

·专题:从信息素养到人工智能素养·

人工智能素养:提出背景、概念界定与构成要素*

尹开国

(1.武汉工商学院管理学院 湖北武汉 430065)

摘要:人工智能(AI)正深刻改变社会生产、生活和组织方式,如何提升个体在智能时代的适应性和创造力成为公众素养研究的新命题。文章从概念演变和技术发展角度解读 AI 素养的提出背景,指出 AI 素养是在人机双边互动技术范式转变环境下对信息素养、数字素养的传承和拓展;在文献回溯基础上,探索性提出 AI 素养的解释性定义,强调 AI 素养是一个多样化、多层次的概念,如何与 AI 协作共生构成 AI 素养的差异化内涵;参考 KSAVE 模型,识别、构建了由 AI 意识、AI 认知、AI 技能、AI 思维、AI 伦理等五个要素构成的通识性 AI 素养框架,并系统阐述各要素内涵、意义及内容架构思路。

关键词:人工智能素养;数字素养;人工智能伦理;人工智能素养教育

中图分类号:TP181 **文献标识码:**A **DOI:**10.11968/tsyqb.1003-6938.2024033

AI Literacy: Background, Conceptual Definition, and Constituent Elements

Abstract As AI rapidly integrates into society, it is profoundly changing the ways of social production, living and organization, and how to enhance individual adaptability and creativity in the intelligent era has become a new proposition in the study of public literacy. The article interprets the background of the proposal of AI literacy from the perspective of conceptual evolution and technological development, pointing out that AI literacy is an inheritance and expansion of information literacy and digital literacy; on the basis of reviewing existing literature, it exploratorily proposes an interpretive definition of AI literacy, emphasizing that AI literacy is a diverse and multi-level concept, and how to collaborate and coexist with AI constitutes the differentiated connotation of AI literacy; referring to the KSAVE model, it identifies and constructs a general AI literacy framework composed of five elements: AI awareness, AI cognition, AI skills, AI thinking, and AI ethics, and systematically elaborates on the connotation, significance, and content architecture ideas of each element.

Key words AI literacy; digital literacy; AI ethics; AI literacy education

人工智能(Artificial Intelligence Literacy, AI)是当今最具挑战性的议题之一。AI 应用加速渗透到各行各业,加速了社会演化进程,将驱动知识生产方式、产品制造模式、社会连接形态、生活休闲场景发生一系列变革^[1]。2024年3月,中国政府工作报告^[2]首次提出要开展“人工智能+”行动,旨在深化大数据、人工智能等研发应用,打造具有国际竞争力的数字产业集群。在带来划时代机遇和巨大便利的同时, AI 也给今天的人类社会带来重重挑战。智能机器将承担大部分生产工作,人与人之间、人与机器之间的竞争压力日益增大^[3]; AI 技术在公平公正、隐私保护、安全和伦理道德等方面也存在着巨大风险,“AI

焦虑”现象蔓延。如果公众对 AI 的意识、理解和技能没有普遍提升,社会可能会经历一个加剧的数字鸿沟,并创造一个不利于劳动力市场和经济增长的新的 AI 鸿沟。关于公众是否准备好理解、接受、信任或适应 AI 世界的问题越来越多^[4],如何全面增强社会成员的适应性和创造力成为公众素养研究的新命题。迄今为止,学界对于如何定义通识性的人工智能素养(以下简称“AI 素养”)尚未展开充分探索。为此,本文尝试从相关技术素养概念演变和 AI 技术发展等角度解读 AI 素养的提出背景,梳理、界定 AI 素养的概念,提出一个通识性 AI 素养框架,并系统阐述 AI 素养核心构成要素的内涵、意义及内容架构思路,

* 本文系国家自然科学基金重大项目“我国政府信息公开到数据开放的理论创新与实践路径研究”(项目编号:22ZD329)研究成果之一。
收稿日期:2024-06-16;责任编辑:刘婷

为理解、提升公众 AI 素养提供一个系统视角。

1 AI 素养的提出背景

素养可以被理解为掌握特定领域的知识或技能,其术语体系随社会发展不断演变。第三次工业革命以来,随着信息技术的深入发展,政府信息公开和数据开放程度不断提升,各类市场化网络信息媒介和数字工具快速普及,相继出现了计算机素养、信息素养、媒介与信息素养、数字素养、数据素养、算法素养、AI 素养等技术素养概念。从公众教育层面来说,联合国教科文组织、欧盟等权威机构对信息素养和数字素养认可度较高。

