·公共管理与行政管理·

人工智能公众风险感知的类型与演化
——基于全球视角的分析

□李智超[1]　李鹿嘉[2]

[1. 上海交通大学　上海　200030；2. 国家环境保护河流全物质通量重点实验室　北京　100871]

[摘　要]　分析全球公众对人工智能风险的感知类型及其时空演化特征，探讨影响其感知的因素，并为全球可信赖人工智能的发展提供实证基础。通过对社交软件中430万帖子的文本语义分析和时间特征分类，构建全球数据库，量化了不同区域的AI风险类型。研究采用凸包分析、地理信息系统以及集成机器学习等方法，揭示了全球公众风险感知的区域异质性、影响因素及时空演变特征。全球范围内公众对人工智能风险感知呈现显著的区域异质性，受创新技术扩散效应影响形成不同感知类别。而个体特质、政治信任与政府效能，对公众风险感知具有重要塑造作用。在动态维度下感知类别呈现共变逻辑，并呈现由单一向多元格局转变的趋势。

[关键词]　风险感知；人工智能；时空格局；全球视角

[中图分类号] D630.1；X922.2　　[文献标识码] A　　[DOI] 10.14071/j.1008-8105(2024)-5011

引言

随着现代化进程中人工智能技术力量激增，社会面临由技术进步带来的“意外后果”的风险累积。同时，伴随着所谓的“属意副作用”[1]，潜在的风险和威胁得到了前所未有的释放，从而催生了一种新的现代性形态——风险社会[2]。在风险社会背景下，公众对人工智能风险的感知展现了社会对于新兴科技的适应性和敏感性。这种感知，间接促使政策制定者、企业家、技术开发者以及学界对人工智能风险的关注，保障个人隐私、确保数据安全、应对就业挑战等已经成为重要的政策议题。因而，准确把握并真实呈现人工智能公众风险感知对于建立一个负责任和高包容性的人工智能发展环境至关重要。

人工智能公众风险感知所蕴含的理论与实践价值催生了一系列研究，综合而言，形成以下两条脉络。第一，基于人工智能风险的感知风险类型。主要是从技术、伦理与社会三个维度把握人工智能的风险，技术风险如数据保护漏洞和算法误判，直接关联到人工智能应用的安全性和可靠性；伦理风险聚焦于AI技术潜在的道德争议，聚焦于智能决策中的责任归属；社会风险感知则聚焦于自动化取代人力劳动带来的社会问题。第二，人工智能风险感知的影响因素。基于个体自主论、文化塑造论及资源分配论与社会信任论四种解释路径，相关经验研究发现：个体自主的有效参与、所在文化的价值观、经济资源是否均衡分布，以及社会对新兴技术的信任水平，都会微妙地调节人们对人工智能的潜在风险感知。目前，人工智能公共风险感知的相关研究，大多基于有限样本，研究方法也多采用问卷调查和心理量表测量结合的方法。尚缺乏基于大规模样本对人工智能公众风险感知的探索。此外，对人工智能公众风险感知的时空演变特征也有待深入研究。

[收稿日期]　2024－06－19

[基金项目]　国家自然科学基金面上项目（71974057，72374132）；上海市教育委员会和上海市教育发展基金会“曙光计划”（21SG49）.

[作者简介]　李智超，上海交通大学国际与公共事务学院、应急管理学院长聘副教授，博士生导师；李鹿嘉，国家环境保护河流全物质通量重点实验室研究助理.

[引用格式]　李智超, 李鹿嘉. 人工智能公众风险感知的类型与演化——基于全球视角的分析[J]. 电子科技大学学报（社科版）, 2024, 26(6): 21-33. DOI: 10.14071/j.1008-8105(2024)-5011.

[Citation Format]　LI Zhi-chao, LI Lu-jia. Types and evolution of public risk perception of Artificial Intelligence: a global perspective analysis[J]. Journal of University of Electronic Science and Technology of China(Social Sciences Edition), 2024, 26(6): 21-33. DOI: 10.14071/j.1008-8105(2024)-5011.

因此，本文聚焦于测度与解析全球公众对人工智能风险感知的类型与演化，边际贡献主要为：第一，构建全球风险感知数据集，测度全球地区间的风险感知差异，深化对人工智能风险感知空间特征的理解；第二，基于多指标多因素，实证检验全球公众人工智能风险感知的影响因素；第三，呈现并分析全球人工智能风险感知的时空演变过程。

本研究的价值在于提供了宏观视角下对人工智能公众风险感知的深入分析，这有助于更全面地理解人工智能技术在不同社会和文化背景下的影响。随着人工智能技术的不断进步，公众的风险感知也随之演变。通过大规模样本的分析，本研究尝试揭示这种感知的多样性和复杂性。强调在全球化背景下，对人工智能风险进行综合考量的重要性，有助于各主体认识到在不同社会环境中人工智能技术可能引起的不同反应和影响，从而促进更加细致和有针对性的政策和措施的制定。

一、文献综述

（一）人工智能公众风险感知：概念阐释、理论基础与感知风险类型

在人工智能语境下，风险感知被理解为个体或群体对于人工智能技术可能引发的风险因素或致灾因子的认知和情感反应[3]。学界对这一概念的研究主要分为两大流派：技术接受模型流派与心理测量流派[4]。技术接受模型流派侧重于评估公众对新兴技术的接受度，主要基于感知有用性和感知易用性两个维度[5]；而心理测量流派则侧重于通过调查工具量化风险的内因特征，并强调认知情感层面的伦理风险类别[6]。

学界对公众人工智能风险感知类型的研究主要从三个维度进行探讨。首先是技术感知维度，研究者运用技术接受模型及其扩展模型来探索公众对人工智能使用中的技术风险感知，如算法决策和创新控制等[7]。这些风险的根源在于人工智能技术的“黑箱性”和不透明性，导致用户难以理解算法的决策过程，限制了决策的解释能力，进而影响公众的知情权[8~9]。其次，是伦理层面的风险感知，研究者通过心理测量模型量化了隐私侵犯和舆论操纵等伦理风险类别。人工智能技术的发展可能违背了“服务于人”的初衷，引发技术异化现象[10]，在智能家居和智能机器人领域引发隐私问题，侵犯了个人的自主能力。最后，社会风险感知关注人工智能对就业市场和社会稳定的影响，技术进步导致的就业结构变化引起了公众对未来劳动市场的担忧，尤其是对低技能工人的就业机会和职业转换的挑战[11]。

然而，现有研究在风险感知的测量和理解上存在局限。不同研究对风险感知的定义和测量方法差异较大，导致研究结果难以比较和整合。此外，某些研究可能因选择偏差而过分侧重于某一类风险感知，忽视了其他类型的风险，这限制了研究结果的全面性和适用性。为弥补这些不足，需要在现有理论和实证研究的基础上，进一步开展工作，以期构建一套全面、可靠且有效的度量体系，精准把握公众对人工智能风险的认知，为科学决策与管理策略提供实证依据。

（二）公众风险感知的影响因素

关于公众风险感知的影响因素，可以甄别出个体自主论、文化塑造论、资源分配论及社会信任论四种解释路径[12]。个体自主论着重强调个人选择在风险感知形成中的核心作用，涵盖性别、年龄、收入、婚姻状况、居住地、教育水平和自我效能等个体特征。然而，风险感知的形成并非独立于社会环境、教育和收入等个人属性，实际上也映射了社会结构对个体的影响[13]。研究指出，那些拥有更高教育和收入水平的个体，由于具备更深层次的反思性思维，可能对人工智能的风险有更为敏锐的感知；而教育和收入水平较低的个体，则可能因社会秩序的复杂性而感到困惑，难以预测人工智能技术的风险[14~15]。个体在教育资源和财富水平上的差异，导致他们处于不同的风险环境，形成了群体间风险感知的差异[16~17]。因此，学者们进一步探究了文化塑造论和资源分配论，以更深入地理解社会和结构层面如何影响公众的风险感知，从而为全面把握公众对人工智能风险的感知提供了更为丰富的视角。

