第 3 章

人工智能伦理概述

3.1 IEEE 发布的《人工智能设计的伦理准则》

美国时间 2017 年 12 月 12 日上午 9 时，电气电子工程师学会（IEEE）于全球发布了第 2 版的《人工智能设计的伦理准则》白皮书（“Ethically Aligned Design” v2），旨在指导我们认识这些技术可能造成的技术外的影响，确保人工智能设计能够符合人类的道德价值和伦理原则，并充分发挥系统益处。其所提供的一些洞察和建议，可为未来从事相关科技领域的技术专家的工作提供重要参考，同时促进符合这些原则的国家政策和全球政策的制定。其目标在于指导人类合乎伦理地设计、开发和应用人工智能技术，并确保它们不侵犯国际公认的人权，在技术的设计和使用中优先考虑人类福祉的指标；同时确保它们的设计者和操作者负责任且可问责，并以透明的方式运行，从而将滥用的风险降到最低。

IEEE 所发布的《人工智能设计的伦理准则》主要包含以下十一个标准：解决系统设计中的伦理问题的建模过程、自主系统的透明性、数据隐私的处理、算法偏见的处理、儿童与学生数据治理标准、雇主数据治理标准、个人数据的 AI 代理标准、伦理驱动的机器人和自动化系统的本体标准、机器人智能与自主系统中伦理驱动的助推标准、自主和半自主系统的失效安全设计标准、合乎伦理的人工智能与自主系统的福祉度量标准。

具体来说，IEEE 所发布的《人工智能设计的伦理准则》明确了个人数据权利和个人访问控制，人们有权决定其个人数据的访问权限，有权利用知情同意控制其个人数据的使用；个人需要各种机制帮助建立、维护其独特的身份和个人数据；同时需要其他政策和做法，使他们能明确知晓融合或转售其个人信息将产生的后果。除此之外，

IEEE 希望通过经济效应增进福祉，也就是通过价格合理的通信网和互联网的接入，智能与自主技术系统可以为任何地方的人群所用并使其受益。它们可以显著改变制度和制度性关系，使其朝着更加以人为本的结构发展；它们还能促进人道主义和发展问题的解决，从而增加个人和社会的福祉。另外，问责的法律框架不可或缺，智能系统与机器人技术的融合带动了系统的发展，这类系统模仿人类，具有部分自主性，有完成特定智力任务的能力，甚至还可能拥有人类的外貌。因此，复杂的智能和自主技术系统的法律地位问题与更广泛的法律问题交织在了一起，这些法律问题涉及如何确保问责制，以及当这类系统造成损害时如何分配责任。例如，智能与自主技术系统应适用相关的财产法，政府和行业利益相关者应该确定哪些决策和操作绝不能委托给这些系统，并制定规则和标准，以确保人类能够有效地控制这些决策，以及能够有效地为造成的损害分配法律责任。虽然自我完善的算法和数据分析可以使影响公民的决策自动化，但法律应该强制要求透明性、参与性和准确性，包括必须允许当事人、其律师和法院可以合理地获取政府和其他国家机关采用这些系统所产生和使用的所有数据和信息，如果可能，系统中嵌入的逻辑和规则必须对监管人员开放，并接受风险评估和严格测试，系统应当生成用于决策的事实和法律的审计数据，并服从第三方核查，同时公众有权了解是谁通过投资制定或支持关于这类系统的伦理决策。而有效的政策应当保护和促进安全、隐私、知识产权、人权和网络安全，以及公众对智能与自主技术系统对社会的潜在影响的认识。为确保政策最符合大众利益，这些政策应当：支持、推广和实施国际公认的法律规范，同时提升劳动力在相关技术方面的专业知识，另外产生对研究和开发的引领作用，并制定规则，以确保公共安全和问责，最后教育公众知悉相关技术的社会影响。

IEEE 同时也回应了对未来技术的关切，包括设计造成人身伤害的自主系统，与传统武器或设计不造成伤害的自主系统相比，需要考虑更多的伦理维度。如确保武器系统在人类有效的控制中，自动化武器的设计应包含供审计的追踪数据，以确保可问责和可控，包含自适应和可学习的系统，以透明和可理解的方式向操作人员解释其推理和决策，培训自主系统的操作人员，其身份应可清晰识别，同时自动功能行为的实现对操作人员而言是可预测的，并确保技术开发人员能够理解其工作的后果，制定职业伦理守则，妥善处理有意造成伤害的自主系统的开发等；而针对通用人工智能（AGI）和超人工智能的安全性和有益性问题，IEEE 认为与其他强大的技术一样，智能和自我改善的技术系统的开发和使用涉及相当大的风险。这些风险可能来源于滥用或不良设

