人工智能应用通识

保世海　李　锐　**主　编**

马　欣　王煜慕　王婷婷　刘　彬　**副主编**

清華大學出版社

北　京

内 容 简 介

本书遵循“循序渐进、螺旋上升”的编写思路，构建了“基础认知—应用实践—专项突破—综合提升—拓展延伸”的五模块立体结构。全书共五篇，内容逐篇深入，第Ⅰ篇“基础认知篇”作为入门基石，系统讲解AI的定义、发展简史、核心技术、应用领域、法律法规与伦理安全，破除技术误解，搭建完整的知识框架；第Ⅱ篇“人工智能应用篇”聚焦AI在文本创作、沟通协作、知识管理等场景的落地应用，解析Transformer架构、大语言模型等关键技术，引入提示工程、AIGC赋能等实践内容；第Ⅲ篇“人工智能语言篇”讲解Python语言，通过实例介绍AI在编程与数据分析两大关键场景中的应用，培养学生的计算思维，提升其对编程的认知水平；第Ⅳ篇“人工智能实践篇”以动手操作为核心，设计看图说话、智慧驾驶模拟、语音识别等实操任务，强化“做中学”的教学效果；第Ⅴ篇“拓展篇(学赛结合)”衔接学科竞赛与产业需求，拓宽学生的创新视野与综合应用能力，实现知识学习与能力提升的闭环。本书强调任务驱动与场景融合，各任务均设计有明确的应用情境，学生可根据自身专业方向选择实践内容。

本书适合作为高等职业院校和中等职业学校人工智能通识类课程的教材，也可作为广大人工智能初学者的自学参考资料。

本书封面贴有清华大学出版社防伪标签，无标签者不得销售。
版权所有，侵权必究。举报：010-62782989，beiqinquan@tup.tsinghua.edu.cn。

图书在版编目(CIP)数据

人工智能应用通识 / 保世海 , 李锐主编 . -- 北京 : 清华大学出版社 , 2026. 2. -- ISBN 978-7-302-70721-9

Ⅰ. TP18

中国国家版本馆 CIP 数据核字第 2026DT0345 号

责任编辑：王　军
封面设计：周晓亮
版式设计：方加青
责任校对：成凤进
责任印制：刘海龙

出版发行：清华大学出版社
网　　址：https://www.tup.com.cn，https://www.wqxuetang.com
地　　址：北京清华大学学研大厦 A 座　　**邮　　编：**100084
社 总 机：010-83470000　　**邮　　购：**010-62786544
投稿与读者服务：010-62776969，c-service@tup.tsinghua.edu.cn
质 量 反 馈：010-62772015，zhiliang@tup.tsinghua.edu.cn
印 装 者：天津鑫丰华印务有限公司
经　　销：全国新华书店
开　　本：190mm×260mm　　**印　　张：**12　　**字　　数：**248 千字
版　　次：2026 年 2 月第 1 版　　**印　　次：**2026 年 2 月第 1 次印刷
定　　价：49.80 元

产品编号：116057-01

教材编写委员会

主　编：保世海　李　锐

副主编：马　欣　王煜慕　王婷婷　刘　彬

参　编：高　丽　张健之　王金金　韩　冰
金　晶　史竣娇　李林海　廖婕辰
宋侠芸

前言

PREFACE

在数字技术重塑全球发展格局的今天，人工智能(AI)已从实验室走向产业一线，成为推动社会进步、产业升级与文明演进的核心驱动力。从自动驾驶的智能决策到医疗诊断的辅助支撑，从个性化推荐的精准适配到生成式AI的创意赋能，AI技术正深度融入经济社会各领域，重构生产方式、生活场景与学习范式，同时对劳动者的跨学科素养、技术应用能力与伦理判断力提出了全新要求。培养具备AI通识认知、创新思维与社会责任感的应用型人才，已成为教育回应时代变革的核心命题。

本书立足人工智能发展的时代浪潮，紧扣“科技向善、知行合一”的教育理念，以培养“懂技术、明伦理、善应用、敢创新”的复合型人才为核心目标。全书内容始终坚持三大导向：一是**时代性导向**，紧跟AI技术前沿与产业实践，将生成式AI、大模型、多模态学习等核心技术与应用场景纳入教学范畴，兼顾技术原理与行业动态；二是**素养导向**，突破“技术工具化”的认知局限，注重培养学生的逻辑思维、跨学科整合能力与AI伦理安全意识，引导学生理性看待技术边界与社会影响；三是**实践性导向**，强化理论与实操的深度融合，通过真实案例讲解、实验与项目设计实践，让学生在实践中理解技术逻辑、掌握应用方法。

本书遵循“循序渐进、螺旋上升”的编写思路，构建了“基础认知—应用实践—专项突破—综合提升—拓展延伸”的五模块立体结构。

第Ⅰ篇“基础认知篇”作为入门基石，系统讲解AI的定义、发展简史、核心技术、应用领域、法律法规与伦理安全，破除技术误解，搭建完整的知识框架。

第Ⅱ篇“人工智能应用篇”聚焦AI在文本创作、沟通协作、知识管理等场景的落地应用，解析Transformer架构、大语言模型等关键技术，通过提示工程、AIGC赋能等实践内容，提升学生的技术应用能力。

第Ⅲ篇“人工智能语言篇”讲解Python语言，通过实例介绍AI在编程与数据分析两大关键场景中的应用，培养学生的计算思维，提升其对编程的认知水平。

第Ⅳ篇“人工智能实践篇”以动手操作为核心，设计智能制造、智慧零售、智能体、行业数据分析等实操任务，强化“做中学”的教学效果。

第Ⅴ篇“拓展篇(学赛结合)”衔接学科竞赛与产业需求，拓宽学生的创新视野与综合应用能力，实现知识学习与能力提升的闭环。

本书既强调核心技术自主可控的家国情怀，又注重引导学生树立正确的技术价值观，理解AI在替代岗位的同时创造新职业的辩证关系，培养终身学习的意识与社会责任感。本书提供电子教案PPT，可扫封底二维码下载。书中“练一练”的答案及解析，仅提供给授课老师，可联系QQ：1815317009索取。

我们期望，无论是AI入门学习者，还是希望提升数字素养的职场人士，都能通过对本书的学习：夯实AI通识基础，明晰技术原理与应用边界；锤炼实践创新能力，将AI工具转化为学习与工作的赋能手段；坚守伦理安全底线，在技术应用中兼顾效率与公平、创新与责任。愿每位读者都能以本书为起点，主动拥抱AI时代的变革与机遇，在技术与人文的融合中成长为推动数字文明进步的积极力量，在人工智能赋能的未来世界中实现自身价值与社会价值的统一。

编 者

2025年末

目录

CONTENTS

第Ⅰ篇　基础认知篇

第 1 章　人工智能基础 / 2

1.1　AI简述 / 4

1.1.1　AI的经典定义 / 4

1.1.2　对AI的误解 / 5

1.1.3　科学理解AI / 5

1.2　AI发展简史 / 6

1.2.1　理论奠基 / 6

1.2.2　达特茅斯会议 / 7

1.2.3　三次发展浪潮 / 7

1.3　AI领域的法律和法规 / 9

1.3.1　AI领域的法律和法规现状 / 9

1.3.2　重点场景下AI技术法律适用的现实困境 / 10

1.3.3　AI合规实践要点与风险防控策略 / 11

1.4　AI核心技术和辅助技术 / 11

1.4.1　机器学习 / 11

1.4.2　深度学习 / 12

1.4.3　自然语言处理 / 13

1.4.4　计算机视觉 / 13

1.4.5　知识图谱 / 14

1.4.6 多模态学习 / 14
1.4.7 AI辅助技术 / 14
1.5 常用AI工具 / 15
1.5.1 聊天对话类AI / 15
1.5.2 内容创作类AI / 16
1.5.3 代码编程类AI / 16
1.5.4 办公辅助类AI / 17
1.5.5 搜索类AI / 17
1.5.6 AI工具的选择 / 18
1.6 数字意识 / 19
1.6.1 数字意识概述 / 19
1.6.2 数字化学习与社会责任 / 23
1.6.3 数字内容创新与安全防护 / 27
1.6.4 安全防护的核心策略 / 30
1.7 任务1：看图说话 / 30
1.8 任务2：智慧驾驶 / 31
1.9 任务3：单音节语音识别 / 31
1.10 本章小结 / 31
1.11 练一练 / 31

第Ⅱ篇 人工智能应用篇

第 2 章 AI 赋能文本创作与沟通 / 36

2.1 AIGC概述 / 40
2.2 AIGC的核心技术 / 41
2.2.1 大语言模型 / 42
2.2.2 扩散模型 / 42
2.2.3 生成对抗网络 / 42
2.2.4 多模态学习 / 42
2.2.5 提示工程 / 43
2.2.6 模型优化技术 / 43

2.3 Transformer架构 / 44
2.3.1 自注意力机制的核心优势 / 45
2.3.2 Transformer的巧妙设计 / 46
2.4 提示工程的方法 / 46
2.5 提示的支柱：CLEAR结构化框架 / 47
2.5.1 背景设定 / 47
2.5.2 逻辑与任务 / 48
2.5.3 范例 / 48
2.5.4 角色扮演 / 48
2.5.5 结果格式 / 49
2.6 高级提示技巧与模板设计 / 49
2.6.1 高级提示技巧 / 49
2.6.2 不同任务场景下的提示模板设计示例 / 52
2.7 AI提示词入门手册 / 55
2.7.1 驯服你的AI助手 / 55
2.7.2 AIGC进阶训练 / 59
2.7.3 结语：AIGC不是“搜寻”，而是“共创” / 64
2.8 任务1：文本创作实践 / 64
2.9 任务2：会议记录智能化 / 64
2.10 任务3：实时翻译 / 65
2.11 本章小结 / 65
2.12 练一练 / 65

第 3 章 AI 驱动学习研究与知识管理 / 68

3.1 学习科学与认知理论 / 72
3.1.1 学习科学 / 72
3.1.2 当代认知理论的发展趋势 / 73
3.1.3 学习科学与认知理论的关联 / 74
3.2 知识获取和提取技术 / 75
3.2.1 知识获取技术 / 75
3.2.2 知识提取技术 / 76
3.2.3 知识获取和提取的工具与框架 / 77
3.2.4 应用场景 / 77

3.2.5 面临的挑战与未来趋势 / 77
3.3 任务1：体验AIGC动态学习伙伴 / 78
3.4 任务2：AI构建个人知识网络 / 78
3.5 本章小结 / 78
3.6 练一练 / 79

