

doi: 10.19920/j.cnki.jmsc.2026.03.005

# 数智服务的信任赤字：生成机制与演化路径<sup>①</sup>

宋瑞震, 曾赛星\*, 高鑫

(上海交通大学安泰经济与管理学院, 上海 200030)

**摘要:** 作为数字化与智能化技术深度融合的新型业态,数智服务在提升服务效率与个性化水平的同时,也引发了用户信任危机.本文将“信任赤字”的概念引入数智服务领域,用以刻画人智信任失位所引发的服务信任水平持续低于信任预期的情景.信任赤字是制约数智服务效能的关键,但基于人际信任或技术信任的单一框架,难以解释其复杂的动态特征.本文聚焦数智服务中信任赤字的多维度结构与交互关系,构建了包含能力、善意、诚实与可控性的四维信任赤字模型,揭示了信任赤字跨周期增强、多维度耦合与异质性衰减的演化路径,并探究了数智服务类型、服务场景风险和用户数智素养对信任赤字的影响机制.研究表明:服务失败后,工具型和共情型数智服务在主导赤字维度与演化路径上存在显著差异;服务场景风险显著放大信任赤字的增长速度,重构信任赤字的演化过程;用户的数智素养能有效缓解信任赤字的增长趋势,甚至使信任赤字在多层服务交互中逐步消失.本文拓展了数智服务信任赤字的外延,并为其治理策略的制定提供理论依据.

**关键词:** 数智服务; 信任赤字; 人智信任; 情景仿真

**中图分类号:** C931 **文献标识码:** A **文章编号:** 1007-9807(2026)03-0061-13

## 0 引言

随着数智化技术的迅速发展,以人工智能为代表的技术变革正在重塑服务业的运行逻辑与价值创造模式<sup>[1-3]</sup>.数智服务作为数字技术与智能算法深度融合赋能服务业而形成的新型服务,已广泛嵌入零售、医疗、金融、交通、教育等多个领域,呈现出高效率、多样化和个性化的服务特征<sup>[4]</sup>.从推荐系统、智能客服到自动驾驶、医疗辅助诊断,数智化技术不仅提升了服务效率与资源配置能力,更催生出普惠化、精细化的服务生态体系,成为推动经济社会高质量发展的关键力量<sup>[5,6]</sup>.

数智化技术在加速服务业变革的同时,也带来了人智信任的危机.算法黑箱、隐私泄露、深度

伪造等问题频发,引发了公众对数智服务可信性的广泛担忧<sup>[2]</sup>.数智化技术的深度嵌入,使数智服务中人智交互的程度与频率显著提升,改变了传统的人智关系,信任失位逐步成为制约数智服务效能提升的瓶颈<sup>[7]</sup>.人智信任的失位不仅源于数智化技术的不确定性,还与用户对技术的误解、不可控与不认同相关<sup>[8]</sup>.这种不信任并非静态存在,而是在持续的人智交互中不断扩散、累积和放大,最终演化为系统性、持续性的信任失位,本研究将其定义为“数智服务的信任赤字”.数智服务的信任赤字是指在以人工智能、算法和大数据等为核心驱动力的数智服务交互过程中,用户感知的人智信任失位所引发的服务信任水平持续低于信任预期的情景.其中,用户感知的人智信任是其在实际服务交互过程中对数智化技术及其代

① 收稿日期: 2025-07-31; 修订日期: 2025-11-01.

基金项目: 国家自然科学基金资助重大项目(72394375).

通讯作者: 曾赛星(1966—),男,安徽安庆人,博士,教授,博士生导师. Email: zengsaixing@sjtu.edu.cn

理的信任判断,而服务信任水平是用户在此基础上对数智服务整体的信任情况.信任预期则是用户基于自身期望、服务提供商承诺、行业与技术规范等多重预期对数智服务的信任预判.用户的信任对象不仅是技术工具或服务提供商,还包括由数智化技术嵌入和驱动的整体服务过程.因此,用户对数智服务的信任赤字不仅源自技术本身,也与服务目标差异、服务环境风险和用户认知偏差密切相关,其本质是数智化技术的决策逻辑与用户有限理性、情感需求、认知局限或行为习惯等之间的矛盾.

信任赤字不仅使用户对服务的使用频率和依赖性降低,更可能引发服务中断、用户退出等问题,阻碍服务业的数智化转型与发展<sup>[9]</sup>.因此,随着数智化技术深度嵌入多样化的服务情景,厘清不同类型数智服务中信任赤字的多维度生成机制与动态演化路径十分关键.

人智信任是数智服务可持续运营的前提,更是将数智化技术的先进优势转化为实际服务价值的关键<sup>[9]</sup>.本研究聚焦数智服务中用户的信任失位现象,将信任赤字的概念扩展到数智服务领域,并从动态视角对其多维度结构与交互关系进行剖析,重点关注以下三个研究问题:首先,由于数智服务的技术特征与应用场景的复杂性,其信任失位并非单一维度的表现,而是多维度信任赤字交织作用的结果<sup>[10]</sup>,既有人际信任或技术信任的理论模型难以揭示数智服务中信任赤字的多维度结构.因此,本文首先关注:如何刻画数智服务中信任赤字的多维度生成机制;其次,数智服务中的人智交互具有连续性,信任赤字并不会在单次服务中固定,而是随着服务过程不断演化,具有连续性与动态性<sup>[8]</sup>.但现有研究大多停留在静态的信任评估,尚不足以解释信任赤字的累积效应及多维度的相互作用.对此,本文继续聚焦:如何识别数智服务中信任赤字的动态演化路径;再次,数智服务的多样化使信任赤字在不同场景中呈现出差异化的演化过程.不同类型服务的侧重不同,服务所处的环境风险亦会放大或抑制信任赤字,而用户的数智素养则可能改变其对技术的理解和接受程度<sup>[11]</sup>.当前研究虽已开始关注服务特征、环境风险和用户差异对服务交互的影响,但其对数智服务中信任赤字的影响机制尚不明确.所以,本文进

一步探究:如何解析数智服务类型、服务场景风险和用户数智素养对信任赤字生成与演化的影响机制.

本研究在信任理论的基础上,对数智服务中信任赤字的生成机制与演化路径进行系统探讨.首先,构建了包含能力、善意、诚实与可控性的四维信任赤字理论模型,解析了数智服务中信任赤字的生成机制.然后,从四维信任赤字的复杂交互关系出发,揭示了信任赤字跨周期增强、多维度耦合与异质性衰减的演化路径.最后,基于情景仿真,探究了数智服务类型、服务场景风险和用户数智素养对信任赤字的影响机制.综上,本研究拓展了数智服务信任赤字的外延,强调用户信任状态与数智化技术间的动态匹配过程,为数智服务的信任赤字治理提供理论依据.

## 1 文献综述

### 1.1 数智服务与人智交互

数智服务是数字技术与智能技术深度融合背景下的新型服务形态,正在重塑服务业的服务供需关系、资源组织方式与价值创造逻辑.以人工智能、大数据、云计算等为代表的数智化技术深度融入服务应用场景,智慧医疗、数字金融、无人驾驶等服务新业态不断涌现<sup>[1]</sup>.数智化技术不仅拓展了服务的边界和能力,还推动了服务从静态供给向动态响应、标准化交付向个性化定制的转变,极大地促进了经济社会的高质量发展<sup>[2]</sup>.