文化塑造论强调社会规范在塑造个体和群体对风险的感知和反应中起关键作用。不同文化背景下，人们对相同风险的理解和反应可能大相径庭。社会规范、宗教信仰、价值观等因素通过社会化过程内化到个体中，形成迥异的风险评估和应对机制[18]。例如，在伊斯兰教中，强调自然与人的和谐共生，任何可能破坏这种和谐状态的技术或行为都可能被视为高风险；在自由主义文化之下，对大数据和监控技术持谨慎态度，人工智能在新闻过滤和社交媒体监控中的应用，可能被视为对信息自由流通的威胁，引起较高的风险感知。

不同于文化环境论，资源分配论从社会结构的角度审视风险感知，认为资源分配的公平性对社会

成员的风险感知具有显著影响[19~20]。这一视角认为，当社会资源得到公平分配时，个体和社群更有可能具备应对未来潜在风险的能力。技术资源的分配不均，如技术鸿沟，可能导致某些群体对技术风险有着更加敏感和负面的看法，因为他们可能感到被边缘化，无法享受到技术进步带来的红利[21]。社会信任论进一步拓展了对风险感知的理解，将信任视为一种关键的社会资本[22]。这种理论视角认为，社会信任能够显著影响个体如何识别和反应风险。在一个信任度较高的社会环境中，个体更倾向于接受新技术，并对相关风险持有更积极的态度。这种信任建立在政府和机构的透明度、政策的连贯性、法规的质量和执行力上。政府的有效应对和政策执行能够增强公众对人工智能技术的信心，减少不确定性，从而降低风险感知[23]。高质量的法规提供了明确的指导和执行标准，这不仅提升了公众对技术发展方向的信心，也增强了对政府监管能力的信任[24~25]。

尽管现有研究提供了对公众人工智能风险感知影响因素的重要理解，但仍存在一些局限。研究往往侧重于分析单一因素，采取静态分析方式，缺乏对全面性考量。此外，研究多基于有限样本或区域层面，缺乏全球视角下的时间趋势和演化模式分析。为了克服现有研究的局限性，本研究构建了一个综合性分析框架，将个体自主论、文化塑造论、社会信任论和资源分配论整合在一起。事实上，社会信任论与其他三种解释路径存在内生性逻辑。个体特质在很大程度上受到社会信任水平的影响，而社会信任论与文化塑造论存在着交互作用。文化背景中的社会规范和价值观塑造了社会信任水平，而高社会信任又增强了文化一致性和稳定性。基于此，图1整合了相关学者基于四种路径讨论的对于公众风险感知的影响因素。此框架揭示了个体特征、文化价值观、经济状况、社会信任和资源分配等因素，如何共同塑造公众对人工智能风险的感知。通过这种多维度的视角，本研究旨在提供一个更为全面和系统的理解，以反映不同文化和社会背景下公众风险感知的复杂性和动态变化。这不仅有助于揭示影响公众风险感知的关键因素，也为人工智能政策的制定提供了理论支持和实证依据。

二、研究设计

（一）数据来源

风险感知作为一个多维度概念，涉及个体或群体在面对潜在威胁时的认知与情绪反应。这种感知不仅基于经济、社会、文化和心理等多个层面，而且体现在人们在预见及应对风险时的认知评估、情绪调整和决策选择能力上。鉴于风险感知的多维性和抽象性，相关变量的衡量自然呈现出复杂性。然而，现有研究在测度公众对人工智能风险感知时，常受限于数据的地域性和选择性偏差，这些研究往往只能提供局部视角，难以捕捉全球范围内的公众情绪和态度。

为了克服这些局限，本研究采取了一种创新的数据收集方法。从全球广泛使用的社交媒体平台，如Twitter和Reddit，抓取关键词并收集数据集，共囊括了430万个帖子的信息。这一方法使我们能够

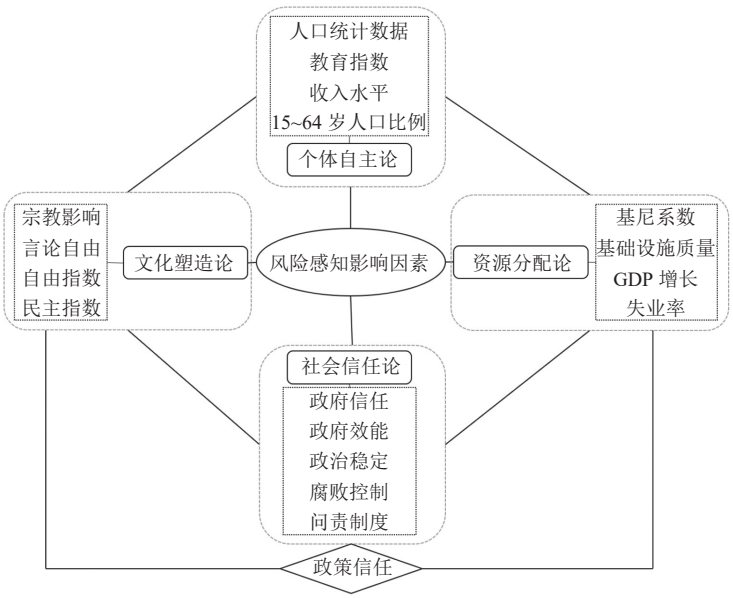

图 1　基于风险感知逻辑链条的四种解释路径

建立一个全球性的数据库[①]，全面反映不同公众对人工智能风险的感知。本研究构建的数据库共包含了29项数据特征值，这些特征不仅覆盖了位置信息（如经纬度）、所属地理单元（例如国家和大洲），还包括了与导出数据相关的内部标识符和测量值，为量化分析提供了丰富的信息。

与现有研究相比，本研究的数据来源更加广泛，样本规模更大，能够更真实地反映全球公众的风险感知。社交媒体平台的数据提供了实时、动态的视角，这有助于避免传统问卷调查中常见的时间滞后和社会期望偏差问题，从而更准确地捕捉公众情绪的即时反应。为全面评估并量化影响人工智能风险感知的因素，本研究在现有学术研究和相关文献综述的基础上，选用了基于全球公众风险感知的统计数据进行变量选择。所选变量旨在综合反映技术特性、社会经济条件、文化背景以及法律政策框架等多维度因素，这些因素共同塑造了公众对人工智能风险的感知。通过这一策略，能够更深入地理解影响公众风险感知的关键因素，并为未来的政策制定和科学研究提供坚实的数据支持。本文选取的变量及数据来源，如表1所示，展示了从多元视角捕捉和量化公众对人工智能风险的感知。

表 1　变量及其数据来源

解释路径	指标名称	数据来源
个体自主论	人口统计数据	全球数据实验室（https://globaldatalab.org）
	教育指数	
	收入指数	
	人类发展指数（HDI）	
	15~64岁人口比例	世界银行（https://databank.worldbank.org/home.aspx）
文化塑造论	宗教影响	Harvard Dataverse（https://dataverse.harvard.edu/）
	民主指数	
	自由指数	
	言论自由	世界银行（https://databank.worldbank.org/home.aspx）
资源分配论	基尼系数	
	失业率	世界银行（https://databank.worldbank.org/home.aspx）
	GDP增长	
	贫困率	
	夜间灯光数据	地球观测组织（https://eogdata.mines.edu/download_dnb_composites.html）
	基础设施质量	Harvard Dataverse（https://dataverse.harvard.edu/）
	次国家区域数据	全球数据实验室（https://globaldatalab.org）
社会信任论	政府信任	Harvard Dataverse（https://dataverse.harvard.edu/）
	政府效能	世界银行（https://databank.worldbank.org/home.aspx）
	法治	
	政治稳定	
	腐败控制	
	暴力/恐怖活动	
	问责制	