第 4 章 AI 生成与编辑内容 / 81

4.1 图像生成基础 / 84
4.1.1 文本编码 / 84
4.1.2 前向扩散 / 85
4.1.3 噪声预测与反向扩散 / 85
4.1.4 图像生成 / 85
4.2 文生图提示词技巧与实践 / 87
4.3 图像编辑与风格迁移 / 88
4.4 语音合成技术原理与实践 / 89
4.4.1 精准类比 / 89
4.4.2 严谨定义 / 89
4.5 视频生成与编辑基础 / 92
4.5.1 文生视频模型 / 92
4.5.2 智能视频编辑技术 / 92
4.6 任务1：视觉内容的商业应用案例 / 93
4.7 任务2：AI音乐生成工具的使用 / 94
4.8 任务3：数字人应用示例 / 94
4.9 本章小结 / 94
4.10 练一练 / 95

第Ⅲ篇 人工智能语言篇

第 5 章 AI 辅助编程与数据分析入门 / 98

5.1 Python基础 / 101
5.1.1 Python概述 / 101

5.1.2 Python的基础内容 / 102
5.1.3 Python在人工智能中的应用简述 / 106
5.2 数据分析与可视化关键技术 / 106
5.2.1 数据分析基础概念与流程 / 106
5.2.2 数据预处理 / 107
5.2.3 机器学习 / 108
5.2.4 数据可视化方法 / 109
5.3 典型工具 / 112
5.3.1 AI辅助编程工具 / 112
5.3.2 数据分析工具 / 113
5.3.3 数据可视化工具 / 113
5.3.4 综合开发环境与平台 / 114
5.3.5 云计算与AI平台 / 114
5.4 任务1：AI编程助手的基本使用 / 114
5.5 任务2：代码生成与调试 / 114
5.6 任务3：数据分析任务自动化 / 114
5.7 本章小结 / 115
5.8 练一练 / 115

第Ⅳ篇 人工智能实践篇

第 6 章 智能制造——产品质量检测方向 / 120

6.1 积木编程大模型API理论 / 124
6.2 积木编程可视化逻辑 / 124
6.3 大模型API机制 / 126
6.4 工业质量标准智能检索 / 127
6.5 自动化质检报告闭环系统 / 128
6.6 生产故障智能诊断助手 / 129
6.7 任务1：工业质量标准智能检索 / 131
6.8 任务2：自动化质检报告闭环系统 / 131
6.9 任务3：生产故障智能诊断助手 / 131

6.10 本章小结 / 132
6.11 练一练 / 132

第 7 章 智慧零售——个性化服务方向 / 134

7.1 提示词设计与参数调优 / 137
7.2 API调用中的安全伦理规范 / 138
7.3 任务1：AI个性化商品推荐系统 / 139
7.4 任务2：全场景智能客服系统 / 139
7.5 任务3：智能促销活动策划师 / 140
7.6 本章小结 / 140
7.7 练一练 / 140

第 8 章 旅游规划——来一场说走就走的旅行 / 142

8.1 智能体概述 / 145
8.1.1 智能体的定义 / 145
8.1.2 智能体的核心特征 / 146
8.1.3 智能体的分类 / 147
8.2 智能体行业应用场景 / 150
8.2.1 教育领域：个性化学习助手 / 150
8.2.2 金融领域：智能投资顾问 / 150
8.2.3 医疗健康：健康监测与辅助诊断系统 / 151
8.2.4 智慧城市：交通与环境智能管控系统 / 151
8.2.5 智能制造：自主控制与协作机器人 / 152
8.2.6 电商与客户服务：虚拟客服与智能推荐系统 / 152
8.2.7 游戏与虚拟世界：交互式NPC与虚拟导览系统 / 152
8.3 任务：旅游规划助手智能体 / 153
8.4 本章小结 / 153
8.5 练一练 / 153

第 9 章 行业应用综合实践 / 156

9.1 商业分析领域应用 / 159
9.1.1 发掘短视频潜力品类 / 159

9.1.2 直播数据可视化分析 / 161
9.2 文创艺术领域应用 / 161
9.2.1 用大模型写优美散文 / 161
9.2.2 用大模型绘画 / 163
9.3 工业与能源科技领域应用 / 164
9.4 医疗健康领域应用 / 165
9.5 任务1：认识短视频数据特征 / 166
9.6 任务2：认识直播数据特征 / 166
9.7 任务3：用大模型写优美散文 / 166
9.8 任务4：用大模型绘画 / 167
9.9 任务5：新能源汽车能耗分析 / 167
9.10 任务6：中医症候关联规则 / 167
9.11 本章小结 / 167
9.12 练一练 / 168

第Ⅴ篇 拓展篇（学赛结合）

第 10 章 人工智能大赛模拟赛 / 172

10.1 竞赛要求 / 173
10.1.1 知识要求 / 173
10.1.2 能力要求 / 173
10.1.3 素养要求 / 174
10.2 竞赛内容 / 174
10.2.1 赛项考查的技术与技能 / 174
10.2.2 涵盖的职业典型工作任务 / 175
10.2.3 检验选手核心能力与职业综合能力 / 175
10.2.4 竞赛规则 / 176
10.3 竞赛样题 / 178

参考文献 / 179



第Ⅰ篇

基础认知篇

第1章　人工智能基础

当前，人工智能(artificial intelligence，AI)技术已经成为推动社会进步和产业升级的关键力量。本章旨在帮助学生直观地理解AI的定义、工作原理与应用领域，深入体验AI技术在实际场景中的应用，培养逻辑思维能力和创新能力，为未来参与更复杂的AI项目奠定坚实的基础。

本章还介绍当前技术生态对跨学科知识的需求。自动驾驶、文本识别等技术的实现不仅依赖计算机科学和AI算法，还涉及机械工程、电子工程、控制理论等多个学科的协同合作。通过本章的学习，学生将接触到此类前沿技术和跨学科知识，拓宽视野，提升综合素养，为满足未来技术发展的多元化需求做好准备。

知识目标

- 理解AI的经典定义，能准确区分强AI与弱AI 的本质差异，理性看待 AI 的应用边界与价值。
- 掌握 AI 发展的关键理论奠基事件，了解达特茅斯会议(1956年召开)的历史地位，掌握 AI 三次发展浪潮的核心特征。
- 理解推荐系统(电商平台、短视频平台)的技术原理，掌握智能助手的技术栈，能够分析人脸识别与美颜技术的实现逻辑。
- 了解与AI相关的法律法规，掌握国内外重要AI法规的核心内容。
- 掌握机器学习的四大类型及应用。
- 能根据具体需求(如撰写文案、生成图像、辅助编程)选择合适的 AI 工具，并说明选择依据。

能力目标

- 能够通过积木式编程完成实践项目(如智能车路径规划、语音识别模块调用)，体验AI技术落地过程。
- 能够分析个性化推荐系统的技术逻辑与潜在问题。

素质目标

- 培养对AI伦理问题的敏感度，理解算法偏见、隐私风险等挑战的成因与应对策略。
- 形成辩证视角，理性探讨AI对社会就业、教育结构的长期影响。

启智与致知

- 培养历史视野和辩证思维，分析AI在替代部分岗位的同时，如何创造新职业；分析AI如何实现个性化学习。引导学生理解，技术革命总会伴随社会结构调整，关键是终身学习、提升不可替代的核心素养，以适应未来发展。
- 树立总体国家安全观。通过讨论AI在军事、经济、舆论等领域的应用，让学生理解AI技术关乎国家安全与发展大局，树立忧患意识和总体国家安全观。

工具准备

人工智能教育平台。

本章内容分析

本章将介绍AI的发展历史，以及AI的经典定义、应用领域、相应的法律法规、核心技术等，具体的能力思维导图和知识思维导图如图1.1和图1.2所示。本章最后通过三个典型任务说明AI如何工作，以及如何对AI的效果进行评价。



图1.1　能力思维导图

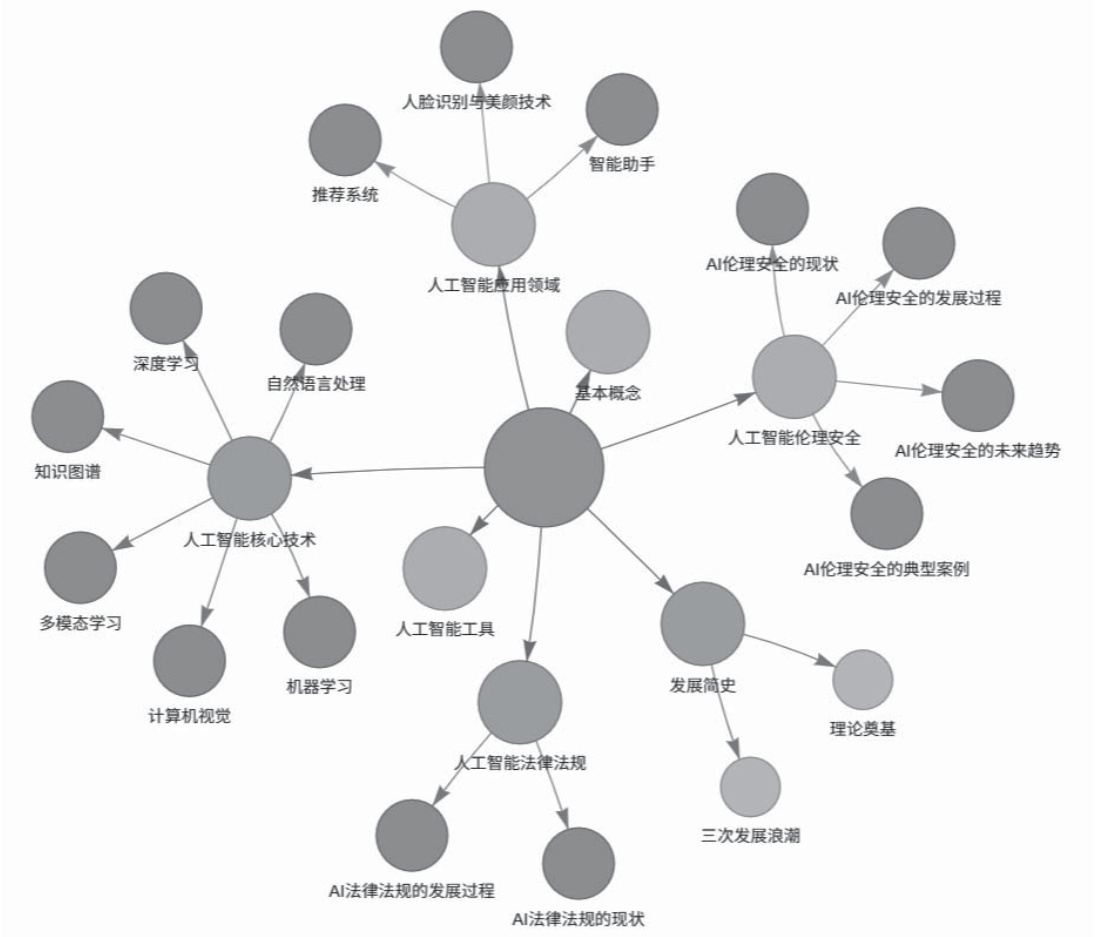


图1.2　知识思维导图

1.1　AI简述

1.1.1　AI的经典定义

AI是通过计算机算法和模型来模拟人类智能的技术；是研究、开发用于模拟、延伸和扩展人的智能的理论、方法、技术及应用系统的新技术科学；也是计算机科学的一个分支。人工智能旨在探究智能的本质，并生产出一种新的能以与人类智能相似的方式做出反应的智能机器，该领域的研究包括机器人、语言识别、图像识别、自然语言处理和专家系统等。AI的基本概念包括算法、模型、数据集和学习算法等。算法是AI的核心，它是一组定义明确的计算步骤，能够自动执行某些任务。模型是对现实世界的简化描述，用于预测和解释数据。数据集是训练模型所需要的大量数据。学习算法是让模型能够自动从数据中学习并改进的算法。AI的应用领域非常广泛，包括但不限于医疗、交通、金融、教育等。