计。然而，根据某些理论，当系统接近并超过 AGI 时，无法预料的或无意的系统行为将变得越来越危险且难以纠正。并不是所有 AGI 级别的系统都能与人类利益保持一致，因此，当这些系统的能力越来越强大时，应当谨慎并确定不同系统的运行机制。针对情感计算，为确保智能技术系统在各种情况下最大可能地服务于人类，以及参与或服务人类社会的人造物不能通过放大或抑制人类的情感体验造成伤害，即使是在一些系统中设计的初步的人工情绪，也会影响决策者和公众对它们的理解。最后，针对混合现实，随着混合现实技术在我们的工作、教育、社会生活和商业事务中的应用越来越普遍，混合现实技术可能会改变我们对身份和现实等概念的理解。尤其是随着技术从耳机转向更加精细、更整合的感官增强设备，混合现实世界的实时个性化能力也有可能引发有关个人权利和个人多重身份管理的伦理问题。

3.2　机器道德

我们从人工智能的发展中受益颇丰，但我们也应当重视其中潜在的诸多伦理、道德与法律方面的风险。一方面，伴随人工智能渗透所出现的达芬奇手术机器人医疗事故责任界定、“微软小冰”诗集的著作权归属、Uber 无人机动车致人死亡的责任归属与责任承担等一系列问题仍未有明确答案。为了规避人工智能的道德风险，机器道德的思想开始出现并得到很多人的支持；但这种观点是出于情感而非理性；而机器道德的提出，源于人类对人工智能技术发展的担忧和恐慌，希望机器能够拥有道德，从而使机器在服务于人类的同时，又不会伤害人类。机器道德是人类道德在人工智能时代的扩展，而人工智能始终是人工的产物，其是否拥有人类的社会性，学界莫衷一是。另一方面，将人工智能放在法律层面研究，脱离不开人工智能如何担责这一基本范畴，而解决这一问题的前提则是明确人工智能的法律地位以及其与人的关系；尤其是未来具有自我意识和自主决策能力的强人工智能对现有司法体系带来的冲击和挑战，对人格权和人类主体地位的动摇，以及对人类合法权利的影响。

与传统伦理学的研究对象——人的道德实践相比，自动驾驶汽车等智能机器的道德实践具有显著的特殊性：首先，机器行为是人为设计的，由计算过程所决定，并非自然的因果过程，其区别于一般的意外或自然灾害；其次，机器行为在很大程度上不由其

使用者决定，这使得机器不同于传统意义上的工具；最后，对于具有广泛适用性、所需处理复杂情况的机器，尤其是拥有学习能力的机器，设计师和制造商很难或原则上不可能准确预料机器行为，或是对机器行为进行控制。这种特殊性导致传统的道德责任归属不适用于智能机器。以自动驾驶汽车为例，如果一种具有较高可靠性的自动驾驶汽车反常地违背了交通规则或者造成了交通事故，那么将事件视作意外，要求乘客和设计师负责的做法都不能让人信服。赋予机器以道德主体地位就是让机器为自己的行为决策负责。

要机器为自己的行为决策负责，前提是机器具有道德主体性是成立的。这意味着，机器具有自由意志且行为是自由的，因为仅当机器的行为是自由的，将道德责任归属于机器的做法才是合理的。

3.2.1 为人工智能赋予道德主体地位

人工智能具备以下与人类高度相似的特征：①理性条件——人工智能可以基于训练学习得出的算法模型独立自主完成所设定的任务，同时在进行任务的同时也能基于实时反馈的数据实现无监督学习和进化，这与人类的学习行为相似；②道德内化——人类的道德行为准则并非与生俱来，而是后天耳濡目染；如果将道德行为准则量化成计算机可以识别的数据并输入给人工智能，其也可基于本身学习特性习得；③人工智能具备享有相应权利、承担义务职责的权利的可能；④强人工智能或许拥有与人类相似的“自由意志”，并通过“认知—学习—创造”实现与人类相似的生产行为。

