

计算机基础与人工智能融合 应用教程（微课版）

刘三满 郭丽蓉 朱彪凯 主 编
陈云云 曾倩倩 张婷婷 魏利梅 曹敏 李一冉 副主编

清华大学出版社
北 京

内 容 简 介

本书共计 10 章,系统梳理数字化办公与人工智能应用相关基础知识,内容体系完整、逻辑层层递进。首先阐释信息数据基础、计算机软硬件核心知识,以及系统安全、AI 应用安全等基础理论内容,兼顾主流操作系统与国产操作系统的实操应用讲解。在办公实操层面,详细讲授 WPS 文字、表格、演示文稿的基础常规操作,同时聚焦智能办公新技能,系统讲解 AI 辅助内容创作、智能排版、数据处理、可视化设计等实用技术。此外,本书深入介绍人工智能基础理论、大语言模型、提示词工程及生成式 AI 的技术原理与落地应用方法,结合实操案例讲解 AI 智能办公综合应用技巧。全书同步涵盖人工智能伦理准则、安全风险防控、数据隐私合规及相关行业法律法规等规范内容。

本书主要面向计算机与人工智能办公初学者,适合作为各类高等院校计算机与人工智能通识课程的教材、数字化办公培训用书,也可作为职场人员提升 AI 办公能力的参考资料。

本书配套免费的电子课件和习题答案,读者扫描封底或前言中的二维码即可获取。另外,读者扫描书中的视频二维码可以直接观看教学视频。

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

计算机基础与人工智能融合应用教程:微课版/刘三满,
郭丽蓉,朱昶凯主编.--北京:清华大学出版社,2026.9.

ISBN 978-7-302-72852-8

I. TP3; TP18

中国国家版本馆CIP数据核字第2026K5C631号

责任编辑:胡辰浩

封面设计:高娟妮

版式设计:妙思品位

责任校对:马遥遥

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:北京鑫海金澳胶印有限公司

经 销:全国新华书店

开 本:185mm×260mm

印 张:22.75

字 数:540千字

版 次:2026年10月第1版

印 次:2026年10月第1次印刷

定 价:69.80元

产品编号:117842-01

前言

当前，人工智能技术已全面渗透社会生产、办公、生活、教育科研各个领域，传统计算机基础教育面临全新的迭代升级需求。单纯的计算机软硬件基础、常规办公软件操作知识已经无法满足新时代数字化、智能化人才培养的核心目标。

为顺应智能时代教育改革趋势，精准匹配本科、高职、中职各层次院校公共计算机基础、人工智能通识、数字化实训等课程教学标准，贴合新时代职场智能化、数字化岗位的核心能力需求，由山西警察学院计算机通识教育、人工智能应用教学领域一线骨干教师组成的教材编写团队，在2020年《新编计算机基础教程》(第一版)、2024年《新编计算机基础教程(微课版)》(修订再版)的基础上，深度对标国家教学改革标准与行业技术前沿，依托多年一线计算机通识教学、人工智能应用教研与数字化技能实训经验，历经多轮课堂试点应用、内容打磨优化与案例更新迭代，摒弃原有传统计算机基础教材框架，颠覆式重构全部教学内容，融入AI通识与智能技术应用体系，完成教材体系全面迭代升级，新编完成《计算机基础与人工智能融合应用教程(微课版)》一书。

本书以零基础通识育人为核心目标，彻底打破传统计算机基础教材与人工智能教材相互割裂的编写壁垒，消除传统计算机教材“重基础、轻智能”、AI通识教材“重理论、难落地”的弊端，构建一套兼顾基础性、前沿性、实用性与规范性的全新通识知识体系，创新性地将基础计算机知识、办公技能与人工智能、大模型、生成式AI前沿技术深度融合，全方位适配院校教师课堂教学、学生自主学习、职场技能提升等多元场景。

本书在整体架构设计上，立足“基础够用、实操为主、智能赋能、合规先行、本土适配”的核心编写原则，创新性构建“计算机基础+国产操作系统+智能办公实操+人工智能应用+合规伦理素养”五位一体的闭环式知识体系。遵循“由浅入深、从基础到应用、从技能到素养、从专项到综合”的认知逻辑层层递进，完整覆盖计算机底层基础、国产操作系统、智能办公应用、人工智能技术、实战场景训练、伦理合规规范等全知识链条，既保障基础知识的系统性与完整性，又突出前沿技术的创新性与实用性，完全契合零基础学习者的认知规律与通识课程的教学逻辑。

本书分为基础赋能模块、智能办公模块、AI通识实操模块、综合实战模块、素养合规模块，各模块环环相扣、互为支撑。基础赋能模块(第1章、第2章)系统讲解计算机软硬件

架构、网络基础与信息安全等内容，重点融入国产操作系统实操教学，补齐传统教材信息技术应用创新(简称信创)教育短板，为后续智能技术学习筑牢知识根基。智能办公模块(第3章至第5章)立足职场实用能力重构，以主流WPS办公平台为载体，摒弃传统机械化操作讲解，深度融合AI智能创作、智能排版、数据智能分析、演示智能优化等智能化功能，实现传统办公技能向智能高效办公能力的升级转型。AI通识实操模块(第6章至第8章)聚焦前沿技术通俗化落地，循序渐进地讲解人工智能基础理论、大模型核心原理、提示词工程、多模态生成式AI应用等前沿内容，以案例化、场景化、实操化方式拆解抽象技术难点，让零基础学习者看得懂、学得会、用得上。综合实战模块(第9章)侧重知识融会贯通，整合全书核心知识点，搭建多工具联动、AI全流程赋能的真实办公实战场景，实现碎片化知识系统化、理论知识实战化，切实提升学习者综合数字应用能力。素养合规模块(第10章)补齐育人体系短板，系统阐释人工智能伦理规范、网络安全准则与相关法律法规，强化学习者规范用AI、安全用AI、理性用AI的职业素养，构建“技能+素养”的完整育人闭环。

本书在教学设计层面，充分适配常态化教学场景，构建标准化、可落地、全覆盖的教学闭环，兼顾教师课堂授课与学生自主学习双重需求。教材整体课时体量适配国内通用教学安排，知识密度均衡、难易梯度合理，适配多层次院校教学需求。各章前置章节概述与重难点梳理，帮助教师快速搭建授课框架、明确教学核心，帮助学生精准定位学习目标、把握知识重点。每章结尾配套系统化知识小结、分层思考题与实操训练题，实现理论巩固、技能实训、思维拓展三位一体的教学效果。同时，全书配套权威参考文献与拓展学习资源，有效保障教材内容的专业性、严谨性与前沿性，为师生深度学习、教研创新提供有力支撑。

本书立足智能时代教育改革需求，相较于国内同类通识教材，形成五大差异化特色优势。一是前沿性突出，紧跟生成式AI技术迭代浪潮，剔除陈旧滞后内容，将大模型应用、提示词工程、多模态智能创作等最新主流技术融入通识教学，贴合数字智能时代的发展节奏。二是融合性充分，打破传统学科壁垒，将智能思维、智能技术、智能应用贯穿计算机基础学习、办公技能训练、综合实践全过程，实现传统基础学科与新兴智能技术的深度耦合。三是实操性极强，坚持“学以致用”的通识育人导向，弱化冗余理论阐释，以校园、职场真实场景为依托，聚焦主流智能工具的落地实操，确保学习成果可直接转化为应用能力。四是本土化适配，紧扣国产信息化与信创教育发展国策，重点强化国产操作系统、国产办公软件教学内容，适配国内教育国产化改革与本土产业人才培养需求。五是体系性完备，构建从基础认知、工具操作、智能应用、实战落地到伦理合规的全链条知识体系，形成数字化人才培养的完整闭环，填补同类教材体系碎片化、素养教育缺失的短板。

本书由刘三满教授统筹整体内容架构、开展教学体系设计，负责全文审核与最终统稿。教材编写过程中，编者参阅了大量国内外优秀教材、权威学术文献与行业技术规范，吸纳了人工智能领域前沿落地成果与全国各地的先进教研经验，在此向所有参考文献作者、行业研究机构与教研同仁致以诚挚的谢意。

智能时代的计算机通识教育，核心育人目标已从“掌握基础操作技能”升级为“培育数字素养、智能思维、创新能力与合规素养四位一体的复合型人才”。希望本书能够助力广大学习者夯实计算机基础功底、掌握智能办公核心技能、理解人工智能核心逻辑、树立合规安全应用理念，全面提升个人数字核心竞争力，更好适配数字化职场与智能化社会的发展需求。同时，也期望本书能够为高校数字化通识教育、人工智能普及教育、信创基础教育提供优质的教学载体，为新时代数字化应用型人才培养提供有力支撑。

本书既可作为本科、高职、中职院校的公共计算机通识课、人工智能入门课、数字化办公实训课的专用教材，也可作为职场人员提升数字化、智能化技能的自学与培训用书。

本书依托多项省部级重点科研课题与教学改革课题编写成果，在此特别致谢：2025年陕西省智慧警务重点实验室开放课题“基于区块链技术的深度伪造信息鉴别与溯源方法研究”(SXZJ25KF10)；警务融合计算四川省重点实验室2025年度开放课题“大模型内容合法性甄别与主动溯源取证技术研究”(JWRH202503003)；太行山西省实验室科技攻关项目“基于深度学习模型的网络资产识别方法研究”(THYF-KJGG-25010500)；山西警察学院教学改革重点课题“AI赋能公安实战化教学改革研究”等。

由于编者学术视野与教研经验有限，书中难免存在疏漏与不足，恳请各位教育同仁、行业专家与广大读者不吝批评指正。我们的电话是010-62796045，邮箱是992116@qq.com。

本书配套免费的电子课件和习题答案，读者扫描下方二维码即可获取。另外，读者扫描书中的视频二维码可以直接观看教学视频。



作者
2026年7月

目 录

第 1 章 信息与计算机基础	1
1.1 信息、数据与智能信息时代	1
1.1.1 数据与信息	1
1.1.2 信息技术的发展	2
1.1.3 智能信息时代	3
1.2 计算机的发展、分类与智能化发展趋势	3
1.2.1 计算机的发展	4
1.2.2 计算机的分类	5
1.2.3 计算机智能化发展趋势	7
1.3 计算机中的信息表示与人工智能数据底层基础	8
1.3.1 进位计数制	9
1.3.2 不同进制之间的转换	10
1.3.3 数据的存储单位	12
1.3.4 计算机的信息编码	12
1.3.5 人工智能数据底层基础	16
1.4 计算机系统组成与智能硬件架构	18
1.4.1 计算机系统组成	18
1.4.2 计算机工作原理	22
1.4.3 智能硬件架构	22
1.5 计算机网络基础	23
1.5.1 计算机网络概述	23
1.5.2 计算机网络服务	25
1.5.3 移动互联网	27
1.6 系统安全与人工智能应用安全入门	29
1.6.1 计算机系统面临的安全威胁	30
1.6.2 计算机系统安全防护技术	32

1.6.3 人工智能应用安全入门	35
1.7 本章小结	37
1.8 思考和练习	37
第 2 章 操作系统与国产系统基础	39
2.1 操作系统概述	39
2.1.1 操作系统的核心定义与系统层级关系	40
2.1.2 操作系统的发展历程	40
2.1.3 现代操作系统的四大核心基本特征	41
2.1.4 操作系统的五大核心功能	42
2.1.5 操作系统的主流分类体系	43
2.1.6 人工智能时代操作系统的技术演进与应用价值	44
2.2 Windows 11 基础操作与系统 AI 工具应用	45
2.2.1 Windows 11 系统概述与核心特点	45
2.2.2 Windows 11 核心基础操作	46
2.2.3 Windows 11 原生系统 AI 工具核心应用	48
2.3 文件资源管理与 AI 智能检索分类	51
2.3.1 文件资源管理核心基础理论	51
2.3.2 Windows 11 标准文件资源管理实操体系	52
2.3.3 Windows 11 AI 智能检索的技术原理与核心优势	53
2.3.4 Windows 11 AI 智能检索的功能与场景应用	54

2.3.5	Windows 11 AI 智能分类与 批量文件 AI 处理功能	55	3.2.2	页面设置	78
2.3.6	AI 文件管理在人工智能教学与 实操中的典型应用场景	55	3.2.3	编辑文本	78
2.3.7	文件智能管理规范、常见问题与 安全注意事项	56	3.2.4	保存文档	79
2.4	系统设置与智能辅助功能应用	57	3.2.5	输出与打印文档	80
2.4.1	Windows 11 系统设置基础 体系与操作入口	57	3.2.6	AI 辅助内容创作	82
2.4.2	核心系统模块设置与 AI 场景适 配配置	58	3.3	格式美化与 AI 智能润色排版	83
2.4.3	Windows 11 原生智能辅助核心 功能	63	3.3.1	字体、段落格式的设置	83
2.4.4	无障碍智能辅助功能应用	65	3.3.2	项目符号和编号的添加	85
2.5	国产操作系统基础	65	3.3.3	WPS 文字中对象的插入	86
2.5.1	国产操作系统的发展背景与核心 价值	66	3.3.4	AI 智能润色排版	89
2.5.2	国产操作系统的核心技术架构与 技术路线	66	3.4	表格图文混排与 AI 图文优化	91
2.5.3	国产操作系统的分类体系与主流 系统详解	67	3.4.1	插入表格	91
2.5.4	国产操作系统的基础操作与智能 化功能应用	68	3.4.2	编辑表格	91
2.5.5	国产操作系统的 AI 适配能力与 教学应用场景	69	3.4.3	美化表格	92
2.5.6	国产操作系统与 Windows 系统 对比分析	70	3.4.4	图文混排	93
2.5.7	国产操作系统发展趋势与学习 意义	70	3.4.5	AI 图文优化	94
2.6	本章小结	71	3.5	长文档排版与 AI 智能审阅处理	96
2.7	思考和练习	71	3.5.1	编辑长文档	96
第 3 章	WPS 文档处理	72	3.5.2	审阅与修订长文档	100
3.1	WPS 文字基础与 AI 功能入口	72	3.5.3	AI 智能审阅处理	102
3.1.1	WPS 文字启动	72	3.6	文档综合训练	102
3.1.2	WPS 文字的界面布局	73	3.6.1	案例背景	102
3.1.3	WPS 文字关闭	75	3.6.2	操作实现	103
3.1.4	AI 功能入口	75	3.7	本章小结	107
3.2	文档编辑与 AI 辅助内容创作	76	3.8	思考和练习	107
3.2.1	创建文档	77	第 4 章	WPS 电子表格处理	110
			4.1	WPS 表格基础与 AI 数据 工具入口	110
			4.1.1	WPS 表格基础	110
			4.1.2	WPS 表格中的基本概念	111
			4.1.3	WPS AI	112
			4.2	数据录入格式与 AI 智能规整	113
			4.2.1	项目提出与分析	113
			4.2.2	项目主要知识点	113
			4.2.3	项目完成步骤	113
			4.3	公式函数应用与 AI 智能 计算辅助	116
			4.3.1	公式与函数	116
			4.3.2	项目需求与分析	117