1.1.2　对AI的误解

目前的AI缺乏人类意识或公共智能，只有执行特定工作内容的能力，缺乏独立思考和推理能力，很难进行跨领域应用，无法理解用户发出的复杂指令，缺乏自我意识和创造力。

强AI(strong AI)指的是具有类似人类的意识、理解能力和自我意识的智能系统。然而，在众多的科幻作品以及媒体的渲染中，强AI经常被人们误解为拥有感情、具备自由意志，甚至可能会毁灭人类的“超级智能”。这种误会使得许多人对AI产生不必要的恐惧。

人们需要客观认识到，当前所有的AI都不具备人的理解力和意识，只能在特定的领域模仿人的行为和思考方式，所以，它们都属于弱AI(weak AI)。例如，聊天机器人虽然可以和人进行沟通与交流，但它并不理解交流的语言的含义，只是模仿人的行为进行计算概率的操作。

在这个时代，强AI仍是理论设想，并没有成为现实，相关的研究也面临哲学、技术和伦理等多方面的挑战。将现有AI与强AI等同，是对AI能力的高估，也是对AI的误解，容易引发技术恐慌，也将影响AI的理性发展。

1.1.3　科学理解AI

当前AI仍依赖预设算法和数据输入，它的自主性实质上是在算法框架下的有限选择。

在AI的发展过程中，公众和媒体经常夸大AI系统的“自主性”，认为它们可以独立思考、自主决策，甚至可以脱离人类控制自主运行。这种“完全自主性”的误区通常源于人类对科学技术的想象而非对现实世界真实技术水平的准确认知。当前的AI系统做出的决定都是基于预设算法和训练数据执行，并非真正意义上的独立判断，都是在人类设定的算法框架内进行计算和选择。

“AI拥有完全自主性”这一认知可能会导致人们对AI的过度恐慌，也给AI的伦理和法律责任的判定带来很大的困难。例如，部分AI产品出现了问题，那么需要负责的是产品拥有者还是产品研发者？

因此，我们应理性看待AI的运行机制并避免为AI赋予夸大的自主性能力，防止干扰对现实AI系统的科学理解。

随着AI的发展，人们担忧自己会被AI所替代，担心失业、技能贬值甚至社会失衡，特别是在自动化、写作、客服、制造等领域。

其实，AI与人类擅长的事情有所不同，AI擅长具有明确的规则或者不断重复的任务，不能完成涉及创造力、情感交流和伦理判断的复杂工作。AI的应用更可能带来“人机协

作”而非“人类被淘汰”。人们应主动适应变化，通过终身学习掌握AI难以替代的创造、情感交流等能力。每次技术革新虽然会导致岗位变化，但也会出现新职业与新技能需求，解决问题的关键在于如何应对转型，而不是恐惧这种改变。

1.2 AI发展简史

1.2.1 理论奠基

1950年，英国数学家、逻辑学家艾伦·图灵(Alan Turing)在其发表的论文《计算机器与智能》(“Computing Machinery and Intelligence”)中提出了一个划时代的问题：“机器能思考吗？”为了回答这个问题，图灵设计了一种检验机器是否具备智能的方法，这就是图灵测试(Turing test)。

图灵测试的核心思想如下：如果一个人和某机器进行交流时并不能分辨出这个机器是不是一个机器，那么这个机器就可以被视为具备智能。在图灵测试中，应该安排一个人分别与真人和机器对话，在不知身份的情况下进行判断，若机器能够通过模仿人类思维与语言交流骗过测试者，就说明其具有类人智能(见图1.3)。

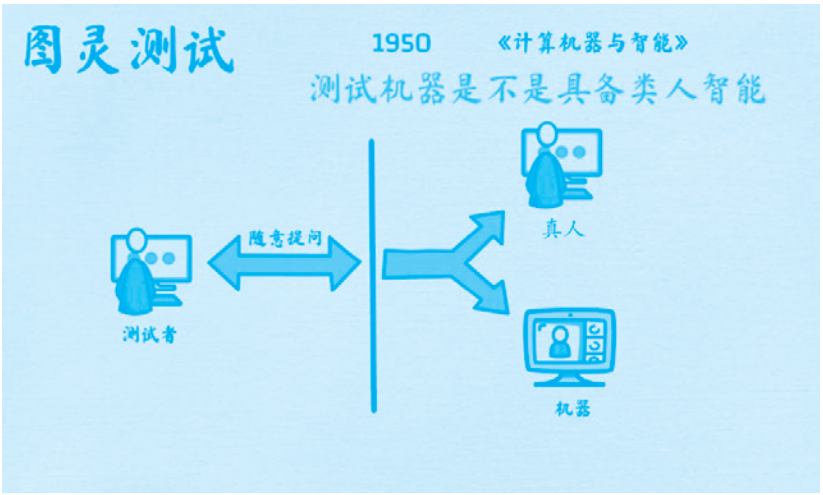


图1.3 图灵测试

图灵测试的意义在于为后来AI的研究提供了理论基础，图灵首次将智能的定义从抽象的哲学问题转变为一种可验证的操作。图灵测试至今仍被广泛讨论，是AI伦理、能力边界和意识模拟等领域研究的出发点之一。

所以，图灵测试不仅是一种技术上的评判方法，更像是一种哲学与科学相融合的思维方式，是AI从假设走向科学探索的第一步。

1.2.2　达特茅斯会议

对于机器是否能够模仿人类的思考方式，不同学科并没有统一的定义或研究方向。在这样的背景下，约翰·麦卡锡、马文·明斯基、克劳德·香农、纳撒尼尔·罗切斯特等人，于1956年在美国达特茅斯学院组织了一次夏季研究会议，图1.4展示了当时会议的地点。会议的核心目的是提出一个框架，用来定义智能是什么，并探讨如何让计算机具备这种智能。

麦卡锡在会议提案中写道：“我们假设，每一种学习或智能的特性原则上都能被精确描述，因此可以由机器模拟。”达特茅斯会议中，约翰·麦卡锡和马文·明斯基等研究人员正式提出了AI学科的名称，并将研究目标定为“通过机器模拟人类的智能行为”，推动了符号主义AI的兴起，开启了人类智能研究的新纪元。



图1.4　达特茅斯会议地点

1.2.3　三次发展浪潮

AI的三次发展浪潮如表1.1所示。

表1.1　AI的三次发展浪潮

阶段	核心技术	典型应用	局限因素
第一次发展浪潮	符号逻辑	专家系统	规则维护成本高
第二次发展浪潮	机器学习	棋类博弈系统	特征工程依赖性强
第三次发展浪潮	深度学习	智能助手、AIGC[①]	算力与数据需求大

注：AIGC 即人工智能生成内容。

1. 第一次发展浪潮：符号逻辑AI为主导

20 世纪50年代至70年代，AI开始了以符号逻辑AI为主导的快速发展阶段。这一阶段的典型应用是专家系统，核心思想是人类思维可以使用符号进行表达，智能行为也可以通过逻辑推理和规则进行模拟。

这次发展浪潮表明AI可以在专业场景中协助决策，但系统依赖人工建模，缺乏适应性。

2. 第二次发展浪潮：机器学习兴起

20世纪90年代，AI领域进入了以机器学习为核心特征的新发展阶段，这一阶段标志着以机器学习为核心的技术开始兴起。在这个阶段，传统的符号主义方法逐渐被基于统计学原理的学习算法所替代。机器学习的核心理念是通过分析大量数据自动发现并提取规律，使计算机能在无须人为编写规则的情况下完成预测和分类任务。

1997年，IBM开发的超级计算机“深蓝”(Deep Blue)在一场六局比赛中战胜了国际象棋世界冠军卡斯帕罗夫，这是历史上AI系统首次在正式比赛中击败世界冠军。这一事件标志着计算机在复杂逻辑与决策领域取得重大突破。

这一阶段的技术发展为日后深度学习、大数据驱动的AI系统打下了坚实基础，AI由此正式进入“从经验中学习”的新时代。

3. 第三次发展浪潮：深度学习突破

进入21世纪，AI迎来了快速发展阶段，并取得了深度学习的重大突破。深度学习利用神经网络的多层结构，能够自动从海量数据中提取复杂特征，大幅提升了图像识别、语音识别、自然语言处理等多种技术的水平。

2016年，AI系统AlphaGo成功战胜围棋世界冠军李世石，在人类与机器的复杂博弈中，机器首次取得了胜利。近年来，ChatGPT等大语言模型展现出了惊人的语言生成和交流能力。卷积神经网络和Transformer模型在图像识别与自然语言处理方面取得了显著进展。

深度学习的成功不仅重塑了AI的技术路径，也极大地扩展了AI技术在教育、医疗、工业和艺术等领域的应用范围。表1.2给出了AI和人类在三个维度上的区别。

表1.2　对比AI和人类

维度	AI的优势	人类的优势
计算能力	高速、精确的数值处理	模糊推理与直觉判断
学习方式	大数据驱动模式识别	小样本迁移学习能力
创造力	组合式创新(如AIGC)	原创性概念生成

