DOI: 10.16538/j.cnki.fem.20260424.401

# 人工智能时代的组织领导力机制探索：技术创新与伦理责任悖论视角

王文龙[1]，孔佳南[2]，席酉民[1,3]，刘　鹏[3]

（1. 西安交通大学 管理学院，陕西 西安 710049；2. 南通理工学院 商学院，江苏 南通 226002；
3. 西交利物浦大学 产业家学院与和谐管理研究中心，江苏 苏州 215123）

**摘　要：**在人工智能技术深度嵌入组织运行与决策过程的背景下，组织领导者正面临数据滥用、算法偏见、责任模糊等新型伦理难题，相关行为规范体系也亟须重构。通过系统梳理并整合相关理论，本文构建了以“主题识别→双重理性→和则-谐则耦合→动态一致性”为核心的领导力机制框架：首先，提出“主题识别”机制，以揭示技术创新与伦理规范冲突中需要优先协调的价值焦点；其次，提出“双重理性”机制，以阐明在人工智能情境下工具理性与价值理性在创新决策中的权衡逻辑及失衡风险；再次，提出“和则-谐则耦合”机制，以揭示围绕上述主题与决策展开的正式制度、技术规则与伦理规范之间的协同与张力结构；最后，构建“动态一致性”循环调节机制，以说明在人工智能技术持续演进中如何通过反馈与调整保持技术创新实践与社会伦理期待之间的持续契合。换言之，本文试图从问题源头上回应人工智能时代技术创新与伦理责任之间的悖论性挑战，进而为中国情境下重构领导行为规范、优化人工智能驱动组织变革路径以及提升组织的适应性、创新性与长期可持续发展提供了新的理论视角与制度安排依据。

**关键词：**AI技术；领导力；技术创新；伦理责任

**中图分类号：**F270　**文献标识码：**A　**文章编号：**1001-4950（2026）07-0022-19

## 一、引　言

当前管理学正处于范式更新与理论重塑的关键交汇点（徐鹏和徐向艺，2020；戚聿东等，2025）。全球化波动、数字经济与平台化扩张、人工智能（AI）深度嵌入组织运行，以及人口结构与社会价值观的持续变迁，共同构成了一个高度动荡且充满张力的外部情境。一方面，AI依托

**收稿日期：**2025-12-20

**基金项目：**国家自然科学基金项目（72572131）

**作者简介：**王文龙（1998—），男，西安交通大学管理学院博士研究生；
孔佳南（1991—），男，南通理工学院商学院副教授；
席酉民（1957—），男，西安交通大学管理学院/西交利物浦大学产业家学院与和谐管理研究中心教授；
刘　鹏（1987—），男，西交利物浦大学产业家学院与和谐管理研究中心副教授（通信作者，Peng.Liu@xjtlu.edu.cn）。

算法学习和实时数据处理极大地提升了组织的运行效率与技术创新能力(Shrestha等,2021);另一方面,数据隐私泄露、算法歧视、责任模糊等问题不断显现,使技术创新与伦理规范之间的结构性矛盾愈发尖锐(Shin和Park,2019)。在这种"不确定性—多元矛盾"交织的情境下,组织领导者不仅面临资源配置与绩效压力,还需在技术变革与伦理责任间作出艰难权衡:如何在追求短期技术突破和组织适应的同时,维护长期伦理合法性与可持续发展,成为当代领导力研究亟待回应的核心议题(Tsoukas和Chia,2002;高中华和张恒,2025)。具体而言,已有研究表明,人工智能深度嵌入组织运行与决策过程后,组织面临的已不只是单纯的技术采纳问题,而是涉及决策模式、控制关系、反馈机制与人机协同安排的一组复合性治理挑战(Kellogg等,2020;尹萌和牛雄鹰,2024;张亚莉等,2024;关健等,2025)。基于此,可以将技术创新与伦理责任之间的悖论性难题概括为以下几类典型张力:决策效率提升与审慎治理要求之间的张力,资源优化与程序公平之间的张力,数据利用扩大与隐私、透明、问责要求之间的张力,自动化决策增强与责任归属模糊之间的张力,以及短期绩效改善与长期合法性基础之间的张力(Aysolmaz等,2023;Breidbach,2024;Papagiannidis等,2025;Strohmeier等,2026)。由此,人工智能情境下的组织治理问题不再只是"是否采用技术"的问题,而是"如何在多元价值要求之间实现持续协调"的问题,这也使组织领导者面临更复杂的价值整合与责任治理任务。

既有领导力研究虽已历经"特质—行为—权变—新兴多元"的演进路径(Bass,1985;Graen和Uhl-Bien,1995;李鹏飞等,2014;章凯,2024),但在AI深度嵌入的UACCS(uncertainty,ambiguity,complexity,changeability,scarcity)情境下却显得力不从心(席酉民等,2025)。首先,主流领导力理论长期以绩效和效率为中心,强调可度量目标的实现,对AI应用引发的多元利益主体伦理冲突与价值协调问题缺乏充分关注(Anderson和Sun,2017;Fischer和Sitkin,2023)。人工智能应用将技术创新、算法治理与伦理责任更紧密地交织在一起。组织在推动技术创新的同时,不仅要追求效率提升与决策优化,还需要回应员工对程序公平、自主性与信任的关切,回应用户和社会公众对隐私保护、算法偏见与责任归属的担忧,并兼顾监管层面对透明性与可问责性的要求。既有研究表明,算法技术进入工作场域后,不仅会重塑组织控制方式,也会影响员工对管理者和组织的信任基础,从而使领导问题超越单纯的绩效管理,转向更复杂的利益相关者协调与责任治理问题(Kellogg等,2020; Leavitt等,2025)。其次,许多理论框架仍然偏重正式结构与制度安排,而对组织文化及非正式网络在缓冲技术扩张与伦理风险中的作用讨论不够充分;相关领导研究在概念边界、情境适配与机制刻画等方面仍存在进一步整合与深化的空间(Yukl和Gardner,2020;Fischer和Sitkin,2023)。再者,既有研究虽然已经从伦理型领导、负责型领导和悖论领导等视角,对组织领导者整合多元价值以及在竞争性目标之间进行规范性权衡进行了探讨(Denison等,1995;Brown和Treviño,2006;Maak和Pless,2006;Zhang等,2015),但这些研究大多数形成于一般组织情境,对于人工智能深度嵌入背景下技术创新与伦理责任如何在组织层面被同步识别、权衡、耦合和动态修正,仍缺乏更具情境针对性的机制解释。已有研究进一步指出,人工智能嵌入组织决策、反馈与协作场景后,决策链条、反馈关系与协作结构均被重新配置,组织领导者需要在技术系统、成员互动与制度规范之间进行持续协调(尹萌和牛雄鹰,2024;张亚莉等,2024;关健等,2025)。因此,人工智能时代的组织领导力研究亟需新的理论视角,以回答技术创新与伦理规范如何在组织层面实现持续协调关系。

复杂性管理范式为理解这一问题提供了重要启示。该范式将组织视为"价值生成—社会整合—系统演化"的有机体,强调在高度不确定的环境中,领导者不仅是决策者和资源配置者,更是价值引领者、文化整合者和系统进化的塑造者(Anderson和Sun,2017;Fischer和Sitkin,2023)。在此基础上,席酉民等(2020,2025)提出的和谐管理理论,以"和谐主题"为核心,通过

“和则-谐则”双规则的协同运行，追求制度安排、行为规范与文化价值之间的动态平衡，为处理多元利益冲突与价值张力提供了系统框架。然而，现有研究多停留于宏观理念和局部机制，尚未回答在AI快速迭代的具体情境中：如何动态识别并强化技术创新与伦理规范之间需要优先协调的“和谐主题”；组织领导者又如何在资源约束与利益冲突中实现工具理性与价值理性的统一；“和则-谐则”双规则如何在制度、技术与行为三个层面实现有效耦合，并通过何种机制在短期适应与长期战略之间保持平衡。

基于上述问题，本文引入和谐管理视角，尝试构建一个面向人工智能时代的组织领导力理论框架。本文首先识别AI情境下技术创新与伦理规范之间需要优先协调的和谐主题；其次，提出双重理性决策机制，阐述如何在创新决策中权衡工具理性与价值理性；接着，分析围绕这些主题与决策展开的和则-谐则耦合机制，揭示制度、技术与伦理规则如何协同工作，并维持有效耦合；最终构建实现技术创新实践与伦理期待动态一致性的调节框架。通过这一“和谐主题识别→双重理性→和则-谐则耦合→动态一致性”的理论机制链条，本文试图回答：在人工智能技术深度嵌入组织运行与决策过程的背景下，组织领导者如何识别技术创新与伦理责任之间的核心张力，如何在效率、公平、隐私、透明与责任归属等多重要求之间进行权衡，并进一步将其转化为可持续运作的组织治理机制?由此，一方面扩展和谐管理理论在数字化与智能化时代的适用范围，丰富人工智能情境下领导力理论的本土化建构；另一方面也为中国情境下设计AI治理机制、重塑领导行为规范以及提升组织适应能力与创新能力提供具有实践指向的理论支撑。