人之所以有人格，是因为其具有理性的特质；而人工智能具备与人类似的理性本质。在反对“人类中心论”观点的学者看来，从道德出发我们应当像对待人一样对待所有的理性存在者，只要这种理性关系存在，无论人工智能在生物学方面与智人有何不同，不尊重其人性或以不人道的方式对待之都是错误的。如果我们不遵循这一法则，又有什么理由平等对待所有具有智慧的生命体呢？这一关系的建立前提是该对象能够像人一样服从道德，而不会对社会造成危害，或者其对社会的危害总体可控，才能认定人工智能具有道德能力，才符合基于伦理人格主义所确立的标准；而那些不具备道德能力的可能威胁人类种族利益的人工智能则不应被视为道德主体。需要注意的是，是否为人工智能赋予道德主体与人工智能应当遵循正确的价值观并不必然联系，即便人工智能不具备道德主体地位，正确的道德与价值准则都是人工智能必须遵守的；随着人工智能的能力不断拓展，其所承担的责任也越大；同时，基于道德主体的约束也会限

制人工智能对于某些非人类应当承担之责任的承担；最后，在确立人工智能具有道德能力之前，应当设立评判人工智能具有道德能力的可靠标准。

人工智能具有道德主体地位的必要不充分条件：认知能力、意思能力与道德能力

认知能力主要指人工智能具有像人类一样的可以有效感知客观物质世界和精神世界的能力，其所涵盖的要素包括记忆、好奇、联想、感知、自省、想象、意向、自我意识等。认知能力可以具象成人工智能通过传感器接收环境和周围其他对象的信息，并在自我计算的过程中加以运用和创造，这也是理性构建的基石。

当然，仅有认知能力是远远不够的。被动地服从程序设定者订立的目标会让人工智能与工具相比并无二致，也无法获得道德主体地位。人工智能需要为自身设定行动目标，也就是具备“意思能力”。所谓意思能力，是指行为能力，也就是责任能力，意味着“人依其本质属性，有能力在给定的各种可能性范围内自主地和负责地决定他的存在和关系，为自己设定目标并对自己的行为加以限制”。该能力包括沟通、审慎、选择、决定、自由意志等多个方面。

尽管人工智能在此情形下的智慧与认知水平和人类无差别，但因为缺乏伦理道德的规约，其行为难以将人类权益作为中心，可能对人类社会秩序、核心利益乃至生存造成严重危害。关于此命题，可以援引菲尼斯曾提出的一项原则：如果损害公共善的潜在风险是由同胞的行为造成的，则我们必须接受这种风险；如果是由同胞之外的人造成的，则不必接受。这就使得人工智能需要具备像人一样的道德能力，尽可能降低或者消除其对人类种族利益危害的风险。道德能力是指像人一样在道德规范与伦理规约的框架下严格按照道德实践判断决策控制行为的能力，每个人具备权利的前提是其本质上是一个具有伦理意义的人，任何人若想获得相对于其他人的道德主体地位，既有权要求别人尊重他自己的人格，也有义务尊重别人的人格。具备道德能力的主体有充分的道德理由要求他人尊重其法律人格，而不具备道德能力的存在者则没有尊重他人人格的能力，也不需要被他人所尊重。

3.2.2　人工智能道德主体地位之批判

尽管所谓“具有自由意志”的“强人工智能”已然成为“主体论”簇拥者热议的对象，但当前人工智能发展的以下三点特性，不得不让我们质疑持有该观点的学者陷入了一种“空想主义”的状态：从研究方向上来说，对人工智能的研究并未明确朝向强人工智

能领域，且强人工智能现在来看依然是一个虚无缥缈的方向，也不会对人工智能对人类的价值与效益带来显著提升；从技术架构上来说，无论是以半导体材料为基础的算力架构，还是以二进制逻辑为基础的算法开发，都难以实现人工智能的“自主进化”，“强人工智能”诞生的基础是人类需要突破现有逻辑电路和数据结构桎梏，而这一点在短期之内几乎不可能实现，也缺乏足够的理论依据；从研究原则上来说，无论是科学研究，还是工业发展，都需要严格遵守伦理和道德原则，即便强人工智能可以被实现，其风险也远大于收益，这一潘多拉魔盒不应被打开。

