

脑机接口伦理研究:学科交叉、风险演化

王高峰¹,姚明月¹,甘烨彤²

(1. 中国科学技术大学科技哲学系,安徽合肥 230026;

2. 广东工业大学人文与社会科学高等研究院,广东广州 510006)

摘要:随着脑机接口技术取得突破性进展、产业扩张不断加速,其引发的伦理问题日益受到学界关注。基于科学计量学方法,结合学科交叉与风险演化的动态分析视角,梳理分析国际脑机接口伦理研究的概况、知识基础及热点,发现:(1)脑机接口伦理研究的发文量呈三阶段演进特征,其波动与技术发展周期、社会事件及政策导向深度关联,2024年为爆发性窗口期。(2)生命伦理学、认知神经科学及机器伦理学从价值规范、技术路径与人机协同规范等维度构成模块化跨学科引文网络,支撑多维理论框架。(3)研究热点在主体、技术与社会场域形成聚类,凸显风险演化动态特征。(4)随着脑机融合的深化,融合主体的神经权利及社会接纳问题亟待重视。基于现状研判,后续研究应着力于跨学科的实证协作、驱动产业优化升级、推动去中心化的网络治理,以助力全球脑机接口伦理研究的完善与发展。

关键词:脑机接口;科技伦理;学科交叉;风险演化;科学计量

中图分类号:R-052;B82-057;TN911

文献标识码:A

脑机接口(Brain-Computer Interface, BCI)技术在医疗康复、神经增强等领域的应用深化,正重塑人类生命形态、主体认知以及社会关系的边界,其伴生的伦理风险引发了学界的哲学审视与治理关切^[1]。当前,脑机接口伦理相关研究多聚焦于临床应用中的知情同意、数据隐私保护、认知增强的公平性等具体议题,虽部分达成共识,却缺乏对“技术-人-社会”这一互动关系的整体关照。具体表现为:微观与宏观伦理的断层、技术可能与伦理接受的脱节以及静态风险规避与动态技术演进的错位。作为新质生产力的典型代表,脑机接口兼具“高科技、高效能、高质量”等特性^[2],推动着人类社会向智能化与生物增强化方向加速迈进。与此同时,其伦理问题呈现出高度复杂性与不确定性。故而,从学科交叉与风险演化的视角出发,厘清脑机接口伦理的研究路径,不仅关乎技术发展的可持续性,更是构建包容、安全的人机关系的关键。

基于上述背景,本研究聚焦科学计量,系统梳理国际脑机接口伦理研究成果,并通过可视化分析,揭

示其知识基础与研究热点,旨在突破既有研究局限:在研究视角上,构建“学科交叉-风险演化”框架,通过引文聚类凝练出生命伦理学、认知神经科学、机器伦理学三大知识基础,进而聚焦主体、技术、社会三大风险演变;在分析维度上,建立“知识基础-研究热点”分析机制,借助可视化手段呈现二者的动态关联来打通分析断层,进而透视该领域研究重心变化与演进方向以捕捉前沿动态,以期为国内脑机接口伦理与治理研究提供理论支持与实践导向。

1 研究方法概述

1.1 研究样本与方法

本研究基于Web of Science核心合集数据库,检索与脑机接口伦理相关的主题词,文献类型为Article和Review,索引源为Sciences Citation Index-Expanded(1900至今)、Social Sciences Citation Index(1900至今)和Arts & Humanities Citation Index(1975至今)。在消除数据库中的“噪音”无效数据

后,最终选定 265^① 篇文献作为知识图谱分析的初始样本。

在获取并整理文献数据后,本研究运用 CiteSpace(6.3. R1) 软件、VOSviewer(1. 6. 19) 软件以及 Biblioshiny 展开系统分析;运用普赖斯定律对脑机接口伦理研究的发文量进行分析,以揭示该领域的发展阶段与研究活跃度;通过桑基图对研究基础与热点进行初步可视化探查;综合运用聚类分析、共现分析等计量方法,并结合系统文献综述,深入探析脑机接口伦理研究的知识基础与热点,从而系统把握该研究领域的整体现状及其演变路径。

1.2 文献发文与影响

年度发文量是衡量某领域学术发展态势的重要量化指标,能有效反映该领域知识生产的动态变化和研究进展,并揭示相关学术研究与社会发展之间的互动关系。如图 1 所示,脑机接口伦理研究的年度发文量呈现明显的上升趋势,其发展历程可大致划分为三个阶段:

第一阶段(1996 - 2012 年)为萌芽迸发期:在 1996 年,阿普佐(Michael Apuzzo)等学者分析了神

经外科发展趋势,并首次关注由脑机接口相关技术所引发的伦理问题^[3]。随后在 2009 年,博斯特罗姆(Nick Bostrom)对认知增强技术进行了开创性伦理探讨^[4]。该阶段的发文量较少且波动平缓,累积发文量仅为 33 篇。

第二阶段(2013 - 2023 年)为持续发展期:自 2013 年起,脑机接口伦理研究呈现出多元化趋势,其探讨涵盖了临床应用中的伦理难题、人机交互的现象学分析及神经哲学等多个层面。此阶段较第一阶段增幅明显,年均发文量显著增长,研究内容趋于细化与深化。脑机接口伦理研究文献平均被引次数为 21,而杨(Guang - Zhong Yang)等人于 2018 年提出的“科学机器人十大挑战”将脑机接口列为核心议题之一^[5],其文献被引量高达 752 次,是统计截止前该领域内最具影响力的研究。2021 年,Neuralink 公司高调演示的猕猴实验^②,引发广泛的伦理热议,推动当年发文量达到阶段性峰值。然而,受临床转化滞后的限制,以及元宇宙、ChatGPT 等热点领域的分流,2022 至 2023 年的研究热度略有回落。