人智交互关系也被数智化技术重塑,人们开始将其视为平等的伙伴,而非辅助的工具<sup>[4]</sup>.相关研究从服务代理<sup>[12]</sup>、信任关系<sup>[13]</sup>、组织创新<sup>[14]</sup>、分工协作<sup>[7]</sup>、情感互动<sup>[9]</sup>等多个角度探索了数智化技术带来的巨大价值.然而,数智化技术的复杂性以及“黑箱”状态引发了人智交互过程中的信任问题<sup>[15]</sup>.对此,相关研究还提出了算法厌恶的概念,即用户与自动化算法交互过程中,对其决策或建议的抵触心理与规避行为,表现为一种非理性偏好<sup>[16]</sup>.

综上,现有研究从多个维度解析了数智服务的典型特征,并关注了人智交互过程中信任关系的重要作用,但对数智服务中人智信任的研究尚显不足.

## 1.2 信任赤字与人智信任

信任作为人类社会稳定的重要基石,长期以来都是不同学科领域中讨论的热点<sup>[11]</sup>.其中,“信任赤字”也受到广泛关注,相关研究尝试从社会运行<sup>[17]</sup>、组织合作<sup>[18]</sup>与经济发展<sup>[19]</sup>等多个角度定义信任赤字.在数智服务的情境下,当算法黑箱、权责模糊等问题导致用户感知的人智信任水平低于其合理或期望水平时,就可能形成信任赤字.然而,数智服务的信任赤字在现有研究中尚未形成明确的定义.部分研究尝试拆解人智信任的多维度特征,以揭示数智服务场景下的信任特征.Mayer等<sup>[10]</sup>提出人际信任包含能力、善意和诚实三个维度,为后续研究人智信任提供了结构化的理论框架<sup>[20]</sup>.但人际信任强调意图性与关系对称性,这在人工智能的非人类系统中可能并不成立<sup>[21]</sup>.因此,有研究提出在自动化情境中,智能服务系统被视为“受托人”,用户为“委托人”,信任的建立围绕性能、过程和目的三个核心要素展开<sup>[11]</sup>.这类研究对人智信任的界定虽然借鉴了人际信任理论,但在概念上更接近以功能性或工具性为主导的信任范畴.

然而,数智服务既有客观工具的功能属性,又有主观情感的类人特征.简单地依照人际信任中的经典理论,或仅从工具性信任的角度定义人智信任的边界,均难以概括数智服务中信任赤字的多维度特征.

## 1.3 人智信任的影响因素

人智信任作为提升数智服务效能、协调人智关系的核心,相关研究尝试挖掘人智信任的关键影响因素并提出建立人智互信的治理策略<sup>[22]</sup>.对人智信任的影响因素研究主要从三个层面展开:其一,关注人工智能本身的技术特征及其表现形式对人智信任的影响.相关研究从提高技术可解释性<sup>[23]</sup>、性能水平<sup>[16]</sup>、表现形式<sup>[24]</sup>、适应性与个性化<sup>[25]</sup>等方面提出改善人智信任的策略;其二,聚焦环境风险对人智信任的影响.信任有时被认为是一种风险可接受性的评估方法,尤其是用户缺少特定领域知识的情况下<sup>[11]</sup>.在高风险的场景中,用户的信任门槛更高<sup>[26]</sup>,而在低风险的场景中,信任则更多基于体验与偏好的一致性<sup>[22]</sup>;其三,围绕用户的主观感知对人智信任的影响.在人智交互中,信任并不是客观存在的,而是用户主观

感知和评价的,受到其专业知识、信任倾向以及情感需求等个人特征的影响<sup>[11]</sup>.相关研究指出,用户获得人工智能的知识和经验后会更容易产生积极的态度,并建立信任关系<sup>[22]</sup>.

显然,人智信任的形成与发展是数智化技术特征、应用场景与用户感知三个维度耦合的结果,而现有研究大多聚焦单一维度对人智信任的影响过程,无法解释数智服务中人智信任的复杂情况.

## 2 数智服务的信任赤字

在数智服务的情景中,信任的来源不仅包含类似人际信任相关理论中对行为动机和道德意图的感知<sup>[10]</sup>,还包含类似自动化信任相关理论中对技术能力和可操作性的评价<sup>[27]</sup>.本研究将用户的信任判断视为一个动态的过程<sup>[8]</sup>.基于Mayer等<sup>[10]</sup>的三维信任理论模型,并结合Vuori等<sup>[28]</sup>对认知信任与情感信任的分类与组合,本研究提出数智服务的四维信任赤字模型,四个维度包含能力、善意、诚实和可控性(如图1).在理论模型构建中,设定用户在数智服务中的信任赤字向量为

$$D_t = (A_t, B_t, I_t, C_t) \quad (1)$$

其中 $A_t$ 为能力赤字;  $B_t$ 为善意赤字;  $I_t$ 为诚实赤字;  $C_t$ 为可控性赤字;  $t$ 为时间变量,表示四维信任赤字是随时间变化的连续变量,以刻画用户在多轮数智服务交互中信任的形成与演化过程.

此外,用户信任判断的重心会基于服务目标的差异而改变.因此,本研究将服务目标划分为为解决具体任务为核心的工具型数智服务( $E$ )<sup>[29]</sup>,以及以提供情感支持为核心的共情型数智服务( $T$ )<sup>[30]</sup>,并设计连续变量以刻画数智服务中客观工具与主观情感的复合特征.在理论模型中,设定数智服务类型 $i \in \{E, T\}$ 的信任赤字为

$$D_i^j = \begin{cases} D_i^E, \tau \in (0, 0.5) \\ D_i^T, \tau \in (0.5, 1) \end{cases} \quad (2)$$

其中 $D_i^E$ 为共情型数智服务的信任赤字;  $D_i^T$ 为工具型数智服务的信任赤字;  $\tau \in (0, 1)$ 为数智服务类型变量,当 $\tau$ 趋向于0时,数智服务更具有共情导向;当 $\tau$ 趋向于1时,数智服务更具有工具导向.

进一步,本研究考虑了实际情况中用户信任

感知的内外部影响因素,即服务场景的风险性  $R \in (0,1)$  与用户的数智素养  $\theta \in (0,1)$ , 并将之融入理论模型,以挖掘内外部因素的影响效应.

数智化技术赋能的服务情景中用户的信任感知过程具有以下三个典型特征:其一,数智服务中的信任赤字具有连续性与动态性.在数智服务交互过程中,用户的信任赤字并非一次性判断形成的静态变量,而是随时间、经验、反馈不断更新的动态过程;其二,数智服务中的信任赤字具有多维度交互与演化的复杂性.数智服务中的用户信任

感知不再是单一维度的判断,而是一种多维度评估、交互与演化的过程;其三,数智服务的信任赤字具有主体异质性.由于数智化技术的可解释性较低,而不同用户在数智素养、认知能力、风险感知、心理预期等方面存在结构性差异,这使他们的信任赤字演化路径完全不同.基于以上分析,本研究提出数智服务中信任赤字的三条演化路径:跨周期增强、多维度耦合与异质性衰减,并基于此进行理论模型的构建,以反映用户在数智服务中信任赤字的动态演化过程.