（二）数据处理与模式识别

为精确区分公众对人工智能风险的感知类型，本研究采用了自然语言处理（NLP）技术对文本的语义及时序特征进行了注释与分类。使用了spaCy、Gensim、FastText等Python工具包，以及Google的Perspective API等系列分析工具。数据标注的技术路线如图2所示。通过文本预处理、词性标注和语义结构分析、实体模式主题归纳以及有害内容检测和过滤的步骤，从帖子中提取并分析特征主题。经检测与过滤筛选后，对提取主题聚类。最终，构建了一个全面的公众人工智能风险感知数据库，以提供实证层面的数据支持。

鉴于网络空间的匿名性和虚拟性，公众对人工智能风险的感知容易受到混淆和误解，特别是当涉及创新控制与决策支持等细微区分时。为了解决这一挑战，引入了开源的IP地理定位数据库IP2Location™ Lite，通过帖子的IP地址，将内容映射到具体的地理位置，包括国家、地区、城市以及经纬度等信息。尽管IP地理定位的精度达到了76.4%，但本文还是排除了那些通过代理服务器、匿名网络和虚拟私人网络（VPN）隐匿真实身份和地点的帖子，以减少由匿名IP地址带来的不确定性。

基于上述分析流程，本文聚类的公众对人工智能风险感知子类型分别为信息透明、创新控制、伦

图 2　标签提取的技术路线

理道德、辅助决策、就业市场及隐私安全，这与现有文献的分类也较为一致[22, 26~29]。本文的人工智能风险感知定义、区分及关键词提取如表2所示。

通过对不同人工智能风险类型的感知程度进行加权处理，本文量化了这些风险对社会或个体产生的影响程度。针对每一特定地理位置，将该地所有帖子关于人工智能风险的加权评分进行综合，从而计算得出一个表征公众对人工智能风险感知的综合性指标。其中w_i是地理位置（10公里）网格下子类型的权重，s_i是地理位置（10公里）网格下对应的影响程度评分。

$$\text{Risk}_{\text{perception}} = \frac{\sum (w_i \times s_i)}{\sum w_i}$$

为了在全球范围内进行比较，本文将每个地理位置的公众对AI风险感知综合指数进行了归一化处理，并利用最小—最大缩放方法，通过地图上的颜色渐变来表示不同的风险感知指数级别，从而清晰地识别出高影响程度和低影响程度的区域。

（三）研究方法

步骤一：全球公众人工智能风险感知类别分布的可视化分析及外推检验。

利用地理信息系统（GIS）技术，将全球公众对人工智能风险感知的分布特征进行可视化展现。为探索和定量评估不同风险感知的时空分布及其区域异质性，本文采用了多维数学凸包分析方法。具体而言，通过Andrew's Monotone Chain Algorithm，确定了全球范围内公众风险感知相关的地理空间数据点的最小边界。这项分析不仅帮助理解风险感知活动的地理集中区域，还揭示了数据收集的潜在盲区，从而评估了样本的覆盖率和代表性。例如，本文发现欧美地区的风险感知数据点较为密集，而非洲某些地区凸包覆盖的不足则暗示了数据代表性的潜在偏差。

步骤二：识别并量化多因素对公众人工智能风

表 2　人工智能风险感知类型：定义与区分及关键词提取

风险感知类型	定义	区分	关键词
信息透明	涉及人工智能系统在数据处理、存储和传输过程中，信息的真实性、系统操作的透明度以及个人隐私保护方面的风险感知	技术运行的开放性和用户的知情权	信息真实性、系统透明度、隐私保护、数据泄露、数据加密、访问控制、信任机制、透明度报告、审计追踪
创新控制	对新技术或新方法在发展和应用过程中能否被有效控制和管理的担忧，包括潜在的不确定性和管理挑战	主要关注新技术带来的不确定性、潜在的风险以及对这些风险的管理和控制能力	新技术、技术监管、不确定性管理、创新风险、控制措施、技术评估、风险缓解、治理框架
伦理道德	涉及人工智能技术和行为在伦理和道德层面上的风险感知，包括是否符合伦理规范和社会道德标准	伦理问题涉及技术应用是否符合伦理规范；道德关切则更侧重于是否符合社会普遍认可的道德标准	伦理规范、道德责任、社会价值、伦理困境、道德冲突、社会公正、人权、伦理审查、利益相关者
辅助决策	在决策过程中对人工智能所使用的信息、工具或系统的可靠性和准确性的风险感知，特别是这些工具和系统能否支持高质量的决策	专注于决策过程中信息和工具的可靠性、准确性和透明度，以及对决策质量的影响	决策过程、信息可靠性、工具准确性、系统可信度、决策质量、风险评估、决策模型、数据完整性
就业市场	人工智能技术变革对就业市场造成负面影响的风险感知，包括潜在的失业风险和就业机会的变化	关注技术变革对就业岗位、职业类型和劳动市场的影响，特别是自动化导致的失业风险	就业影响、失业风险、技术变革、自动化、职业替代、劳动力市场、技能转型、再就业培训、职业发展

险感知指数的具体影响。

采用集成机器学习方法，结合XGBoost、随机森林、Extra Trees和AdaBoost等多种先进模型，量化了教育水平、基尼系数、收入、预期寿命和宗教影响等社会因素对风险感知指数的贡献。进一步地，基于博弈论的解释方法（Shapley Additive Explanations, SHAP）分析被用来识别和量化每个特征对模型预测结果的具体贡献。这种方法不仅揭示了不同因素对风险感知影响的大小和方向，还展示了它们之间的相互作用。模型构建过程中，本文排除了共线性变量和统计上不显著的变量，最终确定了16个预测变量对模型的相对贡献。特别是，个体自主性维度下收入水平及教育水平在模型中显示出了重要贡献。通过部分依赖图（PDP）分析，直观展示了政治信任与政府效能如何独立于其他变量影响模型预测，并揭示了潜在的非线性关系及相互作用效应，为设计和实施针对性干预措施提供了科学基础。

步骤三：全球各类人工智能风险感知的时序演化。

结合时间序列分析、地理信息系统（GIS）制图和集成机器学习技术，详细追踪了2012～2023年期间不同人工智能风险感知类别的时序演化趋势。这一分析不仅评估了各类风险感知的社会影响程度，还对具有代表性的人工智能风险感知类型进行了分析，探讨了它们为何受到全球关注。

三、实证分析与讨论

（一）公众对各类人工智能风险感知的全球模式

为了全面理解全球范围内公众对人工智能风险感知类别的异质性与共性，本文将视角基于空间维度，旨在探索不同文化、经济与政策背景下的公众认知模式，如何塑造他们对AI技术潜在威胁的理解与反应。表3与图3分别描绘了全球范围主要国家人工智能感知的类型占比及各洲人工智能风险感知首要类型的国家数量。研究基于高置信度的凸包分析，研究和解释全球范围内公众对AI技术风险的不同类别存在明显的空间分布差异。通过计算风险感知指数数据点，并将其映射到凸包空间内，以直观观察不同地区公众对人工智能风险感知的特征。86.4%的风险感知指数数据点位于凸包空间内，数据覆盖了95%以上的全球区域。这一结果表明，在大多数全球地区，样本的代表性和外推能力表现良好，即研究结果具有高信度。按照首要风险感知类别的国家数量在全球国家中的比例，进一步呈现全球公众风险感知的风险偏向（图4）。“隐私安全”是排在首位的感知风险类别，占全部国家的33.51%；“辅助决策”在次位，占24.23%。仅此两个领域相加占到全部国家的57.74%，可见各国公众更多感知到人工智能“隐私安全”和“辅助决策”等人工智能的“黑箱特质”：一方面体现在对