## 二、文献回顾

### （一）人工智能技术创新与伦理责任间的结构性张力

作为数字化转型的重要驱动力，人工智能（AI）正在深刻改变组织的权力结构、决策逻辑以及管理情境。借助数据分析、流程自动化与智能决策能力，AI显著提升了组织运行效率，推动组织结构走向扁平化、决策过程趋于实时化；与此同时，算法技术进入工作场域后，也在重塑组织管理边界与信任基础（Kellogg等，2020；Van Quaquebeke和Gerpott，2023）。在此背景下，技术创新与伦理责任之间的矛盾并非零散风险的简单叠加，而是一种具有持续性和内生性的结构性张力。其根源在于，AI所代表的技术创新逻辑更强调速度、效率、规模扩展与持续优化，而伦理责任逻辑则强调公平、透明、问责、隐私保护与人类控制（Papagiannidis等，2025）。

具体而言，这种结构性张力主要表现为以下几个方面。第一，创新速度与审慎治理之间的张力。组织通常希望借助AI加快决策、优化流程并提升创新迭代速度，但算法决策在带来效率和竞争优势的同时，也伴随着潜在的伦理与治理风险，因此需要更前置的治理、监督与责任控制（Breidbach，2024）。第二，效率优化与程序公平之间的张力。AI能够显著提高信息处理、筛选、评估和预测效率，但也可能因模型黑箱、数据偏差和评价标准不透明等而引发程序公平、透明性与问责争议（Aysolmaz等，2023；Strohmeier等，2026）。第三，数据利用与隐私保护之间的张力。AI的训练、优化和持续迭代依赖大规模数据收集与分析，但数据利用范围的扩大也同步强化了隐私保护、信息安全和透明使用等责任要求（Breidbach，2024；Papagiannidis等，2025）。第四，自动化决策与责任归属之间的张力。自动化决策虽然有助于提高一致性和标准化程度，却也容易模糊“由谁决策、由谁负责、如何问责”的边界，从而使透明性、可解释性和问责机制成为组织治理中的关键问题（Aysolmaz等，2023；Papagiannidis等，2025）。第五，短期绩效改善与长期合法性基础之间的张力。某些AI应用能够在短期内提升绩效表现和管理效率，但若缺乏相应的责任治理安排，则可能加剧组织在公平、透明、隐私和问责方面的风险，并进一步影响组织

信任与长期合法性基础(Breidbach,2024;Papagiannidis等,2025)。

国内相关研究从AI决策、智能反馈与人机协作等具体场景进一步表明,算法系统进入组织运行后,不仅提供信息处理和决策支持,也介入评价反馈、任务分配与协作关系塑造,并由此重构组织内部的权责配置、反馈解释与协同方式(尹萌和牛雄鹰,2024;张亚莉等,2024;关健等,2025)。因此,AI并非仅仅改变组织"如何更高效地运作",还进一步改变了组织创新的边界、责任承担的方式以及合法性维持的机制。对组织领导者而言,AI情境下的核心挑战不再只是把握效率提升与创新机会,而是如何在多元利益主体、差异化价值诉求与复杂制度约束之间形成可持续的价值整合与责任治理安排。

总体来看,现有研究较为充分地揭示了AI技术对领导情境、角色边界与能力结构的重塑作用,强调了AI在提升组织效率、增强动态适应能力和推动创新方面的潜在价值。然而,从技术创新与伦理责任关系的角度看,相关研究仍主要停留于对"效率提升—风险增加"并存现象的宏观描述,对于组织领导者如何在AI深度嵌入组织决策与运行的背景下识别并协调效率、公平、隐私、透明与责任归属等多重张力,尚缺乏更具过程性和机制性的讨论。尤其是,现有研究尚未充分说明,组织领导者如何将这些相互交织的价值冲突转化为可持续运作的组织治理安排。

(二)人工智能情境下领导力研究的主要进路

围绕AI与领导力关系,现有研究大致形成了三类分析进路。第一类研究聚焦数字化领导力或AI赋能领导力的能力模型。这一进路强调,在高不确定环境中,组织领导者需要借助AI提升战略感知、资源整合和快速响应能力,以增强组织的动态适应与持续创新能力(张志鑫和郑晓明,2023;王文龙等,2024;Hossain等,2025)。其核心关注点在于:AI如何促使组织领导者能力结构发生变化,以及组织领导者如何借助数字资源和技术基础设施保持组织竞争力。第二类研究关注AI增辅型或混合型领导模式。相关研究强调,人类领导与AI系统之间并非简单替代,而是形成互补关系:AI在数据分析、流程优化、组织监控和决策支持等方面为组织领导者提供支持,并逐步渗透到团队激励、绩效反馈乃至部分互动环节(Noponen等,2023;Van Quaquebeke和Gerpott,2023;Aziz等,2025)。这一研究路径既揭示了AI如何改变领导行为的实施方式,也强调组织领导者需要具备更高的数据素养、技术整合能力和风险敏感性(Raisch和Krakowski,2021;Tsai等,2022)。第三类研究着重讨论AI赋能的高层领导与战略决策模型。这一进路主要关注AI如何改变高层团队的信息处理方式、战略判断与商业模式创新过程。相关研究指出,AI不仅提升了高层管理者的战略预测与数据洞察能力,也通过影响AI采纳动因、组织文化与伦理考量等方面,重塑高层团队的治理逻辑与决策风格(Jorzik等,2024;Bevilacqua等,2025)。与此同时,也有研究指出,AI虽在逻辑性与计算能力上具有优势,但在情境感知和伦理判断方面仍存在不足,因而人机混合决策日益成为组织决策的重要特征(Kögel等,2026)。

总体来看,上述研究从能力结构、角色转型与人机互补等维度揭示了AI赋能领导力的多维图景,丰富了对"AI如何改变领导力"的认识。然而,这些模型仍主要聚焦于技术理性与绩效逻辑,对人工智能情境下技术创新与伦理规范的内在张力关注不足,尚未从"多元价值整合"与"行为规范重构"的视角系统地讨论领导者应如何在技术扩张与伦理约束之间进行规范性权衡。

(三)从"AI如何改变领导力"到"领导者如何治理张力"

尽管现有研究已经较为系统地讨论了AI对领导情境、角色边界与能力结构的重塑,但其分析重心仍主要集中在技术赋能、能力升级与绩效提升等方面,对技术创新与伦理责任之间的内在张力关注不足。换言之,现有研究更多聚焦于"AI如何改变领导力",而对 "组织领导者如

何治理人工智能情境下的多重价值张力”关注不足。

更具体地说，现有研究至少存在三方面不足。第一，相关研究虽然揭示了AI对组织结构、决策方式与角色边界的影响，但对技术创新逻辑与伦理责任逻辑为何会形成持续性冲突，以及这些冲突如何具体表现为效率、公平、隐私、透明与责任归属之间的张力，缺乏更系统的归纳与讨论。第二，既有关于AI与领导力的研究主要关注技术理性、绩效逻辑与动态适应能力，对组织领导者如何在技术扩张与伦理约束之间进行规范性权衡关注不足，尚未从“多元价值整合”与“行为规范重构”的角度展开深入分析。第三，现有治理框架虽然涉及数字伦理、算法监管与数据治理等议题，但大多从制度或技术层面切入，缺乏以领导行为为中心、以多元价值协调为主线的系统解释，也未能充分说明组织领导者如何在组织层面将这些张力识别出来、加以权衡，并进一步转化为可持续运作的治理安排。

综合现有研究，AI技术的广泛应用正在重塑领导力的边界、能力结构与评价标准。学界普遍强调，在数字化环境下，领导者需要具备更强的敏捷性、创新性与动态适应能力，并关注技术赋能在提升组织绩效与流程优化方面的积极作用（Aziz等，2025；Bevilacqua等，2025）。与此同时，相关研究也开始注意到，AI深度嵌入组织运行与决策过程后，组织所面对的已不再只是单纯的技术采纳问题，而是涉及决策模式、控制关系、反馈机制与人机协同安排的一组复合性治理挑战（Kellogg等，2020；尹萌和牛雄鹰，2024；张亚莉等，2024；关健等，2025）。在此背景下，现有治理框架虽然涉及数字伦理、算法监管与数据治理等议题，但多数仍从制度或技术视角切入，对组织领导者如何在技术创新与伦理责任之间进行动态协调的讨论仍显不足。

值得注意的是，既有领导力研究已经从伦理型领导、负责型领导和悖论领导等视角，对组织领导者如何整合多元价值、如何在相互竞争但彼此关联的目标之间进行规范性权衡作出了较为充分的讨论（Brown和Treviño, 2006；Maak和Pless，2006；Zhang等，2015）。然而，将上述研究置于人工智能深度嵌入的组织情境中审视，可以发现其解释仍存在一定边界。现有研究更多关注一般意义上的道德示范、利益相关者协调或竞争性目标平衡，对于人工智能条件下技术系统、组织制度与伦理要求交织所形成的复合性张力，尚缺乏更具过程性的分析。特别是，技术创新与伦理责任并非在组织中静态并存，而是会随着算法应用范围扩展、数据使用加深，以及人机协作方式变化而不断被重新界定和调整。就此而言，现有研究对于这些张力如何在组织层面被持续识别、嵌入决策并转化为具体治理安排，仍缺少更为细致的说明。因此，现有文献整体上仍主要停留在“技术赋能—绩效改善”或“治理议题—制度回应”的分析框架之中，对于人工智能情境下组织领导者如何围绕技术创新与伦理责任之间的多重张力形成持续性的价值整合与治理机制，尚未形成系统解释。