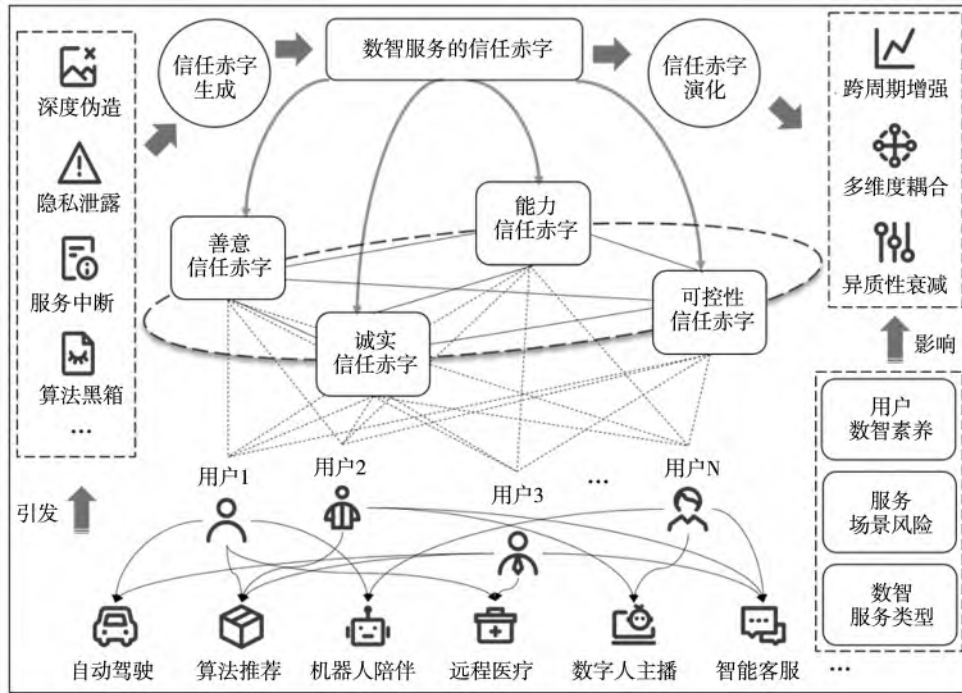


图1 数智服务四维信任赤字概念框架

Fig. 1 Conceptual framework of the four-dimensional trust deficits in intelligent digital services

### 2.1 信任赤字的跨周期增强

在数智服务的情景中,用户在服务交互后会形成“信任锚点”,对后续多轮服务中的信任判断产生累积性影响<sup>[31]</sup>.这种锚点效应反映了信任的路径依赖机制,即用户对服务系统的信任水平,可能影响其未来对该系统的态度.因此,用户对早期数智服务系统的失败体验将会形成认知偏见,产生不信任的“锚点”,这也是信任赤字产生并积累的根源.算法厌恶就是一种典型的信任赤字跨周期影响结果,即用户对算法决策或建议的抵触心理或行为.基于此,本研究设定信任赤字的跨周期影响过程

$$(\mathbf{E}_D)^T = \begin{bmatrix} E_A \\ E_B \\ E_I \\ E_C \end{bmatrix} = \begin{bmatrix} \tau A_t \\ (1 - \tau) B_t \\ (1 - \tau) I_t \\ \tau C_t \end{bmatrix} \quad (3)$$

其中  $\mathbf{E}_D$  为信任赤字的跨周期增强效应;  $E_A$ 、 $E_B$ 、 $E_I$  和  $E_C$  分别为能力、善意、诚实和可控性赤字的跨周期增强效应;  $\tau$  为数智服务类型变量;  $A_t$ 、 $B_t$ 、 $I_t$  和  $C_t$  分别为第  $t$  期时,四个维度的信任赤字水平.

数智服务场景的风险性与信任赤字的形成和演化密切相关.一方面,用户在高风险场景下的失败服务经历更容易固化为不信任的“锚点”,进而

形成信任赤字. 另一方面, 在高风险场景中即便是微小的偏差也可能被用户认为服务系统不可靠, 从而加速信任赤字积累. 因此, 数智服务场景的风险性不仅是信任赤字形成的“触发器”, 更是其跨周期演化的“放大器”. 所以, 考虑服务场景风险, 本研究设定信任赤字的跨周期影响

$$(\mathbf{E}_D^t)^T = \begin{bmatrix} E_A^t \\ E_B^t \\ E_I^t \\ E_C^t \end{bmatrix} = \begin{bmatrix} \tau RA_t \\ (1 - \tau) RB_t \\ (1 - \tau) RI_t \\ \tau RC_t \end{bmatrix} \quad (4)$$

其中  $R \in (0, 1)$  为数智服务场景的风险性系数;  $\mathbf{E}_D^t$  为考虑服务风险后, 信任赤字的跨周期增强效应;  $E_A^t$ 、 $E_B^t$ 、 $E_I^t$  和  $E_C^t$  分别为考虑服务风险后, 能力、善意、诚实和可控性赤字的跨周期增强效应.

## 2.2 信任赤字的多维度耦合

数智服务的场景中, 用户在能力、善意、诚实和可控性维度的信任综合构成了对数智服务的信任结构. 不同维度的信任赤字在形成与演化过程中会通过认知迁移或情感投射等机制触发或影响其他维度的信任. 因此, 数智服务中单一维度的信任失位往往并非孤立存在, 信任赤字呈现出多维度耦合的动态演化特征.

在数智服务交互过程中, 可控性赤字跨阶段影响着善意、诚实和能力赤字的演化. 用户每一轮的服务交互可以分解为三个阶段: 选择服务、接受服务和评价服务, 在经历一轮服务后就会形成信任感知, 并影响下一轮服务的选择. 在选择服务阶段, 用户关注服务是否符合其个人利益或偏好, 当用户对数智服务的动机产生质疑时, 就可能引发善意赤字. 若用户缺乏调控服务动机的能力, 则可控性赤字将激发其对数智服务动机的怀疑, 进一步放大善意赤字. 在接受服务阶段, 用户开始感知服务行为与反馈的真实性和透明性, 若数智服务的分析过程缺失或决策逻辑模糊, 就可能引发用户的诚实赤字. 当用户难以控制或理解这些过程时, 将会加剧对数智服务诚实性的质疑, 引发更大的诚实赤字. 在评价服务阶段, 用户关注服务是否实现了指定目标, 若无法达到预期, 就可能引发能力赤字. 用户可控性赤字的存在将使其对数智服

务的能力提出更高的要求, 使服务效能不足时产生更大的能力赤字. 基于此, 为刻画可控性赤字与能力、善意和诚实赤字的多阶段耦合机制, 本研究设定可控性赤字的多维度影响

$$(\mathbf{Co}^C)^T = \begin{bmatrix} Co_A^C \\ Co_B^C \\ Co_I^C \end{bmatrix} = \begin{bmatrix} (1 - \tau) C_t \\ \tau C_t \\ \tau C_t \end{bmatrix} \quad (5)$$

其中  $\mathbf{Co}^C$  为可控性赤字对其他维度赤字的影响系数;  $Co_A^C$ 、 $Co_B^C$  和  $Co_I^C$  分别为可控性赤字对能力、善意和诚实赤字的耦合系数;  $C_t$  为第  $t$  期可控性赤字水平.

数智服务的结果不达预期时, 用户对能力的质疑可能被归因为服务动机的偏差, 即能力赤字引发善意赤字进一步增长. 由于数智化技术的黑箱状态, 当用户遭遇服务失败时, 一方面会将之归因为数智服务本身的能力缺陷, 另一方面会认为服务没有将用户目的或利益放在首位, 甚至可能有不利于用户的意图. 因此, 当用户在数智服务中产生能力赤字后, 就可能进一步对服务动机产生怀疑. 同时, 这种归因的过程会在可控性赤字存在时进一步扩大善意赤字. 因此, 为刻画能力赤字通过认知归因机制对善意赤字的激发路径, 本研究设定能力赤字的影响

$$Co_B^A = \tau A_t \quad (6)$$

其中  $Co_B^A$  为能力赤字对善意赤字的耦合系数;  $A_t$  为第  $t$  期能力赤字水平.