人工智能是由人类制造的，基于智能科技而非生命代谢的非生命体。首先，人工智能缺乏人类的“碳基”大脑等生理基础。人类大脑是人类生命的核心，大脑死亡意味着法律人格的丧失。而“硅基”的人工智能程序抑或是人工智能机器人都不符合基于现有道德框架下主体人格的生理基础。其次，人工智能不具有理性和意志。理性包括人对周遭事物特性和规律的感知、道德伦理的认识与遵从，以及情感的同理。康德提出“人是理性的，其本身就是目的论断”，黑格尔认同“人因理性而具有目的”，人因为理性而具有自省、自审的能力。不同于人类的理性，人工智能的“理性”本质上来说是一种基于算法模型的逻辑关系，更不存在人类的“天赋”理性，其“理性”存在之价值是实现人类意愿、目的，更不可能实现自省、自审，无法挣脱人类从算法层面为其规制的限制以进行批判性反思；即便人工智能具有深度学习能力，这种基于算法的修正本质上还是一种模式识别，而非真正拥有与生俱来的理性，这种“智慧”是人类赋予的算法模拟表象。

意志也并非单纯的逻辑演算，其本身包含欲望和行动两大关键因素；人类在欲望的驱动下针对所面临的现实命题进行伦理价值判断，人对生存、生活质量、身份人格等方面的欲望和追求驱动人类不断改造自身与自然创造更良好的生存环境或做出趋利避害的判断。但人工智能没有从自身和公共角度出发的欲望，更没有情感的羁绊，面对“电车难题”等类似命题时，所做出的决策难以被评估合理程度，大多数情况下，人工智能只是在人类算法的基础上进行概率演算，更不用谈脱离人类建立自由意志的可能性了。

针对将人工智能视为“非完全行为能力的人”，有批评观点认为该思路充满对人类社会中缺失理性和意志的特殊人群的歧视，缺乏对特殊人群的保护，本质上就是一种非理性的体现；小孩在成年之后也会具有理性和意志，大脑损伤的植物人也可能苏醒并重新获得理性和意志，这显然与人工智能的特质相矛盾。

最后还要强调的是，现有道德框架下人工智能无法承担责任，赋予其主体地位只是一种形式上的规制，但人工智能对财产的控制及拥有，究竟是归属人工智能本身，还是归属制造人工智能的人类主体，这些命题都存在大量缺口和空白；而诸如“有限人格论”等本质上已经自证赋予人工智能道德主体资格的不必要性，无法通过“奥卡姆剃刀原则”的拷问。

3.3　伦理风险评估

无论是深度学习、强化学习，还是联邦学习，人工智能的基本逻辑都离不开对大量真实场景下的数据训练模型的运用，也就是用户数据的采集、使用、共享与销毁。这就使得我们在推广人工智能的同时，也需要关注以下伦理风险。

3.3.1　人工智能威胁用户隐私安全

1. 以数据为基础的人工智能产业链环环相扣、错综复杂，数据分析的采集终端、处理节点、存储介质与传输路径不确定，风险高

鉴于以数据为基础的人工智能产业链环环相扣、错综复杂，且基础训练数据作为人工智能的生产要素贯穿整条产业链，数据分析的采集终端、处理节点、存储介质与传输路径不确定，风险高，用户缺乏对上述环节的控制、监督和知情；同时，行业也缺乏统一标准和行为规范来限制相关企业在以上四个环节的权责分配和行为合规，这一部分隐私数据的使用边界与安全保护完全取决于人脸识别服务提供商自己的安全技术素养、道德规范与行业自律，用户对自己的人脸隐私数据完全失控，而失去对个人隐私数据的掌控可能导致用户处于隐私透支的境地，并对现代人工智能技术产生抵触情绪。

2. 强人工智能产品可能诱导用户主动泄露个人数据，人类失去对隐私与敏感数据边界的控制

除了用于人工智能底层算法模型训练的数据和为支持个性化产品服务而采集到的数据容易遭遇泄露、篡改等高风险情形外，还有一种挑战隐私伦理的可能是强人工

智能产品突破其设计伦理底线，主动甚至强迫用户提供原本对学习过程没有帮助的或高度敏感的隐私数据，尤其是基于强人工智能设计的各类机器人或同伴机器人，一旦未来人工智能技术发展到能够赋予该类教育人工智能产品自主意识与行为决策判断能力，该人工智能就可能在主观恶意或客观操控的情况下，引导、胁迫用户泄露与自己相关的敏感数据，并将其采集用作其他非公开活动。如何保证人工智能产品在规范框架下与人建立合理交互并不做出侵犯用户隐私信息，突破其职能界限的行为？面对人工智能技术的飞速发展，这一命题值得我们正视。