1974年,时任美国信息产业协会主席 Zurkowski 在向美国图书馆与信息科学委员会提交的报告《信息服务环境:关系与优势》^[5]中首次提出了信息素养概念。针对当时社会信息发布的多样化趋势,Zurkowski 提出,信息素养是“利用多种信息工具及主要信息资源使问题得到解决的技术和技能”,涵盖信息检索、信息评估、信息利用等多个方面。报告强调了信息环境中的各种关系和优势,还指出了信息道德的重要性,为人们理解和应对信息社会的挑战提供了宝贵的理论资源和实践指南。此后,在美国图书馆协会信息素养主席委员会、美国国家信息素养论坛等机构推动下,信息素养理论与实践快速发展,并进入国际视野。2013年,联合国教科文组织发布的《媒体与信息素养政策和战略指南》^[6]集学界及实践领域大成,首次将信息素养与媒体素养合二为一,并将其定义为“个人和群体在接触、理解、评估、创造和传播各种形式的媒体与信息的过程中所展现出的认知、技能和态度”。

从二十世纪九十年代开始,个人计算机开始进入更多的家庭和企业,一个更具综合性的素养框架随之出现。1994年,Alkalai 首次提出“数字素养”一词,并于2004年建立了一个由“图片-图素养”、再生产素养、分支素养、信息素养和“社会-情感素养”等五个要素构成的数字素养理论模型^[7]。21世纪以来,越来越多的学者、机构和政府开始采用数字素养概念。其中,尤以欧盟的政策影响最为广泛。2006年,欧盟在《终身学习的关键能力:欧洲框架》^[8]中提

出,数字素养是“个人能充满自信并采取批判性的态度去使用信息社会的各种技术,具备信息通信技术方面的基本技能”。作为“欧洲数字议程”的重要成果,欧盟至今已发布多个版本的《欧洲公民数字能力框架》,2022年3月发布的 DigComp2.2^[9]是其最新版本。在欧盟框架体系中,信息素养和数据素养是构成数字素养的子维度,且在多处子领域中体现了 AI 相关能力要求。

自2022年11月 OpenAI 推出 ChatGPT-3.5 以来,生成式 AI 快速渗透各行各业,全球生成式 AI 用户数量呈指数级增长,AI 大规模商业化时代正式开启。当前,从“人工人类智能”逐步过渡到更高级的“通用人工智能”已是大势所趋,AI 在图像分类、基础阅读理解、视觉推理和自然语言推理等基准测试上超越了人类表现,在复杂认知任务方面与人类的差距不断缩小^[10]。AI 正重塑人类与物理、数字和机器世界的互动。AI 技术区别于非 AI 技术的三个显著特征——自主性、学习和不可解释性,使得传统信息系统理论中关于人类与技术互动的核心假设无效^[11]。如 AI 技术打破了功能一致性的假设,因为它们可以从处理的数据中学习,从而随着时间推移表现出不同行为;语音、视觉等自然交互方式消除了人工界面的必要性,从而使得人类可能没有意识到正与 AI 互动^[12]。人机双边互动技术范式的转变使得修订技术素养体系的需求变得迫切。在此背景下,作为对信息素养、数字素养的传承和拓展,AI 素养进入学界讨论和研究范畴。

2 AI 素养的概念界定

自二十世纪五十年代中期被正式提出以来,AI 已经历了将近七十年发展历程,见证了多次重大技术范式转变。相比之下,AI 素养还是一个新生概念。2015年,AI 素养术语最早在 Konishi 提交的一篇文章中被提及。Konishi 认为,AI 素养不仅是对技术的掌握,更是一种思考和应对未来科技发展的能力^[13]。随后,Kandlhofe 等将 AI 素养定义为理解不同产品、服务背后 AI 基本技术和概念的能力,并确定了七个主题:自动装置、智能代理、图和数据结构、排序、通过搜索解决问题、经典规划以及机器学习^[14]。在此

阶段,由于技术复杂,AI素养教育主要面向高等院校,公众普及度较低。近年来,随着技术的不断成熟和应用场景的拓展,图形模块化编程、虚拟仿真等技术相继出现,AI学习门槛大大降低,AI素养已经转变为面向非专业人士的术语体系,更多关注公众应该拥有的基本AI知识和综合能力。Long和Magerko于2020年提出的定义被广为引用——“AI素养是一系列能力,这些能力使个体能够批判性评估AI技术,有效与AI沟通和协作,并在在线、家庭和工作场所使用AI作为工具。”^[15]该研究标志着AI素养研究进入学界主流视野,围绕AI素养教育的研究开始增多。2020年以来出现了一些较具代表性的AI素养定义(见表1)。

从表中可知相对较早的定义比较关注概念、技术、数据、过程和技能,最近的定义则开始关注个体的态度及在特定情境中应对复杂需求的能力。根据美国国家研究委员会的观点,为了适应社会和时代,一个技术素养概念至少应包括三个独立且互相依存的维度:知识、思维和行动方式、能力^[27]。AI素养不仅涉及专业技能,更包括对AI技术背后的伦理、社会责任以及全球治理等问题的理解和反思^[26],要求构建一种系统性的认知能力,使个体能够全面理解AI工作原理、应用场景以及对社会的影响。

