

对话式人工智能的法律风险及其系统性治理

陈 阳

华东政法大学法律学院, 上海

收稿日期: 2023年12月4日; 录用日期: 2024年1月7日; 发布日期: 2024年1月16日

摘 要

数字时代, 随着算法的飞速发展, 以ChatGPT为代表的对话式人工智能达到了新的技术高度。基于对话式人工智能的基础性与通用性, 其被广泛应用于各行各业, 与此同时, 它也带来了关于数据保护与隐私泄露、知识产权侵权与垄断、法律责任的承担等法律风险。对这些法律风险的认知与分析要求我们深入到自然语言的生成机制之中寻求解决之道。通过此种分析将会发现, 由于对话式人工智能本身具有的高技术性与应用的多样性, 仅凭现有的法律难以对其所带来的法律风险形成全面性治理, 因此, 一种合理的方案是以技术方案作为优先解决手段, 并以技术伦理原则作为治理的基本理念, 同时注重法律规则作为底线的保障作用, 形成多维度治理的合力, 对其所产生的风险进行系统性治理。

关键词

对话式人工智能, ChatGPT, 系统性治理, 技术伦理原则

The Legal Risks of Conversational Artificial Intelligence and Its Systematic Governance

Yang Chen

Law School of East China University of Political Science and Law, Shanghai

Received: Dec. 4th, 2023; accepted: Jan. 7th, 2024; published: Jan. 16th, 2024

Abstract

In the digital age, with the rapid development of algorithms, conversational artificial intelligence represented by ChatGPT has reached a new technical height. Based on the foundation and versatility of conversational artificial intelligence, it is widely used in all walks of life, at the same time, it also brings about data protection and privacy leakage, intellectual property rights infringement and monopoly, legal liability and other legal risks. The cognition and analysis of these legal risks

require us to go deep into the generation mechanism of natural language to seek solutions. Through this analysis, it will be found that due to the high technicality of conversational artificial intelligence itself and the diversity of applications, it is difficult to form a comprehensive management of the legal risks brought by conversational artificial intelligence only by existing laws. Therefore, a reasonable solution is to take technical solutions as the priority solution, and take technical ethical principles as the basic idea of governance. At the same time, we should pay attention to the protection role of legal rules as the bottom line, form the joint force of multidimensional governance, and systematically manage the risks generated by it.

Keywords

Conversational Artificial Intelligence, ChatGPT, Systemic Governance, Ethical Principles of Technology

Copyright © 2024 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 问题的提出

随着大数据、算法、算力的发展，由深度合成技术支持的 AIGC (AI Generated Content)在覆盖范围、智能程度等方面得到极大的提升。深度合成技术指以深度学习、虚拟现实的方式去生成成型的文本、图像、音频或视频等信息的技术。对话式人工智能(Conversation AI)就是通过深度合成技术生成自然语言文本，允许用户通过语言或者文字的形式与它进行交流，交流的形式并不局限于日常问答，用户还可以向对话式人工智能下达指令，以完成如法律文书撰写等更为复杂的任务。对话式人工智能的核心技术包括对自然语言处理的技术、自然语言的生成技术、机器学习与深度学习技术等等。通过这些技术它们具备较高程度的创造能力、在诸多领域中实现了人类的专业水平，推动了相关领域内的发展，显著地提升了人类的生产力。可以说对话式人工智能具备高度理解和执行能力，它们可以理解和生成自然语言，因此，对话式人工智能是基础性的人工智能，它被应用到一系列的下场景之中，换句话说，它能够完成不同领域的差异化任务。如国外 Open AI 公司的 ChatGPT，国内百度的文心一言等，它们具有明确的目标，即能利用算法在海量数据中进行学习从而从无到有地合成逼真的内容，此种“真”意味着使用者难以区分人类与机器，甚至在认知上模糊了人与机器智能的界限[1]。如在法律行业，通过适当的提问，ChatGPT 就能为律师结合特定案件生成相关的专业建议以及标准的文书，甚至能在司法领域促进类案裁判的可普遍化与具体裁判方式的变革，具备广阔的应用前景[2]，在哥伦比亚的北部城市卡塔赫纳巡回法院就首次将 ChatGPT 应用于司法判决。对话式人工智能的强大功能改变了我们的行为模式，进一步改变了我们的生活方式，但这些改变也意味着权利义务的分配在某种程度上的改变，同时也意味风险的产生与变化。