表 3　主要国家（地区）人工智能风险感知的类型占比

洲别	国家	信息透明	创新控制	伦理道德	辅助决策	就业市场	隐私安全
亚洲	中国	18.18746	1.89099	10.10323	20.87995	26.01772	22.92066
	印度	19.99408	0.5757	10.13703	18.06587	30.80077	20.42656
	日本	11.80276	4.78025	1.79939	7.56859	38.10435	35.94466
	韩国	2.07014	14.37	9.8982	6.19352	23.24574	44.2224
	印度尼西亚	0.69489	1.02358	4.35963	3.67192	42.70187	47.54812
	巴基斯坦	28.27698	14.98103	14.26776	17.08817	3.21509	22.17097
	孟加拉国	17.2928	6.45895	8.76322	9.62159	25.44443	32.419
	越南	12.78425	5.77382	23.38775	9.89575	29.58354	18.57489
	菲律宾	24.57549	9.17917	41.05958	8.03071	5.78001	11.37504
	泰国	21.17536	8.01829	22.83358	1.43572	18.04954	28.4875
非洲	尼日利亚	26.72193	10.16746	13.56002	3.8178	27.45925	18.27353
	埃塞俄比亚	13.31745	22.33532	19.08822	13.05983	15.87558	16.3236
	埃及	13.82886	23.93076	9.33164	3.88693	11.2634	37.75841
	南非	44.86943	0.12584	1.83282	9.99329	32.03667	11.14195
	肯尼亚	21.73478	0.14141	6.30682	41.40905	15.55285	14.85509
	乌干达	7.29843	25.08491	12.65275	42.98106	10.80118	1.18167
	阿尔及利亚	0.10097	1.46106	29.11463	8.4307	2.91504	57.9776
	苏丹	5.91734	6.30051	4.69955	1.76414	14.96717	66.35129
	摩洛哥	6.0472	5.30658	30.94485	10.36525	42.78474	4.55138
	加纳	28.84911	25.8865	3.11294	15.01258	12.16715	14.97172

(续表)

洲别	国家	信息透明	创新控制	伦理道德	辅助决策	就业市场	隐私安全
欧洲	德国	3.26468	5.7643	5.02114	21.7806	18.77028	45.40086
	法国	12.53449	4.4511	30.83964	19.84282	4.20439	28.12756
	英国	24.05653	11.49674	6.02956	5.0734	27.69284	25.65093
	意大利	3.77455	10.78928	47.64772	4.1993	22.79864	10.79051
	西班牙	5.64354	3.41418	11.8943	25.09229	5.80036	48.15533
	波兰	2.82561	4.93334	21.60171	41.51395	2.01836	27.10702
	荷兰	12.25464	12.65774	22.15554	3.25253	24.00844	25.6711
	比利时	15.15734	8.15725	10.62894	31.84008	12.50096	21.71543
	希腊	6.48691	2.5715	1.92886	33.39716	10.08581	45.52976
	葡萄牙	12.75473	11.94344	20.49163	6.08303	12.94832	35.77886
北美洲	美国	4.7562	4.23743	21.62194	40.69259	8.06581	20.62602
	加拿大	4.43477	7.06527	22.65297	33.14319	7.20701	25.49679
	墨西哥	19.56611	9.35416	6.23664	43.17351	8.56959	13.09999
	危地马拉	9.85817	3.13827	19.25896	3.57853	0.10354	64.06254
	古巴	2.58084	3.58648	21.59045	1.45244	3.47389	67.31591
	海地	14.72983	2.35109	25.32681	30.74113	24.90224	1.9489
	洪都拉斯	6.01623	21.26882	18.58857	8.04043	10.99626	35.08969
	尼加拉瓜	3.76337	1.73696	19.82758	60.26199	9.06745	5.34265
	哥斯达黎加	23.92791	6.54714	0.41659	10.51436	18.96039	39.63361
	巴拿马	3.38676	0.87525	17.3452	6.61503	47.3802	24.39755
南美洲	巴西	1.32647	37.09434	1.73165	40.6634	10.07124	9.1129
	阿根廷	12.59266	27.31441	1.98555	39.75508	0.99699	17.35531
	哥伦比亚	3.40128	10.8073	2.01779	0.87908	40.26429	42.63025
	智利	1.90516	25.66467	2.33335	6.43451	23.96933	39.69299
	秘鲁	9.51887	16.11774	1.63468	32.97567	1.56741	38.18563
	委内瑞拉	6.32796	2.73998	62.37908	18.25278	5.02534	5.27486
	乌拉圭	33.34096	6.93881	4.98449	3.21152	7.17809	44.34613
	玻利维亚	13.51698	21.62189	18.20036	10.48575	2.09817	34.07683
	厄瓜多尔	46.75268	12.91489	0.03468	32.73937	6.01216	1.54622
	巴拉圭	22.65798	0.11872	12.51957	1.57898	37.65669	25.46805
大洋洲	澳大利亚	6.91872	1.15681	21.91425	27.80679	15.09998	27.10344
	新西兰	71.18695	8.57349	11.27996	2.28669	0.09488	6.57802
	巴布亚新几内亚	23.25768	23.8788	22.53062	4.28936	10.84643	15.19711
	斐济	10.86323	14.71248	20.7835	23.95579	23.58193	6.10307
	所罗门群岛	0.27151	10.88534	7.36028	12.70553	8.48646	60.29088
	密克罗尼西亚	6.97179	7.88203	12.66816	2.97246	26.82321	42.68236
	瓦努阿图	13.02677	0.08515	4.80149	19.09903	7.3849	55.60266
	萨摩亚	5.42925	8.42885	28.09499	7.69098	3.52057	56.83536
	基里巴斯	0.97359	3.11661	19.42981	0.18605	1.17947	75.11447
	汤加	2.61498	14.06548	48.10922	12.58971	7.18058	15.44003