综合已有研究,本文从解释性的角度将AI素养

定义为:在日常生活、学习和工作场景中,个体在接触、理解、评估、利用AI技术和应用过程中体现出的综合能力、行为方式和伦理价值观。从功能性层面,具备AI素养的人应该能够:(1)意识到并理解以AI为驱动力的社会运行机制,了解AI的社会影响和发展趋势;(2)理解AI基本原理,熟悉主流AI工具和应用,灵活利用AI解决问题,在人机共生环境中把握个体主体性和控制力;(3)具备AI伦理意识,积极参与AI政策议题。

AI素养的内涵特色主要体现在以下两个方面:首先,AI素养是一个多样化、多层次的概念。不同教学阶段的学生和教师、企业管理层及各类岗位职员、政府公务员等不同用户群体的AI素养需求内嵌于各自的应用场景。在社会教育实践和个体终身学习层面,培养和提升AI素养的复杂性还体现在对技术理解的深度、应用场景的广度、伦理问题的敏感度等多个层面。在没有界定特定群体时,“AI素养”通常指面向非专业背景的公众AI素养或全民AI素养,与“通识性AI素养”同义。与此最为接近的是美国国家人工智能咨询委员会对AI素养运动受众的界定,即“应优先考虑社区学院、大学、非传统学生、劳动力、代表性不足的社区中的个人,以及最有可能被AI取代的成年人。”^[4]

其次,AI正具备越来越强的社会伦理属性,如何

表1 代表性AI素养定义(按发表年份排序)

来源	定义
Dai等 ^[16]	AI素养是获取和使用与AI相关的知识和技能的能力。
Hermann ^[17]	AI素养是个人对以下五个方面的基本理解:①数据是如何被收集的,以及哪些数据被收集;②数据如何被组合或比较以得出推论、创造和传播内容;③自己决定、行动和反对的能力;④AI对偏见和选择性的敏感性;⑤AI在总体上的潜在影响。
Kong等 ^[18]	AI素养包含三个组成部分:AI概念;使用AI概念进行评估;通过问题解决使用AI概念来理解真实世界。
Laupichler等 ^[19]	AI素养是指能够理解、使用、监控和批判性地反思AI应用的能力。
Ng等 ^[20]	AI素养是在日常生活中有效、道德地使用AI的一组新的技术态度、能力和素质。
UNESCO ^[21]	AI核心素养包括理解AI如何收集、清理、操作、分析数据以及AI算法如何在数据中找到模式和联系的能力。
Wang等 ^[22]	AI素养包括:能够在实际应用中意识到并理解AI技术;能够熟练利用AI技术完成任务;能够分析、选择和批判性地评估AI提供的数据和信息,培养对自己个人责任和相互权利与义务的尊重。
Carolus等 ^[23]	AI素养是以自主和理性的方式与AI技术互动所需的能力。
Pinski和Benlian ^[24]	AI素养是一种社会技术能力,包括在与AI互动中对人类和AI角色的知识、对AI处理步骤(即输入、处理和输出)的知识,以及在AI互动中的经验。
NAIAC ^[4]	AI素养是对AI的好处、风险、机会以及它如何影响日常生活的基本理解。
Cetindamar等 ^[25]	AI素养是技术、工作、人机交互和学习能力的集合。
钟柏昌等 ^[26]	作为适应和驾驭AI的素养,AI素养的本质是通过人的技术化过程,获得人的本质力量的提升,最终实现人与智能技术的和谐发展。

与 AI 协作共生构成 AI 素养的差异化内涵。新一代 AI 在智能形态上将凸显颠覆意义,对社会的影响表现出更广泛的渗透性和更普遍的赋能空间。个体与 AI 互动过程中,社会逻辑超越机器交互逻辑,AI 更多表现为拥有高等智能的协作者、共创者角色。这一差异导致用户评估 AI 产品的标准与评估传统数字产品的标准不同,直接使用与之前数字素养相关的内容描述 AI 素养不再恰当^[22]。AI 在某些特定领域拥有超人类智能,人与智能机器能够在复杂的环境中合作完成更加繁杂的任务。人类的灵活性、想象力、创造性与机器的稳定性、逻辑性可以优势互补。Veloso 等认为,人类与智能系统之间的关系将是“共生自治”^[28]。未来的社会交际不仅仅局限于人类个体之间,更有可能出现在人类和 AI 之间。AI 素养培育要超越技术视角,始终保持人类在人机共生中的主动性、主体性和主导性。

