问题说清楚，不妨先看一个很生活化的例子：垃圾邮件识别。

如果让我们用计算思维来写规则，通常会写出很多“如果……那么……”的条件。例如：如果邮件标题包含“中奖”“免费”“限时”，就判定为垃圾邮件；如果发件人地址异常，就判定为垃圾邮件；如果正文包含大量链接，就判定为垃圾邮件。这样的规则确实能解决一部分问题，但很快会遇到困难：垃圾邮件的写法会变化，关键词会替换，甚至会故意伪装成正常邮件。规则写得越多，维护成本越高，而且总会漏掉新的变种。另一种做法是：不去穷举规则，而是收集大量真实邮件样本（包括垃圾邮件与正常邮件），让系统从数据中学习什么样的邮件更像垃圾邮件。在这种做法中，人类不再把规则写出来，而是把判断依据“藏”在数据中，让模型自己学。这就是典型的数据驱动方式。

但如果我们换一个任务（例如排课、路径规划、资源调度），情况又会不同。排课系统中最核心的不是“学习经验”，而是满足约束：不能冲突、容量要够、实验课要有设备、必修课优先级要高。即使这里有大量历史数据，也无法直接替代这些硬约束。系统更依赖把约束表达清楚，再在约束下寻找可行解。这类做法更接近模型驱动：它强调结构、约束与可解释的推理过程。

这两类任务的差异非常典型：前者更像是从大量样本中学会识别，后者更像是在明确约束下做出可行决策。表 2–3 对这两种常见路径做了一个直观对照。

表 2–3　数据驱动与模型驱动的对比

关键问题	数据驱动（更像“学出来”）	模型驱动（更像“设计出来”）
主要依赖	大量数据与样本	结构、约束、规则与推理
典型任务	识别、理解、推荐	调度、规划、约束满足
直观例子	垃圾邮件识别、人脸识别	排课、路径规划、资源分配
常见难点	数据偏差、泛化能力不足	约束冲突多、多目标求解难度大、最优解计算成本高

通过这两个例子可以看到，人工智能系统并不存在唯一的实现方式。面对不同类型的问题，往往会走两条不同的路径：数据驱动——从经验中学会识别；模型驱动——在约束下做决策。

在真实系统中，这两条路径也常常结合出现（见图 2–4）。例如，智慧校园的门禁识别可能依赖数据驱动，但门禁权限、审批流程、异常报警阈值等部分仍然需要模型驱动或规则约束来保证安全与可控。

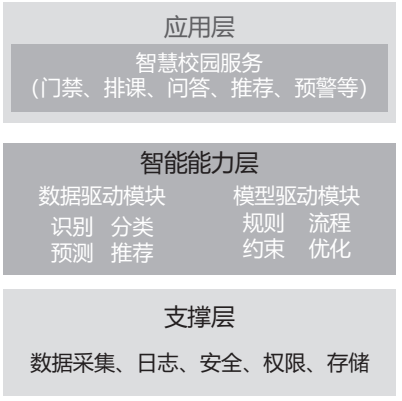


图 2–4　人工智能系统的两类实现路径

2.2.2 核心思想与适用范围

人工智能系统并不是万能钥匙。同样是“智能化”，不同任务对方法的要求差异很大。有些任务可以通过大量数据学习得到能力，有些任务则必须依赖明确的结构、约束与推理过程。对于学习者来说，最重要的不是记住某种算法名称，而是能判断：面对一个具体问题时，应优先依赖数据，还是优先依赖结构与约束？这种判断能力，决定了我们能否真正理解人工智能系统的工作方式。

1. 数据驱动：从大量样本中学习规律

数据驱动的核心思想是：当规则难以穷举时，不再强行写规则，而是收集大量样本，让系统从数据中学习规律。它更像人类的经验学习，见得越多，越容易形成稳定判断。

在数据驱动系统中，数据不仅是输入，更是能力的来源。系统是否能学到有效规律，很大程度取决于数据是否真实、是否覆盖使用场景、是否包含足够多的典型情况与边界情况。表 2–4 所示为将数据驱动的关键要素用通俗方式描述。

