

论人工智能参与下董事信义义务体系的重构

殷静萱

西南民族大学法学院, 四川 成都

收稿日期: 2026年6月5日; 录用日期: 2026年6月28日; 发布日期: 2026年7月7日

摘要

本研究集中于人工智能技术嵌入企业决策场景时, 董事信义义务体系的演变轨迹及其改良途径, 伴随人工智能技术持续更新并逐渐成为现代企业运作的核心推动因素, 它在优化决策速度和改善决策品质方面的作用越发明显, 在这种情形之下, 传统董事信义义务架构正在经历重大改变, 无论是忠实义务还是勤勉责任这两大核心方面均存在显著差别, 特别就忠实义务而言, 由于算法模型变得越发繁杂而且非常倚重大数据分析, 董事个人喜好或者利益联系很难被准确辨别出来, 现存的监督手段或许会遭遇效能下降甚至完全失效的危险。人工智能技术被广泛使用之后, 董事会在执行勤勉义务的时候遭遇了更为繁杂的决策环境, 对履职能力的要求也变得更高, 可以从改进信息披露制度、加强算法透明度、改良勤勉义务评判准则、接纳商业判断原则这些方面着手, 创建起适应时代发展的信义义务体系, 这样的改革举动有益于推动公司法理学理论更新, 也能给我国企业治理模式朝着市场化、法治化、国际化、智能化方向转型给予有力支持。

关键词

人工智能, 公司治理, 董事信义义务, 忠实义务, 勤勉义务

Reconstruction of the Fiduciary Duty System of Directors with the Participation of Artificial Intelligence

Jingxuan Yin

School of Law, Southwest Minzu University, Chengdu Sichuan

Received: June 5, 2026; accepted: June 28, 2026; published: July 7, 2026

Abstract

This study focuses on the evolutionary path and improvement approaches of directors' fiduciary

文章引用: 殷静萱. 论人工智能参与下董事信义义务体系的重构[J]. 争议解决, 2026, 12(7): 55-61.

DOI: 10.12677/ds.2026.127213

duty system amid the embedding of artificial intelligence (AI) technologies into corporate decision-making scenarios. With the continuous upgrading of AI technologies and their growing role as a core driving force for the operation of modern enterprises, such technologies have delivered increasingly prominent effects in accelerating decision-making and elevating decision quality. Against this backdrop, the traditional framework of directors' fiduciary duties is undergoing profound transformations, with marked changes emerging in its two core components: the duty of loyalty and the duty of care. In particular for the duty of loyalty, given the rising complexity of algorithmic models and their heavy reliance on big data analytics, it has become difficult to accurately identify directors' personal preferences or conflicting interests, which puts existing supervision mechanisms at risk of diminished efficiency or even total failure. The widespread adoption of AI has exposed boards of directors to far more complex decision-making environments when fulfilling their duty of care and raised the competency threshold for their performance of duties. A fiduciary duty system compatible with contemporary development can be established through measures including optimizing the information disclosure regime, enhancing algorithm transparency, revising the evaluation criteria for the duty of care, and adopting the business judgment rule. Such institutional reforms help advance the theoretical innovation of corporate law and provide robust support for China's corporate governance to shift toward marketization, legalization, internationalization and intelligentization.

Keywords

Artificial Intelligence, Corporate Governance, Directors' Fiduciary Duties, Duty of Loyalty, Duty of Care

