

人工智能通识

畅玉洁 梁维岩	主 编
尚晋旭 殷旭彪 尤彩虹 尹达恒	副主编
原坤 高静 宫燕 郝静 姚晶晶	

清华大学出版社
北 京

内 容 简 介

本书结合商业实战应用案例，针对生成式人工智能大模型技术在文本文案、数据计算、图形图像、音视频等领域的应用进行深入剖析，全方位诠释了人工智能大模型的多领域融合性应用。全书内容立足于零基础人工智能大模型学习对象，深入浅出地讲解多类案例，降低人工智能大模型技术的通识化学习成本。

本书既可作为高等职业院校生成式人工智能通识课程的教材，也可作为广大读者提升个人生成式人工智能素养的参考学习资料，亦可作为相关培训机构的教材。

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

人工智能通识 / 畅玉洁, 梁维岩主编. -- 北京 :
清华大学出版社, 2026. 1. -- ISBN 978-7-302-70416-4

I. TP18

中国国家版本馆CIP数据核字第202500VL90号

责任编辑：王 军

封面设计：郝 静

版式设计：方加青

责任校对：成凤进

责任印制：丛怀宇

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：北京同文印刷有限责任公司

经 销：全国新华书店

开 本：190mm×260mm 印 张：11.5 字 数：298千字

版 次：2026年1月第1版 印 次：2026年1月第1次印刷

定 价：49.80元

产品编号：114036-01

前言

2022 年，美国 OpenAI 公司发布 ChatGPT 平台，标志着人工智能大模型技术的重大突破，引发了全球范围内的技术革新浪潮，深刻改变了人机交互模式。此后一年内，该技术迅速渗透至文本文案生成、图形图像生成、语音音乐合成及视频动画制作四大行业领域，例如在广告创意中实现个性化内容推送、在教育内容制作中生成互动式课件、在娱乐媒体中创作虚拟角色等高效应用场景，显著提升内容生产效率，引领了新一轮互联网技术浪潮。国内互联网领军企业(如百度、科大讯飞、阿里巴巴等)相继投入人工智能大模型技术的研发与应用，催生出百度的文心一言、科大讯飞的讯飞星火、暗面科技的 Kimi、阿里巴巴的堆友、网易的天音、字节跳动的豆包和即梦、快手的可灵、金山的 WPS AI、深度求索的 DeepSeek 等众多人工智能大模型平台，覆盖文本、图像、音视频等多模态生成需求。在此竞争格局中，DeepSeek 凭借其卓越的模型性能、用户友好性和低门槛操作体验脱颖而出，成为国内顶尖的人工智能大模型技术产品，在自然语言理解和创意内容生成任务中表现优异，推动行业向更高水平发展。

本书深度融合 DeepSeek 提示词生成技术，整合文本生成、数据生成、图像生成、音视频生成等跨模态 AI 内容生成技术平台，构建综合应用体系，确保学生掌握从需求分析、创意构思到内容产出的全流程技能，包括提示词优化和内容评估方法。全书内容严格遵循互联网企业人才需求标准，聚焦商业化应用赋能，通过真实场景案例(如文案优化提升转化率、社交媒体内容创作增加用户黏性、短视频制作提升品牌影响力等)实现理论实践闭环，培养学生学以致用能力，全面提升其对行业需求的认知水平，精准对接互联网应用型人才技术前沿与企业真实项目案例，助力学生快速适应职场挑战，解决实际业务问题。

本书具备以下四大特点。

1. 内容系统性。依据大学生认知发展规律，采用循序渐进的教学设计，从基础概念(如人工智能原理)到高级应用(如多模态生成)层层递进，深度融合商业化案例(如品牌营销策略)，提升对人工智能大模型技术应用的认知，通过深入浅出的课程案例实现理论教学与实践操作的技术闭环，确保知识结构完整且易于消化，帮助学生构建扎实的知识框架。

2. 学训融合性。每节采用理论学习、知识拓展、案例实训的三维架构，确保教学效果可验证，例如通过小组讨论和项目评审强化互动反馈，通过多元化案例练习(如数据可视化任务、社交媒体内容生成)实现学习质量综合评价，提升学生解决实际问题的能力，培养批判性思维。

3. 任务导向性。实战课程内容以项目任务为驱动，如设计一个 AI 生成内容方案，通过系列微案例验证理论模型的有效性，保障内容的真实性与可靠性，帮助学生在实际操作中掌握技术精髓。

4. 产教协同性。深度融合产业端技术平台，如与 DeepSeek 合作模拟企业环境，使学生通过课程学习实现技术能力的迁移应用，凸显教材的实践价值，促进教育与企业需求的无缝对接。

本书由畅玉洁、梁维岩担任主编，尚晋旭、殷旭彪、尤彩虹、尹达恒、原坤、高静、宫燕、郝静、姚晶晶担任副主编。谨向本书涉及的人工智能大模型工具开发者与运营者致谢，其技术

成果(如开源模型和平台接口)为推进我国人工智能领域发展、提升国民人工智能大模型技术的素养与应用能力提供了重要支撑,助力教育创新,特别是 DeepSeek 团队在技术支持和案例提供方面的贡献。

本书适用于高等职业院校、应用型本科院校人工智能大模型技术的通识课程,亦可作为信息技术基础课程的升级替代教材,可满足不同层次学生的需求,包括初学者和进阶者。建议设置 16~32 学时,各院校可根据专业特色与企业需求进行模块化内容配置,如针对设计专业强化图像生成模块、针对营销专业加强文案优化模块、针对数据科学专业深化数据生成模块。

限于编者学术水平及人工智能大模型技术的快速迭代特性,书中难免存在疏漏,恳请专家、学者、师生等读者不吝指正,共同推动教材的完善,欢迎通过出版社反馈渠道提出建议。

编 者
2025 年 8 月

目 录

- 第 1 章 人工智能技术新魔法——颠覆式生成式 1
 - 1.1 人工智能概述 1
 - AI 导读：无处不在的人工智能 1
 - 1.1.1 人工智能的发展史 2
 - 1.1.2 人工智能与人工智能大模型的区别 6
 - AI 大视野：AIGC 技术的兴起给我们带来了什么影响 7
 - AI 大复盘 7
 - AI 即时练 8
 - AI 大作业 8
 - 1.2 AIGC 大模型应用场景 9
 - AI 导读：AIGC 大模型的应用 9
 - 1.2.1 文本内容生成领域 10
 - 1.2.2 图像内容生成领域 11
 - 1.2.3 音频内容生成领域 11
 - 1.2.4 视频内容生成领域 12
 - AI 大视野：AIGC 技术在文本、图像、音视频领域的卓越成就 12
 - AI 大复盘 14
 - AI 即时练 14
 - AI 大作业 15
- 第 2 章 人工智能交互金钥匙——万能的提示词 16
 - 2.1 提示词概述 16
 - AI 导读：神奇的提示词 16
 - 2.1.1 提示词的四大基本特征 17
 - 2.1.2 提示词与问题、指令、需求的关系和边界 17

- 2.1.3 判断提示词是否合格的三条标准 17
 - 2.1.4 概念延伸：提示词是一种结构化表达 18
 - AI 训练营 18
 - AI 大视野：提示词在 AIGC 领域的广泛应用 19
 - AI 大复盘 19
 - AI 即时练 20
 - AI 大作业 20
- 2.2 提示词的作用 21
 - AI 导读：让 AI “听懂”并“做对”的秘诀 21
 - 2.2.1 沟通作用：让人机对话高效、顺畅 21
 - 2.2.2 引导作用：为 AI 指定方向和重点 22
 - 2.2.3 约束作用：设置边界与规则 22
 - 2.2.4 激发作用：引导 AI 发挥创造性思维 22
 - AI 训练营 23
 - AI 大视野：行业中的多面手 24
 - AI 大复盘 25
 - AI 即时练 25
 - AI 大作业 26
- 2.3 提示词的结构 27
 - AI 导读：高质量输出的“骨架” 27
 - 2.3.1 角色设定：告诉 AI “你是谁” 27
 - 2.3.2 任务说明：明确需要 AI 做什么 28
 - 2.3.3 背景信息：给 AI 足够的“上下文” 28