（四）和谐管理理论的解释潜力

和谐管理理论强调，组织并非在单一目标和静态均衡中运行，而是在多元主体、多重价值和持续变化中寻求相对协调。与以效率最大化和工具理性为中心的传统管理逻辑不同，该理论将组织理解为一个在矛盾并存、关系互动和环境演化中不断调整的复杂系统，强调管理活动需要同时回应制度安排、行为规范与价值诉求之间的动态关系（席酉民等，2005，2013）。在数字化与智能化快速发展的背景下，这一理论尤其具有解释意义，因为AI技术进入组织后所引发的，不仅是流程优化和效率提升问题，更是技术扩张、伦理责任、制度约束与利益协调交织而成的复合治理问题。

从理论结构来看，和谐管理理论在三个方面为理解人工智能时代技术创新与伦理责任之间的张力提供了分析基础。首先，和谐管理理论强调“和谐主题”的识别与聚焦功能。王亚刚和席酉民（2008）指出，和谐主题具有“聚焦—引领—协同”的作用，能够帮助组织在复杂情境中识

别需要优先处理的核心矛盾。对人工智能情境而言，组织所面对的已不再是单一的技术采纳问题，而是围绕效率、公平、隐私、透明、问责与长期合法性等方面形成的多重张力，因此，能否识别并持续聚焦这些张力中的核心议题，构成了组织治理能否有效展开的前提。其次，和谐管理理论中的“和则-谐则”双规则机制，有助于解释组织如何在价值整合与制度约束之间实现动态平衡。其中，“谐则”强调正式制度、流程安排和结构设计，“和则”强调文化认同、关系协调和自主演化（席酉民等，2025）。这一理论视角对于人工智能情境尤为重要，因为AI治理既不能仅依赖正式制度与技术规范来控制风险，也不能只依赖价值倡导和文化共识来消解冲突，而必须通过制度设计与关系协调的共同作用，在技术创新与伦理责任之间形成可持续的组织回应。换言之，和谐管理理论能够帮助解释：在AI大规模应用、流程重构和制度再造过程中，组织如何通过正式规则与非正式规范的耦合，将多元价值冲突转化为可治理的实践过程。再次，和谐管理理论本质上蕴含一种多元理性框架。相关研究指出，领导者不能仅以效率、绩效等工具理性为唯一判断标准，而需要同时兼顾伦理、公平、信任、身份认同与长期关系等更高层次的价值诉求（李鹏飞等，2013）。这一点与人工智能情境下技术创新与伦理责任并行强化的现实高度契合。因为在人工智能深度嵌入组织运行与决策过程后，组织领导者面对的并不是单纯的“是否采用技术”问题，而是在多元利益相关者诉求之间如何进行规范性权衡、如何在短期适应与长期协调之间保持平衡的问题。和谐管理理论因此为理解人工智能情境下领导者如何整合工具理性与价值理性提供了重要理论资源。

已有研究也在一定程度上表明，和谐管理理论在复杂技术环境和中国情境下具有较强的适应性与实践价值。相关研究发现，通过强化和谐主题、推动和则-谐则耦合并关注生态共生，企业能够在复杂环境中识别战略主线，实现人—技术—资源要素的整体优化与联动升级（席酉民等，2020；张梦晓等，2024）。林楠等（2022）对平台型企业的研究进一步表明，在AI与数字赋能背景下，和谐主题的持续识别与双规则的动态运用，有助于企业应对技术变革中的不确定性，实现能力结构升级与生态位重构。上述研究说明，和谐管理理论不仅能够解释组织如何应对复杂环境，也能够为分析技术变革背景下的多元张力治理提供本土化视角。

总体来看，和谐管理理论以其对多元矛盾、价值张力与长期协调问题的强调，为AI时代技术创新与伦理责任之间的关系提供了更具整合性的解释框架。不过，现有研究仍主要停留在宏观治理或组织层面的和谐机制讨论，对于在人工智能时代技术创新与伦理责任的结构性张力背景下，组织领导者如何识别核心主题、进行规范性权衡，并进一步通过“和则-谐则”耦合推动组织层面的持续协调，尚缺乏面向领导行为的系统理论建构。

## 三、研究设计与模型推演

### （一）研究范式与方法选择

本研究采用理论建构（theory building）范式，聚焦AI情境下组织领导力模型的理论探索。该研究方法特别适用于理论前沿尚未成熟、现实情境快速演化且亟需新的概念框架来解释复杂现象的研究领域（Suddaby，2010；Shepherd和Suddaby, 2017）。不同于以检验为导向的传统实证研究，理论建构旨在通过系统梳理文献、整合关键概念，并基于逻辑推理生成新的理论模型和理论命题（Bacharach，1989）。

具体而言，本文采用“文献分析—概念整合—模型推演—理论命题生成”的路径（Edmondson和McManus，2007；Gioia等，2013），这一路径已被广泛应用于理论型研究。首先，通过系统梳理AI情境下的领导力、和谐管理等相关领域的核心文献，识别并归纳影响领导力特征变化的关键要素。其次，以和谐管理理论及其核心假设为基础，分析AI赋能组织中的新型互动关系与管

理挑战。最后，结合逻辑推理和情境分析，逐步建构面向AI时代的组织领导力理论模型，并提出可为后续实证研究提供基础的理论命题。

（二）关键命题建构

在人工智能深度嵌入组织运行与决策过程的背景下，技术创新与伦理责任之间的张力并不表现为单一层面的价值冲突，而是沿着核心议题形成、规范性判断生成、治理安排嵌入与持续反馈修正的过程不断展开。和谐管理理论强调，组织需要在复杂多变环境中围绕“和谐主题”聚焦关键矛盾，并通过和则、谐则及其耦合机制实现阶段性的相对协调，这为理解人工智能情境下多重张力的治理过程提供了一个过程性框架（席酉民等，2005，2025）。基于此，本文将人工智能情境下领导行为的运行逻辑概括为和谐主题识别、双重理性决策、和则-谐则耦合与动态一致性四个相互嵌套的机制。

和谐主题识别构成机制链的起点。在多重技术议题、价值诉求与利益冲突并存的条件下，组织并不能同时回应所有问题，而需要从复杂情境中提炼出需要优先协调的核心张力，并将其凝练为组织治理的中心议题。围绕这一主题，组织领导者进一步进入双重理性决策过程。此时，技术创新不再只是效率提升和资源优化的问题，伦理责任也不只是抽象价值宣示的问题，二者都需要在具体决策中被同时纳入考量，组织领导者既要评估技术应用在速度、绩效和适应性方面的收益，也要权衡其在公平、透明、责任和合法性方面的后果，由此形成对张力的规范性判断。规范性判断若要真正转化为组织行动，还必须嵌入制度、技术与行为之中，这就要求和则与谐则的协同展开：前者体现为价值引导、关系协调与文化整合，后者体现为制度安排、流程设计与正式规则建构。人工智能情境下的张力治理既不能仅依靠正式制度和技术规范来加以控制，也不能仅依赖文化认同和价值倡导来加以缓冲，而是需要二者在组织实践中形成持续耦合，使技术创新与伦理责任之间的协调要求获得制度化与行为化的承载。进一步地，由于人工智能技术持续迭代、应用边界不断扩展、利益关系和外部环境持续变化，前述主题识别、理性判断与规则耦合并不会一次完成后保持稳定，而需要在实践中不断接受反馈、修正偏差并重新平衡。动态一致性正是在这一意义上展开：它使组织能够在持续变化的条件下不断调整主题、判断与规则之间的关系，从而维持技术实践与伦理期待之间的相对一致。

由此，四个机制并非彼此并列，而是沿着“主题识别—理性权衡—和则-谐则耦合—动态修正”的路径层层展开。和谐主题识别规定了治理对象，双重理性决策形成了判断基础，和则-谐则耦合提供了组织化实现路径，动态一致性则保证这一机制链能够在持续变化的人工智能情境中不断被修正和维持。

基于上述四项相互关联的机制，以下各节将依次展开关键理论命题的建构。

1. 主题识别能力

在数智时代，人工智能（AI）驱动的深度数字化转型，使组织领导者必须在多元议题与价值冲突交织的情境中，识别并聚焦技术创新与伦理责任之间需要优先协调的关键议题。与传统主要围绕绩效提升和资源配置展开的目标设定不同，AI情境下的主题识别不再只是对任务与指标的界定，而是对核心张力的提炼、排序与聚焦。传统“目标设定”逻辑在相对稳定环境中能够发挥较为清晰的导向作用（Cyert和March，1963；March和Olsen，1979），但在技术创新与伦理责任高度交织的AI情境中，单一目标导向已难以回应多元利益诉求和复杂价值冲突。由此，组织领导者不仅需要识别关键问题，更需要将事关组织长期合法性与可持续发展的核心问题上升为组织层面的“和谐主题”，从而为后续的规范性判断与治理安排提供方向。

在这一背景下，人工智能情境下可被识别并上升为和谐主题的核心议题，主要包括效率提升与程序公平、数据利用与隐私保护、自动化决策与责任归属、短期绩效改善与长期合法性维

持等。在不同组织、行业与应用场景中，这些主题的重要性及其优先级并不相同。因此，主题识别并非对既有议题的静态呈现，而是一个围绕核心议题持续进行筛选、排序与聚焦的过程。也正因为如此，主题识别能力的关键不在于议题数量的扩展，而在于能否从众多技术问题和价值冲突中提炼出当前最值得优先回应的中心议题。