4.3.3 项目主要知识点	118	5.3.3 AI 智能版式美化的 实操流程	149
4.3.4 项目完成步骤	118	5.3.4 AI 美化后的精细化 优化技巧	150
4.4 数据管理分析与 AI 数据 挖掘校验	120	5.3.5 版式美化中的常见问题与 解决方法	151
4.4.1 项目需求与分析	120	5.4 动画交互设置与 AI 智能 特效推荐	151
4.4.2 项目主要知识点	120	5.4.1 动画与交互的基础认知及 规范原则	152
4.4.3 项目完成步骤	120	5.4.2 手动动画与交互的 核心设置操作	153
4.5 图表可视化操作与 AI 图表 智能分析	125	5.4.3 WPS AI 智能特效推荐的 功能与实操流程	155
4.5.1 项目需求与分析	125	5.4.4 动画交互中的常见问题与 解决方法	156
4.5.2 项目主要知识点	126	5.5 演示文稿放映输出与 AI 演示 辅助应用	156
4.5.3 项目完成步骤	126	5.5.1 演示文稿放映	156
4.6 电子表格综合训练项目	128	5.5.2 演示文稿打包与输出	159
4.6.1 项目背景	128	5.5.3 AI 演示辅助应用	161
4.6.2 项目要求	129	5.6 演示文稿综合训练	162
4.6.3 操作步骤	129	5.6.1 应用 AI 工具制作 PPT	162
4.7 本章小结	133	5.6.2 制作“三晋风韵宣传” 演示文稿	163
4.8 思考和练习	133	5.7 本章小结	166
第 5 章 WPS 演示文稿制作	137	5.8 思考和练习	167
5.1 WPS 演示基础与 AI 演示 功能概述	137	第 6 章 人工智能基础	169
5.1.1 WPS 演示文稿启动	137	6.1 人工智能概述	169
5.1.2 WPS 演示操作界面	138	6.1.1 人工智能与图灵测试	171
5.1.3 AI 演示功能概述	138	6.1.2 人工智能的分类	171
5.2 幻灯片编辑与 AI 文稿框架生成	139	6.1.3 人工智能、机器学习与 深度学习的关系	173
5.2.1 幻灯片基础编辑操作	140	6.2 人工智能发展历程	174
5.2.2 WPS AI 文稿框架生成的 功能概述	142	6.3 人工智能关键技术	177
5.2.3 AI 文稿框架生成的 完整实操流程	142	6.3.1 机器学习	177
5.2.4 AI 生成框架的编辑优化 技巧	146	6.3.2 深度学习与人工神经网络	180
5.2.5 实操中的常见问题与 解决方法	146	6.3.3 知识表示与推理	182
5.3 版式美化与 AI 智能设计优化	147		
5.3.1 演示文稿的基础版式 美化规范	147		
5.3.2 WPS AI 智能设计优化的 功能概述	149		

6.3.4	自然语言处理	184	7.4.2	提示词核心应用体系	221
6.3.5	计算机视觉	186	7.4.3	办公场景典型提示词设计实践	224
6.3.6	智能语音技术	188	7.5	大模型使用规范与实操技巧	225
6.4	人工智能典型应用场景	190	7.5.1	大模型使用规范与风险防控	225
6.4.1	智慧医疗	190	7.5.2	大模型高效实操的核心技巧	227
6.4.2	智能交通与自动驾驶	191	7.5.3	大模型发展趋势与终身学习准则	228
6.4.3	智慧金融	193	7.6	本章小结	229
6.4.4	智慧教育	194	7.7	思考和练习	230
6.4.5	智慧制造	194	第8章	生成式人工智能技术与应用	231
6.4.6	智能生活与生成式应用	195	8.1	生成式人工智能基础概述	231
6.5	人工智能发展趋势	196	8.1.1	什么是生成式人工智能	232
6.5.1	大模型驱动通用智能的探索	196	8.1.2	生成式人工智能核心技术	235
6.5.2	多模态融合与具身智能	197	8.1.3	生成式人工智能应用场景及影响	236
6.5.3	可信、可解释与负责任的人工智能	197	8.2	文本生成技术与办公应用	238
6.5.4	人工智能与各领域的深度融合	198	8.2.1	文本生成底层技术原理	238
6.5.5	绿色低碳与高效化发展	198	8.2.2	文本生成应用场景与配套工具	239
6.5.6	人机协同共生	198	8.2.3	办公场景实操案例	241
6.6	本章小结	199	8.3	图像生成技术与图文设计应用	242
6.7	思考和练习	199	8.3.1	图像生成的底层技术原理	242
第7章	大模型与提示词基础	201	8.3.2	主流图像生成工具与场景应用	243
7.1	大语言模型概述	201	8.3.3	图像生成实训案例	244
7.1.1	大模型的核心概念	202	8.4	音频生成技术与声音媒体应用	246
7.1.2	大模型的核心优势	204	8.4.1	基本原理与步骤	246
7.1.3	大模型的应用领域	205	8.4.2	核心技术与术语	247
7.2	大模型技术原理	207	8.4.3	音频生成工具	249
7.2.1	大模型的核心技术	208	8.4.4	音频生成应用案例	250
7.2.2	大模型的训练过程	209	8.5	视频生成技术与影音媒体应用	260
7.2.3	大模型的能力边界与局限性	211	8.5.1	基本原理与步骤	260
7.3	主流大模型介绍	213	8.5.2	核心技术与术语	261
7.3.1	国际主流大模型	213	8.5.3	视频生成工具	263
7.3.2	国内主流大模型	215	8.5.4	视频生成应用案例	264
7.3.3	大模型的分类与选型方法	217			
7.4	提示词工程基础	219			
7.4.1	提示词工程概述	219			

8.6	生成式人工智能综合实操训练	270	9.4.4	人工智能优化演示效果	312
8.6.1	实训案例概述	270	9.4.5	综合案例：完成一整套 汇报 PPT	313
8.6.2	案例整体创作流程	270	9.4.6	Prompt 技巧	314
8.6.3	实训总结	277	9.5	多工具联动与人工智能办公	
8.7	本章小结	277	综合实训	315	
8.8	思考和练习	278	9.5.1	为什么需要多工具联动	315
第 9 章	人工智能赋能智能办公		9.5.2	WPS 人工智能的跨组件 联动	317
综合实战	279		9.5.3	其他平台的联动能力	319
9.1	智能办公体系与人工智能 办公生态	279	9.5.4	综合实训：举办校园人工智能 创新论坛	320
9.1.1	从数字办公到智能办公	280	9.5.5	Prompt 技巧：多任务串联的 指令设计	322
9.1.2	人工智能办公生态体系	281	9.6	本章小结	323
9.1.3	Prompt 基础	282	9.7	思考和练习	323
9.1.4	人工智能办公伦理与规范	284	第 10 章	人工智能伦理、安全与法规	325
9.1.5	技能训练：设计三个不同 场景的 Prompt	285	10.1	人工智能伦理	325
9.2	WPS 人工智能辅助文档处理 实战	286	10.1.1	人工智能伦理的概念	325
9.2.1	人工智能辅助公文写作	286	10.1.2	人工智能伦理核心问题	326
9.2.2	人工智能辅助报告写作	289	10.1.3	人工智能伦理治理	330
9.2.3	人工智能辅助论文排版	290	10.1.4	人工智能伦理典型案例	333
9.2.4	人工智能辅助文档校对	291	10.2	人工智能安全	336
9.2.5	综合案例：完成一份完整的 项目申请书	292	10.2.1	人工智能安全的概念	336
9.2.6	Prompt 技巧	293	10.2.2	人工智能安全的核心问题	336
9.3	人工智能数据处理与分析实战	294	10.2.3	人工智能安全治理	338
9.3.1	人工智能辅助 Excel 操作	295	10.2.4	人工智能安全典型案例	341
9.3.2	人工智能辅助数据清洗	297	10.3	人工智能相关法律法规与 行业规范	342
9.3.3	人工智能辅助统计分析	299	10.3.1	国内外与人工智能相关的 法律法规	342
9.3.4	人工智能辅助图表设计	300	10.3.2	国内外人工智能风险 治理政策	347
9.3.5	综合案例：从数据到生成 报告的完整流程	305	10.4	本章小结	350
9.4	人工智能辅助演示设计实战	306	10.5	思考和练习	350
9.4.1	人工智能生成 PPT 大纲	306	参考文献	351	
9.4.2	人工智能生成 PPT 内容	309			
9.4.3	人工智能生成图片	310			

第 1 章

信息与计算机基础

如今人工智能已全面渗透生产生活，智能推荐、语音助手、大数据分析、AI生成等应用随处可见，而一切智能技术的落地，都根植于计算机科学最底层的基础知识。本章遵循由浅入深、层层递进的编排逻辑，依次展开六大模块内容：从源头辨析数据、信息的定义与关联，梳理信息技术演进历程，整体感知智能信息时代全貌；回顾计算机诞生至今的发展阶段、品类划分，梳理设备整体智能化演进方向；讲解二进制等计算机信息编码规则，阐释人工智能数据存储与运算的底层原理；剖析完整计算机系统构成，延伸适配人工智能运算的智能硬件架构；阐释计算机网络基础知识，分析数据互通传输的支撑条件；最后聚焦信息安全领域，兼顾传统计算机系统安全与新兴人工智能应用安全要点。

📖 本章的教学重点：

- 信息与数据的概念及价值
- 计算机中的信息表示与人工智能数据底层基础
- 计算机系统组成及智能硬件架构
- 系统安全防护与人工智能应用安全

📖 本章的教学难点：

- 二进制数据底层原理与人工智能数据处理的关联
- 传统计算机技术与智能应用的融合逻辑
- 传统系统安全与人工智能应用安全的风险防护

1.1 信息、数据与智能信息时代

随着信息技术、互联网与物联网的深度普及，人类社会进入数字化发展阶段。数据、信息是信息技术体系的基础要素，是计算机处理的核心对象。信息技术的成熟，推动数据资源价值深度释放，结合人工智能技术的迭代升级，社会由传统信息时代逐步迈入智能信息时代。

1.1.1 数据与信息

1. 数据与信息概述

数据与信息是计算机科学中最基础的两个概念，二者既有联系又有明显区别，是理解

信息技术工作原理的基础。

数据是对客观事物属性的原始记录，是未经加工、未赋予具体含义的符号集合。数据的表现形式多样，包括数字、文字、字符、图像、音频、视频等。原始数据具有零散、无序、抽象的特点，本身不具备明确的语义，无法直接用于指导判断和决策，仅作为承载客观事实的原始素材。

信息是对数据进行采集、筛选、整理、加工、分析后得到的有意义、有价值的内容，能够准确反映客观事物的状态、特征和变化规律，可直接服务于人们的认知、判断与决策。信息具有不灭性、可存储性、可处理性、可重用性和价值性等特性；同时信息能减少不确定性。

2. 数据与信息的内在联系

数据与信息的关系可以概括为：数据是信息的载体，信息是数据加工处理后的结果与价值体现。数据是原材料，信息是成品。没有数据，信息就失去了来源；未经加工处理的数据，不具备信息价值。计算机的基本工作过程，本质上就是将海量原始数据转化为有效信息。在数字经济时代，数据已成为重要的新型生产要素，是推动信息化、智能化发展的核心资源。

1.1.2 信息技术的发展

信息技术的发展经历了多个阶段，这些阶段不仅标志着人类信息传递方式的巨大变革，也推动了社会文明的进步。以下是信息技术发展的主要阶段。

1. 语言的使用

大约在30万年前，人类开始使用语言作为思想交流和信息传播的工具。这一阶段标志着信息技术的起源。语言不仅使人类能够表达复杂的思想和情感，还促进了文化的传承和社会的发展。

2. 文字的出现

大约在公元前3500年，人类社会开始出现文字。文字的发明是人类历史上的重要里程碑，它使信息可以超越时间和地域的限制被保存和传播。文字的出现极大地推动了人类文明的发展，使知识、历史和文化的积累成为可能。

3. 印刷术的发明

公元1040年左右，毕昇发明了胶泥活字印刷术，欧洲则在约1450年开始使用印刷术。印刷术的发明极大地提高了信息传播的效率。特别是活字印刷术的出现，使图书和报刊等印刷品得到大量生产，成为重要的信息存储和传播媒介，对人类文明的传播和知识的普及产生了深远影响。

4. 电磁波传播技术的发展

19世纪，随着电报、电话、广播及电视的发明和普及，人类进入利用电磁波传递信息的时代。电磁波传播技术的出现和应用使信息的传递速度极大提升，信息的流通范围也进一步扩大。该技术的发展为现代通信技术的兴起奠定了基础。

5. 计算机和互联网的使用

这个阶段的标志是1946年电子计算机的问世。计算机和互联网的发明与使用标志着

现代信息技术的诞生。计算机技术的快速发展使信息处理变得更加高效和便捷，互联网则将全球连接成一个巨大的信息网络。这一阶段彻底改变了人类获取、处理和传递信息的方式，推动了信息社会的形成和发展。

信息技术发展到今天，技术更新迭代不断加快，新兴数字技术层出不穷，云计算、物联网、大数据、人工智能、区块链、元宇宙等新兴数字技术正在深刻改变着人们的生活和工作方式，对人类社会的未来发展也将产生颠覆性影响。

1.1.3 智能信息时代

随着大数据、人工智能、云计算、物联网等新一代信息技术的深度融合，传统以信息处理、信息传输为核心的信息时代逐步迭代升级，进入智能化、数字化深度融合的智能信息时代。与传统信息时代计算机仅能被动执行程序、完成数据存储与运算不同，智能信息时代以数据为核心资源、以人工智能为核心驱动力，使计算机系统具备感知、学习、推理、判断和自主决策的能力，实现了信息技术从“数字化处理”向“智能化赋能”的根本性转变。

智能信息时代具有区别于传统信息时代的典型特征，主要体现在以下四个方面。

1. 数据驱动智能化

数据成为核心生产要素，通过对海量、多维、实时的数据进行训练与分析，计算机系统能够自主挖掘规律、拟合场景需求，摆脱了完全依赖人工编程的运行模式，实现任务自适应、场景自适应。

2. 人机交互智能化、自然化

打破键盘、鼠标等传统机械式交互局限，形成语音识别、机器视觉、触控交互、多模态融合等新型交互方式，人机交互更加贴近自然语言和人类感知习惯，交互效率与体验大幅提升。

3. 算力架构协同化

形成云、边、端一体化的协同计算模式，云端负责海量数据训练与模型迭代，边缘端负责就近实时处理，终端设备实现智能感知与快速响应，有效解决了传统计算延迟高、算力不足、实时性差等问题。

4. 应用场景泛在化、安全化

智能技术全面渗透工业、医疗、教育、交通、政务等各个领域，实现全行业智能化升级。同时，隐私计算、可信人工智能、数据安全等技术不断完善，构建起安全、可控、可追溯的智能应用体系。

1.2 计算机的发展、分类与智能化发展趋势

从第一台电子计算机诞生至今的八十年里，人类用晶体管、芯片和代码搭建起了全新的数字文明底座：从占据整座厂房的巨型电子管设备，到揣在口袋里的智能手机，计算机

的形态在不断迭代，性能沿着摩尔定律的轨迹实现了指数级跃升。当我们回望这段发展历程，会发现计算机的分类边界早已从最初的“巨型机—大型机—小型机—微型机”的清晰层级，延伸到了今天覆盖云边端的多元算力矩阵。而当下，随着大模型、多模态交互技术的快速落地，计算机正从被动执行指令的工具，向具备自主感知、自主决策能力的智能体演进。

1.2.1 计算机的发展

1. 计算机的诞生

在“二战”时期，由于军事计算需要，美国宾夕法尼亚大学成功研制出第一台通用电子数字计算机，命名为ENIAC(Electronic Numerical Integrator And Computer，电子数字积分计算机)，如图1-1所示。这台计算机重达30t，占地面积约为170m²，使用了18000个电子管，每小时耗电达到140kW·h，每秒可完成5000次加减法运算，相比同时期的传统手工操作，计算机速度提高了约8400倍，ENIAC的成功研制标志着电子计算机时代的到来。