当前，AI发展仍处于感知智能向认知智能过渡的阶段，需要理性看待其技术边界。

1.3　AI领域的法律和法规

1.3.1　AI领域的法律和法规现状

AI领域的法律和法规是随着AI技术发展而逐渐完善的，旨在规范AI相关活动，保障其健康、安全、有序发展。目前，世界各地已初步构建起涵盖法律、行政法规、部门规章及地方性法规等多层次，以及数据、算法、应用等多维度的人工智能法律法规框架。

1. 我国的法律法规

《互联网信息服务算法推荐管理规定》自2022年3月1日起施行，旨在规范互联网信息服务算法推荐活动，维护国家安全和社会公共利益，保护公民、法人和其他组织的合法权益。该规定要求算法推荐服务提供者应当坚持主流价值导向，履行算法安全主体责任，建立健全算法机制机理审核、科技伦理审查、用户注册、信息发布审核、数据安全和个人信息保护、反电信网络诈骗、安全评估监测、应急处置等管理制度和技术措施。

《互联网信息服务深度合成管理规定》自2023年1月10日起施行，对深度合成服务进行了规范，明确了深度合成服务提供者和使用者的权利、义务和责任。该规定要求深度合成服务提供者建立健全管理制度和技术保障措施，加强内容管理，规范标识等，以防范深度合成技术被滥用带来的风险。

《生成式人工智能服务管理暂行办法》是我国针对生成式人工智能领域的重要法规，2023年由国家网信办等七部门联合发布，自 2023年8月15日起施行。该办法明确了利用生成式人工智能技术向中华人民共和国境内公众提供生成文本、图片、音频、视频等内容的服务适用此办法。该办法规定提供和使用生成式人工智能服务，应当遵守法律、行政法规，尊重社会公德和伦理道德，如不得生成违法有害内容、防止产生歧视、尊重知识产权和他人合法权益等；同时，对生成式人工智能服务提供者在训练数据处理、服务规范、监督检查等方面提出了具体要求，如使用合法来源数据、保护用户信息、对生成内容进行标识等。

《网络数据安全管理条例》于2024年9月公布，设置了生成式人工智能专条，明确加强对训练数据和训练数据处理活动的安全管理，完善了生成式人工智能数据安全制度。

《中华人民共和国民法典》明确禁止利用信息技术手段伪造他人肖像等侵害他人肖像权的行为，并参照肖像权保护的有关规定对自然人声音进行保护。《中华人民共和国反垄断法》禁止利用算法实施垄断行为。《中华人民共和国电子商务法》《中华人民共和国个人信息保护法》等对个性化推荐、自动化决策进行了规定。

此外，我国还颁布了《新一代人工智能发展规划》《关于推动未来产业创新发展的实施意见》《人工智能安全治理框架》等政策或文件，全方位、多维度促进并规范人工智能的发展。

2. 其他国家或地区的法律法规

欧盟于2023 年通过《人工智能法》，该法规成为全球人工智能治理的制度蓝本。该法规旨在通过规范人工智能技术的开发和应用，确保人工智能的发展符合人类的价值观和利益，如强调人工智能系统的安全性、透明度和可问责性等，对不同风险等级的人工智能系统采取不同的监管措施。

美国作为AI技术领先国，目前并不急于进行专门立法，而是有意利用“制度真空期”快速占领全球市场，后续再通过市场贸易向世界输出美国规则，但在一些具体领域(如数据保护、算法偏见等)也有相关的政策和监管行动。

1.3.2　重点场景下 AI 技术法律适用的现实困境

在人工智能技术快速落地的诸多场景中，生成式人工智能与深度合成技术领域的法律适用问题尤为突出，其技术特性与现行法律框架之间的张力日益凸显，成为当前司法实践与监管工作的核心难点。

对于生成式人工智能而言，其法律适用的争议主要集中在三个层面。

首先，作品权属认定面临法理挑战。当AI生成的文字、图像、音乐等内容达到独创性标准时，能否构成著作权法意义上的“作品”存在理论分歧。

其次，侵权责任界定存在制度空白。当AIGC产出的内容涉及侵犯他人著作权、肖像权或名誉权时，责任分配机制尚不清晰。

最后，内容合规监管面临技术瓶颈。AIGC具有生成速度快、内容规模大、表现形式隐蔽等特点，传统人工审核模式已难以满足监管需求。如何在保障内容安全的同时维护创作自由，如何在提高审核准确性与保护用户隐私之间取得平衡，成为监管部门亟须解决的重要问题。

深度合成技术的法律适用困境则主要体现在真实性认定与责任追溯两个维度。在真实性认定维度，深度合成生成的视频、音频等内容逼真度极高，普通公众的辨别能力严重不足，这为虚假信息传播、网络诈骗等违法犯罪行为提供了技术温床。

在责任追溯维度，深度合成技术链条涉及模型开发、数据采集、算法训练、内容生成、传播分发等多个环节，形成复杂的责任主体网络。当发生侵权事件时，受害者往往面临举证困难、责任主体多元、因果关系复杂等障碍。

为应对这些挑战，需要在立法层面明确AIGC与深度合成技术的法律属性，建立适应技术特点的权属认定规则和责任分配机制；在监管层面推动技术治理创新，开发更高效的内容识别与溯源工具；在司法层面完善证据规则和认证标准，为权利救济提供更有效的程序保障。只有通过多方协同治理，才能在促进技术创新的同时，有效防范其可能带来的法律风险。

1.3.3　AI合规实践要点与风险防控策略

对于企业而言，构建系统化的AI合规体系与风险防控机制已成为应对日益复杂的AI法律环境的必然选择。基于当前全球AI立法趋势和行业最佳实践，企业应当从合规管理体系建设、风险识别与防控、合规能力提升三个方面构建全方位的AI合规体系。

在合规管理体系建设方面，企业需要构建覆盖决策层、管理层和执行层的三级治理架构。

在风险识别与防控方面，企业应当基于AI应用生命周期建立四维防护体系。

在合规能力提升方面，企业需要在技术工具、人才队伍和生态协作三个方向持续投入。

通过构建治理、风控、能力三位一体的合规体系，企业不仅能够有效应对当前的监管要求，更能在快速演进的AI法律环境中保持主动，实现创新与合规的平衡发展。这种系统化的合规建设，将成为企业在AI时代的核心竞争力之一。

1.4　AI核心技术和辅助技术

AI核心技术是支撑其实现感知、理解、推理、决策、生成等智能能力的底层技术体系，涵盖从数据处理到模型构建、从单一任务到跨模态交互的多个维度。其核心技术有6种，辅助技术有6种。

1.4.1　机器学习

机器学习(machine learning，ML)是 AI 的“基础引擎”，是让机器通过数据自主学习规律，而非依赖人工编写的固定规则。其本质是从数据中拟合出映射关系(输入→输出)，并用于预测或决策。根据学习方式，机器学习可分为四大类。

1. 监督学习

监督学习利用标注数据(输入+对应标签)训练模型，学习从输入到标签的映射。典型任务包括分类(如垃圾邮件识别、图像分类)、回归(如房价预测、温度预测)。主要算法有逻辑回归、支持向量机(SVM)、决策树、随机森林、梯度提升树(GBDT、XGBoost)等。

2. 无监督学习

无监督学习仅用无标注数据训练，目标是发现数据内在结构(如聚类、降维)。典型任务包括聚类(如用户分群)、异常检测(如信用卡欺诈识别)、维度约减(如 PCA 降维)。主要算法有K-Means 聚类、DBSCAN、主成分分析(PCA)、自编码器(Autoencoder)等。

3. 半监督学习

半监督学习结合少量标注数据和大量无标注数据训练，解决标注成本高的问题(如医疗影像分析，标注数据稀缺)。

4. 强化学习

强化学习(reinforcement learning，RL)通过试错学习最优策略——智能体(agent)与环境(environment)交互，通过动作(action)获得奖励(reward)，逐步优化行为以最大化累积奖励。这种方法的核心框架是“状态→动作→奖励→新状态”的循环，主要应用有AlphaGo、机器人控制(如机械臂抓取)、游戏 AI(如星际争霸AI)等。

1.4.2 深度学习

深度学习(deep learning，DL)是机器学习的分支，基于深层神经网络(多隐藏层的神经网络)实现对复杂数据的自动特征提取，突破了传统机器学习对人工特征工程的依赖，是当前 AI 爆发的核心驱动力。

1. 神经网络

神经网络主要由输入层(接收原始数据)、隐藏层(提取特征)、输出层(输出结果)组成，层与层之间通过神经元连接，每个连接有权重，通过训练优化权重。它的激活函数赋予网络非线性能力(否则深层网络等效于单层线性模型)，如 ReLU(修正线性单元)、Sigmoid、Tanh等。

2. 关键深度学习模型

卷积神经网络(CNN)专门为处理网格结构数据(如图像、音频)设计，通过卷积层(局部感受野 + 权值共享)高效提取局部特征(如边缘、纹理)，主要应用包括图像分类(如ResNet)、目标检测(如YOLO、Faster R-CNN)、医学影像分析等。

3. 循环神经网络

循环神经网络(RNN)主要用来处理序列数据(如文本、语音、时间序列)，通过“循环连接”保留历史信息(如句子中前词对后词的影响)，改进模型是长短期记忆网络(LSTM)、门控循环单元(GRU)等，解决 RNN 的长程依赖遗忘问题。RNN的主要应用包括语音识别、机器翻译(早期)、时间序列预测等。

4. Transformer 模型

基于自注意力机制(让模型关注输入内容中的重要部分，如句子中的“他”指代“小明”)，该模型的并行计算能力远超 RNN，是当前自然语言处理和计算机视觉的核心架构，其核心特征是多头注意力和前馈网络。它的衍生模型包括BERT(双向预训练，擅长理解)、GPT(自回归生成，擅长文本生成)、LLaMA 等大语言模型。

1.4.3　自然语言处理

自然语言处理(natural language processing，NLP)是让机器理解、生成、交互人类语言的技术，核心是架起人类语言与机器可处理形式的桥梁。它的核心任务是基础理解、生成与交互。基础理解包括分词、词性标注(如“苹果”是名词、品牌)、命名实体识别(NER，如识别“北京”是城市)、句法分析(如分析句子结构)、语义理解(如“打太极”的隐喻)等。生成与交互则是完成机器翻译(如中英互译)、文本生成(如写诗、写代码)、问答系统(如ChatGPT 问答)、情感分析(如评论褒贬判断)等。NLP的关键技术主要包括以下几种。