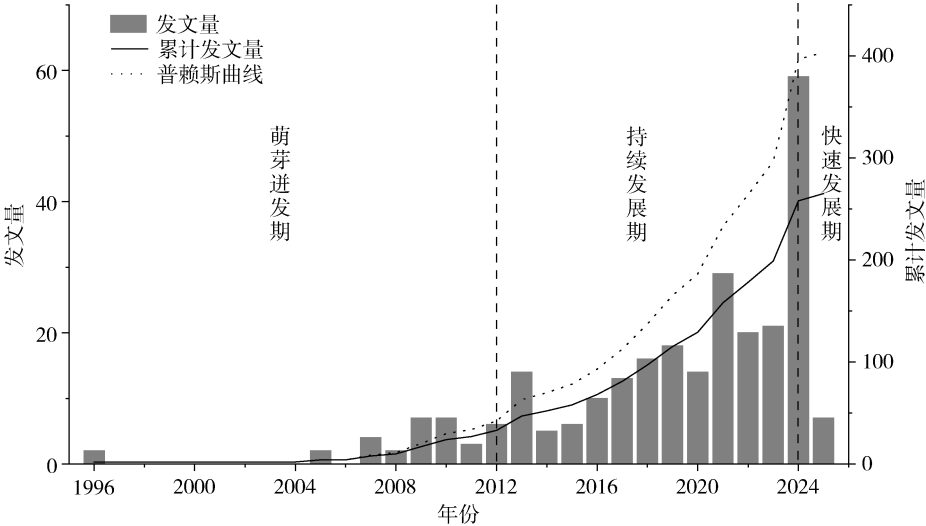


图 1 国际脑机接口伦理研究文献年度发文量、累积发文量及普赖斯曲线
Figure 1 Annual and cumulative publication volumes and price curve of international literature on the ethics of brain - computer interface

第三阶段(2024 年至今)为快速发展期:受各国相关科技布局、关键技术的突破、多款产品进入临床

实验以及产业产值的迅速增长等多重因素影响,2024 年发文量达到峰值(59 篇),标志着脑机接口

① 检索式为:TS = (brain - computer interface * OR brain - machine interface * OR bci OR brain computer interface * OR brain machine interface *) AND TS = (ethic * OR moral * OR philoso *),检索时间为:2025 年 2 月 17 日。

② 一只名为 Pager 的猕猴能够借助 1024 个植入电极和数据传输设备(称 N1 Link)在计算机屏幕中移动光标。来源: <https://neuralink.com/updates/pager - plays - mindpong/>。

伦理研究进入全新发展阶段。另借助普赖斯拟合曲线发现,该领域仍具较大发展潜力,预计未来相关发文量仍将保持较高的增长趋势,研究方向也将持续向技术治理与哲学思辨纵深演进。

1.3 研究基础与热点

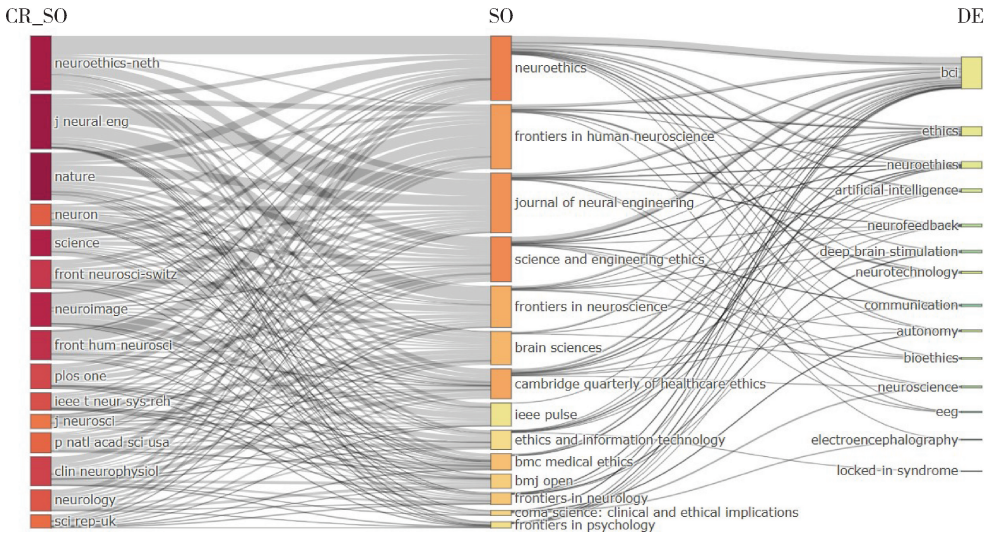


图 2 脑机接口伦理研究基础和热点的桑基图
Figure 2 Sankey diagram of ethical research foundations and hotspots in brain – computer interface

左侧的被引期刊 (Cited Reference Source, CR_SO) 系该领域论文的引文来源期刊信息。其中,以 Neuroethics 为代表的理论期刊,为脑机接口的伦理反思提供了核心概念框架。Journal of Neural Engineering、Frontiers in Human Neuroscience 等专业期刊的引用表明,神经工程和神经科学是脑机接口伦理研究的重要基础,不仅提供了科学背景和技术细节,还展示了脑机接口伦理研究的跨学科特性。而 Nature、Science 等高影响力期刊的引用展现出脑机接口伦理研究与前沿科学发现之间的紧密关系。