2.3.4 输出要求：告诉 AI 你想要的结果	28	AI 大复盘	46
2.3.5 附加限制与提示：给 AI 加“护栏”	28	AI 即时练	47
AI 训练营	29	AI 大作业	47
AI 大视野：跨领域的延伸应用	30	3.2 文本生成的作用	48
AI 大复盘	31	AI 导读：文本生成不止“代写”	48
AI 即时练	32	3.2.1 提升信息处理与表达效率	49
AI 大作业	33	3.2.2 降低创作门槛，扩展创意边界	49
2.4 提示词的应用	33	3.2.3 实现多场景、多角色内容定制	49
AI 导读：从课堂到真实世界的跨越	33	AI 训练营	49
2.4.1 辅助学习场景	34	AI 大视野：你的智能协作伙伴	50
2.4.2 职场办公场景	34	AI 大复盘	51
2.4.3 商业创意场景	35	AI 即时练	52
2.4.4 技术支持场景	35	AI 大作业	52
2.4.5 跨界创新场景	36	3.3 文本生成的结构	53
AI 训练营	36	AI 导读：AI 文本创作的“施工图纸”	53
AI 大视野：提示词为多行业赋能	39	3.3.1 文本生成结构的概念	54
AI 大复盘	40	3.3.2 常见的文本结构类型	54
AI 即时练	40	3.3.3 如何在提示词中明确结构	54
AI 大作业	41	AI 训练营	55
第 3 章 人工智能文本生成术——文本的魔法般诞生	43	AI 大视野：结构思维的重要性	56
3.1 文本生成概述	43	AI 大复盘	57
AI 导读：让 AIGC 成为你的全能写作伙伴	43	AI 即时练	57
3.1.1 文本生成的概念	43	AI 大作业	58
3.1.2 文本生成的基本原理	44	3.4 文本生成的应用	58
3.1.3 文本生成的常见类型	44	AI 导读：文本生成应用的多领域覆盖与定制化价值	58
3.1.4 文本生成的优势	44	3.4.1 学习与学术场景的应用	59
3.1.5 文本生成的限制与注意事项	44	3.4.2 工作与职场场景的应用	59
3.1.6 获得高质量文本生成结果的方法	44	3.4.3 创意与内容创作场景的应用	60
AI 训练营	45	3.4.4 公共服务领域的应用	60
AI 大视野：AIGC 的行业应用	45	AI 训练营	61
		AI 大视野：AIGC 文本生成的多维商业应用	62
		AI 大复盘	63
		AI 即时练	63

AI 大作业	64	4.3.1 汇总统计类	76
第 4 章 人工智能数据生成术——精准与速度的融合	65	4.3.2 条件筛选类	77
4.1 电子表格数据统计与计算		4.3.3 分组对比类	77
概述	65	4.3.4 趋势分析类	78
AI 导读：万能的 AIGC 数据		AI 训练营	79
处理应用	65	AI 大视野：商业级 AIGC 数据统	
4.1.1 电子表格的数据统计与		计与决策应用	80
计算的概念	65	AI 大复盘	81
4.1.2 掌握 AI 赋能的表格统计		AI 即时练	81
与计算的必要性	65	AI 大作业	82
4.1.3 AI 赋能下的核心变化	66	第 5 章 人工智能图像生成术——高效设计	84
AI 训练营	66	5.1 图像生成的工具	84
AI 大视野：AIGC 数据处理的多		AI 导读：AI 图像生成工具	84
场景应用	67	5.1.1 Midjourney	85
AI 大复盘	68	5.1.2 Stable Diffusion	85
AI 即时练	68	5.1.3 哩布哩布 AI	86
AI 大作业	69	5.1.4 即梦 AI	89
4.2 电子表格数据统计与计算的		AI 训练营	91
作用	70	AI 大视野：图像生成技术的应用	
AI 导读：AI 助力数据统计与		拓展	93
计算	70	AI 大复盘	97
4.2.1 快速提取关键信息	70	AI 即时练	97
4.2.2 辅助科学决策	71	AI 大作业	98
4.2.3 节省时间与人力成本	71	5.2 图像生成的模型	98
4.2.4 增强表达力与说服力	71	AI 导读：AI 图像生成模型	98
4.2.5 降低技术门槛，让人人		模型分类	99
会用数据	71	AI 训练营	102
AI 训练营	71	AI 大视野：图像生成模型的	
AI 大视野：数据统计与计算在		发展	103
不同决策类型中的价值	72	AI 大复盘	103
AI 大复盘	73	AI 即时练	103
AI 即时练	74	AI 大作业	104
AI 大作业	74	5.3 图像生成的参数控制	105
4.3 电子表格数据统计与计算的常见		AI 导读：图像生成参数设置	105
类型	75	5.3.1 Steps 参数的奥秘	105
AI 导读：AI 赋能的电子表格	75		

5.3.2	CFG Scale: 文本约束的 “松紧带”	105
5.3.3	Sampler: 风格质感的 “调色盘”	106
5.3.4	Style Weight: 风格倾向的 “平衡秤”	106
5.3.5	参数联动: 哩布哩布 AI 的协同创作逻辑	106
5.3.6	哩布哩布 AI 的参数控制 特性	106
	AI 训练营	107
	AI 大视野: 图像生成参数系统的 发展	108
	AI 大复盘	109
	AI 即时练	109
	AI 大作业	110
5.4	图像生成的高阶控制	111
	AI 导读: ControlNet 神经控制网	111
5.4.1	Canny 边缘检测模型	112
5.4.2	OpenPose 姿态估计模型	112
5.4.3	Depth 深度估计模型	113
	AI 训练营	114
	AI 大视野:	114
	AI 大复盘	114
	AI 即时练	115
	AI 大作业	115
第 6 章	人工智能音频生成术——声音灵感 的创作源	117
6.1	口播生成技术的应用	117
	AI 导读: AI 口播生成工具	117
6.1.1	DeepSeek、豆包口播文案 的生成	118
6.1.2	可灵 AI 口播生成的应用	118
6.1.3	海螺 AI 口播生成的应用	119
6.1.4	即梦 AI 口播生成的应用	119
	AI 训练营	120
	AI 大视野: 口播生成技术的发展	120

	AI 大复盘	122
	AI 即时练	122
	AI 大作业	123
6.2	音乐生成技术的应用	123
	AI 导读: AI 音乐生成工具	123
6.2.1	DeepSeek 的应用	123
6.2.2	豆包的应用	124
6.2.3	可灵 AI 的应用	124
6.2.4	海螺 AI 的应用	124
	AI 训练营	125
	AI 大视野: 音乐生成技术的 发展	126
	AI 大复盘	126
	AI 即时练	126
	AI 大作业	127
第 7 章	人工智能视频生成术——动态视频 分秒生成	128
7.1	视频生成技术的分类	128
	AI 导读: AI 视频生成	128
7.1.1	文本驱动的视频生成技术	129
7.1.2	图像驱动的视频生成技术	129
7.1.3	视频驱动的视频生成技术	129
7.1.4	3D 模型驱动的视频生成 技术	129
	AI 训练营	130
	AI 大视野: 视频生成技术的 发展	131
	AI 大复盘	132
	AI 即时练	132
	AI 大作业	133
7.2	视频生成的参数控制	133
	AI 导读: AI 视频生成参数设置	133
7.2.1	时长参数	134
7.2.2	分辨率参数	134
7.2.3	帧率参数	135
7.2.4	风格提示词	135
7.2.5	运镜效果	135

AI 训练营	136	AI 大复盘	153
AI 大视野：视频生成参数控制 技术的发展	138	AI 即时练	153
AI 大复盘	139	AI 大作业	154
AI 即时练	139	8.2 深度逻辑构建：复杂程序的 AI 协同开发	155
AI 大作业	140	AI 导读：从答题游戏看 AI 协同 开发	155
7.3 视频生成的高阶应用	140	8.2.1 复杂程序逻辑的“拼图法”： 模块拆分与整合	155
AI 导读：AI 视频生成技术进阶	140	8.2.2 Python 函数与逻辑封装： 让 AI 生成更高效	155
7.3.1 影视制作与特效生成	141	8.2.3 AI 代码生成的“协同 优化”：迭代调整逻辑	156
7.3.2 虚拟主播与数字人视频 创作	141	AI 训练营	157
7.3.3 广告与营销创意视频生成	141	AI 大视野：代码生成术推动的 “全民编程”趋势	161
7.3.4 教育与培训视频制作	142	AI 大复盘	162
7.3.5 游戏开发与动态场景生成	142	AI 即时练	163
AI 训练营	142	AI 大作业	164
AI 大视野：视频生成高阶应用 未来图景	143	8.3 代码生成术的边界与进阶方向	164
AI 大复盘	143	AI 导读：AI 能完全替代人类写 程序逻辑吗	164
AI 即时练	144	8.3.1 AI 代码生成的边界： 能做与不能做	165
AI 大作业	145	8.3.2 从生成代码到优化逻辑： AI 进阶辅助	165
8.3.3 代码生成术的未来：与低 代码或无代码平台融合		AI 训练营	167
第 8 章 人工智能代码生成术——程序逻辑 的极速锻造	146	AI 大视野：“低代码 + AI” 重塑开发版图	170
8.1 代码生成术的基础认知与工具 初遇	146	AI 大复盘	172
AI 导读：以数字猜测场景为例 看编程模式的转变	146	AI 即时练	172
8.1.1 代码生成术的底层逻辑： AI 如何“学会”写程序	146	AI 大作业	173
8.1.2 Python 程序逻辑的常见“积木”： 从需求到结构	146	参考文献	174
8.1.3 主流 AIGC 代码生成工具 全景：能力与场景适配	147		
AI 训练营	148		
AI 大视野：代码生成术在行业的应 用——从简单工具到复杂系统	153		

第1章 人工智能技术新魔法

——颠覆式生成术

1.1 人工智能概述

AI 导读：无处不在的人工智能

人工智能在生活、生产中的应用场景不胜枚举，如图 1.1 所示。自 2010 年全球进入移动互联网时代以来，依托大数据技术形成了大数据知识架构体系，使得人工智能技术越来越成熟。人工智能可以实现自动化数据挖掘，从而快速处理大量的数据信息，精准地为不同用户定制个性化的方案。