1. 词向量

词向量(word embedding)将词语转化为低维稠密向量(如 300 维)，使语义相关的词向量距离更近(如“国王−男人+女人≈女王”)。典型代表是Word2Vec(如Skip-gram、CBOW)、GloVe和FastText。

2. 预训练语言模型

预训练语言模型(PLM)先在海量文本(如互联网语料)上预训练通用语言能力，再针对具体任务进行微调，大幅提升效果。典型代表是BERT(双向编码，适合理解任务)、GPT 系列(自回归生成，适合生成任务)、T5(Text-to-Text统一框架)等。

3. 语言模型

参数量达百亿甚至万亿级的预训练模型(如 GPT-4、文心一言)通过涌现能力实现复杂推理、多轮对话、跨领域任务处理。

1.4.4　计算机视觉

计算机视觉(computer vision，CV)让机器“看懂”图像和视频，核心是从像素中提取语义信息(如识别物体、场景、动作)。它的核心功能是图像理解、视频理解和图像生成。其中，图像理解主要是完成分类(如“这是猫还是狗”)、目标检测(定位并识别物体，如“图中有 3 辆汽车，坐标分别是……”)、图像分割(像素级分类，如“哪部分是草地，哪部分是天空”)等。视频理解则是实现目标跟踪(追踪物体在视频中的运动)、行为识别(如“人在跑步”)、视频生成(如“根据文本生成视频”)等功能。图像生成主要基于文本/图像生成新图像，如生成对抗网络(GAN)、扩散模型(如 Stable Diffusion、Midjourney)等。CV实现的核心技术包括特征提取和跨模态对齐。特征提取是通过 CNN 或 Transformer(如ViT，将图像分割为“图像块”输入 Transformer)提取图像的语义特征(如形状、颜色、纹理)。而跨模态对齐则是将图像与文本关联(如 CLIP 模型，学习“图像-文本”共同表示空间，实现“通过文字搜图”)。

1.4.5　知识图谱

知识图谱(knowledge graph)是结构化的语义知识库，以“实体-关系-实体”三元组(如“李白-职业-唐代诗人”)存储知识，本质是“机器可理解的知识网络”，用于增强 AI 的推理能力。其核心功能是解决 AI知其然不知其所以然的问题，提供背景知识支撑(如问答系统中，通过知识图谱推理“李白是哪个朝代的诗人”)。核心技术包括实体识别(从文本中提取实体)、关系抽取(判断实体间关系)、知识融合(合并重复知识)、知识推理(如“父亲的父亲是祖父”)等。主要应用于智能问答、推荐系统(如“喜欢《三体》的人可能喜欢《球状闪电》，因同属刘慈欣作品”)、反欺诈(识别关联风险)等领域。

1.4.6　多模态学习

多模态学习(multimodal learning)处理文本、图像、音频、视频等多种类型的数据，实现跨模态的理解、关联与生成，是 AI 向“类人感知”进化的关键。主要功能是学习不同模态的“统一表示”(如文本“红色苹果”与苹果图像映射到同一向量空间)。多模态学习主要应用于跨模态检索(如“输入文字搜相似图像”“输入图像搜相关文本”)和多模态生成(如“输入图像生成描述文字”“输入文本生成图像”“输入语音生成字幕”)等。当前的代表模型包括CLIP(图文对齐)、DALL·E(文本生成图像)、GPT-4(支持文本、图像输入)。

1.4.7　AI辅助技术

除了上述6项核心技术，AI技术还包括如下辅助技术。

1. 优化算法

优化算法，比如梯度下降，用于训练模型的基石，包括随机梯度下降(SGD)、Adam等。贝叶斯优化是实现超参数调优的自动化方法。

2. 大数据与分布式计算

大数据与分布式计算包括数据预处理(含特征工程、数据增强等)，用于分布式训练：使用框架TensorFlow、PyTorch，结合GPU、TPU加速。

3. 边缘AI

边缘AI(edge AI)主要在终端设备(如手机、IoT)部署轻量级模型(如MobileNet、TinyML)。

4. 自动机器学习

自动机器学习(AutoML)实现自动化模型选择、调参(如AutoKeras、Google AutoML)。

5. 类脑计算

类脑计算主要模拟生物神经网络的硬件架构，如脉冲神经网络SNN。

6. 联邦学习

联邦学习实现分布式数据隐私保护训练(如医疗、金融场景)等。

AI 核心技术显然是一个协同体系，其中机器学习(尤其是深度学习)提供“从数据中学习”的能力，自然语言处理和计算视觉实现对语言与视觉的感知，知识图谱和强化学习增强推理与决策，多模态学习打破模态壁垒，而数据与算力是支撑这一切的基础。这些技术的融合与突破，正推动 AI 从专用智能(如人脸识别)向通用智能(AGI)逐步演进。

1.5　常用AI工具

常用的 AI 工具种类繁多，涵盖聊天对话、内容创作、代码编程等多个领域。按照功能不同，AI工具可以分为五大类。

1.5.1　聊天对话类AI

当前主流的聊天对话类AI如表1.3所示。

表1.3　聊天对话类AI

名称	开发公司或团队	特点
ChatGPT	OpenAI	通用型对话模型，支持文本生成、问答、编程辅助等，版本迭代快，如GPT-4
文心一言	百度	中文理解能力强，适合内容创作、知识问答，本土化优化好
通义千问	阿里云	支持多轮对话、文本生成，可集成到阿里云业务场景
Claude	Anthropic	强调安全性和逻辑性，适合长文本分析和复杂推理

ChatGPT是OpenAI 开发的语言模型，目前在业内处于领先地位，能回答各种类型的问题，提供广泛的知识和信息，还可用于文本生成、对话交互等多种场景。文心一言则是百度推出的语言模型，对中文的理解较为到位，可提供实用的回答，在内容创作、知识问答等方面有不错的表现。通义千问是由阿里开发的人工智能大模型，具备自然语言处理能力，能进行对话交流、文本生成、问题解答等，可应用于多种业务场景。Claude是由Anthropic 开发的一系列高性能且智能的 AI 模型，它基于 Transformer 架构，在大规模文本数据上进行预训练，具备强大的自然语言处理能力。

1.5.2　内容创作类AI

内容创作类AI工具相对较为专业，影响力较小，受众要少一些。根据创作内容的区别，内容创作类AI又可以分为文本创作类、图像创作类、音/视频创作类，如表1.4所示。

表1.4　内容创作类AI

分类	名称	特点
文本创作类	智谱清言	基于ChatGLM4，擅长文章撰写、文案生成，支持多轮对话
	Notion AI	集成在Notion中，可自动整理笔记、生成文档大纲
图像创作类	Midjourney	艺术风格突出，需要通过Discord使用，适合设计师、艺术家使用
	DALL·E 3	OpenAI出品，与ChatGPT结合，可以生成更精准的图像
	Stable Diffusion	开源模型，可本地部署，自定义性强
音/视频创作类	Runway ML	AI视频编辑工具，支持绿幕抠图、自动剪辑等
	ElevenLabs	高拟真AI语音合成，支持多语言和情感语调

智谱清言基于 ChatGLM4 模型，支持多轮对话，具有内容创作、信息归纳与总结等能力，可帮助用户撰写文章、生成文案等。Midjourney属于文生图工具，生成的图片视觉效果出色，常用于艺术创作、设计配图等领域，只需要输入文字描述，就能生成相应的图像。DALL·E3则是依据OpenAI 的图像生成模型，与 GPT 相结合后使用更加便捷，无需复杂的指令就能生成高质量的图像。

1.5.3　代码编程类AI

代码编程类AI如表1.5所示。

表1.5　代码编程类AI

名称	特点
Cursor	AI驱动代码编辑器，支持自动补全、代码解释、错误修复
GitHub Copilot	基于OpenAI，提供实时代码建议，支持多种编程语言
代码小浣熊	商汤出品，覆盖需求分析、代码生成等，支持90多种语言
Tabnine	本地化AI代码补全工具，注重隐私保护

Cursor是由 Anysphere 公司开发的由 AI 驱动的代码编辑器，专为开发者设计，可提升编程效率，受到众多公司工程师的信赖；代码小浣熊则是基于商汤大语言模型的软件智能研发助手，覆盖软件需求分析、架构设计、代码编写等环节，支持多种主流编程语言。Tabnine则是一款代码补全工具，旨在帮助开发者更高效地编写代码，它通过机器学习模型

分析上下文，预测并推荐接下来的代码片段，支持多种编程语言和集成开发环境。

1.5.4　办公辅助类AI

办公辅助类AI应该是当前最为流行的AI工具，其典型代表如表1.6所示。

表1.6　办公辅助类AI

名称	特点
豆包	字节跳动出品，支持文档处理、语音转写、日程管理
DeepSeek	通用型人工智能助手类工具
秘塔写作猫	专注中文文本纠错、润色，适合论文、公文等写作场景
办公小浣熊	自动将数据转化为有意义的分析和可视化结果
ChatPDF	上传PDF后可通过对话快速提取关键信息
Fireflies	会议录音转文字，自动生成摘要和待办事项

豆包是由字节跳动推出的多功能人工智能助手，可协助撰写各类文档、分析市场数据等，还具备文档智能处理、语音转文字、日程管理等特色功能，覆盖日常生活、学习、办公等场景。DeepSeek是一个综合性的通用人工智能模型，是由杭州深度求索公司开发的高性能人工智能大模型系列，在全球AI领域以卓越的推理能力和开源策略著称。DeepSeek可以进行对话、文本生成、深度思考、联网搜索、文件处理、代码生成与调试；其应用已渗透到教育、研究与开发、企业办公、产业融合等领域；未来的重点方向是开发能自主规划并执行多步骤复杂任务的AI智能体，采用“高性能+开源、免费”的模式，让中小企业甚至个人都能使用顶尖AI技术，推动AI普惠化。秘塔写作猫则专注于写作辅助，可进行文本纠错、语法检查、内容优化等，有助于提高写作效率和质量，适用于论文、文案等多种写作场景。办公小浣熊是基于商汤大语言模型的大模型原生数据分析产品，可通过自然语言输入需求，自动将数据转化为有意义的分析和可视化结果，能完成数据清洗、趋势分析等常见数据分析任务。

1.5.5　搜索类AI

搜索类AI是大众常用的一类工具，典型代表如表1.7所示。

表1.7　搜索类AI

名称	特点
秘塔AI搜索	无广告、结构化结果，支持学术文献检索和思维导图生成
Perplexity AI	结合大模型的搜索引擎功能，提供来源引用和简明答案
新必应	微软整合GPT-4的搜索，支持聊天式交互和多语言回答