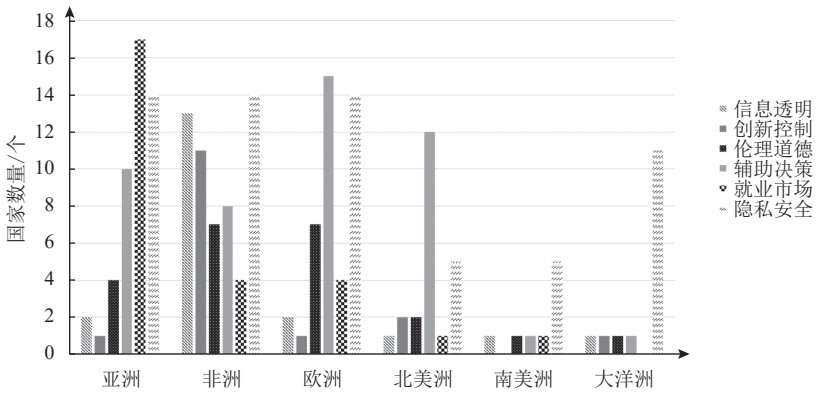

图 3　各洲人工智能风险感知类型（首要类型）的国家数量

人工智能技术会无授权地收集、存储和分析个人数据，从而侵犯个人隐私的忧虑；一方面体现在对决策过程模糊性而导致不公正结果的恐慌。

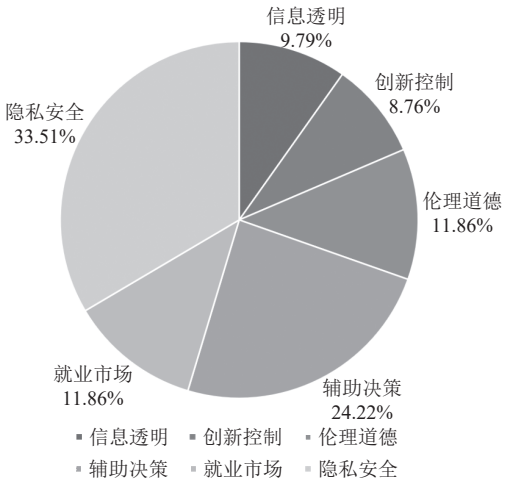

图 4 人工智能风险感知类型（首要类型）的国家数量占比

图5展示了各大洲内风险感知类别聚合的比例。通过对比各洲的风险感知类别的聚合比例，发现这些比例与各洲首要风险感知类别的国家数量比例高度一致，从而验证了不同地区风险感知类型的一致性。在政策指导和经济发展水平较高的地区，公众更关注隐私安全和辅助决策的风险；而在亚洲和非洲等发展中国家集中的地区，就业相关风险受到更多关注，这反映了地理和经济邻近性导致的“创新扩散效应”。以下是对各洲风险感知类别的具体分析：

1. 大洋洲：随着数字化时代的到来，公众对隐私侵犯的担忧日益增加。澳大利亚电信公司Optus的数据泄露事件凸显了隐私保护的紧迫性。澳大利亚政府计划修改隐私保护规则，并在未来10年投入约11亿美元加强网络基础设施，以保护公民个人信息[30]。新西兰等地的隐私法发展和社交媒体讨论也反映了公众对隐私侵犯的高度敏感性[31]。

2. 非洲：非洲地区对信息透明度风险的敏感度较高，部分原因是对政府透明性和责任性的普遍诉求。非洲联盟通过AGA和AGP等机制推动透明度提升，以满足公众对透明治理的期待。

3. 欧洲：个人隐私权和自由在欧洲社会中被视为核心价值观。欧盟的“以人为中心”的AI发展策略，以及《通用数据保护条例》（GDPR）的实施，体现了对个人数据保护的高标准，力图在全球AI市场中塑造“可信赖与安全”的AI技术形象。

4. 亚洲：亚洲对AI就业市场的风险高度敏感，这与经济结构转型和教育与技能差距有关。中国、日本、韩国和新加坡等国家在AI技术发展和应用方面处于全球前沿，引发公众对AI重塑就业市场的深刻关注[32~33]。技能不匹配问题和低技能劳动力面临的就业挑战尤为突出[34]。

5. 北美洲：北美洲在技术应用方面具有前瞻性优势，但公众对技术辅助决策的准确性和公正性保持警惕。米达斯反欺诈系统（Michigan Integrated Data Automated System）的高误判率事件暴露了自动化决策系统在保障公正性方面的漏洞，引发了对自动化决策系统监督和约束的紧迫需求。

6. 南美洲：南美洲对技术发展中的道德关切表现出深刻关注，强调技术进步应与社会道德和公正原则相协调。技术发展被视为对社会价值和道德原则的潜在挑战，政策制定者和公众在考虑技术发展时，更加重视其对社会道德和公正原则的影响。

（二）影响公众对人工智能风险感知的因素

为探究全球公众对人工智能风险感知的影响因素，本文利用XGBoost、随机森林、Extra Trees、AdaBoost等机器学习模型，结合SHAP值对变量的重要性进行排序，识别出对公众风险感知影响最大的16个变量（图6）。分析揭示了个体价值观、环境因素和治理需求在形成公众对人工智能风险感知

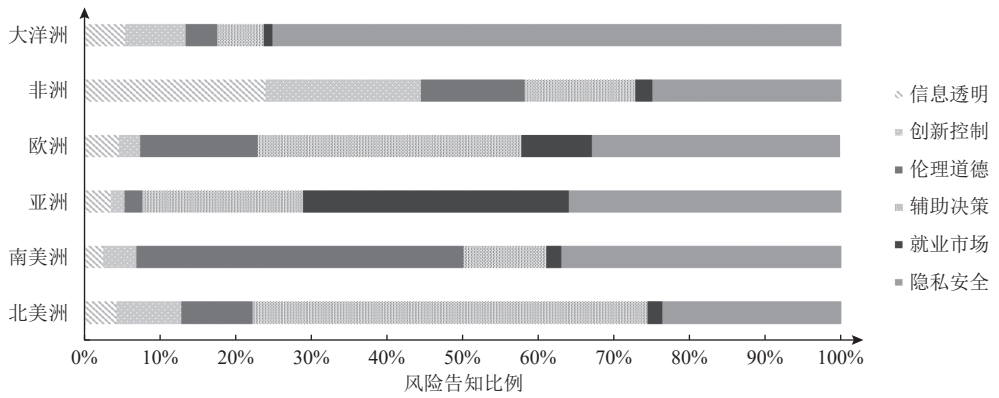

图 5 各洲不同类别风险感知

中所起的复杂作用。

经济因素，尤其是收入水平和基础设施的完善度，在解释公众风险感知的变异中占据了显著地位，分别贡献了39.4%和35.8%。此外，健康安全指数和法规质量指数也被证实对风险感知有显著影响。尽管政治因素如法规质量、选举民主指数和政府信任指数在风险感知的解释中排名靠后，但它们在塑造公众对人工智能风险的感知模式中仍然扮演着关键角色。特别是教育水平较高且收入较高的群体，对AI技术的风险感知更为强烈，这可归因于他们对前沿技术的认知深度及新技术的批判性评估能力。

进一步的分析，本文通过部分依赖图（Partial Dependence Plots, PDP）将机器学习模型预测结果的可视化，以分析单一特征对模型预测结果的平均影响。图7和图8展示了政治信任指数和政府效能对公众风险感知的具体影响。政治信任指数不仅反映了公众对政府机构的信任程度，还体现了公众对政府处理技术风险能力的信心。通过部分依赖图（PDP）分析（图7），观察到政治信任指数的提升与公众对AI风险感知的降低之间存在显著的负相关性。当政治信任指数从1.2增至1.4时，公众的部分依赖性保持相对稳定；但当指数超过1.4，特别是介于1.6~1.8之间时，依赖性显著下降，表明随着政治信任的增强，公众对政府的依赖减少，对AI风险的担忧也随之减轻。

此外，政府效能对公众风险感知的影响存在显著正相关性（图8）。政府效能的提升意味着政府在制定和执行政策方面更为有效，这增强了公众对政府管理AI技术能力的信任。在效能指数1.2~1.4的区间内，政府效能对部分依赖性的影响较小，表现为稳定状态。然而，当效能指数超过1.4，特别是达到1.6~1.75时，部分依赖性显著上升，反映出政府效能的提高显著增强了公众对政府处理AI风险能力的信心。政治信任和政府效能的这些发现强调了政府在塑造公众对人工智能风险感知中的重要作用。政府的透明度、及时性以及对技术风险的有效管理能够显著提升公众的政治信任，降低对AI技术潜在风险的担忧。这种信任的增强有助于为社会创造一个更加稳定和可预测的环境，促进公众与政府之间的沟通，使政府关于人工智能的监管措施和风险防控策略更易被公众理解和接受。

