doi:10.16060/j.cnki.issn2095-8072.2026.06.006

智慧仲裁时代的人机交互障碍及其克服*

武振国

（北京科技大学文法学院，北京 100083）

摘　要：随着智慧仲裁系统的广泛应用，人类在仲裁决策过程中的互动地位被显著弱化，人机交互障碍日趋突出。特别是，当智慧仲裁系统一步生成结果时，人机交互障碍使得当事人本应享有的互动体验被大幅压缩，由此产生的“损失厌恶”加剧了人机之间的信任隔阂，使当事人对相同的客观结果产生更强烈的质疑与抵触。传统以“结果公平、程序公平”为核心的客观公平分析框架，难以有效回应因人机交互障碍所导致的信任危机，因为它未能认识到互动体验是人民群众在每一个司法案件中切实感受到公平正义的重要载体。立足于组织公平理论的谱系演进，本文将“互动公平”作为破解人机交互障碍的核心切入点。互动公平所蕴含的信息公平与人际公平有助于破解智慧仲裁中的人机交互障碍，重塑当事人的互动体验与情绪认同，从而切实提升人类对智慧仲裁的信任度与接受度。

关键词：互动公平；智慧仲裁；算法黑箱；人工智能

中图分类号：D997.4/DF974　　**文献标识码：**A　　**文章编号：**2095－8072(2026)06－0088－10

一、问题的提出

随着人工智能（artificial intelligence，下文简称AI）技术的快速应用，仲裁系统正在从人工化驱动模式向自动化驱动模式转型升级（张圣翠和田洋，2024）。在AI技术与各行业的深度融合下，各主流仲裁机构纷纷开始部署“AI仲裁秘书”，实现从立案、庭审记录到裁决生成的全链路自动化（贺林平，2023）。2025年新修订的《中华人民共和国仲裁法》（以下简称新《仲裁法》）第11条明确鼓励在线仲裁，支持新技术在仲裁领域的应用。然而，越来越多的研究表明，仲裁程序自动化水平与当事人接受度提升之间并未呈现正向同步关系。尤其在涉外智慧仲裁这一特定场域，法系背景差异、语言表达习惯、跨文化沟通礼仪等多重因素叠加，极易在境外当事人与智慧仲裁系统之间形成人机交互障碍。例如，来自英美法系的当事人习惯通过当庭陈述、交叉询问、即时辩论等方式来强化程序参与度，而高度自动化的智慧仲裁系统则会大幅压缩人机之间的互动空间，导致当事人产生程序疏离感。正因如此，在人机互动供给不足的智慧仲裁审理架构下，自动化程度越高，人机交互障碍反而愈发凸显，当事人与智慧仲裁系统之间的信任隔阂亦同步加深。可以说，智慧仲裁系统所面临的危机已不再是传统语境下因客观公平供给不足所引发的信任危机，而是由人机交互障碍所催生的新型信任危机。

对于这场由人机交互障碍所引发的新型信任危机，以“结果公平、程序公平”为核心的客观公平分析框架难以提供合理解释（项军，2019）。客观公平分析框架预设客观公平能够自动转化为当事人对裁决的主观认可，这一逻辑在涉外智慧仲裁催生的跨文化人机交互障碍中，将会遭遇根本性冲击。事实上，不同法域与文化背景下的纠纷主体并非纯粹的理性决策个体，其对智慧仲裁的接受度很大程度上来源于人机交互过程中的人格尊重、理由释明与沟通交流。具言之，当智慧仲裁系

* 基金项目：本文受中国博士后科学基金第75批面上资助项目“科学数据商业化利用的法律规制研究”（项目编号：2024M750168）和中央高校基本科研业务经费专项资金资助项目“新技术时代的技术治理与育人研究”（项目编号：FRF–BD–26–026）的共同资助。

统通过统一模板自动生成裁决草稿时，境外当事人与仲裁庭之间的直接互动渠道被弱化，难以有效理解算法决策的内在逻辑与裁判形成过程，其通过程序参与和面对面沟通所建立的信任纽带难以维系。党的二十大报告中所倡导的“努力让人民群众在每一个司法案件中感受到公平正义”理念，也从价值层面印证了交互体验在纠纷解决中的基础性意义。由此可见，当事人对智慧仲裁系统的信任度和接受度不能只依赖传统的客观公平体系，亟需引入新的理论维度破解人机交互障碍。

从组织公平理论的演进看，互动公平之所以被单列为独立于结果公平、程序公平的第三维度，是因为它能为破解智慧仲裁时代的人机交互障碍提供关键理论支撑（孙东山，2020）。互动公平所内含的信息公平与人际公平，构成了“努力让人民群众在每一个司法案件中感受到公平正义”的信任纽带。这意味着，智慧仲裁系统面临的本质挑战，并非单一提升算力，而在于如何通过有效的人机互动机制消解交互障碍。基于此，本文提出一个基本命题：互动公平是破解人机交互障碍的核心维度。为此，本文以互动公平理论为分析视角，构建相应的制度化解决方案，以期回应智慧仲裁时代因人机交互缺失所引发的信任困境。