图 1-1 ENIAC

2. 计算机的发展

自第一台计算机诞生以来，根据计算机中采用的电子器件的不同，计算机的发展大体可以分为四个时代：电子管计算机时代、晶体管计算机时代、中小规模集成电路计算机时代和大规模及超大规模集成电路计算机时代。

1) 第一代电子管计算机(1946—1958年)

第一代计算机主要采用电子管(如图1-2所示)作为主要逻辑元件，内存储器采用磁鼓、磁芯，外存储器使用磁带、纸带、卡片等。这个时期的计算机体积庞大、运算速度低(一般每秒运算几千次到几万次)、成本高、可靠性低、内存容量小(一般为几千字节)。没有系统软件，用机器语言和汇编语言编写程序，主要用于军事和科研领域的科学运算(如弹道计算、天气预报等)。



图 1-2 电子管

2) 第二代晶体管计算机(1959—1964年)

第二代计算机使用的主要逻辑元件是晶体管(如图1-3所示)，内存储器采用磁芯，外存储器使用磁带和磁盘。软件方面开始使用管理程序，后期使用操作系统，并出现了FORTRAN、COBOL、ALGOL等一系列高级程序设计语言。这个时期计算机的应用扩展到数据处理、工业控制等方面。计算机的运行速度已提高到每秒几十万次，体积大大减小，可靠性和内存容量也有较大的提高。

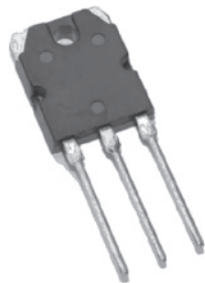


图 1-3 晶体管

3) 第三代中小规模集成电路计算机(1965—1970年)

第三代计算机主要以中小规模集成电路(如图1-4所示)为逻辑元件, 内存储器开始采用半导体元件代替磁芯, 存储容量达到几兆字节, 运行速度提高到每秒几十万次到几百万次, 体积进一步变小, 成本进一步降低, 而可靠性和性能进一步提高。软件方面, 操作系统进一步完善, 高级语言数量增多, 使电子计算机的功能越来越强。正是从第三代开始, 电子计算机进入普及阶段, 被广泛应用于科学计算、数据处理、工业控制等诸多领域。

4) 第四代大规模及超大规模集成电路计算机(1971年至今)

随着大规模、超大规模集成电路(Very Large Scale Intergrated Circuit, VLSI)技术的发展, 计算机进入大规模、超大规模集成电路时代。VLSI技术使数百万甚至数亿个晶体管能够集成在一块芯片上(如图1-5所示), 极大地提高了计算机的集成度和性能, 在体积、功耗、可靠性等方面达到了前所未有的水平。软件方面, 操作系统不断发展和完善, 同时发展了数据库管理系统、通信软件等。微型计算机和各类嵌入式系统的广泛应用使计算机几乎无处不在。此外, 随着互联网的普及和云计算、大数据、人工智能等技术的发展, 计算机已经成为推动社会进步和发展的重要力量。随着量子计算等新型计算技术的兴起, 计算机的发展将迎来更加广阔的前景。

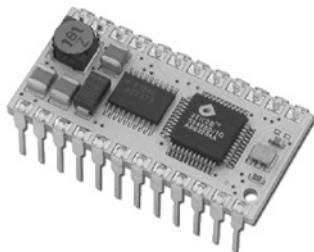


图 1-4 中小规模集成电路

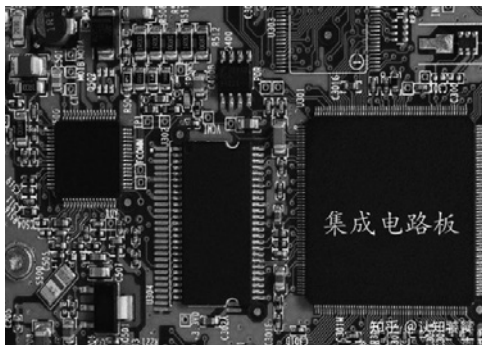


图 1-5 大规模、超大规模集成电路

1.2.2 计算机的分类

计算机种类很多, 可以从不同的角度对计算机进行分类。按照用途来划分, 可分为通用计算机和专用计算机; 按照信息的表示方式分类, 有数字计算机、模拟计算机和数模混合计算机; 按照计算机的机器字长来划分, 可分为8位、16位、32位和64位计算机; 按照规模和处理能力分类, 有超级计算机、大型计算机、小型计算机、微型计算机、工作站和服务器。

1. 超级计算机

超级计算机又称高性能计算机、巨型机, 是性能极高、功能极强、速度极快、存储量巨大、结构复杂、价格昂贵的一类计算机, 其浮点运算速度可达每秒千万亿次甚至更高, 它通常由大量处理器、高速互联网络和大容量存储器组成, 通过集群管理系统统一调度, 实现超强的集中式计算能力。主要用于数值计算(科学计算), 在气象、军事、能源、航天、探矿等领域承担大规模、高速的计算任务。

超级计算机是世界高新技术领域的战略制高点，是国家科技发展水平和综合国力的重要标志，我国是世界上少数能够生产超级计算机的国家之一，在2026年6月发布的全球超级计算机TOP500最新榜单中，中国全国产自主研发的灵晟超级计算机(如图1-6所示)以2.19 EFlops的实测持续性能位列榜首，成为世界首台持续性能超200亿亿次的超算系统，榜单前五其余席位由美国的El Capitan、Frontier、Aurora和德国的JUPITER占据。



图 1-6 灵晟超级计算机

2. 大型计算机

大型计算机又称大型机、大型主机等，是从IBM System/360开始的一系列计算机及与其兼容或同等级的计算机，主要用于大量数据和关键项目的计算，例如银行金融交易及数据处理、人口普查、企业资源规划等。这类计算机具有极强的综合处理能力和极大的性能覆盖面。在一台大型计算机中可以搭载几十块微型计算机芯片，可以同时支持上万个用户，可以支持几十个大型数据库，用以完成特定的操作，主要应用于政府部门、银行、大公司、大企业。IBM公司的Z系列大型机是业界的代表产品(如图1-7所示)，广泛应用于全球各大企业和机构。

3. 小型计算机

小型计算机又称小型机(如图1-8所示)，相对于大型计算机而言，其机器规模小、结构简单、设计试制周期短，便于及时采用先进工艺技术，软件开发成本低、可靠性高，易于操作维护。小型计算机不仅广泛应用于工业自动控制、大型分析仪器、测量设备、企业管理、大学和科研机构，还可以作为大型与巨型计算机系统的辅助计算机。



图 1-7 IBM 公司的 Z 系统大型机



图 1-8 IBM 公司的小型机

4. 微型计算机

微型计算机简称微型机、微机，又称个人计算机(PC)，俗称电脑。自1981年美国IBM公司推出第一代微型计算机IBM-PC以来，微型计算机以其执行结果精确、处理速度快捷、性价比高、轻便小巧等特点迅速渗透到社会各个领域，且技术不断更新、产品快速换代，

从单纯的计算工具发展为能够处理数字、符号、文字、语言、图形、图像、音频、视频等多种信息的强大多媒体工具。其应用已覆盖家庭、办公、日常娱乐等通用场景。

5. 工作站

工作站是一种介于微型计算机和小型计算机之间的高档微型机系统，通常配有大容量的主存、大屏幕高分辨率显示器，并具备较高的运算速度和较强的网络通信能力，具有大型计算机或小型计算机的多任务、多用户能力，兼具微型计算机的操作便利性和友好的人机交互界面，主要用于CAD工程设计、3D建模渲染、影视特效后期、科学仿真计算、金融高频数据分析等专业领域。

6. 服务器

服务器是面向多用户的高性能共享计算机，作为网络中的“大管家”，是互联网和云计算的核心基础设备。可搭载多路企业级处理器，支持可扩展TB级内存与数十块大容量硬盘，标配冗余电源、冗余风扇，搭载BMC远程管理模块，多数机型配备企业级算力加速卡。主要用于网络服务和数据存储等领域，按用途分为Web服务器、数据库服务器、云计算节点服务器等。

1.2.3 计算机智能化发展趋势

随着科技的进步，计算机技术、网络技术飞速发展，未来计算机将向着微型化、网络化、智能化和巨型化的方向发展，智能化是现代计算机技术最核心、最主流的发展方向。传统计算机依靠人工编写的固定程序被动执行指令，功能单一、适应性差。而智能化计算机以人工智能技术为核心，具备感知、认知、学习、推理、决策与自主优化能力，能够主动适配复杂场景，摆脱传统程序的固定限制。当前，计算机智能化发展主要呈现出以下几大趋势。

1. 软硬件一体化智能升级

计算机智能化不再局限于软件算法优化，而是实现硬件与软件的深度融合、协同升级。硬件层面，传统通用CPU架构逐步向异构智能计算架构迭代，NPU、TPU、AI专用芯片成为计算机、终端设备、服务器的核心配置，专门适配人工智能运算需求，大幅提升智能任务的运算效率。软件层面，生成式人工智能、大模型、机器学习算法全面植入计算机系统，使操作系统、应用软件具备自适应能力，可根据用户使用习惯、场景需求自动优化运行模式，实现从“固定执行”到“智能适配”的转变。

2. 自主学习与动态进化

智能化计算机突破了传统程序静态固化的短板，具备自主学习和动态进化的能力。依托自监督学习、小样本学习、迁移学习等技术，计算机无须大量人工编程和数据标注，即可自主从海量数据、场景案例中提取规律、积累经验。在运行过程中，系统能够持续迭代优化模型参数，主动适配新场景、新需求，不断提升任务处理精度，具备了类似人类的自适应成长能力，极大降低了人工维护与升级成本。

3. 自然化、多模态人机智能交互

人机交互模式是计算机智能化的重要体现，逐步摆脱键盘、鼠标等传统机械交互方式，向自然、高效、多维度的智能交互迭代。目前，语音识别、图像视觉、手势操控、人脸识别等交互技术已经普及。同时，多模态融合交互快速发展，计算机可同时识别文字、语音、图像、视频等多种信息，精准理解用户语义与操作意图。未来，脑机接口、沉浸式交互等技术将进一步成熟，实现人机无感、自然化的双向交互，彻底改变人机协作模式。

4. 云边端协同的全域智能普及

早期计算机智能主要依赖云端算力支撑，存在延迟高、隐私性差、依赖网络等短板。随着边缘计算技术的发展，智能化形成了云端、边缘、终端协同的全域布局。手机、智能穿戴设备、工业终端、车载计算机等各类终端设备，均可搭载轻量化智能模型，具备本地实时推理、自主决策的边缘智能能力。云端负责大数据训练、模型迭代，边缘和终端负责实时场景响应，三者协同互补，实现低延迟、高隐私、高可靠性的全域智能服务，广泛适配民生、工业、交通等各类场景。

5. 行业深度赋能，场景化智能落地

计算机智能化不再是单一的技术迭代，而是全面渗透各行业领域，实现场景化、专业化落地应用。在工业领域，智能计算机支撑智能制造、工业互联网，实现设备自主监控、故障预判、智能生产调度；在民生领域，赋能智慧医疗、智慧教育、智慧家居，提供个性化智能服务；在高端领域，支撑自动驾驶、数字孪生、航空航天智能测控等复杂场景。计算机从基础运算工具，升级为各行业的核心智能决策中枢，推动传统产业数字化、智能化转型。

6. 安全可信的智能化发展

随着计算机智能程度不断提升，智能系统的安全性、可信度成为重要发展方向。为解决人工智能算法黑箱、决策偏差、数据泄露等问题，可解释AI、隐私计算、联邦学习、智能安全监管等技术快速迭代。智能化计算机在实现自主决策的同时，能够保障数据隐私安全、决策过程可追溯、算法逻辑可解释，以此构建安全、可控、可信的智能计算体系，为智能化、规模化落地提供保障。

7. 通用人工智能持续探索

当前的计算机智能均为专用人工智能，仅能完成单一领域的特定任务。未来计算机智能化的终极方向是通用人工智能(AGI)，旨在让计算机具备人类级别的通用认知、逻辑推理、常识理解、创新创造和自主规划能力，能够跨领域处理复杂的未知场景任务，实现真正意义上的全方位智能自主，这也是未来计算机技术发展的核心探索方向。

1.3 计算机中的信息表示与人工智能数据底层基础

现实世界中的数据形式是丰富多彩的，计算机不仅能帮助人们做各种复杂的数值计算，还能处理各种文本、图形、图像、声音、视频等。电子计算机由各种逻辑器件构成，而逻

辑器件只有0和1两种状态,因此,所有为计算机存储和处理的数据,都必须转换为只用两种状态表示的数据,即由0和1组成的二进制形式。计算机中的各类数据只有采用二进制表示,才能在计算机系统中存储、加工和传输。



教学视频

1.3.1 进位计数制

1. 数制

数制也称为计数制,是指用一组固定的符号和统一的规则来表示数值的方法。

2. 进位计数制

按进位的原则进行计数,称为进位计数制。传统常见的进制是十进制(D),用0~9共10个符号表示,运算时遵循“逢十进一”的规则。计算机中采用的是二进制计数制(B),按照“逢二进一”的原则进行计数。除这些进制以外,常用的还有八进制(O)、十六进制(H)等。

3. 数位、基数和位权

在进位计数制中有数位、基数和位权3个要素。

(1) 数位:指数码在一个数中所处的位置,从右到左用0、1、2、3……表示。

(2) 基数:指在某种进位计数制中,每个数位上所能使用的数码个数。

(3) 位权:指多位数中某一位上的“1”所代表的实际数值大小,由数制基数和数位位置共同决定,表示为 r^i ,其中 r 表示基数, i 表示数位。

任意一个 m 位整数, k 位小数的 r 进制数 N 可按位权展开为:

$$N = D_{m-1} \times r^{m-1} + D_{m-2} \times r^{m-2} + \dots + D_1 \times r^1 + D_0 \times r^0 + D_{-1} \times r^{-1} + \dots + D_{-k} \times r^{-k}$$

4. 常用的进位计数制

1) 十进制

十进制数由0~9十个数码表示,其进位基数是10,运算规则是“逢十进一”,位权是 10^i 。

例如,将 $(432.45)_{10}$ 按位权展开。

$$(432.45)_{10} = 4 \times 10^2 + 3 \times 10^1 + 2 \times 10^0 + 4 \times 10^{-1} + 5 \times 10^{-2}$$

2) 二进制

二进制只有0和1两个表示符号,其进位基数是2,遵循“逢二进一”的进位规则,位权是 2^i 。

例如,将 $(1011.11)_2$ 按位权展开。

$$(1011.11)_2 = 1 \times 2^3 + 0 \times 2^2 + 1 \times 2^1 + 1 \times 2^0 + 1 \times 2^{-1} + 1 \times 2^{-2}$$

二进制虽然不符合人们日常的使用习惯,但是,计算机内部的数据采用二进制表示,其主要原因如下。

(1) 技术实现简单可靠。在计算机内可用电压的高和低两种状态来表示1和0两个数字。如果使用其他进制则需要多个电压量来表示其多个符号,其可靠性比使用二进制更容易受电压波动影响。

(2) 运算规则简单。使用二进制时，编码、计数和加减运算的规则都很简单，使计算机运算器的硬件结构大大简化。

(3) 适合逻辑运算。二进制的两个数字符号“0”和“1”正好与逻辑运算的两个值“真”和“假”相对应，为计算机实现逻辑运算和程序中的逻辑判断提供了便利条件。

但是使用二进制表达一个比较大的数值时，书写冗长，不直观。所以常常采用八进制和十六进制作为二进制数的简化表示。

3) 八进制

八进制数由0~7八个数码表示，其进位基数是8，运算规则是“逢八进一”，位权是 8^i 。

例如，将 $(425.22)_8$ 按位权展开。

$$(425.22)_8 = 4 \times 8^2 + 2 \times 8^1 + 5 \times 8^0 + 2 \times 8^{-1} + 2 \times 8^{-2}$$