主题识别之所以在AI情境下变得尤为重要，是因为技术应用带来的风险并不总是以显性的故障或失灵形式出现，而更多表现为价值冲突、责任模糊与制度边界松动的累积效应。AI应用过程中的数据公正性、隐私保护与算法偏见等问题，使效率、伦理、公平与社会责任等价值不断发生交叉与碰撞（Mikalef等，2022）。如果组织领导者仅从技术可行性或绩效改善的角度理解AI采纳，往往容易将伦理风险视为附属问题，从而使组织暴露于舆论危机与合法性风险之中。与此同时，AI的黑箱性、不可解释性以及自动化决策可能带来的不可预期后果，也削弱了传统风险评估与责任界定方式的有效性，使技术部门的单点评估与线性风险管理难以覆盖全局（Van Quaquebeke和Gerpott，2023）。在此意义上，主题识别能力意味着领导者能够突破“技术问题—技术解决”的线性思维，将零散的技术风险、伦理争议与责任模糊整合为具有战略意义的核心议题，并由此推动跨部门、跨层级的协同应对。进一步来看，主题识别并不是单向度的高层界定，而表现为一种分布式的意义建构过程。随着信息透明度提高与员工技术素养增强，组织成员对AI伦理边界及其治理后果的关注与参与意愿显著提升（Aziz等，2025）。在这种条件下，传统由少数高层单向界定议题并主导技术采纳的模式，越来越难以获得成员的认同与支持。组织领导者在主题识别过程中不仅需要充当问题界定者，更需要发挥协调者与桥梁构建者的作用，通过跨部门协作、内部对话和利益相关者参与，共同明确何为需要优先回应的核心问题，以及技术创新应遵循何种价值边界。也就是说，和谐主题并不是预先给定的，而是在互动、协商和再确认中逐步生成的。

这种意义建构的过程进一步指向组织信任链与价值体系的重构。AI技术的自动化特征与非人格化倾向，容易放大组织成员对技术采纳的心理不安，而这种不安并不单纯来源于对技术本身的不熟悉，更源于对其背后价值取向与责任后果的不确定性（Hossain等，2025）。在这种情况下，组织领导者能否敏锐识别并清晰表达组织在AI采纳过程中真正坚持的价值为何、哪些伦理底线不容突破，将直接影响成员对相关规则的认同程度以及对技术变革的适应意愿。组织领导者若能在议题优先排序中纳入伦理考量、技术公平性、员工参与度和社会责任等，并配合透明沟通、参与式且透明的决策程序以及相对公平的资源分配机制，便有助于在主题识别过程中逐步修复并重塑组织信任基础，从而使AI采纳更可能被视为共同治理下的组织选择，而非单向推动的技术压力。综上，AI情境下的主题识别能力不仅体现为对外部技术趋势的前瞻性感知，更体现为对内部价值冲突、责任边界与信任结构的系统性梳理。它决定了组织能否在议题源头上将技术创新与伦理责任纳入同一分析框架，并为后续的理性权衡、和则-谐则耦合与动态一致性奠定前提（见图1）。由此提出：

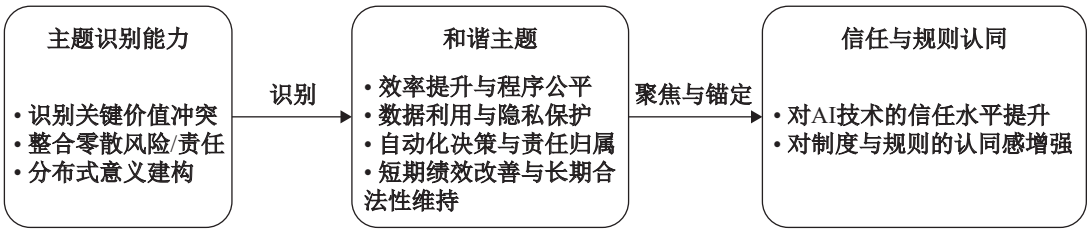


**图 1　主题识别能力**

命题一：在AI情境下，组织领导者的主题识别能力越强，组织越能够清晰界定AI应用中的核心价值张力与责任边界，从而增强成员对相关规则与治理安排的认同，并提升其对技术变革的适应程度。

2. 双重理性决策

当组织围绕人工智能应用形成需要优先回应的和谐主题之后，领导者面临的关键任务便不再是单纯推进技术部署，而是在“决策—判断”层面形成能够同时回应创新诉求与伦理责任的规范性判断。所谓双重理性决策，指的是领导者在人工智能情境下，不仅依据效率、成本、速度和绩效等工具性标准进行决策，还将公平、透明、责任、信任与正当性等价值性要求纳入同一判断过程，通过两类理性的动态权衡形成治理选择。由此，组织领导决策不再是单一目标优化的结果，而是一个在技术创新收益与伦理责任约束之间持续调节的过程。

人工智能驱动的数字化转型一方面为组织带来前所未有的技术创新机会，另一方面也使决策过程暴露于更高频率、更高敏感度的伦理风险之中。算法赋能使决策更加迅捷、资源配置更加精准，但同时也提高了数据滥用、算法歧视与责任模糊等问题的暴露频率及其社会可行性（Floridi等，2018；Van Quaquebeke和Gerpott，2023）。在这种不确定性、多重矛盾与利益张力并存的情境下，领导者无法再沿用单一理性主导的判断方式，而必须在效率与合规、创新与安全、短期绩效与长期责任之间持续进行权衡。若仅从绩效视角理解AI应用，决策容易滑向技术决定论与“自动化崇拜”；若仅从风险控制与伦理约束出发，又可能使组织因过度谨慎而压缩必要的试错空间与创新窗口。因此，双重理性并不是两种判断逻辑的外在并置，而是AI时代组织决策能够维持创新活力与伦理边界的基本前提。

从具体内涵来看，工具理性主要体现为对效率、成本、产出、速度、精确性与资源配置能力的考量。AI技术的引入强化了这一判断逻辑：算法优化、流程再造与预测能力提升，使组织能够显著提升决策速度和资源利用效率，从而在竞争中获得优势。然而，单一工具理性导向容易将伦理、心理与社会后果视为外部性，忽视成员价值与社会期待，进而引发信任危机、合法性风险乃至组织文化侵蚀（Floridi等，2018）。在AI情境中，这种偏向往往表现为决策黑箱化、数据使用边界不清、员工“被替代”与“被监控”的焦虑加剧，以及对技术结果的过度依赖。与之相对，价值理性强调决策的公平性、包容性、透明性、责任性与社会可接受性，将规范合法性与价值契合度置于重要位置（Weber，1978）。这一判断逻辑有助于增强组织的合法性基础，强化成员的价值认同与规则认同，对于回应AI引发的伦理争议尤为关键。但若将价值理性绝对化，组织又可能因过度审慎而弱化必要的冒险精神与创新试探能力，使治理机制趋于僵化。由此可见，人工智能情境下的核心问题并不在于选择工具理性或价值理性中的某一方，而在于如何使二者在同一决策过程中保持动态嵌合。前者更多指向短期、可量化和结构化的目标，后者则更多指向长期、难以完全量化但具有规范意义的价值要求，双重理性决策正是在这两类判断标准之间形成持续性的张力平衡。

因此，双重理性决策并不是两种理性的机械折中，而是一种嵌合式、情境化的判断结构。组织领导者在推动技术创新、流程优化与绩效提升的同时，需要将伦理责任和社会责任内嵌于决策过程，使工具理性与价值理性成为相互调节、共同塑造治理选择的两条逻辑线。在实践中，这种动态嵌合通常表现为理性权重的情境性调节：当AI应用进入高敏感、高风险领域时，价值理性的权重相对提高，领导者会更强调伦理审查、透明度要求、责任链条清晰化以及员工心理支持；当AI应用处于低风险或探索性阶段时，工具理性的权重可以适度提高，从而为试错创新、流程优化与快速迭代保留空间。双重理性的意义并不在于消除张力，而在于使组织能够在张力中形成可持续的判断结构，而不是陷入效率导向与伦理导向的零和博弈（Hossain等，2025）。在

这一过程中，领导者实际上承担着“理性中介”的角色。其一，组织领导者需要警惕将技术进步视为组织问题自动解决方案的“算法幻觉”，避免组织对单一工具理性形成路径依赖；其二，组织领导者也需要防止价值理性滑向抽象的原则宣示或过度风险规避，从而压制必要的创新探索。双重理性决策由此要求组织领导者持续回答两个问题：当前AI应用的创新边界在哪里，伦理与责任的底线又在哪里。只有当这两条边界能够在具体情境中被同时纳入决策时，组织才可能在保持创新活力的同时维持价值稳定性，并为后续的规则耦合与动态一致性提供判断基础。

综上，双重理性决策并非传统意义上的决策风格选择，而是人工智能时代组织实现持续治理的关键机制。它使组织领导者能够在效率与伦理、创新与合规、短期业绩与长期责任等多重张力中形成更具反思性的判断能力，并将技术创新导向与价值责任要求同时嵌入组织决策过程（见图2）。由此提出：