二、智慧仲裁时代人机交互障碍的核心表征

在智慧仲裁时代，人机之间若缺少解释、回应和确认，互动缺失便不只是表面现象，而会成为人机交互障碍的核心表征。尤其是在跨文化仲裁场景下，人机交互障碍将放大法律文化差异，形成沟通阻隔，致使当事人觉得自己的诉求未被真正理解，对裁决的接受度亦会随之降低。可见，自动化程度越高，越要为人机交互留出充足空间。

（一）人机交互的信息不对称障碍

在智慧仲裁时代，信息不对称既是造成互动信息缺失的重要原因，也是致使当事人互动体验缺失的首要诱因。具体而言，智慧仲裁系统为互动公平维度下的信息沟通制造了双重困境：一方面，智慧仲裁系统所依赖的算法具有黑箱属性，通常直接输出运算结论而缺乏充分的裁判逻辑释明与决策依据展示，致使不同法律文化背景的当事人难以理解推理基础；另一方面，数据链条缺乏可溯源性，将导致算法偏差无法被有效核验与追溯。上述交互障碍在跨文化涉外仲裁语境下表现得尤为突出，共同构成互动信息缺失的核心表征。

1. 算法黑箱化的信息不对称障碍

在智慧仲裁时代，真正棘手的制度风险并非算法预测是否准确，而是算法黑箱化带来的深层信息不对称（吴维锭，2024）。智慧仲裁系统所依赖的深度学习模型，往往由高维度参数和复杂非线性结构构成，其决策逻辑既在机器语言中运行，又被封装进难以直接读取、难以直观解释的计算框架。于是，算法怎样理解案件事实、怎样筛选证据、怎样给不同因素赋权，基本都以技术语言完成，而不会自动转化为当事人能读懂的自然语言。尤其是在涉外仲裁场景中，这种技术性遮蔽往往与文化差异所形成的理解鸿沟相互叠加，会进一步放大人机交互障碍。比如，普通法系当事人常常希望系统围绕争议焦点逐项回应，想知道哪些事实被采纳、哪些证据被排除、理由如何形成等。若系统只给出简化的裁决结论，这种互动体验显然难以达到其预期，也会明显削弱当事人对仲裁结果的信任感和接受度（邱威棋，2025）。

算法黑箱化的风险不仅局限于信息被遮蔽的表层问题，更核心的危害在于其切断了人机之间的语言互动链条。在跨文化仲裁中，囿于表达习惯和沟通文化的不同，这一问题尤为突出。当事人无法确认自己的陈述是否被算法正确解读，更无法判断智慧仲裁系统是否理解同一表达方式在不同

文化语境下的差异化含义。譬如，境外当事人的肢体语言和语气性表达更能代表其真实意思，但智慧仲裁系统在自然语言和机器语言的转译过程中压缩甚至删除了这些信息，就会使人机之间产生误判，引发新的紧张关系（刘一帆，2025）。

2. 数据不溯源的信息不对称障碍

在智慧仲裁时代，数据不溯源带来的信息不对称，已不只是技术瑕疵，更是人机交互中的一道实际障碍。大模型要靠海量训练数据提高判断精度，可这些数据从何而来、由什么构成、经过怎样的预处理，当事人通常看不到。当事人难以判断其中嵌入的立法导向或文化偏好，在个案中是否都是适当的。放到跨文化争议中，这种数据的不透明更容易被境外当事人理解为人机交互障碍，进而加深其对智慧仲裁系统的不信任（武振国和周健，2024）。

更深一层的风险，还不止于数据来源不透明，而在于它可能直接撼动仲裁制度最有特色的个案正义逻辑。传统仲裁中，裁决本应牢牢锚定于当事人为个案选定的仲裁规则、准据法和仲裁地，再结合个案事实，提供有针对性的、个性化定制的法律服务（宋连斌，2025）。然而，在智慧仲裁中，数据不溯源可能使算法将训练阶段吸收的偏差固化为自身决策逻辑，并以“技术中立”为外衣，将与本案无关联性，甚至不应适用的他案仲裁规则及他国法律规范，隐性嵌入当前案件的裁判过程。如果智慧仲裁系统在后台私自对来自不同文化背景的当事人设置差异化的可信度评分，显然就会形成一种隐性歧视。在此情况下，缺乏清晰的数据溯源体系，不仅意味着智慧仲裁系统可能演化为一种“不可追问”的算法利维坦，还会使当事人本应具有的陈述权和申辩权被系统性弱化，放大人机之间的交互障碍（张爱军，2021）。可以说，数据不溯源虽然在形式上提高了仲裁效率，但却可能因隐性适用他案规则、忽视表达差异而偏离个案正义轨道，最终破坏当事人对智慧仲裁的信任度和接受度。

（二）人机交互的地位不对等障碍

智慧仲裁越是自动化，人机之间的权力配置越有可能被重塑。传统仲裁中，仲裁员的居中裁判、程序引导等互动，并非可有可无的形式，而是确认当事人程序主体地位的关键载体。也正因有这些互动，不同文化背景的当事人才更可能感受到程序尊严与主体价值。智慧仲裁时代，人机交互障碍削弱了这种体验。当事人不仅难以通过互动充分感知自身的主体性，反而更容易觉得自己只是被系统处理的对象。因此，自动化不只关乎效率，也关乎程序尊严与主体地位如何得到保障。