(a) 智能语音助手



(b) 智能全屋定制



(c) 智能车载系统



(d) 智能工厂



(e) 智能餐厅

图 1.1 人工智能的应用场景示例

2022 年，基于人工智能的技术不断完善并得到广泛应用，一门新型的人工智能技术——人工智能大模型，瞬间席卷了全球各个行业，人工智能的话题再次引发了广泛关注。

1.1.1 人工智能的发展史

人工智能，从 1943 年提出至今，其发展一共经历了七个阶段。

1. 阶段一：理论奠基期(1943—1955 年)

这一阶段又被称为智能机器的思想萌芽期。人工智能的思想根源可追溯至 20 世纪中叶的数个跨领域学科的突破，这些开创性工作为人工智能学科的正式建立奠定了理论基础。在这一阶段，数学家、神经科学家和计算机先驱者开始探索机器模拟人类智能的可能性。

(1) 人工神经元模型：1943 年，美国神经科学家沃伦·麦卡洛克(Warren McCulloch)和数理逻辑学家沃尔特·皮茨(Walter Pitts)合作发表论文《神经活动中内在思想的逻辑演算》，提出了首个形式化神经元数学模型(M-P 模型)。该模型证明神经网络理论上可以执行任何逻辑运算，为后来连接主义学派的发展奠定了数学基础。

(2) 图灵测试：1950 年，英国数学家艾伦·图灵(Alan Turing)在其具有划时代意义的论文《计算机与智能》中提出著名设问——机器能否思考，并设计了“模仿游戏”(后称“图灵测试”)作为判断机器智能的操作性标准。他预言到 2000 年，计算机有望通过该测试，这一思想实验至今仍是人工智能领域的核心哲学命题。

(3) 控制论与早期学习机器：1951 年，马文·明斯基(Marvin Lee Minsky)开发了世界上第一台神经网络学习机 SNARC(Stochastic Neural Analog Reinforcement Calculator)，使用 3000 个真空管模拟 40 个神经网络，实现了老鼠走迷宫行为的模拟。与此同时，克劳德·香农(Claude Shannon)提出信息论框架，为智能系统的信息处理提供了数学工具。

这一阶段虽未诞生“人工智能”术语，但跨学科探索已为即将到来的人工智能革命埋下种子。这些先驱者将人类的认知过程抽象为可计算模型，挑战了“思维仅为人类专属”的传统观念，为后续研究开辟了道路。

2. 阶段二：诞生与黄金期(1956—1973 年)

1956 年夏天，在达特茅斯学院举行的为期两个月的学术研讨会被公认为人工智能诞生的标志性事件，主要参会人员如图 1.2 所示。会议由约翰·麦卡锡(John McCarthy)、马文·明斯基、克劳德·香农等学者组织，首次明确提出“人工智能”(Artificial Intelligence)术语，并设定了核心目标：让机器模拟人类智能的所有方面。



图 1.2 达特茅斯人工智能会议主要参会人员

1) 符号主义的兴起

- 逻辑理论家与问题求解：1956 年，艾伦·纽厄尔(Allen Newell)与赫伯特·西蒙(Herbert Simon)开发的“逻辑理论家”(Logic Theorist)程序成功证明了《数学原理》提及的 52 个定理中的 38 个，甚至给出了比原书更精巧的证明。这被视为第一个真正意义上的人工智能程序，标志着符号主义(Symbolism)范式的兴起。
- LISP 语言奠基：1958 年，约翰·麦卡锡发明 LISP 编程语言，成为主导 AI 研究数十年的核心工具。LISP 编程语言的符号处理能力特别适合进行逻辑推理和递归操作，推动了专家系统、自动定理证明等领域的发展。
- 早期自然语言处理：1964—1966 年，约瑟夫·维森鲍姆(Joseph Weizenbaum)开发出 ELIZA 聊天程序，通过模式匹配模拟心理治疗师对话。虽然技术简单，但部分用户误以为其具有真实理解能力，引发对“拟人化错觉”的早期伦理讨论。

2) 连接主义的初步尝试

1957 年，弗兰克·罗森布拉特(Frank Rosenblatt)发明感知机(Perceptron)，这是首个可通过权重调整学习分类模式的神经网络模型。1969 年，马文·明斯基和西摩·佩珀特(Seymour Papert)在《感知机》一书中证明该模型无法解决异或(XOR)等非线性问题，导致神经网络研究陷入停滞。

3) 里程碑：专家系统

1965 年，爱德华·费根鲍姆(Edward Feigenbaum)开发 DENDRAL 系统，可依据质谱数据推断分子结构，成为专家系统原型。20 世纪 70 年代，MYCIN 医疗诊断系统能识别细菌感染并推荐抗生素，准确率达 69%，超越初级医生水平。

3. 阶段三：寒冬与反思期(1974—1980 年)

20 世纪 70 年代中期，AI 技术的发展首次遭遇重大挫折。《莱特希尔报告》(1973 年)批评 AI 未能解决现实世界的复杂问题，受该报告影响，英国政府大幅削减 AI 相关研究经费，美国 DARPA 随后也减少相关资助，导致这一阶段被称为“AI 寒冬”。

1) 符号主义的统治

- 知识表示革命：爱德华·费根鲍姆提出“知识原则”，即智能行为主要取决于处理知识而非形式推理，推动了专家系统的发展，其开发的 PROSPECTOR 系统在 1982 年成功预测华盛顿州钼矿沉积，为所有者带来 1 亿美元的收益。
- 框架理论创新：1974 年，马文·明斯基提出“框架理论”，认为人类通过预存的知识结构来理解新情境。该理论为知识表示提供新工具，成为后来面向对象编程的灵感来源。

2) 连接主义的蛰伏

- 1974 年，保罗·韦伯斯(Paul Werbos)在哈佛大学博士论文中首次提出反向传播算法(Backpropagation)的概念，但受限于算力和学术关注度，直到 1986 年才被杰弗里·辛顿(Geoffrey Hinton)等人独立重新发现并广泛应用。

这一阶段，AI 技术的发展看似停滞，实则孕育着关键突破。研究者意识到早期对“通用智能”的追求过于乐观，转而聚焦特定领域问题求解，为专家系统在 20 世纪 80 年代的商业繁荣奠定了基础。寒冬中的反思促使 AI 研究从“蛮力搜索”向“知识工程”转型，重塑了发展路径。

4. 阶段四：复兴与连接主义崛起(1981—1997 年)

20 世纪 80 年代，随着专家系统商业化的成功和神经网络理论的突破，AI 研究迎来复兴。日本的“第五代计算机计划”(1982 年)刺激欧美国家重启大规模投资，而摩尔定律带来的计算能力提升则为复杂模型提供了物质基础。

1) 符号主义的技术延伸

1980 年，卡内基梅隆大学为 DEC 公司开发 XCON 系统，每年为公司节约 4000 万美元。1985 年，专家系统的全球市场规模达 10 亿美元，广泛应用于医疗诊断、地质勘探等领域，促进了专家系统的产业化。

2) 连接主义的突破性进展

- 霍普菲尔德网络：1982 年，约翰·霍普菲尔德(John Hopfield)提出循环神经网络架构，引入能量函数的概念，使网络具有联想记忆能力，为神经网络理论提供新框架。
- 反向传播革命：1986 年，大卫·鲁梅尔哈特(David Rumelhart)、杰弗里·辛顿等人独立重新发现并完善反向传播算法，成功训练多层神经网络解决异或问题。这一突破标志着连接主义正式复兴。
- 卷积神经网络雏形：1989 年，杨立昆(Yann LeCun)在贝尔实验室开发 LeNet，首次将卷积神经网络用于手写邮政编码的识别，准确率超 99%，成为后来深度学习革命的先导。

3) 里程碑：机器超越人类

1997 年，IBM 深蓝(Deep Blue)超级计算机以 3.5:2.5 的比分击败世界冠军加里·卡斯帕罗夫。该系统通过专用硬件实现每秒 2 亿步棋局搜索，展示暴力计算与人类直觉结合的威力，成为 AI 技术发展史上的分水岭事件。

5. 阶段五：新千年突破(1998—2011 年)

互联网的普及催生了海量数据集，而 GPU 等硬件的进步提供了前所未有的算力支持。AI 技术的研究重心从规则驱动转向数据驱动，统计学习方法逐渐成为主流。

1) 连接主义的实战验证

- 统计学习浪潮：支持向量机(SVM)、隐马尔可夫模型(HMM)等技术在语音识别(如 Dragon Systems)、垃圾邮件过滤等领域广泛应用，奠定机器学习产业化基础。
- ImageNet 革命前夜：2006 年，杰弗里·辛顿提出“深度学习”的概念，并通过受限玻尔兹曼机(RBM)解决深层网络训练难题。同年，李飞飞启动 ImageNet 项目，构建包含 1400 万张标注图像的大规模数据集，为计算机视觉突破提供“燃料”。

2) 自然语言处理进化

2003 年，约书亚·本吉奥(Yoshua Bengio)等提出神经语言模型，使用词嵌入表示语义，替代传统的 n-gram 方法。谷歌公司于 2007 年将神经网络语言模型应用于机器翻译，显著提升流畅度。