3 AI 素养的构成要素

与 AI 素养概念化研究相应,目前还没有比较公认的 AI 素养通用框架。比较有代表性的是,Long 和 Magerko 基于人机交互的视角,从认知 AI、理解 AI、使用 AI 等角度,提出包含 17 项要素的 AI 素养能力框架^[15];在此基础上,Heyder 和 Posegga 将能力结构化为三个概念块:功能性 AI 素养、批判性 AI 素养和社会文化 AI 素养^[29];Ng 等根据布鲁姆的教育目标分类法构建了一套 AI 核心素养编码框架,通过文献综述确定了培养 AI 素养的四个方面:知道和理解 AI、使用和应用 AI、评估和创建 AI、AI 伦理^[20];Wang 等以“技术—认知—伦理”模型为基础,从意识、使用、评估和伦理四个方面构建了 AI 核心素养框架^[22];Chiu 等提出的框架包括五个关键组成部分:技术、影响、伦理、协作和自我反思^[30];联合国教科文组织在 2021 年 12 月举行的“教育中的人工智能”主题会议上将 AI 素养界定为关于 AI 的知识、理解、技能和价值观^[31]。可见,当前对 AI 素养框架的研究涵盖了意识、知识、技能、思维、态度、伦理、价值观、评估、创建、社会文化等诸多要素,同一要素概念在不同研究和实践背景下内涵外延各有差异,各要素间的逻辑关系还不甚明确。

基于学习进阶理论及社交网络观点,Wilson 等提出了名为 KSAVE 的信息技术与通信(ICT)素养模型,由知识、技能、态度、价值观和伦理构成^[32]。其中,“态度”所指与美国国家研究委员会的技术素养观强调的“思考方式”相近,可以理解为适应动态技术环境的思维方式;价值观和伦理都反映了人文主义因素,在联合国教科文组织、欧盟等权威机构政策文件中,一般用“伦理”概念统驭。参考前述文献及 KSAVE 模型,从 AI 素养概念内涵出发,基于通用性、可操作性、稳定且兼具灵活性原则,强调 AI 意识、AI 思维在适应当前环境剧烈变革和保持终身学习中的独特重要性,本文提出了一个包括 AI 意识、AI 认知、AI 技能、AI 思维、AI 伦理等五个主要构成要素的通识性 AI 素养框架(见图 1)。这五个要素既有各自相对独立的内涵和边界,又相互联系、相互影响。其中,AI 意识是先导和基础;AI 认知在 AI 意识基础上实现对 AI 概念原理、主流应用深入、全面的理解;AI 技能将 AI 意识和认知转化为实践能力,用于解决问题;AI 思维强调高阶思维模式和方法论构建,以支撑 AI 协同创新创造、应变决策和复杂决策;AI 伦理则贯穿始终,构成所有要素的底层支撑,确保个体在提升 AI 意识、认知、技能和思维的过程中贯彻“以人为本”原则,实现自我发展和社会价值的统一。与 AI 动态发展一致,个体应保持终身学习,AI 认知、AI 技能和 AI 思维的提升触发更高层次的 AI 意识,进而驱动新一轮认知学习、实践应用、思维成长和行为习惯,形成

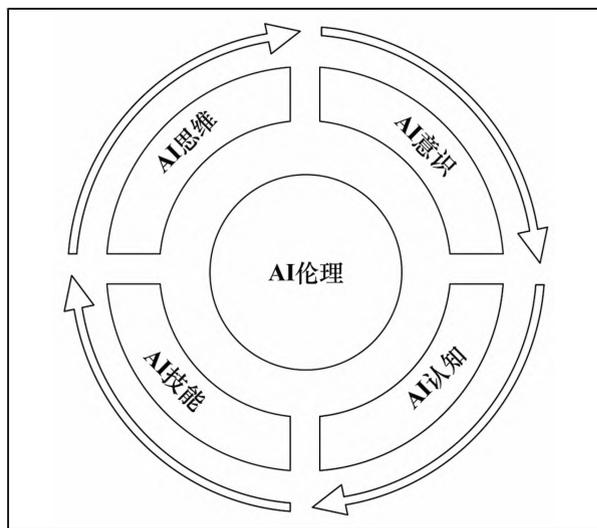


图 1 AI 素养框架

一个螺旋式上升的 AI 素养发展循环。

3.1 AI 意识

AI 意识是指对 AI 发展的敏感度和对 AI 技术存在、应用及社会影响的基本感受和认识,与形成对 AI 的态度密切相关。鉴于前文所提到的人机双边互动范式转变带来的影响,这种心理准备是公众主动使用 AI 作为解决现实生活中问题的工具所必需的^[16],以往的研究往往对此重视程度不高。AI 正从最底层改变人类社会经济与生活运作方式,相关应用的爆发式、非连续性增长导致在社会意识层面存在严重断层,先行先用和无知无觉并存。提高公众 AI 意识,避免对 AI 技术的盲目排斥或过度依赖,是构建一个更加公平、可持续和创新驱动的智能社会的关键步骤。