1. 主体倾听地位弱化

在智慧仲裁时代，主体倾听地位的弱化已成为实现有效人机互动的重大阻力。在传统商事仲裁中，倾听并非可有可无的礼节，而是仲裁员恪守正当程序的核心义务。它也是保障境外当事人程序参与权、弥合跨法系与跨文化差异的关键路径。进入智慧仲裁时代，以人际倾听为核心的互动模式被自动化流程替代，当事人的程序参与度在主体表达、意见传递、诉求认知等层面均遭到系统性压缩，其对智慧仲裁的信任基础也随之动摇，具体表现为以下两个方面。

其一，交互渠道从双向对话退化为技术主导的单向输入。系统会把上传资料统一翻译为机器语言并加以运算，原本依赖面对面交流形成的叙事性、情境化陈述，被迫改写成符合系统解析标准的标准化文本。这种单向信息流既剥夺了当事人主动补充案件事实、澄清争议焦点的动态参与机会，也削弱了其在跨文化互动中借双向对话确认程序主体地位的可能，境外当事人的主体意志与互动地位因此难以充分彰显（李贤森和李妍妍，2025）。其二，倾听从实质性对话滑向形式化审查。尽管部分平台使用了在线听证或视频会议形式，试图增加人机之间的互动机会，但部分仲裁员对当事人的开场陈述和庭审策略的把握显著下降，使“听取意见”异化为形式审查。在由香港高等法院审查

的*Song Lihua v Lee Chee Hon*案中，[①]其中一位仲裁员虽通过视频形式远程参与线上庭审，但庭审视频显示其不断走动、频繁与他人互动、未专注屏幕，且中途坦言正赶往高铁站且信号中断，实质上未能真正参与庭审，导致境外当事人的程序参与权无法得到有效保障。当事人在互动地位无法得到保障的情况下，极易产生自身诉求未被重视的负面认知，进而会对仲裁程序的公正性产生质疑。

2. 情绪回应机制失效

智慧仲裁时代，情绪回应机制的失效并不是边角问题，它其实是互动缺失在人机交互层面的具体投影，也是当事人参与度下滑、互动地位被侵蚀的关键信号。传统仲裁程序之所以能维系信任，很大程度上靠仲裁员对情绪的感知和同理心回应。语气语调的调整、微表情的捕捉、安抚性的阐释、澄清性的追问，这些看似细碎的动作，能缓解当事人的程序性焦虑，回应其被理解、被尊重的需要，并在互动中织出信任纽带（杜焕芳和李贤森，2020）。然而，智慧仲裁系统的回应机制呈现出明显的去人格化特征，容易使境外当事人产生被边缘化的孤立感与疏离感。尤其在跨文化争议中，系统既难精准识别情绪倾向，也难提供有针对性的情绪回应，原本由仲裁员承担的情绪沟通与安抚功能便被系统性弱化。一旦情感纽带缺位，当事人对仲裁结果的评价很容易走向负面。

具体而言，高语境文化与低语境文化在信息表达上本就存在差异。东方文化背景下的当事人往往重意会，属于典型高语境文化群体，很多信息并不直接说破，而是内化于交往场景、关系和语气之中。因而，东方当事人在庭审中往往倾向间接、含蓄，核心考虑是维护关系和谐。若对方态度强硬、步步紧逼，东方当事人可能在庭审结束后才明确提出诉求，要求补充书面材料来完善主张。而西方文化背景下的当事人则重言传，属于低语境文化群体，习惯明确、直白地陈述意图。两种表达逻辑碰到一起，智慧仲裁系统若仍以标准化、单向度方式处理，就容易把含蓄视为沉默、把间接视为回避，从而加深误解，削弱互动信任（武振国和何新宇，2025）。对于此种文化差异，智慧仲裁系统仅能基于字面信息与固定算法模式提取当事人意图，难以识别高语境文化中内嵌的模糊语义、隐含诉求，亦无法认可东方当事人事后追补主张的合理性，进而易导致高语境文化当事人的申辩未被系统识别为有效抗辩，引发其对仲裁程序不公的强烈不满。相比之下，人类仲裁员可凭借自身的文化敏感性与仲裁实践经验，精准理解此类跨文化表达背后的交往逻辑与真实诉求，避免作出武断性裁判结论。针对上述问题，坎贝尔教授进一步指出，在线争议解决平台若缺乏有效情绪回应，用户好感度往往会明显下降（Campbell，2020）。换句话说，智慧仲裁中情绪回应机制的缺失，并不只是技术功能上的短板，它更像是跨文化人机沟通链条断裂的一个信号。沟通一旦断在这里，当事人对智慧仲裁的信任感和接受度自然会被拉低。

三、智慧仲裁时代人机交互障碍的理论破局

智慧仲裁中的人机交互障碍，带来的不只是人与机器之间的信任隔阂，还会引发不同法律文化和程序理念之间的碰撞。在此背景下，智慧仲裁的理论破局必须突破传统客观公平分析框架的局限，引入涵盖互动公平的新型公平理念，强调公平正义在个案中的可感知性。

（一）新旧公平理论的演进脉络

根据组织公平理论，可以梳理出新旧公平理论从“结果公平、程序公平”的二元分析结构向“结果公平、程序公平、互动公平”的三元结构的演进脉络。这个变化不只是概念层面的增加，它背后是公平评价从客观维度向主观维度的延伸。也正因如此，互动公平才可能为破解智慧仲裁中人

① *Song Lihua v Lee Chee Hon*, [2023] HKCFI 2540.