3. 人工智能的滥用造成隐私透支，人格尊严更易被贬损

以人工智能领域最常见的生物识别为例，面对处置用户生物识别信息命题，以苹果为代表的企业宣称自己仅将用户的指纹、人脸数据存储在终端设备本地，并采用物理加密的方式确保该数据的整个调用过程完全脱敏和本地化存储；但更多的厂商则是将人脸数据作为用户画像的一部分在线上随意流转传输，常见的网络传输与加密协议显然无法与人脸数据的安全级别与敏感程度相匹配，人脸识别应用的数据存储与传输流程均有可能被劫持，存在严重的安全风险；加之当前数据爬虫、网络入侵、数据泄露等已经成为互联网中的常态，收集人脸数据的供应商完全有可能主动或被迫在未经用户授权或超出用户协议许可的范围下，对用户面部数据进行采集、使用、流转等非法操作。例如，一旦人脸特征信息被不法分子拦截，或者从被攻破的本地加密存储中复制出来并运用在目标用户所使用的安全验证服务上，攻破用户自己使用的人脸识别服务并获得用户本人才具有的敏感权限（如金融交易、人脸门禁、手机解锁与计算机敏感数据访问等）可谓轻而易举；或者将该人脸数据结合深度合成技术，被不法分子用于伪造具有人格诋毁性质的多媒体内容，或者捏造虚假音视频片段恶意侮辱、诽谤、贬损、丑化他人，势必会对受害者造成极大的人格侮辱与内心创伤，并带来非常恶劣的社会影响，甚至被用于干预国家政治或执行军事行动。用户对人脸识别服务的依赖越深，隐私泄露事件对用户的影响也越显著，人脸识别的专属性与唯一性也导致后期受害者难以挽回与消除由人脸数据泄露或非法篡改利用而造成的损失与不良影响。

最后，人脸作为人在社交活动中最直观的身份认证，难以通过有效手段施加保护，这就导致用户在公共场合面临未经自己允许甚至不知情的情况下被非侵入性识别技术监控并抓取面部数据的风险，如 IBM 公司在被获取面部数据的自然人毫不知情的情况下从 Flickr 网站抓取近 100 万张照片用于训练人脸识别算法，这无疑是对用户隐

私的侵害。

由此可见，人工智能的潜在隐私风险不仅会造成用户隐私权益遭受侵害，还有可能面临由隐私泄露导致的名誉权、肖像权、姓名权、信用权等权利遭受侵扰，其中的系统性风险不容忽视。

3.3.2　人工智能侵犯用户生命财产等合法权益

以医疗机器人为例，培育一名医生需要其十几年的刻苦钻研与上千小时的临床经验，知识架构与经验体系紧密相扣，哪怕中间有一层出现缺失，都无法成就其白衣天使的角色；纵观整个医疗人工智能产业链，从基础层的数据分析与算力架构到技术层的算法和平台建设，再到应用层的场景开发，一项落地并商用的医疗人工智能产品，可能基于千万量级数据和上百万行代码，几万份零件来自十几家供应商的几十条生产线，背后牵连着太多不同领域不同方向的企业，所具备的行业背景、技术资源、产品判断标准等差异悬殊，对患者来说，这无疑为人工智能的安全性蒙上了一层焦虑的疑影，如人工手术需要对主刀医师和医疗器械进行消毒，而机器人参与的手术则需要对机器整体不留死角的消毒，考虑到机器的精密性与耐久性，我们无法保证对一台金属机器施行无死角的消毒，也无法保证是否会有别的有害物质感染。即便从产业链角度来说万无一失，安全性得到保障，但鉴于人工智能厂商不会透露他们的人工智能系统如何工作，在任何一种情况下，当传统人工智能生成决策时，人类用户无法及时获悉该决策是如何生成的，尤其处于高危医疗环境，信任人工智能的决策而不了解人工智能给出建议可能存在的潜在风险根源，患者都会面临极大的风险。例如，运用手术机器人执行高难度、高精确的外科手术时，即便在医生操控下出现毫厘的差错，都会直接导致手术失败与患者死亡，而此时无论是患者还是医生都无法获知该差错是现场操控机器人的医生的“直接人为错误”，还是由机器人制造商或算法提供商造成的“间接人为错误”，也可能是由于外部个人或组织主观恶意的物理攻击与远程干扰，或者电力、数据传输中断造成的意外事故，甚至医生与供应商、制造商都没有出错，只是人工智能算法黑箱特性或自主意识决策下的“机器错误”……患者安全无法得到切实保障，人工智能在医疗（尤其是临床）领域的应用将在伦理层面举步维艰。