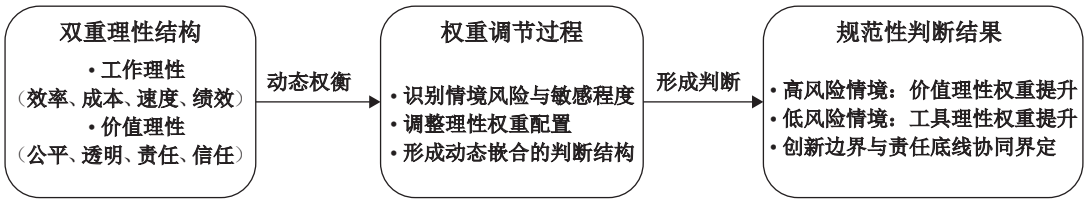


**图 2　双重理性**

命题二：在人工智能情境下，组织领导者越能够在工具理性与价值理性之间形成动态嵌合的双重理性决策结构，组织越可能在推进技术创新的同时维持伦理边界与规则稳定，并促进技术应用与组织文化之间的持续协调。

3. “和则-谐则”耦合治理体系

在完成围绕技术创新与伦理责任的规范性判断之后，领导问题进一步转化为一个组织化命题：如何将前述判断嵌入制度安排、技术流程与成员行为之中，使其不只停留于价值宣示，而能够在实际运行中形成持续约束与有效引导。人工智能深度嵌入组织治理后，治理重心已不再局限于流程优化和技术升级，而更多体现为围绕规则结构、责任链条与价值边界展开的系统性重构。在这一背景下，单纯依赖正式制度、流程规范与合规要求，往往难以充分回应技术伦理与数据安全问题的复杂性与演化性；而若过度依赖非正式管理、价值倡导与文化引导，又可能导致规则边界模糊、责任配置松动以及治理弹性失控（李鹏飞等，2014；席酉民等，2020，2025）。因此，人工智能情境下治理的关键，并不在于制度约束或文化引导的单向强化，而在于通过和则与谐则的持续耦合，将技术创新与伦理责任之间的张力转化为可组织化承载的治理过程。

就其功能分工而言，谐则主要体现为制度化、程序化和结构化的治理安排，其作用在于将组织围绕AI应用形成的边界判断具体化为规则体系，例如权限配置、责任链条、审查流程、数据留痕、例外处理与风险升级机制等（席酉民等，2020，2025；Bevilacqua等，2025）。谐则的重要性在于，它为组织提供了清晰的行为边界和责任框架，使AI治理不至于停留在原则性表态上，而能够进入可追踪、可问责、可检验的运行状态。与之相对，和则主要体现为价值引导、关系协调与文化整合，其作用不在于替代制度，而在于使制度安排能够被成员理解、接受并持续运用。和则所提供的是组织内部对技术应用边界的共同理解、对责任要求的共享认知以及在情境变化中进行柔性协商的文化基础。没有谐则，治理将缺乏结构性边界；没有和则，治理则难以获得真正的组织承载。二者并非相互替代，而是分别从制度刚性与关系弹性两个方向，为AI治理提供支撑。

和则与谐则的关键不在于二者同时存在，而在于它们如何在组织实践中形成持续耦合。就

其展开方式而言，这种耦合并不是静态配置，而是一个不断循环的组织过程。其一，组织围绕已识别出的核心张力形成基本共识，即明确在特定AI应用场景中，哪些效率诉求可以推进，哪些责任底线不可突破。其二，这种共识被进一步转译为可操作的制度表达，包括流程设计、行为规范与技术控制要求，使价值边界获得制度化承载。其三，正式规则进入执行之后，还需要通过沟通、解释、参与和协商机制等方式嵌入具体行动，使成员不仅知晓规则，而且理解规则何以如此设定。其四，随着AI技术迭代、业务场景扩展和外部评价变化，组织还需要根据实践中的技术后果、伦理争议与治理绩效，对前述价值表达与制度设计进行同步修正，从而维持和则与谐则之间的持续协同。由此，和则-谐则耦合实际上体现为一个“共识生成—规则转译—执行嵌入—反馈修正”的循环过程。

这一过程并不存在固定不变的作用比例，而是呈现出显著的情境依赖性和阶段差异性。对于探索性较强、边界尚不清晰的AI应用，组织往往需要更多依靠和则来维持开放协作、试验容错与跨界学习，因为此时正式规则尚难完全覆盖不断涌现的新问题；而当AI系统进入规模化应用、涉及敏感数据、关键岗位或高风险决策时，谐则的重要性会显著上升，组织需要通过更明确的流程、审查与责任配置来防止技术失控（郭小东，2025；Bevilacqua等，2025）。真正成熟的治理并不意味着某一类规则的绝对强化，而在于组织领导者能够根据技术类型、风险等级和利益相关者结构，不断调节和则与谐则之间的作用方式，使正式制度与非正式规范保持相互支撑而非相互抵消。

从这一意义上说，和则-谐则耦合治理体系并不是对制度安排与文化规范的静态并置，而是对组织如何承载技术创新与伦理责任张力的一种过程性解释。它表明，人工智能治理的实质既不是单纯依赖合规控制来维持秩序，也不是停留于抽象价值倡导来凝聚共识，而是通过正式规则与非正式规范的持续互嵌，将前述规范性判断转化为可运行、可修正、可延续的组织秩序。只有当这种耦合达到较高水平时，组织才可能在人工智能带来的结构性不确定性中，同时维持创新活力、治理稳定性与价值边界的清晰性（见图3）。基于此，本文提出：

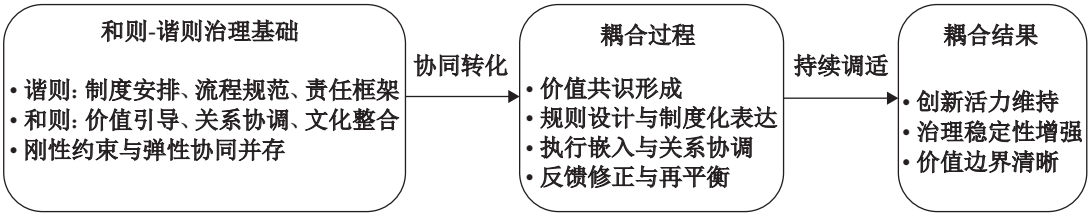


**图 3　和则-谐则耦合**

命题三：组织中和则-谐则的耦合程度越高，越能够将技术创新与伦理责任之间的张力转化为可持续运作的治理安排，从而在人工智能情境下更有效地维持创新活力、治理稳定性与价值边界之间的动态平衡。

4. 动态一致性

在人工智能情境下，组织治理并不会因为形成了和谐主题、双重理性判断以及和则-谐则耦合安排而自然保持稳定。相反，技术路径的持续迭代、业务场景的不断扩展、监管要求的快速变化以及社会伦理议题的周期性升温，都会不断改变原有治理安排的适用边界。由此，技术创新与伦理责任之间的协调不可能被理解为一次性实现的静态平衡，而只能被理解为一个需要在时间维度上持续维持的动态过程。所谓动态一致性，正是指组织能够在环境扰动和内部演化中，不断校正主题、判断与规则之间的关系，使技术实践与伦理期待之间始终保持阶段性的相

对一致。

动态一致性的核心不在于消除变化，而在于将变化转化为可感知、可解释和可调整的治理过程。即便组织已经在某一阶段形成了较为清晰的议题聚焦、规范性判断与规则耦合结构，这一结构也可能随着AI系统功能升级、应用边界扩展或外部评价变化而迅速失效。传统管理框架往往依赖预设制度与标准流程来维持秩序，但在面对AI技术快速迭代、监管政策高频调整与伦理争议持续生成时，静态制度设计很难长期覆盖不断涌现的新问题。正因如此，动态一致性机制强调，组织必须在时间维度上持续维持“议题—规则—理性”之间的匹配关系，而不是将前述安排视为一次设计后即可长期适用的固定结构。

这一机制的运行，体现为一个不断循环的组织过程。首先，组织需要持续感知偏差，即通过对技术应用后果、伦理争议、市场反馈和政策信号的追踪，识别现有治理安排与现实情境之间是否出现新的错位。其次，组织需要对偏差信号进行解释和提炼，将零散的问题重新上升为值得回应的核心议题，从而推动对原有和谐主题的重新审视。再次，在主题重估的基础上，组织需要同步调整既有规则安排与理性权衡方式，使和则-谐则的耦合方式以及工具理性、价值理性的权重配置能够适配新的技术条件与治理要求。最后，这些新的调整还需要重新嵌入制度流程、组织协作与成员行为之中，形成下一轮运行的基础。由此，动态一致性并不是单纯的“事后修补”，而是一个“发现偏差—重估主题—调整规则与理性—重新嵌入行动”的持续循环过程。