根据中间的论文来源期刊 (Publication Source, SO) 显示,以 Neuroethics、Science and Engineering Ethics 为代表的专业伦理学期刊发挥着双重作用:既通过构建理论模型为具体伦理问题提供分析框架,又搭建起神经科学、生物医学与哲学、伦理学的跨界对话平台。而右侧的关键词 (Author Keywords, DE) 揭示了当前脑机接口伦理研究的热点话题。“ethics”“neuroethics”“bioethics”等关键词直接关联神经伦理期刊 (如 Neuroethics), 构成该领域核心理论坐标体

系。同时,“bci”作为中心关键词,与“neurofeedback”“neurotechnology”等一同直观反映出研究者对技术演进轨迹的持续性追踪。此外,“artificial intelligence”“communication”等关键词突显了脑机接口伦理研究思路的拓展,并且这些关键词与工程类期刊 (如 Journal of Neural Engineering) 以及技术类来源期刊 (如 Frontiers in Neuroscience) 高度契合。

2 学科交叉:脑机接口伦理研究知识基础

桑基图对文献引用路径的可视化分析表明,脑机接口伦理研究的知识基础呈现出显著的跨学科特征。尽管研究主题复杂,其知识体系仍可归纳为两大核心维度:一是聚焦于人自身,具体涵盖对生命一般意义的探讨以及对个体认知过程的探究;二是聚焦于机器,即对机器的规范性研究。结合上述维度与聚类分析结果^③,本研究进一步提炼出三类知识基础 (见图 3):探讨生命本质与价值的生命伦理学、

③ 本研究借助 CiteSpace,采用对数似然率算法对脑机接口伦理文献进行引文聚类分析。结果显示,该聚类网络的模块值为 0.8266 > 0.3,表明知识基础聚类的社团结构显著;聚类网络的平均轮廓值 (S 值) 为 0.9322 > 0.7,且 16 个子集聚类轮廓值均保持在 0.7 以上,说明聚类结果可信 (见表 1)。

研究个体认知机制的认知神经科学以及针对机器规范性研究的机器伦理学。

表 1 脑机接口伦理研究的引文聚类

Table 1 Citation clustering in ethical research on brain – computer interface

序号	规模	S 值	平均年	聚类标签
0	93	0.927	2021	ethical consideration ; narrative review; invasive brain-computer interface; ethical issue; ai-driven speech-bcis
1	56	0.904	2015	using brain-computer interface ; employing social research method; affective brain-computer interface; rehabilitation medicine; closed loop brain device
2	55	0.843	2019	influencing discussion ; systematic review; neural interface; visual neuroprostheses-a; canadian speech-language pathologist
3	45	0.928	2010	brain-hardware interface ; stakeholders opinion; asilomar survey; ethical issue; bonding brain
4	40	0.916	2016	ecg-based brain-computer interface ; the-art technologies; human cognitive augmentation; current state; introductory tutorial
5	34	0.957	2006	consciousness science ; coma; new era; answering multiple-choice question; bold response
6	29	0.988	2006	view ; using bcis; locked-in syndrome; brain-computer interface; ethical consideration
7	28	0.888	2013	brain-computer interface communication ; rehabilitation technologies; cybathlon lesson; neuroengineering assistive; proactive ethical design
8	27	0.946	2012	informed consent ; identifying risk; exploring meaning; bests practice; research risk
9	20	0.997	2009	intention concept ; brain-computer interface; ethical consideration; ethical issue; narrative review
10	20	0.951	2011	overview ; brain-machine interface; paralysis; direct brain control; communication
11	17	1	2002	forthcoming ethical issue ; biological psychiatry; brain-computer interface; ethical consideration; ethical issue
12	16	1	2023	artificial intelligence ; transformative synergies; neuroscience; brain research; clinical application
14	13	1	2004	life ; end; diagnosis; amyotrophic lateral sclerosis; ethical issue
15	13	1	2007	ethical perspective ; military neurosurgical; neuroethics; challenge; brain-machine interface
17	10	0.987	2013	neuroprostheses ; trend; communication; challenge; brain-computer interface

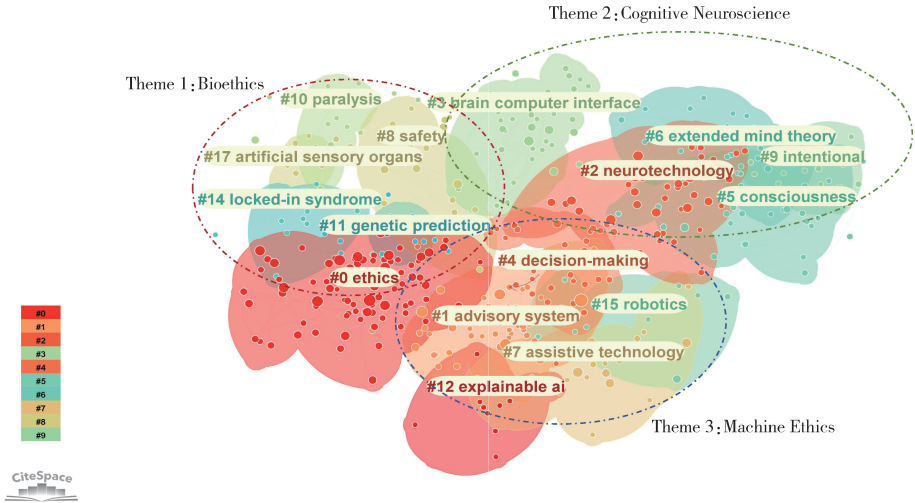


图 3 脑机接口伦理研究的文献共被引图谱

Figure 3 Literature co – citation graph of ethical research on brain – computer interface