机交互障碍提供理论支撑（Muir et al.，2022）。最早的“结果公平”强调投入与产出之间具有正相关性，是一种典型的客观公平指标。此后，随着学界对程序问题的系统深耕，程序公平也成为组织公平理论的核心分析维度。然而，“结果公平、程序公平”的二元体系仍属于客观公平的分析范式，如果仲裁公平只是到此为止，便是预设客观公平可以自动生成主观公平。但如上所述，这一预设在智慧仲裁时代遭遇现实冲击，充分暴露了客观公平分析框架的局限性。特别是在跨文化争议中，智慧仲裁系统即便严格按照程序要求和法定分配规则执行，也未必能让境外当事人满意。文化背景不同、价值理念迥异，当事人表达方式一旦未被系统理解，或者系统回复不合其文化预期，他们就很可能觉得这套程序缺乏人文关怀、不尊重主体地位，甚至没有沟通诚意。这意味着，客观公平并不能自动转化为跨文化语境中当事人的主观公平体验，二者之间存在明显的转化鸿沟，而这一鸿沟正是传统客观公平框架无法解释与弥合的。

在传统客观公平分析框架难以解释主观公平体验缺失、无法弥合客观公平与主观公平间的鸿沟的现实背景下，互动公平作为组织公平理论的第三维度被正式提出，并在此后数十年间经学界持续研究，迅速发展为衡量当事人主观公平感的重要理论支柱，为破解上述理论困境提供了新的分析视角。互动公平揭示了公平感的生成并不单纯依赖客观层面的结果分配与程序设计，而是深度依赖互动过程中的信息沟通质量与人格尊重体验。在智慧仲裁时代，这一点尤为突出。系统有技术优势，形式上的程序公平要求往往能一一满足。可它很难精准识别跨文化场景里的语言差异、语气偏好和情绪表达，也缺少人类仲裁员那种同理心、情境理解与人文关怀。结果容易导致境外当事人在人机交互中反复碰壁、感到被情感忽视，于是仲裁公平的可体验性便被严重削弱。也就是说，合规不等于可感，程序正确也不等于让当事人觉得被尊重。若只看客观规则是否落实，而不看互动中是否被理解、被回应，跨文化仲裁中的公平感就可能落空（Wei & Moser，2020）。因此，立足互动公平理念破解智慧仲裁中的人机交互障碍，重建仲裁公平的可感知性和可体验性，正是破解当前实践难题的理论出路。

（二）程序公平和实体公平理论的适用困境

在智慧仲裁时代，以“程序公平、实体公平”为核心的客观公平分析框架，逐渐暴露出其解释力不足的问题。在跨文化智慧仲裁场景中，当事人的核心关切往往并不止于结果与程序本身，还包括自己有没有被公平地对待、能否感受到公平。恰恰是这种主体性体验成了客观框架难以照见的盲区，这也是智慧仲裁在人机交互层面陷入信任危机的重要原因。

“程序公平、实体公平”分析框架的适用困境主要体现在以下两个方面：从实体公平的角度看，智慧仲裁系统在客观层面确有助于提升裁决的一致性与可预测性，但并未消除当事人对仲裁过程的质疑（Rajendra & Thuraisingam，2022），系统即便在结果上做到公平，也可能因互动缺失，让当事人主观上产生强烈不公平感。从程序公平看，系统形式上可以严守程序规则，可高度自动化流程往往压缩表达空间与对话渠道，使程序参与权在跨文化情境中实质性萎缩。尤其在语言能力、法律文化与表达习惯差异明显的涉外争议中，当事人更依赖仲裁员的解释、追问和情境化沟通来弥补理解落差，进而有效参与仲裁程序。而智慧仲裁系统通常只给结构化选项和标准化回复，缺少针对具体文化语境的解释机制与沟通渠道，难以回应差异化需求。究其根本，问题不在于系统是否合规，而在于合规之后，当事人是否仍觉得被听见、被理解、被尊重。若这一层落空，客观公平再完整，其体验感也可能在跨文化人机交互中大打折扣。这并非否定客观公平的意义，而是说它不足以完全承担起跨文化仲裁的信任建构功能。在此情况下，在形式上完备的智慧仲裁程序中，当事人在体验层面的参与缺席与权利虚化，使程序公平机制丧失其应有的心理安抚与认同建构功能，进而加剧境外当事人的主观不公平感。

（三）互动公平理论的破局优势

在智慧仲裁时代，原本靠语言、文化情境支撑的双向互动，慢慢被单向的技术处理所取代。即使系统在结果公平和程序公平上做得无可挑剔，也未必能弥补因跨文化沟通不畅、文化误读和情感忽视造成的主观公平断裂。正是在这个意义上，互动公平理论的价值才凸显出来。互动公平关注的不是单纯的结果和程序，而是主观公平。这使其跳出了传统客观公平框架的局限，为破解人机交互障碍提供了有针对性的分析工具。