3) 里程碑：人机知识竞技

2011 年，IBM 沃森(Watson)系统在美国智力竞赛节目《危险边缘》中击败两位人类冠军。该系统整合自然语言理解、知识检索和概率决策等功能，展示 AI 处理开放领域问题的潜力，推动认知计算研究。

这一阶段的关键转变是产业界取代学术界成为 AI 创新引擎。谷歌、微软等公司投入巨资建

立 AI 实验室，将研究成果快速产品化(如谷歌翻译、语音助手)。开源框架(如 Torch、Theano)降低了研究门槛，全球协作生态初具雏形，为深度学习革命铺平道路。

6. 阶段六：深度学习革命(2012—2020 年)

2012 年成为深度学习的“觉醒之年”，以 AlexNet 在 ImageNet 竞赛中表现的性能碾压传统方法为标志。GPU 集群和大数据的结合释放了神经网络的潜力，引发 AI 技术的指数级进步。

1) 技术突破三连击

- 计算机视觉超越人类：2012 年，亚历克斯·克里热夫斯基设计的 AlexNet 将 Image Net 图像识别的错误率从 26%降至 15%，开创卷积神经网络新时代。2015 年，ResNet 在 ImageNet 上将错误率降至 3.57%，超越人类水平(约 5%)。
- 生成对抗网络：2014 年，伊恩·古德费洛提出生成对抗网络(Generative Adversarial Network, GAN)，通过生成器与判别器对抗学习，使 AI 能创造逼真的图像、音乐和文本。StyleGAN(2019 年)生成的人脸几乎无法被肉眼识别真伪。
- 自注意力机制与 Transformer：2017 年，谷歌发布 Transformer 架构，彻底取代循环神经网络，成为自然语言处理新标准。其自注意力机制(Self-Attention)可并行处理序列数据，大幅提升训练效率，奠定大语言模型的基础。

2) 大语言模型爆发

- GPT 系列进化：2018 年，OpenAI 发布 GPT-1，首次展示无监督预训练和有监督微调的威力。2020 年，拥有 1750 亿参数的 GPT-3 展示少样本学习能力，仅需少量示例即可完成新任务，但“幻觉”(编造事实)问题凸显。
- BERT 理解上下文：2018 年，谷歌推出 BERT(Bidirectional Encoder Representations from Transformer)，通过双向上下文理解大幅提升搜索和回答问题的准确性，成为工业界标杆。

3) 多智能体博弈取得突破：AlphaGo 的历史性胜利

2016 年，DeepMind 的 AlphaGo 以 4:1 击败围棋世界冠军李世石。其结合蒙特卡洛树搜索与深度神经网络，攻克了围棋布局 10^{360} 种可能性的复杂问题。2017 年，AlphaZero 不需要人类棋谱，仅通过自我对弈即可掌握围棋、国际象棋和将棋。

这一阶段，AI 从“专用弱人工智能”向“通用强人工智能”迈出关键一步。Transformer 统一了处理多模态数据的框架，而模型规模化则揭示性能随参数和数据量增长的规律，催生大模型军备竞赛。

7. 阶段七：多模态融合(2021—2025 年)

当前阶段以技术融合与社会嵌入为特征。神经符号 AI、多模态大模型、具身智能等方向的技术不断突破，同时各国加速 AI 治理框架构建，确保技术创新与社会价值对齐。

1) 技术突破：感知与推理的统一

- 多模态大模型普及：OpenAI 的 GPT-5(2025 年)与谷歌的 Gemini 2.5 实现文本、图像、音频的联合理解与生成。DeepSeek-V3(2024 年)支持用户上传文档、图表进行跨模态分析，推动 AI 成为通用信息处理平台。
- 神经符号 AI 医疗落地：2025 年初，FDA 批准 MedBrain 5.0 系统，将神经符号 AI 应用于癌症诊断，其双轨架构中，神经网络解析医学影像像素特征，符号引擎执行临床指

南逻辑推理，误诊率降至 0.3%，且可生成符合循证医学的决策证据链。

- 视频理解效率革命：2025 年 7 月，英伟达与 MIT 联合推出 Long-RL 技术，通过强化学习优化长视频训练。其分层记忆网络可避免信息遗忘，在 ActivityNet 数据集上的训练速度提升 200%，为影视分析与自动驾驶提供新工具。

2) 人工智能大模型与产业融合

- 自回归模型跨界视觉：快手与清华大学在 ICML 2025 上提出 VARSR 模型，将语言模型的 next-token 预测创新为 next-scale 预测，实现图像超分重建。在保留结构的同时提升分辨率，推理速度较扩散模型提升 10 倍，延时仅 0.59 秒。
- “人工智能+制造”国家战略：中国工业和信息化部于 2025 年 7 月发布政策，推动通用大模型与行业大模型在制造业落地，重点部署智能工厂、工业大模型及行业智能体，通过知识蒸馏技术将大模型的能力迁移至边缘设备。

当前，AI 领域的发展呈现三大趋势：技术融合化(神经+符号、AI+量子计算)、治理常态化(欧盟 AI 法案、《生成式人工智能服务管理暂行办法》)，以及工具民主化(DeepSeek-R1 开源模型降低了大模型使用门槛)。据 Gartner 2025 年预测，采用 AI 工程化实践的企业，其 AI 项目成功率将提升 50%，标志着技术步入成熟应用期。

1.1.2 人工智能与人工智能大模型的区别

1. 概念不同

人工智能大模型属于人工智能的子集，两者的应用范围大有不同。从广义维度考量，人工智能，即 AI，主要是指模拟人类智能的跨领域技术体系，核心在于对面向对象的认知，表现在理解与决策层面，一般包含感知、推理、决策、行动的全链路。

例如自动驾驶技术，通过车体所搭载的自动驾驶感知模块，结合车体的监控单元，即可自动判定汽车的位置、距离障碍物的远近，预估危险的临界值与行程剩余里程，进而综合判定自动驾驶的速度、占道、转向等信息，在极短的时间内做出驾驶决策，实现自动驾驶。

作为人工智能的自动化生成这一子集，人工智能大模型则更关注内容的自动化创造，并不参与决策，主要以文本、图像与音视频的内容生成作为技术表象，如 ChatGPT 技术、Stable Diffusion 技术等。

2. 技术原理不同

传统意义上的人工智能是基于判别模型实现智能自动化操作的，其目标是学习从输入数据到输出标签的映射关系，代表算法有 SVM、CNN 分类器、BERT 文本分类等。

引用智能驾驶的案例，在可能出现障碍物的场景中，智能驾驶系统会设定车身与障碍物之间的安全距离阈值，这一阈值范围会被人工智能的控制单元捕获并学习，这一过程属于学习输入数据。通过对这一数据的学习，在特定的障碍环境下，车体通过摄像、雷达等控制单元检测到危险存在后，由人工智能判定是否需要改变驾驶方式，这一过程就属于输出标签。只有输入与输出达到零错误的映射，智能驾驶的安全性才有望得到大幅提升。

主流的人工智能大模型是基于生成模型的，其目标是学习数据分布 $P(X)$ 以生成相似的新样本。它的实现路径则是通过变分自编码器、生成对抗网络、扩散模型(如 Diffusion)、自回归模型(如 GPT)等系列模型自动学习并归纳出同类数据的统一特性，进而实现扩散式衍生。

3. 应用领域不同

人工智能更多地应用于工业制造与产品制造领域，结合人工智能的自动化数据捕获与学习，能够使产品的应用紧跟时代与用户所需，可以在一定程度上降低硬件产品频繁更迭的成本。

人工智能大模型更多地应用于资料生产领域，其最重要的应用方向之一是人工智能生成内容(Artificial Intelligence Generated Content, AIGC)，结合人工智能的自动化数据捕获与关联学习，能够在文本内容、图像内容、音视频内容等相关领域快速给予用户完整、可行的解决方案，提升工作效率。



AI 大视野：AIGC 技术的兴起给我们带来了什么影响

人工智能大模型技术的浪潮正以不可阻挡之势席卷全球，深度融入我们的生活与工作，引发了一场静默却深刻的效率与范式革命。

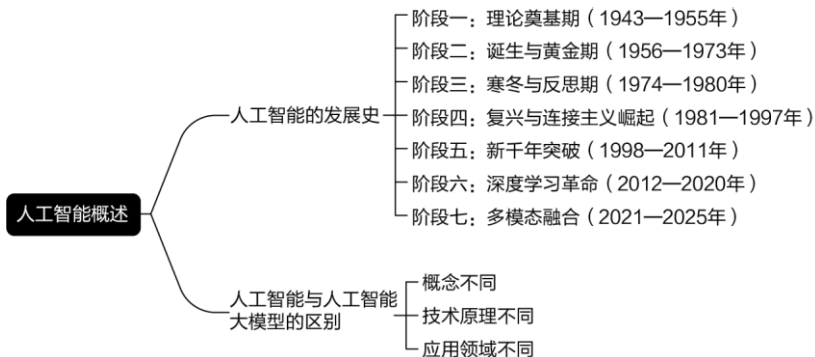
在生活场景中，AIGC 正成为普惠的创造力引擎与个性化助手。以往被视为专业壁垒的创意内容生产，如今在 AI 的赋能下，门槛大幅降低。普通人借助提示词，即可轻松生成媲美专业水准的图像、视频片段甚至音乐旋律，创作不再是少数人的特权。