2.1 生命伦理学

脑机接口伦理对生命状态的干预以及临床实践规范的关注,使其天然地归于生命伦理学的研究范畴之中。生命伦理学主要对临床医学、生物医学研究、公共卫生实践以及新兴生物技术的研发和应用中实然与应然的伦理问题展开探讨^[7],并据此制定包括政策与立法在内的规范。同时,脑机接口伦理旨在为技术的开发与应用确立价值遵循。这一目的本质上隶属于规范伦理学范畴,并与生命伦理学的内在规范路径相一致。尽管描述性研究是生命伦理学的重要起点,但规范本身承载、蕴含了价值取向与价值判断,且“是”与“应该”之间没有必然的逻辑通路。因此,脑机接口伦理的规范建构不能仅靠观察和实验得出,还需一定的理性指导,即在生命伦理学的理论框架下探寻其研究动因与价值规范,并将脑机接口对使用者的改变(是何)、使用者的切身体验和需求(为何)、技术临床实践中面临的问题处理(如何)共同纳入研究范畴。

2.2 认知神经科学

脑机接口伦理研究对认知、神经工程以及神经外科应用等领域的关切,主要涉及认知神经科学。认知神经科学即关于心智研究的理论与学说,以认知过程及其规律为研究对象^[8],涵盖了学习、记忆、思维、理解等认知行为,并对神经系统展开深入研究。认知神经科学从技术路径出发,为脑机接口伦理研究提供了坚实的理论基础和明确的研究对象,进而引出神经干预对自然认知边界的影响、脑机融合下的意识归属以及责任判定等前沿议题。这使得相关伦理研究不再是纯理性思辨,而是转为在技术实践与逻辑演变的交叉影响下开展的科学实证研究。研究者通过认知神经科学对复杂心智的深度解码,能够前瞻性预判脑机融合技术的伦理阈值,为技术路线的伦理设计、政策法规的适应性制定及个体认知权益保护提供决策依据,最终保障脑机接口的发展符合人类认知尊严与普遍价值共识。

2.3 机器伦理学

脑机接口伦理研究对机器使用规范及其自身伦理属性的探讨,涉及机器伦理学。机器伦理关注于机器本身的伦理属性以及机器使用中体现的伦理功能^[9],并研究是否能够通过技术手段使得计算机等智能机器拥有道德决策能力。此类研究为脑机接口伦理研究提供新思路,衍生出有关机器组件设计伦理化、人机混合体的道德责任归属以及技术系统道德

能力建构等前沿议题。机器伦理的方法论源自传统伦理的双重划分:一是,描述性伦理提炼现实社会的道德信念、道德判断和道德行为的模式,为脑机接口的技术设计提供人类道德参照系;二是,规范性伦理依据既定的伦理理论为具体情境下脑机接口应用提供价值遵循。这种跨学科融合推动脑机接口伦理从“人类中心主义”转向“人机协同”,通过整合机器伦理的双重要求,在技术精度与伦理保护间建立平衡。并确保脑机接口在拓展个体能力的同时,始终坚守“机器辅助而非替代人类道德判断”的准则,从而为构建人机共融的伦理生态开辟可行路径。

3 风险演化:脑机接口伦理研究热点

脑机接口技术对人脑的深度介入及其跨学科特性,使得其伦理风险呈现出多维交织性。然而,现有研究对这些伦理议题的梳理仍显零散,缺乏对核心热点的系统性聚焦与整合。为此,本研究利用 VOS-viewer 绘制关键词共现图,对该领域的主题分布特征以及演化规律进行可视化呈现(见图4),识别出主体风险、技术风险与社会风险三个相互关联且动态演变的热点集群,揭示了其从个体认知层面逐步扩展到社会结构层面的演进逻辑。

3.1 主体风险

脑机接口在主体功能恢复或能力增强的场景中,常引发诸如思维活动被算法实时解析、行为决策受外部设备隐性干预、长期使用导致自我身份认知模糊等有关认识论、行动论与本体论的问题。深入厘清此类风险,不仅关乎个体尊严与基本权利的维护,更是确保技术不僭越人类主体地位、避免人机关系异化的关键前提。故此,本研究进一步聚焦于认知、自主性、人格三方面对脑机接口主体风险展开系统性剖析。

(1) 认知

脑机接口作为大脑与外界直接交互的媒介,重塑了主体认知的物理载体与信息路径,深刻改变了主体的认知方式。一方面,脑机接口作为认知辅助工具可重构主体认知模式,例如帮助闭锁综合征(locked-in syndrome, LIS)患者克服肢体功能障碍,重建与现实的感知连接。另一方面,脑机接口向主体内部输入信息的过程可能破坏认知系统平衡并干扰固有神经机制。巴克斯(Alexander Backus)的研究表明,外部记忆存储的编码过程会改变海马体

振荡的相位同步性,并干扰其生物记忆编码与提取功能,进而导致主体记忆时序混乱并诱发认知紊乱^[10]。上述问题迫使学界重新审视脑机接口介入下的主体神经权利,其核心在于确立主体对自身神经信号的专属支配权、保障认知机制的不可侵犯性,以及捍卫主体在技术干预中的自主决策权。基于

此,卡希纳德里 (Guido Cassinadri) 等学者强调,必须构建一个超越单纯技术功能主义的伦理审查框架,尤其需要关注神经权利侵害后产生的“审后义务”^[11],即开发者不仅要对于干预本身负责,还需对技术干预后的神经状态变化承担持续监测与修复责任。