一方面，互动公平理论直面智慧仲裁的核心危机，即当事人的主体性缺失问题。按组织公平理论的经典分类，互动公平可分为“信息公平”与“人际公平”。前者强调决策过程中信息是否被充分披露，后者关注互动中当事人的人格尊严是否被尊重、主体地位是否被明确认可。放到跨文化仲裁场景中，不同法律文化传统、语言表达习惯与价值理念，对信息公平、人际公平的具体内涵和实现方式有显著差异。这种差异若被系统忽视，便会成为跨文化人机交互障碍的重要诱因。当智慧仲裁系统仅依据字面语义进行意图识别与信息处理，缺乏对不同文化表达习惯的适配能力时，东亚当事人的含蓄表达就可能被误判为未明确提出主张或放弃权利，进而引发当事人的主观不公平感与对系统的不信任。可以说，互动公平集中关注当事人在跨文化沟通过程中的心理体验、身份确认与情感满足，因而能够更精准地解释“形式公正却体验不佳”“结果合理却不被接受”的跨文化仲裁困境，这也是其相较于传统客观公平框架的核心优势所在（塞拉，2024）。

互动公平的价值，不只在于解释因语言差异、文化冲突与沟通不畅所造成的互动断裂，更在于通过信息公平、人际公平的制度化设计，重新连接当事人与智慧仲裁系统之间的信任关系，因此能够成为破解人机交互障碍、推动智慧仲裁规范化发展的关键理论基石。

四、智慧仲裁时代人机交互障碍的克服路径

智慧仲裁时代，人机交互障碍的克服路径不能只靠技术修补。原本依赖仲裁员的文化理解、语言解释与情感回应来维系的双向沟通机制，如今被压缩为标准化信息输入与算法自动化输出的单向流程，人文关怀和情境适配性随之流失，当事人的主观不公平感也被放大。要走出这个困境，互动公平理论所强调的信息公平与人际公平，恰好能提供制度化的解决方案。

（一）构建信息公平实现机制

信息公平是实现互动公平的重要维度，其核心在于化解跨文化语境下人机之间的信息不对称。在智慧仲裁中，文化价值差异以及表达方式上的跨语境误读，往往比单纯技术故障更直接地阻断沟通链条。因此，信息公平的实现不能只理解为技术透明化，它还要求算法逻辑与数据应用在多元文化和多语言背景下能够被清晰理解、充分解释与有效沟通，使技术理性与跨文化沟通需求真正适配。具体而言，智慧仲裁系统需要聚焦算法原理的可理解性与数据来源的可追溯性两大核心，为破解互动信息缺失问题提供制度化支撑。

1. 推进算法可视化

在智慧仲裁场景中，算法黑箱问题并非孤立存在，而是与语境差异深度交织，形成双重信息壁垒。技术黑箱叠加跨文化认知差异，人机之间原本就脆弱的有效沟通链条很容易被切断，程序不信任感也随之产生。值得注意的是，当事人真正不满的，未必是裁决结果在实体上不公正。更常见的情况是，他们看不懂系统作出裁决的推理过程和判断依据，于是形成决策不透明的主观认知，并进

一步产生智慧仲裁不公平的体验。面对这种情况，为契合互动公平的核心要求，智慧仲裁系统就不能只把自己当作运算工具，而应承担必要的转换义务和解释责任，把算法转化为可被人类理解，甚至具有可视化性质的表达（武振国，2025）。

为了有效支撑上述目标，智慧仲裁系统有必要改造源代码，构建算法可视化机制，以匹配不同参与主体的认知能力与程序需求（贺小勇和刘真宇，2026）。一方面，面向当事人，应建立通俗化解释机制。图形化操作界面、流程动态示意、情境化案例说明等多元形式，可以把抽象算法决策转译为具有清晰因果结构、又符合不同文化认知习惯的通俗表达，使不同法律文化背景的当事人至少能理解系统的决策路径。另一方面，面向仲裁员，则需要专业化解释机制。通过通俗易懂的方式把技术细节呈现出来，为仲裁员理解算法逻辑提供依据。同时，还要在算法语言、专业语言与通俗语言之间建立逻辑映射关系，让不同文化背景的参与者围绕同一决策逻辑进行有效沟通（王晓霞，2024）。此举不仅能够避免算法可视化过程陷入二次黑箱化的困境，更能确保仲裁员和当事人都能在自身认知能力范围内实现对算法逻辑的有效理解与顺畅沟通，并消除信息不对称，强化境外当事人的程序认同感与主观公平感。

2. 保障数据可溯源

在智慧仲裁里，数据溯源机制是一种制度工具，用来重新安放当事人的主观公平感，并在跨文化误解容易滋生的地方，给沟通留下可验证的线索。由于不同国家的法律文化、证据规则与商业习惯存在显著差异，当智慧仲裁系统以标准化的数据模型进行事实抽取与权重分配时，极易引发境外当事人的制度性焦虑。境外当事人往往担忧系统把他国法律环境、文化背景下的案例逻辑和证据规则，直接套到自己的案子里，形成一刀切的隐性推定。这种情况一旦成立，就会背离当事人合理预期与认知习惯，裁决结果的可接受性自然被削弱。

