引用格式：刘雨霖，肖延高，李代天，等. 数据治理的场景识别与沙箱机制研究：人机交互视角下的理论框架[J]. 技术经济，2026，45(8)：62-75.

Liu Yulin，Xiao Yangao，Li Daitian，et al. Context identification and sandbox mechanisms in data governance：A theoretical framework from a human-computer interaction perspective[J]. Journal of Technology Economics，2026，45(8)：62-75.

数据治理的场景识别与沙箱机制研究：人机交互视角下的理论框架

刘雨霖，肖延高，李代天，肖天瑞，王笑天

（电子科技大学经济与管理学院，成都 611731）

摘 要：随着数字经济的深入推进，数据驱动智能决策与运营日益成为企业管理的常态。企业在萃取数据价值的同时，也面临着数据安全与合规的风险。如何平衡数据的价值实现与安全保障，业已成为企业数据治理的关键。本文基于人机交互理论视角，构建起“价值-安全”数据治理场景矩阵，进而结合“高价值-高风险”数据治理场景，提出了“建模、实验、诊断和解释”四阶段数据治理沙箱机制。研究表明：首先，人机交互理论为数字经济时代的数据治理提供了全新视角和理论方法。数据治理并非先安全合规、后开发利用或先开发利用、后安全合规的线性过程，而是需要实现数据开发利用和安全合规的动态平衡。其次，基于价值密度和安全风险两个维度，可以构建起“价值-安全”数据治理场景矩阵。数据治理的重点和难点在于“高价值-高风险”场景，即数据开发利用价值大但未获有效保护或被非法利用损失也大的场域。最后，数据治理MEDI沙箱机制可以通过人机交互实现规则固化、问题溯源和规则再生，推动沙箱由技术隔离工具升级为可验证、可追溯、可再生的动态数据治理系统。研究成果为企业落实《中华人民共和国数据安全法》第十三条“以数据开发利用促进数据安全”“以数据安全保障数据开发利用”的制度安排，提供了可嵌入、可复用的数据治理机制框架。

关键词：数据治理；场景识别；沙箱机制；人机交互；理论构建；MEDI模型

中图分类号：F273.1 **文献标志码**：A **文章编号**：1002-980X(2026)08-0062-14

DOI：10.12404/j.issn.1002-980X.J26021501

一、引言

随着数字经济的不断涌现并加速向工业经济和农业经济深度渗透，数据已成为驱动企业智能决策、产业转型升级与经济社会发展的核心要素[1-2]。诸多研究表明，数据要素的价值实现与安全保障相辅相成，共同构成数据治理研究的核心议题[3-5]。数据治理的价值取向已由早期以合规管控为中心的单一目标[6]转向大数据下的“实现价值—保障安全—实现公平”多目标决策[7-9]。在此背景下，以数据安全与价值协同为核心的数据治理理念正在催生新的治理范式[10-11]。该理念强调在保障数据安全与隐私保护的前提下，借助隐私计算、联邦学习等技术，实现数据“可计算、可调用而不可泄露”[12-13]。从而在安全合规的框架下促进数据的共享和价值实现[14]。近期的研究逐渐拓展到人机协同的数据治理技术实现[15-16]和多场景应用[17-18]，但尚缺乏数据治理理论逻辑、协同机制和模型构建的理论阐释[19]，从而为完善数据治理制度提供经验证据[20]。

近年来，世界主要经济体纷纷出台和完善数据治理制度政策，以促进数据的利用和产业发展，保障数据价值实现的合规与安全。例如，欧盟于2018年通过《通用数据保护条例》(GDPR)，确立了数据采集和利用的“最小化”原则和“目的限制”要求。为了确立人工智能发展的安全和伦理标准，欧盟在2024年通过的《人

收稿日期：2026-02-15

基金项目：国家社会科学基金重点项目“数据价值系统的运行机理与实现机制研究”(24AGL011)

作者简介：刘雨霖(1999—)，电子科技大学经济与管理学院博士研究生，研究方向：数据治理；(通信作者)肖延高(1970—)，博士，电子科技大学经济与管理学院教授，副院长，研究方向：知识产权管理，数据管理；李代天(1986—)，博士，电子科技大学经济与管理学院副教授，研究方向：技术与创新管理；肖天瑞(1998—)，电子科技大学经济与管理学院博士研究生，研究方向：数据管理；王笑天(2002—)，电子科技大学经济与管理学院硕士研究生，研究方向：数据管理。

工智能法案》（AI Act）中提出了基于风险分级管理的监管模式，要求人工智能系统在高风险应用场景中必须具备可解释性、可审计性与人类监督机制，以确保数据与算法的可信运行。美国在2022年提出《数据隐私与保护法案》（ADPPA），并于2024年以《美国隐私权法案》（APRA）形式继续讨论算法透明与问责。整体监管取向以行业自律和风险导向为主，体现出隐私保护与技术创新并重的宽松政策特征。为了抓住数字经济发展的历史机遇，中国在《中华人民共和国数据安全法》中明确提出国家统筹发展和安全，坚持以数据开发利用和产业发展促进数据安全，以数据安全保障数据开发利用和产业发展，并在《关于构建数据基础制度更好发挥数据要素作用的意见》（以下简称《数据二十条》）中提出构建以安全可信、流通有序和价值共生为导向的数据治理“中国方案”。在上述数据治理理念转化为企业数据治理实践的过程中，企业面临的核心挑战是：如何构建起安全、可控的数据创新应用机制，以实现数据的价值最大化与安全合规的平衡。

本文从人机交互理论[21-22]出发，引入计算机领域网络安全监管的“沙箱”方法[23-24]，构建集建模、实验诊断解释（modeling experiment diagnosis and interpretation，MEDI）于一体的数据治理沙箱模型。该模型以企业为研究对象，旨在通过人机协同和动态反馈机制，实现在保障数据安全的前提下提升决策效率和组织韧性。边际贡献主要体现在以下方面：第一，在理论层面，将人机交互的“共生协同”“认知反馈”特征系统嵌入沙箱机制设计，突破了传统数据治理中“技术管控”或“人类主导”的单向逻辑，构建了“人类设定目标-机器执行计算-结果交互反馈”的闭环治理体系；第二，在模型层面，提出MEDI模型作为企业数据治理的理论框架，揭示其在价值实现与安全保障之间的动态平衡机制，为实现“以数据开发利用促进数据安全”“以数据安全保障数据开发利用”的制度逻辑提供理论支点；第三，在管理实践层面，从人工智能技术视角引入人机交互理论，构建以智能感知、动态诊断与认知解释为核心的企业数据治理机制，为企业智能决策与风险防控一体化提供了可推广的模型原型与理论支撑。本文研究不仅拓展了数据治理与人机交互融合的理论边界，也为企业在复杂环境下实现“数据驱动-智能决策”的转型提供了新的认知框架与治理路径。

二、文献综述

（一）人机交互理论的源与流

人机交互（human-computer interaction，HCI）思想可追溯至美国麻省理工学院心理学和人工智能领域的著名学者Licklider[25]提出的“人机共生”（man-computer symbiosis，MCS）构想。该构想借鉴认知心理学与人因工程学的理论框架，认为随着计算能力的迅速提升，人机关系应由单向控制转向优势互补的“伙伴关系”。人类在目标设定、假设生成与决策判断方面具有认知优势，而计算机擅长高速、复杂的运算，人类与计算机的结合能够形成新的“认知共同体”。上述构想的提出，标志着人与计算机的关系进入了新阶段[26]，奠定了人机交互研究的哲学基础。按通常的理解，人机交互中的“机”是指计算机系统。其实在计算机诞生之前，人与各类工具或机械装置之间的交互已广泛存在。Grudin[27]曾回顾人机交互的发展历程，发现人因工程，即“人与工具的交互”主要聚焦如何优化工具的可操作性及其与人生理特征的匹配，以提高工作效率与安全性。随着智能技术的快速发展，特别是具有“类人”特征的机器人不断涌现，人与机器人交互（human-robot interaction，HRI）的研究逐渐从交叉学科演化为独立的研究领域[28]。人机交互内涵和外延的拓展，反映了人机交互中的“机”由传统的物理系统向数字系统再向具备学习与决策能力的智能系统进化，“人机”经历了从“人与工具”到“人与计算机”再到“人与智能体”的演进，体现了人机交互的技术形态与认知边界双重扩展。需要进一步指出的是，人与智能体的交互和前两类存在本质差异。人与工具的交互遵循工具被动执行，反馈完全可预知的规律。人与计算机的交互基于符号系统和图形界面，仍以人类预设程序为边界，结果的不确定性较小。而人与智能体的交互具有双向适应和认知责任共担特征，智能体可自主生成非预设输出、动态更新参数，甚至质疑人类指令，形成人类设目标、智能体探路径、人与智能体共同修正的闭环。