本文写作思路如下：第二部分我将从对话式人工智能的自然语言生成机制中分析它所带来的法律风险，在此基础上，第三部分我将分析现有的法律文件如何对其进行规制，并指出由于对话式人工智能本身的复杂性，仅通过这些法律规制，仍存在许多不足之处。因此，第四部分，我将在上述不足的基础上，指出应该选择一条有机整合技术手段、技术伦理原则与法律规则的系统性治理方案。

2. 对话式人工智能的法律风险

目前，对话式人工智能已广泛应用于客户服务、营销推广、智能虚拟助手、医疗保健等各个领域，

它们的跃迁式发展导致许多新的问题与挑战，尤其以 ChatGPT、文心一言等为代表的知识增强的大语言模型，它们在与人们的对话互动，回答问题，协助创作的过程中提供更加智能的服务的同时，带来更多隐蔽式的风险，如在对话内容隐藏了社会偏见、仇恨思想，生成错误的内容对人认知与行动产生误导、可能对知识产权产生侵犯等等。但对于这些风险，并不能孤立地看，而应该深入到对话式人工智能的运行逻辑中进行探讨。对话式人工智能对文本的生成与理解主要依赖于大语言模型，大语言模型可以通过对海量的文本数据进行训练，学习语言的结构、规则和语义，从而可以生成具有自然语言风格的文本或回答自然语言问题。大语言模型的建构主要包括三个阶段：预训练、微调以及强化学习。预训练阶段指用一些已有的大规模数据来训练一个模型，这些数据为模型提供知识，这些知识帮助模型更好地理解自然语言文本的语法、语义和语境信息。微调阶段是指模型在经过预训练阶段的基础上，再使用一个新的数据集进行模型的调整和优化，以便引导模型在特定的任务上表现出更好的性能。强化学习阶段则是使其在与环境的交互过程中通过学习策略以达成回报最大化或实现特定目标，换言之，通过建立符合人类偏好的决策序列(奖励模型)，其能根据环境的状态和可能的动作选择最优的动作，以最大化从环境中获得的累积奖励，从而实现其生成内容的优化。

2.1. 数据保护与隐私泄露

如前所述，对话式人工智能技术在预训练阶段以及微调阶段都需要海量的数据支持，这些数据包括用户的个人信息、对话数据等。这些数据具有强烈的个人身份性与私密性，往往涉及用户的隐私权，如果未经用户的同意使用或者不能对其妥善保管，将有很大的风险导致数据隐私泄露。具体而言，在数据收集与爬取阶段，提供给模型的原始数据是否具备合法性？这些海量的数据要想征询每一个用户的同意是十分困难的事情，在技术上也许能攻克该难题，但是在此种严格的限制下，对话式人工智能还能得到多大的发展？因为对话式人工智能的基础在于数据、算力与算法，如果没有足够多的数据，即使有强大的算力，最优化的算法，也无法支撑对话式人工智能的智能化运行。除了原始数据的提供，对话式人工智能系统在运行和使用阶段也需要收集用户的语音、文字等信息，用于后续的处理和分析。由于对话式人工智能卓越的服务能力，其被用户坚定地信赖，因此，相比于其他类型的人工智能，对话式人工智能有更多的机会获取用户的偏好与隐私。但未经授权与同意的数据收集行为违反了《中华人民共和国网络安全法》与《中华人民共和国个人信息保护法》中对于网络活动、数据保护、个人隐私信息的相关规定。