基于上述实践困境，数据溯源机制成为智慧仲裁系统优化完善的重要方向。通过引入以区块链为代表的可信计算技术，智慧仲裁系统可以构建全流程、可验证、不可篡改的数据溯源体系，实现数据应用的全程留痕（付大学，2026）。更重要的是，不可篡改的溯源结构还能在技术层面提升仲裁证据链条的完整性、真实性与合法性，使技术理性与法律理性不至于各行其是。而一旦溯源链条出现断裂、记录异常或语境不匹配，系统便可预判为数据不适格，仲裁庭也能据此行使裁量权，排除风险数据，确保数据应用的合理性与适配性。

值得注意的是，在智慧仲裁系统中，数据可溯源绝非单纯的技术透明化工具，它更应被定位为一种跨文化沟通与信任建构机制。其核心价值在于帮助不同文化、不同法系背景的当事人理解机器决策如何形成，从而在多元文化语境之间搭建起共同的信息认知基础，化解跨文化信息误解。这个过程，也恰好契合哈贝马斯交往行动理论所强调的主体间性。不同文化背景的参与者，只有在共享可理解、可验证的信息结构基础上，才能通过理性沟通、平等对话，形成对智慧仲裁的共同认同（孙绍勇，2022）。可以说，数据可溯源在智慧仲裁中承载着制度正当性与跨文化信任重建的双重功能，其通过重塑主观公平的生成机制，能够推动仲裁正义在跨文化理性交往中的持续演进与完善。

（二）完善人际公平实现机制

人际公平作为互动公平的核心维度，要求智慧仲裁系统以充分倾听、有效回应、尊重人格为基本准则，平等对待仲裁当事人。在智慧仲裁时代，不同法律文化、沟通习惯与表达方式的差异，使当事人更容易在人机交互场景中陷入主体地位失衡的弱势处境。部分当事人因语言风格、表达逻辑

或礼貌规范与系统内置标准化语料库存在适配偏差，其陈述容易被智慧仲裁系统误读，进而使当事人产生被漠视的负面感知。基于此，本文秉持人机协同共治基本逻辑，从多元互动渠道扩容、拟人化回应机制建构等路径出发，系统性破除跨文化语境下人机交互的主体地位壁垒，推动涉外智慧仲裁人际公平规范化、制度化落地。

1. 拓宽多元化互动渠道

跨法系立法理念、沟通习惯与思维模式的差异是造成境外当事人诉求表达无法被智慧仲裁系统完整识别、精准吸纳的核心原因。为纾解这一困境，应着力拓宽人机交互渠道，以技术手段增强不同文化表达方式在程序中的可感知性。智慧仲裁系统应搭建双向闭环沟通通道，实现直接对话衔接，保障不同法域、不同表达习惯的主体意见均可在智慧仲裁系统中得到完整呈现（姚梦莹，2022）。为此，智慧仲裁系统可以引入多模态输入接口，搭建双向闭环沟通渠道，强化当事人与仲裁庭之间的直接对话衔接，保障不同法域、不同表达习惯的主体意见均可在数字化仲裁程序中完整呈现、有效流转。智慧仲裁系统还应设立智能答疑窗口，针对跨文化表达歧义、法律术语差异、语义理解分歧等问题，自动生成回复并推送给当事人和仲裁员，及时完成观点补正与事实厘清。

为了避免智慧仲裁系统的工具化属性将仲裁法意义上的“程序倾听”窄化为单纯的数据抓取与机械解析，仲裁庭有必要强化对智慧仲裁系统的控制权与实质审查权。这不仅是实现互动公平的需要，也是仲裁庭行使仲裁权的必然要求。仲裁庭作为仲裁权唯一合法行使主体，若无法对智慧仲裁系统抓取的案件信息、筛选的关键诉求、整合的交互内容开展实质性核验与法律解释，便不必将系统输出的结论纳入裁决中（杜焕芳，2025）。在这一背景下，新《仲裁法》第67条规定，仲裁庭在作出裁决时应当说明理由，这进一步强化了仲裁庭在此方面的责任。质言之，唯有将技术赋能的多元交互渠道，与仲裁庭的核心职权深度绑定，方能符合“倾听当事人意见”的规范内涵，避免当事人在智慧仲裁中沦为数据供给客体的被动情况。

2. 引入拟人化回复机制

在跨文化语境下，不同礼貌规则与表达习惯的差异往往会放大智慧仲裁系统语言风格所带来的误解风险，加剧境外当事人的程序弱势地位。在此情况下，引入拟人化回复机制，便有了现实针对性。它不是让系统简单模仿人类的表达，而是让回应更懂情感、更有适配度，从而提升人机交互的共情与包容。当然，此种机制的落地实施不能停留于抽象设想，尚需具体的技术支撑与制度配套。它需要跨语种、跨文化双向校准的情绪识别算法，还要结合自适应深度学习，尽可能捕捉当事人陈述中隐含的情绪倾向和诉求立场。在此基础上，系统再依据情感生成模型，合理调节回复语气，补充情境化法律解释，增设安抚性表述，在保证法律表述严谨准确的前提下，让智能回复带一点人文温度。通过柔性化、适配化的语言输出，智慧仲裁系统有望缓解人机交互中的疏离感，增强当事人对智慧仲裁的认可度与信任度（周斌和李守石，2021）。