Copyright © 2026 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. 引言

人工智能技术飞速发展促使它慢慢变成重要的决策支撑工具, 2014 年风险投资公司 Deep Knowledge Ventures 率先把 VITAL 机器学习算法融入董事会决策过程, 这被看作是 AI 技术踏入企业战略规划领域的关键标志, 而且引起很大关注, 之后很多企业效仿, 利用智能化手段改良决策流程、改善经营效能、加强竞争优势。南方电网凭借知识图谱创建智能调度系统, 操作票审核准确率达到 97.6%, 东风汽车借助“擎天 AI 智算调度管理平台”达成异构计算资源的高效整合, 给自动驾驶技术研发给予有力支撑。目前, 部分企业已经把人工智能技术深入到高级管理层的战略决策支持系统当中, 在业务操作方面, 人工智能因为其高效的数据处理能力而突出, 它能够很快地对大量信息进行整合分析, 并且在不断变化的市场环境中给管理层给予即时回馈, 从而改善决策速度, 这项技术可以做到部分重复性或者规则驱动的任务自动化处理, 比如数据录入、文档归类以及报告制作等等, 这样就能改进人力资源调配的效率, 促使整个组织运作水平得到提升。

人工智能技术的深度结合给传统公司治理法律架构带来深刻改变, 董事信义义务的内涵出现明显变动, 在传统的公司治理结构里, 董事属于关键决策主体, 既要承担忠实义务也要承担注意义务, 忠实义务着重于守护公司整体利益并规避可能的利益冲突, 注意义务重视董事在决策时的谨慎态度和勤勉精神, 目的在于做到企业价值最大化以及股东权益保障。伴随人工智能慢慢渗入决策流程当中, 董事的角色界定及其责任范围产生巨大变化, 这使当前信义义务理论体系及其实际操作遭遇不少亟待处理的新状况。在完善生成式人工智能治理架构并健全中国特色现代企业制度, 大力弘扬企业家精神的大环境下, 深入

探究人工智能技术渗入决策流程之时董事信义义务的实际范围问题，渐渐变成了法学研究领域的重点议题，由于技术革新持续推动，《公司法》正在从重视经营效率慢慢向重视合规监管转变，这就对监管部门制订有关“算法失灵”现象的专业化规范性文件提出了更高的要求，而且还要进一步完善董事及高管的责任体系，从而符合新时代发展的必要条件。

本研究试图塑造契合人工智能时代特征的企业治理法律架构，促使技术同治理模式达成深度融合，进而改进治理效能、改善组织架构，通过深入剖析董事信义义务的关键含义，给公司法范畴的制度革新给予能提供帮助的理论支撑和实践路径。

2. 人工智能决策参与的法律地位及角色演变

2.1. 法律地位界定

人工智能主体地位的探讨一直是法学界的重要话题，“主体说”与“客体说”成为主要争议点。主体说认为，当人工智能表现出自主意识，具备行为能力，且结合特定智能属性时，可以视作本体论上的主体，从而取得独立法人资格。佛罗里达大学莱文法学院教授瑞依(Robert J. Rhee)就主张，当 AI 获得自我意识、能动性和独特智能时，就将获得本体论人格(ontological personhood)，法律可以赋予 AI 法人地位，就像过去赋予传统商业公司人格地位一样，而不再将 AI 作为财产纳入公司法边界之内[1]。从客体论角度来说，虽然人工智能表现出明显的智能化特征，但因为缺少人类特有的心智意识和精神内涵，所以至今没有得到独立的法律主体资格，仍旧被当作辅助人类活动的工具性存在，在责任归属方面，由于不具备独有的专属性财产权益，也不能具备成为完整法律主体的必要条件。

学界一般认为，人工智能虽然在某些方面表现出了自主决策和行为执行的能力，但是否真正具备自我意识尚存争议，虽然有学者预测强人工智能时代即将到来[2]，由于目前的技术尚未攻克人工智能的关键难题，将其纳入法律主体的范畴存在不少现实阻碍。因此，本研究主要针对弱人工智能，探究其在决策过程中所具有的功能特点以及给社会带来的潜在影响。

2.2. 角色转型：从董事协助到顾问履职

从法律客体理论视角来看，现阶段的人工智能在董事会决策中并没有实现主体地位的独立化，而是作为辅助工具存在的，随着技术革新与应用场景的不断拓宽，其在董事决策过程中存在两种典型的运用模式：

协助董事模式：董事会辅助方面，人工智能主要以工具化的方式出现，其核心功能包括数据采集，整合以及初步分析等基础环节，目的在于为决策提供信息支持，并提升决策速度，在此模式之下，人工智能所体现出来的自主决策能力或者决策功能几乎可以忽略不计[3]。