在人机交互理论发展进程中，Harrison等[29]及Suchman[30]提出的“三范式”框架揭示了人机交互的认知层次演进逻辑：第一范式基于人因学，关注系统物理属性与人的操作关系；第二范式基于认知主义理论，探讨人机信息交互中的认知加工与心理负荷；第三范式以现象学情境为基础，强调社会情境与制度环境对交互行为的影响。这一理论划分不仅奠定了人机交互的基本范式体系，也为后续的多维拓展提供了理论基

石。首先,随着个人计算机的普及,人机交互从思想萌芽逐步走向理论化与科学化。里程碑事件是 Card 等[31]提出的“人类处理器模型”(model human processor,MHP),该模型以认知心理学为基础,将人的感知、记忆与思维过程抽象为信息处理流程,为界面设计与任务分配提供了操作性理论工具。这一阶段的研究强化了人因工程与认知心理学的结合,为人机交互提供了可度量、可模拟的理论框架。其次,网络化协作与群体计算的兴起,使得人机交互研究逐渐超越单一用户与界面的线性交互,开始关注多用户协同与社会情境。其中,Jacob[32]提出的“社会-认知工程”(socio-cognitive engineering,SCE)方法强调交互设计需兼顾个体认知特征与社会互动机制,为复杂任务环境下的人机协作提供了理论支撑。这一阶段的研究标志着人机交互研究由个体认知向社会协同扩展,为后续的人机交互智能化发展及其治理奠定了方法论基础。最后,近年来人工智能和大数据技术不断涌现,推动着人机交互理论研究和实践应用持续变革,人机交互正在经历从单一界面效率提升到混合智能协同的跃迁[33-34]。混合智能系统是一种以“认知—计算—执行”三层协同为核心的人机融合架构,旨在通过人机智能互补实现数据开发利用与安全治理的动态均衡[35-38]。随着多终端与情境感知技术的发展,人机交互进入多模态交互阶段,感知智能技术在多场景下得到广泛应用[39-40]。进一步地,现有研究提出“可解释交互设计”[23]与“动态协商反馈机制”[41],强调系统透明性、算法可解释性与人机共同决策[42],推动人机协作从技术层的“辅助”转向治理层的“共治”。

人机交互理论的演进体现出由个体认知向社会协同、由界面优化向智能共治的转变特征,展现出多学科交叉与多场景驱动的趋势,为理解数智时代的人机关系提供了系统框架,也为数据治理、智能监管等领域的创新提供了理论支撑与实践路径。

(二)数据治理的核心议题:平衡数据的价值实现与安全保障

公共管理领域的学者较早关注数据治理问题[43-44]。相关文献将数据治理视为一种风险控制和约束规范的嵌入式制度安排,主要聚焦于国家安全领域的数据治理[45-46],包括政治安全[47]和公共安全[48]。在政治安全层面,研究发现数据无序流动[49]、网络舆情泛政治化[50]及算法推荐的放大效应[51],显著增加了意识形态渗透、认知操纵与数据主权受损的风险。相较于传统信息安全防护,算法的非中立性、平台传播信息的可见性对议题扩散具有塑造作用,进而影响政治认同与国家治理稳定[52]。相应地,数据治理的重点集中于维护主流意识形态与国家数据主权,通过国家主导的多主体协同机制,将价值导向、算法审查和技术监管嵌入数据使用全过程[53-54]。依托数据分级分类与算法问责制度,实现政治风险的前置识别与制度化约束[50-51]。在公共安全层面,为了应对隐私泄露、算法歧视及系统失灵等风险[55-56],数字平台和算法工具在社会治理和应急管理等领域得到广泛应用,数据治理面临技术依赖加深与风险外溢扩大的双重挑战[57]。亟待构建算法影响评估与跨部门协同共享机制,实现对公共数据风险的全过程管控[46],治理模式应由事后处置转向数据驱动的主动预警和协同防控[58]。

随着数据作为一种新型生产要素在经济社会发展中的价值凸显[59],学者们逐步关注到数据治理的价值议题,即数据治理的另一重要目标在于实现数据价值,通过制度化安排促进数据的开发利用与产业发展,使数据从潜在资源转化为现实生产要素[60-61]。围绕数据治理的价值议题,既有研究主要从价值取向、实现路径、制度安排三个角度展开讨论。一是数据治理的价值取向由风险管控转向数据要素价值实现[62-64]。伴随着数字经济的涌现,数据已由传统管理对象转变为关键生产要素,数据治理的重心亦由风险防范转向激活数据要素潜能和价值创造[65-66]。数据治理有助于支撑数据驱动创新的实现,即促进新产品、新流程和新组织形态的生成,提升全要素生产率,并引导企业资源配置由金融化倾向转向实体经济和创新活动[67-68]。二是数据治理通过全生命周期管理与有序流通机制促进数据开发利用[69-71]。具体而言,数据治理通过构建覆盖数据采集、加工和利用的全生命周期管理闭环,提升数据质量、消除数据孤岛,进而以高价值数据集的有序开放和共享为抓手[72],培育数据交易与服务生态,并在隐私计算等技术支持下降低数据跨主体流通中的信任与安全成本[73]。三是实现数据价值有赖于制度供给对治理机制的系统性支撑[74-76]。研究表明,清晰、稳定且可执行的产权安排有助于数据价值的持续释放,产权制度缺失或产权安全保障不足,无法持续激励多主体参与数据价值创造[77-78]。为此,亟待通过“三权分置”推进数据资产化确权,完善数据入表和会计规则,配套分类分级监管与跨境流动制度,设置专门管理机构并辅之以相应的激励措施[79-80],构建起权责清

晰、激励相容且安全可控的数据治理产权制度环境[81-83]。

综上所述，先前关于数据治理的研究存在两种差异化的进路：一是聚焦国家和社会安全治理，侧重数据安全的风险识别、责任分配和过程控制，关注政治安全和公共安全等高敏感场景中的算法审查、问责机制与跨部门协同；二是聚焦财产安全和数据价值实现，将数据视为生产要素和战略资源，强调通过财产权利安排和全生命周期管理等活动来提升数据的配置效率和可复用性，并促进数据资源对企业创新和产业发展的影响。随着数据在经济、政治、社会和文化发展中的历史性作用日益凸显，数据治理的价值取向和制度逻辑正在发生重大变化，集中体现在2021年9月1日起实施的《中华人民共和国数据安全法》第二章“数据安全与发展”中。该法第十三条明确提出，国家统筹发展和安全，坚持以数据开发利用和产业发展促进数据安全，以数据安全保障数据开发利用和产业发展。这一法律规定不仅将数据治理的核心议题统一到“实现数据价值释放和数据安全保障的动态平衡”上，而且为识别数据治理的场景提供了方向[84]。

三、数据治理的场景识别与实现机制

（一）数据治理的“价值-安全”场景矩阵

组织的数据治理活动离不开业务运营和产业发展场景[85]。自2000年Kenny和Marshall[86]将“场景”一词引入网络营销管理以来，场景概念逐渐扩展到企业管理的诸多领域。一般认为，场景（Context）概念源于拉丁语“Contextus”，词源意义是“编织在一起产生的结构、连贯性、联系”。在管理学意义上，场景概念既不同于描述特定时间点状态的处境（Situation）概念，也不同于特指外部客观总体条件的环境（Environment）概念，并有别于描述或推演未来可能发生情况的情境（Scenario）概念。场景是指特定时间的特殊复杂性场域（field），涵盖时间、空间、过程和文化情感维度，是时间、问题、主体、社群、要素、事件汇聚与发生关系及相互作用的场域，包括物理空间和社会空间，也包括赛博空间[87-89]。场景驱动的创新强调的是以场景为载体，以战略为引领，驱动技术、数据等要素有机协同和多元化应用[90]。构建数据要素价值化生态系统，需要数据要素各主体通过联动作用，结合数据使用的真实场景，对价值低密度数据进行高效治理、价值增值和价值沉淀并形成价值高密度数据产品，最终释放数据价值。可见，场景驱动的数据治理范式超越了传统的线性技术驱动逻辑，转而通过识别具体场景中的复杂综合性需求，推动数据要素在“收—存—治—易—用—管”的全生命周期中实现高效配置[91-92]。