在数据存储和使用阶段，对话式人工智能系统需要将收集到的用户数据进行存储和处理，以便进行后续的数据分析和挖掘，如果这些数据被未经授权访问或泄露，将会对用户的隐私和安全造成巨大的影响，进一步而言，这些被存储的用户数据可能在其他的交互和沟通中生成新的信息，因此存在被用于商业推广、政治宣传等情况，而这些内容的生成也会对用户的权益造成侵犯。此外，对话式人工智能并不局限于人与人的交互，还存在人工智能与人工智能的交互，这种交互共享可能存在于一个机构或企业中，换言之，一个模型合法收集的数据，可能经历“辗转”成为其他模型的数据来源。因此，对话式人工智能造成隐私泄露的风险面将成倍增加。在数据收集与爬取阶段、存储和使用阶段、数据交互共享阶段，对话式人工智能都存在隐私泄露与数据保护的风险。

2.2. 知识产权侵权与垄断

如果作为商用式的对话式人工智能未对数据的来源以及生成的内容进行审查，那么将极有可能侵犯他人的知识产权。由于模型存储的海量数据，对话式人工智能的角色已经慢慢靠近知识的提供者，通过微调与强化训练，可信、可靠、权威的知识被模型视为优先解，在这一过程中，各个领域的专业知识被模型所挑选出来，但这也使得模型面临着严重的知识产权侵权问题。该问题可以划分为两个阶段，一是

预训练阶段与微调阶段的数据来源是否存在对著作权的侵害的问题，二是在其内容生成阶段，其生成的内容是否侵犯了他人的著作权，其本身是否具备著作权的问题。但从行为后果而言，第二个阶段的问题是根本的，只有当对话式人工智能生成了在知识产权上具有争议性的内容时，我们才能合理地主张其数据来源可能存在对知识产权的侵害。商用式的对话式人工智能追求知识的“正确性”，因此，在数据爬取过程中，可能未经授权地使用他人的作品从而侵犯他人的知识产权，或者在与用户进行交互，回答问题或者提供服务时，因其生成的内容的商用性质，以及它的本质只是在一定抽象程度对已有的知识，即语料、文本规则、语法进行重新组合[3]。因此，就难以避免其侵犯他人的知识产权或对个人数据进行非法编造和错误加工。

此外，考虑到对话式人工智能的发展条件，它也可能带来垄断，并间接导致许多法律上的不平等问题。在数据的来源与输入上，难以避免歧视与偏见，这些偏见将对人们的合法权益造成威胁。偏见的根源在于不平等，这种不平等并不由机器决定，对此，需要回到人的世界之中。仅从一般的技术层面就难以保障对话式人工智能技术的可及可用与均等化，首先在可及层面上，这意味着拥有，但受制于经济科技等因素，高新技术对于许多亚非拉国家的民众是遥不可及的，数字鸿沟确实存在。其次在可用层面上，受制于操作系统的复杂性、个人的智力水平、年龄、文化程度等因素，数字弱势群体对于新兴技术无阻碍运用可能性极低。最后，即使技术可及可用，但在一定意义上也难以实现均等化，如不平等访问的存在。就以生成式 AI 的发展为例，它需要海量的训练样本与强大的计算能力，这就导致其对服务器硬件设备提供的基础算力以及用于训练的数据提出了很高的要求，此种重负将导致技术的集中与垄断，此种情况技术、数据共享几乎不具可能性，因此 AI 的跃迁式发展极可能导致难以弥合的社会分裂境况，如若形成垄断，那么对话式人工智能所生成的内容将符合设计者和发布者的价值观与取向，由其所引发的行为也将侵害人们的合法权益，如地域歧视、性别歧视、职业歧视等，这些歧视将直接或间接地侵犯人们在就业、受教育等领域的合法权益。