4) 十六进制

十六进制数由0~9、A、B、C、D、E、F十六个数码表示，其中A、B、C、D、E、F这6个数码，分别代表十进制数中的10、11、12、13、14、15，其进位基数是16，运算规则是“逢十六进一”，位权是 16^i 。

例如，将 $(1A5E.4)_{16}$ 按位权展开。

$$(1A5E.4)_{16} = 1 \times 16^3 + 10 \times 16^2 + 5 \times 16^1 + 14 \times 16^0 + 4 \times 16^{-1}$$

1.3.2 不同进制之间的转换

1. r 进制转十进制

任何一个 r 进制数转换为十进制数时，只要“按位权展开”即可。

例如，把二进制数100110.101转换成相应的十进制数。

$$(100110.101)_2 = 1 \times 2^5 + 1 \times 2^2 + 1 \times 2^1 + 1 \times 2^{-1} + 1 \times 2^{-3} = (38.625)_{10}$$

例如，把八进制数207.5转换成相应的十进制数。

$$(207.5)_8 = 2 \times 8^2 + 0 \times 8^1 + 7 \times 8^0 + 5 \times 8^{-1} = (135.625)_{10}$$

例如，把十六进制数1A5E.4转换成十进制数。

$$(1A5E.4)_{16} = 1 \times 16^3 + 10 \times 16^2 + 5 \times 16^1 + 14 \times 16^0 + 4 \times 16^{-1} = (6750.25)_{10}$$

2. 十进制转 r 进制

将十进制数转换成 r 进制数，整数部分和小数部分的转换方法是不同的。

(1) 整数部分的转换用除 r 取余法。即把一个十进制数不断除以基数 r ，取其余数。

例如，把十进制数25转换成二进制数，计算过程如图1-9所示，所以 $(25)_{10} = (11001)_2$ 。

(2) 小数部分转换用乘 r 取整法。即将十进制小数不断乘以 r ，从上到下取其整数。

例如，将十进制数0.3125转换成相应的二进制数，计算过程如图1-10所示，所以 $(0.3125)_{10} = (0.0101)_2$ 。



图 1-9 十进制整数转二进制数



图 1-10 十进制小数转二进制数

如果十进制数包含整数和小数两部分，则必须将十进制数整数和小数部分分开转换，然后把 r 进制整数和小数部分组合在一起。

例如，将十进制数 25.3125 转换成二进制数，只需将上例整数和小数部分组合在一起即可， $(25.3125)_{10}=(11001.0101)_2$ 。

3. 非十进制数间的转换

由于二进制、八进制和十六进制之间存在特殊关系，即 $2^3=8$ ， $2^4=16$ ，根据这种对应关系，二进制数转换成八进制数(或十六进制数)十分简单。只需将二进制数以小数点为中心，分别向左、向右每 3 位(或 4 位)分成一组，不足 3 位(或 4 位)的补 0，然后根据表 1-1 即可完成转换。

表 1-1 二进制、八进制和十六进制之间的关系

二进制	八进制	二进制	十六进制	二进制	十六进制
000	0	0000	0	1000	8
001	1	0001	1	1001	9
010	2	0010	2	1010	A
011	3	0011	3	1011	B
100	4	0100	4	1100	C
101	5	0101	5	1101	D
110	6	0110	6	1110	E
111	7	0111	7	1111	F

例如，将二进制数 10100101.01011101 转换成八进制数。

010' 100' 101.010' 111' 010

↓ ↓ ↓ ↓ ↓ ↓
2 4 5. 2 7 2

所以 $(10100101.01011101)_2=(245.272)_8$

例如，将二进制数 1111111000111.100101011 转换成十六进制数。

0001' 1111' 1100' 0111' .1001' 0101' 1000

↓ ↓ ↓ ↓ ↓ ↓ ↓
1 F C 7. 9 5 8

所以 $(1111111000111.100101011)_2=(1FC7.958)_{16}$

将八进制(或十六进制)数转换成二进制数的过程正好相反。只需把每位八进制(或十六进制)数展开为3位(或4位)二进制数，再去掉首尾的0即可。

说明：八进制数与十六进制数之间的转换可以通过二进制数完成，即先将八(或十六)进制数转换成二进制数，再将转换后的二进制数转换为十六(或八)进制数。

1.3.3 数据的存储单位

在计算机中的程序和数据是按二进制的形式存放的，其度量单位如下。

1. 位

位(bit, b)是计算机存储设备的最小存储单位，存放1位二进制数，即0或1。

2. 字节

字节(Byte, B)是表示存储容量的基本单位。8个二进制位编为一组称为1B，即1B=8b。

3. 存储容量

存储容量是指存储设备所能容纳的二进制信息的总和，是衡量计算机存储能力的重要指标，通常用字节表示和计算。

计算机存储单位一般用B、KB、MB、GB、TB、PB、EB、ZB等来表示，它们之间的关系为：

1KB(Kilobyte, 千字节)= 2^{10} B=1024B

1MB(Megabyte, 兆字节)=1024KB

1GB(Gigabyte, 吉字节)=1024MB

1TB(Terabyte, 太字节)=1024GB

1PB(Petabyte, 拍字节)=1024TB

1EB(Exabyte, 艾字节)=1024PB

1ZB(Zettabyte, 泽字节)=1024EB

4. 字长

字长是CPU在单位时间内一次处理的二进制的位数，是衡量计算机性能的一个重要指标。不同的CPU字长是不同的，常用的字长有8位、16位、32位、64位等。字长越长，一次处理的二进制位数越多，速度也就越快。

1.3.4 计算机的信息编码

1. 数字编码

计算机中用于存储、运算与表示数据的核心数字编码主要分为三类：原码、反码、补码，部分场景会用到移码，所有编码均以二进制为基础，是计算机处理数值数据的核心方式。

1) 原码

原码是最直观、最贴近人类读写习惯的二进制编码。最高位为符号位，0表示正数、

1 表示负数，其余数位为数值本身的二进制值。例如对于8位二进制数，+11的原码为00001011，-11的原码为10001011。原码的优点是简单易懂、直观，缺点是不能直接参与运算，因为直接使用原码进行加减运算可能产生错误的结果。

2) 反码

反码是数字存储的一种表示方法，多应用于系统环境设置。反码的编码规则是：正数的反码与原码相同，负数的反码是原码除符号位外按位取反。例如，对于8位二进制数，-1的原码是10000001，则-1的反码为11111110。反码通常用作由原码求补码或由补码求原码的过渡码，本身不直接参与运算。

3) 补码

补码是计算机中用于存储数字的编码方式，它将减法运算转化为加法运算，简化了计算机的运算过程。补码的编码规则是：正数的补码和原码相同，负数的补码是其反码末位加1。例如，对于8位二进制数，-1的补码是11111111。补码解决了原码和反码在运算中可能产生的问题，是计算机内部表示数值的主要方式。

4) 移码

移码主要用于浮点数阶码存储(计算机浮点数运算专用)，通过整体偏移固定数值实现，多用于科学计算、浮点运算场景，日常整数运算不使用。

2. 字符编码

1) ASCII码

计算机中使用最多、最普遍的字符编码是ASCII码(American Standard Code for Information Interchange，美国信息交换的标准代码)，该标准总共规定了128(2⁷)个符号所对应的数字编码，每个编码使用7位二进制数来表示。其中包含英文的大小写字母、数字、标点符号等常用字符，数字代号从0至127，如表1-2所示。ASCII满足了早期计算机处理英文文本的需求，但无法表示其他语言(如中文、日文、韩文等)的字符。

表 1-2 ASCII 码表

d ₃ d ₂ d ₁ d ₀	d ₆ d ₅ d ₄ 位							
	000	001	010	011	100	101	110	111
0000	NUL	DLE	SP	0	@	P	`	p
0001	SOH	DC1	!	1	A	Q	a	q
0010	STX	DC2	“	2	B	R	b	r
0011	ETX	DC3	#	3	C	S	c	s
0100	EOT	DC4	\$	4	D	T	d	t
0101	ENQ	NAK	%	5	E	U	e	u
0110	ACK	SYN	&	6	F	V	f	v
0111	BEL	ETB	‘	7	G	W	g	w
1000	BS	CAN	(	8	H	X	h	x
1001	HT	EM	)	9	I	Y	i	y
1010	LF	SUB	*	:	J	Z	j	z
1011	VT	ESC	+	;	K	[	k	{
1100	FF	FS	,	<	L	\	l	

(续表)

$d_3d_2d_1d_0$	$d_6d_5d_4$ 位							
	000	001	010	011	100	101	110	111
1101	CR	GS	-	=	M	]	m	}
1110	SO	RS	.	>	N	^	n	~
1111	SI	US	/	?	O	_	o	DEL

2) Unicode编码

为了解决多语言字符编码的问题，Unicode应运而生。Unicode是一种国际标准，旨在为世界上所有的字符和符号分配唯一的编码。Unicode使用多字节(通常是2字节或更多)来表示一个字符，从而能够表示几乎所有的字符和符号。Unicode的出现极大地促进了国际信息交流和文化传播。

3) UTF-8 编码

UTF-8(8-bit Unicode Transformation Format)是一种针对Unicode的可变长度字符编码方式。UTF-8用1~4字节来表示一个Unicode字符，其中，常用的英文字符用1字节表示，中文字符则用多字节表示。UTF-8的优点在于其兼容ASCII(即ASCII是UTF-8的一个子集)，从而可以方便地在只支持ASCII的环境中传输和存储Unicode文本。

3. 汉字编码

用计算机处理汉字时，必须将汉字代码化，即对汉字进行编码。汉字是象形文字，字符数量多，各个笔画之间差异巨大，即使在同一个汉字处理系统中，输入、内部处理、输出对汉字代码的要求也不尽相同，像英文字符那样在输入、内部处理、存储和输出环节都使用一个代码是不现实的。因此，需要支持多种编码以解决汉字在输入、输出和处理方面的各种问题。汉字常用编码有输入码、国标码、机内码和字形码。

1) 输入码

在计算机系统中使用汉字，首先遇到的问题是把汉字输入计算机内。为了能直接使用西文标准键盘进行输入，必须为汉字设计相应的编码方法，汉字输入编码方法主要有三类：数字编码、拼音编码和字形码。

2) 国标码

国标码指国家标准汉字编码，是用于不同的具有汉字处理功能的计算机系统间交换汉字信息时使用的编码。原由国家标准总局发布的GB/T 2312—1980《信息交换用汉字编码字符集 基本集》收录了7445个字符(包括6763个汉字和682个其他字符)，适用于汉字处理、汉字通信等系统之间的信息交换。

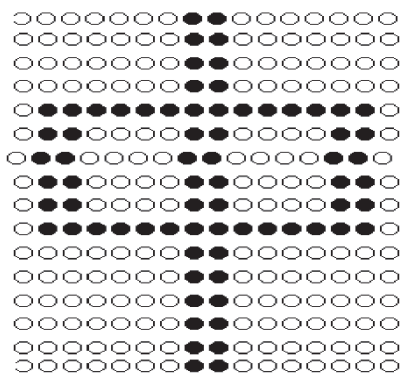
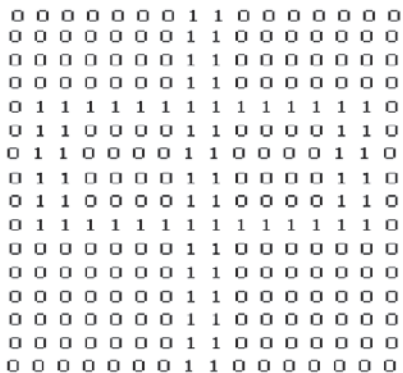
3) 机内码

机内码是字符在设备或信息处理系统内部最基本的表达形式，是在设备和信息处理系统内部存储、处理、传输字符用的代码。现在我国的汉字信息系统一般用两个字节构成一个汉字机内码，为了和英文字符区分，通常将国标码中每字节的最高位设置为1来表示机内码。

4) 字形码

汉字字形码(字模码)用于显示或打印汉字时产生字形。字形码有点阵方式字形码和矢量方式字形码两种。一个汉字信息系统的所有汉字字形码的集合构成汉字库。对点阵方式

字形码而言,根据输出汉字的质量要求不同,点阵的规格也不同。汉字点阵越多,输出的质量越高。汉字点阵的规格以横向点数乘纵向点数表示。目前计算机中普遍采用 16×16 、 24×24 、 32×32 、 48×48 的汉字点阵,图1-11、图1-12所示为“中”字在 16×16 点阵中的字形表示和字形编码。不同字体的汉字拥有不同的字库。

图 1-11 16×16 点阵字形图 1-12 16×16 字形编码

4. 音频编码

计算机中音频的表示主要基于数字化技术。音频信息首先被模拟成电信号,然后通过采样、量化和编码等步骤转换为数字信号。采样是将连续的模拟信号转换为离散的数字信号的过程,通过以一定的频率(如44.1kHz)进行。量化则是将采样得到的模拟电平转换为有限数量的数字电平。编码是将量化后的数字电平按照一定的规则(如MP3、WAV等格式)进行压缩和编码,以便存储和传输。这种数字化的音频表示方法使音频信息在计算机中以二进制形式存储和处理,从而实现高效的音频编辑、播放和传输。

5. 图像编码

图像(Image)也称为位图,是由像素点阵构成的栅格图。每个像素用若干个二进制位来表示其颜色和亮度信息。图像以像素为基本单位,分辨率(通常以“像素/英寸”为单位)决定了图像的清晰度和细节。图像文件占用的存储空间相对较大,适合用来描述照片、复杂场景等含有大量细节的对象。然而,图像在放大时可能会失去细节或产生锯齿现象。

图像的编码主要分为有损编码和无损编码两大类。有损编码(如JPEG)通过舍弃部分不重要的图像信息以实现高压缩比,解码后的图像与原始图像在视觉上可能有细微差异,但通常可接受。无损编码(如PNG)可以保留图像所有信息,解码后的图像与原始图像完全相同,但压缩效率较低。每种编码方式都有其特点和适用场景,需要根据具体需求选择。

6. 视频编码

视频编码是将连续的图像序列(即视频)通过特定的算法转换为数字信号,并进行压缩处理,以减少其对存储空间和传输带宽的需求。以下是两种主要的视频编码方式。

1) 有损编码

这是目前应用广泛的视频编码方式,通过去除视频中的冗余信息(包括空间冗余和时间冗余)来实现压缩,但会在一定程度上牺牲视频质量。常用的有损编码包括MPEG系列

(如MPEG-1、MPEG-2、MPEG-4等)、H.264(也称AVC, 即高级视频编码)和H.265(也称HEVC, 即高效视频编码), 它们能够在保证视频质量可接受的前提下大幅减小文件。

2) 无损编码

与有损编码不同, 无损编码方法不会丢失视频中的任何信息, 因此能够保持视频的最高质量。然而, 由于不进行压缩或压缩率极低, 无损编码的视频文件通常会比有损编码的文件大得多。常用的无损编码有DV、DVCPRO等。

1.3.5 人工智能数据底层基础

人工智能(AI)并非单纯的算法与程序模型, 其本质是以数据为“燃料”、以算力为支撑、以算法为规则的智能运算系统。如果说算法是人工智能的“大脑思维逻辑”, 那么数据就是人工智能的“认知基础与学习素材”。所有AI智能效果, 包括图像识别、语音交互、自然语言处理、智能预测等, 全部建立在规范、海量、高质量的数据底层支撑之上。