(1)AI 影响意识。个体应对 AI 应用保持好奇心,保持开放和包容心态。深度嵌入电子产品、家居设备、办公软件、社交媒体、交通工具、生产装备的 AI 应用已成为社会生态系统的一部分。只有感受、关注、意识到 AI 在当下经济和社会中的普遍性存在和影响,才能主动提升 AI 认知,具备利用 AI 解决学习、生活和工作问题的动机。

(2)AI 需求意识。如同前文在 AI 素养概念内涵中讨论到的,AI 是一种使能性构建。个体应具备与 AI 优势互补的主体意识、合作意识和共创意识,善于将 AI 与实际生活、学习和工作联系起来。个体持续、成功利用 AI 解决现实问题能够增强自我效能。

(3)AI 安全意识。与后续将要讨论的 AI 伦理相关,个体应当注意到 AI 的局限性、潜在风险,警惕 AI 产生的负面影响,保护自身与他人的利益、隐私和安全。如短视频和社交媒体中的 AI 算法往往被设计成多巴胺最大化器,容易带来成瘾性,对未成年人身心健康产生重大威胁,需要受到社会、家庭和未成年人自身的关注。

(4)AI 自我评估意识。AI 意识还包括元认知素养,自我反思心态应该被纳入 AI 素养概念构成^[30]。个体应该能够意识到自身 AI 素养的水平和状态,恰当地衡量自身 AI 素养与所处职业、学习和生活环境需求的匹配程度。

3.2 AI 认知

AI 认知是 AI 素养的核心组成部分,涉及对 AI

基本概念、技术原理、应用场景的理解以及对 AI 技术发展趋势的洞察,要求个体具备更深层次的理解、分析、预判 AI 技术和应用的能力。

(1)AI 基本概念和技术原理认知。概念理解是公众在日常生活中应用 AI 的基础。除了作为 AI 用户,还应该理解背后的基础技术原理。理解的程度是区分专业教育和公众教育的关键,原则上,不需要任何先验知识的 AI 概念构成了 AI 素养的最基本水平^[20]。以下三个方面的基本概念和技术原理应予以关注:①AI 模型。能够理解主流技术范式中 AI 算法、无监督学习算法和神经网络的基本逻辑;理解不同 AI 模型的不可解释性、学习能力、自主性程度;理解 AI 模型如何运行和做出决策,以及如何响应环境;了解知识在 AI 中的表示方式以及这些知识表示的局限性;理解不同类型、不同大小 AI 模型的优劣势。②用于 AI 的数据。由于 AI 的引入,数据与技术之间的关系发生了变化,因为 AI 可以从数据中学习,不再完全依赖专家编码。AI 越来越能够处理非结构化数据,这给数据治理带来了新挑战^[33]。数据知识基础包括数据结构、数据清洗等基本概念,以及理解数据对不同 AI 应用的意义。个体应理解如何收集和处理好数据;能够适当解释 AI 输出的数据;能意识到输出的数据可能存在的偏见^[34]。③AI 接口。个体应该能够理解 AI 与其环境互动的基本原理;认识到作为 AI 交互虚拟或实体界面的传感器的类型、功能、性能及分布;理解不同 AI 技术如何通过它们各自的接口类型接收和输出。

(2)AI 工具及应用场景认知。各类 AI 工具在工业、教育、交通、医疗、金融证券、科研、文化娱乐等多个领域的应用不断深入。个体应该能够理解不同 AI 模型和常见 AI 工具之间的联系;理解 AI 在不同领域的应用工具及其基本原理、功能范畴和典型案例;了解如何正确操作 AI 工具;能够敏锐识别 AI 工具的存在与运行状态;了解如何在与 AI 工具互动时管理自己的数据和隐私。

(3)AI 趋势认知。AI 技术正以前所未有的速度发展,为适应不断变化的环境和需求,个体应持续学习以保持对新兴技术的洞察,能判断 AI 技术、应用发展的主要趋势。

3.3 AI技能

AI技能的培养侧重于实践操作层面,要求个体能够将理论知识应用于识别、评估、应用和整合AI技术、工具和资源以达成特定目标,解决数字生活、学习、工作和社交场景中的实际问题,提升个人效率、优化决策过程,成为AI世界的积极参与者。

(1)AI选择评估。AI产品逐步多样化。基于实际问题,个体应能评估不同AI工具的功能、优势和局限性;能够综合考虑前述AI认知中关于模型、数据、工具、接口、硬件等多方面的因素,识别、选择适合需求的AI工具;能够评估AI项目风险和收益,制定合理的管理策略。