表 2–4　数据驱动范式的关键要素：系统如何“学出能力”

关键要素	通俗理解	在系统中表现	举　例
数据样本	经验材料	图片、语音、文本、行为记录	邮件正文、签到记录、消费记录
标签或反馈	参考答案	分类标签、评分、点击 / 购买	正常 / 垃圾、好 / 差、点击 / 未点击
学习过程	反复训练	调整模型参数	多轮训练直到更准确
泛化能力	举一反三	新样本上的效果	新型垃圾邮件也能识别

从表 2–4 可以看到，数据驱动的本质是“用样本替代规则”。在这种范式下，人类不必提供判断依据，但必须让系统看到足够多、足够真实的样本。因此，数据驱动最典型、最成功的应用集中在“感知与理解”类任务中。这些任务的共同特点是：输入复杂、规则难写，但可以通过大量样本让系统学到稳定模式。表 2–5 列出了数据驱动最常见的应用类型。

表 2–5　数据驱动范式最常见的应用类型

任务类型	典型输入	系统输出	常见应用
图像识别	图片 / 视频帧	类别、位置、特征	人脸识别、车牌识别、医学影像分析
语音识别	语音信号	文本 / 命令	语音转文字、语音助手
文本理解	句子 / 段落	分类、抽取、判断	垃圾邮件、情感分析、舆情分类
推荐预测	行为记录	概率、排序	课程推荐、内容推荐、商品推荐

数据驱动带来的优势很明显，它能够处理人类难以穷举规则的复杂任务。但它的短板也同样明显：很多问题并不在于“模型不够强”，而在于“数据不够好”。在实际应用中，数据驱动系统面临的困难通常集中在以下四类。

（1）数据缺乏代表性：训练数据与真实使用环境差异过大，导致模型在实验中效果很好，上线后性能明显下降。

（2）数据分布不均衡：某些类别样本过多，而某些类别样本过少，导致模型“偏科”，对少数类别的识别能力较弱。

（3）标签不可靠：标注错误、标注标准不一致，或者标签本身具有主观性，都会让模型学到错误规律。

（4）数据随时间变化：场景与行为模式发生变化，模型逐渐失效，需要持续更新与再训练。

可以看到，数据驱动并不是“堆数据就行”，它更像一种工程能力——需要持续保证数据的质量与代表性，才能让系统稳定工作。

科技中国

从“写规则”到“训模型”：国家为什么要布局算力与大模型？

当规则难以穷举时，系统往往需要从数据中学习规律。进入大模型时代，这种变化进一步加速——许多能力不再靠人工写规则，而依赖海量数据与大规模训练获得。换句话说，智能系统的“学习驱动”越来越依赖两个底座：算力与数据。

这也是为什么我国在人工智能发展中，特别强调国家级基础设施与顶层设计。一方面，《新一代人工智能发展规划》等政策将人工智能上升为国家战略，推动关键核心技术攻关；另一方面，“东数西算”等重大工程通过优化算力布局，为大规模训练与产业应用提供底层支撑。对大模型而言，算力不只是“跑得快”，更直接决定了能否训练更大规模的模型、能否支撑更多行业落地。

更重要的是，国家战略的意义不仅在于“把模型训出来”，还在于“把系统用起来”。在航天遥感、能源电网、智能制造、乡村振兴等重大任务中，人工智能往往以融合形态落地：感知与预测依赖数据驱动，流程控制与安全审计依赖规则与约束。由此可以看到，人工智能的竞争不仅是算法竞争，更是“算力—数据—平台—应用生态”的系统竞争。