此外，因为存在不同的数据库与不同的训练方式，因此其所带来的风险存在变异、变形的可能性。

2.3. 法律责任的承担

对话式人工智能在其运行过程不仅会自发地产生对于他人隐私、名誉、知识产权等权益的侵犯，而且它所生成的内容也容易被滥用，从而再次造成法律风险。就对话式人工智能运行过程中所暗含的侵权责任、违约责任与网络安全责任风险都涉及不同的主体，造就了复杂的法律关系。传统的网络侵权关系依然是一种双方或者三方主体的关系，即网络服务提供者与受害者，或者再加上网络用户。但是就对话式人工智能而言，从其生成机制来说，数据拥有者、数据提供者、算法设计者、对话式人工智能技术的提供者、用户，甚至政府都可能在对话式人工智能的各个阶段承担责任。此外，由于对话式人工智能本身所具备的复杂性、应用的多样性等性质，使得法律责任的划分、认定与承担都是困难的。

以生成的虚假信息举例，对话式人工智能与我们生活的不同侧面都已经开始实现深度交融，我们就不得不考虑这种虚假信息所造成的严重性后果，ChatGPT 在结果的生成与输出上，可以依照其算法确保其生成内容是最大化概率上的真，但即使这意味着 90% 的真，但依然存在 10% 的假，如果将其运用在医疗、法律建议这种特殊领域，这 10% 的假造成损失则是 100%。在司法上，国内外就已经开始探索对话式人工智能如何运用于司法，而这种虚假信息的可能性在司法上也将对公民生命、财产、自由等法益带来巨大风险。此外，由于此种存在虚假信息的可能性也将导致它被进一步的不道德的使用，如制造虚假新闻和信息、假冒身份信息。但这一系列因为虚假信息所产生风险却又存在责任承担方面的争议，一方面这与对话式人工智能是否具备主体地位有关，另一方面，是基于对话式人工智能算法运行的黑箱特质，从其内部而言，进行深度学习的人工智能究竟从数据中习得何种知识在技术层面上是难以被他者所观察

的，从其外部而言，算法决策的规则常常被开发者所隐蔽，因此缺乏透明性，也就很难指导决策的过程与逻辑[4]。正是因为对话式人工智能的模型是复杂的深度学习模型，其内容的生成过程缺乏可解释性，此种生成内容的具体逻辑甚至无法被研发者所理解。因此，这种不确定、多方面的复杂情况以及由对话式人工智能所衍生的诸多下游应用将导致各类主体之间的责任难以划分[5]。在这些侵权情况下，我们很难通过传统的过错责任来认定法律责任，因为算法的运行已经出现某种程度的不可控性，除非我们认为对话式人工智能具备法律主体地位，并使之承担法律责任。但对此存在激烈的纷争，并且涉及了主体性的判断标准的哲学讨论，至今并未形成定论。

3. 关于对话式人工智能的规范性文件及其不足

为了保障人工智能的健康快速发展，我国近年来相继出台了政策、白皮书、标准指引等文件。如2017年国务院印发的《新一代人工智能发展规划》，明确了新一代人工智能发展的战略目标，旨在推动人工智能技术持续创新，并使之与经济社会深度融合；新一代人工智能治理专业委员会在2019年的《发展负责任的人工智能：我国新一代人工智能治理原则发布》中指出，人工智能的研发和应用过程要坚持和谐友好、公平公正、包容共享等原则，这些原则旨在推动人工智能的可持续发展。全国信息安全标准化技术委员会在2021年发布的《网络安全标准实践指南——人工智能伦理安全风险防范指引》中强调了对于个人数据的保护，关注人工智能系统在决策过程中的公平性和透明度，重视其安全性等内容。2021年由国家新一代人工智能治理专业委员会发布的《新一代人工智能伦理规范》中强调了诸如增进人类福祉、促进公平公正、保护隐私安全、确保可控可信、强化责任担当、提升伦理素养的基本伦理要求，以此为从事人工智能相关活动的自然人、法人和其他相关机构等提供伦理指引。2021年颁布的《中华人民共和国个人信息保护法》也为保护个人信息权益，规范个人信息处理活动，即信息的收集、分析、使用做出了翔实的规定。此外，2021年颁布的《中华人民共和国数据安全保护法》不仅建立了数据分类分级保护制度，也明确了数据安全责任。国家互联网信息办公室、工业和信息化部、公安部和市场监管总局四部门联合在2021、2022年发布了《互联网信息服务算法推荐管理规定》与《互联网信息服务深度合成管理规定》，这两部规定同样是对对话式人工智能技术形成约束和限制的重要规范性文件。在2023年4月更是起草了《生成式人工智能服务管理办法(征求意见稿)》，7月发布了《生成式人工智能服务管理暂行办法》(简称《办法》)。在宏观层面上，人工智能监管和治理规范已经形成初步治理框架；在微观层面上，《办法》对生成式AI展开了全方位的综合治理。而在国际上，则以2021年联合国教科文组织磋商的《人工智能伦理问题建议书》为范例，这些规范都对对话式人工智能的发展具有重要意义。