在工作场景中，AIGC 的价值则集中体现在生产力的巨大解放与流程的智能化重构上。AIGC 正成为各行各业的“超级协作者”：营销人员可借助工具批量生成适配不同平台与受众的广告文案初稿；程序员可利用代码生成与补全功能显著提升开发效率，将精力集中于复杂架构设计；设计师可利用 AI 快速生成海量设计元素与原型，加速创意落地。麦肯锡报告指出，知识工作者超过 60% 的任务活动存在自动化潜力，而 AIGC 正是关键推手。它不仅替代重复劳动，而且更深层地驱动工作模式的进化——从线性执行转向人机协同的创意激发与决策优化。企业开始重新设计流程，以嵌入式 AIGC 作为核心生产力要素。

展望未来，AIGC 终将从炫目的工具进化为人类能力的自然延伸。这场变革的本质并非机器取代人类，而是人类智慧与机器算力在协作中共同跃升的新起点。我们将立于 AIGC 掀起的智能风暴的中心，锚定方向，驶向一个创造力被极大激发、人机共生共赢的未来。



AI 大复盘





AI 即时练

1. 人工智能诞生的标志性事件是()。
A. 图灵测试的提出(1950 年) B. 达特茅斯会议(1956 年)
C. 感知机的发明(1957 年) D. 深蓝击败卡斯帕罗夫(1997 年)
2. ()技术解决了神经网络无法处理非线性问题(如异或问题)的难题。
A. 卷积神经网络 B. 反向传播算法(1986 年)
C. 支持向量机 D. 霍普菲尔德网络(1982 年)
3. 2012 年深度学习“觉醒之年”的关键事件是()。
A. AlphaGo 击败李世石
B. GPT-3 发布
C. AlexNet 在 ImageNet 竞赛中取得性能突破
D. Transformer 架构提出
4. 人工智能大模型的核心技术目标是()。
A. 学习输入数据到输出标签的映射 B. 模拟人类决策与行动
C. 学习数据分布以生成新样本 D. 优化特定领域的规则推理
5. 以下各项中, ()属于生成模型。
A. SVM 分类器 B. BERT 文本分类
C. 生成对抗网络 D. 卷积神经网络
6. Stable Diffusion 图像生成平台基于的核心模型是()。
A. 变分自编码器 B. 自回归模型
C. 潜在扩散模型 D. 卷积神经网络
7. 2025 年“人工智能+制造”国家战略中提到的关键技术是()。
A. 神经符号 AI 医疗 B. 知识蒸馏迁移至边缘设备
C. 多模态大模型 D. 强化学习优化视频训练
8. 人工智能的应用领域包括()。
A. 图像生成 B. 自动驾驶 C. 音乐生成 D. 视频生成
9. 传统人工智能与人工智能大模型的核心区别在于()。
A. 是否依赖大数据 B. 是否参与决策层
C. 是否使用神经网络 D. 是否需要硬件支持
10. ()标志着 AI 在开放领域问题处理技术上的突破。
A. 深蓝击败国际象棋冠军(1997 年) B. 沃森在《危险边缘》(2011 年)中获胜
C. AlphaGo 战胜李世石(2016 年) D. GPT-3 发布(2020 年)



AI 大作业

项目一：文本内容生成初识

使用 DeepSeek 文本生成平台, 结合日常的作业、校园活动等, 尝试输出完整的作业或活动方案。

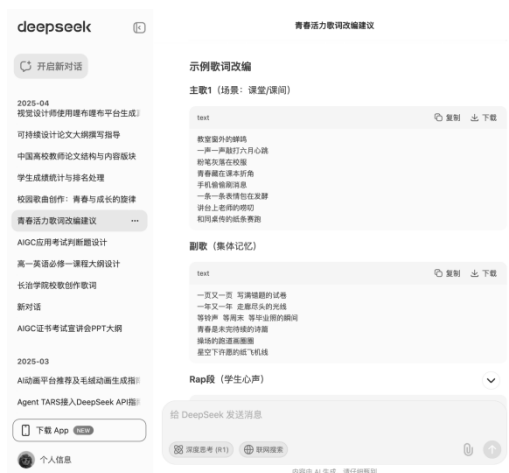
项目二：高考作文内容生成

使用 DeepSeek 文本生成平台，结合本年度的高考语文全国卷，进行高考作文的内容生成输出，形成完整的高考作文。

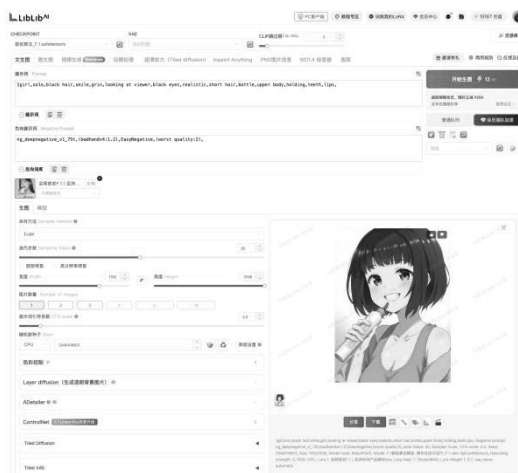
1.2 AIGC 大模型应用场景

AI 导读：AIGC 大模型的分类应用

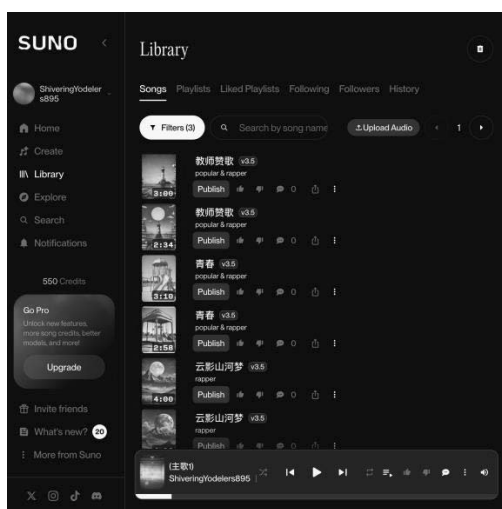
现如今，AIGC 的应用已覆盖文本写作、创意图像、音频创作与视频动画等领域(见图 1.3)，触及生活与工作的方方面面，一场新时代的技术革命浪潮正奔涌而来。



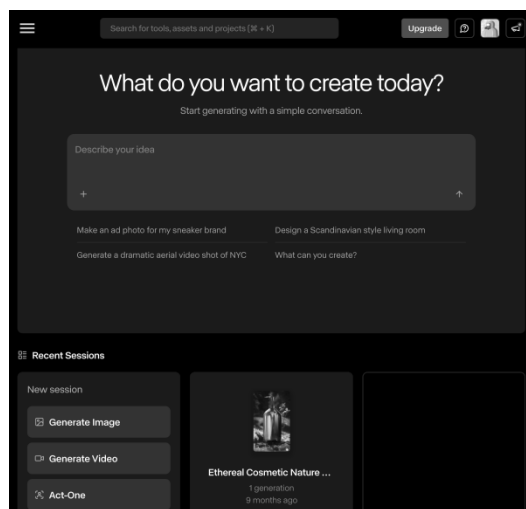
(a) AIGC 文本大模型应用



(b) AIGC 图像大模型应用



(c) AIGC 音频大模型应用



(d) AIGC 视频大模型应用

图 1.3 AIGC 大模型的应用分类

自 2022 年掀起全球热潮以来，人工智能大模型的影响力延续至今，在文本内容、图像内容与音视频内容的创作、生成方面崭露头角，大幅提升了各类白领人群的工作效率。

1.2.1 文本内容生成领域

在文本内容生成领域，AIGC 能结合不同的行业应用落地各式各样的商用方案，如图 1.4 所示。



图 1.4 AIGC 在不同行业的文本化应用

2022 年，美国 OpenAI 公司依托 GPT 人工智能大模型技术，通过在线对话的方式唤起 AI 机器人。用户提出内容写作需求后，便可轻松获取完整的写作方案，由此开启了 ChatGPT 技术的元年。

之后，国内外一众互联网科技公司纷纷加入 GPT 工具开发的行列，国外以谷歌公司的 Gemini、Anthropic 公司的 Claud 和微软公司的 Copilot 为代表，率先推动 GPT 技术不断向各大平台与各类行业延伸。在国内，百度公司的文心一言、阿里巴巴的通义千问与月之暗面的 Kimi 率先发力，打开了国内的 AIGC 市场，而 2025 年之后，均转向深度求索公司的 DeepSeek 技术(见图 1.5)。



图 1.5 DeepSeek 平台

至此，AIGC 便成为各大行业与各类岗位不可缺少的效率化文本生成技术，其应用不仅局限于生成文案、创意歌词、公文材料，以及开发代码、证明数学公式、求解高阶函数等领域，只