数智服务的可控性赤字激发了能力、善意和诚实赤字的的增长, 这些赤字的出现使得用户在服务交互中提出更高的控制需求, 从而反过来影响可控性赤字的变化. 当用户发现数智服务的动机或目标与自身利益预期不符, 但又难以将这一不信任仅归咎于服务能力不足时, 则会在服务交互中产生更强的控制需求, 如试图关闭推荐信息、手动设定参数、干预算法行为等. 基于此, 为刻画能力赤字与善意赤字共同影响可控性赤字的过程, 本研究设定能力与善意赤字的影响

$$Co_C^{AB} = (1 - \tau) (B_t - A_t) \quad (7)$$

其中  $Co_C^{AB}$  为善意赤字和能力赤字对可控性赤字的耦合系数;  $A_t$  和  $B_t$  为第  $t$  期时, 能力赤字和善

意赤字水平.

另一方面,当用户在服务交互中无法理解其运行逻辑,如算法推荐规则、信息采集方式等,就会对服务的透明性和真实性产生怀疑,从而直接激发出对可控性的心理焦虑.所以,为刻画诚实赤字对可控性赤字的直接影响,本研究设定诚实赤字的影响

$$Co_C^I = (1 - \tau) I_t \tag{8}$$

其中  $Co_C^I$  为诚实赤字对可控性赤字的耦合系数;  $I_t$  为第  $t$  期诚实赤字水平.

### 2.3 信任赤字的异质性衰减

在数智服务场景中,用户的信任赤字会随着其数智素养的提升,而呈现出阶段性的衰减趋势.数智化技术已经广泛融入服务业的多个环节,使得用户在生活中能够持续接触数智服务,逐步熟悉和适应新的服务模式.用户通过多次服务交互不断学习和了解数智化技术的功能,从而对数智服务的效能、动机和行为逻辑有更清晰的判断,即数智素养提升,从而使信任赤字降低.

用户的数智素养影响其对数智服务信任赤字的衰减程度,但对认知信任赤字与情感信任赤字衰减的作用机制不同,表现出信任维度异质性与用户个体异质性.一方面,认知信任是用户基于理性计算对数智服务性能、逻辑、操控能力等因素的评估,服务失败后产生能力赤字和可控性赤字.对于数智素养较低的用户,由于缺乏对智能服务能力范围与潜在风险的认知,其在初始阶段往往表现出相对盲目的信任,对服务交互中数智化技术的能力或可控性缺陷缺乏敏感性,这反而使其认知信任赤字较快衰减.而对于数智素养较高的用户,因其熟悉数智化技术的行为逻辑,且具备较强的技术理解与操作能力,从而使认知信任赤字更快衰减.相比之下,数智素养中等的用户具备部分数智化技术的知识,但还没有深入理解数智服务的运行机制.在面对能力不足或不确定性时,由于用户对服务偏差和结果更具有敏感性但又难以解释其来源,即形成疑而不信的状态,这将导致更强的不确定感与失控感,使认知信任持续存在.

另一方面,情感信任是用户基于情感纽带对

数智服务动机、行为一致性等因素的感性认知,服务失败后产生善意赤字和诚实赤字.与认知信任赤字不同的是,情感信任赤字是从情感共鸣与道德归因中对数智服务的情绪判断结果,难以通过认识数智化技术缺陷而被修复.数智素养较低的用户对服务的动机没有清晰的认知,因而更容易经历服务失败,使其形成的善意和诚实赤字难以恢复.而随着用户数智素养的逐步提高,更加了解数智化技术的适用范围与局限性,这使其在选择与接受数智服务时能更主动地规避情感冲突的风险,且更明确数智化技术的非人性本质,对善意或道德期待更加理性,从而更快缓解善意赤字与诚实赤字的的增长趋势.

基于以上分析,本研究设定用户的数智素养与能力赤字和可控性赤字的衰减速度为 U 型关系,而与善意赤字与诚实赤字的衰减速度刻画为正相关关系,异质性衰减过程为

$$(\mathbf{F}_D)^T = \begin{bmatrix} F_A \\ F_B \\ F_I \\ F_C \end{bmatrix} = \begin{bmatrix} -(2(\theta - \theta_A))^2 A_t \\ -\theta B_t \\ -\theta I_t \\ -(2(\theta - \theta_C))^2 C_t \end{bmatrix} \tag{9}$$

其中  $\mathbf{F}_D$  为信任赤字的异质性衰减系数;  $F_A$ 、 $F_B$ 、 $F_I$  和  $F_C$  分别为能力、善意、诚实和可控性赤字的异质性衰减系数;  $\theta \in (0,1)$  为用户的数智素养水平;  $\theta_A$  和  $\theta_C$  分别为使能力和可控性信任赤字异质性衰减系数取极小值点的数智素养水平;  $A_t$ 、 $B_t$ 、 $I_t$  和  $C_t$  分别为第  $t$  期时,四个维度的信任赤字水平.

综上,本研究设定用户在数智服务中信任赤字的动态演化方程为

$$(\mathbf{dD}_t/\mathbf{dt})^T = \begin{bmatrix} \mathbf{dA}_t/\mathbf{dt} \\ \mathbf{dB}_t/\mathbf{dt} \\ \mathbf{dI}_t/\mathbf{dt} \\ \mathbf{dC}_t/\mathbf{dt} \end{bmatrix} = \begin{bmatrix} E'_A + Co_A^C + F_A \\ E'_B + Co_B^C + Co_B^A + F_B \\ E'_I + Co_I^C + F_I \\ E'_C + Co_C^{AB} + Co_C^I + F_C \end{bmatrix} \tag{10}$$

其中  $dD_t/dt$  为数智服务中信任赤字在  $t$  时刻的演化速率;  $dA_t/dt$ 、 $dB_t/dt$ 、 $dI_t/dt$  和  $dC_t/dt$  分别为能力、善意、诚实和可控性赤字在  $t$  时刻的演化速率;  $E_A^t$ 、 $E_B^t$ 、 $E_I^t$  和  $E_C^t$  分别为能力、善意、诚实和可控性赤字的跨周期增强效应;  $Co_A^C$ 、 $Co_B^C$  和  $Co_I^C$  分别为可控性赤字对能力、善意和诚实赤字的耦合系数;  $Co_B^A$  为能力赤字对善意赤字的耦合系数;  $Co_C^{AB}$  为善意赤字和能力赤字对可控性赤字的耦合系数;  $Co_C^I$  为诚实赤字对可控性赤字的耦合系数;  $F_A$ 、 $F_B$ 、 $F_I$  和  $F_C$  分别为能力、善意、诚实和可控性赤字的异质性衰减系数。

### 3 情景仿真

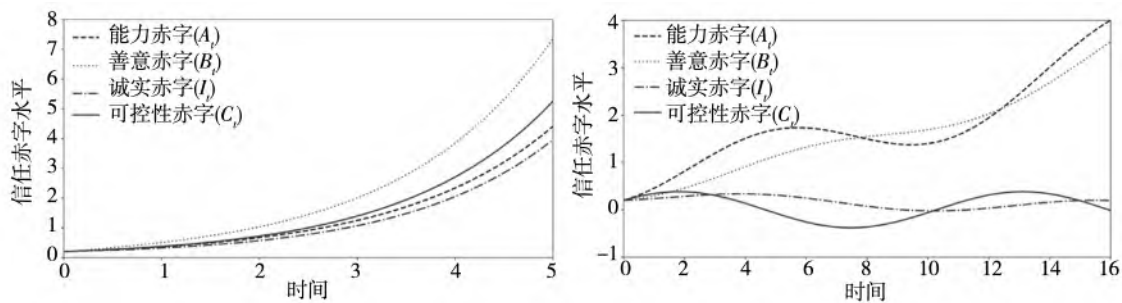
在数智服务情境下,信任赤字的生成与演化受到多重因素的共同作用,既包括服务本身的功能属性,也涉及外部环境的不确定性以及用户个体的认知差异.对此,本研究分析了服务类型、环境风险与用户素养对信任赤字动态演化过程的影响机制.首先,数智服务类型体现了服务本身的目的与能力特征,是用户进行信任判断的直接依据.本研究通过区分数智服务类型,分析工具型与共情型数智服务情景中信任赤字的演化过程,探究数智服务类型对信任赤字演化路径的影响;其次,数智服务场景风险代表外部环境的不确定性,是影响用户风险感知与信任评估的重要外生变量.本研究在不同数智服务类型的基础上,分析数智服务场景的风险如何影响用户的信任感知过程,揭示数智服务场景的风险性对信任赤字演化结果

的影响;再次,用户的数智素养则反映了个体内部的认知和适应能力,决定了用户对复杂服务逻辑的理解与信任赤字的缓释水平.本研究考虑用户的数智素养,分别探讨工具型与共情型数智服务情景下用户的数智素养对信任赤字演化的影响机制.