2. 模型驱动：用结构与约束表达问题

与数据驱动不同，模型驱动更强调在解决问题之前，先把问题的结构表达清楚。它关注的不是“从样本中学会识别”，而是“在明确约束下做出决策”。

模型驱动的典型任务往往具有三个特点：目标清晰，约束明确，过程必须可解释、可追溯、可控制。

排课、路径规划、资源调度都属于这一类任务。系统要解决的不是“识别出什么”，而是“在规则下找到一个可行方案，甚至找到最优方案”。

模型驱动的核心并不在于“从数据中学会识别”，而在于“把问题设计清楚”。它强调先明确目标与约束，再在规则空间中推理或求解得到结果。从表 2–6 可以看到，模型驱动的最大

优势在于可靠、可控、可解释，特别适合那些对安全性和责任追溯要求很高的场景。

表 2–6　模型驱动范式的关键要素：系统如何“在约束下求解”

关键要素	通俗理解	在系统中体现为	举　例
结构化表达	把问题写成“可推理形式”	变量、关系、约束	课程 – 教师 – 时间 – 教室
约束优先	先满足规则，再追求更好	硬约束、冲突检测	不冲突、容量必须够
推理或求解	在规则空间中找答案	搜索、推理、优化	找到可行课表
可解释与可控	结果来源清晰，可调整	推理链、规则兜底	为什么这样安排可说明

表 2–7 展示了模型驱动更常见的任务类型。

表 2–7　模型驱动范式更常见的任务类型

任务类型	典型特点	常见输出	典型例子
排课 / 调度	约束多、必须满足规则	可行方案 / 最优方案	排课、排班、考场安排
路径规划	目标清晰、约束明确	路径 / 策略	导航、机器人路径
资源分配	资源有限、优先级不同	分配方案	教室分配、设备调度
风险控制（规则层）	责任明确、可追溯	是否触发规则	金融风控、权限控制

模型驱动也有其局限性：当输入高度复杂、规律难以显式表达时，规则往往难以覆盖。例如在图像识别、语音识别、自然语言理解等任务中，试图靠规则穷举会迅速陷入“规则爆炸”。这也是为什么现代智能系统通常不会用单一范式解决全部问题。

为了在实际任务中更快做出初步判断，可以从以下五个角度进行区分。

（1）看输入是什么：如果输入主要是图像、语音、文本等复杂信息，往往更偏向数据驱动；如果输入主要是表格数据、明确变量与约束，往往更偏向模型驱动。

（2）看规则能否写清楚：如果规则难以穷举、难以覆盖各种变化，通常更适合数据驱动；如果规则可以明确表达，通常更适合模型驱动。

（3）看任务更像什么：识别、理解、预测类任务更适合数据驱动；规划、调度、约束满足类任务更适合模型驱动。

（4）看是否必须可解释：如果结果必须可追溯、可审计、可解释，通常需要模型驱动或规则约束；如果允许一定误差并可通过置信度控制风险，则更适合数据驱动。

（5）看是否可容错：在很多感知任务中允许一定误差（例如推荐、识别），数据驱动更常见；在硬约束任务中误差往往不可接受（例如排课冲突、权限错误），模型驱动更常见。

从这些判断角度可以得到一个实用结论：当任务更像感知与识别时，优先考虑数据驱动；当任务更像约束决策时，优先考虑模型驱动。现实系统通常位于两者之间，需要组合使用，这种融合形态，才是现代人工智能系统最常见、也最可落地的形态。

智领未来

2.2.3 融合与协同

在 2.2.1 和 2.2.2 小节中，我们已经看到：数据驱动擅长处理识别、理解、预测等“从经验中学会判断”的任务；模型驱动擅长处理规划、调度、约束满足等“在规则下做出决策”的任务。但在现实世界中，真正可落地、可长期运行的智能系统，往往很少只依赖其中一种方法。原因很简单：现实系统不仅要“聪明”，还要“可靠”。

仅靠数据驱动模型，系统可能会出现不可预测的误判；仅靠模型驱动规则，系统又可能缺乏对复杂信息的理解能力。于是，大多数成熟的人工智能应用都会采用一种更常见的形态：在同一个系统中组合两条路径，让它们各自承担最擅长的部分。