对话式人工智能技术的发展与应用在个人信息、算法、数据、深度合成、伦理等多个维度都受到了相关的规范与约束。虽然现阶段有法可依，但依然没有统摄性规定，也未形成多元有序的规制体系，存在规制主体多且分散的情况。从上述的规范性文件的颁布主体与其核心任务可知，对对话式人工智能的治理依然是呈现一种纵向治理模式，即在特殊领域内进行专业化的特殊立法。适应不同的场景变化，依托模型产业结构进行专业化的特殊治理有其合理之处，但是这种模式下将造成规制主体过多，规制成本过高的情况。此外，在人工智能技术性能不断进行实质性跃升的背景下，法律所具有的滞后性使得现行的法律并不能十分有效且妥当地应对新生的模型或其产品。因此，只能通过不断颁布新的法律法规来应对可能存在的漏洞和新生的风险，但这也容易造成规制主体与对象的混乱。更为重要的是，法律本身具有的保守性和滞后性无法兼容技术的迅速迭代与更新，此种张力将造成巨大的立法压力。由于存在多个规范性文件，又不存在统筹性的规定，因此，对对话式人工智能治理的协调性和能动性不足，纵横交错的法律规范之间可能会存在一定程度的冲突与抵牾，难以形成风险治理的合力。这会导致立法预期与实践效果产生较大的差异，因此，需要增强相关规范的实践性。但毕竟法律规定与技术解决方案属于两

类性质不同事物，这之间的张力使得法律无法深入到日新月异技术领域做出翔实的规定，从这方面而言，相关规范的实践性受限似乎是必然结果。从另一方面而言，技术方案是一种客观的事实，而法律是一种规范性实践，它旨在指导人类的行为[6]，这两者无法混为一谈，我们既不能主张通过法律规定要去解决那些还未攻克的技术难题，因为法律无法为不可能之事设定义务，其次我们也不能主张法律在对话式人工智能治理中是不重要的，因为法律是现代生活中最重要的行动准则，维护着社会的底线。

此外，除了技术方案与法律外，伦理规范也在对话式人工智能的治理中发挥着重要作用，因为对于对话式人工智能的治理的困难来源于它的复杂性与广泛的应用性，换言之，是来自它已经近似“主体”能力，不可否认，它在某些方面已经对我们社会结构形成了剧烈的冲击并带来诸多的法律与道德风险，这些风险一部分是因为人的不当使用带来的(如利用生成式 AI 生成虚假信息诈骗)，对于这一部分，我们可以通过法律规定加以约束；一部分是因为技术的固有缺陷所导致的(如算力、算法、数据库)，这一部分通过技术的发展和突破就能解决；但还存在许多法律伦理风险来自对话式人工智能的复杂性以及人类价值观的多元性，这要求我们从消除特定的风险要逐步转向构建全方位的系统性治理体系。基于此，我们必须一般在一般层面上为法律在对对话式人工智能的治理中确定一个妥当的位置，使其能与对话式人工智能领域内的技术规则、技术伦理相匹配、相统合，形成对对话式人工智能的系统性治理。这就要求我们确定各种规范在此种系统性治理中所扮演的角色与其规制的限度。