秘塔 AI 搜索基于自研大语言模型 MetaLLM，主打无广告干扰、结构化信息展示、学术研究支持等核心功能，能自动生成思维导图、大纲等，还支持中英文文献检索。Perplexity AI 是一款问答和搜索引擎，旨在通过自然语言处理技术提供精准、可靠的答案。它结合了大型语言模型(如GPT)和实时网络搜索能力，以对话式交互帮助用户高效获取信息。新必应则是微软推出的搜索引擎，结合了AI 技术，具有聊天交互功能，可通过自然语言提问获取更精准的搜索结果，还能提供一些内容创作方面的帮助。

1.5.6 AI工具的选择

人工智能技术正以前所未有的速度重塑各行各业的发展格局，从智能制造到金融服务，从医疗健康到城市管理，各类AI工具正在成为推动产业变革的核心驱动力。随着人工智能技术的快速发展，AI工具已经渗透到各行各业的不同应用场景中。面对琳琅满目的AI工具，如何根据具体需求挑选最合适的解决方案成为关键点。

从技术架构来看，当前AI工具主要呈现三大发展路径：通用大语言模型、垂直领域专业工具和多模态融合平台。以ChatGPT为代表的通用大语言模型展现出强大的综合能力，在知识广度、推理能力和多模态支持方面表现突出；而以讯飞星火、商汤AI为代表的专业工具则在特定领域深耕细作，在教育场景和计算机视觉等领域形成了独特优势。此外，腾讯混元这样的本土化模型，在中文理解和本土化适配方面展现出显著特色。

在应用场景适配性方面，各工具呈现出明显的差异化特征。教育领域对工具的易用性、安全性和教学适配度要求较高，讯飞星火凭借其深耕教育行业多年的经验积累，在课堂教学、作业批改等场景表现优异；在创意设计领域则更注重工具的艺术表现力和创新能力，Midjourney在图像生成质量和风格多样性方面继续保持领先地位；企业级应用场景则强调数据安全、定制化能力和系统集成性，Minimax和阿里云百炼在这些方面提供了专业级的解决方案。

成本效益分析显示，当前AI工具市场已经形成了多层次的价格体系。个人用户和中小型机构可以充分利用各类工具的免费版本满足基础需求，而大型企业和专业机构则可以根据具体场景选择相应的付费服务或定制化解决方案。值得注意的是，成本不仅包括直接的使用费用，还包括部署成本、培训成本和系统集成成本等间接投入。

在技术风险评估方面，数据安全和合规性成为选型过程中的重要考量因素。海外工具在技术先进性方面具有一定优势，但在数据出境、本地化服务等方面存在一定风险；国产工具在数据合规、本地化服务方面表现更好，但在前沿技术跟进和创新能力方面仍有提升空间。此外，版权问题、技术依赖性和系统稳定性等也是需要重点关注的因素。

综合来看，AI工具的选择并非简单的好坏评判，而是需要基于具体的使用场景、技术

需求、预算限制和风险承受能力进行多维度的综合考量。理想的解决方案往往是多个工具的组合使用，通过优势互补来实现最佳的应用效果。随着技术的不断发展和市场的持续成熟，将出现更多专业化、场景化的AI工具，为不同领域的数字化转型提供更有力的支持。

1.6　数字意识

1.6.1　数字意识概述

数字意识(digital awareness)是当代数字化社会与教育领域的一个核心概念，通常指个体或群体对数字技术、数据环境及其社会影响的自觉认知与批判性理解。它不仅是技术使用能力的前提，更是智能时代公民素养的重要组成部分。下面从基本定义、核心内涵与延伸维度三方面进行系统阐述。

1. 基本定义

数字意识是指个体在数字社会与智能环境中，能够感知、理解、反思并主动参与数字技术发展的综合意识状态。它体现了人类在面对人工智能、大数据、算法推荐、虚拟现实等技术时的认知觉醒与价值判断能力。

与传统的信息素养不同，数字意识不仅关注“如何使用技术”(操作与应用层面)，更强调“理解技术意味着什么”(哲学与社会层面)。它要求个体能够超越工具理性，洞察数字技术背后的逻辑与影响——理解算法如何重塑知识结构、数据如何重构社会关系、平台如何重组权力格局，以及技术创新如何引发新的伦理挑战。

简言之，数字意识是一种面向未来的数字生存能力，它连接着技术理解力、社会判断力与伦理自觉性。其核心结构可以表示为

数字意识 = 对数字世界的理解力 + 对技术行为的反思力 + 对数据影响的判断力

(1) 理解力：认识数字技术的运行机制、逻辑原理与社会功能，理解算法、数据、智能系统等如何构成数字世界的基本秩序。

(2) 反思力：能够对自身使用技术的方式与社会的技术依赖进行批判性思考，识别潜在的偏见、操控与信息失真。

(3) 判断力：面对数据共享、隐私保护、算法决策等议题，具备价值判断能力与责任感，能够平衡效率与人本原则。

从教育视角来看，培养数字意识的目标不只是让学生掌握技术，更是让他们学会与技术共处、与算法共思、与数据共生，可以帮助个体从“被技术塑造”走向“理解并塑造技术”，从而在数字化进程中保持自主性与创造力。

数字意识的形成，是数字公民素养的核心标志，也是构建数字文明社会的思想基础。

2. 核心内涵

数字意识的核心内涵包括以下四个方面。

1) 数字认知

数字认知(digital cognition)是数字意识的基础，指个体对数字技术及其运行机制的科学理解与理性认识，包括对算法逻辑、平台机制、数据流动等核心概念的理解。例如，认识人工智能的“学习”并非真正的思考，而是基于数据模式的统计推理；理解算法推荐系统在塑造信息获取与公共舆论中的作用；意识到技术设计并非中立，而会体现开发者的价值取向与商业逻辑。

具备数字认知的人，能够超越用户视角，进入理解者视角，既能使用工具，又能思考技术背后的“黑箱机制”。这种认知帮助个体在面对复杂的数字环境时保持清醒，理解技术如何影响我们的决策、认知与社会关系。

2) 数字使用意识

数字使用意识(digital use awareness)是指个体在学习、工作与生活中，能够主动、理性、策略性地使用数字工具的意识与能力。它强调的不仅是“会用”，更是“懂用”“善用”。例如，能够根据任务目标选择合适的AI助手或数据分析平台，能够在网络协作中有效使用云端工具以提高效率；同时具备风险防范意识，懂得保护个人隐私、数字资产与网络安全。

在教育情境中，数字使用意识还意味着学习者能以批判性与创造性的方式使用AI工具，避免过度依赖或机械模仿。它是一种兼具操作技能与判断思维的能力，体现了数字时代“能动的学习者”形象。

3) 数字伦理意识

数字伦理意识(digital ethics awareness)是数字意识的灵魂部分，强调个体在技术使用中的道德反思与价值判断。它要求人们认识到，技术发展不仅关乎效率与创新，更涉及公平、公正与人类尊严。关键议题包括数据安全、隐私保护、算法偏见、信息透明度与AI可解释性等。

例如，当学生使用生成式AI进行写作时，应思考“作者身份”与“原创性”的边界；在社交媒体发布AI生成的内容时，应考虑言论对他人的潜在影响；使用AI决策系统时，应关注算法是否延续社会不平等。

数字伦理意识使个体能够在面对技术诱惑或利益冲突时，坚持人本立场，平衡技术可行性与伦理可接受性。

4) 数字公民意识

数字公民意识(digital citizenship awareness)是数字意识的社会延伸，体现了个体在数字公共空间中的责任感与参与意识。具备数字公民意识的人，能够理解自己在网络社会中同

时是信息生产者、传播者、监督者和治理者。他们不仅维护自身权利，也关注数字环境的公平与可持续性。

例如，在社交平台上理性表达，拒绝网络暴力与谣言传播；在数据社会中关注算法治理与隐私立法；积极参与公共数字议题讨论，如AI监管与数字鸿沟问题。

数字公民意识的培养是建设健康网络生态与数字文明社会的关键途径。它将“个体意识”转化为“社会行动”，使数字技术的发展更符合公共利益与人类价值。

数字认知、数字使用意识、数字伦理意识与数字公民意识构成了数字意识的四个相互支撑的维度：数字认知提供理解基础、数字使用意识落实行动实践、数字伦理意识确立价值方向、数字公民意识实现社会共建。它们共同塑造了数字时代的理性个体与负责任的公民，核心目标是培养兼具数字素养与社会责任感的公民。

3. 延伸维度

数字意识并非孤立存在的概念，它与数字素养、数据素养和AI素养密切相关，共同构成数字时代公民应具备的核心素质体系。它们之间既有区分，又存在递进与融合的关系。

1) 数字素养

数字素养(digital literacy)强调个体在数字环境中有效使用数字技术与工具的能力。它主要关注“做什么”“怎么做”，即技术操作、信息获取、网络沟通与数字创作等方面的能力。其核心目标是让个体具备在数字环境中生存和工作的能力，包括：

(1) 熟练操作各类软件与平台；

(2) 能够识别、筛选与评估在线信息的可靠性；

(3) 具备数字协作与沟通的能力。

然而，单纯的数字素养仍属于工具理性范畴，容易陷入只会用、不深思的局限。而数字意识在此基础上进一步提升，强调对数字世界的理解、反思与判断，关注技术背后的社会逻辑与伦理后果。

换言之，即数字素养回答“如何使用技术”，数字意识回答“为什么这样使用技术”。

2) 数据素养

数据素养(data literacy)是指个体理解、分析、解释并合理使用数据的能力，包括掌握数据收集、清洗、分析与可视化等基本技能，同时具备从数据中提取信息、形成洞见、支持决策的能力。

例如，能够判断数据来源是否可靠，理解统计结果的局限性，或识别算法输出中隐藏的偏差。

在数据驱动的社会中，数据素养是“与数据共生”的基本能力。而数字意识则在此基础上更进一步，强调对数据社会的整体认知与伦理反思：

(1) 数据如何塑造社会结构与个人身份；

(2) 数据被谁掌握、如何被使用；

(3) 大数据与隐私保护之间的张力。

因此，数据素养是方法论层面的理性分析力，而数字意识则是价值层面的社会洞察力。

3) AI素养

AI素养(AI literacy)是指个体对人工智能原理、应用和影响的理解能力与对人工智能的使用能力。它包括以下三个层面。

知识层面：理解AI的基本机制(如机器学习、神经网络、算法推理)。

技能层面：能使用AI工具进行辅助学习、创作或决策。

态度层面：认识AI的局限性与潜在风险。

AI素养帮助人们在智能时代实现“与机器共智”，而数字意识不仅关注“人如何用AI”，还关注“AI如何改变人”。AI素养要求学习者反思人工智能对劳动结构、教育公平、文化表达与人类身份的深远影响。