例如，在智慧校园场景中，既包含大量需要感知与识别的任务（例如门禁识别、课堂签到、学习行为分析），也包含大量必须严格遵守规则的任务（例如权限审批、教务流程、宿舍管理、安全审计）。如果只使用一种范式，系统要么不够智能，要么不够可靠。

表 2–8 展示了智慧校园系统中常见的功能模块，并说明它们分别更依赖哪一条路径，以及两者如何协同。

表 2–8　智慧校园系统的融合

功能模块	更依赖数据驱动的部分	更依赖模型驱动的部分	为什么需要融合
门禁识别	人脸识别模型、活体检测	权限规则、黑名单、阈值策略	识别要智能，放行要安全
课堂签到	人脸 / 姿态识别、异常检测	课程表、学生名单、签到规则	识别负责判断“是谁”，规则负责判断“算不算到”
学习预警	行为数据建模、风险预测	预警规则、人工复核流程	预测有误差，流程要可控
智能推荐（课程 / 资源）	推荐模型、兴趣建模	推荐约束（必修优先、学分要求）	推荐要个性化，也要符合培养方案
智能客服	意图识别、文本理解	业务流程、知识库结构、权限控制	能回答问题，也要遵守业务边界
校园安全巡检	视频异常识别、目标检测	告警等级、处置流程、日志审计	识别负责发现，规则负责处置

从表 2–8 可以看到，一个智能系统通常可以被理解为“上下两层”的组合。

◎ 上层负责智能能力：识别、理解、预测、推荐等任务更多依赖数据驱动。

◎ 下层负责可靠运行：权限、流程、约束、日志、审计等任务更多依赖模型驱动与规则体系。

这也解释了一个常见现象：许多系统的“智能模块”看起来很像人工智能，而系统的“主体框架”仍然是传统软件工程与计算思维。人工智能并不是取代系统，而是嵌入系统，成为系统的一部分。图 2–5 展示了融合式智能系统的典型结构：智能模块提供识别与预测能力，系统框架负责流程控制、权限安全与运行保障。

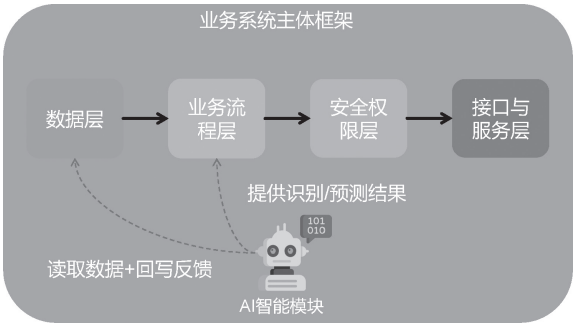


图 2-5　融合式智能系统的结构示意图

融合范式不仅能让系统更可靠，也能让系统更容易落地。因为很多关键场景并不允许完全自动化。例如门禁系统中，模型识别再准确，也需要阈值与权限控制；学习预警中，模型预测再灵敏，也需要人工复核与合理的流程；智能客服中，模型回答再流畅，也需要业务边界与权限限制。通过规则与流程对智能模块进行约束，系统才能在复杂现实环境中长期稳定运行。

因此，在理解人工智能系统时，一个非常重要的观念是：人工智能并不是孤立的模型，而是一种可嵌入系统的能力模块。数据驱动与模型驱动并不是竞争关系，而是互补关系。前者让系统具备感知与学习能力，后者让系统具备结构化的可靠运行能力。两者结合，才是现代人工智能系统最常见、也最可落地的形态。

“人工智能 + 专业”的融合路径

2.3.1　核心逻辑：专业问题的人工智能改造

当我们理解了数据驱动与模型驱动在智能系统中的分工后，一个现实问题随之出现：不同专业的核心任务、数据来源、约束条件及评价标准各不相同。所谓“人工智能 + 专业”，绝不是简单地把一个模型套用到所有领域，而是要在具体场景中判断：哪些环节适合用数据驱动获得感知与预测能力，哪些环节必须依赖模型驱动（规则与约束）来保证可靠性，以及如何把两者组织成可落地的系统方案。