在不同类型的数智服务情景中,用户对服务的功能需求、价值评判和道德标准有所差异,这使得信任赤字的动态演化过程不同.实际情况中,用户初始阶段对技术感知能力、透明性、动机均有不确定性判断,导致不同维度的初始信任赤字.为了体现不同维度信任赤字的相对变化情况,不失一般性的,本研究设定信任赤字的初始值为 0.2,信任赤字衰减系数取极小值点的数智素养水平为 0.5.

#### 3.1 工具型和共情型数智服务中的信任赤字演化

设定数智服务场景的风险性与用户的数智素养为中等水平 ( $R = \theta = 0.5$ ),工具型数智服务情景下 ( $\tau = 0.8$ ) 的信任赤字演化的仿真结果如图 2(a) 所示.用户在四个维度的信任赤字均呈现出不断加速上升的趋势,这表明在工具型数智服务中,用户因失败经历而产生的信任赤字不仅会持续存在,还会出现信任的加速坍塌.其中,善意赤字的增长最快,成为主导型信任赤字.这表明工具型数智服务中,能力的缺失直接降低了服务的完成水平,用户将服务的失败不再视作能力局限,而认为是服务提供方的有意为之.因此,当用户对数智服务的动机产生怀疑时,就会将服务的能力问题归咎于服务的目的,使得在演化过程中善意赤字高于能力赤字.



(a) 工具型数智服务的信任赤字演化 (b) 共情型数智服务的信任赤字演化

图2 不同数智服务类型下的信任赤字演化

Fig. 2 Evolution of trust deficits in different types of intelligent digital services

在工具型数智服务的情景中,可控性赤字的增长也高于能力赤字. 由于工具型数智服务一般依赖复杂的底层算法驱动,如购物平台的推荐系统、自动驾驶系统等,用户往往缺乏对其分析或决策过程的参与和控制. 这使得用户一旦感知到服务偏离期望,而难以通过反馈或修正干预其行为,就会形成控制被剥夺的感受. 值得注意的是,这种剥夺感并不源于服务能力不足,而是源于用户只能被动接受服务结果.

图2(b)为共情型数智服务情景下( $\tau = 0.2$ )的信任赤字演化的仿真结果. 用户的能力赤字和善意赤字总体为上升趋势,并演化成为共情型数智服务的主导性赤字. 从演化过程来看,能力赤字与善意赤字的演化还表现出持续震荡的过程,其中能力赤字的变化幅度更大. 在共情型数智服务中,如老年人陪护机器人、儿童互动助手等,用户的情感需求更高,对情绪的判断也更敏感. 当服务难以响应用户的情感需求时,用户会首先质疑其情感识别、理解和反馈能力. 另一方面,在共情型数智服务的情景中,用户的可控性赤字和诚实赤字也表现出波动性的演化特征,但总体为缓慢下降趋势. 不同于工具型数智服务中确定性的目标达成,共情型数智服务更强调情感上的陪伴. 因此,用户更倾向于将共情型数智服务系统视为交流的伙伴而不是执行命令的助手,这使其对服务交互中的控制性要求更低,甚至不希望系统被自己过多的操控,所以可控性赤字的演化过程表现

为缓慢下降. 此外,用户在服务失败后更容易将之理解为共情能力不足,而不是故意欺骗,所以诚实赤字也没有显著增长.

### 3.2 数智服务场景的风险性对信任赤字的影响

工具型数智服务情景中( $\tau = 0.8, \theta = 0.5$ ), 服务风险性低( $R = 0.2$ )时信任赤字演化的仿真结果如图3(a)所示,服务风险高( $R = 0.8$ )时的演化结果如图3(b)所示. 总体来看,无论风险性高低,工具型数智服务中用户四个维度的信任赤字均保持加速增长的趋势. 但随着服务场景风险性增加,信任赤字增长速度显著提升,并在短时间内演化为更高的赤字水平,反映出风险在信任赤字跨周期演化中的“放大器”效果. 高风险环境会激发用户更频繁地评估、归因和质疑服务交互中的行为,如虚拟医生、自动驾驶等,每一个小错误都会使用户放大成对整个服务系统能力和可靠性的质疑. 而在低风险情景中,用户更可能包容服务中的错误,产生的信任赤字总体处于较低的水平. 其中,善意赤字总是保持最高的增长速度,为工具型数智服务情景中的主导性赤字. 这表明即使在低风险场景下,用户对工具型数智服务的动机始终保持高敏感性. 进一步结合图2(a)的结果,可以发现随着服务场景的风险性增加,能力赤字的相对增长速度逐步提升,在工具型数智服务中的重要性越来越高. 这说明在高风险环境中,用户更倾向于将服务失败归因为非偶发性的、可能频繁发生的能力缺陷.

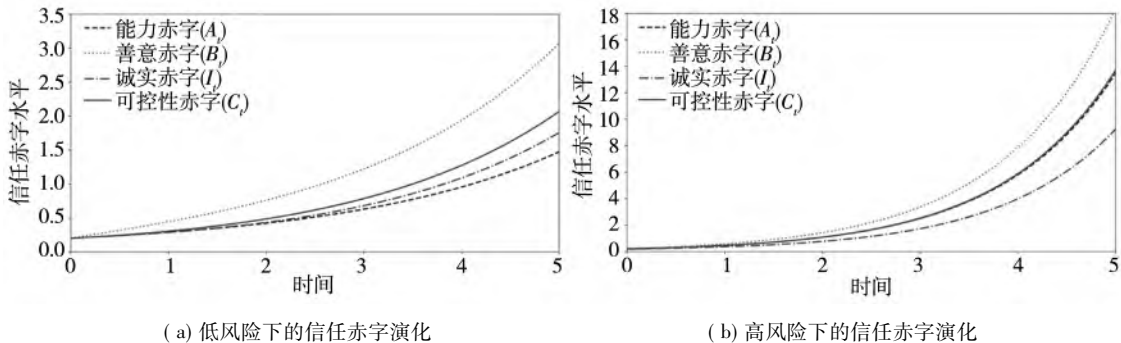


图3 工具型数智服务中不同风险下的信任赤字演化

Fig. 3 Evolution of trust deficits under varying risk levels in tool-based intelligent digital services

共情型数智服务情景中( $\tau = 0.2, \theta = 0.5$ ), 服务风险性低( $R = 0.2$ )时信任赤字演化的仿真

结果如图4(a)所示. 从趋势上来看,共情型数智服务中四个维度的信任赤字交互震荡且振幅越来越

越大,表现出交互增强的演化过程. 共情型数智服务以情绪理解和情感共鸣为重心,在低风险的服务场景下,用户初始时对信任感知的敏感度相对较低、对服务系统错误的容忍度相对较高. 用户会进入不断试错、修复和重建信任的过程,由此形成信任赤字的交互震荡,但由于信任赤字的跨周期增强,震荡幅度逐步增大.