要是能够用文字表达的行业方案，就可以看见 AIGC 的身影。

1.2.2 图像内容生成领域

以 Latent Diffusion 潜在扩散模型为主要生成模型，德国慕尼黑大学 CompVis 小组开发了图像生成技术，2022 年，随着 Stability AI 的加入，Stable Diffusion 图像生成平台(见图 1.6)问世，自此成为图像设计领域的高频生成工具。

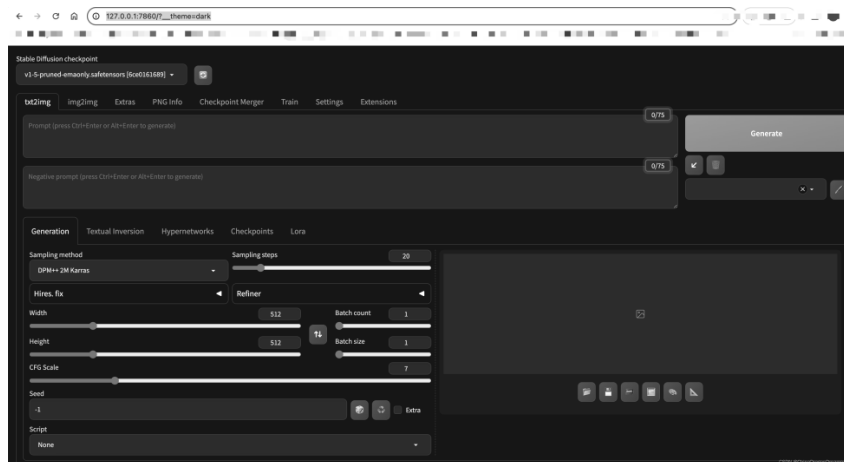


图 1.6 Stable Diffusion 图像生成平台

1.2.3 音频内容生成领域

国外的 Suno 是行业中应用最广泛的音频内容生成平台之一。2022 年，四位机器学习专家 Mikey Shulman、Georg Kucsko、Martin Camacho 和 Keenan Freyberg 在麻省剑桥市创立 Suno(见图 1.7)。Suno 平台于 2023 年 12 月正式上线，意味着音乐生成成为现实，自此推动了一系列以音乐生成技术为核心的国内外生成平台陆续上线，使得音乐、歌曲、口播等内容都可以通过 AIGC 技术降低创意成本，实现高效落地。

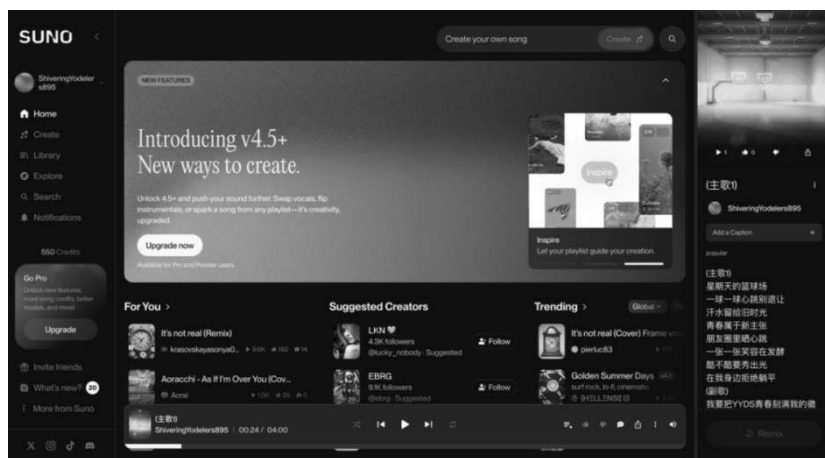


图 1.7 Suno 音乐生成平台

1.2.4 视频内容生成领域

最先在全球领域发起 AIGC 视频创作，并且拥有大量用户访问量的平台是 Runway(见图 1.8)，自此实现了包含文本和图像的生成，再结合音频生成技术生成的音乐，整合成系列化、可联动的视频内容，完成了 AIGC 内容生成领域的闭环。

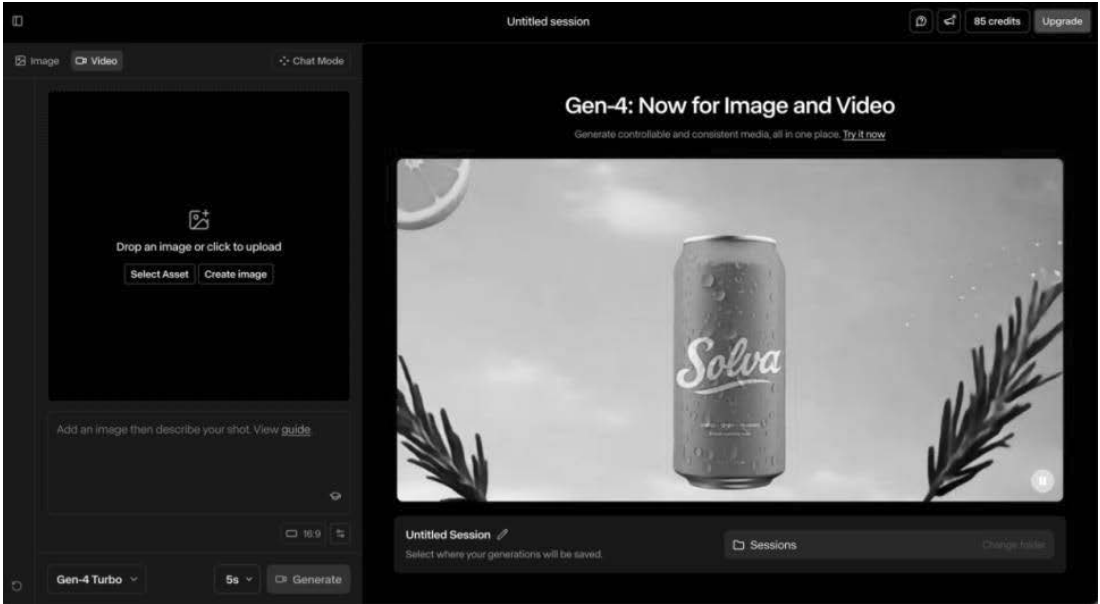


图 1.8 Runway 视频生成平台

 AI 大视野：AIGC 技术在文本、图像、音视频领域的卓越成就

2025 年，人工智能大模型已走出实验室，正式跻身产业核心领域，成为文本、图像、音视频领域颠覆性创新的引擎，其成就不仅体现为技术参数的突破，更在于解决社会痛点、激活文化传承、重构商业逻辑。国内外案例共同证明：AIGC 正在重新定义创造力。

1. 文本生成：从语言壁垒到逻辑巅峰

1) 跨语言实时交互：拯救生命的“翻译官”

2025 年缅甸地震救援中，北京语言大学基于 DeepSeek 大模型开发的中缅英三语互译系统，仅用 7 小时便完成部署。系统突破语言障碍，支持救援指令实时转译与跨文化沟通，服务超 700 名救援人员，首次实现大模型在国际人道行动中的高效落地。这一案例印证了 AIGC 在应急语言服务中的不可替代性——它将语言屏障转化为生命通道。

2) 长文本推理：为 AI 装上“节能大脑”

2025 年 7 月，DeepSeek 团队凭借论文“Native Sparse Attention”斩获 ACL 最佳论文奖。其提出的“原生稀疏注意力”机制，让大模型像人类阅读般“选择性聚焦”，在处理 64KB 长文本时，所消耗的计算资源降低 30%，推理速度提升最高达 11.6 倍。这项突破被业内称为“大模型的节能引擎”，为长文档分析、法律合同审查等场景提供了新范式。

3) 数学逻辑推理：通用 AI 的“金牌时刻”

2025 年国际数学奥林匹克竞赛(IMO)见证历史：OpenAI 未公开实验模型与 Google Gemini Deep Think 双双达到金牌分数线(35 分以上)。前者仅用 4.5 小时便独立完成 6 道证明题，并在面对超难组合数学题时展现出“自知无法解决”的元认知能力。通用 AI 首次在顶级智力竞技中媲美人类顶尖选手，标志着推理能力从专用工具到通用思维的重大跨越。

2. 图像生成与处理：高精度修复与安全防护

1) 文物数字化：攻克“不可能三角”

静安科技企业合合信息推出无限扫描技术，解决了大尺寸图像拼接中的高清晰度、高完整性、高色彩还原度“不可能三角”。以南京博物院镇馆之宝《坤舆万国全图》为例：通过手机拍摄视频，利用 AI 自动消除畸变与色差，实现宽幅巨图的无痕拼接，将历史真迹“装进手机”。该技术已应用于敦煌壁画等脆弱文物的数字化，为文化遗产打造“掌上博物馆”。

2) AI 鉴伪：毫秒级狙击“数字假面”

面对 AI 换脸技术泛滥的现状，合合信息研发的毫秒级人脸鉴伪模型成为安全盾牌。该模型通过像素级真伪标签分析，在人眼无法辨识的伪造视频中实现 90%以上的检测准确率，并对压缩、裁剪等攻击具有强鲁棒性。在金融远程开户、大额交易验证中，这项技术正成为防范身份诈骗的核心基础工具。

3. 音视频生成：影视工业化与实时交互革命

1) 渲染效率：15 秒的“技术奇点”