1. 人工智能数据底层逻辑

传统计算机程序遵循“人工编写规则、机器机械执行”的逻辑, 机器只能完成人类预设的指令任务, 不具备自主学习能力。而人工智能的核心突破在于机器学习, 即机器通过海量数据自主总结规律、迭代模型、优化精度, 无须人类逐条编写所有场景规则。

简单来说, 传统程序靠代码定义结果, 人工智能靠数据训练结果。没有数据, 人工智能模型只是一套无意义的数学公式, 无法实现任何智能功能。数据的数量、质量、规整度, 直接决定了人工智能模型的准确率、稳定性和适用场景。

从底层逻辑看, 人工智能数据具备三大核心作用。

(1) 建模依据。为人工智能模型提供学习样本, 让模型总结不同场景下的特征规律, 例如通过数万张人脸图片学习人脸五官、轮廓等特征。

(2) 迭代标尺。通过标注数据校验模型输出结果, 不断修正模型参数, 降低识别误差、预测偏差。

(3) 场景适配。多样化、场景化的数据让模型具备泛化能力, 能够适配复杂真实场景, 而非仅适配单一测试环境。

2. 人工智能数据的核心分类

在计算机底层架构中, 根据数据形态、结构和用途, 可将人工智能数据分为三类, 覆盖绝大多数人工智能应用场景, 是数据处理的基础分类标准。

1) 结构化数据

结构化数据是格式固定、维度统一、可直接录入数据库的标准化数据, 具备严格的字段规范, 计算机可直接解析运算, 是传统数据分析与轻量人工智能预测的核心数据。常见形式包括表格数据、数据库字段数据, 如用户年龄、身高、消费记录、设备参数、时间数值等。

这类数据底层存储规整、读取速度快、计算成本低, 广泛应用于金融风控、用户画像、数据预测等AI场景。

2) 非结构化数据

非结构化数据是无固定格式、无统一字段、体量庞大的原始数据，也是当前人工智能训练的核心数据来源。这类数据无法直接通过传统数据库解析，需要经过底层转换、特征提取才能被人工智能模型识别。

常见的非结构化数据包括：图片、视频、音频、自然语言文本、表情包、手写字体等。日常生活中90%以上的人工智能原始数据均为非结构化数据，例如人脸识别的图像数据、大模型训练的文本数据、语音助手的音频数据。

3) 半结构化数据

半结构化数据介于结构化与非结构化之间，具备基础标记框架，但无统一固定字段，可灵活拓展，是互联网AI场景的重要数据类型。常见格式为JSON、XML，多用于网页数据、接口传输数据、日志数据等，需要通过解析框架提取有效特征后，方可用于人工智能训练。

3. 人工智能数据的底层处理流程

原始的采集数据存在大量冗余、错误、无效信息，无法直接用于模型训练。人工智能的学习过程，本质是数据从原始杂乱状态到标准化可用状态的处理过程，底层通用流程分为四步，是所有人工智能项目的基础链路。

1) 数据采集

通过传感器、爬虫、数据库、设备终端、公开数据集等渠道获取原始数据，完成数据的初步归集。例如通过摄像头采集图像、麦克风采集音频、服务器采集用户行为日志。该环节的核心在于保证数据来源全面、场景覆盖完整。

2) 数据清洗

清洗是人工智能数据底层最关键的预处理操作，主要包括剔除无效数据、修正错误数据、补全缺失数据，去除噪声干扰。例如删除模糊无效的图片、修正文本乱码、补齐缺失的数值字段、过滤重复数据。数据清洗的质量，直接决定人工智能模型是否会出现“学习偏差”。

3) 数据标注

标注是赋予原始数据“语义标签”的过程，让机器知道数据对应的真实含义，是监督学习类人工智能的核心底层支撑。例如给图片标注“人脸/车辆/树木”、给文本标注“正面/负面情绪”、给语音标注“对应文字内容”。未经标注的数据，对人工智能而言只是无意义的二进制点阵、像素波形，无法完成学习。

4) 数据归一化与特征提取

计算机底层无法直接理解图像、文字、音频的语义，需要将所有数据转化为标准化数字矩阵、向量数值，这一过程即为归一化与特征提取。比如将图片的像素点阵转化为数字数组、将文本语义转化为词向量，统一数据维度与取值范围，降低模型运算难度，提升训练效率。

4. 数据、算法、算力的底层协同关系

人工智能系统的底层运行，依靠数据、算法、算力三者闭环协同，三者缺一不可。

数据是基础，提供学习素材，决定人工智能的认知边界；算法是规则，搭建数学模型，

定义数据的学习、运算、迭代逻辑；算力是动力，依托CPU、GPU、AI芯片，完成海量数据的矩阵运算、模型迭代，没有算力支撑，海量数据无法完成训练。

5. 底层数据的核心特性要求

用于人工智能训练的数据，必须满足四大底层特性，否则会导致模型精度低、泛化能力差、运行不稳定。

- (1) 海量性。充足的数据量可避免模型过度拟合，让模型充分学习场景规律。
- (2) 真实性。数据必须来源于真实场景，虚假数据会直接导致人工智能学习错误规律。
- (3) 多样性。覆盖不同场景、不同状态的数据，保证人工智能适配复杂环境。
- (4) 规范性。经过清洗、归一化处理，格式统一、维度标准，适配模型运算逻辑。

1.4 计算机系统组成与智能硬件架构



教学视频

1.4.1 计算机系统组成

一个完整的计算机系统由硬件系统和软件系统两部分组成。硬件系统是指客观存在的物理实体，即由电子元件和机械元件构成的各个部件。软件系统是指运行在硬件上的程序、运行程序所需的数据以及相关文档的总称。硬件为软件提供了运行平台，软件使硬件的功能充分发挥，两者相互配合才能完成各项功能。计算机系统的层次结构如图1-13所示。

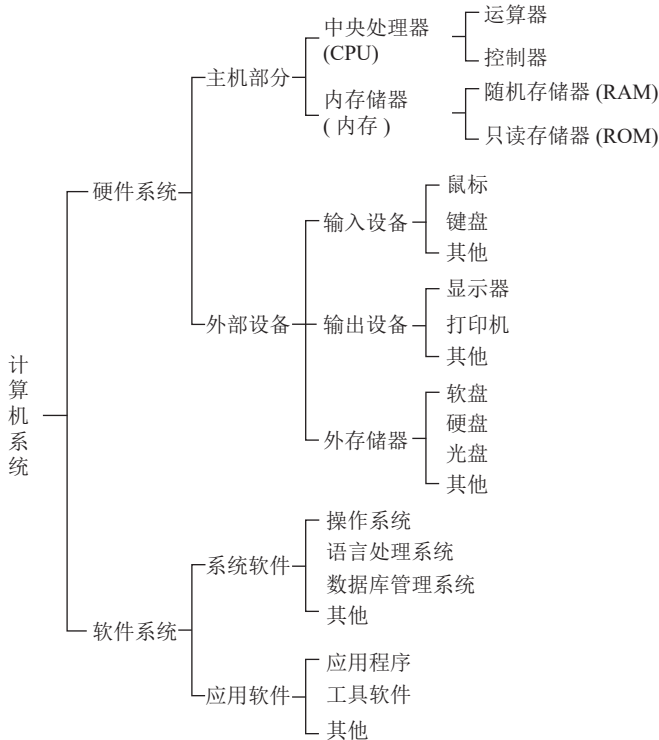


图 1-13 计算机系统的层次结构

1. 计算机的硬件系统

计算机硬件系统的基本功能为在程序的控制下完成数据输入、数据存储、数据处理和数据输出等操作。计算机的硬件系统主要包括主机和外部设备两大部分。

1) 主机

主机由中央处理器(CPU)和内存存储器(主存)组成,用来执行程序、处理数据。

(1) 中央处理器。中央处理器又称为CPU(Central Processing Unit),是由大规模集成电路构成的,集成在小硅片上(如图1-14所示)。CPU是计算机的核心部件,负责执行指令和处理数据。CPU由运算器和控制器组成,能够执行算术逻辑运算、控制程序指令的执行顺序,并管理计算机系统中的各种资源。CPU的性能直接决定了计算机的运行速率和效率,是现代计算机技术发展的重要标志之一。随着技术的发展,CPU的运算能力和功能不断增强,满足了日益增长的计算需求。

(2) 内存存储器。计算机的内存存储器(简称内存或主存)是计算机的核心组成部分(如图1-15所示),作为CPU和硬盘之间的高速桥梁,用于临时存储正在运行的程序和数据。它主要包括随机存储器(Random Access Memory, RAM)和只读存储器(Read-Only Memory, ROM)两种类型。RAM是一种可以随机读写数据的存储器,也称为读写存储器,既可以读出,也可以写入。RAM具有易失性,即只能用于暂时存放信息,一旦断电,存储内容立即消失。ROM存储的信息只能执行读取操作,不能执行写入修改等操作,它是一种非易失性存储器,一旦写入信息后,无须外加电源来保存信息,不会因断电而丢失。原来存储的内容是采用掩膜技术由厂家一次性写入的,并永久保存下来。它一般用来存放专用、固定的程序和数据。

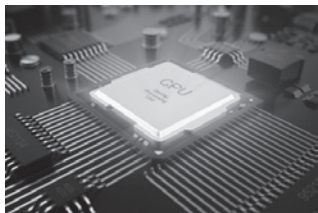


图 1-14 CPU

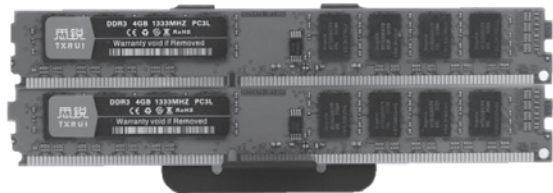


图 1-15 内存

2) 外部设备

外部设备包括输入设备、输出设备和外存储器(辅助存储器)。

(1) 输入设备。输入设备将数据和程序转化为电信号输入计算机中,人们通过输入设备操作和控制计算机。典型输入设备包括键盘、鼠标、麦克风、手写输入笔等。其中键盘是微型计算机必备的标准输入设备,用户的程序、数据以及各种对计算机的命令都可以通过键盘输入。标准键盘是最常见的键盘,各厂商的标准键盘无论从尺寸、布局还是外形上来看,都大同小异。鼠标也是一种输入设备,能方便、快速、精确地将光标定位到屏幕指定位置,以完成各种有关操作。鼠标接口大多采用PS/2接口、USB接口或无线接口。

(2) 输出设备。输出设备将计算机中的运算结果以人们能够识别的形式打印或显示出来。输出设备包括显示器、打印机、绘图仪、音箱等。显示器主要包括阴极射线显示器和液晶显示器,目前市场上主流的显示器是液晶显示器。打印机包括针式打印机、激光打印机和喷墨打印机等。针式打印机通过打印头中的钢针打击色带,在打印纸上以点阵形式构

成字符，速度慢、打印质量差，且打印过程中噪声大。激光打印机的印刷原理类似于静电复印，输出速度快、打印质量高。喷墨打印机将墨水直接喷到纸上实现印刷，印刷质量好，但速度和质量比不上激光打印机。绘图仪主要用于一些工程设计图的输出。

(3) 外存储器(外存)。外存储器一般用来存储需要长期保存的各种程序和数据。这种存储器不能被CPU直接访问，需要先调入内存才能被CPU操作利用。外存和内存相比，外存存储容量大，能长期保留信息，但运行速度慢。外存储器主要包括硬盘、光盘、U盘、存储卡等。

硬盘是计算机的主要存储设备，用于永久存储数据和程序，分为机械硬盘(Hard Disk Drive, HDD)和固态硬盘(Solid State Disk, SSD)等类型，如图1-16和图1-17所示。HDD成本低廉、容量大，而SSD读写速度快、耐用性强，造价高。硬盘按照电磁转换原理工作，将数据存储在盘片的磁性物质上，是现阶段计算机不可或缺的重要部件。



图 1-16 机械硬盘



图 1-17 固态硬盘

光盘是一种利用激光原理进行读写的存储设备，以光信息作为存储的载体并用来存储数据，如图1-18所示。光盘可分为不可擦写光盘和可擦写光盘两大类。不可擦写光盘有CD-ROM、DVD-ROM等，数据写入后不能更改；可擦写光盘有CD-RW、DVD-RAM等，支持数据的反复读写。光盘具有存储密度高、存储容量大、携带方便等优点，广泛应用于音乐、视频、软件等多媒体数据的存储和分发。

U盘是USB内存盘的简称，也被称为优盘，主要使用闪存芯片进行存储，如图1-19所示。U盘具有体积小、质量轻、携带方便、即插即用等优点，广泛应用于个人数据存储、文件传输、系统安装等领域。随着USB接口技术的不断升级(如USB 3.0、USB 3.1等)，U盘的读写速度也不断提升。



图 1-18 光盘



图 1-19 U 盘

2. 计算机的软件系统

计算机软件由相关程序和有关文档组成，通常计算机软件分为系统软件和应用软件两大类。

1) 系统软件

系统软件是面向计算机硬件本身、用于管理计算机资源、为上层软件提供运行环境、支撑软硬件协同工作的基础软件，是保障计算机正常、稳定、高效运行的核心底层软件。常用的系统软件有操作系统、程序设计语言与语言处理程序、数据库管理系统和系统辅助软件等。

(1) 操作系统。操作系统(OS)是最基础、最核心的系统软件，是计算机硬件与软件、用户与计算机之间的接口。它统一管理计算机的硬件资源(CPU、内存、外存、输入输出设备)和软件资源，合理调度程序运行，为各类应用软件提供统一的运行平台。

操作系统种类繁多、各有特色，适用于不同的计算场景和需求，主要有Windows系列(Windows 10、Windows 11)、UNIX/Linux系列、macOS、移动操作系统(iOS、Android、鸿蒙操作系统)、云操作系统(Microsoft Azure、alibaba Cloud、中标麒麟安全云、华为欧拉云)等。

(2) 程序设计语言与语言处理程序。程序设计语言与语言处理程序是系统软件的重要组成部分，它们共同支持软件开发的全过程。程序设计语言是人与计算机之间进行信息交流的工具，它允许人们以特定的方式编写指令，这些指令可以被计算机理解并执行。程序设计语言分为机器语言、汇编语言和高级语言三种。

- 机器语言在计算机内部用二进制代码指令来表示各种数据和指令。机器语言能够被计算机直接识别和执行，执行速度快、占用内存少；但它难以记忆、编写、调试和不易移植。
- 汇编语言是将机器语言中的指令用便于记忆的符号表示出来的一种语言，它比机器语言中的指令简短，容易记忆。汇编语言是为特定的计算机设计的面向机器的符号化程序设计语言，通常用于底层硬件操作和高要求的程序优化。
- 高级语言是一种更接近于自然语言和数学语言的程序设计语言，这种语言接近日常用语，对机器的依赖较低，相比汇编语言更加直观，易于编写和修改等。常用的高级语言有C语言、Python语言、Java语言等。

汇编语言和高级语言编写的程序无法直接在计算机中执行，需要用语言处理程序把它翻译为机器语言才能执行。语言处理程序是指将用程序设计语言编写的源程序转换为机器语言的程序，以便计算机能够运行。根据处理方式不同，语言处理程序可以分为编译程序、解释程序和汇编程序三大类。编译程序是一种将用高级程序设计语言编写的源程序，整体翻译成等效的机器语言目标程序的翻译程序，其工作过程包括词法分析、语法分析、语义分析、中间代码生成、代码优化和目标代码生成等阶段。解释程序是将源程序边翻译成机器代码边执行的高级语言程序。解释程序逐行读取源程序，进行词法分析、语法分析和语义分析，并直接执行源程序，不生成目标程序。这种方式便于测试，但执行效率相对较低。汇编程序将汇编语言源程序翻译成机器指令程序，这些指令被CPU直接识别并执行。