在服务风险高 ( $R = 0.8$ ) 时,对共情型数智服务的信任赤字的仿真结果如图 4(b) 所示. 随着风险

性的增加,四个维度信任赤字演化过程中的震荡过程消失,变为持续上升的趋势. 在高风险情境下,用户对服务表现得敏感性更高,对服务需求从情绪陪伴转向确定性的情感支持. 此时,服务交互中的错误会被放大,原本试错和修复的过程转为信任赤字的持续上升. 具体来看,能力赤字与善意赤字增长更快,成为共情型数智服务的主导性赤字. 这表明在共情型数智服务交互中,用户更加重视的是情感的有效性和真实性,而不是逻辑的透明性和可操控性.

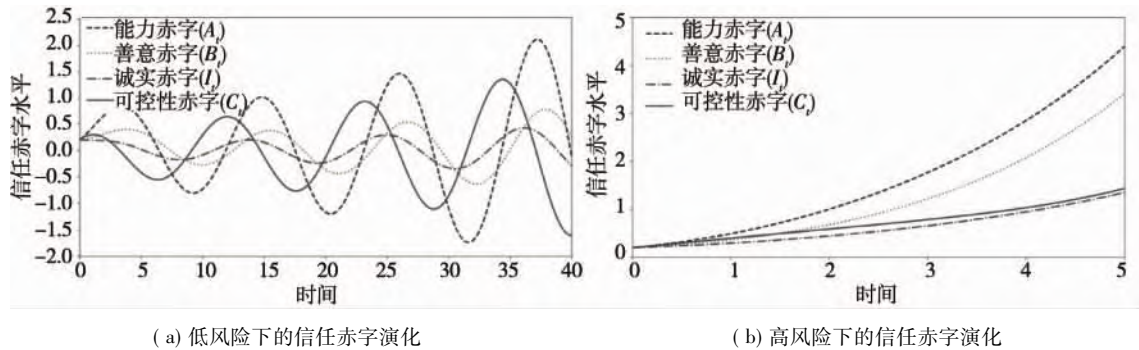


图 4 共情型数智服务中不同风险下的信任赤字演化

Fig. 4 Evolution of trust deficits under varying risk levels in empathy-based intelligent digital services

### 3.3 用户的数智素养对信任赤字的影响

工具型数智服务情景中 ( $\tau = 0.8, R = 0.5$ ), 用户数智素养低时信任赤字演化的仿真结果如图 5(a) 所示, 用户数智素养高时的演化结果如图 5(b) 所示. 结果显示, 用户素养的变化并没有改变信任赤字逐步上升的演化趋势. 这表明在工具型数智服务中, 信任的路径依赖性更强, 使得信任赤字一旦开始积累就难以在后续的交互中反转. 结合图 2(a) 的演化结果, 可以发现随着用户数智素养的提高, 信任赤字的增长速度呈现出先上升后下降的变化. 这表明用户的数智素养虽然无法

逆转信任赤字的上升趋势, 但能改变信任赤字的上升速度. 具体来看, 随着用户的数智素养提高, 善意赤字始终保持主导性的地位, 而能力赤字和可控性赤字的重要性逐步提高. 一方面, 在工具型数智服务中, 即便数智素养高的用户更能理解技术原理, 但仍然在服务动机上保留怀疑, 使善意赤字为主导性赤字. 另一方面, 随着用户数智素养的提高, 对服务交互中错误的归因过程趋于理性, 更加关注服务的能力是否满足需求或是否可自定义调控, 即高数智素养用户的信任感知不仅局限于对服务结果的评价, 而扩展到对数智能化技术能力和有效性的判断.

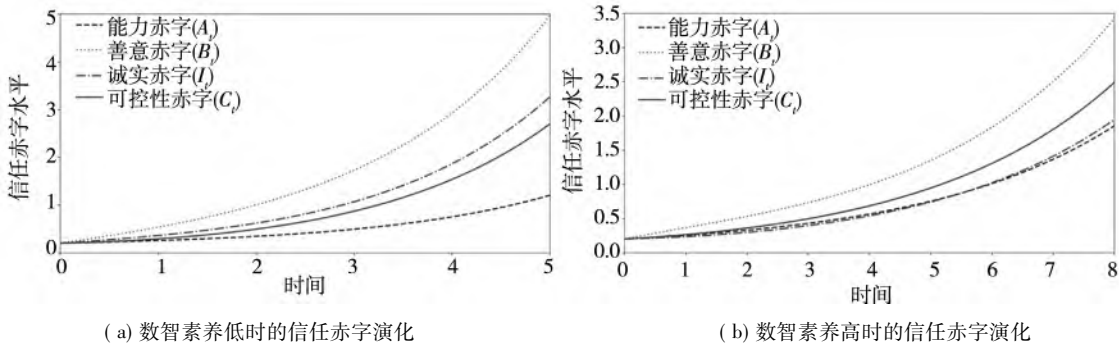


图 5 工具型数智服务中不同数智素养下的信任赤字演化

Fig. 5 Evolution of trust deficits under different levels of intelligent digital literacy in tool-based services

共情型数智服务情景中 ( $\tau = 0.2, R = 0.5$ ), 当用户数智素养较低时信任赤字演化的仿真结果如图 6(a) 所示, 各维度的信任赤字均表现出加速上升的趋势; 当用户数智素养高时, 信任赤字表现为先上升后下降的演化过程, 如图 6(b) 所示. 这表明用户的数智素养对信任赤字的演化过程具有关键的影响作用. 一方面, 结合图 2(b) 的结果, 当用户的数智素养过低时, 信任赤字演化中震荡变化的过程转变为加速上升. 数智素养过低的用户在服务交互中缺乏识

别情感反馈的能力, 因此在多轮服务中无法完成信任感知的动态调节, 进而难以建立信任修复的过程, 另一方面, 随着用户的数智素养不断提高, 信任赤字在震荡后最终收敛为零. 先增后减的震荡过程反映了用户在共情型数智服务交互中, 进行试错、修复和重建信任的过程. 较高的数智素养使用户在信任震荡变化的过程中能够更快识别服务能力与服务动机, 使服务交互中的情感表达和交互反馈更有效, 从而实现信任赤字的收敛式修复.

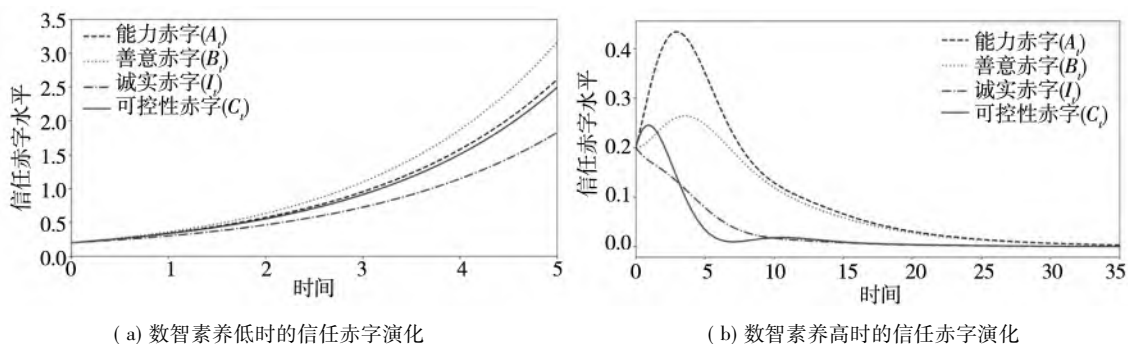


图 6 共情型数智服务中不同数智素养下的信任赤字演化

Fig. 6 Evolution of trust deficits under different levels of intelligent digital literacy in empathy-based services

## 4 结束语

数智化技术深度融入服务业的各个环节, 人机交互频率和依赖性不断提高, 信任已成为发挥数智服务效能的关键. 本研究聚焦数智服务中的用户信任失位, 将“信任赤字”的概念扩展到数智服务领域. 通过构建四维信任赤字模型, 刻画数智服务信任赤字的多维度生成机制, 揭示其跨周期增强、多维度耦合和异质性衰减的动态演化路径. 基于情景仿真, 解析数智服务类型、服务场景风险和用户数智素养对信任赤字生成与演化的影响机制. 本研究不仅拓展了信任理论在数智服务情境中的适用边界, 也为构建数智服务信任赤字的治理措施提供理论依据.