董事顾问模式：董事顾问机制推动的智能决策系统能够把搜集到的数据变成具体的执行方案，做到无需董事会随时参与就能独立完成决策，该系统在没有外界干扰的时候可以自行运作并做出判断。有学者指出，该模式下的人工智能可以“通过提出和回答问题以及构建和模拟情景的方式，为更复杂的问题解决和情况决策提供支持”[4]。不得不承认，该模式下的人工智能，如智能决策，虽然有一定的自主性，但其自主性仍依附于董事，并未取代自然人[3]。大型企业在数据资源足够充裕的情形下，可以购置通用型的人工智能设备，并且按照行业的特点数据加以调整优化，再融入到企业的管理架构之中，小型企业由于缺乏足够的数据积累，一般会选取像 DeepSeek 或者 ChatGPT 这样的商业软件工具，来满足日常经营的需求。

3. 人工智能对传统信义义务的挑战

现代公司制以所有权与控制权相分离的两权分立结构为特征，这种结构虽提升了管理效率，却也导

致了股东与高级管理人员之间的利益冲突。为了控制代理成本、最大化公司利益、并鼓励董事积极履行职责，传统信义义务体系通常以商业判断标准为基础，要求董事进行经营决策时遵守忠实义务和勤勉义务。我国新《公司法》¹第 180 条对忠实义务和勤勉义务做出了具体规定，其中，忠实义务要求董事“采取措施避免自身利益与公司利益冲突，不得利用职权牟取不正当利益”，勤勉义务则要求董事“以管理者通常应有的合理注意”履行其职责。传统信义义务体系是以自然人董事为规制对象的，然而，人工智能的介入在一定程度上颠覆了公司的决策过程，给忠实义务和勤勉义务的判断带来了一系列挑战。

3.1. 忠实义务：自利行为的隐蔽性转换

在现代公司治理体系中，传统忠实义务的核心要义在于防止自然人董事利用关联交易或自我交易谋取私利，董事顾问制度的兴起以及算法决策技术的广泛应用，使得董事顾问制度正面临着前所未有的挑战。从理论上讲，人工智能依靠代码和算法处理信息并做出判断，其运作方式具有高度独立性且“中立”，这种技术特性为董事规避忠实义务提供了可能，若董事完全依赖人工智能生成的决策方案，理论上其行为可被解释为遵循客观数据的结果，从而难以被认定为利益输送或损害公司利益^[5]。虽然人工智能缺乏传统自然人董事的“自我利益倾向”和主观动机，但是目前主流仍为弱人工智能，且受到“算法黑箱”问题、“数据偏差”等问题的影响，由算法生成的决策很难完全符合忠实义务的基本要求。

当前最终决策权仍掌握在自然人手中，人工智能的客观中立并不意味着人类不会想办法逃避忠实义务的约束^[3]。人类主体依然可以通过多种方式规避忠实义务的约束，当董事与技术研发人员存在利益关联时，董事可能会在源代码中植入偏见性逻辑，或者在数据采集阶段筛选并隐藏不利信息，从而造成人工智能生成的决策结果产生偏差，即使监管机构发现决策偏离公司利益，“算法黑箱”的存在使监管机构无法完全了解内部运行机制，忠实义务的履行情况也难以被有效评估和验证，这种技术环境大大降低了忠实义务的实际约束力，给相关方逃避法律责任创造了有利条件。

3.2. 勤勉义务：“知情”判断环境的复杂性分析

勤勉义务要求董事在公司治理过程中遵循“普通管理者的合理注意标准”来履行职责，其执行效果依赖于董事对相关事务的认知水平，“合理注意”概念存在很强的灵活性，所以它具体的表现形式会因为董事的专业背景、社会经验以及特定情境等因素而存在明显的差别，为了推动企业可持续发展，董事应当积极获取并细致剖析对企业运营至关重要的核心信息(比如财务状况、市场定位和行业竞争态势)，然后凭借这些数据制订出科学严谨的战略规划。