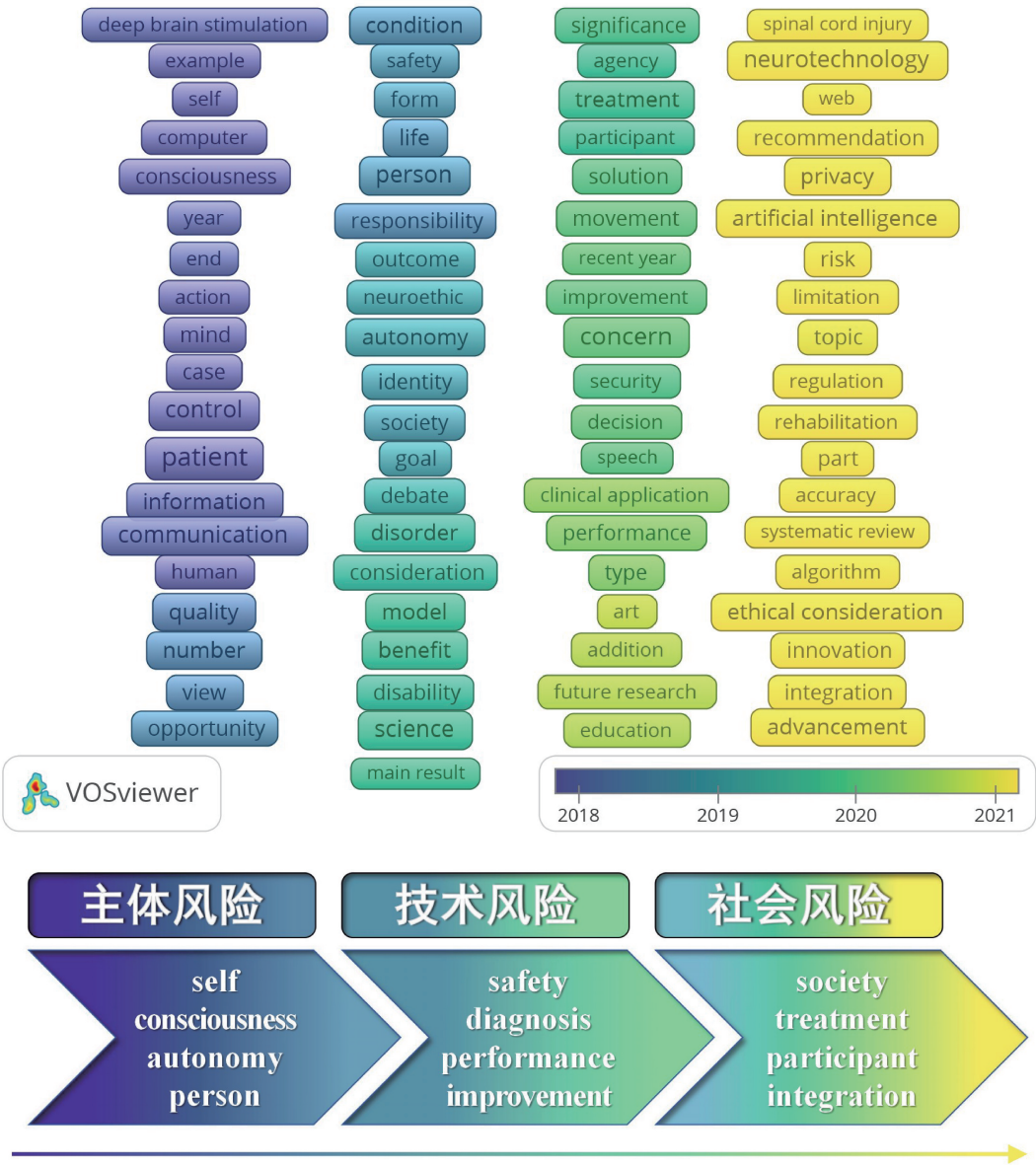


图 4 脑机接口伦理研究主题分布与演化

Figure 4 Thematic distribution and evolution in ethical studies on brain – computer interface

(2) 自主性

自主性意味着主体能依照自我意愿行事,包括运用信息与知识生成行为动因的能力、确保预期行动有效实现的控制能力,以及在具体关系和情境中实施意图的能力^[12]。脑机接口伦理对主体自主性

的探讨,聚焦于技术介入及算法分类引发的三重能力的双向变化上,即自主性的增强与削弱。这种双向影响在实践中呈复杂性:一方面,脑机接口可通过功能恢复、反馈等途径增强主体自主性。但主体若过度依赖这种“优化后的自主”,可能导致自然决策

能力退化。例如,当算法为提升任务完成效率而预筛选项时,用户生成行为动因的能力也将随之减弱。另一方面,因算法限制与控制权转移导致的自主性削弱,将使主体从“自主的人”降格为“被控的人”,进而冲击其道德主体性:主体既难以承担行为的道德责任,也无法在伦理实践中确立自身的主体性地位。例如,技术失控导致的不良后果,常因神经信号溯源困难、责任边界模糊而陷入追责困境。因此,技术的前瞻性伦理设计需以保护自主性三重能力为核心,构建全周期的自主权保障机制。