(2)AI交互使用。由于AI的技术特性,AI的产出效率很大程度上取决于用户实践水平。如面对同一个大语言模型对话应用,以不同方式构建提示词,可以引导AI沿着数据集中的不同路径前进,从而得到不同水准的输出。个体应该能够熟练使用与自己面对的日常生活、工作应用场景密切相关的主流AI智能助手、语言处理、办公辅助、多媒体制作、推荐系统、智能设备等应用;能够整合使用不同类型的AI工具。

(3)AI产出评估。AI技术的复杂性和AI运行过程的“黑箱”特征强化了产出评估的重要性。执行基于AI的解决方案后,个体应该能够谨慎衡量和评估结果,对解决方案的优点和不足形成明确的见解,以完善当前解决方案或设计新的解决方案^[35],不断提高AI交互和问题解决能力。

3.4 AI思维

AI思维是AI认知在逻辑和方法论层面的进阶。狭义的AI思维是对算法、模型的应用,属于AI技术的子集,其基本概念、原理已涵盖在前述AI认知要素。此处所指AI思维主要为个体在使用AI技术、适应AI社会环境过程中应具备的思维模式和方法论,强调运用底层逻辑、系统思维进行创新思考、协同共创,以支撑应变决策和复杂决策。

(1)计算思维和数据思维。二者是一体两面的关系。计算思维在AI素养框架下主要是一种跨学科的通用思维方式,通过一系列步骤、规则,逻辑、系统地分析和处理信息,以找到最优解决方案,包括分解、模式识别、抽象和算法设计等核心要素。计算思维包

含了算法思维,理解以算法为支撑的信息处理过程有助于个体在智能时代实现对信息的管理、控制和表达。数据是新一代AI深度学习的“石油”。突破传统的直觉或经验驱动,数据思维要求个体能够理解、应用基于数据驱动的决策方法。公民应能理解数据质量、来源以及数据的潜在用途和局限性。

(2)批判性思维。AI虽然在某些任务上表现出色,仍然受限于设计意图、开发水平和训练数据集质量。近年来兴起的大语言模型往往具有百亿、千亿甚至万亿级参数,解释其运算、输出过程和可靠性更为困难,系统产生的“幻觉”、错误极有可能造成严重后果;物理真实与数字虚拟之间的界限愈加模糊,给应用安全和社会治理带来重大挑战。具备AI思维的个体应该能够批判性地评估与AI互动的需求、过程和结果,保持独立判断。

(3)设计思维。相比人类个体,AI具有无限学习和认知能力,对历史数据依赖性很强,是一种典型的连续型学习,最大优势是知识广度和产出速度,智能“涌现”能力仍待发展。人类除了会从经验中学习,还会创造,即跳跃型学习。设计思维是指将AI视为设计创新合作伙伴,充分发挥主客体优势,创造性形成多种解决问题的思路与方案。

3.5 AI伦理

AI伦理关注有道德地使用AI概念和应用,与智能社会中的法律、文化、道德和社会价值问题密切相关,其复杂性在于需要平衡技术与人类价值的关系。人类行为本质上受价值驱动,以人类利益为导向来构建和应用AI至关重要。AI更擅长的是输出“是什么”,而在解释“为什么”的时候则有着相当明显的价值观倾向^[1]。AI的加速发展可能带来不可测、不可控的危险,对信息安全、人权和隐私构成重大风险^[31]。Fei-Fei和Etchemendy指出,“与核能和生物技术等少数‘两面性’技术不同,AI开发和使用遍布全球,进入门槛相对较低,我们无法控制如此分散的事物。”^[3]AI开发体系下的“AI对齐”要求AI系统目标要和人类价值观相一致,同样,个体使用AI时应该能够充分考虑其社会影响。

(1)伦理识别。AI技术的发展带来了新的伦理和法律问题,类似于AI灰产、数据和网络安全、隐私和

IP 侵犯、偏见和歧视、网络诈骗、虚假新闻和舆论操纵、剧烈的就业替代、对劳动力的过度虚拟控制、对多样性的损害、用于创建生物、网络和自主武器等领域的滥用等既有被认知的风险引起了人们的广泛担忧,智能助手、数字人、脑机接口、服务机器人的渗透可能改变传统社会关系和伦理文化。个体应该能够识别、描述与上述 AI 相关的主要伦理议题、伦理风险。

(2) 伦理利用。在伦理识别的基础上,个体应以积极、理性的态度面对 AI 带来的影响,探索 AI 合理应用,采取有效措施消解、规避、管理负面后果,提升识别虚假信息、防范 AI 欺诈、保护个人数据和隐私安全、依法维权的能力,恰当应对 AI 应用中遇到的安全问题和道德困境;严格遵守《生成式人工智能服务管理暂行办法》^[36]等有关法律法规,成为负责任的 AI 用户,避免侵害他人合法权益和社会福利。