“人工智能 + 专业”最常见的误解，是把它理解为“给专业加一个模型”。更准确的理解是：把专业任务拆分成多个环节，在不同环节选择更合适的方法。数据驱动更擅长从样本中学会识别与预测，而模型驱动更擅长在规则与约束下做出可控决策。多数行业方案并不是二选一，而是把两者融合在同一个系统中。

从工程视角来看，许多专业任务都可以抽象为“输入—处理—决策—执行—反馈”的链条。人工智能通常嵌入其中的某些关键环节，而不是替代整个专业流程。表 2–9 给出了一个通用拆解框架，用于帮助读者快速判断：哪里适合引入学习能力，哪里必须保留规则约束，哪里需要人机协同。

表 2–9 “人工智能 + 专业”的通用拆解框架

任务环节	环节功能	数据驱动能力	规则 / 约束要点	典型输出
信息采集与感知	多源数据获取与初步感知	图像 / 语音 / 文本识别、异常检测	采集规范、权限控制、数据合规	结构化数据、识别结果
信息理解与分析	语义提取与特征表达	语义理解、特征提取、分类聚类	术语体系、标准定义、可追溯机制	关键字段、分析结果
预测与评估	状态预测与风险评估	预测模型、评分模型、排序推荐	阈值策略、评价指标、复核规则	风险等级、推荐列表
决策与规划	方案生成与策略选择	数据辅助决策	约束满足、流程规则、优化求解	可行方案、最优策略
执行与控制	任务执行与过程控制	自适应控制	安全策略、审批机制、责任审计	执行记录、审计日志
反馈与迭代	运行反馈与持续改进	数据回流、再训练更新	版本管理、灰度发布、审计机制	模型更新、规则更新

从表 2–9 可以看到，“人工智能 + 专业”的关键不是让系统“全自动”，而是让系统在合适的环节变得更智能，同时在关键环节保持可控。例如在医学场景中，影像识别可以由学习模型承担，但诊疗流程必须遵循严格规范；在金融场景中，风险评分可以由模型提供，但最终决策必须符合规定与审计要求；在工程制造场景中，视觉质检可以由模型完成，但生产排程仍需要约束与优化。

图 2–6 展示了“人工智能 + 专业”系统的典型融合结构：数据驱动模块负责识别与预测，模型驱动模块负责流程与约束，两者共同支撑专业应用。

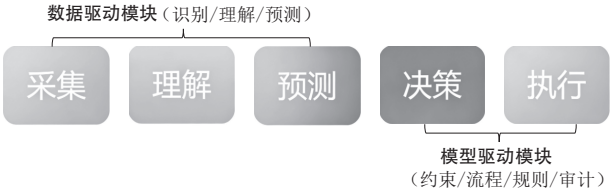


图 2–6 “人工智能 + 专业”融合框架图

2.3.2 学科门类的融合方向

有了表 2–9 所示的通用拆解框架，就可以进一步思考：在不同专业中，哪些任务环节更可能由数据驱动模块提供能力，哪些环节必须由规则与约束保证可靠。由于各学科的研究对象不

同、数据形态不同、约束条件不同，“人工智能 + 专业”的落点也会呈现明显差异。

表 2–10 按学科门类分类，总结了“人工智能 + 专业”常见的融合方向与典型任务。我们可结合自身专业特点，快速定位本专业中人工智能最常应用于哪些任务环节，系统最需要哪些类型的数据支持，以及哪些约束机制是不可缺少的。