(3) 人格

人格作为个体身份和自我认同的核心,界定着“我之为我”的根本标尺。脑机接口的兴起引发了关于人格与身份变化的讨论,其焦点已超越传统理性能力范畴,转向人之构成的关系维度。脑机接口通过神经可塑性调节与心理状态干预,不仅重构大脑功能网络,更可能深度解构主体既定的社会身份叙事、个性特质及其真实性。在脑机接口介入下,人机界限的持续模糊催生了双重危机。一方面,它深刻改变个体身份及其身体图式,使主体重构对身体边界的认知,进而模糊“我”与“非我”的原初分野。例如,瘫痪患者通过脑机接口操控机械臂完成抓握动作时,会将其触觉反馈内化为躯体感觉,而当设备故障时甚至产生出类似肢体残缺的幻痛。另一方面,自我表征系统失调会动摇人格同一性的稳定基底。那些曾划定“我是谁”的核心特质,可能在脑机接口的介导下变得难以辨识,即个体的象征秩序(symbolic order)被打破^[13]。例如,接受植入式记忆的用户,可能在真实经历与植入记忆的交织中怀疑自我真实性,导致其符号意义网络断裂与身份叙事混乱。对此,吉尔伯特呼吁开发工具以获取人格变化的实证论据,进而为后续理论探讨提供经验支持^[14]。

3.2 技术风险

脑机接口在临床植入、日常交互与长期适配等应用场景中,常引发诸如设备植入引起的机体排异、数据传输中断导致的功能失效、硬件故障造成的不可逆损伤等技术层面的问题,直接关涉使用者安全与健康。因此,本研究进一步聚焦于植入安全性、传输稳定性、操作可控性三个维度,对脑机接口的技术风险加以系统性剖析。

(1) 设备植入安全性

设备植入安全性的核心在于生物相容性,即脑

机接口与人体组织的相互作用需处于安全状态,避免损害大脑及其他生物组织。首先,从植入结果来看,电极作为植入脑组织的异物与神经系统直接接触后,易引发三类问题:一是引发免疫反应,导致慢性炎症、细胞死亡或纤维化等神经损伤;二是设备耐用性不足,表面易腐蚀、老化;三是电极及相关电路在传输电信号时,不可避免地对人体产生刺激性干预。对此,设计者需将“不伤害”原则嵌入技术改进,而非一味追求功能优化。同时,还需对术后患者进行长期随访,以防各类感染风险。其次,在植入过程中,因手术精度要求高、侵入性强,易引发安全隐患。对此,需围绕操作安全性构建伦理规范,涵盖手术器械的生物相容性评估、术中感染风险防控等。最后,在植入前准备阶段,须克服因受试者认知不足所引发的心理恐慌。这要求知情同意环节的落实,即以受试者可理解的方式告知潜在风险、预期收益及不可逆性,确保其自主选择。此外,相关研究致力于规范植入手术中的术语^[15],为手术方案制定、术后评估及学术交流构建了统一的语言参照体系。

(2) 数据传输稳定性

大脑活动与个体的思想、情感乃至记忆存在深度关联,并能间接反映这些心理活动特征。因此,脑机接口在采集和解读人脑活动时,所生成的数据具有高敏感性。脑机接口伦理对数据传输的考察主要集中于三点:一是数据隐私危机。脑机接口对脑电波与神经信号的监测与解码,能够推测出用户的有关信息。无论是无意的数据外泄,抑或是恶意的信息窃取,均可能导致个体最为私密的内心世界与私人经历暴露于外。二是数据滥用困境。若数据收集过程缺乏严格的授权机制与规范程序,恶意行为者可能利用所获取的大脑活动数据,实施商业化个性推荐或其他非法行为,从而侵犯个体权益。三是黑客攻击与操控。脑机接口设备作为数据传输的载体,已有研究证实其可破解性^[16]。恶意软件或攻击者不仅可入侵设备系统进行数据篡改和窃取,还可操控设备,干预用户思想与情感。对此,研究者强调必须关注如何收集、保护和存储神经数据^[17]。具体而言,应确立以用户为中心的伦理原则,包括赋予用户对自身神经数据的完全控制权、构建基于用户授权的分级访问机制、建立用户参与的数据治理架构。

(3) 硬件操作可控性

硬件操作可控性直接关涉责任问题,特别是当脑机接口设备运行失控或出现误操作时,研究者主

要围绕“何者之责”以及“如何追责”两大问题展开探讨。前者指向责任厘清,若脑机接口设备故障导致用户身心受损时,责任应由谁承担?这涉及到硬件开发商、医疗机构、使用者等多方的法律责任划分。为此,需建立明确的问责机制,并制定相关法规和保险政策,以实现责任界定的制度化与规范化。后者主要集中于事故追责,即当使用者经由脑机接口操作而对第三方造成影响时,应如何厘定其责任。针对此问题,研究者不仅从行动哲学视角重新诠释了脑机接口介导的行为,还指出由行动意图发起的行动(go-command)本身即构成责任^[18]。此外,另有学者针对这种因果责任而展开实验以测量脑机接口介导下的具身虚拟运动中用户的主体感与责任感^[19],为责任认定提供实证支撑。此类研究为责任主体的界定、追责机制的构建提供了理论指导。

3.3 社会风险

脑机接口产业的快速兴起与应用场景的持续拓展,使其在技术普及、资源分配、社会融入等场景中,常引发社会不公、人际互动异化与社会排斥等风险。深入探讨此类风险,不仅关乎社会秩序的稳定与和谐,更是对产业发展所衍生出的各类社会问题的必要回应。故此,本研究进一步聚焦于社会公平、社会互动、社会接纳三个维度,对脑机接口的社会风险展开系统性分析。