4. 对话式人工智能的系统性治理

4.1. 技术解决手段与法律策略

要形成对对话式人工智能的系统性治理，首先要解决的便是技术解决手段与法律策略的关系。主流观点认为，面对科学技术不断更新迭代与法律滞后性之间张力[7]，对法律进行广义的理解，将科学技术和政府规制纳入法律的概念范围内来应对技术发展带来的问题[8]。这种观点可能会变为对特定技术或特定场景进行专业化的法律治理，但无论是哪一种版本，主流观点的核心是将技术解决手段纳入法律的范畴之中。但是这种做法有多少合理性？正如上文所提到的，法律是一种规范性实践，但是技术解决手段并非一种规范，而只是作为一种事实存在，在算法层面上我们能进行优化，但这仅意味着一种真实情况，即客观上，我们通过科学方法与研究得出了一个技术上的更有效手段。当这一方案取代了旧有的算法时，它也是作为客观的事实存在，而不是任何能作为任何能被违背的规范而存在。简言之，作为技术解决手段的事实直接控制了人工智能的行为，而非规制它们的行为。因为“……我所遇到的不再是命题中通常的‘是’与‘不是’等连系词，而是没有一个命题不是由一个‘应该’或一个‘不应该’联系起来的…因为这个应该或不应该既然表示一种新的关系或肯定，所以就必须加以论述和说明……” [9]，质言之，在“是”与“应当”、“事实”与“价值”的领域之间存在一条无法逾越的鸿沟。因此，技术解决手段与法律策略是两种不同性质的事物，不能以某一种吸收另一种。

如果技术解决手段无法作为规范存在，也无法被法律策略涵盖，那么它们两者存在何种联系？一种可靠的方案是拒绝对法律进行广义式的理解，而是寻求技术解决手段的使用。法律虽然无法纳入事实，但可以规定技术解决手段的使用。具体而言，在对对话式人工智能进行治理时，我们的治理是出于良善的目的，这种目的正当性使得我们应该去采取某些行动，这些行动涉及工具的使用，那么我们应当去使用有效率的工具以便实现我们的良善目的便是关乎人的行动的，因此，技术解决手段的使用能被法律所规定，也符合法律是调整行动的基本看法(因为该工具能不被使用)。但是，进一步的问题是，法律如何调整技术解决手段的使用行为？如果我们主张法律应该即时且翔实地规定所有更新的技术手段的使用，那么这不仅违背法律的保守性和稳定性特征，而且在实践中也不具有操作可能性，毕竟立法者的工作还是

关于社会的，而人工智能专家的工作是关于自然的，两者无法形成相同程度的认知。此外，我们还需要留给对话式人工智能足够的发展空间。

基于此，通过技术解决可能存在的法律风险需要作为优先的方案，为保障技术方案的使用，法律可以对此进行规定。此种规定就意味着法律在处理新技术所带来的风险上的角色转变，即从风险治理的直接技术干预者转变为技术使用的间接监管者。因此，法律不应该贸然地进入技术领域内，具体的技术解决方案交由专业人士。就法律层面而言，应该进一步明确服务者与用户的责任，并且注重软法的指导性、号召性、宣示性作用。此外，还需要突破纵向规制模式，要建立起多元有序的规制体系，注重不同规制主体的协调性。但正如前面所指出的，法律与技术是两种性质截然不同的事物，在他们之间构建联系取决于上文所说的良善的目的。那么这一良善目的是什么？它在系统性治理中又发挥着什么作用？