(3) 伦理政策参与。由于存在不可预测的风险和挑战,诸如自动驾驶等很多关键 AI 应用的商业化进展主要并不取决于技术,而取决于社会公众的认知、参与和支持程度。个体应该能够理解 AI 伦理相关政策议题,以确保未来 AI 技术的使用和发展符合包容性、公平获取和最小化偏见原则^[22]。随着 AI 技术在

教育、医疗、交通等公共服务、民生领域的应用,公众还需要具备参与政策讨论和建议的能力,积极参与道德环境维护、弱势群体关怀和伦理政策建设,履行智能时代的积极公民责任。

4 结语

AI 是引领新一轮科技革命和产业变革的引擎和重要驱动力,正深刻改变社会生产、生活和社会组织方式。全社会都应该为技术进步做好准备,确保 AI 技术得到负责任的发展和普惠应用。正如 OpenAI 首席执行官 Altman 在参加斯坦福大学企业思想领袖论坛交流时所说,“让社会与技术共同发展。让社会告诉我们,无论是集体还是个人,他们希望从这项技术中获得什么。”^[37]显然,只有具备基本 AI 素养的个体以及由这些个体构成的社区和组织有能力真正参与其中。AI 素养将是智能时代个人竞争力的重要体现,也是国家和地区科技进步、产业发展和社会福利提升的关键驱动力。全面构建面向社会各个层次的 AI 素养教育体系是当务之急。后续可进一步研究面向多样化群体的 AI 素养整合模型构建、培育体系设计和评测系统开发。

参考文献:

- [1] 陈一,徐衡.共识、共舞、共进:智联社会的人工智能素养[J].青年学报,2023(3):38-42.
- [2] 国务院.2024 年政府工作报告[EB/OL].[2024-05-02].https://www.gov.cn/yaowen/liebiao/202403/content_6939153.htm.
- [3] Fei-Fei L, Etchemendy J. A common goal for the brightest minds from Stanford and beyond: putting humanity at the center of AI[EB/OL].[2024-05-20].<https://hai.stanford.edu/news/introducing-stanfords-human-centered-ai-initiative#SnippetTab>.
- [4] NAIAC. Recommendations: Enhancing AI Literacy for the United States of America[EB/OL].[2024-05-22].https://ai.gov/wp-content/uploads/2023/12/Recommendations_Enhancing-Artificial-Intelligence-Literacy-for-the-United-States-of-America.pdf.
- [5] Zurkowski P G. The information service environment relationships and priorities[EB/OL].[2024-05-07].<https://files.eric.ed.gov/fulltext/ED100391.pdf>.
- [6] UNESCO. Media and information literacy: policy and strategy guidelines[EB/OL].[2024-05-07].<https://unesdoc.unesco.org/ark:/48223/pf0000225606>.
- [7] Eshet-Alkalai Y. Digital literacy: A conceptual framework for survival skills in the digital era[J]. Journal of educational multimedia and hypermedia, 2004, 13(1):93-106.
- [8] European Parliament, Council of the European Union. Recommendation on key competences for lifelong learning[EB/OL].[2024-05-16].<https://www.britishcouncil.org/sites/default/files/youth-in-action-keycomp-en.pdf>.
- [9] European Commission, Joint Research Centre, Vuorikari R. et al. DigComp2.2, The Digital Competence framework for citizens[EB/OL].[2024-05-25].<https://op.europa.eu/en/publication-detail/-/publication/50c53c01-abeb-11ec-83e1-01aa75ed>