总体来说，AI素养培养“会与AI共事的人”，而数字意识塑造“能理解AI时代意义的人”。数字素养、数据素养、AI素养与数字意识之间的关系见表1.8。

表1.8　关系总结表

相关概念	关注焦点	主要特征	与数字意识的关系
数字素养	技术使用与信息处理	强调操作与应用能力	数字意识是对其进行价值反思与批判深化的结果
数据素养	数据理解与分析	强调数据处理与决策支持	数字意识扩展为对数据社会与伦理的反思
AI素养	AI的原理与应用	强调对智能技术的理解与使用	数字意识将其包含在内，关注AI的社会文化后果
数字意识	对数字世界的理解与反思	兼具认知、伦理与社会维度	数字意识是数字素养、数据素养、AI素养的整合与升华

数字意识标志着人类由“会用数字工具”迈向“理解并驾驭数字世界”的新阶段。它不仅是一种认知能力，更是一种文化态度和伦理自觉。在教育和职业训练中，强化数字意识具有深远意义。

- 培养批判性思维与创造性思维：使学习者能够在信息洪流中保持独立判断，并用创造性思维解决新型数字问题。
- 防止被技术逻辑所裹挟：帮助人们识别算法操控与信息偏见，守住人类价值底线。
- 促进数字文明建设：推动个体从技术消费者转向技术共建者，形成负责任、可持续的数字文化生态。

数字意识的教育目标是让人类不再被动适应技术，主动理解、反思与塑造技术，使数

字文明的发展更具人本精神、更能彰显社会责任。

1.6.2　数字化学习与社会责任

1. 数字化学习的时代转向

数字化学习不仅是学习方式的技术革新，更是认知模式、价值观与社会关系的全面转型。这一转型并非突然而至，而是顺应了信息社会的三次浪潮：网络化学习、数据化学习、智能化学习。在人工智能、云计算与数据驱动的学习环境中，学习者不再是被动的知识接收者，而是成为主动建构知识、创造内容，并与智能系统共生共学的主体。数字化学习因此被认为是“第四次学习革命”，它重新定义了“学什么”“怎么学”“与谁学”以及“为何而学”的教育命题。

在人工智能、云计算、区块链与5G等技术的共同作用下，数字化学习表现出四个显著特征。

1) 智能化

学习系统利用AI算法实现个性化推荐、自适应反馈与认知诊断。例如，学习分析平台可根据学生的行为数据识别学习困难点，并动态调整教学资源；ChatGPT类智能体可在学习过程中充当“认知伴侣”，提供即时指导与语言生成支持。智能化学习让“教”与“学”实现了双向智能互动。

2) 数据化

学习不再是“黑箱行为”，而是可被记录、分析、建模的过程。学习管理系统与学习分析技术能追踪学习路径、评估学习效果，并为教育决策提供依据。通过数据反馈形成的学习闭环使教育从经验驱动转向证据驱动。

3) 开放性

数字化学习突破了传统教育的时空界限，学习资源通过MOOC、OER(开放教育资源)、在线知识社区等方式实现全球共享。学习不再局限于课堂，而是延伸至社会、企业与虚拟空间，形成无边界学习生态。这种开放性不仅促进了知识民主化，也催生了新的学习文化与跨界合作模式。

4) 协同性

在数字环境中，学习成为一种人机共智与群体共创的过程。学习者与AI系统、同伴、社区共同参与知识生产，通过协作平台与创作工具(如GitHub、Notion AI、Dify、CoLab等)形成多主体共学网络。这种协同性强调“共学”而非“独学”，是社会建构主义在数字时代的延伸。

尽管数字化学习带来了前所未有的便利与创新，但也引发了新的伦理与社会风险。因

此，数字化学习不能仅停留在技术赋能层面，而应与数字意识深度结合，使学习者在掌握技术的同时具备技术反思力与伦理判断力。

数字化学习的最终目标不仅是更高效地学习，而且包括以学习促进人类的可持续发展与社会共同进步，代表了一种以学习者为中心、以人文精神为核心的教育哲学转向。正如2023年联合国教科文组织在《生成式AI与教育未来报告》中指出的：**“数字化学习的核心任务，不是让机器更聪明，而是让人类更有智慧。”**因此，教育者的使命不仅是传授数字技能，更在于引导学习者理解技术、反思技术、驾驭技术，从而在数字社会中形成自主、理性与有温度的学习主体性。

2. 数字化学习中的社会责任意识

在数字社会中，学习者不仅是知识的吸收者与创新者，更是数字公民。数字化学习不仅仅是个人的技能提升与知识积累的途径，也是一种社会实践与文化行动的形式。学习者的每一次信息检索、内容创作与数字互动，都会影响整个社会的信息生态与价值导向。因此，数字化学习的核心目标之一是培养兼具技术能力与社会责任的学习者，让学习回归“以人为本”的教育本质。

数字化学习的社会责任可从三个层面展开：数据责任、学术诚信与创作伦理、数字公民责任。

1) 数据责任

数据已成为数字社会的“新资源”，而学习数据则是教育智能化的基础。学习者在使用数字平台、AI助教、在线测评或协同工具时，会不断生成、分享与分析数据，因此，数据责任不仅是一种技术规范，更是一种伦理立场。

(1) 尊重与保护隐私。学习者需要理解“个人数据即个人权利”，避免在学习社群中随意上传、转发他人信息，增强个人隐私保护意识。例如，在提交作业或上传作品时，应关注授权范围与可见性设置。

(2) 理解算法逻辑，警惕偏见操控。应认识到AI推荐与数据分析可能存在隐性偏差。比如，学习平台的算法可能优先推送热门资源，而忽视了少数群体的学习需求。学习者应具备算法素养，学会质疑与反思推荐机制背后的逻辑。

(3) 参与数据治理与公共监督。学习者可通过教育社区反馈机制、数据伦理讨论等方式，参与学习平台的制度建设，推动教育数据使用的透明化、公平化与可追溯性。这体现了学习者从“被动使用者”到“负责任参与者”的角色转变。

2) 学术诚信与创作伦理

生成式AI的出现使知识生产与创作边界被重新定义。AI可以辅助学习、提供灵感、生成文本，也带来滥用、抄袭、责任模糊等新问题。

因此，学习者必须确立AI参与学习的伦理边界，以保障学术公平与原创精神。

(1) 明确AI的辅助定位。AI是协同伙伴，而非替代者。学习者应主动说明AI在作品创作中的参与程度，恰当引用或标注AI生成内容的来源。例如，在研究报告中注明“部分内容由AI模型辅助生成”。

(2) 坚持学术诚信与原创精神。无论是写作、编程还是设计，学习者都应以原创思考为核心。AI可以加速创作过程，但不能代替思维过程。教育者可通过“AI使用声明”制度强化这一意识。

(3) 反思“人机共创”的伦理边界。AI生成内容虽然具备形式完备性，却缺乏情感与责任感。学习者应思考：当AI能“写”，人为何仍需要“思”？这种反思正是AI时代学术伦理教育的核心要义。

3) 数字公民责任

在开放、互联的数字学习环境中，学习行为本身也是社会参与的一种形式。学习者通过发布评论、分享知识、参与在线课程等行为，正在塑造新的公共空间。

(1) 理性与负责任地表达。在社交平台与学习社区中，学习者应以事实为基础发表观点，避免传播谣言、极端言论或歧视性信息。这不仅有利于维护网络空间的健康，也体现了公民的社会担当。

(2) 关注数字鸿沟与教育公平。数字化学习的普及不能忽视弱势群体。学习者应关注不同地区、不同年龄层在技术获取与教育机会上的差距，倡导教育资源持有者和相关部门开放和完善教育资源、真正落实全民数字素养提升计划。

(3) 践行共享与协作精神。真正的数字学习不是孤立的过程，而是协同共创的过程。学习者可通过开源项目、知识共享平台(如GitHub、Wikibooks等)参与集体智慧建设，形成“共学、共创、共益”的学习共同体。

数字化学习中的社会责任是技术与人文的交汇点。它要求学习者在思考“能做什么”的同时，思考“应该做什么”“如何更好地做”。

3. 数字意识在数字化学习中的实践路径

数字意识的形成不是自然发生的，它需要在学习实践中被唤醒、体验与建构。在数字化学习环境中，教师与学生应通过一系列有意识的教学设计，将技术使用与价值反思相结合，使学习不仅提升技能，更促进思考与责任感培养。以下三条实践路径代表了数字意识在学习中的核心落点：批判性技术理解、伦理反思嵌入式学习、公共参与共创学习。

1) 批判性技术理解

批判性技术理解强调学习者不仅要会用技术，更要理解技术。它要求学习者在使用AI、算法平台或学习分析系统的过程中，主动探究其背后的运行逻辑、隐性假设与社会影响。这种理解力不仅是一种认知能力，更是一种数字理性，即对技术的意识化分析与反思性理解。

教育实践策略如下。

(1) 算法可视化体验：通过AI模拟平台(如TensorFlow Playground、Google Teachable Machine)直观理解模型训练与偏差来源。

(2) 算法观察任务：让学生追踪推荐系统(如短视频、学习平台)如何影响信息选择与兴趣形成。

(3) 批判性讨论环节：围绕“算法中立”“AI决策公平”等主题进行课堂辩论或案例分析。

学生通过此类学习，能从“工具使用者”成长为“技术观察者”，理解技术并非中立，而是社会权力与伦理结构的产物，从而形成更为成熟的数字判断力与公共意识。

2) 伦理反思嵌入式学习

伦理反思嵌入式学习强调在学习中反思技术伦理，而非课后补充。它将伦理讨论自然融入项目式学习、问题导向学习或研究性任务的全过程，使学生在“做中学”的同时在“思中行”。其核心目标是让学生理解技术选择的过程也是价值选择的过程。

教育实践策略如下。

(1) AI写作反思环节：在生成式AI写作任务中，讨论AI署名、原创性与责任归属。

(2) 数据隐私议题探讨：在教育数据分析任务中，让学生模拟学校数据治理的决策过程。

(3) 反思日志与同伴评议：学生定期撰写学习反思日志，评估自己在AI与数据使用中的责任与判断。

这种模式将伦理教育从抽象宣讲转化为情境化体验。学生在真实任务中理解“如何合理使用AI”，“什么是数据可持续性”，从而内化伦理判断力，形成数字时代的道德主体意识。

3) 公共参与共创学习

数字化学习并非仅限于个体层面，而是连接社会与文化的学习行为。公共参与共创学习倡导学习者走出课堂，进入开放的数字生态，通过参与社区项目、知识共享与公益实践，理解数字公民身份与社会责任。