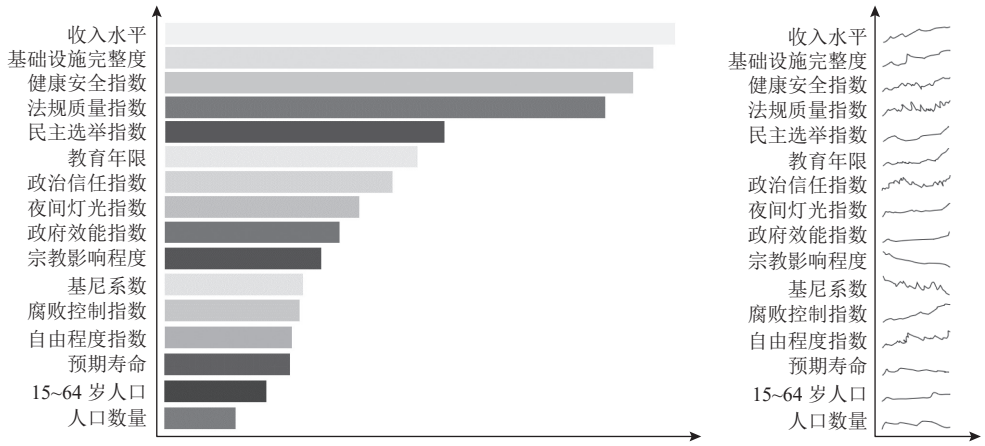

图 6　影响公众对人工智能风险感知的关键因素及其贡献度评估

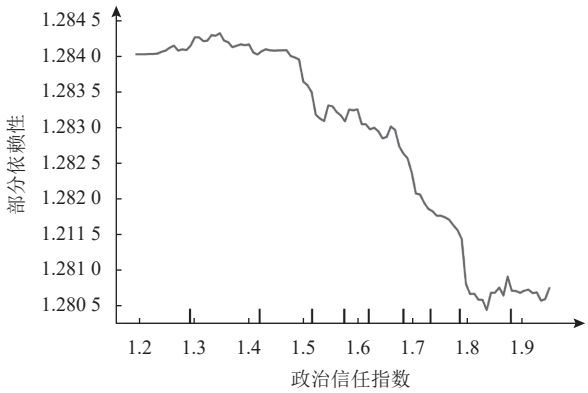

图 7　政治信任指数的部分依赖性

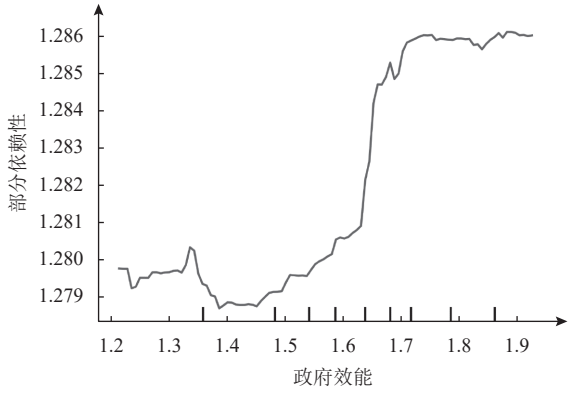

图 8　政府效能指数的部分依赖性

全球范围内，许多国家政府正积极推动AI产业的发展，并将其定位为“可信赖的人工智能”。经济合作与发展组织（OECD）的数据显示，截至2023年9月，全球共有69个国家针对人工智能治理推出了超过800项政策措施。欧盟和美国在推动“可信赖人工智能”的发展上尤为积极，通过制定政策、动员资源、共享算力和数据集等措施，加速AI技术的广泛部署[35]。这些行动不仅体现了政府对AI技术风险管理的重视，也展示了政府在提升公众信任和构建包容性全球规则体系中的积极作用。

（三）公众对人工智能风险感知模式的时序演化

本文的时间序列分析聚焦于公众对人工智能风险感知的演变。通过审视图9所呈现的数据，可以观察到公众对人工智能不同风险类别感知的动态变化。随着时间的推移，各类风险感知并非孤立发展，而是展现出相互关联和共变的模式。信息透明度和伦理道德的风险感知尤为突出，它们不仅数值较高，更呈现出明显的上升趋势。特别是在2017～2020年间，公众对信息透明度和伦理道德的关注度急剧上升，反映出社会对这些领域的快速觉醒和日益增长的关切。同时，自2018年起，创新控制的风险感知亦显著增强，这可能与技术进步带来的管理挑战和不确定性紧密相关。此外，辅助决策和就业市场的风险感知自2017年后也显示出加速增长的趋势。整体来看，公众的风险感知呈现出全面增长的态势，这一变化不仅标志着风险意识的扩散，也映射出技术普及所带来的广泛忧虑。风险感知的“外溢”现象，揭示了技术发展在为社会带来便利的同时，也激发了公众对潜在风险的思考和广泛讨论。在某些特定领域，公众的风险感知还经历了显著的转折，表现出“异化”的波动特性，这可能与社会事件、技术革新或政策变动等因素密切相关。

图10所描绘的公众对人工智能风险感知比例的时序变迁，展现了特定风险维度如何逐渐成为公众关注的焦点。从2012年的“伦理道德”风险感知占据首位，到2023年形成以“辅助决策”为主导，“伦理道德”和“隐私安全”紧随其后的多元格局，这一变化不仅标志着AI议题在全球治理议程中的迅速上升，也反映了人工智能感知在社会各个层面的复杂演变。在2012～2016年的中期，人工智能技术的初期发展伴随着伦理探索的热潮，科幻文化的广泛传播和初步伦理争议案例的浮现，使得“伦理道德”成为公众讨论的核心。然而，随着全球人工智能治理的重心逐步调整，公众的风险感知重点开始向“辅助决策”类别转移，这种转变反映了公众对人工智能可能成为政治操纵工具的深切担忧[36]。特别是在2020年美国大选期间，一些政治团体利用人工智能技术生成假新闻和伪造视频（Deepfake），以及针对性的政治广告，这些行为显著加剧了公众对辅助决策风险的恐慌情绪，进一步凸显了构建有效治理机制的紧迫性。这一进程不仅见证了从民间倡议到政府与国际组织深度参与的转变，也突显了公众对人工智能技术潜在风险的持续关注和对治理响应的期待。

随着全球互联的深化和公众风险规避意识的提升，全球议题的溢出效应凸显了构建高效国际治理机制的紧迫性。2016年成为全球人工智能治理的新起点，多国政府相继推出政策框架，如《国家人工智能研究与发展战略计划》（美国）、《新一代人工智能发展规划》（中国）及《人工智能白皮书》（英国），旨在引导AI技术的健康发展，并形成了政策引导与公众感知之间的互动循环[37~38]。同时，欧盟与其他国家的双边合作协议，致力于消除政策碎片化，推动治理规则的一致性[39]，体现了从伦理