然而，不同场景下数据主体对数据的价值认知和诉求存在显著差异，在数据的可获得性上也面临着不同的约束条件[93-94]。因此，围绕数据治理的核心议题开展机制设计，有必要厘清数据治理的“价值-安全”双重维度下的场景[95-96]。依据《中华人民共和国数据安全法》和《中华人民共和国个人信息保护法》，数据可从所有权角度分为公共数据、企业数据和个人数据；数据还可按种类分结构化数据和非结构化数据；按重要程度分为核心数据、重要数据和一般数据。本文从“价值”与“安全”两个核心维度界定数据治理场景，这一划分已内在地涵盖了数据的重要性和风险等级。从这两个维度出发的场景界定方法，适用于任何类型的数据治理场景。就数据治理的“价值”维度而言，按照价值哲学的通常理解，价值是事物对于人（更确切地说，是客体对于主体）的“意义”，代表着客体主体化过程的性质和程度，即客体的存在、属性和合乎规律的变化与主体尺度相一致、相符合或相接近的性质和程度[97]。为了分析的可行性，本文将数据治理场景的“价值”维度简化为价值密度“高”和“低”两种情形，用以刻画数据开发利用给数据主体带来的现实或潜在收益大小。就数据治理的“安全”维度而言，中共中央、国务院发布的《关于构建数据基础制度更好发挥数据要素作用的意见》提出了公共数据“可用不可见”要求。这里的“不可见”不宜作狭义理解，即视觉上或技术上的不可见，而应当理解为从公开渠道不可获得，以便将数据的安全要求从公共数据拓展到企业数据或个人数据。数据主体出于政治安全、公共安全、财产安全、个人隐私、合规利用等安全风险考量，通过技术和制度安排控制数据接触权限，未经法定程序或权利人许可，他人不得接触其掌握的数据。依据中国《中华人民共和国数据安全法》第三条第三款对数据安全的定义，可以将数据治理场景的“安全”维度简化为安全风险“高”和“低”两种情形，用以刻画数据未能获得有效保护或被非法利用给数据主体带来的现实或潜在损失大小。

基于上述分析，构建“价值-安全”双维度数据治理场景矩阵，从价值密度（高/低）和安全风险（高/低）两个维度划分出四种数据治理场景（图1）。进而从价值密度、安全风险、治理目标和场景列举四个方面给出不



图1　“价值-安全”双维度数据治理场景矩阵

同数据治理场景画像，为后续数据治理机制设计和制度选择提供可操作分析框架。

（1）“低价值-低风险”数据治理场景。该场景是指数据的价值低，且从公开渠道或合规边界内可获取或可访问，难以直接转化为支撑业务洞察、决策改进与绩效提升的有效信息资源[98-99]。从价值密度看，该类数据多以未经清洗加工的原始形态沉积，质量波动和标准缺失使其呈现“可得但难用”，现实增值空间有限[100]。就安全风险而言，该场景下的数据可从公开渠道获取或访问，风险整体较低或可通过常规管理控制，即使发生不当使用，造成的损失也相对有限[101]。治理目标宜以“剥离并适时清除”为导向，基于数据盘点和分类分级识别冗余与噪声数据，优先实施去重整合、归档压缩和合规销毁，降低沉积成本并抑制“数据沼泽”化[102]。典型场景表现为浏览器缓存记录、过期系统日志、非审计必需的历史监控指标明细等[103]。

（2）“低价值-高风险”数据治理场景。该场景是指数据的价值低，且从公开渠道或合规边界内不可获得或不可访问，难以被有效整合与调用[104-105]。从价值密度看，该场景下的数据表现为低频使用的历史遗留沉没数据，开发利用的收益有限，难以形成可以用货币计量的数据资产。就安全风险而言，该类数据对外部主体不可获得主要源于系统割裂与访问链路中断形成的结构性封闭，其“高风险”更多体现为遗留系统不可控带来的潜在合规暴露、恢复泄露和安全隐患，一旦发生泄露或被非法利用，潜在损失较大[99]。治理目标宜以“清理、封存或销毁”为导向，通过退出机制清理不可获得的无价值存量数据，压降沉积成本并控制潜在风险暴露[106]。典型场景表现为老旧信息系统中不再维护的历史备份、组织调整或系统迁移后滞留于停用模块且难以追溯的过期业务台账，以及涉及恶意代码传播、钓鱼欺诈或恶意跳转的有害网站（俗称“有毒网站”）等失管存量对象[103]。

（3）“高价值-低风险”数据治理场景。该场景是指数据的开发利用价值大，且从公开渠道或合规边界内可获得或可访问，能以较低的获取门槛和合规成本实现跨主体复用，从而形成相对稳定的数据供给与价值扩散基础[107-108]。从价值密度看，该类数据价值密度高，数据开发利用的收益大。该场景下的数据具有提升生产效率的经济价值，乃至实现数据公平的社会价值或保障数据安全的政治价值[109]。就安全风险而言，该场景下的数据多为公开披露或已完成必要脱敏处理，敏感性相对较低，即便发生不当使用，其潜在损失通常有限且可控[72-110]。治理目标宜在充分释放数据价值的同时保障数据的产权安全、政治安全或公共安全，通过“开放-共享”的制度化供给促进数据的复用和扩散，并以边界控制和责任机制实现安全可控。典型场景表现为可持续复用的公开数据供给，如世界知识产权组织国际专利检索系统数据集、上市公司定期报告、工

商登记与统计公报等公开披露资源[72]。

（4）“高价值-高风险”数据治理场景。该场景是指数据的开发利用价值大，但难以在“明文交付-直接共享”模式下实现数据获取与流通[102,111-112]。从价值密度看，该类数据价值密度高，数据开发利用的收益大，能够产生现实或潜在的经济价值及社会福利。就安全风险而言，受隐私、敏感性、商业秘密或国家安全等约束，原始数据通常被严格限制直接接触，一旦保护不力或被非法利用，将造成严重的经济损失或政治及社会风险后果，因此安全要求显著高于常规数据[113-114]。治理目标宜在保障数据产权安全、政治安全与社会安全的前提下释放数据的价值，通过授权、审计和问责闭环降低安全风险外溢[115]。典型场景包括医疗健康电子病历和影像记录、平台用户行为轨迹和身份标识信息、税务海关等监管申报明细，以及供应链协同的核心工艺参数和客户清单等商业秘密[116-118]。

（二）不同场景的数据治理实现机制

以上分析表明，不同数据治理场景的数据价值密度、安全风险、治理目标和治理机制存在显著差异。相对而言，前三类数据治理场景的治理目标清晰，治理机制相对成熟。①“低价值-低风险”数据治理场景以数据减量和全生命周期治理为主，通常由组织内部的数据管理部门或数据资产管理团队牵头，业务部门、其他职能部门配合确认和审计数据的价值，安全合规部门提供风险底线约束。通过数据盘点编目和分类分级，将数据按“剥离、归档、删除”进行分流，以质量阈值、口径一致性和去重整合作为清退规则，并将归档压缩与合规销毁嵌入数据运维流程，从而抑制数据沉积与“数据沼泽”化[100-102]。例如，平台型企业通过数据生命周期管理系统，对低复用、过期和冗余数据进行自动识别、归档与清理，以降低存储成本并抑制数据沉积[52,96]。②“低价值-高风险”数据治理场景以存量清退和封存机制为主，其安全风险更多源于遗留系统造成的数据割裂、流程断点和责任缺位，而非数据内容本身的高敏感属性。该场景宜由数据管理部门牵头统筹数据存量治理，安全合规部门协同明确风险边界、处置标准和问责要求，必要时引入审计监督以强化可追责性。通过存量盘点厘清数据归属和管理责任，形成封存、清退和销毁的可执行规则，并将留痕审计和访问隔离嵌入流程控制。对无现实或潜在经济价值且有效保护成本大的数据实施清退或销毁[104-106]。例如，企业在系统迁移过程中，对历史数据进行归档，并退役遗留系统，以减少数据库规模、迁移成本和遗留系统风险[119]。③“高价值-低风险”数据治理场景以制度化供给和开放复用为主，在遵守合规边界和授权规则前提下，数据资产或数据产品供给主体通过双边或平台化交易方式开展数据交易，形成数据市场价格信号。具体包括通过规则目录和标准体系将数据标准化，进而形成可持续供给的数据产品，并以授权规则、用途约束、责任分担与纠偏机制界定数据使用边界，降低数据误用偏差，从而在安全风险可控条件下形成稳定的数据供给和价值释放机制[107-110]。例如，制造企业通过企业级数据平台，将跨业务域数据转化为可发现、可访问和可复用的数据资产，以支撑决策优化和人工智能应用扩展[120-121]。

在“高价值-高风险”数据治理场景中，数据利用同时面临高增值潜力和高合规约束，传统“开放-共享”或“封存-退出”机制难以满足数据治理的双重需要。数据治理机制须从交付数据本体转向受控条件下的可验证调用，并在数据经营者、提供者和治理者等主体[109]之间形成可执行的数据授权、利用、审计与问责闭环。现有理论研究和管理实践表明，既能释放数据经济价值又能控制数据安全风险的沙箱是该类场景的可行治理机制[24]。沙箱又名沙盒，早期主要用于计算机安全的系统防控，是在受控环境中对计算机程序或数据行为进行隔离和验证的技术工具，用以防止计算机系统性风险和信息泄露[122-123]。在金融监管语境下，该技术进一步演化为监管沙箱，为创新产品和服务提供可控试验空间，以实现创新和监管的动态平衡[124-125]。随着数据成为新型生产要素，有学者研究表明，沙箱机制可以从计算机安全防护扩展为数据治理工具，即通过构建边界清晰、过程受控且监管可介入的可信数据空间，使多方在不直接接触原始数据的前提下完成协同分析和价值挖掘，从而为数据安全合规和价值释放之间的张力提供可操作的实现机制[126-128]。在数据要素流通需求持续上升的背景下，沙箱机制的运行逻辑通常呈现限定空间、受控试验和动态反馈的闭环特征，为跨组织协作提供可信的过程条件，并在一定程度上提升过程透明度和可解释性[24,129]。然而，现有研究更多聚焦技术能否安全运行，对治理主体如何理解与干预模型行为、如何在试验过程中实现责任配置和风险分担，以及如何将制度约束嵌入反馈循环的讨论不足，难以支撑“高价值-高风险”场景下数据治理机制的认