4.2. 技术伦理原则作为系统性治理的基本理念

对良善目的的理解要回到技术向善的基本原则中予以理解，具体而言，任何技术或者科技都需要服务人，服务于人类的共同善，即增进人类的共同福祉。迄今为止，人类的大多数活动也都为了增进人类福祉，这是我们不可怀疑的出发点。这种人类中心主义观念对人工智能的设计与发展也提出了价值对齐的要求，即要求高度自主的人工智能系统的设计应确保其目标和行为在整个运营过程中与人类价值观保持一致。在《办法》中第4条也规定了“提供生成式人工智能服务应当尊重社会公德和伦理道德”。基于服务于人类共同善的基础原则，可以衍生出其他的一些技术伦理原则，这些原则有一部分是实质性的，另一些则是形式性的，前者如公平原则、正义原则、非恶意原则等；后者如透明性原则，安全性原则，人类可控原则等等。因此，基于共同善所形成的原则不仅考虑到了技术层面，又能覆盖人类普世的道德价值，实现了两者相统一。这些原则本身不仅约束了对话式人工智能服务者、设计者的行动，并且它们也共同构成了系统性治理的基本理念，指引着法律的制定与技术方案的实现。而法律反过来可以将这些原则以法律规则的形式规定，从而保障这些原则的实现。

至此，以法律作为底线，以技术伦理原则作为基本理念，以技术解决手段作为优先方案的系统性治理框架已形成，在其中，法律与技术方案截然区分又统合在一起，法律通过规定技术方案的使用与其产生联系，技术方案的使用与法律的调整都以是否符合技术伦理原则为界限，技术伦理原则并非单一的原则，而是具有不同内容的实质性原则与形式性原则组合而成，而这些原则的唯一基础便是服务于人类的福祉这一根本目的。而法律在其中作为底线，保障着技术伦理原则与符合这些原则的技术方案的使用，因此，我们得到了一个系统性的层级治理框架，该框架的最顶层是人类的福祉以及由此派生的诸多原则，在中层是由这些原则为指导并具有重要意义的法律规则，在底层是在实践中大量存在的各类技术解决方案。

5. 结语

对话式人工智能的出现带来了一场广泛与深刻，波澜壮阔的科技革命，此种影响不亚于互联网的发明，但在推动人类文明大步前进时，对话式人工智能不可避免的带来了诸多法律与伦理风险，这些风险已经无法单独地依靠技术伦理、技术方案、法律规范来予以解决，而应该在一般层面上寻找稳固的出发点，统合这三者，形成治理的合力，而不应以一种方案取代或吸收其他诸种方案。通过本文的分析与论证，至少能确定各种方案在系统性治理中的位置与限度。通过区分法律策略与技术解决手段，并以人类福祉为基点，本文逐步构建了系统性治理的层级框架，即一种金字塔式的三级规范体系。该体系以技术伦理原则作为基本理念，以相应的法律规范作为中层规范体系，以具体的技术方案作为底层的治理体系，层层相扣，进而推动人类与社会的良性与可持续发展。

参考文献

- [1] 张凌寒. 深度合成治理的逻辑更新与体系迭代——ChatGPT 等生成式人工智能治理的中国路径[J]. 法律科学(西北政法大学学报), 2023, 41(3): 38-51.
- [2] 郑曦. 生成式人工智能在司法中的运用: 前景、风险与规制[J]. 中国应用法学, 2023(4): 81-87.
- [3] 於兴中, 郑戈, 丁晓东. 生成式人工智能与法律的六大议题: 以 ChatGPT 为例[J]. 中国法律评论, 2023(2): 1-12.
- [4] 毕文轩. 生成式人工智能的风险规制困境及其化解: 以 ChatGPT 的规制为视角[J]. 比较法研究, 2023(3): 155-172.
- [5] 徐继敏. 生成式人工智能治理原则与法律策略[J]. 理论与改革, 2023(5): 72-83.
- [6] [美]安德瑞·马默, 亚历山大·萨奇. 法律的性质[J]. 张民全, 译. 北大法律评论, 2019, 20(1): 293-321.
- [7] 郭春镇. 生成式 AI 的融贯性法律治理——以生成式预训练模型(GPT)为例[J]. 现代法学, 2023, 45(3): 88-107.
- [8] [美]罗杰·布朗斯德沃德. 法律 3.0: 规则、规制与技术[M]. 毛海栋, 译. 北京: 北京大学出版社, 2023: 2-3.
- [9] [英]休谟. 人性论[M]. 关文运, 译. 北京: 商务印书馆, 2016: 505.