- 71a1/language=en.
- [10] Stanford HAI. Artificial Intelligence Index Report 2024[EB/OL]. [2024-05-26]. https://aiindex.stanford.edu/wp-content/uploads/2024/04/HAI_AI-Index-Report-2024.pdf.
- [11] Pinski M, Benlian A. AI literacy for users—A comprehensive review and future research directions of learning methods, components, and effects[J]. *Computers in Human Behavior: Artificial Humans*, 2024, 2(1):1–22.
- [12] Schuetz S, Venkatesh V. Research perspectives: The rise of human machines: How cognitive computing systems challenge assumptions of user–system interaction[J]. *Journal of the Association for Information Systems*, 2020, 21(2):460–482.
- [13] Konishi Y. What is Needed for AI literacy? Priorities for the Japanese economy in 2016[EB/OL]. [2024-05-08]. https://www.rieti.go.jp/en/columns/s16_0014.html.
- [14] Kandlhofer M, Steinbauer G, Hirschmugl–Gaisch S, et al. Artificial intelligence and computer science in education: From Kindergarten to University[C]. 2016 IEEE Frontiers in Education Conference (FIE), 2016:1–9.
- [15] Long D, Magerko B. What is AI literacy? Competencies and design considerations[C]. *Proceedings of the conference on human factors in computing systems proceedings*, 2020:1–16.
- [16] Dai Y, Chai C–S, Lin P–Y, et al. J. Promoting students’ well–being by developing their readiness for the artificial intelligence age[J]. *Sustainability*, 2020, 12(16):6597.
- [17] Hermann E. Artificial intelligence and mass personalization of communication content——an ethical and literacy perspective[J]. *New Media & Society*, 2021, 24(5):1258–1277.
- [18] Kong S–C, Cheung M–Y W, Zhang G. Evaluation of an artificial intelligence literacy course for university students with diverse study backgrounds[J]. *Computers & Education: Artificial Intelligence*, 2021, 2:100026.
- [19] Laupichler M C, Aster A, Schirch J, et al. Artificial intelligence literacy in higher and adult education: A scoping literature review[J]. *Computers & Education: Artificial Intelligence*, 2022, 3:100101.
- [20] Ng D T K, Luo W, Chan H M Y, et al. Using digital story writing as a pedagogy to develop AI literacy among primary students[J]. *Computers & Education: Artificial Intelligence*, 2022, 3:100054.
- [21] UNESCO. K–12 AI curricula: a mapping of government–endorsed AI curricula[EB/OL]. [2024-05-23]. <https://unesdoc.unesco.org/ark:/48223/pf0000380602>.
- [22] Wang B, Rau P L P, Yuan T. Measuring user competence in using artificial intelligence: Validity and reliability of artificial intelligence literacy scale[J]. *Behaviour & Information Technology*, 2023, 42(9):1324–1337.
- [23] Carolus A, Augustin Y, Markus A, et al. Digital interaction literacy model—conceptualizing competencies for literate interactions with voice–based AI systems[J]. *Computers & Education: Artificial Intelligence*, 2023, 4:100114.
- [24] Pinski M, Benlian A. AI literacy—towards measuring human Competency in artificial intelligence[C]. 56th Hawaii international conference on system sciences, 2023.
- [25] Cetindamar D, Kitto K, Wu M, et al. Explicating AI Literacy of Employees at Digital Workplaces[J]. *IEEE Transactions on Engineering Management*, 2024, 71:810–823.
- [26] 钟柏昌, 刘晓凡, 杨明欢. 何谓人工智能素养: 本质、构成与评价体系[J]. *华东师范大学学报(教育科学版)*, 2024, 42(1):71–84.
- [27] Pearson G, Young A T, National Research Council(U.S.). *Technically speaking: Why all Americans need to know more about technology*[M]. Washington, DC: The National Academies Press, 2002.
- [28] Veloso M, Biswas J, Coltin B, et al. Symbiotic–Autonomous Service Robots for User–Requested Tasks in a Multi–Floor Building[EB/OL]. [2024-05-18]. <https://www.cs.cmu.edu/~mmv/papers/12irosw-VelosoEtAl.pdf>.
- [29] Heyder T, Posegga O. Extending the foundations of AI literacy[EB/OL]. [2023-05-12]. https://www.researchgate.net/publication/357511112_Extending_the_foundations_of_AI_literacy_Research-in-Progress.
- [30] Chiu T K F, Ahmad Z, Ismailov M, et al. What are artificial intelligence literacy and competency? A comprehensive framework to support them[J]. *Computers and Education Open*, 2024, 6:100171.

- [31] UNESCO.International Forum on Ai and Education:Ensuring AI as a Common Good to Transform Education [EB/OL].[2024-05-12].<https://en.unesco.org/sites/default/files/ai-in-education-forum-2021-cn-en.pdf>.
- [32] Wilson M,Scalise K,Gochyyev P.Rethinking ICT Literacy:From Computer Skills to Social Network Settings [J].Thinking Skills and Creativity,2015,18:65-80.
- [33] Vazhayil A,Shetty R,Rao B,et al.Focusing on teacher education to introduce AI in schools:perspectives and illustrative findings [C].Proceedings of the 2019 IEEE tenth international conference on technology for education (T4E),2019:71-77.
- [34] Melsión G I,Torre I,Vidal E,et al.Using explainability to help children understand gender bias in AI [C].Proceedings of the 20th Annual ACM Interaction Design and Children Conference,2021:87-99.
- [35] Schraw G,Crippen K J,Hartley K.Promoting self-regulation in science education:Metacognition as part of a broader perspective on learning [J].Research in Science Education,2006,36(1):111-139.
- [36] 国家互联网信息办公室.生成式人工智能服务管理暂行办法 [EB/OL].[2024-05-13].https://www.gov.cn/zhengce/zhengceku/202307/content_6891752.htm.
- [37] Stanford eCorner.The Possibilities of AI [EB/OL].[2024-05-03].<https://ecorner.stanford.edu/videos/the-possibilities-of-ai-entire-talk/>.

作者简介:尹开国(1979-),男,武汉工商学院管理学院副教授,研究方向:创业信息素养、数字伦理、数字商业模式。