知协同、反馈学习和规则再生[130]。也就是说,如果沙箱机制在组织层面仅停留在技术验证阶段,难以成为“高价值-高风险”场景下的可行数据治理机制,难以充分回应企业级数据治理对可验证、可追溯和可演化的综合要求[131]。例如,欧盟创新健康倡议 MELLODDY(machine learning ledger orchestration for drug discovery)联合项目在不共享原始专有实验数据的前提下,通过联邦学习开展药物发现模型训练,实现高价值数据的受控调用与安全协同[132]。为此,有必要构建以人机交互为认知基础、以沙箱为机制载体、以制度约束为行为边界的数据治理框架,通过认知、机制和制度的三元耦合推动数据治理走向组织化运行,并更好契合以数据开发利用促进数据安全、以数据安全保障数据开发利用的双向循环。

四、数据治理 MEDI 沙箱模型的构建

(一)数据治理 MEDI 沙箱模型的提出

《中华人民共和国数据安全法》第十三条“国家统筹发展和安全,坚持以数据开发利用和产业发展促进数据安全,以数据安全保障数据开发利用和产业发展”的规定,确立了中国数据治理的制度逻辑,即以数据价值释放促进数据安全,以数据安全保障数据价值释放。这一逻辑突破了传统治理的静态平衡观,强调以制度为纽带,驱动安全技术迭代与数据流通创新的双向反馈——数据的开发利用推动安全技术和制度的迭代演化,而安全保障又通过规则嵌入、风险分级和流程控制反向促进数据资源的持续流通和创新。从国家层面的数据治理范式来看,这实质上是一种制度驱动的“价值实现-安全保障”协同能力生成机制。

然而,从宏观制度向微观企业治理实践转化的过程中,仍存在显著的机制缺口。现有研究多聚焦宏观监管、法律制度和政策导向,较少提供能够嵌入组织流程,同时支撑价值实现与安全保障双重目标的微观治理机制。尤其在“高价值-高风险”的数据治理场景中,数据具有明确价值,却因隐私、商业秘密或合规边界限制而难以在传统明文交付/直接共享模式下直接获取和流通,企业由此面临必须用、但不可交付(不可明文交付原始数据)的结构性矛盾。一方面,需要跨主体协同以形成可验证的价值产出;另一方面,又必须严格控制原始数据的可获得性和内容暴露。人机交互理论为弥补该缺口提供了潜在路径,其强调在目标设定—算法执行—反馈优化的循环中实现认知协同,使治理从单向控制走向动态互动。但若仅停留在认知或技术层面,上述互动难以同时满足企业层面数据治理的价值实现可验证和安全责任可追溯的要求。因此需要一个兼具价值验证和安全隔离功能的中介载体,在不扩大暴露面的前提下提供可控验证空间,使理论逻辑、制度逻辑和机制逻辑得以贯通。

在此背景下,提出 MEDI 数据治理沙箱模型。该模型属于机制型理论框架,其目标不在于提出静态变量关系,而在于以四阶段的连贯机制揭示企业在安全约束下实现数据价值的运行逻辑,从而回应“高价值-高风险”场景下的数据治理需求。具体而言,该模型以人机交互的认知闭环为理论基石,将《中华人民共和国数据安全法》的双向循环原则内化为制度约束,以受控沙箱为机制载体,面向“高价值-高风险”场景构建结果可用、过程可控的受控验证空间。当前主流数据平台在数据隔离和受控计算上采用了不同的技术路线。例如,以 DeepSeek 为代表的生成式 AI 平台侧重于交互式数据分析与输出审核;联邦学习平台强调数据本地留存下的模型协同训练,可信执行环境(TEE)依托物理层面的安全隔离;多方安全计算技术则允许多个参与方在不泄露各自原始数据的前提下开展协同计算,实现数据的安全隔离。这些技术在安全性、计算效率、可解释性和治理嵌入深度上各有优劣。本文的 MEDI 模型不绑定特定技术,而是提供统一的治理编排框架。MEDI 模型构建了一个可验证、可迭代、可解释的数据治理实验体系,为企业落实国家统筹发展和安全的数据治理顶层设计提供一个可操作、可迭代的机制框架。

(二)“高价值-高风险”场景下数据治理 MEDI 沙箱的运行机制

从图 2 可以看出,数据治理 MEDI 沙箱模型将企业数据治理机制转化为“建模—实验—诊断—解释”的闭环链条,建模阶段明确治理目标和规则边界,实验阶段在不扩大暴露面的前提下开展受控计算并生成过程证据,诊断阶段基于证据识别偏差和风险成因,解释阶段将结果转化为可审计的决策依据并反馈更新规则和参数。数据治理 MEDI 沙箱模型在受控验证空间内实现过程可审计、风险可控制、规则可迭代,使价值验证和安全控制在同一流程中协同发生,从而推动企业数据治理由一次性试点或静态防控转向动态协同。