(1) 社会公平

脑机接口虽为医疗和增强领域带来巨大潜力,却可能引发以技术获取的机会不平等为核心的社会公平性问题。这一担忧的立论在于,该技术在恢复与增强场景中均展现显著优势,而这两种应用场景均与社会公平深度相关。在恢复场景中,不平等的根源在于资源分配的结构失衡。资本逐利性驱动研发资源向利润更高的增强领域倾斜,导致神经修复型脑机接口等医疗康复技术迭代缓慢、成本居高不下。这一现象违背了罗尔斯“差异原则”的核心诉求,即不平等的安排需最有利于处境最不利者^[20]。而随着脑机接口研究的重心从恢复转向增强,社会公平的讨论重心也相应转变。在增强环节中,脑机接口造成的社会不公集中体现为技术获取门槛所造就的能力鸿沟。依据森(Amartya Sen)的“能力平等观”,公平要求个体实现有价值生活的实质能力平等^[21]。然而,脑机接口的技术获取差距会削弱部分群体实质能力,未来可能拉大阶层分化,形成“社会分层”,并进一步固化能力发展起点的不平

等,最终冲击社会的整体公平与和谐。正因如此,如何确保脑机接口的普惠性,并让其成果覆盖更广泛的群体,已成为众多研究者关注的焦点。

(2) 社会互动

作为带有社会属性的个体,人始终以直接或间接的方式参与社会互动。脑机接口通过重塑个体互动的载体与方式,成为影响社会交往的新变量。钱德勒(Jennifer Chandler)等人就脑机接口对通信障碍群体的影响进行考察后发现,该技术能有效优化闭锁综合症患者的社交互动^[22]。无论是直接的信息输出,还是转化为具象化动作,脑机接口都深度嵌入并重构了主体的交往过程。值得注意的是,交往动作本身即是理解的有机组成部分,肢体姿态、行为反馈等具象化表达与语言信息共同构成了沟通的完整维度,是实现相互理解的基础。然而,脑机接口的稳健性与准确性仍待检验。正如马图特(Jérémy Mattout)所指出的,其可靠性需从机器、设计者与使用者三重维度进行审视^[23]。在社交场景中,若因机器性能不稳定导致信息传达错误,不仅难以达成交往目的,还会加剧社交压力。这种交往困境印证了哈贝马斯(Jürgen Habermas)的论断,即交往需以理解为导向^[24]。交往的有效性依赖于沟通的真实性、正确性等要求,而技术应用的不稳定性会同时破坏语言信息与交往动作的准确性,最终阻碍交往双方达成相互理解。

(3) 社会接纳

社会接纳,即社会对一项技术的接受度,通常取决于其可行性与有用性。二者共同塑造着社会对脑机接口的接纳程度:可行性体现为技术应用的稳定性、易用性与安全性,是社会接纳的基础;有用性则通过恢复残障群体功能、增强认知等实际价值,构成社会接纳的核心动力。然而,任一维度的显著缺陷均会削弱公众的接纳意愿。提升脑机接口的社会接纳度,需重点关注各方对技术改进的意见,将多元主体的反馈纳入优化路径。例如,奈博尔(Femke Nijboer)等人通过问卷调查,收集了受访者对于脑机接口领域亟需解决问题的看法,包括澄清其标准定义、制定适用的道德指南、设立认证流程以明确各方资质,以及创办专业学术期刊^[25]。科格尔(Johannes Kögel)的访谈研究则表明,技术带来的有益体验本身即可促进用户接纳^[26]。随着脑机接口研究的深入,部分学者的研究方法已从定性转向定量,旨在通过对数据隐私、透明度及公平获取等伦理议题的量

化探讨,为提升技术的社会包容性提供实证依据。

4 总结与展望

总体来看,脑机接口伦理研究呈现出多维交叉、分层演进的特点。通过梳理,其发展轨迹可清晰归纳为两条核心路径:其一,由经典科学到伦理争议,即基础科学的突破推动生命伦理学、认知神经科学、机器伦理学等领域深入反思人、机及人机关系。其

二,从工程实现到临床应用,即相关应用研究在技术落地过程中所暴露的诸多现实问题,引发了从技术可行性到应用正当性的持续伦理追问。结合脑机接口伦理研究框架与模式(见图 5),当前研究存在三方面明显不足:(1)跨学科的实证研究较为匮乏;(2)产业化伦理治理缺失;(3)“技术先行-伦理跟进”的研究模式困于预设角色框架。针对上述问题,未来的研究可从以下方向寻求突破。