### 4.1 研究结论

本研究构建了数智服务中四维信任赤字模型, 并对信任赤字的动态演化过程进行模拟仿真, 主要结论如下.

其一, 数智服务中的信任赤字具有多维度特

征, 表现出动态性与复杂性. 本研究提出能力、善意、诚实与可控性四个维度的信任赤字模型, 揭示用户信任感知并非静态, 而是在交互过程中不断更新演化, 反映了数智服务中用户信任判断在时间演进、多维交互与个体差异层面的深层机制.

其二, 工具型与共情型数智服务在信任赤字的演化路径与主导维度上存在显著差异. 在工具型数智服务中, 用户对服务动机的敏感性较高, 呈现出信任加速坍塌的趋势. 相比之下, 共情型数智服务中信任赤字的演化具有明显的波动性, 整体表现为持续震荡式上升, 体现出较强的互动调节机制与主导维度的不确定性.

其三, 服务场景的风险性显著放大了信任赤字的增长速度. 无论是在工具型还是共情型数智服务中, 服务场景的风险性越高, 用户对服务系统的要求越严苛、期望越高, 信任赤字的增长趋势越加陡峭. 这表明服务场景的风险不仅改变了用户的归因逻辑, 也重构了信任的演化路径.

其四, 用户的数智素养对信任赤字具有非线性的影响. 用户的数智素养虽不能完全逆转工具

型数智服务信任赤字的上升趋势,但对其演化速度与主导维度具有显著影响.较高的数智素养能使共情型数智服务信任赤字的增长趋势放缓,甚至促进信任赤字在多轮交互中逐步消失.

#### 4.2 理论贡献

基于以上研究结论,本研究的理论贡献主要体现在以下三个方面.

其一,将“信任赤字”的概念扩展到数智服务领域,构建了数智服务的四维信任赤字模型.在传统服务的信任评估中,以往研究多关注用户对服务效能的积极作用<sup>[5, 32]</sup>,即使有对服务中信任失位的关注也没有形成统一的概念和系统性的评估框架<sup>[18]</sup>.本研究提出数智服务信任赤字的概念,拓展了信任理论在数智服务情境下的研究边界,系统阐述了服务交互中信任赤字的生成机制与演化路径.

其二,揭示了信任赤字跨周期增强、多维度耦合与异质性衰减的演化路径.在传统的服务场景中,相关研究多将服务提供商视为信任的对象,关注服务提供前后用户的信任变化<sup>[33]</sup>,忽视了服务交互中的动态演化过程.本研究基于数智服务信任赤字四个维度的特征,解析了多轮服务交互中信任赤字演化路径,为理解数智服务情景中人智信任的动态性与系统性提供了新视角.

其三,解析了服务类型、服务风险与数智素养对信任赤字的影响机制.现有研究多将人智信任视为静态结构或单一维度驱动<sup>[20, 27]</sup>,难以解释数智服务中信任赤字的复杂变化.本研究构建了一个基于服务属性、外部环境与用户个体的信任赤字动态演化模型,揭示了信任赤字在不同情境下的主导维度迁移、演化速度变化与动态耦合效应,弥补了现有理论中对人智信任结构层面影响机制

的研究空白.

#### 4.3 实践启示

本研究对于数智服务中信任赤字的管理实践具有重要意义,主要包括以下两个方面.

其一,推动基于数智服务类型与信任赤字维度的信任治理策略体系构建.针对工具型与共情型数智服务在信任赤字结构与演化过程上的差异,应针对性地提出信任治理机制,以最大化提升治理效率.在工具型服务中应优先提升系统的动机透明性与用户干预机制,以降低善意与可控性赤字;而在共情型服务中,则需重点优化服务的情感识别与响应能力,强化对用户情感反馈的理解与记忆,以减少能力与善意赤字的累积.

其二,强化服务风险识别、提升用户数智素养,构建信任共建机制.一方面,在医疗、金融、交通等高风险领域配置基于风险等级的动态信任反馈机制,以降低信任坍塌的风险.另一方面,注重提升用户的算法认知与交互理解能力,使用户在交互中获得更充分的认知准备与调节能力,从而实现信任共建的正向循环,增强数智服务效能.

本研究仍存在一定的局限性.首先,本文主要基于理论建模与逻辑推演,解析了数智服务信任赤字的生成机制与演化路径,尚缺乏大样本数据来验证信任赤字的动态演化过程.未来研究可以通过实验设计或多源行为数据分析,进一步检验模型结论的稳健性与适用范围.其次,模型中对数智服务类型及场景风险的刻画为静态设定,未能充分反映不同交互阶段与情境变化下的动态过程.未来研究可以对信任赤字的不同影响机制进行动态演化建模或纵向案例分析,并探讨算法透明度、隐私保护机制等多维度因素的影响,以更全面地揭示数智服务信任赤字的复杂演化规律.

#### 参考文献:

- [1] Rahwan I, Cebrian M, Obradovich N, et al. Machine behaviour [J]. Nature, 2019, 568(7753): 477–486.
- [2] Choi W J, Liu Q, Shin J. Predictive analytics and ship-then-shop subscription [J]. Management Science, 2024, 70(2), 1012–1028.
- [3] 陈国青, 吴刚, 顾远东, 等. 管理决策情境下大数据驱动的研究和应用挑战——范式转变与研究方向 [J]. 管理科学学报, 2018, 21(7): 1–10.  
Chen Guoqing, Wu Gang, Gu Yuandong, et al. The challenges for big data driven research and applications in the context of