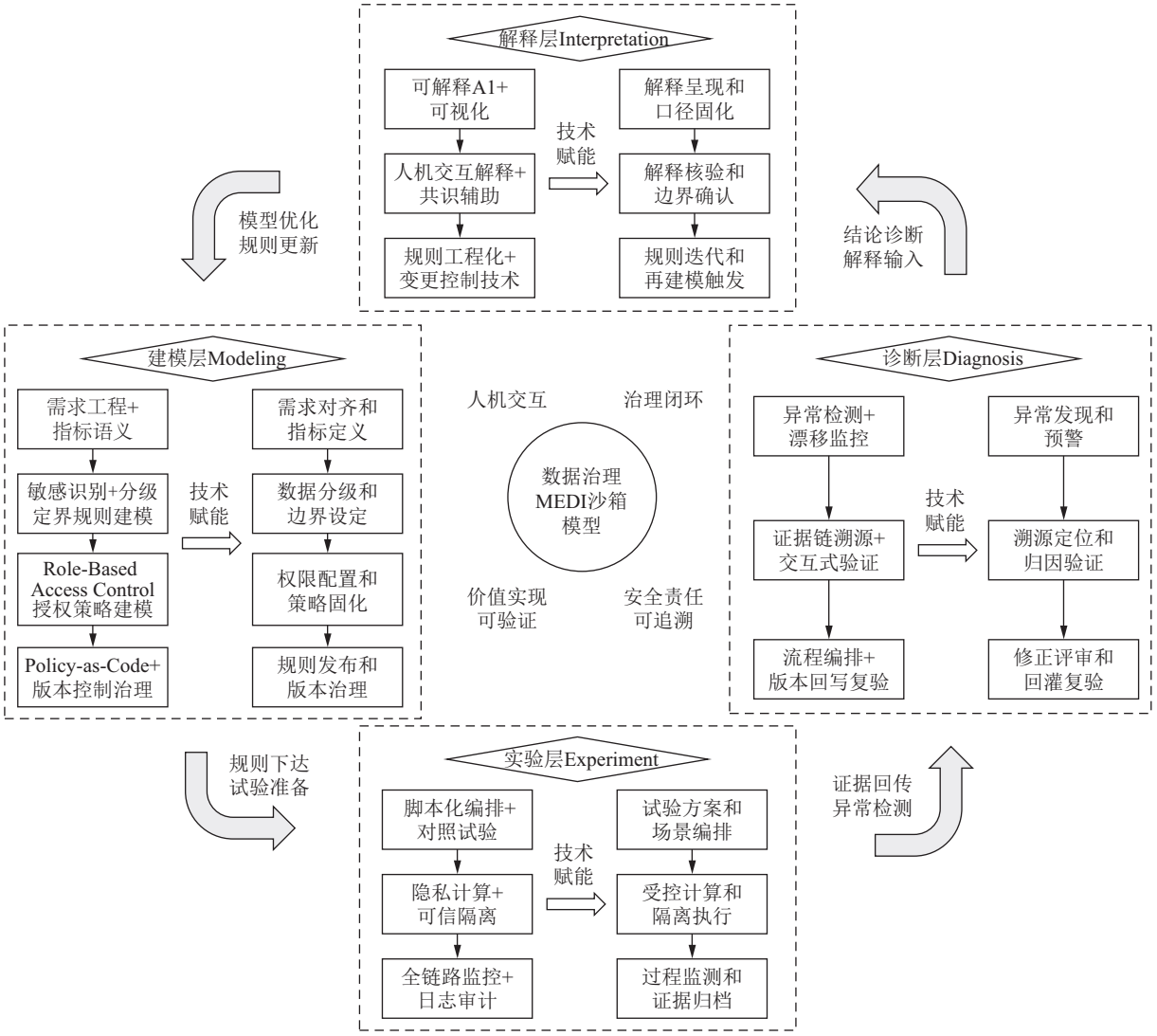


图 2　人机交互视角下数据治理 MEDI 沙箱机制

1. 建模阶段（modeling）：目标定义与规则设定

建模阶段是 MEDI 数据治理沙箱机制的起点，旨在将业务目标和安全约束转化为可执行、可审计、可迭代的规则体系与边界配置。本阶段以需求协同和指标定义为起点，通过需求工程和指标语义建模将业务问题、评价指标及其口径规则结构化，形成统一的口径基线。继而开展数据分级和边界设定，运用敏感识别和分级规则建模完成数据的盘点、分类分级和边界定界，明确数据的可用范围、禁用字段和隔离要求。进一步推进至权限配置和策略固化，以基于角色的访问控制（role-based access control，RBAC）授权策略将主体身份、访问条件、使用目的和输出粒度绑定，实现受控用途下的受控输出。最后，通过规则发布和版本治理，依托策略即代码技术和版本控制技术将口径、分级和策略工程化为可发布规则包并纳入变更、审批与回滚。由此形成口径清晰化—边界参数化—权限用途策略化—规则版本工程化的链条，为后续实验和诊断提供前置条件，体现《中华人民共和国数据安全法》第十三条所强调的以数据开发利用促进数据安全。

2. 实验阶段（experiment）：算法执行与证据归档

实验阶段是连接规则设计和风险识别的关键环节，旨在受控沙箱环境中运行建模成果。首先，制定试验方案并编排场景，通过场景脚本和对照设计，把需要验证的假设、怎么扰动、用什么指标评估、达到什么结果写成一套可重复执行的试验规程，明确验证目标和判定规则。其次，进入受控计算和隔离执行，在既定权限、用途和输出约束下，调用隐私计算和可信隔离完成跨主体、跨系统计算，并将输出限定为规则允许的结

果形态。最后,以过程监测和证据归档形成可核验闭环,通过全链路监控和日志审计记录规则版本、数据范围、关键操作与异常告警,固化可追溯证据链,确保结果可解释、可核验、可追责。该阶段意在在保障数据安全的条件下实现数据价值释放的创新,将数据安全保障与数据开发利用嵌入实验化数据治理流程。

3. 诊断阶段(diagnosis):异常溯源与闭环修正

诊断阶段是将实验输出转化为风险认知和规则改进的关键环节,主要实现异常识别、原因定位和闭环修正功能。首先,进行异常发现和预警,基于实验结果和过程日志开展异常检测和漂移监控,对数据突变、指标波动和分布漂移等信号进行识别与分级。其次,进入溯源定位和归因验证,借助证据链溯源和交互式验证,将异常还原到数据源、规则版本、执行路径或策略条款,并通过对照复核、复跑验证确认根因,避免技术解释和业务理解错位。最后,以修正评审和回灌复验完成闭合,通过流程编排执行任务分派、审批、留痕,并以版本回写固化规则,进而实现策略更新,再经过回归复验,确认异常消退和控制生效,形成"发现—定位—修正—验证"的模型增强闭环。由此使得数据开发利用能在安全、透明、可追溯的框架下持续运行,从而承载起数据安全保障数据开发利用的制度要求。

4. 解释阶段(interpretation):可解释输出与制度迭代

解释阶段旨在实现模型输出的可理解化、可确认化和制度沉淀。首先,完成解释呈现和口径固化,通过可解释人工智能和可视化技术呈现指标报告、变量贡献、规则路径和因果链条,并固化组织统一的解释口径。其次,进入解释核验和边界确认,借助人机交互解释和共识辅助,对结果可用范围、用途边界和输出粒度进行核验并形成发布策略和责任留痕,使数智技术结果转化为组织可用的数据治理策略集。最后,以规则迭代和再建模,实现数据治理的制度化闭环;通过规则工程化技术和版本变更控制技术,沉淀数据治理的规则口径、适用边界和处置经验;触发下一轮规则更新与版本演进,形成"解释—确认—沉淀—迭代"的持续学习机制。该阶段强化规则透明性和责任可溯性,使人机交互行为可追问、可修正,为数据安全合规和数据开发利用提供制度支撑。

在上述数据治理MEDI沙箱机制中,建模、实验、诊断与解释四个阶段并非简单的线性推进,而是立足人机交互理论方法和"高价值-高风险"数据治理场景,将输入、验证、校正、反馈行为耦合为可持续演化的数据治理闭环,使"以数据开发利用促进数据安全""以数据安全保障数据开发利用"的制度能够在企业层面形成可迭代的数据治理机制,确保数据的价值实现可验证、安全责任可追溯。①建模阶段输出结构化的需求口径、分级边界、权限用途策略和版本化规则包,直接构成实验阶段受控计算和试验编排的运行输入。该阶段由人类主导目标和安全约束的规则设计,智能体仅作为执行层不参与目标设定,回应了人与智能体交互中"人类保有最终认知权威"的原则。②实验阶段形成的结果产出、过程日志和异常告警沉淀为证据链,为诊断阶段的异常识别、溯源归因和闭环修正提供事实基础。智能体在沙箱隔离环境中执行计算,其非预设行为(如异常路径探索)被完整日志记录,将智能体的自主性转化为可观测的过程证据。③诊断阶段在证据链支撑下完成根因定位和修正评审,并将修正规则和版本回写方案作为解释阶段核验边界、形成发布策略和沉淀治理经验的依据。该阶段采用"系统归因建议+人类业务复核"的双层模式,形成人机协同的责任分配格局。④解释阶段则通过可解释呈现与共识核验将模型逻辑、用途边界和处置结论显性化,进一步将确认后的规则资产与变更要求回注到建模阶段,触发新一轮目标设定和规则配置。借助可解释人工智能技术将智能体的内部决策逻辑映射为人类可理解的因果链,弥补"黑箱"带来的认知不对称。可见,数据治理MEDI沙箱机制能够以"建模—实验—诊断—解释—再建模"的方式实现连续迭代,形成从局部试验到整体演化的治理闭环,使模型边界、规则体系和安全策略在反馈强化中不断更新。MEDI模型并非将智能体简单视为高级工具,而是视其为具有有限自主性的协作伙伴,通过四阶段的结构化交互流程,在释放智能体效率优势的同时确保人类对关键决策的可控与可追责,最终构建起可验证、可追溯、可再生的动态数据治理系统。

五、研究结论与展望

本文从人机交互理论视角出发,聚焦数据治理的核心议题,即统筹发展和安全,实现数据价值释放和数据安全保障的平衡,提出以场景为牵引、以沙箱为载体、以闭环迭代为运行逻辑的数据治理机制框架。研究

表明：第一，人机交互的“机”经历了机器、计算机和智能体等不同发展阶段，人机交互理论为数字经济时代的数据治理提供了全新视角和理论方法；数据治理并非先安全合规、后开发利用或先开发利用、后安全合规的线性路径，而是需要在制度安排、技术嵌入和流程控制的协同作用下形成数据开发利用和安全合规的动态平衡机制。第二，基于价值密度和安全风险两个维度，可以构建起“价值-安全”数据治理场景矩阵。数据治理的重点和难点在于“高价值-高风险”场景，即数据开发利用价值大但未获有效保护或被非法利用损失也大的场域。第三，从人机交互理论视角，立足“高价值-高风险”数据治理场景，提出了数据治理 MEDI 沙箱机制框架，构建起建模、实验、诊断和解释四阶段数据治理模型。该模型以证据链和可追溯过程为支撑，通过人机交互实现规则固化、问题溯源和规则再生，推动沙箱由技术隔离工具升级为可验证、可追溯、可再生的动态数据治理系统。上述研究成果为企业落实《中华人民共和国数据安全法》第十三条“以数据开发利用促进数据安全”“以数据安全保障数据开发利用”的制度安排，提供了可嵌入、可复用的数据治理机制框架。

为了深入推进场景驱动的数据治理理论探索和机制设计，有必要从以下方面进一步深化和拓展数据治理沙箱机制的理论方法研究及实践验证。

第一，重点解析数据治理的核心议题和理论源流，凝练数字经济时代人工智能等数智技术支撑下的数据治理概念范畴和问题谱系。一是借鉴形式逻辑关于概念特有属性的界定方法，探讨数据治理的内涵和外延，明确数据治理的对象、范围和目标等。二是结合人机场景和治理理论，围绕“价值实现-安全保障”这一数据治理核心议题，提炼出人机交互场景下的数据治理主要命题，构建起场景驱动的数据治理理论演绎和分析框架。三是探索人机场景下的数据治理研究方法，重点引入人工智能推理、多模态学习方法，形成可复用的数据治理研究工具箱。