教育实践策略如下。

(1) 在线共创项目：参与开源软件、维基百科编辑、公共数据可视化等项目，体验知识的社会再生产。

(2) 数字公益实践：组织AI助老、数字无障碍设计、数据科普传播等社会行动。

(3) 跨界合作社群：建立学生、教师、社区、企业协同创新小组，共同解决社会问题。

学生在参与过程中获得技术向善的体验，认识到数字学习的社会维度。他们不再是孤立的学习者，而是社会创新的共同体成员，具备以数字行动促进公共利益的意识与能力。数字化学习的未来，不仅在于掌握新技术，更在于形成技术理性与人文理性的平衡。

4) 从数字学习者到数字公民，教育目标的跃迁

数字意识的培养标志着教育从“教会学生使用技术”迈向“引导学生以技术塑造社会”的时代转型。在这一过程中，学习者逐步完成从“数字技术的使用者”到“数字社会的塑造者”的角色跃迁。这种跃迁体现了教育的价值升维：

- 从技能导向走向责任导向，强调技术理性与人文理性的平衡；
- 从个体成功走向公共福祉，以学习成果回馈社会、反哺文化；
- 从知识消费走向知识共创，以学习推动社会创新与可持续发展。

(1) 教师：数字伦理的引导者。教师不仅是知识的传递者，更是数字文化的导航者。

他们应在教学中引导学生识别AI与数据技术的社会意义，鼓励质疑、反思与创新。例如，教师可在课堂上提出“算法能否公正”“AI是否应具备情感”等开放性问题，让学习成为价值探究的过程。

(2) 学校：责任学习的生态建构者。学校应构建一个开放、包容、负责任的学习生态，既提供数字技术支持，又强化伦理审查与人文教育。

可通过设立“AI伦理实验室”“数字素养成长档案”“校园数据透明计划”等机制，推动数字教育治理的规范化与人本化。

(3) 社会：终身数字学习的支持者。社会是数字教育的延伸空间，应提供持续的学习资源与政策保障，建立终身数字学习体系，包括开放教育资源共享平台、公共AI学习空间、跨领域学习认证机制等，使公民能够不断提升数字、数据与AI素养。最终目标是形成一种以数字意识为核心的人本数字文化，让学习者不止于个人成就，能够成为推动社会进步与数字文明建设的力量。

数字化学习与社会责任是数字意识的现实体现。它提醒我们：真正的数字化学习，不仅是让人更聪明，更是让人更有责任感、更具判断力、更有温度。只有当学习者具备反思技术、守护伦理、服务社会的意识时，数字教育才能真正实现“以人为本、智用为善”的目标。

1.6.3　数字内容创新与安全防护

1. 数字内容创新的时代背景与核心内涵

在人工智能、云计算与AIGC的推动下，数字内容正经历由“人工创作”向“智能共创”的深度转型。数字内容创新不仅是技术应用的结果，更是数字意识在文化创造中的具体体现。它意味着学习者、教育者和社会公众正在共同进入一种以算法、数据与创造力为驱动的新型知识生态。

数字内容创新是指基于数字技术和智能算法，对文字、图像、视频、音频等信息进行

智能生成、整合与传播的过程。它体现了人机协同创作的思维模式，推动知识传播方式从“生产—接受”走向“共创—共享”。

数字内容创新的本质不是机器替代人类，而是拓展人类创造力的边界。在此背景下，数字内容创新的核心内涵主要体现在以下三个方面。

1) AI办公能力

这是数字创新在教育与工作中最具普及性的应用。学习者通过掌握AI办公工具，能够实现知识整理自动化、文本生成智能化、任务执行高效化。例如，在课程报告撰写过程中，学生可利用AI进行资料检索与初步生成，再通过人工编辑完善内容；在团队协作中，AI可辅助生成会议纪要、分析数据与进行任务追踪，提高整体学习效率与组织能力。AI办公能力不仅提升了生产力，更促使学习者反思人机协作的合理边界，理解自动化背后的人文与伦理价值。

2) 图像与视频媒体处理能力

随着生成式媒体工具的成熟，图像与视频已成为表达思维与传播知识的重要形式。学习者需要掌握基本的视觉编辑、AI图像生成与视频合成技能。例如，在学习项目中，学生可以通过AI视频生成技术将研究成果可视化，或以数字影像重构历史场景，从而实现知识的再呈现与意义再造。这种能力不仅涉及技术操作，更体现了视觉表达与信息审美的融合，是数字时代创新思维的重要组成部分。

3) 创新思维表达能力

在数字内容创新中，思维的创造性表达比工具的使用更为关键。AI提供了丰富的生成资源与思维激发机制，使学习者能够在算法启发下进行联想、改写与重构。例如，通过AI完成散文生成、数字故事叙事、跨模态表达(文字→音乐→图像)等任务，学生可以探索多元化的表达路径，打破传统知识边界，形成独特的“人机共创思维”。这种创新思维的核心不在于“AI替我写”，而在于“我借AI思”。它帮助学习者从被动模仿转向主动构建，学会以技术为思维工具，以表达为创造过程。

数字内容创新的核心不仅是掌握新型技术，更是以人文精神引导技术应用、以创造性实践实现认知升级。它要求学习者在AI与多媒体的支持下，既具备技术操作能力，又保持批判思维、审美意识与社会责任感，使创新真正成为数字学习的内在动力。

2. 数字内容创新的价值与教育意义

在数字化学习环境中，内容创新不仅是技术演进的结果，更是教育理念变革的重要体现。它改变了知识的生产、传播与学习方式，使学习者能够在智能系统的支持下，从被动的知识接收者转变为主动的知识建构者与文化创造者。数字内容创新的价值，不仅在于让学习更高效、更生动，更在于通过技术赋能实现学习体验、思维方式与社会责任的综合提升。

1) 数字内容创新重塑了学习体验

借助AI创作工具，如文生图、文生视频、智能写作助手等，学习者能够以视觉化、叙事化的方式表达思想，增强学习过程的沉浸感与参与度。这种智能化的表达方式，使学习不再局限于传统文本，而是扩展为多模态的认知实践，激发了学生的创造兴趣与探索欲。

2) 数字内容创新促进了跨学科整合与思维迁移

在AI与数据驱动的创作环境中，语言、艺术、技术、设计等多种学科边界被打破，学习者需要在技术逻辑与人文审美之间建立联系。这种融合式学习不仅拓展了知识应用的广度，也促进了综合素养和跨域思维的形成，为未来创新型人才的培养奠定基础。

3) 数字内容创新有助于培养学习者的创造能力与审美判断力

AI辅助创作的过程并不是让技术替代思考，而是引导学生理解算法背后的逻辑与限制，学会在与智能系统的互动中激发新的创意。通过这种人机共创的过程，学习者能够反思“创作”与“智能”的关系，从而在技术依赖与人文表达之间形成更成熟的创造意识。

4) 数字内容创新也是文化传承与再生的重要途径

数字化表达让传统文化得以以更具吸引力和传播力的形式呈现，例如通过AI生成诗画、数字文物复原、交互式非遗展示等方式，让文化教育走向开放、包容与可持续。学习者在这一过程中，不仅学习知识，更在数字环境中参与文化的再创造与传播，强化了文化认同与数字文化素养。

数字内容创新的教育意义体现在“以技术促创造、以创造促思考、以思考促价值”的教育逻辑中。它要求教育者在教学设计中，不仅关注技术的应用层面，更要引导学生理解创新背后的社会价值与伦理意义，使数字化学习成为一种兼具创造性、反思性与责任性的学习实践。

3. 数字内容创新的安全与伦理风险

一方面，AI内容生成的“低门槛”让创造更加自由；另一方面，也带来了信息失真、版权不清与数据泄露等问题。数字内容创新的高速发展伴随着多重风险，主要风险类型见表1.9。

表1.9　主要风险类型

风险类型	具体表现	潜在影响
虚假信息与深度伪造	AI可生成虚假的图片、伪造的声音与视频	误导公众、损害信任体系
版权与署名争议	AI生成内容的作者身份模糊	存在侵权风险、学术诚信受损
隐私与数据泄露	创作过程中需要上传文本或图像数据	个人信息暴露、数据滥用
算法与偏见问题	模型训练数据存在文化与性别偏见	加剧社会刻板印象与歧视
创造力依赖	过度依赖AI生成而缺乏原创思维	创造性退化、批判力下降

1.6.4 安全防护的核心策略

为了在创新与风险之间取得平衡，教育系统与学习者需要构建数字安全防护的认知与实践体系。

1. 版权保护意识

(1) 理解AI生成作品的知识产权边界；

(2) 规范引用AI生成内容的方式，注明模型与版本信息；

(3) 学习开源许可与署名规则，使用合法素材库。

2. 隐私与数据安全

(1) 避免向AI系统上传敏感数据(如身份信息、隐私图片)；

(2) 学会识别平台隐私条款、设置访问权限；

(3) 教育学生使用加密与防追踪技术，保护个人数据。

3. 信息辨识与媒介素养

(1) 通过AI真伪识别实验，培养识别虚假内容的能力；

(2) 鼓励学生查证信息源，验证AI生成内容的可靠性；

(3) 建立多平台交叉验证机制，提高批判性信息素养。

4. 伦理守则与规范教育

(1) 在教学中嵌入AI创作伦理准则；

(2) 讨论AI的价值中立性与人文导向性的问题；

(3) 引导学生在创作中体现公平、尊重与技术向善理念。

数字内容创新是未来社会的重要创造力来源，也是数字意识教育的关键实践场域。

创新的真正意义不在于“技术领先”，而在于“价值引领”。人们应该清楚地认识到：**“负责任的创造力”才是数字时代的核心竞争力**。

因此，数字内容创新与安全防护的平衡，不仅是技术问题，更是文明选择，只有当创新与安全、自由与规范相互支撑时，数字学习才能成为培养有创造力、有判断力、有伦理感公民的真正场所。

1.7 任务1：看图说话

本任务基于商汤人工智能教育平台，结合计算机视觉与自然语言处理技术，让 AI 系统通过分析图像内容，理解图像中的关键元素(如物体、人物、场景、动作、情感等)，生成符合逻辑、流畅、自然的描述文本，实现从“看图”到“说话”的精准转化。详细内容请扫码阅读。