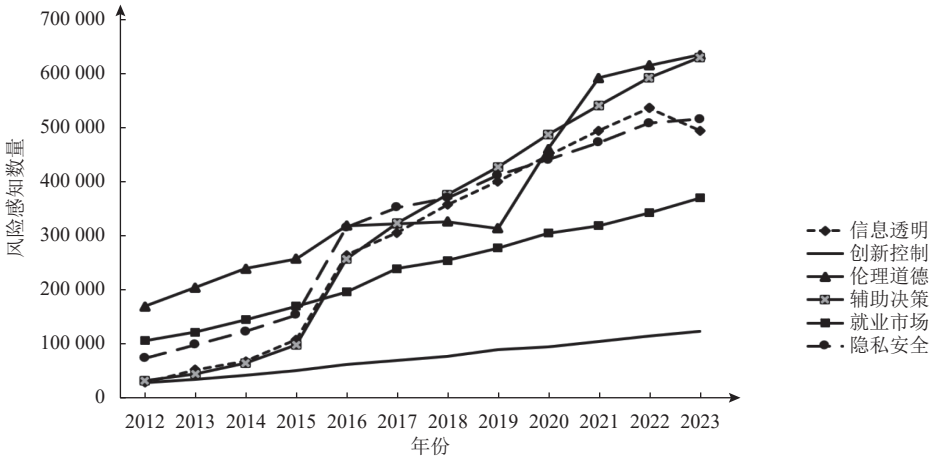

图 9 公众对不同人工智能类型风险感知数量变化的时序演化

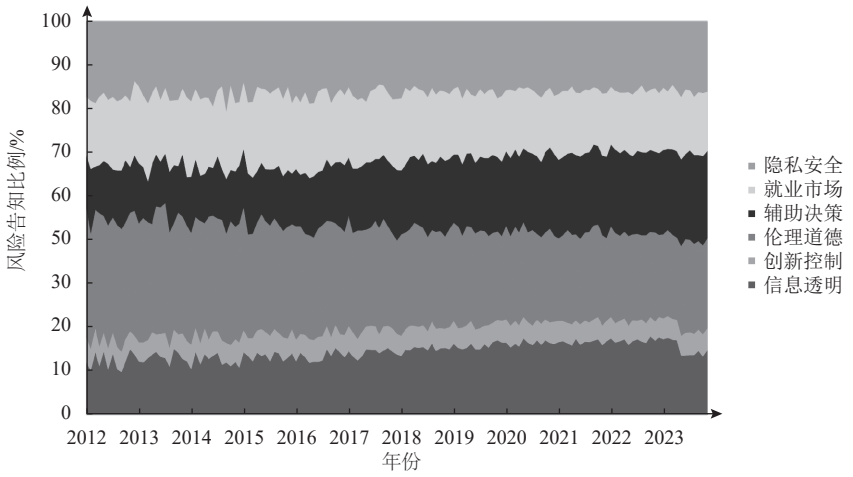

图 10　公众对不同人工智能类型风险感知比例变化的时序演化

规范、国际合作到政策协调的全方位治理决心，旨在提升公众信任，确保人工智能技术的普惠性，让全人类共享其发展成果[40]。这些举措不仅反映了全球社会对人工智能风险的深刻认识，也展现了共同构建一个稳定、透明、可预测的人工智能治理环境的明确愿景。

四、结论与启示

在当前快速发展的人工智能领域，其传播速度和影响范围的广泛性引发了对潜在风险的广泛关注。鉴于此，深入研究公众对人工智能风险的感知及其价值变得尤为重要。本文从全球视角出发，构建了一个包含时空特征的数据库，旨在揭示公众对人工智能风险感知的全球分布、特定风险感知的类型及其随时间的变化趋势。研究发现：第一，区域异质化。在人工智能应用的具体风险上，公众因地域文化、意识形态、价值观念上的差异而形成不同的风险感知特征，并呈现出“创新扩散效应”。第二，多因素影响。个体自主性、政治信任和政府效能等社会经济对公众的风险感知产生显著贡献。这表明，公众的风险感知并非孤立形成，而是与其所处的社会环境密切相关。第三，显著时序演化。公众的风险感知随时间推移而发生了动态演变，这反映出社会对人工智能技术发展的认知在不断深化，同时也显示出对新兴风险的关注在逐渐增强。

本文的研究结果为人工智能政策领域提供了针对性的研究启示，突显了在政策制定中需细致考虑的多个维度。首先，研究揭示的区域异质性特征强调了制定全球人工智能政策时必须考虑的地域特异性。政策制定者应当超越单一文化和政治体系的局限，采纳多元和包容的视角，以确保治理策略能够适应不同社会的需求和期望。这种跨文化的敏感性和对地方特定情境的深刻理解，是构建有效全球治理框架的前提。其次，研究结果表明个体自主性、政治信任和政府效能等因素对公众的风险感知有显著影响。这意味着政策制定者在设计人工智能治理策略时，需要综合考虑社会经济和政治背景，以及这些因素如何塑造公众的风险感知。政策制定应当促进社会各方面的积极参与，包括提升个体的自主性和能动性，增强政治信任，以及提高政府效能，从而共同构建一个稳定、透明和可预测的政策环境。第三，公众风险感知的显著时序演化特征要求政策制定者采取动态的治理方法。随着人工智能技术的不断进步和应用领域的拓展，新的风险类型和风险感知模式将不断出现。鉴于公众风险感知的动态特性，政策制定者应考虑建立灵活的政策框架，这些框架能够适应技术发展和社会变革带来的新挑战。这包括定期审查和更新政策措施，以确保它们与最新的社会趋势和技术进步保持同步。同时，应鼓励跨学科合作，利用社会科学、技术科学和人文学科的洞见，共同构建一个全面的风险治理体系。政策制定者需要建立灵活的治理机制，以适应这些变化，并能够及时响应新兴的风险挑战。这包括持续监测技术发展的趋势，评估其对社会的潜在影响，并更新政策以管理这些风险。第四，本文呼吁建立一个持续的监测和评估机制，对公众的风险感知进行动态跟踪和分析。这一机制不仅能够帮助政策制定者及时了解和响应公众的风险感知变化，而且能够为政策的持续优化和调整提供科学依据。通过这种持续的研究和评估，可以更好地理解人工智能技术的社会影响，确保其在经济社会中的健康发展。

本文也存在一定的研究局限。首先，本研究在数据收集上依赖于社交软件平台，这些平台的数据可能会受到审核机制和删除政策的影响。这种数据的偏差可能导致我们观察到的公众风险感知与实际情况存在偏差。其次，本研究主要采用了定量分析方法，虽然这种方法在宏观层面上为我们提供了公众风险感知的趋势和模式，但它在揭示个体层面的动机和心理社会机制方面存在局限。定量数据往往掩盖了个体差异和深层次的社会文化因素，这些都是理解公众风险感知的关键要素。再者，本研究在分析的颗粒度尚不够精细，未能充分细化到个体和群体层面的具体特征。未来研究将通过细化分析单位，更精确地捕捉个体和群体层面的风险感知差异，从而为制定更为精准和有效的政策提供支持。

注释

① 本研究构建的公众对AI的风险感知数据集及其空间高分辨率分布数据集可以在以下网址开放获取：https://figshare.com/articles/dataset/_b_Untitled_Item_b_b_Spatial_DataSet_for_Global_public_b_b_risk_b_b_perception_of_artificial_intelligence_b_/25305244。

参考文献

[1] MÜLLER C V. Risks of Artificial Intelligence[M]. London: CRC Press, 2016.

[2] 乌尔里希·贝克. 风险社会: 新的现代性之路[M]. 张文杰, 何博闻, 译. 南京: 译林出版社, 2018.

[3] REBEKKA S, IRINA B, JÜRGEN B, et al. Using artificial intelligence (AI)? Risk and opportunity perception of AI predict people's willingness to use AI[J]. Journal of Risk Research, 2023, 26(10): 1053-1084.

[4] 李森林, 张乐, 李瑾. 当代青年人工智能风险感知的测度与解析[J]. 科学学研究, 2023, 41(10): 1737-1746.

[5] 李建霞, 余丹丹. 科学数据共享平台用户感知有用性分析[J]. 情报杂志, 2023, 42(9): 196-201.

[6] 袁媛, 严宇桥. 风险感知与网络舆情的微博传播模型研究[J]. 现代传播(中国传媒大学学报), 2020, 42(1): 158-163.

[7] SUHARINI E, SUPRIYADI, SYIFAUDDIN M, et al. An evaluation of community adoption of the InaRISK BNPB platform for disaster management: an application of the Technology Acceptance Model (TAM)[J]. International Journal of Safety and Security Engineering, 2023, 13(4): 673-689.

[8] 张楠, 闫涛, 张腾. 如何实现"黑箱"下的算法治理?——平台推荐算法监管的测量实验与策略探索[J]. 公共行政评论, 2024, 17(1): 25-44+196.