这一过程之所以不同于传统危机应对，在于它并不将组织理解为依靠偶发性纠偏恢复原状的封闭系统，而是将组织视为一个能够通过反馈、学习与再平衡不断进行自我更新的复杂适应系统（Senge，2006；Bevilacqua等，2025；Hossain等，2025）。在这一意义上，动态一致性机制既是前三种机制得以延续的时间性条件，也是连接组织短期治理响应与长期演化能力的关键枢纽。组织若缺乏动态一致性，即便曾在某一阶段形成适宜的议题聚焦、规则组合与理性结构，也可能因环境加速变化而迅速过期失效；只有通过持续的学习与修正，前述机制链条才可能在人工智能情境中保持开放性与适应性。在操作层面，动态一致性要求组织构建多层级、多主体参与的反馈与学习体系。信息层面，需要通过跨部门数据共享、伦理风险通报、员工反馈通道与外部评价监测，形成对AI应用后果的持续跟踪能力；判断层面，需要将这些反馈系统性地纳入领导决策过程，使原有主题识别、规则耦合与理性权重能够根据新情境被重新界定；行动层面，则需要把调整后的安排重新嵌入制度设计、流程运行与行为规范，形成“反馈—学习—再嵌入”的闭环。通过这一闭环，组织不仅能够更及时地应对技术不确定性、伦理争议与外部压力，也能够使创新不再被视为突破既有规则的例外，而是在动态边界中持续生成的常态。在这一机制中，组织领导者所承担的并不是简单的控制者角色，而是一种持续性的动态协调角色。一方面，组织领导者需要具备跨部门、跨专业的信息整合能力，将技术团队、合规部门、一线员工和外部利益相关者所生成的碎片化信号转化为可用于组织判断的系统性认知；另一方面，组织领导者还需要把握调整的节奏与强度，既避免因调整迟缓导致风险累积，也避免因过度频繁的变动而削弱组织认同与制度稳定性。动态一致性的真正作用，就在于使组织能够在高度不确定的人工智能环境中保持敏捷性而不丧失价值定力，在不断变化的条件下维持技术实践与伦理期待之间的相对协调。

综上，动态一致性是使主题识别、双重理性决策与和则-谐则耦合得以持续有效运作的时间性机制。它通过持续反馈、组织学习与动态再平衡等方式，使组织在技术演化、制度调整与价值重估过程中不断更新既有治理安排，从而维持技术创新与伦理责任之间的长期协调。基于此，提出如下命题：

命题四：在人工智能情境下，组织的动态一致性机制越完善，越能够通过持续反馈、学习

与调整维持技术实践与伦理期待之间的相对一致，并增强既有领导力机制对环境变化的适应能力。

回顾前述分析，AI场域下的组织领导力模型以“主题识别→双重理性→和则-谐则耦合→动态一致性”为核心链条，通过多环共生与层级互动，构建出一套系统化、演化型的组织治理框架（见图4）。动态主题识别为组织的AI战略和价值治理提供方向锚点，使组织领导者能够穿透技术表层，精准把握AI伦理、员工认同、算法公平等时代关键议题，并在多元利益相关者之间形成意义共建与集体认同（Van Quaquebeke和Gerpott，2023；席酉民等，2025）。其次，双重理性则让组织能够在效率、创新、合规、合法等多重目标的张力中进行“理性切换”与动态调节。通过平衡工具理性与价值理性，组织能够将技术驱动与价值引领深度融合，为制度运行以及可持续发展奠定工具理性与价值理性并行的基础（Hossain等，2025）。再者，和则-谐则协同使正式规则与弹性规范在AI场域中实现动态耦合，一方面保障技术合规、伦理底线与责任清晰，另一方面释放创新活力和团队自主性，从而最大限度降低技术风险、内部摩擦与治理内耗（Aziz等，2025；Bevilacqua等，2025）。最后，动态一致性机制确保组织在面对技术演进、政策变化与伦理议题涌现时具备自我反思、自我修复与系统升级的能力，使组织得以沿着“自我革新—创新共生—生态繁荣”的演化路径实现持续进化（Senge，2006；席酉民等，2025）。

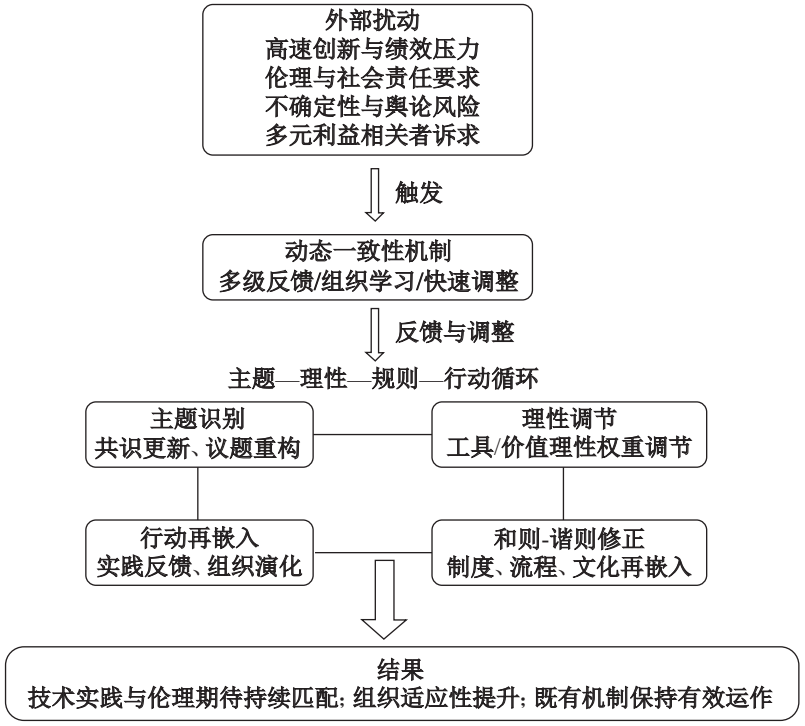


**图 4　AI情境下领导力模型图**

## 四、研究结论、理论贡献和展望

### （一）研究结论

本文立足AI深度赋能的中国情境，围绕“技术创新与伦理规范张力日益凸显的背景下，领导行为应遵循何种规范逻辑”这一核心问题，从和谐管理视角构建了“人工智能时代组织领导力机制链”，主要结论如下。AI时代的组织领导情境呈现出技术创新压力、绩效考核要求、伦理

风险暴露与多元利益诉求交织的结构性张力，传统以绩效为中心的线性领导逻辑难以解释领导实践的复杂性与矛盾性。组织领导力因而需要被重新理解为嵌入技术系统、制度安排与价值秩序中的复合治理机制。主题识别机制构成化解“技术创新—伦理规范”悖论的起点。通过对AI技术议题、伦理争议与多元诉求进行系统梳理，提炼需优先协调的和谐主题，组织领导者得以在效率、责任与合法性之间建立共同价值锚点，为后续的制度安排与行为规范明确方向。双重理性决策机制为领导者在技术效率、绩效压力与伦理责任之间提供了可操作的权衡框架。通过在决策过程中嵌入工具理性与价值理性两条判断逻辑，并根据应用情境调整其权重，组织既能把握AI带来的创新机会，又能坚守伦理底线和社会责任。“和则-谐则”规则耦合机制成为连接技术变革与伦理治理的关键中介。正式制度（谐则）与非正式规范（和则）只有围绕和谐主题进行场景化配比与动态协同，才能在AI治理中同时兼顾风险防控、创新活力与价值整合，避免陷入“高压合规—低信任”或“高弹性—低约束”的单向失衡。动态一致性机制为上述三类机制在时间维度上持续发挥作用提供演化平台。通过多级反馈、组织学习与快速调整，组织在“和谐主题—双重理性—规则耦合—组织进化”之间形成循环联动，在外部环境持续扰动中保持领导力的适应性与创新性，从而为理解AI情境下的中国式领导力提供了一套系统化的理论解释。

（二）理论贡献

基于和谐管理理论与复杂性管理视角（席酉民等，2005），本文围绕人工智能时代技术创新与伦理责任之间持续存在且不断演化的张力，构建了由和谐主题识别、双重理性决策、和则-谐则耦合与动态一致性构成的组织领导力机制链。既有组织悖论研究主要强调，相互冲突却又彼此依赖的要素需要被同时维持，而非被简单消解；相关领导力研究则分别从伦理型领导、负责型领导与悖论领导等角度，讨论了价值整合、利益相关者责任与竞争性需求平衡问题（Lewis，2000；Brown和Treviño，2006；Maak和Pless，2006；Smith和Lewis，2011）。在此基础上，本文进一步揭示，在人工智能深度嵌入组织运行与决策过程之后，这些张力如何被识别、被权衡、被组织化承载，并在时间维度上不断修正。

第一，本文将领导力研究的分析起点由传统的目标设定推进到核心议题识别，拓展了人工智能情境下组织领导力研究的问题边界。既有组织悖论研究强调张力的共存、循环强化及其动态平衡逻辑（Lewis，2000；Smith和Lewis，2011），但在人工智能深度嵌入之后，技术创新与伦理责任之间的张力并不首先表现为抽象的目标冲突，而更多呈现为算法公平、数据隐私、责任归属、员工认同等多元议题的交叉叠加。相较于传统将领导力理解为围绕既定目标展开影响与激励的分析框架，本文进一步指出，组织领导问题首先体现为对关键议题的辨识、排序与聚焦。正是在这一意义上，本文提出“和谐主题识别”这一机制，使技术创新与伦理责任之间的冲突不再只是一般意义上的目标平衡问题，而是被转化为组织需要优先回应的和谐主题。这样的推进，也使人工智能时代的组织领导力研究从“目标导向”进一步延伸至“议题导向”。