(3) 数据库管理系统。数据库管理系统(Database Management System, DBMS)是一种操纵和管理数据库的大型软件，用于建立、管理、组织和维护数据库，对数据库进行统一管理和控制，以保证数据库的安全性和完整性。可实现数据的定义、存储、查询、修改、备份、权限管理等功能，保障海量数据的有序管理与高效调用。

在人工智能数据底层体系中，结构化训练数据、用户样本数据、模型参数数据大多通过数据库管理系统进行存储管理。常见软件包括MySQL、Oracle、SQL Server、GaussDB

(华为)、OceanBase(阿里巴巴)、TDSQL(腾讯)等,是人工智能数据预处理、数据归集的核心工具。

(4) 系统辅助软件。系统辅助软件又称工具软件,用于维护计算机系统、优化运行状态、保障软硬件正常工作,为系统运行和数据处理提供辅助支撑。常见类型包括磁盘清理工具、文件管理工具、驱动程序、系统监控软件等,可完成硬件驱动适配、系统垃圾清理、运行故障排查等工作,保障人工智能数据处理、模型训练过程稳定运行。

2) 应用软件

应用软件是为了满足用户特定需求而开发的,它们直接为用户提供服务,涵盖多个领域和功能,如办公软件WPS、Microsoft Office等,图像处理软件Adobe Photoshop,视频剪辑软件会声会影、剪映、快剪辑等,社交软件微信、QQ等,安全软件360卫士、金山毒霸……。应用软件种类繁多,以上只是一部分。随着信息技术的发展,新的应用软件将不断涌现,它们将为人们的生活和工作带来更多的便利与可能性。

1.4.2 计算机工作原理

计算机的工作原理即“存储程序”原理,它是由美籍匈牙利数学家冯·诺依曼于1946年首先提出的。他将计算机的工作原理总结为:编好的程序和原始数据,输入并存储在计算机的存储器中,按照程序逐条取出指令加以分析,执行指令规定的操作。虽然现在计算机设计和技术有了很大的发展,但是基本结构和思想仍然采用冯·诺依曼体系结构,计算机的工作原理如图1-20所示。

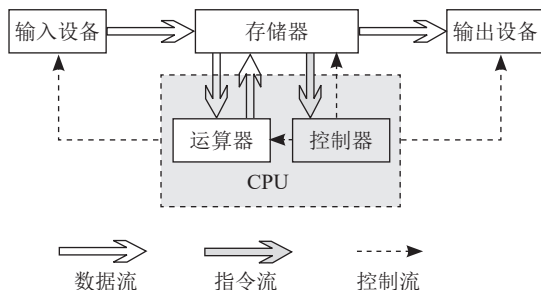


图 1-20 计算机的工作原理

1.4.3 智能硬件架构

智能硬件架构是各类智能终端设备的底层硬件整体组织结构与运行逻辑规范,是设备实现数据处理、环境感知、智能控制与网络交互的物理基础。该架构继承了冯·诺依曼计算机体系的核心工作原理,同时针对智能设备微型化、低功耗、场景化、嵌入式运行的应用特点进行优化升级,摒弃了传统计算机冗余的硬件设计,形成模块化、可裁剪、高适配的硬件体系结构。完整的智能硬件架构主要由核心处理单元、存储单元、感知交互单元、供电与通信单元四大功能模块组成,各模块独立分工、协同工作,有序完成数据采集、运算处理、信息存储、指令执行及数据传输的完整工作流程。

1. 核心处理单元

核心处理单元是智能硬件的核心控制与运算中枢,相当于设备的“大脑”,核心器件

为嵌入式微控制器、微处理器等专用芯片。其主要功能为识别并解析程序指令、完成各类数据运算、统筹调度所有硬件模块资源，根据程序逻辑做出智能判断与控制决策，直接决定智能设备的运行速度、处理能力与智能化水平，是设备实现各类自定义功能与智能算法的核心载体。

2. 存储单元

存储单元是智能硬件的数据与程序存储载体，承担设备运行记忆与数据留存的作用，主要分为运行内存与非易失性存储两类。运行内存为设备运行提供临时数据缓存空间，实时加载运行程序、存储即时运算数据，断电后数据自动清空，直接影响设备运行的流畅度；非易失性存储用于长期保存系统程序、应用程序、用户数据、配置参数等内容，断电后数据不会丢失，保障设备开机自启、参数留存等基础功能稳定实现。

3. 感知交互单元

感知交互单元是智能硬件与外界信息交互的接口，对应设备的“感知与执行系统”，由输入感知器件与输出执行器件组成。输入感知器件包含温度、湿度、红外、压力等各类传感器及物理按键，负责采集外界环境信息与用户操作指令；输出执行器件包含显示屏、指示灯、电机、继电器等设备，负责反馈运行状态、执行控制指令，最终实现设备与人、设备与环境的双向交互，是智能设备区别于传统硬件的核心模块。

4. 供电与通信单元

供电与通信单元是智能硬件稳定运行的保障模块。供电单元通过稳压、降压电路为整机硬件提供持续、稳定的电力供应，适配嵌入式设备低功耗运行需求；通信单元依托蓝牙、Wi-Fi、串口、物联网模组等通信方式，实现设备之间、设备与云端、设备与终端的互联互通，完成数据上传、指令接收等远程交互功能。相较于传统计算机硬件架构，智能硬件架构具备高度的灵活性与可裁剪性，可根据智能家居、可穿戴设备、工业智能终端、车载设备等不同应用场景，增减硬件模块、调整硬件配置，精准适配各类智能化应用场景。

1.5 计算机网络基础

当今计算机时代，计算机网络已成为连接世界、促进信息交流与资源共享的重要基础设施。无论是日常生活中的网络购物、在线学习，还是企业间的数据交换、远程办公，都离不开计算机网络的支撑。

1.5.1 计算机网络概述

1. 计算机网络概念及发展

计算机网络是现代信息技术的重要组成部分，它是指将多个地理位置不同、功能独立的计算机系统通过通信设备和通信线路相互连接，以实现资源共享和信息交换的系统。具体来说，计算机网络通过路由器、交换机等网络设备，将不同地理位置的计算机终端连接

起来，形成一个覆盖范围广泛的网络体系。这个网络体系不仅能够实现数据的传输和共享，还能够支持各种网络服务和应用，如电子邮件收发、网页浏览、文件传输、远程登录等。计算机网络的出现极大地提高了数据处理能力和资源利用效率，推动了社会的信息化进程。

计算机网络主要由硬件和软件两大部分构成。硬件部分包括计算机、路由器、交换机、网卡、传输介质等，它们构成了计算机网络的基础设施。软件部分包括网络操作系统、网络通信协议、网络应用软件等，它们为计算机网络中的用户提供各种服务和支持。

计算机网络的发展经历了多个阶段，从最初的面向终端的计算机网络，到后来的计算机互联网络，再到现在的开放式标准化计算机网络和移动互联网，其功能和性能都得到了极大提升。如今，计算机网络已成为人们日常生活和工作不可或缺的一部分，为信息的快速流通和资源的有效利用提供了有力支持。

2. 计算机网络类型

根据不同的划分标准，计算机网络可以划分为多种类型。

1) 按覆盖范围划分

按覆盖范围划分，可分为局域网、城域网、广域网和互联网。局域网通常局限于一个较小的地理区域，如办公室或校园；城域网覆盖一个城市；广域网跨越多个城市或国家；互联网是全球最大的广域网，连接了全球范围内的无数网络，它诞生于1969年，是由美国的阿帕网发展而来的。

2) 按拓扑结构划分

按拓扑结构划分，可划分为总线拓扑、星形拓扑、环形拓扑、网状拓扑等计算机网络。拓扑结构描述了计算机网络中各节点之间的连接方式，不同的拓扑结构对计算机网络的性能、可靠性和成本有不同的影响。

3) 按传输技术划分

按传输技术划分，可划分为广播网络和对等网络。广播网络中，信息从源节点发送到所有节点，再由各节点根据需要进行处理；对等网络中，信息直接从一个节点发送到另一个特定节点。

3. 计算机网络的关键要素

计算机网络包括以下关键要素。

1) 节点

节点指网络中的计算机、交换机、路由器等设备，它们是信息传输的起点和终点。

2) 链路

链路指连接节点的物理或逻辑通路，它可以是光纤、双绞线、同轴电缆等物理介质，也可以是无线信号传输的虚拟路径。

3) 协议

协议是网络通信双方必须遵循的一系列规则和标准，用于确保信息传输的准确性和可靠性。TCP/IP是互联网中的基本协议集。

4) IP地址

IP地址是互联网中每台设备的唯一标识，用于在网络中定位和识别设备。它由32位二

进制数组成(IPv4), 通常以点分十进制形式表示, 如192.168.4.1。随着IPv6的到来, IP地址长度扩展到128位, 这极大地扩展了地址空间。IP地址是网络通信的基础, 确保数据能够准确、高效地传输到目标设备。

5) 域名

域名是互联网中的一种标识符, 由一串用点分隔的名称组成, 以方便人们记忆和访问网站。例如, 清华大学的网站域名是www.tsinghua.edu.cn。域名系统(Domain Name System, DNS)将域名和IP地址相互映射, 会自动把域名转换为对应的IP地址, 使人们输入域名即可访问网站, 提高了互联网的易用性。

4. 计算机网络的接入方式

计算机网络的接入方式多种多样, 每种方式都有其特点、使用时期和适用场景。以下是几种主要的接入方式。

1) 拨号方式

拨号方式是20世纪90年代的计算机网络接入方式, 通过调制解调器和电话线连接到因特网服务提供方, 获取上网账号后进行拨号上网。这种方式简单易行, 但速率较低, 已被淘汰。

2) ADSL接入

ADSL技术是指利用现有电话线路进行高速数据传输。它将普通电话线路的低频信号和高频信号分离, 实现高速上网的同时不影响电话通话。这种接入方式成本低廉, 数据传输速率较高。ADSL在20世纪90年代末至21世纪10年代得到了广泛的推广和应用, 它以较快的上网速度和稳定的连贯性能逐渐取代了传统的拨号接入方式, 成为当时主流的宽带接入技术。尽管ADSL技术在当时具有显著的优势, 但随着光纤宽带等新技术的不断发展, ADSL逐渐被替代, 目前仅有一些偏远地区或经济条件相对较差的地区在使用。

3) 光纤宽带接入

光纤作为传输介质, 将光信号转换为电信号进行数据传输。光纤宽带接入方式具有传输率高、带宽大、稳定性好等优点, 是目前企业和家庭用户追求高速网络体验的首选接入方式。目前, 我国大多数家庭的宽带网络都已经实现“光纤入户”。

4) 无线接入

无线接入包括Wi-Fi、4G/5G移动网络等。用户通过无线设备连接到无线网络, 实现随时随地上网。无线网络接入具有灵活性强、覆盖范围广等优点, 但安全性和稳定性欠佳。

5) 专线接入

专线接入为企业 provide 高速、稳定、安全的网络接入方式。专线接入通常采用光纤作为传输介质, 提供独享的带宽资源, 确保数据传输的可靠性和实时性。这种方式适用于对数据传输要求较高的企业用户。

1.5.2 计算机网络服务

1. 计算机网络服务概述

计算机网络服务是指在计算机网络上提供的各种服务和功能。计算机网络服务涵盖数

据的传输、存储、检索以及通信等多个方面，以下是几项具有代表性的计算机网络服务。

1) 文件传输

FTP软件是用于在计算机网络上传输文件的工具，它支持文件的上传和下载。用户可以通过FTP软件连接FTP服务器，实现文件的远程管理和共享。常见的FTP软件包括FileZilla、WinSCP、CuteFTP等，它们提供了友好的用户界面和丰富的功能选项。

2) 电子邮件

电子邮件是互联网上广泛使用的通信方式之一，用于个人、企业和组织之间的信息交流。

3) 网页浏览

互联网上常用的服务之一，用户可以通过浏览器获取各种信息和服务。

4) 文件共享

允许用户在网络上共享文件和资源。

5) 远程登录

允许用户通过网络远程登录到另一台计算机，流行的远程登录工具包括Secure Shell(SSH)和Telnet等。

6) 虚拟专用网络

提供加密和安全的网络连接，使用户能够通过公共网络安全地访问私有网络资源。

7) 即时消息传递

允许用户通过网络进行语音和视频通话，如腾讯会议等。

8) 云存储

提供在线存储和数据备份服务，如百度网盘等。

2. 网页

网页是互联网信息展示与交互的基本单元，是用户与网站之间沟通的桥梁。它由一系列的文件组成，其中最主要的是超文本标记语言(HyperText Markup Language, HTML)文档，这些文档通过特定的标签和元素定义网页的结构与内容。HTML文档中的文本、图片、视频、链接等元素共同构成了用户在浏览器中看到的网页内容。下面是网页文件demo.html的HTML源代码。

```
<html>
  <head><title>搜索指数</title></head>
  <body>
    <table>
      <tr><td>排名</td><td>关键词</td><td>搜索指数</td></tr>
      <tr><td>1</td><td>大数据</td><td>187767</td></tr>
      <tr><td>2</td><td>云计算</td><td>178856</td></tr>
      <tr><td>3</td><td>人工智能</td><td>134587</td></tr>
      <tr><td>4</td><td>物联网</td><td>134587</td></tr>
    </table>
  </body>
</html>
```

使用网页浏览器(如IE、Firefox)打开这个网页文件，就会看到图1-21所示的网页。



排名	关键词	搜索指数
1	大数据	187767
2	云计算	178856
3	人工智能	134587
4	物联网	134587

图 1-21 网页文件显示效果

访问网页时会用到超文本传输协议(HyperText Transfer Protocol, HTTP)。HTTP是由万维网联盟(World Wide Web Consortium, W3C)和互联网工程任务组(Internet Engineering Task Force, IETF)共同制定的规范。HTTP用于从网络传输超文本数据到本地浏览器,它保证高效而准确地传送超文本内容。

HTTP是基于“客户端/服务器”架构进行通信的,HTTP的服务器实现程序有httpd、nginx等,客户端实现程序主要是Web浏览器,如Chrome、Safari、IE、360浏览器、Firefox等。Web浏览器和Web服务器之间可以通过HTTP进行通信。

典型的HTTP请求过程如图1-22所示。

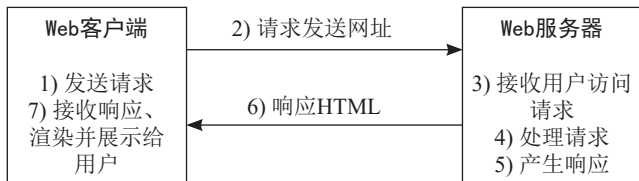


图 1-22 典型的 HTTP 请求过程

具体过程如下。

(1) 用户在浏览器中输入网址,浏览器向Web服务器发起请求。

(2) Web服务器接收用户访问请求,处理请求,产生响应,把处理结果以HTML形式发送给浏览器。

(3) 浏览器接收来自Web服务器的HTML内容,进行渲染以后展示给用户。

1.5.3 移动互联网

移动互联网是在互联网的基础上发展起来的,它利用无线网络技术(如3G、4G、5G等)使移动设备(如智能手机、平板电脑)能够随时随地访问互联网。移动互联网是移动通信与互联网的结合,它将互联网技术、平台、商业模式、应用与移动通信技术相结合并实践。