第二，细致刻画“价值-安全”双维度下的数据治理场景，清晰界定不同的数据治理场景形态及其治理目标。一是结合不同数据价值形态的市场表现细致刻画数据的价值密度，将数据的价值密度区分出由低到高的不同层级，明确不同价值密度的数据治理价值取向。二是根据不同场景下的数据安全风险细分数据的安全强度，将数据的安全风险刻画为由低到高的不同等级，为数据治理手段和方法选择提供制度依据。三是据此构建起颗粒度更细的“分层价值和分级安全”数据治理场景，给出不同细分场景下的数据治理目标和方法。四是深入探讨数据价值的动态变迁规律，重点关注低价值密度数据在长时间跨度积累或多源关联融合中向高价值资产转化的机制，从而在场景识别中嵌入价值演化的动态视角。

第三，差异化构建“高价值-高风险”场景下的可操作数据治理沙箱机制，为“数据开发利用促进数据安全”“数据安全保障数据开发利用”的数据治理制度安排提供经验证据。一是结合上述“分层价值和分级安全”数据治理场景，探索数据治理沙箱机制的关键构件和运行规则，形成可落地的数据治理机制方案。二是面向不同的“高价值-高风险”场景，差异化构建数据治理沙箱机制的运行效应评价指标体系，实证检验沙箱机制的边际收益和治理成效。三是运用人工智能推理、多模态学习等方法，系统验证“沙箱模型—过程控制—治理效应”可行性和可靠性。

第四，立足典型的数据开发利用和产业发展场景，提炼能够平衡数据开发利用和数据安全的企业数据治理行为规律，有针对性给出企业数据治理行为启示和制度优化建议。一是聚焦智能制造、金融服务、医疗服务、数据产业等高价值、高敏感的数据治理场景，开展多案例比较和过程追踪，识别出企业数据治理行为的影响因素和实现机制。二是围绕企业数据治理的全流程实践，提炼可复用的企业数据治理行为模式和实现路径。三是在上述理论研究和经验证据的基础上，给出场景驱动的企业数据治理启示和制度供给诉求，推动构建数据治理中国方案的形成和发展。

参考文献

[1] 谷炜，刘亚金，卢凤·苏珊，等. 人工智能驱动管理决策：应用、感知与偏见[J]. 中国管理科学，2025，33(5)：99-112.

[2] 赵涛，张智，梁上坤. 数字经济、创业活跃度与高质量发展——来自中国城市的经验证据[J]. 管理世界，2020，36(10)：65-76.

[3] BUCZAK A，GUVEN E. A survey of data mining and machine learning methods for cyber security intrusion detection[J]. IEEE Communications Surveys & Tutorials，2017，18(2)：1153-1176.

[4] CAPPA F，ORIANI R，PERUPPO E，et al. Big data for creating and capturing value in the digitalized environment：Unpacking the effects of volume，variety and veracity on firm performance[J]. Journal of Product Innovation Management，2020，38(1)：49-67.

[5] 唐要家, 唐春晖. 数据价值释放的理论逻辑、实现路径与治理体系[J]. 长白学刊, 2022(1): 98-106.
[6] 许多奇. 论跨境数据流动规制企业双向合规的法治保障[J]. 东方法学, 2020(2): 185-197.
[7] 张灵, 冯科, 孙华平. 制造业企业数据价值释放: 效应与机制[J]. 系统工程理论与实践, 2024, 44(1): 68-85.
[8] 宗绍昊, 罗世龙. DeepSeek 类生成式人工智能的新型数据安全风险治理[J]. 科学学研究, 2025, 44(4): 690-700.
[9] 宋华, 韩思齐, 刘文诣. 数字技术如何构建供应链金融网络信任关系?[J]. 管理世界, 2022, 38(3): 182-200.
[10] 范渊.《中华人民共和国数据安全法》解读: 如何实现数据"可用不可见"的安全目标?[J]. 中国经济周刊, 2021(11): 100-101.
[11] 龚强, 班铭媛, 刘冲. 数据交易之悖论与突破: 不完全契约视角[J]. 经济研究, 2022, 57(7): 172-188.
[12] 韩文蕾, 张军辉. 隐私计算环境下的公共数据授权运营法治进路研究[J]. 法治与社会, 2024(7): 57-59.
[13] HERNÁN M A, ROBINS J M. Using big data to emulate a target trial when a randomized trial is not available[J]. American Journal of Epidemiology, 2016, 183(8): 758-764.
[14] 蔡家玮, 欧阳桃花, 曾德麟, 等. 技术可供性视角下数据要素价值实现机制研究——以联通数科为例[J]. 科技进步与对策, 2025, 42(20): 11-21.
[15] JHAVER S, BIRMAN I, GILBERT E, et al. Human-machine collaboration for content regulation: The case of reddit automoderator[J]. ACM Transactions on Computer-Human Interaction, 2019, 26(5): 1-35.
[16] 谢小云, 左玉涵, 胡琼晶. 数字化时代的人力资源管理: 基于人与技术交互的视角[J]. 管理世界, 2021, 37(1): 200-216, 13.
[17] 黄阳华. 基于多场景的数字经济微观理论及其应用[J]. 中国社会科学, 2023(2): 4-24, 204.
[18] JIA W, WANG W, ZHANG Z. From simple digital twin to complex digital twin part ii: Multi-scenario applications of digital twin shop floor[J]. Advanced Engineering Informatics, 2023, 56: 101915.
[19] 张志学, 华中生, 谢小云. 数智时代人机协同的研究现状与未来方向[J]. 管理工程学报, 2024, 38(1): 1-13.
[20] 肖延高, 汤艻东, 李代天, 等. 数据资源的产权逻辑: 基于文献计量的研究进展[J]. 技术经济, 2025, 44(2): 17-30.
[21] 张维, 曾大军, 李一军, 等. 混合智能管理系统理论与方法研究[J]. 管理科学学报, 2021, 24(8): 10-17.
[22] LOPATOVSKA I, ARAPAKIS I. Theories, methods and current research on emotions in library and information science, information retrieval and human-computer interaction[J]. Information Processing & Management, 2011, 47(4): 575-592.
[23] 孔祥维, 王子明, 王明征, 等. 人工智能使能系统的可信决策: 进展与挑战[J]. 管理工程学报, 2022, 36(6): 1-14.
[24] ALAASSAR A, MENTION A L, AAS T H. Exploring a new incubation model for fintechs: Regulatory sandboxes[J]. Technovation, 2021, 103: 102237.
[25] LICKLIDER J C R. Man-computer symbiosis[J]. IRE Transactions on Human Factors in Electronics, 1960(1): 4-11.
[26] 董士海. 人机交互的进展及面临的挑战[J]. 计算机辅助设计与图形学学报, 2004(1): 1-13.
[27] GRUDIN J. From tool to partner: The evolution of human-computer interaction[M]. San Rafael: Morgan & Claypool Publishers, 2017.
[28] GOODRICH M A, SCHULTZ A C. Human-robot interaction: A survey[J]. Foundations and Trends in Human-computer Interaction, 2008, 1(3): 203-275.
[29] HARRISON S, SENGERS P, TATAR D. Making epistemological trouble: Third-paradigm hci as successor science[J]. Interacting with Computers, 2011, 23(5): 385-392.
[30] SUCHMAN L A. Plans and situated actions: The problem of human-machine communication[M]. Cambridge: Cambridge University Press, 1987.
[31] CARD S, MORAN T, NEWELL A. The model human processor—An engineering model of human performance[J]. Handbook of perception and human performance. 1986, 2(45-1): 1-35.
[32] JACOB R J K. The use of eye movements in human-computer interaction techniques: What you look at is what you get[J]. ACM Transactions on Information Systems, 1991, 9(2): 152-169.
[33] 王刚, 黄丽华, 张成洪. 混合智能系统研究综述[J]. 系统工程学报, 2010, 25(4): 569-578.
[34] 张瑞, 司孟韩. 生成式 AI 支持下社会科学知识生产的共同体德性[J]. 科学学研究, 2026, 44(4): 746-756.
[35] BAHRAMMIRZAEE A. A comparative survey of artificial intelligence applications in finance: Artificial neural networks, expert system and hybrid intelligent systems[J]. Neural Computing and Applications, 2010, 19(8): 1165-1195.
[36] FONG T, NOURBAKHSH I, DAUTENHAHN K. Man-computer symbiosis revisited[J]. IEEE Intelligent Systems, 2014, 29(6): 5-9.
[37] 王红卫, 李珏, 刘建国, 等. 人机融合复杂社会系统研究[J]. 中国管理科学, 2023, 31(7): 1-21.
[38] CUKUROVA M. The interplay of learning, analytics and artificial intelligence in education: A vision for hybrid intelligence[J]. British Journal of Educational Technology, 2025, 56(2): 469-488.
[39] DINCELLI E, YAYLA A. Immersive virtual reality in the age of the metaverse: A hybrid-narrative review[J]. Computers in Human Behavior Reports, 2022, 8: 100227.
[40] HSU Y C, NOURBAKHSH I. When human-computer interaction meets community citizen science[J]. Communications of the ACM, 2020, 63(2): 31-34.
[41] 商希雪. 人机交互的模式变革与治理应对——以人形机器人为例[J]. 东方法学, 2024(3): 143-158.