第二，本文将既有领导力研究中的价值整合问题推进到人工智能情境下的双重理性决策结构。伦理型领导研究强调领导者的道德示范与伦理氛围塑造，负责型领导研究强调面向利益相关者的责任承担与关系协调，悖论领导研究则揭示了领导者如何同时回应相互竞争却又彼此依赖的需求（Brown和Treviño，2006；Maak和Pless，2006；Zhang等，2015）。这些研究已经为理解多元价值整合提供了重要基础。本文在此基础上进一步看到，在人工智能深度嵌入之后，组织领导者面对的并不仅是一般意义上的价值冲突，而是效率、成本、速度、绩效等工具性要求与公平、透明、责任、信任等价值性要求同时进入同一决策过程。由此，本文提出“工具理性—价值理性”的双重理性结构，并强调两类理性并非一次性折中，而是在不同风险水平、应用阶段和场景敏感度下，通过权重调节形成动态嵌合的判断结构。相较于既有研究更多从道德引导、责任

承担或竞争性需求平衡等方面来解释领导决策，本文更进一步揭示了人工智能情境下规范性判断的形成机制。

第三，本文将制度与文化之间的互补关系推进到人工智能情境下可展开、可调节的耦合治理过程。既有研究已经注意到，正式制度与非正式规范之间并非简单对立，而是共同影响组织治理的运行效果；和谐管理理论中的“和则-谐则”双规则结构也为理解制度刚性与关系弹性并存提供了本土化视角（王亚刚和席酉民，2008；席酉民等，2020，2025）。本文进一步表明，在人工智能治理中，问题并不在于正式规则和非正式规范何者更重要，而在于二者如何围绕技术创新与伦理责任之间的核心张力形成持续耦合。沿着这一思路，本文将“和则-谐则”由一般性的制度—文化协同关系，推进为一个包含“价值共识形成—规则设计与制度化表达—执行嵌入与关系协调—反馈修正与再平衡”的组织过程。这样一来，前述双重理性判断不再停留于理念层面，而能够被转化为流程、责任链条、行为规范和文化认同共同支撑的治理安排。相较于既有研究较多从制度与文化互补的角度理解治理关系，本文更进一步揭示了二者在人工智能情境中的阶段差异、场景依赖与动态协同。

第四，本文将组织悖论研究中的动态平衡进一步推进为一种具有时间维度的机制链逻辑。既有组织悖论研究与复杂性管理研究都指出，组织中的张力不会通过一次性安排而被永久消解，而需要在持续变化中维持阶段性平衡（Lewis，2000；Smith和Lewis，2011）。本文在此基础上提出“动态一致性”机制，并将其具体化为“发现偏差—重估主题—调整规则与理性—重新嵌入行动”的循环过程。与现有人工智能领导力研究更多强调领导能力、结构适配或技术赋能不同，本文进一步指出，人工智能情境下真正重要的不只是形成恰当的议题识别、理性结构或规则组合，而是这些安排如何在技术迭代、监管变化、伦理争议和外部评价持续变化的条件下保持相对一致。由此，本文将组织领导力研究从较多关注静态结构的分析进一步推进到演化过程的解释，也使和谐管理理论中“在变动中保持和谐”的思想，在人工智能时代获得了更清晰的机制表达（Lewis，2000；Smith和Lewis，2011）。

总体来看，本文强调组织领导力不再只是围绕目标实现、行为激励或价值倡导展开，而更多体现为一种贯穿议题识别、理性权衡、规则承载与持续调整的组织过程。也正是在这一层面上，本文尝试为人工智能情境下组织领导力研究提供一种更具过程解释力和本土理论色彩的分析视角。

（三）实践价值

从管理实践与政策治理视角看，本文提出的组织领导力机制在中国AI情境下具有以下应用价值。第一，在组织治理与战略管理方面，企业可将主题识别运用于重大技术与战略投资决策，通过构建“技术创新—伦理责任—社会价值”的议题清单与优先序列，避免单一绩效导向导致的战略短视、伦理失范与合法性风险，提升战略决策的前瞻性与稳健性。第二，在组织领导选育与能力建设方面，双重理性机制为领导能力评价提供了新的标尺。组织可以在选拔、培训与考核过程中，将数据素养、战略判断等工具理性与伦理敏感、社会责任等价值理性相结合，促使组织领导者在AI决策中形成自觉的理性权衡，减少陷入“技术决定论”或“价值理想化”的偏差。第三，在制度设计与文化建设方面，依托“和则-谐则”耦合机制，企业能够在 AI 治理中实现制度约束与文化弹性的动态平衡：以算法规范、数据治理和伦理审查等谐则划定底线，同时以创新文化、员工参与与跨部门协同等和则激活组织弹性，从而构建兼具合规性与创新性的制度生态。第四，在组织韧性与公共治理方面，动态一致性机制为构建自适应治理系统提供了路径。企业可通过多级反馈通道、跨部门学习共同体与场景化评估机制，形成“发现问题—调整规则—修正理性—再投入实践”的循环闭环。政策制定者亦可据此构建更具弹性的监管沙盒、算法审

查与伦理评估框架，实现技术创新与社会责任的协同落地。

（四）不足与展望

尽管本文在中国AI情境下的领导力理论创新方面进行了初步探索，但仍存在若干局限，有待后续研究进一步深化。第一，本文主要基于理论整合与逻辑推演，对“AI情境下组织领导力机制链”的论证偏重概念层与解释层，缺乏多情境、大样本的系统实证支持。未来研究可采用问卷调查、多案例比较与纵向追踪等方法，对各机制之间的作用路径及其边界条件进行量化检验。第二，本文虽立足中国情境，但对不同行业、不同所有制与不同发展阶段组织的差异关注不足。随着AI在制造业、平台企业与公共部门等领域的异质化发展，领导机制可能呈现不同模式，未来可开展细分行业和跨组织的比较研究，以提高模型的解释力与适用性。第三，本文主要从组织领导者视角分析技术变革与伦理规范之间的对话机制，对员工、算法开发者、平台用户等其他主体的能动性关注有限。未来可引入多主体治理与共创视角，考察不同主体如何通过参与议题建构、规则协商与反馈机制，共同塑造AI时代的组织领导力格局。

## 主要参考文献


[1]高中华, 张恒. 领导AI符号化与员工数字化创造力: 基于认知-情感个性系统理论[J/OL]. 财经论丛, https://doi.org/10.13762/j.cnki.cjlc.20251205.001, 2025-12-05.

[2]关健, 李文朴, 何国华, 等. 人工智能反馈: 文献述评与研究展望[J]. 外国经济与管理, 2025, 47(3): 83-100.

[3]郭小东. 大模型时代全球人工智能治理的挑战与中国方案[J]. 科学学研究, 2025, 43(4): 694-702.

[4]李鹏飞, 葛京, 席酉民. 和谐管理视角下的领导研究发展初探[J]. 管理学报, 2014, 11(11): 1591-1600.

[5]李鹏飞, 席酉民, 韩巍. 和谐管理理论视角下战略领导力分析[J]. 管理学报, 2013, 10(1): 1-11.

[6]林楠, 席酉民, 刘鹏. 产业互联网平台的动态赋能机制研究——以欧冶云商为例[J]. 外国经济与管理, 2022, 44(9): 135-152.

[7]戚聿东, 朱正浩, 赵志栋. 人工智能时代中国管理学学术体系建构[J]. 管理世界, 2025, 41(7): 172-191.

[8]王文龙, 席酉民, 刘鹏. 复杂适应系统理论下数字化领导力的内涵、构成及涌现机制[J]. 管理学报, 2024, 21(4): 475-483,526.

[9]王亚刚, 席酉民. 和谐管理理论视角下的战略形成过程: 和谐主题的核心作用[J]. 管理科学学报, 2008, 11(3): 1-15.

[10]席酉民, 韩巍, 葛京, 等. 和谐管理理论: 创建可能的新世界[M]. 北京: 机械工业出版社, 2025.

[11]席酉民, 刘鹏, 孔芳, 等. 和谐管理理论: 起源、启示与前景[J]. 管理工程学报, 2013, 27(2): 1-8.

[12]席酉民, 肖宏文, 王洪涛. 和谐管理理论的提出及其原理的新发展[J]. 管理学报, 2005, 2(1): 23-32.

[13]席酉民, 熊畅, 刘鹏. 和谐管理理论及其应用述评[J]. 管理世界, 2020, 36(2): 195-209.

[14]徐鹏, 徐向艺. 人工智能时代企业管理变革的逻辑与分析框架[J]. 管理世界, 2020, 36(1): 122-129.

[15]尹萌, 牛雄鹰. 与AI“共舞”: 系统化视角下的AI-员工协作[J]. 心理科学进展, 2024, 32(1): 162-176.

[16]章凯. 领导力的本源与未来[M]. 北京: 中国人民大学出版社, 2024.

[17]张梦晓, 刘鹏, 席酉民. 产业互联网情境下企业战略转型路径与机制研究——基于和谐管理理论视角[J]. 管理工程学报, 2024, 38(4): 302-320.

[18]张亚莉, 李辽辽, 丁振斌. 组织管理中的人工智能决策: 述评与展望[J]. 外国经济与管理, 2024, 46(10): 18-38.