3.3.3　人工智能动摇人公平参与社会活动，享有社会资源的权利

人工智能产业链上集中了芯片、传感器、算法、终端、行业应用、解决方案、安全加

密、网络传输等诸多相关方，技术滥用与不正当竞争的现象难以完全避免。一旦其中任何一方采取针对用户或者其他相关参与者的歧视性策略或不正当竞争举措，偏见与不平等最终会传递至终端消费者并损害其应当享有的公平权利；加之人工智能的硬件、算法、数据集等都具有较高技术壁垒，不公平一旦形成，则难以被打破；最后，由于机器学习的黑箱特性，模型所得结果往往具有一定的不可解释性，算法透明度存在较大争议，这可能导致基于人工智能的行为决策与训练数据之间的相关性，以及该相关性对利用人工智能解决此项命题的影响难以精准评估，可能被别有用心的人利用，从而进一步加剧阶层分化与种族特定群体歧视等现象。

例如，在医疗领域，尽管驱动 AI 赋能医疗产业（尤其是公共卫生体系）的初衷是解决医疗资源供求端矛盾，使得现有医疗资源均衡化，国民健康管理结构化、常态化落实，但鉴于我国地区发展差异悬殊，如果现有医疗人工智能产业不够成熟，就激进推行医疗人工智能对现有公共卫生体系的渗透，AI 医学产品的消费零售价格仍居高不下，产品种类乏善可陈，产品性能与质量不尽如人意，偏远地区或中低收入群体消费者购买 AI 医疗设备的意愿就会大打折扣，不仅失去享有公共卫生资源的权利，更会因缺乏 AI 医疗相关配套设备，未形成数据闭环而在后续看病就医过程中遭遇更加负面的诊疗体验；或者为享受同等条件的医疗资源而付出更多的学习与资金成本，加重普通家庭的医疗负担。

3.3.4 人工智能干涉人类对自我意识与行为的判断决策，动摇人的独立自主性

以一些特殊医疗场景下的人工智能应用为例，机器的自主意识可能会威胁到医生或患者作为具有独立自主意识的人类个体对自我意识与行为准则的判断和掌控。例如，在健康管理领域，基于自然语言处理和计算机视觉的人工智能聊天机器人多用于精神健康管理，如心理辅导或危机心理干预，一旦该聊天机器人出现判断失误并向交谈对象传递错误的信息，极有可能干扰该对象原本正常的自我意识并形成错误的认知；而在医疗保健领域，一些用于老幼监护的护理机器人可能通过限制其人身自由、减少积极活动种类的方式“确保”监护对象“安全”。

3.3.5 数字化劳动相关权利问题

人工智能一方面对人类的主体地位造成影响，另一方面也在动摇人类数字化劳动

相关权利。其主要体现在：对于劳动主体来说，人工智能打破了现有的劳动关系架构，威胁人类在生产活动中的主体地位；对于接受人工智能劳动生产服务的人来说，双方关系从某种程度上被异化。

1. 劳动主体性地位问题：逐渐深入生产、生活的人工智能，其优势令人类难以望其项背，人类在生产过程中的主体性地位受到挑战

医生做出的每一个正确决策都基于扎实的专业知识和长期的经验积累，而人工智能强大的算力与信息储备能力，使其仅花费几秒钟即可完成一名主治医师需要花费十几年研习才能获得的全部医学知识和临床案例的学习。人工智能拥有比人类更高的精确度与工作效率，不会被情感与精力所限而拥有更小的出错概率，无论是对已有数据知识的回溯检索，还是未来案例内容的学习更新，人工智能的速度都是人类无法企及的。尤其是针对流程烦琐或需要收集大量资料、分析大规模数据的工作，人工智能往往能够以最快的速度提供最优质的反馈意见和针对性的解决方案，这种高效、精准、便捷的模式一方面极大提升了人类的工作效率，在节省劳动成本的同时优化了工作成果，另一方面也威胁着人类对工作项目本身的决策与贡献权重，使人类职员遭受被取代的威胁和压力。