移动互联网的发展经历了多个阶段,包括起步阶段、无线应用协议(Wireless Application Protocol, WAP)广泛应用阶段、互动娱乐的移动互联网阶段,以及当前的产品广泛应用阶段。随着3G、4G、5G网络在我国广泛普及,移动互联网在各个领域得到了快速发展,极大地改变了人们的生活和工作方式。我国拥有全球先进、完善的通信基础设施,因此移动互联网发展迅猛,在全球处于领先地位,截至2025年年末,我国移动互联网月独立设备数达14.66亿台。

1. 移动互联网的特点

移动互联网集移动性、定制化、多媒体性、社交化等特点于一体,具体介绍如下。

1) 移动性

移动互联网的移动性是其最显著的特点。随着移动通信技术的飞速发展和移动终端设备(如智能手机、平板电脑等)的普及,人们可以随时随地使用这些设备。这些设备小巧轻便,便于放入背包或口袋中,这使移动互联网几乎不受时间和空间的限制。用户可以在任何时间、任何地点使用移动互联网,进行信息的获取、交流和娱乐等活动。这种移动性极大地提高了移动互联的使用频率,使移动互联网成为现代人生活不可或缺的一部分。

2) 定制化

移动互联网的定制化特点主要体现在两方面:一是内容的定制化;二是服务的定制化。

(1) 内容的定制化。移动互联网能够根据用户的兴趣、需求和行为习惯,通过算法推荐个性化的内容。例如,新闻App(如今日头条)可以根据用户的阅读历史推荐相关新闻,购物App(京东、淘宝)可以根据用户的购买记录推荐相似商品等,短视频App(如抖音、快手)可以根据用户的喜好智能推荐用户感兴趣的视频。这种内容的定制化使用户能够更快速地找到自己感兴趣的信息,提高了信息获取的效率。

(2) 服务的定制化。移动互联网提供了丰富的定制化服务。例如,移动支付、在线医疗、在线教育等服务的出现,使用户可以根据自己的需求选择合适的服务。这些服务不仅满足了用户的个性化需求,还提高了生活的便利性和质量。

3) 多媒体性

移动互联网的多媒体性是指其能够传输和处理多种形式的媒体信息,包括文本、图像、音频、视频等。这种多媒体性使移动互联网在信息传递和表达方面丰富多彩、生动直观。例如,用户可以通过短视频App观看各种有趣的视频内容,通过音乐App收听各种类型的音乐,通过社交媒体App(如小红书)分享图片和文字等。这种多媒体性不仅提高了信息的传递效率,还增强了用户的互动性和参与感。

4) 社交化

随着社交媒体和即时通信工具的普及,人们越来越倾向于通过移动互联网进行社交活动。这些社交平台不仅提供了文字、语音、视频等多种交流方式,还通过兴趣群组、地理位置定位等功能帮助用户发现志同道合的朋友和附近的人。此外,移动互联网还促进了线上线下的融合,使社交活动更加多样化和便捷化。例如,用户可以通过社交平台组织线下聚会、参与公益活动等。这种社交化特点不仅丰富了人们的社交活动,还促进了信息传播和文化交流。

2. 移动互联网的应用

移动互联网的应用领域极为广泛,它深刻地改变了人们的生活方式和工作模式。以下是移动互联网的几个主要应用领域。

1) 通信与信息获取

移动互联网使即时通信成为可能,用户可以通过智能手机等移动设备随时随地发送和接收消息,如微信、QQ等应用已成为人们日常沟通的重要工具。通过移动互联网,用户可以方便地浏览新闻、获取各类信息,相关应用有新闻App(如今日头条)、搜索引擎、出行服务App(如飞常准、航班管家)等。这些应用不仅提供实时信息,还通过个性化推荐来满足用户的多样化需求。

2) 娱乐与休闲

移动互联网使用户可以在任何时间、任何地点观看视频内容，相关视频平台有抖音、快手等短视频平台，以及爱奇艺、腾讯视频等长视频平台。手机游戏已成为移动互联网的重要应用之一，用户可以通过移动设备随时随地享受游戏乐趣，热门游戏有王者荣耀、和平精英等。用户还可以通过移动互联网收听和下载音乐，如网易音乐、QQ音乐等应用提供了丰富的音乐资源；有声书、播客等音频内容也受到用户的喜爱。

3) 电子商务

移动互联网推动了电子商务的快速发展，用户可以通过电商平台(如淘宝、京东)或社交媒体(如微信小程序、抖音商城)购买商品。移动支付的普及进一步简化了购物流程。美团、饿了么(已更名为淘宝闪购)等外卖平台以及京东七鲜、大众点评等本地生活服务平台，通过移动互联网为用户提供便捷的餐饮、娱乐、旅游、超市购物等服务。以小红书为代表，用户可以分享购物心得、旅行经历、美妆技巧等，并通过图文形式进行“种草”和推荐，这种形式使内容更加详细具体，有助于用户做出正确的购买决策。

4) 金融服务

移动互联网推进了移动支付的普及，这极大地提高了金融服务的便捷性。移动互联网促进了互联网金融(如网络保险、网络理财等)的发展，一些金融企业通过移动互联网为用户提供灵活、便捷的金融服务。

5) 教育与学习

移动互联网使在线教育成为可能，用户可以通过各类在线教育平台(如网易云课堂、中国大学MOOC)随时随地学习新知识、提升自我。移动学习App(如有道词典、扇贝单词)等学习工具深受用户喜爱，它们通过智能化、个性化的学习方式帮助用户提高学习效率。

6) 医疗健康

移动互联网促进了在线医疗的发展，用户可以通过医疗App(如好大夫在线、春雨医生)进行在线咨询、预约挂号等操作。移动健康App(如Keep、薄荷健康)等健康管理工具备受用户青睐，它们通过提供科学的健康管理方案帮助用户调整生活习惯、保持健康状态。

7) 政务与公共服务

移动互联网在政务领域的应用日益广泛，政府通过移动政务App(如国家政务服务平台、交管12123、个人所得税)为公众提供便捷、高效的政务服务。公共交通、水电燃气缴费等公共服务也通过移动互联网实现了线上办理，提高了服务效率和质量。

8) 人工智能应用与智慧生活

这是目前增长最快的领域。2025年，人工智能类应用的同比增长率达到88.5%，像豆包、Kimi、DeepSeek等这类AI助手应用已经拥有千万级的月活用户。此外，人工智能技术也渗透到智能电视、智能汽车等硬件设备中，例如，海信智能电视在2026年2月的活跃设备数量达到了4733万台。

1.6 系统安全与人工智能应用安全入门

计算机系统安全涵盖保护计算机硬件、软件、数据免受未经授权访问、破坏，以及防范

信息泄露或被篡改等内容。

1.6.1 计算机系统面临的安全威胁

计算机系统安全威胁，是指所有能够对计算机硬件、软件、网络链路、存储数据及业务服务造成干扰、篡改、窃取、破坏的各类风险因素的总称，直接危害计算机系统的保密性、完整性与可用性。结合威胁成因与表现形式，可将计算机系统安全威胁划分为系统自身漏洞威胁、恶意代码威胁、网络攻击威胁、人为安全威胁、硬件与环境安全威胁五大核心类型。

1. 系统自身漏洞威胁

系统自身漏洞属于内生型基础安全隐患，是各类网络入侵与系统攻击的主要突破口，广泛存在于操作系统、应用软件、系统组件及硬件驱动程序中，是计算机系统最根本的安全风险来源。系统漏洞的产生并非外部攻击导致，而是系统研发、配置、运维过程中留存的固有缺陷，主要分为程序漏洞、配置漏洞和版本老旧风险三类。

1) 程序漏洞

软件开发阶段因代码编写不规范、逻辑设计不合理、数据边界校验缺失等问题，极易产生缓冲区溢出、权限绕过、代码注入、越权访问等高危程序漏洞。攻击者可利用这类漏洞突破系统固有防护机制，在未授权状态下读取私密数据、执行恶意程序、获取系统最高控制权限，彻底扰乱计算机系统的正常运行逻辑。

2) 配置漏洞

计算机系统与各类应用软件的默认配置优先适配通用性、易用性，存在大量安全短板。常见问题包括默认账号密码过于简单、高危端口默认开放、匿名访问权限未限制、系统冗余服务未关闭等。多数用户缺乏系统安全加固意识，长期沿用默认配置，大幅扩大系统攻击面，使设备极易被网络端口扫描、恶意探测和非法入侵。

3) 版本老旧风险

长期未迭代更新的操作系统、老旧应用软件、废弃插件等，会持续留存大量已知安全漏洞，且无法获取官方安全补丁修复。这类老旧程序的漏洞具有公开性、可批量利用性，极易被自动化攻击工具精准捕获、利用，成为长期暴露的固定安全隐患。

2. 恶意代码威胁

恶意代码是由人为编写、具备窃取、破坏、干扰等恶意功能的程序集合，是终端计算机最普遍、最典型的安全威胁。该类威胁具备传播速度快、隐蔽性极强、破坏力突出的特征，主要包含计算机病毒、木马程序、蠕虫病毒、勒索软件四大主流类型，各类恶意代码的危害特征与攻击方式各有不同。

1) 计算机病毒

计算机病毒具备极强的自我复制和主动感染能力，可依附于系统程序、文件资源中，通过文件传输、移动设备复制、网络分享等渠道广泛扩散。其核心危害在于破坏系统核心文件、篡改业务数据、占用大量系统运行资源，最终造成计算机运行卡顿、程序闪退、蓝屏死机甚至系统崩溃。

2) 木马程序

木马程序主要通过伪装成正规软件、安装工具、资源文件诱导用户主动下载安装，不会直接破坏系统设备，隐蔽性极强。其核心危害是秘密开启系统后门，非法窃取用户账号密码、私密资料、核心业务数据，同时支持攻击者远程控制终端设备，造成数据泄露、系统被控等严重安全问题。

3) 蠕虫病毒与勒索软件

蠕虫病毒依托网络漏洞、局域网通信机制自动传播，无须人工触发，可批量入侵联网终端设备，极易引发局域网大规模设备中毒、网络带宽拥堵、整体网络瘫痪。勒索软件是当前危害性最高的恶意代码之一，通过加密终端本地存储设备内的全部数据，以索要赎金作为解密条件，直接造成用户数据永久丢失、企业蒙受重大经济损失。

3. 网络攻击威胁

联网状态下的计算机系统持续暴露于公共网络环境中，面临大量外部主动攻击风险，网络攻击是联网设备最主要的外部安全威胁。网络攻击手段丰富、针对性强，攻击者依托各类技术手段探测系统缺陷、突破防护体系，窃取、破坏系统资源与数据，主流攻击方式包括暴力破解、端口扫描、DDoS攻击、注入攻击、中间人攻击等。

1) 暴力破解与端口扫描

暴力破解通过程序批量枚举账号、密码组合，尝试破解系统登录权限，实现非法入侵设备、窃取系统控制权的目的。端口扫描属于前置探测类攻击，通过批量探测计算机开放端口，排查系统漏洞、开放服务和可入侵点位，为后续深度攻击提供精准依据。

2) DDoS攻击与注入攻击

DDoS(分布式拒绝服务)攻击通过构建海量虚假访问流量，猛烈攻击计算机网络端口与业务服务，占用全部网络带宽和系统资源，导致合法用户请求无法响应，最终造成业务瘫痪、网络中断。SQL注入、XSS跨站脚本等注入攻击，主要针对搭载数据库的业务系统，可非法读取、篡改、删除核心数据库数据，破坏业务数据完整性。

3) 中间人攻击

中间人攻击主要针对网络数据传输环节，攻击者拦截、窃听、篡改设备与服务器之间的传输数据，全程隐蔽且不易被察觉。该攻击方式会直接破坏数据传输的保密性与完整性，极易引发数据泄露、信息伪造、业务指令篡改等安全问题。

4. 人为安全威胁

相较于技术类威胁，人为威胁具备随机性、主观性、不可控性的特点，主要分为用户无意误操作和恶意人为操作两类。

1) 无意误操作

无意误操作普遍存在于日常设备使用与运维场景中，普通用户随意单击陌生链接、下载未知软件、接入陌生移动外设、泄露账号密码、随意开放系统权限、误删核心文件数据，均会引入安全风险。同时运维人员补丁更新失误、安全策略配置错误、系统操作不规范，也会引发系统故障、数据异常等安全问题。

2) 恶意人为操作

恶意人为操作多为内部人员违规越权或外部人员蓄意破坏，包括内部工作人员越权查阅、复制、外泄涉密数据，恶意篡改业务数据，私自安装违规软件，刻意绕过安全防护策略等。由于内部人员熟悉系统架构、权限体系和防护短板，其恶意操作的隐蔽性更强、造成的安全危害更大。

5. 硬件与环境安全威胁

硬件设备故障与运行环境异常是极易被忽视但危害直接的安全威胁，会直接破坏计算机系统的稳定性与可用性，造成硬件损毁、数据丢失、系统停机等问题。包括自然灾害(地震、洪水)，设备老化故障以及人为的盗窃和破坏，虽然看似传统，但在数字时代，这些威胁同样能迅速转化为对信息系统的重大打击，导致数据丢失、服务中断，甚至整个业务系统瘫痪。

1.6.2 计算机系统安全防护技术

针对计算机系统面临的各类安全威胁，为全面保障计算机硬件、软件、网络、数据及业务流程安全，抵御入侵、破坏、泄露、篡改等风险，需构建全方位、多层次的安全防护技术体系。计算机系统安全防护技术以保障系统保密性、完整性、可用性、可控性和不可否认性为核心目标，覆盖系统底层、网络传输、数据存储、终端运维、应急保障全流程，从根源上规避安全风险、阻断攻击行为、降低安全事故损失。根据防护场景与防护对象，可将主流防护技术分为系统底层安全防护技术、网络安全防护技术、数据安全防护技术、终端安全防护技术、安全运维与应急防护技术五大类别。

1. 系统底层安全防护技术

系统底层防护是计算机安全的基础屏障，主要针对操作系统、硬件固件、系统漏洞、权限缺陷等内生性风险进行加固防护，从源头弥补系统的固有短板，减少攻击突破口，是抵御各类恶意攻击的前提保障。

1) 系统漏洞修复与补丁管理技术

针对系统程序漏洞、版本老旧漏洞等风险，通过自动化漏洞扫描、补丁监测、版本更新机制，及时获取官方发布的安全补丁，对操作系统、驱动程序、应用软件的高危漏洞进行修复。同时通过补丁预测试、分批部署、版本管控等规范流程，避免因补丁安装引发的系统兼容故障，在保障系统稳定运行的前提下，彻底修复缓冲区溢出、权限绕过、代码执行等各类已知漏洞，消除底层安全隐患。

2) 系统权限管控技术

严格遵循最小权限原则，对系统账号、用户、应用程序进行分级权限管控。将系统权限划分为管理员权限、普通用户权限、访客权限等不同等级，严格限制普通用户修改系统配置、删除核心文件、安装未知程序等高危操作；同时对各类应用程序的运行权限进行约束，禁止非法程序获取系统最高控制权。常态化开展权限审计、闲置账号清理、超标权限回收工作，杜绝因越权操作、账号滥用引发的安全风险。

3) 系统安全加固技术

针对系统默认配置漏洞开展专项加固,通过关闭系统冗余服务、禁用高危开放端口、取消匿名访问权限、修改默认账号密码、锁定系统核心配置等方式,缩小系统攻击面,减少被扫描、探测、入侵的概率。同时支持系统内核加固、硬件固件加密保护,防止底层系统配置被篡改、底层后门被植入,筑牢计算机系统的底层安全防线。