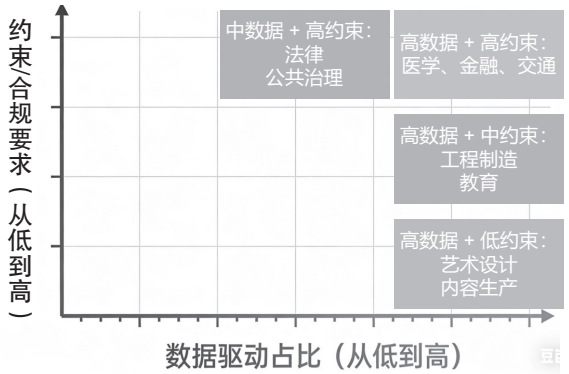
表 2–10 “人工智能 + 专业”常见的融合方向与典型任务

学科门类	典型任务场景	数据驱动切入点	规则 / 约束要点	常见成果形式
计算机与软件类	智能应用开发、系统优化	识别 / 理解模型训练、检索与推荐	架构设计、安全权限、性能约束	智能系统、平台工具
自动化与控制类	设备控制、预测维护	状态识别、故障诊断、寿命预测	控制约束、安全策略、实时性	控制系统、工业方案
机械与材料类	质量检测、材料设计	视觉质检、缺陷识别、性能预测	工艺流程、检测标准、可靠性	质检系统、设计辅助
电气与能源类	能耗优化、电网监测	负荷预测、异常检测、风险预警	调度规则、稳定性约束、合规要求	监测平台、优化策略
土木与建筑类	工程监测、施工管理	结构监测、图像巡检、风险预测	工程规范、安全标准、责任审计	监测系统、预警平台
交通与物流类	路径规划、运力调度	流量预测、需求预测、异常识别	约束优化、时空规则、安全约束	调度系统、智能导航
医学与生命科学类	影像辅助、健康管理	影像识别、风险预测、文本抽取	诊疗规范、隐私合规、人工复核	辅助诊断、健康平台
生物与农业类	育种与病虫害识别	病虫害识别、表型分析、产量预测	采样标准、农业规范、可解释	识别系统、决策支持
数学与统计类	建模分析、推断评估	数据建模、预测分析、模式发现	指标定义、模型验证、可解释	分析报告、评估模型
物理与化学类	实验分析、模拟加速	光谱 / 显微图像识别、性质预测	实验条件约束、机理一致性	分析工具、预测模型
经济与金融类	风险管理、市场分析	风险评分、欺诈检测、趋势预测	合规规则、审计追溯、阈值策略	风控系统、决策支持
管理与公共事务类	运营分析、公共治理	舆情分析、行为预测、资源评估	政策规则、流程规范、责任边界	分析平台、辅助决策
法学类	法律检索、类案分析	文本检索、要素抽取、文书辅助	法律条文约束、责任主体、可追溯	检索系统、辅助工具
新闻传播类	内容生产、舆情监测	内容分类、情感分析、热点发现	事实核查流程、传播规范	监测平台、辅助写作
外语与语言类	翻译与写作辅助	机器翻译、写作润色、语音评测	语域规范、评价标准、版权边界	学习工具、教学系统
教育学类	学习分析、教学支持	学习行为建模、个性化推荐	教学目标、评价体系、隐私保护	教学平台、学习助手

续表

学科门类	典型任务场景	数据驱动切入点	规则 / 约束要点	常见成果形式
心理学类	行为分析、风险预警	行为模式识别、量表预测	伦理规范、知情同意、人工复核	评估工具、辅助系统
艺术与设计类	设计生成、创意辅助	图像生成、风格迁移、内容推荐	版权边界、审美目标、人工把关	创作工具、设计方案
体育与健康类	训练评估、动作分析	姿态识别、训练评估、风险预测	训练规范、安全阈值、隐私保护	评估系统、训练建议

从表 2–10 可以看到，尽管不同学科的研究对象差异很大，但“人工智能 + 专业”的融合模式具有共同规律：数据驱动更常用于感知、理解、预测等环节，模型驱动更常用于流程、约束、责任与合规等环节。对于本科生而言，理解这种分工比记住某种模型名称更重要，因为它决定了人工智能在专业场景中能做什么、不能做什么，以及应该如何落地，如图 2–7 所示。