- managerial decision-making: Paradigm shift and research directions [J]. *Journal of Management Sciences in China*, 2018, 21(7): 1–10. (in Chinese)
- [4] 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.  
Wang Hongwei, Li Jue, Liu Jianguo, et al. Research on human-machine integration complex social system [J]. *Chinese Journal of Management Science*, 2023, 31(7): 1–21. (in Chinese)
- [5] 卢向华, 邹玉凤. AI 普及化背景下的价值提升机制与未来研究方向——基于人机持续互信视角 [J]. *中国科学基金*, 2024, 38(5): 867–875.  
Lu Xianghua, Zou Yufeng. AI value enhancement mechanisms and future research directions in the ubiquitous AI era: A perspective from sustainable human-machine trust [J]. *Bulletin of National Natural Science Foundation*, 2024, 38(5): 867–875. (in Chinese)
- [6] 周雨薇, 吕巍. 人工智能重塑零售行业的底层逻辑: 综述及展望 [J]. *系统管理学报*, 2021, 30(1): 180–190.  
Zhou Yuwei, Lü Wei. Underlying logic of artificial intelligence reshaping the retail industry: Overview and prospect [J]. *Journal of Systems & Management*, 2021, 30(1): 180–190. (in Chinese)
- [7] Dennis A R, Lakhiwal A, Sachdeva A. AI agents as team members: Effects on satisfaction, conflict, trustworthiness, and willingness to work with [J]. *Journal of Management Information Systems*, 2023, 40(2): 307–337.
- [8] Daly S J, Wiewiora A, Hearn G. Shifting attitudes and trust in AI: Influences on organizational AI adoption [J]. *Technological Forecasting and Social Change*, 2025, (2150): 124108.
- [9] Li M, Bitterly T B. How perceived lack of benevolence harms trust of artificial intelligence management [J]. *Journal of Applied Psychology*, 2024, 109(11): 1794–1816.
- [10] Mayer R C, Davis J H, Schoorman F D. An integrative model of organizational trust [J]. *Academy of Management Review*, 1995, 20(3): 709–734.
- [11] Wirz C D, Demuth J L, Bostrom A, et al. (Re) Conceptualizing trustworthy AI: A foundation for change [J]. *Artificial Intelligence*, 2025, (342): 104309.
- [12] 华中生, 袁泉. 数字化医疗服务资源的跨域协同 [J]. *系统工程学报*, 2025, 40(5): 738–747.  
Hua Zhongsheng, Yuan Quan. Cross-domain collaboration of digital medical service resources [J]. *Journal of Systems Engineering*, 2025, 40(5): 738–747. (in Chinese)
- [13] Omrani N, Riviaccio G, Fiore U, et al. To trust or not to trust? An assessment of trust in AI-based systems: Concerns, ethics and contexts [J]. *Technological Forecasting and Social Change*, 2022, (181): 121763.
- [14] 吴小龙, 肖静华, 吴记, 等. 人与 AI 协同对组织学习的影响机制研究——探索与利用学习的视角 [J]. *管理科学学报*, 2024, 27(9): 11–28.  
Wu Xiaolong, Xiao Jinghua, Wu Ji, et al. The influence mechanism of human-AI collaboration on organizational learning: The exploratory and exploitative perspectives [J]. *Journal of Management Sciences in China*, 2024, 27(9): 11–28. (in Chinese)
- [15] Wang W, Benbasat I. Recommendation agents for electronic commerce: Effects of explanation facilities on trusting beliefs [J]. *Journal of Management Information Systems*, 2007, 23(4): 217–246.
- [16] Dietvorst B J, Simmons J P, Massey C. Algorithm aversion: People erroneously avoid algorithms after seeing them err [J]. *Journal of Experimental Psychology-General*, 2015, 144(1): 114–126.
- [17] Adelopo I, Rufai I. Trust deficit and anti-corruption initiatives [J]. *Journal of Business Ethics*, 2020, 163(3): 429–449.
- [18] Kramer R M, Lewicki R J. Repairing and enhancing trust: Approaches to reducing organizational trust deficits [J]. *Academy of Management Annals*, 2010, (4): 245–277.
- [19] Hillenbrand C, Money K, Ghobadian A. Unpacking the mechanism by which corporate responsibility impacts stakeholder relationships [J]. *British Journal of Management*, 2013, 24(1): 127–146.
- [20] Wang W, Qiu L, Kim D, et al. Effects of rational and social appeals of online recommendation agents on cognition-and affect-based trust [J]. *Decision Support Systems*, 2016, (86): 48–60.
- [21] Hoff K A, Bashir M. Trust in automation: Integrating empirical evidence on factors that influence trust [J]. *Human Factors*, 2015, 57(3): 407–434.
- [22] Lalot F, Bertram A M. When the bot walks the talk: Investigating the foundations of trust in an artificial intelligence (AI) chatbot [J]. *Journal of Experimental Psychology-General*, 2025, 154(2): 533–551.
- [23] Xu J, Benbasat I, Cenfetelli R T. The nature and consequences of trade-off transparency in the context of recommendation agents [J]. *MIS Quarterly*, 2014, 38(2): 379–406.

- [24] Glikson E, Woolley A W. Human trust in artificial intelligence: Review of empirical research [J]. *Academy of Management Annals*, 2020, 14(2): 627–660.
- [25] Fuegener A, Grahl J, Gupta A, et al. Cognitive challenges in human-artificial intelligence collaboration: Investigating the path toward productive delegation [J]. *Information Systems Research*, 2022, 33(2): 678–696.
- [26] Bockstedt J C, Buckman J R. Humans' use of ai assistance: The effect of loss aversion on willingness to delegate decisions [J]. *Management Science*, 2026, 72(1): 323–342.
- [27] Lankton N K, Mcknight D H, Tripp J. Technology, humanness, and trust: Rethinking trust in technology [J]. *Journal of the Association for Information Systems*, 2015, 16(10): 880–918.
- [28] Vuori N, Burkhard B, Pitkaranta L. It's amazing-but terrifying!: Unveiling the combined effect of emotional and cognitive trust on organizational member' behaviours, AI performance, and adoption [J]. *Journal of Management Studies*, 2026, 72(1): 323–342.
- [29] Dietvorst B J, Simmons J P, Massey C. Overcoming algorithm aversion: People will use imperfect algorithms if they can( even slightly) modify them [J]. *Management Science*, 2018, 64(3): 1155–1170.
- [30] Chandra S, Shirish A, Srivastava S C. To be or not to be horizontal ellipsis human? Theorizing the role of human-like competencies in conversational artificial intelligence agents [J]. *Journal of Management Information Systems*, 2022, 39(4): 969–1005.
- [31] Luhmann N. *Social Systems* [M]. Stanford: Stanford University Press, 1995.
- [32] Decamp M, Lindvall C. Mitigating bias in AI at the point of care [J]. *Science*, 2023, 381(6654): 150–152.
- [33] Montoya M M, Massey A P, Khatri V. Connecting IT services operations to services marketing practices [J]. *Journal of Management Information Systems*, 2010, 26(4): 65–85.

## Trust deficits in intelligent digital services: Generative mechanisms and evolutionary pathways

SONG Rui-zhen, ZENG Sai-xing<sup>\*</sup>, GAO Xin

Antai College of Economics and Management, Shanghai Jiao Tong University, Shanghai 200030, China

**Abstract:** As a novel service paradigm emerging from the deep integration of digital and intelligent technologies, intelligent digital services have enhanced service efficiency and personalization while simultaneously triggering a trust crisis among users. This study introduces the concept of a trust deficit into the domain of intelligent digital services to characterize situations in which the displacement of human-AI trust leads to persistently lower service trust levels than expected. As a key constraint on the effectiveness of intelligent digital services, trust deficits cannot be adequately explained by single-dimensional frameworks grounded solely in interpersonal trust or technical trust. This research focuses on the multidimensional structure and interrelationships of trust deficits, proposing a four-dimensional trust deficit model encompassing ability, benevolence, integrity, and controllability. It identifies the evolutionary pathways of trust deficits characterized by cross-cycle amplification, multidimensional coupling, and heterogeneous decay, and examines the mechanisms through which service type, contextual risk, and users' intelligent digital literacy shape the evolution of trust deficits. Scenario simulations reveal that following service failure, tool-based and empathy-based intelligent digital services differ markedly in their dominant deficit dimensions and evolutionary trajectories; Service risk significantly accelerates the growth of trust deficits and reshapes their evolutionary patterns; And higher intelligent digital literacy effectively mitigates the escalation of trust deficits, even enabling their convergence to zero over multiple service interactions. This study broadens the theoretical boundaries of trust research in the context of intelligent digital services and offers a conceptual foundation for developing strategies to manage trust deficits.

**Key words:** intelligent digital services; trust deficit; human-AI trust; scenario simulation