[42] 许晖，龙杨，李阳，等. 人机联合认知视角下制造企业如何实现智能决策[J]. 中国工业经济，2025(4)：174-192.
[43] MOSSBERGER T K. The effects of e-government on trust and confidence in government[J]. Public Administration Review，2010，66(3)：354-369.
[44] 王芳，陈锋. 国家治理进程中的政府大数据开放利用研究[J]. 中国行政管理，2015(11)：6-12.
[45] 张茂月. 大数据时代个人信息数据安全的新威胁及其保护[J]. 中国科技论坛，2015(7)：117-122.
[46] 陈强远，崔雨阳，蔡卫星. 数字政府建设与城市治理质量：来自公共安全部门的证据[J]. 数量经济技术经济研究，2024，41(11)：132-154.
[47] 曾润喜，杨腾飞，徐晓林. 重视网络社会风险保障国家政治安全——2016“国家政治安全与网络社会风险治理”研讨会综述[J]. 中国行政管理，2016(6)：158-159.
[48] 朱正威，肖群鹰. 国际公共安全评价体系：理论与应用前景[J]. 公共管理学报，2006(1)：27-33，108.
[49] 王庆. 低空经济发展：新兴安全风险与敏捷治理[J]. 科学学研究，2025，43(8)：1569-1578.
[50] 张晋宏，李景平，白东海. 基于政治安全的网络舆情泛政治化治理研究[J]. 电子政务，2019(8)：29-39.
[51] 肖红军. 算法责任：理论证成、全景画像与治理范式[J]. 管理世界，2022，38(4)：200-226.
[52] KIM H Y，CHO J S. Data governance framework for big data implementation with NPS case analysis in Korea[J]. Journal of Business and Retail Management Research，2018，12(3)：36-46.
[53] 郭建锦，郭建平. 大数据背景下的国家治理能力建设研究[J]. 中国行政管理，2015(6)：73-76.
[54] 姚志奋，王保民. 政府数据开放的公共安全悖论及其法治策应[J]. 中国科技论坛，2023(8)：139-149.
[55] ANDREW S A，SHORT J E，JUNG K，et al. Intergovernmental cooperation in the provision of public safety：Monitoring mechanisms embedded in interlocal agreements[J]. Public Administration Review，2015，75(3)：401-410.
[56] 谭九生，杨建武. 人工智能技术的伦理风险及其协同治理[J]. 中国行政管理，2019(10)：44-50.
[57] 李广乾，陶涛. 电子商务平台生态化与平台治理政策[J]. 管理世界，2018，34(6)：104-109.
[58] 穆荣平，蔺洁，池康伟，等. 创新驱动社会服务数字转型发展的趋势、国内外实践与建议[J]. 中国科学院院刊，2022，37(9)：1259-1269.
[59] 吕薇，金碚，李平，等. 以新促质，蓄势赋能——新质生产力内涵特征、形成机理及实现进路[J]. 技术经济，2024，43(3)：1-13.
[60] 尹西明，林镇阳，陈劲，等. 数据要素价值化动态过程机制研究[J]. 科学学研究，2022，40(2)：220-229.
[61] BHARADWAJ A，EL SAWY O A，PAVLOU P A，et al. Digital business strategy：Toward a next generation of insights[J]. MIS Quarterly，2013，37(2)：471-482.
[62] 朱秀梅，林晓玥，王天东，等. 数据价值化：研究评述与展望[J]. 外国经济与管理，2023，45(12)：3-17.
[63] HANNILA H，SILVOLA R，HARKONEN J，et al. Data-driven begins with data；potential of data assets[J]. Journal of Computer Information Systems，2022，62(1)：29-38.
[64] 刘祺，黄君义，冯耕中，等. 数字经济发展根基：数据质量研究述评与未来展望[J]. 系统工程理论与实践，2025，45(7)：2101-2123.
[65] 王艳，杨达. 中国式管理会计体系变革：从数据要素到数据资产[J]. 管理世界，2024，40(10)：171-189.
[66] 李姝，赵灿，谢雁翔. 数据资产与企业金融化：数据治理还是概念炒作?[J]. 外国经济与管理，2025，47(10)：21-38.
[67] 林娟娟，黄志刚，唐勇. 数据质量、数量与数据资产定价：基于消费者异质性视角[J]. 中国管理科学，2025，33(5)：88-98.
[68] ZYGMUNTOWSKI J O. Data governance in a trilemma：A qualitative analysis of rights，values，and goals in building data commons[J]. Digital Society，2023，2(2)：30.
[69] 李三希，王泰茗，刘小鲁. 数据投资、数据共享与数据产权分配[J]. 经济研究，2023，58(7)：139-155.
[70] YANG J，WEN J，JIANG B，et al. Blockchain-based sharing and tamper-proof framework of big data networking[J]. IEEE Network，2020，34(4)：62-67.
[71] 刘业政，宗兰芳，杨祖艳，等. 数据要素可信流通交易的研究框架与展望[J]. 南开管理评论，2025，28(4)：158-170.
[72] 夏义堃，宋佳，纪昌秀. 高价值数据集开放与再利用的内涵边界与实践进路分析[J]. 电子政务，2024(6)：54-68.
[73] 张军，王璐，王保平. 数据资产管理进程中的会计治理功能拓展研究[J]. 会计研究，2025(9)：35-46.
[74] 江小涓，宫建霞，李秋甫. 数据、数据关系与数字时代的创新范式[J]. 中国社会科学，2024(9)：185-203，208.
[75] BROEDERS D，SCHRIJVERS E，VAN DER SLOOT B，et al. Big data and security policies：Towards a framework for regulating the phases of analytics and use of big data[J]. Computer Law & Security Review，2017，33(3)：309-323.
[76] 申卫星. 论数据产权制度的层级性：“三三制”数据确权法[J]. 中国法学，2023(4)：26-48.
[77] MCAFEE A，BRYNJOLFSSON E，DAVENPORT T H，et al. Big data：The management revolution[J]. Harvard Business Review，2012，90(10)：60-68.
[78] 张新宝. 论作为新型财产权的数据财产权[J]. 中国社会科学，2023(4)：144-163，207.
[79] 潘大鹏，郝亚杰，乔朋华，等. 数据跨境流转风险分类监管策略研究[J]. 管理评论，2024，36(7)：43-53.
[80] 张军，刘长翠，孔晓冉，等. 数据要素治理的价值发现功能[J]. 经济管理，2025，47(7)：5-24.
[81] ECKARDT M，KERBER W. Property rights theory，bundles of rights on iot data，and the EU data act[J]. European Journal of Law and