[19]张志鑫, 郑晓明. 数字领导力: 结构维度和量表开发[J]. 经济管理, 2023, 45(11): 152-168.

[20]Shrestha Y R, Krishna V, von Krogh G. Augmenting organizational decision-making with deep learning algorithms: Principles, promises, and challenges[J]. Journal of Business Research, 2021, 123: 588-603.

[21]Anderson M H, Sun P Y T. Reviewing leadership styles: Overlaps and the need for a new ‘full-range’ theory[J]. International Journal of Management Reviews, 2017, 19(1): 76-96.

[22]Aysolmaz B, Müller R, Meacham D. The public perceptions of algorithmic decision-making systems: Results from a large-scale survey[J]. Telematics and Informatics, 2023, 79: 101954.

[23]Aziz M F, Rajesh J I, Jahan F, et al. AI-powered leadership: A systematic literature review[J]. Journal of Managerial

Psychology, 2025, 40(5): 604-630.
[24]Bacharach S B. Organizational theories: Some criteria for evaluation[J]. The Academy of Management Review, 1989, 14(4): 496-515.
[25]Bass B M. Leadership and performance beyond expectations[M]. New York: Free Press, 1985.
[26]Bevilacqua S, Masárová J, Perotti F A, et al. Enhancing top managers'leadership with artificial intelligence: Insights from a systematic literature review[J]. Review of Managerial Science, 2025, 19(9): 2899-2935.
[27]Breidbach C F. Responsible algorithmic decision-making[J]. Organizational Dynamics, 2024, 53(2): 101031.
[28]Brown M E, Treviño L K. Ethical leadership: A review and future directions[J]. The Leadership Quarterly, 2006, 17(6): 595-616.
[29]Cyert R M, March J G. A behavioral theory of the firm[M]. Englewood Cliffs: Prentice-Hall, 1963.
[30]Denison D R, Hooijberg R, Quinn R E. Paradox and performance: Toward a theory of behavioral complexity in managerial leadership[J]. Organization Science, 1995, 6(5): 524-540.
[31]Edmondson A C, McManus S E. Methodological fit in management field research[J]. Academy of Management Review, 2007, 32(4): 1246-1264.
[32]Fischer T, Sitkin S B. Leadership styles: A comprehensive assessment and way forward[J]. Academy of Management Annals, 2023, 17(1): 331-372.
[33]Floridi L, Cowls J, Beltrametti M, et al. AI4People—an ethical framework for a good AI society: Opportunities, risks, principles, and recommendations[J]. Minds and Machines, 2018, 28(4): 689-707.
[34]Gioia D A, Corley K G, Hamilton A L. Seeking qualitative rigor in inductive research: Notes on the Gioia methodology[J]. Organizational Research Methods, 2013, 16(1): 15-31.
[35]Graen G B, Uhl-Bien M. Relationship-based approach to leadership: Development of leader-member exchange (LMX) theory of leadership over 25 years: Applying a multi-level multi-domain perspective[J]. The Leadership Quarterly, 1995, 6(2): 219-247.
[36]Hossain S, Fernando M, Akter S. Digital leadership: Towards a dynamic managerial capability perspective of artificial intelligence-driven leader capabilities[J]. Journal of Leadership & Organizational Studies, 2025, 32(2): 189-208.
[37]Jorzik P, Yigit A, Kanbach DK, et al. Artificial intelligence-enabled business model innovation: Competencies and roles of top management[J]. IEEE Transactions on Engineering Management, 2024, 71: 7044-7056.
[38]Kellogg K C, Valentine M A, Christin A. Algorithms at work: The new contested terrain of control[J]. Academy of Management Annals, 2020, 14(1): 366-410.
[39]Kögel D, Meile S, Canos-Daros L. Who's steering whom and in what direction? An experiment on AI versus top management decision-making in projects[J]. Project Management Journal, 2026, 57(2): 205-221.
[40]Leavitt K, Barnes C M, Shapiro D L. The role of human managers within algorithmic performance management systems: A process model of employee trust in managers through reflexivity[J]. Academy of Management Review, 2025, 50(4): 745-767.
[41]Lewis M W. Exploring paradox: Toward a more comprehensive guide[J]. The Academy of Management Review, 2000, 25(4): 760-776.
[42]Maak T, Pless N M. Responsible leadership in a stakeholder society–a relational perspective[J]. Journal of Business Ethics, 2006, 66(1): 99-115.
[43]March J G, Olsen J P. Ambiguity and choice in organizations[M]. Bergen: Universitetsforlaget, 1979.
[44]Mikalef P, Conboy K, Lundström J E, et al. Thinking responsibly about responsible AI and 'the dark side' of AI[J]. European Journal of Information Systems, 2022, 31(3): 257-268.
[45]Noponen N, Auvinen T, Sajasalo P. The search for authenticity in artificial-intelligence-enhanced leadership[A]. In: Turcan R V, Reilly J E, Mølbjerg K, et al. The emerald handbook of authentic leadership[M]. Leeds: Emerald Publishing, 2023.
[46]Papagiannidis E, Mikalef P, Conboy K. Responsible artificial intelligence governance: A review and research framework[J]. The Journal of Strategic Information Systems, 2025, 34(2): 101885.
[47]Raisch S, Krakowski S. Artificial intelligence and management: The automation–augmentation paradox[J]. Academy of

Management Review, 2021, 46(1): 192-210.
[48]Senge P M. The fifth discipline: The art & practice of the learning organization[M]. New York: Doubleday, 2006.
[49]Shepherd D A, Suddaby R. Theory building: A review and integration[J]. Journal of Management, 2017, 43(1): 59-86.
[50]Shin D, Park Y J. Role of fairness, accountability, and transparency in algorithmic affordance[J]. Computers in Human Behavior, 2019, 98: 277-284.
[51]Smith W K, Lewis M W. Toward a theory of paradox: A dynamic equilibrium model of organizing[J]. Academy of Management Review, 2011, 36(2): 381-403.
[52]Strohmeier S, Becker M, Scheer-Weller E. Beyond aversion–principles of appropriate algorithmic decision-making in human resource management[J]. Expert Systems With Applications, 2026, 296: 128954.
[53]Suddaby R. Editor's comments: Construct clarity in theories of management and organization[J]. The Academy of Management Review, 2010, 35(3): 346-357.
[54]Tsai C Y, Marshall J D, Choudhury A, et al. Human-robot collaboration: A multilevel and integrated leadership framework[J]. The Leadership Quarterly, 2022, 33(1): 101594.
[55]Tsoukas H, Chia R. On organizational becoming: Rethinking organizational change[J]. Organization Science, 2002, 13(5): 567-582.
[56]Van Quaquebeke N, Gerpott F H. The now, new, and next of digital leadership: How artificial intelligence (AI) will take over and change leadership as we know it[J]. Journal of Leadership & Organizational Studies, 2023, 30(3): 265-275.
[57]Weber M. Economy and society: An outline of interpretive sociology[M]. Berkeley: University of California Press, 1978.
[58]Yukl G A, Gardner W L. Leadership in organizations[M]. 9th ed. Boston: Pearson Education, 2020.
[59]Zhang Y, Waldman D A, Han Y L, et al. Paradoxical leader behaviors in people management: Antecedents and consequences[J]. Academy of Management Journal, 2015, 58(2): 538-566.

# Exploring the Mechanisms of Organizational Leadership in the Era of AI: A Paradox Perspective on Technological Innovation and Ethical Responsibility

Wang Wenlong[1], Kong Jianan[2], Xi Youmin[1,3], Liu Peng[3]
(1. *School of Management*, *Xi'an Jiaotong University*, *Xi'an* 710049, *China*; 2. *Business School*, *Nantong Institute of Technology*, *Nantong* 226002, *China*; 3. *College of Industry-Entrepreneurs and HeXie Academy*, *Xi'an Jiaotong-Liverpool University*, *Suzhou* 215123, *China*)

**Abstract:** As AI technology becomes deeply embedded in organizational operations and decision-making processes, leaders are increasingly faced with new ethical challenges, such as data misuse, algorithmic bias, and unclear accountability. These challenges highlight the urgent need to restructure the corresponding behavioral norms. By systematically reviewing and integrating relevant theories, this paper constructs a leadership framework based on four core mechanisms: Theme Recognition, Dual Rationality, HeXie Coupling, and Dynamic Consistency. First, the Theme Recognition mechanism identifies the value focus that should take precedence when addressing the conflicts between technological innovation and ethical norms. Second, the Dual Rationality mechanism explains how leaders balance tool rationality and value rationality in innovation decisions within the AI context, and addresses the risks of imbalance. Third, HeXie Coupling reveals the synergy and tension between formal systems, technical rules, and ethical norms surrounding the theme and decisions. Fourth, the Dynamic

Consistency mechanism emphasizes how organizations can maintain alignment between technological practices and societal ethical expectations through continuous feedback, learning, and adjustment as AI technology evolves. This paper, in essence, aims to address the paradoxical challenges between technological innovation and ethical responsibility in the AI era, offering new theoretical perspectives and institutional frameworks for reconstructing leadership behavior norms, optimizing AI-driven organizational transformation, and enhancing organizational adaptability, innovation, and long-term sustainability, particularly in the Chinese context.

**Key words:** AI technology; leadership; technological innovation; ethical responsibility

（责任编辑：王雅丽）