智领未来

图 2–7　学科门类的 AI 落点分布图

2.3.3　规范、边界与责任

微课视频

“人工智能 + 专业”的价值在于提升效率、提高质量、增强能力，但它并不是一个可以随意套用的技术标签。越是触及专业核心环节，人工智能系统越需要面对一个现实问题：系统输出可能带来真实后果。这些后果不仅包括准确率低，还包括隐私泄露、偏差歧视、责任不清、误用滥用等风险。

因此，“人工智能 + 专业”能否真正落地，往往不取决于模型是否足够强，而取决于系统是否具备可控、可追溯、可审计的机制。尤其在医学、金融、法律、公共治理等高风险领域，智能能力必须被嵌入规则、流程与责任体系之中。

从落地视角来看，“人工智能 + 专业”的边界主要体现在以下五个方面。

1）数据合规是底线

许多专业场景的核心数据天然敏感，例如医疗数据、学生行为日志、个人画像、企业经

营数据等。系统即使能够训练出更高准确率，也不能以牺牲隐私与合规为代价。实践中通常需要遵循“最小必要原则”：该采集的数据才采集、该共享的数据才共享，并对数据进行权限分级、脱敏处理与安全审计。

2）模型输出必须可控

数据驱动模型的输出具有概率性，容易出现误判，因此很多场景并不允许“凭一次判断就直接执行”，尤其是门禁放行、风险拦截、诊疗提示、舆情处置等关键任务。工程上常见的做法是设置置信度阈值、引入多模型校验，并在关键决策点保留人工复核机制，从而避免小概率错误造成严重后果。

3）责任边界必须清晰

当 AI 进入专业流程后，最容易出现的问题之一是“系统建议出了错，谁负责”。因此，成熟系统通常不会把 AI 输出当作最终裁决，而是把它定位为“辅助判断”。同时，需要明确哪些环节由系统自动执行，哪些环节必须由专业人员签字确认，哪些环节需要留下可追溯的日志记录。

4）公平与偏差需要被看见

数据驱动系统学习的是历史数据，而历史数据往往并不完美。它可能隐含对某些群体的偏差，例如不同地区、不同性别、不同年龄、不同学校背景的样本分布差异。如果缺乏检查与纠正，系统可能在不知不觉中放大不公平。因此，在涉及筛选、评估、推荐、风险评分等任务时，必须关注数据分布，评估模型对不同群体的表现差异，并对偏差进行修正。

5）安全与误用不可忽视

智能系统不仅可能出错，还可能被“利用”。例如识别系统可能遭遇对抗攻击，文本系统可能被注入诱导指令，生成系统可能被用于伪造内容或传播虚假信息。专业场景越敏感，对安全与误用的防护要求越高。常见措施包括输入过滤、安全测试、访问权限控制、输出审查与内容溯源等。

归纳起来可以看到，“人工智能 + 专业”的核心不是“把模型接进来”，而是“把能力放进去，并把边界守住”。只有在合规、安全、责任与流程机制之上，人工智能能力才能稳定地融入专业核心任务，并长期产生价值。

榜样力量

李飞飞——让数据推动人工智能发展

李飞飞是国际知名的人工智能科学家，也是计算机视觉领域的重要研究者。她在斯坦福大学任教，并长期致力于推动人工智能基础研究的发展。为了推动机器能够“看懂世界”，李飞飞主持建立了大规模图像数据集 ImageNet，为计算机视觉研究提供了关键的数据基础。

2012 年，研究团队利用 ImageNet 数据训练深度学习模型，在图像识别竞赛中取得突破性成果，这一成果被认为是深度学习发展的重要里程碑。李飞飞的研究表明，高质量的数据资源对于机器学习模型的能力提升具有重要意义，也推动人工智能从依赖规则的系统向数据驱动的学习模式转变。

她的科研经历不仅体现了长期坚持基础研究、推动开放合作的科学精神，也为人工智能技术的发展作出了重要贡献。


实训项目2-1


实训项目2-2


实训项目2-3


学以致用


综合评价表