拟人化回复的目标不止在于改变表达方式，更在于重构人机交互的对话结构，改变跨文化语境下当事人的弱势地位。通过情感识别算法与自适应生成模型的协同作用，系统可以从被动的应答工具变成主动的共情对话者，让当事人切实感到自己的意见被倾听，而不仅仅是被处理。换句话说，系统要回应的不只是一句法律结论，还有当事人身处陌生程序中的不安与期待。当然，为确保拟人回复不逾越法律文书的客观中立底线，系统应嵌入合法性审查子模块，对情感化表述进行合规性过滤。遇到情绪激动或事实误解类陈述，可以先以标准化法律语言确认核心事实，再以拟人化语气附加解释性、安抚性表述，使法律理性与交流感性不至于互相打架。通过子模块的嵌入确保拟人回复机制能够在不牺牲法律严谨性的前提下，实质性提升当事人对智慧仲裁系统的信任度和接受度。

（三）建立人机互动纠偏机制

在智慧仲裁系统中，即便信息公平与人际公平在技术设计层面已初步落实，智慧仲裁系统在跨文化语境下仍可能因技术偏差而输出偏离实质公正的结果。比如，某种在特定文化中具有合理性的情绪表达，被系统误判为攻击性语言，当事人陈述在系统中的可信度就会受到影响。这种技术性误判若不及时矫正，信息层面的偏差便会传导至人际层面，造成人格尊严受损与程序地位弱化，甚至引发仲裁责任。因此，有必要引入人机互动纠偏机制，把它作为连接信息公平与人际公平的制度桥梁，即时矫正系统输出的失当结论，使二者从各自独立的目标状态，升华为相互协同的运作过程。

至于纠偏机制的设计，主要包括三条路径，其中前两条尤为关键。第一，通过预设语义规则实现技术自动修正，对明显源于文化表达差异的系统误判进行实时筛查与调整。这条路径侧重信息公平的补强，同时为人际公平提供准确的认知前提。第二，充分尊重当事人意思自治，允许境外当事人通过合意选择、限制或禁用特定智慧仲裁系统，或要求使用符合其语言与文化背景的交互模块（Moreira & Zhang，2025）。例如，依据2025年《中国国际经济贸易仲裁委员会关于在仲裁中使用人工智能技术的指引（试行）》第3.1条之精神，当事人有权对智慧仲裁系统的使用方式作出约定，也可就仲裁中使用智慧仲裁系统的披露等事项作出约定。该路径赋予当事人程序选择权，使人际公平通过意思自治获得实质性保障。第三，仲裁庭作为程序主导者，在主动发现或受理当事人针对智慧仲裁系统的跨文化误判异议后，不能只做旁观者，而应协调技术力量对系统进行语境校准，必要时还得行使最终决定权。这种人工纠偏行为既是对算法偏见的必要矫正，也是维系仲裁权依法独立行使、实现信息公平与人际公平深度协同的制度保障（胡玉鸿，2025）。只有这三条路径联动起来，人机互动纠偏机制才能在个案中持续弥合信息处理与人机互动之间的裂隙，使智慧仲裁系统同时满足技术理性与人文关怀的双重要求。

五、结论

在智慧仲裁时代，迈向互动公平是续造当事人主观公平感、克服人机交互障碍的关键路径。智慧仲裁确实提升了程序效率，但也可能在不同语言、法律传统与沟通习惯之间制造新的信息不对称，使当事人在系统中难以获得充分倾听、有效回应与文化尊重，主观公平体验随之被削弱。这类困境，往往并非源于客观裁决结果不公，而是源于跨文化语境下互动体验的缺失。未来主观公平的制度建构，应超越传统以“结果公平、程序公平”为中心的客观公平分析框架，转而将互动公平中的信息公平与人际公平确立为核心分析维度。具体来说，信息公平为人际公平提供事实认知的前提，人际公平为信息公平划定价值尊重的边界，二者再通过人机互动纠偏机制实现动态协同。也只有把新的公平保障理念真正嵌入智慧仲裁系统，技术才不至于凌驾于仲裁的公正本质之上，智慧仲裁才能在数字化浪潮中向以人为本、文化包容的现代化转型。

参考文献

[1] 杜焕芳,李贤森.仲裁员选任困境与解决路径——仲裁员与当事人法律关系的视角[J].武大国际法评论,2020(2):39–59.
[2] 杜焕芳.仲裁法应有的包容性开放性融通性[J].政法论坛,2025(1):71–80.
[3] 付大学.可信数据空间的法律定位与制度构建[J].法学家,2026(4):62–74.
[4] 贺林平.探索提供更公平高效的仲裁服务[N].人民日报,2023–11–17(11).
[5] 贺小勇,刘真宇.源代码条款的碎片化困境及中国因应[J].上海对外经贸大学学报,2026(1):82–97.
[6] 胡玉鸿.人工智能时代法律职业“变”中的“不变”[J].法律适用,2025(10):113–130.
[7] 李贤森,李妍妍.人工智能赋能国际商事仲裁的风险应对与新型人智混合决策体系构建[J].商事仲裁与调解,2025(3):25–53.
[8] 刘一帆.人工智能应用于国际商事仲裁的限度[J].华南理工大学学报(社会科学版),2025(3):25–37.