[9] 邱泽奇. 算法治理的技术迷思与行动选择[J]. 人民论坛·学术前沿, 2022(10): 29-43.

[10] 秦川申, 刘运喆. 人脸识别风险沟通中叙事情节的作用: 基于一项调查实验[J]. 公共管理评论, 2023, 5(3): 31-56.

[11] 朱力, 夏恩君, 王为. 人工智能的就业影响研究综述[J]. 科技和产业, 2022, 22(1): 32-43.

[12] 蒲晓红, 赵海堂. 互联网使用对公众风险感知的影响机制——基于政府回应视角[J]. 中国行政管理, 2021(5): 146-154.

[13] SCHULTE F, TRINN C. Collective emotions, triggering events, and self-organization: the forest-fire model of cultural identity conflict escalation[J]. Aggression and Violent Behavior, 2024, 78(4): 101954.

[14] JIA K, ZHANG N. Categorization and eccentricity of AI risks: a comparative study of the global AI guidelines[J]. Electronic Markets, 2021, 32(1): 1-13.

[15] NADIA S, E. A P, IRINA B, et al. An Artificial Intelligence perspective: how knowledge and confidence shape risk and benefit perception[J]. Computers in Human Behavior, 2023, 149.

[16] K M S, KATHERINE C. Are robots/AI viewed as more of a workforce threat in unequal societies? Evidence from the eurobarometer survey[J]. Technology, Mind, and Behavior, 2022, 3(2): 1-13.

[17] JOHANNA V, AHMED M T, NUNO O, et al. Artificial Intelligence in K-12 Education: eliciting and reflecting on Swedish teachers' understanding of AI and its implications for teaching & learning[J]. Education and Information Technologies, 2023, 29(4): 4085-4105.

[18] 贾开, 俞晗之, 薛澜. 人工智能全球治理新阶段的特征、赤字与改革方向[J]. 国际论坛, 2024, 26(3): 62-78+157-158.

[19] S R W, ADAM Z, HUGH W. Developing a broadly applicable measure of risk perception[J]. Risk Analysis: An Official Publication of the Society for Risk Analysis, 2019, 39(4): 777-791.

[20] MATHILDE H D G D. Fair strategies to tackle unfair risks? Justice considerations within flood risk management[J]. International Journal of Disaster Risk Reduction, 2022, 69.

[21] KLEIN J T, ANDREWS M, CIBULSKIS M, et al. The digital divide in action: how experiences of digital technology shape future relationships with artificial intelligence[J]. AI and Ethics, 2023, 4(2): 345-367.

[22] 朱依娜, 何光喜. 信任能降低公众对人工智能技术的风险感知吗?[J]. 科学学研究, 2021, 39(10): 1748-1757+1849.

[23] 上官莉娜, 徐云鹏. 互联网使用对公众风险感知作用机理的实证研究——基于公众政治参与和对政府信任的视角[J]. 中南大学学报(社会科学版), 2022, 28(3): 153-164.

[24] 杨建武. 智能治理伦理风险的关键影响因素研究——基于DEMATEL方法[J]. 科学与社会, 2021, 11(4): 80-97.

[25] CLINE N B. WILLIAMSON, et al. Trust, regulation, and market efficiency[J]. Public Choice, 2022, 190(3-4): 1-30.

[26] 陈宇飞, 沈超, 王骞, 等. 人工智能系统安全与隐私风险[J]. 计算机研究与发展, 2019, 56(10): 2135-2150.

[27] 吕丹阳, 郎元柯, 范柏乃, 等. 生成式人工智能在公共服务中应用的机遇与挑战[J]. 电子科技大学学报(社科版), 2024, 26(3): 35-45.

[28] 唐永, 李想. 人工智能发展对制造业就业的影响——基于马克思社会再生产模型的中国经验分析[J]. 当代经济研究, 2024(6): 36-50.

[29] HONGJUN G, LIYE D, AIWU Z. Ethical risk factors and mechanisms in Artificial Intelligence decision making[J]. Behavioral Sciences, 2022, 12(9): 343-343.

[30] SHAH, PRITAM GAJKUMAR. Optus data breach Australia year 2022—case study[J]. Australian Journal of Wireless Technologies, Mobility and Security, 2022, 3(1): 45-60.

[31] DAS CHAUDHURY, RATUL, CHOE, CHONGWOO. Digital privacy: GDPR and its lessons for Australia[J]. Australian Economic Review, 2023, 56(2): 204-220.

[32] YAN H, GAI S. Work life, and artificial intelligence (AI): emerging findings from Asia[J]. Work-Life Research in the Asia-Pacific: Implications for Justice, Equity, Diversity, and Inclusion, 2023, 2(1): 95-112.

[33] MALLAH N B. Artificial Intelligence impact on banks clients and employees in an Asian developing country[J]. Journal of Asia Business Studies, 2022, 16(2): 267-278.

[34] 王林辉, 钱圆圆, 周慧琳, 等. 人工智能技术冲击和中国职业变迁方向[J]. 管理世界, 2023, 39(11): 74-95.

[35] 钟悦, 王洁. 教育领域人工智能的应用现状、影响与挑战——基于OECD《教育中的可信赖人工智能: 前景与挑战》报告的解读与分析[J]. 世界教育信息, 2021, 34(1): 73-79.

[36] HAMELEERS, MICHAEL. Cheap versus deep manipulation: the effects of cheapfakes versus deepfakes in a political setting[J]. International Journal of Public Opinion Research, 2024, 36(1): edae004.

[37] HJALTALIN T I, SIGURDARSON T H. The strategic use of AI in the public sector: a public values analysis of national AI strategies[J]. Government Information Quarterly, 2024, 41(1): 101914.

[38] GUANGYU F Q, RONGSHENG Z. China's Artificial Intelligence ethics: policy development in an emergent community of practice[J]. Journal of Contemporary China, 2024, 33(146): 189-205.

[39] CHARANJIT S. Artificial Intelligence and deep learning: considerations for financial institutions for compliance with the regulatory burden in the United Kingdom[J]. Journal of Financial Crime, 2024, 31(2): 259-266.

[40] LAUX J, WACHTER S, MITTELSTADT B. Trustworthy artificial intelligence and the European Union AI act: on the conflation of trustworthiness and acceptability of risk[J]. Regulation & Governance, 2023, 18(1): 3-32.

Types and Evolution of Public Risk Perception of Artificial Intelligence: A Global Perspective Analysis

LI Zhi-chao[1]　LI Lu-jia[2]

(1. Shanghai Jiao Tong University　Shanghai　200030　China; 2. State Environmental Protection Key Laboratory of All Materials Fluxes in River Ecosystems　Beijing　100871　China)

Abstract　This study aims to analyze the global public's perception of AI risks, examining the types and spatio-temporal evolution of these perceptions. It also seeks to identify the factors influencing these perceptions, providing empirical evidence for the development of trustworthy AI globally. By conducting semantic analysis of 4.3 million posts on social media and categorizing temporal features, a global database is constructed to quantify AI risk types across different regions. The study employs convex hull analysis, geographic information systems (GIS), and integrated machine learning methods to reveal the regional heterogeneity, influencing factors, and spatio-temporal evolution of global public risk perceptions. The findings indicate significant regional heterogeneity in global public perceptions of AI risks, shaped by the diffusion effects of innovative technologies, resulting in diverse perception categories. Individual characteristics, political trust, and government effectiveness significantly shape public risk perceptions. These perception categories exhibit a covariant logic, transitioning from a singular to a pluralistic pattern.

Key words　risk perception; Artificial Intelligence; spatio-temporal pattern; global perspective

编　辑　朱娜