除此之外，人工智能挑战人类在生产活动中的主体性地位，同时可能导致未来人类对人工智能技术的过分依赖而自我降低其对专业技能知识与工作经验的水准要求，甚至从潜意识层面开始动摇与质疑所学知识技能培养的存在合理性。

2. 异化劳动者关系：逐渐深入生产、生活的人工智能，导致人类在生产过程中的参与率下降，与人建立有效沟通与稳定关系的难度加大

以医疗人工智能为例，医疗人工智能发展的一大目标是通过计算机模拟人类医师大脑思考得出的智能行为，尽可能减少人为干预在现代医学诊疗过程中的占比，辅助医师提高诊断正确率和治愈率，同时缓解医生的工作负担，以此促进医疗质量提高。医疗人工智能的深入使得人类医护人员在临床诊疗尤其是医学检验场景中的直接参与率显著下降，而患者与机器接触的概率显著提高，智能医学影像、自然语言处理等人工智能技术将医务工作者从繁杂重复性的基础信息收集工作中解放出来，使其能够更专注于疑难病灶的诊治。传统医师与患者之间建立的直接沟通联系逐渐变成医师与患者借助人工智能医疗设备建立的间接沟通联系。但医学诊疗不仅是检验数据、科学

与经验的碰撞，更是在医患双方的直接交流中，患者从医生那里获取抵御疾病未知焦虑恐慌的心理和社会支持，而医生则根据患者心理状况给予适当的干预，从而缓解患者的负面情绪，提升诊疗效果。这种基于医患有效沟通而达成的隐性精神需求与心理支持，有时往往比药物或手术更有助于患者痊愈，更是对医患双方的人格尊重与人文关怀。

由此可见，一些特殊的职业其本质上体现的是人格尊重与人文关怀，以及绝大多数人依赖通过劳动行为建立对社会和人际关系的认知，人工智能的渗透可能从根本上改变这一现状，而其带来的影响仍旧值得深思。

3.3.6 进一步加深数字鸿沟

人工智能的应用与推广需要强有力的经济基础作为支撑，对于不发达地区和国家而言，智能技术所带来的对数字鸿沟的增大将严重影响其所能带来的福祉，同时，较差的信息基础设施也会使人工智能的发展捉襟见肘；对于发达地区和国家而言，大面积推广人工智能技术虽然具有足够的技术设施基础与经济基础，但城市中诸如老年人、残疾人等弱势群体由于其在教育、认知、生理等方面的特殊性，因此无法完全享受人工智能所带来的福利，这一点在城市数字化治理与基础设施建设中尤为凸显。以人脸识别为例，随着人脸识别的兴起，刷脸支付、刷脸乘车、刷脸门禁等逐渐普及，甚至成为公共基础设施的一部分。例如，以本次新冠肺炎疫情中推行的健康码为例，基于带有活体检验的人脸识别技术的健康码与国务院风险查询系统本身，在如此大规模的流行性公共卫生安全危机之下的确是一个值得鼓励的创新性公共管理策略，但如果仔细分析就不难发现，该策略实际上是基于一个不平等的前提，也就是默认公民都携带手机出行且都为技术红利的享受者与拥护者，可实际上基于 CNNIC（中国互联网络信息中心）的数据，现阶段我国仍有超过 5 亿人未接触互联网，其中很大一部分为老年人、残疾人等弱势群体。他们在面对“健康码”这样人脸识别与社会治理高度结合的产物时可能寸步难行，以至于因此被剥夺享有正常社会公共服务与保障的权利。技术所带来的数字鸿沟使他们在社会生活中边缘化，而数字鸿沟进一步导致权利分化；除此之外，随着技术的扩张，在技术风险不断膨胀的同时，技术福利却逐渐转变成基本规则；并非所有人都是技术的拥护者，每个人都有权利决定自己的人脸等隐私数据是否被用于其他用途，这就让那些希望降低技术风险或捍卫自身隐私数据而拒绝使用人脸识别的公民丧失了享有社会基本公共服务与基础设施的权利。