[9] 邱威棋.人工智能商事仲裁的隐忧与治理[J].华南理工大学学报(社会科学版),2025(3):10–24.
[10] 塞拉.精简司法:在线法院如何应对本人诉讼的挑战[J].司法改革论评,2024(2):5–12.
[11] 宋连斌.司法支持仲裁的底层逻辑解析[J].政法论坛,2025(1):81–88.
[12] 孙东山.从“群体正义”到“个体正义”:组织公平转向的实践逻辑[J].清华大学学报(哲学社会科学版),2020(3):182–190.
[13] 孙绍勇.交往理性的主体间性向度解析及当代审思——以哈贝马斯交往范式与交往实践旨趣为论域[J].山东社会科学,2022(7):57–65.
[14] 王晓霞.数字司法的基本特征与未来构想[J].数字法治,2024(6):43–56.
[15] 吴维锭.人工智能仲裁裁决:兴起、法律效力和法律回应[J].海南大学学报(人文社会科学版),2024(5):196–206.
[16] 武振国,何新宇.社交距离差序格局下仲裁员公正性的制度重塑[J].商事仲裁与调解,2025(1):30–44.
[17] 武振国,周健.Generative AI生成司法决策的可靠性困境及其应对[J].天津大学学报(社会科学版),2024(6):512–520.
[18] 武振国.人工智能辅助仲裁庭审阅证据的可控性研究[J].上海对外经贸大学学报,2025(2):85–95.
[19] 项军.客观“获得”与主观“获得感”——基于地位获得与社会流动的视角[J].社会发展研究,2019(2):135–153.
[20] 姚梦莹.他山之石:人工智能辅助仲裁借镜智慧法院建设的“未来仲裁”探索[J].商事仲裁与调解,2022(6):38–48.
[21] 张爱军.“算法利维坦”的风险及其规制[J].探索与争鸣,2021(1):95–102.
[22] 张圣翠,田洋.人工智能在商事仲裁中的应用：价值、困境与路径[J].上海财经大学学报,2024(2):122–136.
[23] 周斌,李守石.还原:跨文化传播过程中的文化落差弥合[J].社会科学文摘,2021(7):118–120.
[24] Campbell, R. W.,“Artificial Intelligence in the Courtroom: The Delivery of Justice in the Age of Machine Learning”, *Colorado Technology Law Journal*, 2020, 18(2): 323–350.
[25] Moreira, J. I., and J. Zhang,“ChatGPT as a Fourth Arbitrator? The Ethics and Risks of Using Large Language Models in Arbitration”, *Arbitration International*, 2025, 41(1): 71–84.
[26] Muir, C. P., E. N. Sherf, and J. T. Liu,“It’s Not Only What You Do, But Why You Do It: How Managerial Motives Influence Employees’ Fairness Judgments”, *Journal of Applied Psychology*, 2022, 107(4): 581–603.
[27] Rajendra, J. B., and A. S. Thuraisingam,“The Deployment of Artificial Intelligence in Alternative Dispute Resolution: The AI Augmented Arbitrator”, *Information & Communications Technology Law*, 2022, 31(2): 176–193.
[28] Wei, D., and G. Moser,“Human Arbitrators (The Undisputed Champion) v Robots (The Challenger)”, *Hong Kong Law Journal*, 2020, 50(1): 215–241.

【作者简介】武振国：北京科技大学文法学院副教授。研究方向：仲裁法。

Human-machine Interaction Barriers in the Era of Smart Arbitration and Their Overcoming

WU Zhen-guo

(School of Humanities and Social Sciences, University of Science and Technology Beijing, Beijing 100083, China)

Abstract: With the widespread application of smart arbitration systems, the interactive status of humans in arbitration decision-making has been significantly weakened, and human-machine interaction barriers have become increasingly prominent. In particular, when a smart arbitration system generates outcomes in a single step, human-machine interaction barriers substantially compress the interactive experience to which parties should be entitled. The resulting “loss aversion” intensifies the trust gap between humans and machines, leading parties to develop stronger doubts about and resistance to identical objective outcomes. The traditional objective fairness analytical framework centered on “distributive justice, procedural justice” is unable to effectively respond to the trust crisis caused by human-machine interaction barriers, because it fails to recognize that interactive experience is an important vehicle through which the people genuinely feel fairness and justice in every judicial case. Grounded in the genealogical evolution of organizational justice theory, this paper takes “interactional justice” as the core point of entry for overcoming human-machine interaction barriers in smart arbitration. The informational justice and interpersonal justice embodied in interactional justice can overcome human-machine interaction barriers in smart arbitration, reshape the parties’ interactive experience and emotional identification, and thereby effectively enhance human trust in and acceptance of smart arbitration.

Keywords: interactive fairness; smart arbitration; algorithmic black box; artificial intelligence

（责任编辑：马莹）