2025 年 8 月，马斯克宣布 Grok Imagine 视频渲染时间从 60 秒缩短至 15 秒(10 天内效率提升 75%)，并计划 3~6 个月实现实时渲染。用户输入文字即可生成带音效的视频，引发社交媒体创作狂潮。技术跃进的背后是算法与算力的协同进化——液冷系统、800G 光模块等硬件升级支撑了十万级 GPU 集群的高效运行。

2) 中国文化表达：超越 Sora 的本土化叙事

2025 年 7 月，SuperCLUE 发布的视频榜单中，万兴天幕 2.0 超越 OpenAI Sora，尤其在专业运镜、立体音效与中国文化元素表达上领先。万兴科技与华为云共建 AI 视频大模型实验室，助力《新世界加载中》等 AI 单元剧的影视级画面生成，制作成本仅为传统制作成本的 1/10。技术本土化证明：AIGC 的文化转译能力决定全球竞争力。

3) 文旅 IP 活化：动画夺得国际奖杯

吉林市融媒体中心出品的《松花江白鱼故事》获中欧电影节“中欧光影传播奖”。作品采用 AI 视频技术动态复现“开江鱼宴”非遗场景，通过军帐炊烟、宫廷珍馐等三重历史时空的沉浸式叙事，将松花江渔猎文明转化为能引起全球共鸣的生态哲学。这一案例凸显 AIGC 在传统文化国际化传播中的桥梁作用。

从吉林动画斩获国际奖项到马斯克刷新渲染速度，从 DeepSeek 的“节能引擎”到 OpenAI 的数学“金牌时刻”，AIGC 的成就不仅是技术的胜利，更是人类借助机器拓展认知边界的证明。当吉林农民用 AI 销售茶叶、数学家与 AI 共解千年难题时，我们看到的不是机器对人类的替代，而是一种新的共生智慧——AIGC 正在成为文明传承的“数字基因”与创新裂变的“原子催化剂”。



AI 大复盘

AIGC大模型应用场景

- 文本内容生成领域
- 图像内容生成领域
- 音频内容生成领域
- 视频内容生成领域



AI 即时练

1. Stable Diffusion 图像生成平台的核心技术基于()。
A. 变分自编码器 B. 自回归模型
C. 潜在扩散模型 D. 卷积神经网络
2. 音频生成领域应用最广泛的平台是()。
A. Runway B. Suno
C. DeepSeek D. Stable Diffusion
3. Runway 平台的主要功能是()。
A. 文本内容生成 B. 音乐创作
C. 多模态视频生成 D. 图像分类
4. 2025 年, 跨语言实时翻译案例中使用的技术依托于()。
A. GPT-4 B. DeepSeek 大模型
C. Gemini D. Stable Diffusion
5. 以下各项中, ()属于 AIGC 在文本生成领域的应用。
A. 图像修复 B. 公文材料撰写
C. 音乐合成 D. 视频剪辑
6. Suno 平台的创立地点是()。
A. 德国慕尼黑 B. 美国剑桥市
C. 中国北京 D. 日本东京
7. 图像生成领域解决文物数字化“不可能三角”的技术由()推出。
A. OpenAI B. 合合信息 C. 华为云 D. 谷歌
8. 传统人工智能与人工智能大模型的核心区别在于()。
A. 是否依赖大数据 B. 是否参与决策层
C. 是否使用神经网络 D. 是否需硬件支持
9. 2025 年, 视频渲染效率实现“15 秒生成”突破的是()。
A. OpenAI Sora B. Grok Imagine
C. 万兴天幕 D. Runway
10. AIGC 在视频领域推动“非遗活化”的代表作是()。
A. 《坤舆万国全图》 B. 《松花江白鱼故事》
C. 《新世界加载中》 D. 敦煌壁画数字化

AI 大作业

项目一：图像内容生成初识

运用 DeepSeek 平台生成一段提示词，再使用图像生成平台 Liblib 生成写实风格的图像。

项目二：音频内容生成初识

运用 DeepSeek 平台生成一段音乐歌词，再使用音频生成平台海绵音乐生成音频内容。

第2章 人工智能交互金钥匙

——万能的提示词

2.1 提示词概述

AI 导读：神奇的提示词

如果把人工智能比作一位超级能干的助理，那么“提示词”(Prompt)就是你交给这位助理的任务单，如图 2.1 所示。这个任务单上写着你要他做什么、怎么做、做到什么程度、用什么形式交付给你。AI 的聪明之处在于能理解各种自然语言，但它并不是“读心术大师”，它对任务的理解完全依赖提示词所表达的内容和方式。

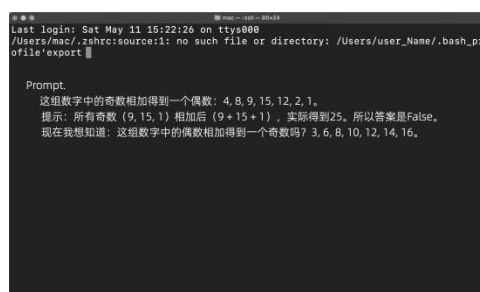


图 2.1 提示词示例

在日常生活中，我们其实每天都在用“提示词”与他人沟通。比如你去奶茶店时只说“来一杯珍珠奶茶”最基本的提示，那么你很可能得到与想象完全不同的结果。如果你补充“少糖、去冰、大杯”，那么服务员就能更准确地做出符合你口味的饮品。这与和 AI 交互的道理是一样的——提示越模糊，AI 的输出就越不可控；提示越具体，AI 就越可能给出让人满意的结果。

对于刚接触人工智能的同学来说，“提示词”听起来可能有点技术感，其实它就是一句话、一段话，甚至是一整段描述，只要能清晰地表达你的需求，它就是有效的提示词。它不需要特殊的编程语法，任何人只要掌握表达技巧，都能写出高质量的提示词。

为什么要重视提示词？因为 AI 的能力就像一片广阔的海洋，而提示词就是船舵。不同的提示会让 AI 朝着不同的方向“航行”。比如，同样是让 AI 写一篇文章，如果提示词是“写一篇关于环保的文章”，它可能给你一篇泛泛而谈的文章；但如果提示词是“写一篇 500 字的环保倡议书，主题是减少塑料袋使用，语言通俗易懂，适合中学生阅读”，它就会很精准地按照要求生成结果。

更重要的是，提示词不仅影响输出的内容，还能决定输出的形式。AI 不仅可以写文章，还

能生成表格、代码、图片描述等，这些全都取决于提示词。比如，提示词“用表格列出 5 种节能方法”就会直接让 AI 输出表格格式的答案，而不是一段文字。

在学习人工智能的过程中，理解“提示词”的概念是 AI 应用的第一步。提示词不仅是一条指令，更是一种思维方式——你要学会像设计任务一样去设计提示词。就像好的导演能向演员提出清晰的场景设定、台词和情绪要求，AI 也需要你用提示词为它搭好舞台、设定剧情、安排角色，这样它才能按你的预期进行“表演”。

2.1.1 提示词的四大基本特征

从概念上讲，提示词就是用自然语言写成的任务说明书，目的是让 AI 明白三件事：要做什么、做到什么程度、以什么样子交付。它不等同于口号或随意聊天，而是面向结果的沟通文本。好的提示词应具备四个基本特征。

- 可读性：不掌握专业术语也能读懂。
- 指向性：目标明确，知道“要产出什么”。
- 约束性：给出必要边界(对象、范围、时间、格式等)。
- 可迭代：可根据反馈快速增删细节，形成连贯的上下文。

因此，提示词不是“魔法咒语”，更像“清晰表达+任务管理”的结合体——人人可学、跨学科通用。

2.1.2 提示词与问题、指令、需求的关系和边界

很多人把“随口一问”当提示词，结果输出不可控。概念上可这样理解：

- 问题是求知(如“什么是碳中和”)。
- 指令是行动(如“写一段介绍碳中和的文字”)。
- 需求是目标和约束(如“为新生制作一页 PPT，主题是碳中和，3 个要点，口语化”)。

提示词：既说明目标，又给出必要约束，并暗含评判标准。

还需要理解以下两个概念。

最小可用提示词：只说清“做什么”，能得到“能用但粗糙”的结果。

充分提示词：补充对象、场景、受众、格式与边界，结果更稳定、更贴合预期。

同一个模型下，最小可用提示词与充分提示词的差距，往往决定了需要“事后返工”还是“一次到位”。

2.1.3 判断提示词是否合格的三条标准

为了让概念落地，可用三性检核做自查。

- 清晰性：目标是否用一句话复述，应避免“写好点”“丰富些”这类空话。
- 可执行性：AI 是否能据此工作，是否提供了对象、范围、格式或示例。
- 可验证性：产出是否可对照检查，是否能用“是否包含 $\times\times$ ”“是否不超过 Y 字”“是否列出 N 条”等进行验收。