人工智能技术飞速发展之际，董事会把它融入决策体系，实际上重新构建了传统治理模式的主要构成要素。人工智能的介入让董事知情的判断环境变得更为复杂，增大了董事履行注意义务的难度^[3]，关于董事会是否应当强制应用人工智能技术，关键在于董事是否尽到了勤勉义务。一般来说，董事在决策时是否引入人工智能应当由董事自行判断，虽然有学者主张，人工智能在某些特定领域拥有一般人无法比拟的优势，因此董事有义务聘请人工智能^[6]，但是从当前立法来看，并没有要求董事必须使用人工智能的规定。仅在一些特殊的交易领域，基于交易惯例以及行业的特殊性，数字工具的使用可能成为强制性要求^[4]，比方说智能投顾。

董事会凭借人工智能辅助决策的时候，可以部分把职能交给智能化系统，但由于人的惰性和对人工智能的盲目信任，董事决策的任务最终可能会在很大程度上被人工智能取代。尽管名义上自然人拥有最终决策权，但若董事仅仅是执行人工智能的意见，而没有加以审慎监督，那么这将是一种瑕疵的知情状态，

¹http://www.npc.gov.cn/c2/c30834/202312/t20231229_433999.html

难以认定其符合注意义务的要求，从而增大了董事规避注意义务的可能性[3]。再加上人工智能技术的复杂性和不可解释性，如何判断人工智能董事是否尽到了合理的注意义务同样存在困难，带来了决策逻辑的不透明问题。也就是说，人工智能的决策或许是正确而明智的，但其很难向自然人董事和利益相关者清晰解释其决策过程或决策逻辑，这使得利益相关者难以理解其如何平衡不同利益并作出决策[5]。

4. 重构路径

人工智能技术的深度应用推动信义义务体系产生重大变革，当下，我国及其相关国家把信义义务归入“义务束”范畴，但其内在逻辑结构存在明显瑕疵，规范界限也不够明晰，信义义务作为以忠实、审慎、合规为关键要素的制度体系，表现出很强的灵活性和适应性，人工智能迅猛发展时这种特性更为突出，要顺应新的治理模式需求，传统信义义务框架必须革新改良，才能做到制度的持续发展和效能的提升。

新《公司法》的颁布也为这一重塑提供了契机。新《公司法》引入了英美公司法中的单层制治理结构，允许公司选择是否设立监事会。第69条规定，有限责任公司可以在董事会中设置由董事组成的审计委员会，行使监事会的职权，不设监事会或监事。同样，第121条也允许股份有限公司以设立审计委员会的方式替代监事会。这种新的监督模式旨在适应董事日益增加的监管需求，增强了董事会信义义务中的监督职能。

具体而言，人工智能参与决策和董事监管职能增强的背景下，信义义务体系的重构体现在以下方面：

4.1. 忠实义务的重构：强化信息披露与算法透明度

在人工智能推动的企业治理架构里，“算法黑箱”效应以及“数据偏差”现象明显加强了管理层自我利益行为的隐秘程度，这就给董事会规避忠实义务赋予了便利条件，要解决这个难题，就要改善信息披露制度，提升透明度监管的效能，从而有效地找到并制止利益输送的风险，虽然技术更新改变了自我利益行为的呈现形式，但其背后的逻辑没有发生本质上的变化，在这种情况下，管理层可能会用如下手段来执行自我利益操纵：形成研发团队同高层管理者之间的利益联结网络、篡改源代码以控制程序输出结果的方向。