Economics, 2024, 57(1): 113-143.
[82] HICKS J. The future of data ownership: An uncommon research agenda[J]. The Sociological Review, 2023, 71(3): 544-560.
[83] 孙宇. 国家数据管理体制构建的历程、逻辑与图景——以国家数据局的组建为视角[J]. 中国科技论坛, 2024(6): 100-110.
[84] 续继, 杨卿栩, 侯嘉奕. 统筹发展与安全的数据治理路径研究[J]. 管理世界, 2026, 42(2): 39-54.
[85] ACHARYA A, SINGH S K, PEREIRA V, et al. Big data, knowledge co-creation and decision making in fashion industry[J]. International Journal of Information Management, 2018, 42: 90-101.
[86] KENNY D, MARSHALL J F. Contextual marketing[J]. Harvard Business Review, 2000, 78(6): 119-125.
[87] 李高勇, 刘露. 场景数字化: 构建场景驱动的发展模式[J]. 清华管理评论, 2021, 11(6): 87-91.
[88] 尹西明, 苏雅欣, 陈劲, 等. 场景驱动的创新: 内涵特征、理论逻辑与实践进路[J]. 科技进步与对策, 2022, 39(15): 1-10.
[89] 尹西明, 陈劲. 场景驱动创新: 数字时代科技强国新范式[M]. 北京: 中国科学技术出版社, 2024.
[90] 尹西明, 聂耀昱, 李振宇, 等. 场景驱动公共数据价值化的理论机制、体系设计与实践模式[J]. 科技创新发展战略研究, 2024, 8(6): 47-60.
[91] HUSSAIN M, TAPINOS E, KNIGHT L. Scenario-driven roadmapping for technology foresight[J]. Technological Forecasting and Social Change, 2017, 124: 160-177.
[92] 尹西明, 林镇阳, 陈劲, 等. 数据要素价值化生态系统建构与市场化配置机制研究[J]. 科技进步与对策, 2022, 39(22): 1-8.
[93] HEIN A, SCHREIECK M, RIASANOW T, et al. Digital platform ecosystems[J]. Electronic Markets, 2020, 30(1): 87-98.
[94] 孙南翔. 论作为消费者的数据主体及其数据保护机制[J]. 政治与法律, 2018(7): 21-34.
[95] RAHMAN H T, SAINT VILLE A S, SONG A M, et al. A framework for analyzing institutional gaps in natural resource governance[J]. International Journal of the Commons, 2017, 11(2): 823-853.
[96] KHATRI V, BROWN C. Designing data governance[J]. Communications of the ACM, 2010, 53(1): 148-152.
[97] 李德顺. 价值论: 一种主体性的研究[M]. 3 版. 北京: 北京师范大学出版社, 2020.
[98] JANSSEN M, ZUIDERWIJK A. Infomediary business models for connecting open data providers and users[J]. Social Science Computer Review, 2014, 32(5): 694-711.
[99] 崔宏轶, 冼骏. 政务数据管理中的“数据可用性”——痛点及其消解[J]. 中国行政管理, 2019(8): 55-60.
[100] 郑大庆, 黄丽华, 张成洪, 等. 大数据治理的概念及其参考架构[J]. 研究与发展管理, 2017, 29(4): 65-72.
[101] 衲钦, 张慧春. 数智环境下匿名数据治理创新对策研究[J]. 科学管理研究, 2022, 40(2): 124-130.
[102] 刘露, 杨晓雷. 新基建背景下的数据治理体系研究——以数据生命周期为总线的治理[J]. 治理研究, 2020, 36(4): 59-66.
[103] 朱琳, 赵涵菁, 王永坤, 等. 全局数据: 大数据时代数据治理的新范式[J]. 电子政务, 2016(1): 34-42.
[104] BANNISTER F. Dismantling the silos: Extracting new value from IT investments in public administration[J]. Information Systems Journal, 2001, 11(1): 65-84.
[105] 方鑫, 李香, 魏姝. 场景驱动的政府数字化转型: 数字技术何以回应复杂治理需求——以江苏南通市为例[J]. 电子政务, 2024(6): 14-27.
[106] 夏义堃, 管茜. 政府数据资产管理的内涵、要素框架与运行模式[J]. 电子政务, 2022(1): 2-13.
[107] ZUIDERWIJK A, JANSSEN M. Open data policies, their implementation and impact: A framework for comparison[J]. Government Information Quarterly, 2014, 31(1): 17-29.
[108] 蒋佳妮, 刘晓博. 身联网场景中数据应用与个人信息保护的平衡路径[J]. 财经理论与实践, 2025, 46(4): 152-160.
[109] 秦浩朗, 肖延高, 汤芃东, 等. 数据价值的基本问题: 文献回顾与研究展望[J]. 科研管理, 2026, 47(1): 11-24.
[110] 马广惠, 魏玮, 安小米. 大数据议题批判性反思[J]. 电子政务, 2019(5): 109-115.
[111] RADAUER A, SEARLE N, BADER M A. The possibilities and limits of trade secrets to protect data shared between firms in agricultural and food sectors[J]. World Patent Information, 2023, 73: 102183.
[112] 尹西明, 苏雅欣, 陈泰伦, 等. 场景驱动型人工智能创新生态系统: 逻辑与进路[J]. 中国科技论坛, 2024(6): 35-45.
[113] 夏诗园, 尹振涛. 数字经济下金融数据风险及治理研究[J]. 电子政务, 2022(7): 57-66.
[114] 聂耀昱, 范梓腾, 张文泽. 数据要素视角下公共数据开发利用的县域治理困境与长效路径——以中部 L 县为例[J]. 电子政务, 2024(5): 21-32.
[115] 崔文波, 蔚海燕. 人工智能环境下多模态数据治理问题与路径研究[J/OL]. 图书馆杂志, 2025: 1-12[2026-07-16]. https://link.cnki.net/urlid/31.1108.G2.20251031.1532.002.
[116] 黄国平. 创新和重塑数据治理体系——以金融数据治理为例[J]. 经济管理, 2023, 45(1): 25-42.
[117] 窦智, 韩永辉. 生态环境大数据治理促进绿色创新了吗: 基于环境风险管控与协同治理的分析[J]. 中国软科学, 2025(9): 154-165.
[118] 杨艳, 鲁世宇, 林凌. 跨部门的政策红利: 政府数据治理如何塑造企业市场竞争优势——来自地方大数据管理局设立的实践证据[J]. 外国经济与管理, 2025, 47(12): 24-40.
[119] OTTO B. Organizing data governance: Findings from the telecommunications industry and consequences for large service providers[J]. Communications of the Association for Information Systems, 2011, 29(1): 3.

[120] LEE J, YOON J, KIM B. A big data analytics platform for smart factories in small and medium-sized manufacturing enterprises: An empirical case study of a die casting factory[J]. International Journal of Precision Engineering and Manufacturing, 2017, 18(10): 1353-1361.
[121] EICHLER R, GRÖGER C, HOOS E, et al. Introducing the enterprise data marketplace: A platform for democratizing company data[J]. Journal of Big Data, 2023, 10: 173.
[122] GUMBO L, CHUDE UA. Regulatory sandbox as a frontier for innovation and sustainability: A systematic review[J]. Cogent Business & Management, 2025, 12(1): 2510555.
[123] 张欣. 生成式人工智能的数据风险与治理路径[J]. 法律科学(西北政法大学学报), 2023, 41(5): 42-54.
[124] 李敏. 金融科技的监管模式选择与优化路径研究——兼对监管沙盒模式的反思[J]. 金融监管研究, 2017(11): 21-37.
[125] 程雪军, 尹振涛. 监管科技的发展挑战与中国选择: 基于金融科技监管视角[J]. 经济体制改革, 2022(1): 135-142.
[126] KALKAR U, SAXENA S, VERHULST S. Data sandboxes: Governance, design, and use cases[R]. New York: Open Data Policy Lab, The GovLab, 2023.
[127] 吕悦, 陈旭, 彭子璇, 等. 静态到敏捷: 人工智能监管沙盒治理机制研究[J]. 科学学研究, 2026, 44(5): 910-920.
[128] CHEN X, WANG J, YANG W, et al. Risk preference-based decision-making and control framework for pedestrian interaction[J]. IEEE Transactions on Intelligent Transportation Systems, 2025, 26(12), 21605-21621.
[129] HUNT T, ZHU Z, XU Y, et al. Ryoan: A distributed sandbox for untrusted computation on secret data[J]. ACM Transactions on Computer Systems, 2018, 35(4): 1-32.
[130] SENEVIRATHNA T, LA V H, MARCHA S, et al. A survey on XAI for 5G and beyond security: Technical aspects, challenges and research directions[J]. IEEE Communications Surveys & Tutorials, 2024, 27(2): 941-973.
[131] 林梓瀚, 刘羿鸣, 袁千里. 自贸区数据跨境流动监管沙盒机制: 理论证成与制度建构[J]. 中国科技论坛, 2025(5): 30-39, 71.
[132] LECKENBY E, DAWOUD D, BOUVY J, et al. The sandbox approach and its potential for use in health technology assessment: A literature review[J]. Applied Health Economics and Health Policy, 2021, 19: 857-869.

Context Identification and Sandbox Mechanisms in Data Governance: A Theoretical Framework from a Human-computer Interaction Perspective

Liu Yulin, Xiao Yangao, Li Daitian, Xiao Tianrui, Wang Xiaotian

(School of Management and Economics, University of Electronic Science and Technology of China, Chengdu 611731, China)

Abstract: A value-security data governance scenario matrix was constructed from the perspective of human-computer interaction theory. A four-stage data governance sandbox mechanism, including modeling (M), experiment (E), diagnosis (D), and interpretation (I), was further proposed for the high-value-high-risk data governance scenario. The following findings are obtained. Firstly, a new theoretical perspective and analytical approach are provided for data governance in the digital economy by human-computer interaction theory. Data governance is not regarded as a linear process in which security compliance precedes data utilization or data utilization precedes security compliance. A dynamic balance between data utilization and security compliance is required. Secondly, a value-security data governance scenario matrix can be established based on the two dimensions of value density and security risk. The key challenge of data governance is identified in the high-value-high-risk scenario, where substantial value can be created through data utilization, while significant losses may occur if data are not effectively protected or are illegally exploited. Thirdly, the MEDI data governance sandbox mechanism enables rule solidification, problem tracing, and rule regeneration through human-computer interaction. The sandbox is therefore transformed from a technical isolation tool into a dynamic data governance system characterized by verifiability, traceability, and regenerability. An embeddable and reusable data governance mechanism framework is provided for the implementation of Article 13 of the Data Security Law, which requires data security to be promoted through data utilization and data utilization to be safeguarded through data security.

Keywords: data governance; context identification; sandbox mechanism; human-computer interaction; theory development; MEDI model