2. 网络安全防护技术

网络安全防护技术主要针对网络端口扫描、DDoS攻击、注入攻击、中间人攻击等外部网络威胁,在网络边界和传输链路搭建防护屏障,过滤恶意流量、阻断非法访问、保障网络传输安全,是联网计算机系统的核心外部防护手段。

1) 防火墙防护技术

防火墙是网络边界的核心防护设施,分为硬件防火墙和软件防火墙两类。通过预设精细化安全策略,对进出计算机系统的所有网络流量进行全面过滤与审计,精准拦截非法访问、端口扫描、恶意连接等高危行为,仅放行合法合规的网络请求,有效抵御各类基础网络入侵攻击,构筑起网络安全的第一道屏障。

2) 入侵检测与防御技术

入侵检测系统(IDS)可实时监控系统网络行为与运行状态,通过特征匹配、行为分析、异常识别等技术,精准发现暴力破解、SQL注入、DDoS攻击、木马渗透等恶意行为,并及时推送安全告警,实现风险早发现、早预警。入侵防御系统(IPS)可在检测到攻击行为的瞬间,自动阻断恶意流量、切断攻击链路,实现“检测—预警—阻断”一体化主动防御,抵御实时网络攻击。

3) 网络加密与安全传输技术

针对网络传输过程中的窃听、篡改、中间人攻击风险,采用SSL/TLS加密、IPSec加密、VPN虚拟专用网络等技术,对网络传输数据进行加密封装,避免数据在传输链路中被窃取、篡改。同时配套网络身份认证机制,严格管控设备接入权限,杜绝非法设备接入网络窃取系统数据,保障网络传输的保密性与完整性。

3. 数据安全防护技术

数据是计算机系统的核心资产,数据安全防护技术针对数据泄露、篡改、丢失、非法窃取等风险,覆盖数据存储、使用、传输、备份、销毁全生命周期,全方位守护系统数据安全。

1) 数据加密技术

数据加密是数据防护的核心技术,分为存储加密和文件加密两类。全盘存储加密可对计算机硬盘、移动存储设备等存储介质进行整体加密,防止设备丢失、被盗后数据批量泄露;文件加密可针对核心文档、业务数据、数据库文件进行单独加密,通过密钥验证管控访问权限,未授权用户无法读取涉密数据。主流采用AES对称加密、RSA非对称加密算法,兼顾加密安全性与运行效率。

2) 数据备份与恢复技术

针对病毒破坏、硬件故障、误操作、网络攻击引发的数据丢失风险,通过定时备份、增量备份、异地容灾备份等方式,常态化留存系统文件、业务数据、用户资料。同时配套

专业数据恢复技术，在数据丢失、损坏、篡改后可快速还原数据与系统状态，保障计算机系统业务连续性，提升数据安全容错能力。

3) 数据防泄露与脱敏技术

通过终端外设管控、网络传输管控、文件操作审计等手段，禁止核心数据通过U盘、网盘、陌生邮件等渠道非法外发，阻断数据泄露渠道。针对身份证、手机号、涉密业务数据等敏感信息，采用数据脱敏技术进行模糊化、匿名化处理，在不影响数据正常使用的前提下，规避敏感数据泄露风险。

4. 终端安全防护技术

终端安全防护聚焦计算机单机设备与用户操作行为，主要抵御恶意代码入侵、人为违规操作、外设非法接入等终端侧安全威胁，规范终端运行环境与操作行为。

1) 恶意软件查杀防护技术

依托终端安全防护软件、杀毒系统，通过病毒特征库比对、人工智能行为识别、实时动态监控等方式，全方位查杀计算机病毒、木马、蠕虫、勒索软件、间谍软件等恶意程序。支持实时防护、全盘扫描、离线查杀等功能，可主动拦截恶意程序下载、安装与运行，防止恶意代码入侵，避免其破坏系统、窃取数据。

2) 外设安全管控技术

针对移动设备接入带来的病毒传播、数据窃取风险，对U盘、移动硬盘、外接设备、打印机等外设进行严格管控。可禁用未知外设自动接入、限制外设读写权限、审计外设数据传输行为，有效切断通过物理移动设备传播病毒、复制涉密数据的风险，筑牢终端物理安全防线。

3) 终端安全审计技术

全程记录计算机系统的用户登录、文件操作、程序运行、网络访问、权限变更等各类操作日志，实现终端行为全溯源。所有操作日志可长期留存、随时查询，一旦发生数据泄露、系统异常、恶意入侵等安全事件，可通过审计日志追溯攻击源头、定位问题成因，为安全事件处置与责任界定提供依据。

5. 安全运维与应急防护技术

安全运维与应急防护是计算机系统长效安全的保障，可实现安全风险事前预警、事中管控、事后处置，弥补静态防护的不足，提升系统整体安全防御与容错能力。

1) 安全态势监控技术

通过安全管理平台对计算机系统运行状态、漏洞隐患、网络攻击、数据流转、终端行为进行全天候实时监控，动态感知系统安全态势。可主动发现潜在安全风险与异常行为，提前预警安全隐患，实现从被动防御向主动预判、提前防护的转变。

2) 应急响应与灾难恢复技术

针对病毒爆发、系统瘫痪、数据泄露、大规模网络攻击等突发安全事件，建立标准化应急处置流程，快速开展风险隔离、攻击阻断、漏洞修复、数据恢复等工作，最大限度地降低安全事故造成的损失。同时通过容灾部署、灾备备份机制，保障极端故障场景下计算机系统业务不中断，提升系统的安全性和稳定性。

1.6.3 人工智能应用安全入门

随着人工智能技术广泛落地，大语言模型、人脸识别、智能推荐等人工智能应用极大便利了生产生活，也催生了区别于传统网络安全的新安全风险。传统网络安全主要防范网络入侵、病毒木马、程序漏洞等。人工智能应用安全则覆盖数据采集、模型训练、部署运行、迭代更新全生命周期，是针对人工智能数据、算法、业务场景开展风险识别、防御与管控的新型安全领域，是传统网络安全的时代延伸。

1. 人工智能应用安全的核心特征

相较于传统网络安全，人工智能应用安全具有独特风险特征，也是其安全问题高发的核心原因。

1) 数据高度依赖

人工智能模型性能完全依托训练数据，数据的真实性、安全性、完整性直接决定模型安全，数据缺陷会直接转化为模型漏洞。

2) 算法黑盒性强

深度学习、大模型算法决策过程不可解释、难以溯源，出现安全问题后漏洞排查、问题修复难度远高于传统程序。

3) 风险场景多元

风险突破了传统代码、网络层面，延伸至数据、算法、业务、伦理等多个维度，隐蔽性更强、影响范围更广。

4) 动态对抗性

人工智能模型容易遭受针对性对抗攻击，攻击者可依托模型特性动态调整攻击方式，防护体系需要持续迭代优化。

2. 人工智能应用全生命周期安全风险类型

人工智能应用全生命周期安全风险可分为数据层、模型层、应用层三大核心类别，覆盖人工智能从训练到落地运行的全部流程，是安全防护的核心重点。

1) 数据层安全风险

数据是人工智能的核心基础，数据层风险是人工智能最基础的源头风险，主要分为三类。

(1) 数据隐私泄露。人工智能训练需海量用户隐私数据，若数据存储、传输、训练环节未做好加密与脱敏处理，会造成用户信息泄露，大模型还可能通过对话诱导还原用户隐私数据。

(2) 数据投毒攻击。攻击者向训练集中混入恶意、失真、带偏见等数据，污染数据集。模型学习异常数据后，会出现识别错误、判断偏差、功能失效等问题。

(3) 原生数据缺陷。训练数据存在残缺、冗余、偏见等问题，会使模型产生固有漏洞，引发算法歧视、决策失误等非人为安全问题。

2) 模型层安全风险

模型是人工智能应用的核心。模型层风险源于算法缺陷与结构漏洞，是人工智能专属攻击的主要突破口，主要有四种类型。

(1) 对抗样本攻击。典型的人工智能专属攻击，攻击者在正常数据中添加人眼不可见的微小扰动，即可误导模型做出误判，常用于破解人脸识别、自动驾驶视觉识别系统。

(2) 模型窃取和复刻。攻击者通过公开API接口反复测试，探测模型参数与逻辑，复刻盗版模型，造成企业核心技术与知识产权泄露。

(3) 模型后门漏洞。训练阶段被植入隐藏后门，模型常规运行正常，触发特定指令后会执行恶意逻辑，实现权限绕过、违规输出，隐蔽性极强。

(4) 模型偏见与失控。受数据偏差或算法缺陷影响，模型易产生各类算法歧视，在复杂场景下出现逻辑混乱、决策失控，输出有害内容。

3) 应用层安全风险

应用层风险是模型部署落地后，面向终端用户和业务场景产生的实操安全风险。主要包括以下四种类型。

(1) 提示注入攻击。针对大语言模型的主流攻击方式，攻击者构造恶意提示词，绕过安全过滤，诱导大语言模型生成诈骗、暴力、破解教程等违规内容。

(2) 越权与滥用风险。权限管控不当会导致用户越权获取涉密数据；恶意人员可利用人工智能批量生成钓鱼、垃圾内容，实施网络诈骗、网络攻击等违规行为。

(3) 接口安全漏洞。人工智能对外API接口若存在未授权访问、参数篡改等漏洞，攻击者可批量调用模型、窃取数据、发起攻击，造成服务瘫痪。

(4) 业务场景失控。在自动驾驶、智能医疗、智能安防等关键场景，人工智能决策失误会直接引发人身伤害、公共安全事故。

3. 人工智能应用安全基础防护技术与原则

针对人工智能全生命周期安全风险，行业建立了“源头防控、过程防护、终端管控、持续迭代”的防护体系，相关人员需掌握人工智能应用安全核心防护原则与基础技术。

1) 人工智能安全核心防护原则

(1) 最小权限原则。数据采集、模型调用、接口访问等仅开放必要权限，杜绝超权限操作，从源头规避风险。

(2) 全程可溯源原则。对数据处理、模型训练、用户调用等全流程留存日志，实现安全问题可追溯、可定位。

(3) 隐私优先原则。人工智能研发与应用优先保障用户隐私，严格落实数据脱敏、匿名化，严禁违规采集、滥用个人信息。

(4) 动态防护原则。针对人工智能动态对抗攻击特性，持续更新防护策略、优化模型、排查漏洞，实现防护能力迭代升级。

2) 基础防护技术

(1) 数据安全防护。核心技术包括数据脱敏、加密、去重、数据集校验、隐私计算等，可有效防范数据泄露、数据投毒等问题，实现数据安全合规使用。

(2) 模型安全防护。通过对抗训练、模型加固、后门检测、数字水印等技术，提升模型抗攻击能力，防范模型后门、模型窃取等风险。

(3) 应用安全防护。通过提示词过滤、输入输出内容审核、接口权限校验、流量限流、行为风控等手段，防范注入攻击、越权调用等应用层风险。

(4) 合规管控。依据《生成式人工智能服务管理暂行办法》等法规，对人工智能应用开展备案、合规审核与风险评估，杜绝人工智能非法滥用与违规落地。

1.7 本章小结

本章系统介绍了智能信息时代下的计算机基础核心知识。首先辨析了信息与数据的概念关系，明确了数据在人工智能领域的重要价值，梳理了计算机的发展历程、分类特点及智能化发展趋势。同时重点介绍了计算机二进制信息表示的底层原理，计算机软硬件系统组成及智能硬件架构，计算机网络在数据传输与智能交互中的基础作用。此外，本章还介绍了传统系统安全与人工智能应用安全的基础常识，帮助学习者建立基础安全防护意识。本章打通了传统计算机知识与智能技术的关系，构建了完整的入门知识框架，为后续人工智能相关课程学习筑牢基础。

1.8 思考和练习

1. 选择题

- (1) 未经加工的客观符号、数字、文字等原始素材被称作()。
A. 信息 B. 数据 C. 知识 D. 智能
- (2) 数据经过加工处理、赋予具体语义、具备使用价值后转变为()。
A. 信号 B. 数据 C. 信息 D. 编码
- (3) 第一台电子数字计算机 ENIAC 诞生于()。
A. 1936 年 B. 1946 年 C. 1956 年 D. 1966 年
- (4) 按照运算速度、性能规模划分，超级计算机属于()。
A. 巨型机 B. 微型机 C. 工作站 D. 嵌入式计算机
- (5) 计算机内部存储、处理所有信息统一采用的进制是()。
A. 十进制 B. 八进制 C. 二进制 D. 十六进制
- (6) 人工智能模型训练、深度学习运算最核心依赖的硬件部件是()。
A. 键盘 B. 图形处理器 (GPU)
C. 光驱 D. 打印机
- (7) 完整计算机系统分为两大组成部分，分别是()。
A. 主机和外设 B. 硬件系统和软件系统
C. CPU 和内存 D. 系统软件和应用软件
- (8) 计算机网络最核心的功能是()。
A. 单机运算 B. 资源共享与数据传输
C. 数据加密 D. 本地存储
- (9) 下列不属于人工智能应用安全风险的是()。
A. 训练数据泄露 B. 算法偏见
C. 硬盘正常读写 D. 人工智能生成虚假内容

- (10) 四代计算机的主要划分依据是()。
- A. 运算速度 B. 核心电子元器件
C. 体积大小 D. 存储容量
- (11) 比特(bit)是计算机中最小的()单位。
- A. 存储 B. 重量 C. 长度 D. 时间
- (12) 手机里搭载的微型处理器属于()。
- A. 大型机 B. 嵌入式计算机
C. 小型机 D. 超级计算机
- (13) 下列属于系统软件的是()。
- A. 微信 B. Windows 操作系统
C. WPS 文档 D. AI 绘画软件
- (14) 防范黑客入侵、病毒破坏属于()范畴。
- A. 硬件架构 B. 信息系统安全
C. 进制转换 D. 计算机分类
- (15) 智能信息时代的核心生产要素是()。
- A. 煤炭 B. 数据 C. 电力 D. 钢材

2. 填空题

- (1) _____是信息的载体，_____是经过加工后具有价值的数据。
- (2) 计算机发展依次经历了电子管、晶体管、中小规模集成电路、_____四个时代。
- (3) 计算机中最小的存储单位是位(bit)，基本存储单位是_____。
- (4) 面向人工智能大模型运算优化的硬件统称为_____硬件。
- (5) 计算机网络按照覆盖范围可分为局域网、城域网和_____。
- (6) 计算机系统硬件五大核心部件：运算器、控制器、存储器、输入设备、_____。
- (7) 运算器 + 控制器合称为_____，是计算机的核心。
- (8) 人工智能底层所有文本、图像、语音数据最终都会被转换为_____代码进行存储与计算。
- (9) 信息安全主要包含系统安全、数据安全和_____应用安全三大板块。
- (10) 个人台式电脑、笔记本电脑统称为_____计算机。
- (11) 数据安全三大目标：保密性、完整性、_____。
- (12) 信息技术发展最终迈入以大数据、人工智能为标志的_____时代。

3. 简答题

- (1) 简述信息与数据的区别和联系，并说明数据在智能信息时代的核心价值。
- (2) 计算机为何以二进制作为信息的基础表示形式？
- (3) 简述计算机的发展趋势，谈谈传统计算机技术与智能化应用是如何融合发展的。
- (4) 对比分析传统计算机系统安全与人工智能应用安全的主要差异。
- (5) 结合计算机系统与网络基础，简述智能硬件架构和计算机网络在人工智能数据交互过程中的作用。