董事借助人工智能系统辅助决策时，如果故意剔除掉那些对其不利的数据信息，就可能使输出结果背离预期目标，为了提升决策过程的透明度，防止董事凭借技术手段逃避忠实义务，董事应当把算法模型，数据来源以及处理流程等关键要素全部披露出来，还要清楚表明存在利益冲突的情形，立法机关要大力推动人工智能系统可解释性方面的研究工作，保证其决策逻辑具备追溯与解析的能力，这样既有益于加强公众对人工智能决策的信任程度，又给忠实义务的履行给予了有力的技术保障。创建“算法影响评价”机制，企业须在部署人工智能系统之前，开展全面的风险评价和社会效应分析，把评价报告呈交董事会审议。同时，可以借鉴《欧盟人工智能法案》的做法，对人工智能系统进行分级管理，根据其风险等级制定不同的披露和监管要求。对于高风险的人工智能应用，如涉及财务决策或重大战略规划的系统，应要求更详细的信息披露和更严格的算法审计[7]。

4.2. 注意义务的重构：更新判断标准并引入商业判断规则的应用

勤勉义务的界定标准要依照人工智能技术的发展趋向来动态调节，这样才能符合复杂多变的决策环境。参考商业判断规则，商业判断规则的核心在于，只要董事在决策过程中遵循了合理的程序，且不存在恶意或重大过失，即使决策结果不尽如人意，董事也不应承担个人责任[8]。所以，在评判勤勉义务的时候，重点要看董事对于人工智能系统的监督和管控能力，而不是单纯依靠它提供的信息，董事要保证所用的工具数据真实可信，还要凭借自己的专业知识，经验和人工智能的建议做综合分析，从而达成科学严谨的决策目的。

要推动人工智能技术在企业治理层面深入应用，还要提升董事会职能协同效率，就迫切需要构建起“人工智能决策支持平台”的认证体系，这个体系不仅要考虑系统的可靠性和安全性，还要让董事在使用人工智能时，定期接受相关培训，以提高其对人工智能技术的理解和应用能力[9]。“人工智能董事”由于具备“理性智能体”的特性，在信息整合以及决策分析方面有着明显的优势，其履职成效远远好于传统的自然人董事，在理想状态下且没有外界干扰时，“人工智能董事”可以很好地完成数据收集与处理工作，极大地减少了因粗心大意或者误判而产生的注意义务违约风险。在这种情况下，人工智能董事的注意义务范围将收窄，且主要集中于“谨慎决策”层面[5]。认定董事是否违反勤勉义务，决定性判断要素并非董事是否有义务使用人工智能，而是在董事决策时一个谨慎负责的管理者在作出决定时应当使用哪些信息[4]。

4.3. 忠实与注意义务协同：构建综合性监督机制

传统信义义务理论主要围绕忠实与勤勉这两项要素展开论述，但是随着企业治理结构向着多元化、分散化方向发展，董事监督职能的重要性日益突出，并呈现出动态扩展态势，当董事会把经营决策权交给管理层之后，董事能否有效地执行监督任务，就成为评判其保护股东权益成果好坏的一项关键指标。有些学者表示，人工智能技术在企业管理中被运用起来可以看作是新出现的一种权力分配形式，借助提高决策速度并促进资源合理调配水平来帮助企业达成价值最大化的目标，这样的做法与现代董事会向高层管理者移交部分治理权限的趋势非常吻合。现代企业治理体系当中，董事监督职能加强变得愈发重要，这对于董事避免推诿责任有着深刻意义，借助人工智能技术达成责任分散和履职规范优化十分关键，形成包含忠实义务和勤勉义务的协同监管体系，要依靠以人工智能为主的信息化监督平台，通过算法审核制度，外部专家咨询系统，专题培训等手段来推动，这样一种综合方案既能促使监督网络顺畅运行，又能明显改善董事履职效率，而且还能给董事忠实履职给予有效的技术支持。

5. 结论

人工智能技术的深度融入正在促使企业的治理结构与运作模式产生深刻变革，这给传统的董事信义义务理论体系带来了新的挑战，依靠对人工智能在企业决策过程中的应用状况及其未来走向展开系统探究，本文深入剖析人工智能对企业董事信义义务制度所引发的影响机理，并给出相应的完善途径与改进意见。