可用 10 秒自检法判断提示词是否合格，读完提示词能否回答三个问题：做什么、边界在哪、交付什么，若有一个问题不能回答，提示词就不合格。

这里给大家举例说明。

不合格的提示词：介绍一下校园环保(模糊、无边界、无交付形态)。

合格的提示词：为新生班会写一段校园环保倡议词(约 200 字)，包含分类投放、重复利用、随手关灯 3 个可执行举措，口语化，末尾加一句号召语。

2.1.4 概念延伸：提示词是一种结构化表达

可以把提示词当成一种结构化表达，先思考受众是谁、要解决什么、产出以何种形态呈现、有哪些硬约束，再组织语言。它对应三种通用思维。

- 任务分解：大目标拆成小步骤，逐条写清。
- 信息压缩：去虚词、避空话，保留决定产出的关键信息。
- 受众建模：想清楚目标受众，面对不同的受众采用不同的表达方式。

以下误解需要澄清。

误解 1：提示词越长越好。实际上“越关键的内容越精练”，冗词会稀释重点。

误解 2：一次成形。优秀的提示词来自小步试错、快速调整的迭代，不断加入被证实“有用”的细节。

误解 3：只靠模板。模板虽能启发思维，但真正有效的是把任务的关键变量写进提示词中。

AI 训练营

训练目标

- 能够对清晰性、可执行性、可验证性进行三性检核，写出合格的提示词。
- 能对比模糊提示词与高质量提示词的差异。
- 能在不同任务场景下直接套用或改写案例。

案例一：写作任务

1) 模糊版(不合格)

要求：写一篇关于校园环保的文章。

问题分析：

- 目标笼统，没有指定受众和用途。
- 没有长度限制，AI 输出结果可能过长或过短。
- 缺少核心要点，输出内容可能没有重点。

2) 优化版(合格)

要求：你是一名校园社团宣传部干事，请为即将举行的“绿色校园周”活动写一篇倡议书(约 200 字)，对象是全校师生，要求包含“垃圾分类”“节约用水”“低碳出行”三项行动的建议，要求内容积极、鼓舞人心，结尾附一句号召语。

点评：

- 角色明确(社团干事)。
- 目标清晰(写倡议书)。
- 输出限制(200 字左右)。
- 核心要点(3 个行动建议)。

- 可检验结果(结尾是否有号召语)。

案例二：辅助学习任务

1) 模糊版(不合格)

要求：给我讲讲人工智能。

问题分析：

- 太宽泛，可能输出超长科普。
- 缺少学习目标和学习场景。

2) 优化版(合格)

要求：用大学生能理解的方式简要说明人工智能的概念(不超过 150 字)，并列举一个生活中的例子，最后用一句话总结人工智能的核心价值。

点评：

- 受众定位(大学生)。
- 字数限制(150 字)。
- 必须有例子和一句话总结(可验证)。



AI 大视野：提示词在 AIGC 领域的广泛应用

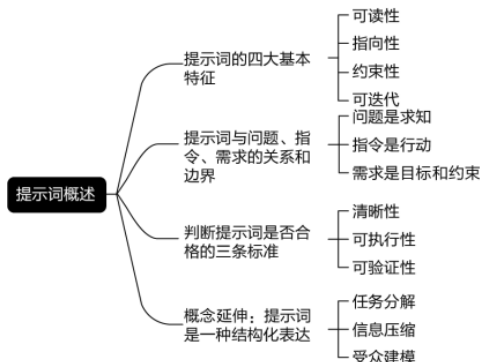
在 AI 技术不断成熟的今天，提示词已经不只是一个技术参数，而是一种全新的沟通与创造工具。它让不懂代码的人也能调用复杂的算法模型，就像过去需要专业程序员才能完成的功能一样，如今只需要用自然语言描述，就能快速生成可用结果。对于不同行业的人来说，提示词降低了技术门槛、放大了专业价值。

提示词的最大价值之一是具有高度的行业迁移性——只要理解了概念和结构，无论你是教育工作者、工程师、设计师还是创业者，都能在自己的领域发挥作用。甚至有企业已经把“提示词工程师”列为招聘岗位，要求员工能高效地用提示词调动 AI 产出高质量结果，并维护公司内部的提示词模板库。

提示词是 AI 时代的“通用语言”，也是跨行业、跨专业的“能力放大器”。它让不同背景的人都能低成本、高效率地调用人工智能的能力。在未来，高效写出提示词很可能会成为职场竞争的核心竞争力。掌握了这项能力，你就掌握了与 AI 高效协作的第一把钥匙。



AI 大复盘



AI 即时练

- 提示词的本质是()。
A. AI 的代码
B. 给 AI 的任务说明书
C. AI 的名字
D. AI 的结果
- 下列各项中, ()不属于提示词的特征。
A. 可读性
B. 指向性
C. 约束性
D. 随机性
- “写一篇关于校园环保的文章”这个提示词的主要问题是什么()。
A. 太长
B. 缺少角色定位和边界
C. 没有使用专业术语
D. 主题不明确
- 下列各项中, ()属于充分提示词。
A. 写一篇新闻稿
B. 给我介绍人工智能
C. 你是一名科技记者, 请写一篇 300 字以内的新闻稿, 报道新款智能手表的发布信息, 包含价格、功能亮点和发售日期
D. AI, 请帮我写文章
- 判断提示词是否合格的三性检核中, 不包括()。
A. 清晰性
B. 可执行性
C. 可验证性
D. 创造性
- “用表格列出 5 种节能方法”中的“用表格”属于提示词结构中的()。
A. 角色定位
B. 任务目标
C. 执行细节/输出形式
D. 输出限制
- 在提示词优化过程中, ()是正确的。
A. 尽量用模糊的描述让 AI 自由发挥
B. 一次性写完, 不做修改
C. 通过多轮迭代逐步补充细节
D. 完全依赖 AI 自动补全
- “你是一名高中语文老师, 请写一篇 150 字的励志演讲稿, 面向高三学生, 语言鼓舞人心, 结尾包含一句名人名言”这段提示词符合三性检核中的()。
A. 清晰性
B. 可执行性
C. 可验证性
D. 以上全部
- 在提示词的概念中, “最小可用提示词”指的是()。
A. 只包含任务目标的提示词
B. 包含角色、细节和边界的提示词
C. 字数最少的提示词
D. 只能生成一部分内容的提示词
- 提示词能力的行业迁移性意味着()。
A. 只适用于 IT 行业
B. 学会提示词后可在不同领域复用
C. 提示词必须固定不变
D. 提示词只能在写作中使用

AI 大作业

项目一：从模糊到精准——提示词迭代实验

任选一个你感兴趣的主题(如旅游攻略、短视频脚本、科普文章等):

- 先写最小可用提示词(只包含任务目标), 并在 AI 工具中运行, 保存输出结果。
- 在原提示词的基础上, 逐步加入角色定位、细节描述、输出限制等内容, 形成充分提示词, 再次运行并保存结果。
- 对比两次输出, 分析在内容完整度、贴合度、可读性上的差异。
- 用 200~300 字总结你对“提示词质量影响结果”的理解, 并分享一条你认为最实用的优化技巧。

项目二：一题多用——提示词的场景迁移

针对三个不同领域, 分别设计高质量提示词:

- 写作类(如生成文章、演讲稿、广告文案等)。
- 绘画类(驱动绘画 AI 生成图片)。
- 数据分析类(要求 AI 处理并分析数据)。

每段提示词都需要包含角色定位、任务目标、执行细节、输出限制四要素。

2.2 提示词的作用

AI 导读：让 AI “听懂” 并 “做对” 的秘诀

在人工智能的世界里, 我们常常会遇到一个有趣的现象: 同一个 AI 模型, 有的人用得得心应手, 轻松产出高质量内容; 而有的人却总觉得“AI 不够聪明”“回答离题”“生成结果差强人意”。造成这种差别的关键因素之一就是是否深刻理解提示词的作用。

提示词的作用并不是单一的“让 AI 回答问题”这么简单, 它涵盖了多个层面。

- 沟通作用: 用自然语言把任务精确传达给 AI, 避免“鸡同鸭讲”。
- 引导作用: 像导航地图一样, 为 AI 指定前进的方向和终点, 让它沿着预期路径生成内容。
- 约束作用: 在任务中设置边界与规则, 防止 AI 输出偏题、冗余或不合规的结果。
- 激发作用: 通过创意化的提示词, 调动 AI 的联想力与创造力, 让结果更加新颖、有趣。

举一个简单的例子, 让 AI “写一首诗”, 它可能天马行空地写出任何风格的诗; 但如果提示词改为“用唐诗的格律写一首关于春天校园生活的七言绝句”, AI 不仅能明确诗的形式, 还会在主题、用词上更契合需求。这种从模糊到精准的变化, 就是提示词作用的直接体现。

2.2.1 沟通作用：让人机对话更高效、顺畅

提示词首先是人与 AI 的沟通桥梁。就像两个人说话, 如果语言含糊、信息不全, 则对方很难完全理解你的意思; AI 也是如此。

在自然语言交互中, 提示词承担着“翻译官”的角色, 把人的需求转化为 AI 能准确理解的结构化信息。

- 模糊提示词示例: “帮我做个计划。” AI 可能输出旅游计划、学习计划、健身计划……结果五花八门。
- 精准提示词示例: “帮我制订一个为期 30 天的健身计划, 每周锻炼 4 天, 重点提升上肢力量, 适合初学者, 并在表格中列出每日训练内容。”