人工智能在企业决策领域地位与功能演变显示，虽然人工智能在现代商业生态系统里占有重要地位，不过当下仍处在弱人工智能阶段，不具备自主决策能力，在这种情况下，人工智能更多是以辅助工具或者专业咨询的形式发挥作用，很难取代传统董事会的核心决策职能，这种状况促使学界和实务界从法律角度重新思考并界定人工智能的合法地位，进而推进人工智能在公司治理体系里的规范化应用和可持续发展。

人工智能技术飞速发展给传统信义义务理论体系带来重大改变，忠实义务与勤勉义务是其关键部分，在忠实义务方面，算法逻辑嵌入和数据驱动特性明显加大了企业内部关联交易的隐秘程度，现存监管手段遭遇极大考验，针对此状况，此项研究提议改良信息披露机制，提升算法透明度，创建“算法影响评判”制度，接纳独立第三方审核组织，保证人工智能决策流程的公正性与可追踪性，在勤勉义务层面，信息技术被普遍应用之后，董事获取信息的途径变得更多元化，这便对董事的履职水平提出了更为严格的要求，就这个问题而言，本文提倡重新塑造勤勉义务的评价准则，采用商业判断原则，而且创建起“人工智能决策支撑系统”认证体系，做到技术创新与合规经营的协调统一。

本研究设计出一套多维协同监管机制的创新架构，成立专门的人工智能监管机构，改良算法审查流程，采用外部专家咨询体系，加大董事会成员的专业技能培训力度等详细举措，其核心目的就是借助忠

实义务和勤勉义务的协同作用，促使企业在治理过程中规范使用人工智能技术，给董事履职给予技术支持，进而改进董事履行信义责任的水平和效率。

人工智能技术的加入给传统的董事信义义务体系带来了很大的改变，给现代公司的治理结构改进赋予了新的想法，通过全面重新塑造信义义务规则，可以很好地应对智能化时期所带来的挑战，促使技术和管理更好地融合，从而改善公司的治理水平和经营品质，这样的改革举动有益于推进公司法理论更新和实际操作的发展，也为我国迈向市场化，法治化，国际化以及数字化转型给予了关键的制度保障。

参考文献

- [1] Rhee, R.J. (2024) Do AIs Dream of Electric Boards? *Northwestern University Law Review*, **119**, 1007-1070.
- [2] 何立民. 人工智能系统智能生成机理探索之六: 从弱人工智能、强人工智能到超人工智能[J]. 单片机与嵌入式系统应用, 2020, 20(8): 87-89.
- [3] 程威. 人工智能介入董事会的董事义务与责任重释[J]. 东北大学学报(社会科学版), 2022, 24(2): 100-108.
- [4] 孟大淇. 新《公司法》视域下董事利用人工智能决策的挑战与应对[J]. 湖北大学学报(哲学社会科学版), 2024, 51(6): 136-145.
- [5] 刘成杰. 人工智能董事法律地位及其信义义务构想——基于强人工智能视角[J]. 中州学刊, 2024(2): 79-86.
- [6] 陈忠杰. 人工智能对公司法的影响: 挑战与应对[J]. 法制与社会, 2020(4): 247-248.
- [7] Official Journal of the European Union (2024) Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 Laying Down Harmonised Rules on Artificial Intelligence and Amending Regulations (EC) No 300/2008, (EU) No 167/2013, (EU) No 168/2013, (EU) 2018/858, (EU) 2018/1139 and (EU) 2019/2144 and Directives 2014/90/EU, (EU) 2016/797 and (EU) 2020/1828 (Artificial Intelligence Act) (Text with EEA Relevance).
<https://eur-lex.europa.eu/legal-content/EN/ALL/?uri=CELEX:32024R1689>
- [8] 程威. 大模型嵌入公司治理的法理省思与规范调适[J]. 东方法学, 2025(3): 31-46.
- [9] 刘威. 人工智能介入董事决策下董事信义义务制度研究[D]: [硕士学位论文]. 长春: 吉林大学, 2024.