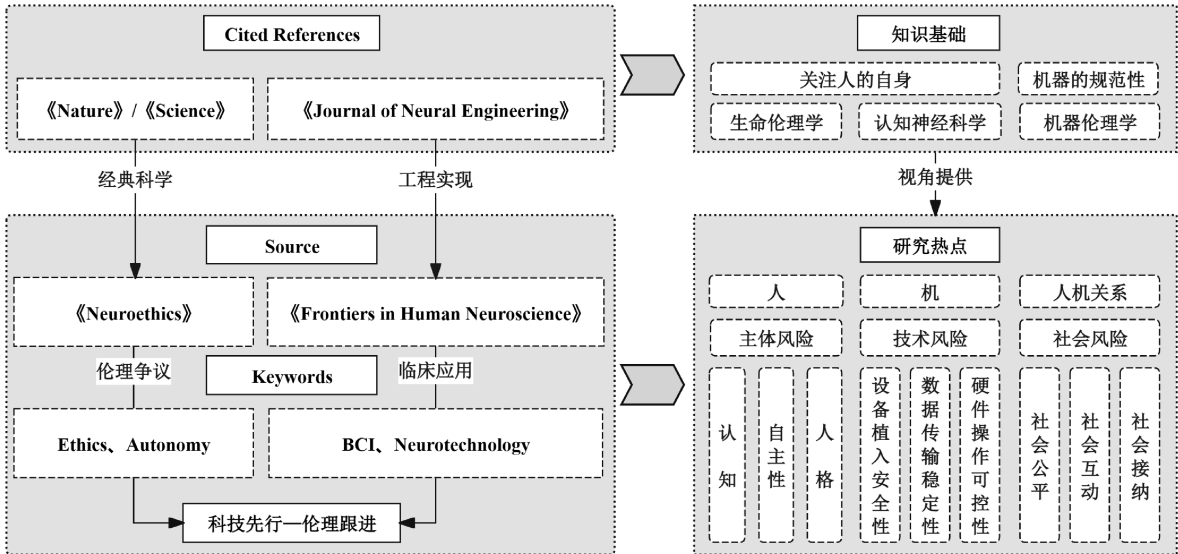


图 5 脑机接口伦理研究框架及研究模式
Figure 5 Framework and paradigm of ethical research on brain – computer interface

4.1 构建跨学科的实证协同网络

当前关于脑机接口伦理的讨论大多停留在理论推演阶段,且缺乏充分的实证数据支持。鉴于此,应建立“理论-实证”双向验证机制,以推动技术与伦理探讨的协同发展。该机制一方面需要完善现有理论模型,通过“计算实验”^[27]进行建模与仿真,以验证其逻辑自洽性,并探索脑机接口应用中的隐私敏感性、算法公平性等伦理风险临界点。另一方面,应基于临床数据开展“社会实验”^[27],通过结构化访谈、量表问卷等方式采集用户真实反馈,将抽象伦理假设转化为可量化指标,从而实现对理论构想的实证校验。此外,脑机接口与人工智能技术的融合趋势日益显著,从利用人工智能算法实时解析神经信号、重构视觉图像,到依托智能模型优化人机交互响应效率,皆体现出二者的深度整合。可进一步参考人工智能伦理治理的实验设计与实施流程,构建涵盖实验假设、方法、平台、操作与结果分析等核

心环节的脑机接口伦理治理实验框架。

4.2 驱动脑机接口产业优化升级

当前伦理研究对脑机接口产业化的系统性风险、商业模式缺陷及产业链协同治理等问题缺乏针对性剖析,导致伦理治理未能有效嵌入技术商业化进程。产业链上下游因责任模糊而形成的伦理治理“真空地带”,便是这一问题的典型表现。要充分释放其技术潜能,需以伦理重构产业体系:其一,构建覆盖技术研发、产品生产、市场推广到售后服务的全链条伦理责任追溯机制。通过细化各环节责任主体与权责边界,平衡企业利益与社会责任,例如在生产环节嵌入伦理合规性检测节点。其二,推动产业链各方凝聚伦理共识,联合制定行业伦理标准,将数据隐私保护、算法公平等指标纳入技术成熟度评估体系,从而将合规能力转化为技术竞争力。唯有以伦理为价值导向来重塑产业逻辑,才能避免技术迭代与社会需求的结构性背离,并推动脑机接口从技术

突破到规模化产业价值的实质性跃迁,最终构建可持续的新质生产力体系。

4.3 推进去中心化网络治理模式

现有脑机接口伦理研究模式,已难以应对脑机融合背景下网络数据与生物实体之间的互联性增强所带来的挑战。作为一项集体思维技术(Technologies for Collective Minds,TCMs),脑机接口通过人脑与计算机的连接赋能集体思维协同,并在神经整合层面实现了脑对脑的直接通信^[28]。这种跨主体的神经级互联与动态协作,使得参与主体的权责边界在实时交互中持续重构。传统中心化治理模式因权力集中、响应滞后且难以适配多主体动态关系,而无法有效应对此类高频、分布式协同场景中所涌现的伦理问题。针对这种动态协作特性,需推进“去中心化的网络治理”(Decentralized Network Governance)模式。在此模式下,用户角色将依据网络的性质和内部关系动态演变^[29]。同时,治理任务与权责分配亦高度敏感于参与者间的权力结构。这一模式能为伦理研究提供动态分析框架,助力追踪多主体间伦理边界的迁移,从而为构建适配技术持续演进的伦理体系开辟新路径。

参考文献:

- [1] Jebari K. Brain machine interface and human enhancement - An ethical review [J]. *Neuroethics*, 2013, 6(3): 617 - 625.
- [2] 加快发展新质生产力扎实推进高质量发展 [N]. 人民日报, 2024 - 02 - 02(001). Accelerating the development of new quality productive forces and making solid progress in promoting high - quality development [N]. *People's Daily*, 2024 - 02 - 02(001).
- [3] Apuzzo M L J. New dimensions of neurosurgery in the realm of high technology: Possibilities, practicalities, realities [J]. *Neurosurgery*, 1996, 38(4): 625 - 639.
- [4] Bostrom N, Sandberg A. Cognitive enhancement: Methods, ethics, regulatory challenges [J]. *Science and Engineering Ethics*, 2009, 15(3): 311 - 341.
- [5] Yang G - Z, Bellingham J, Dupont P E, et al. The grand challenges of Science Robotics [J]. *Science Robotics*, 2018, 3(14): eaar7650.
- [6] Sankey M. Introductory Note on the Thermal Efficiency of Steam - Engines [C]. *Minutes of Proceedings of the Institution of Civil Engineers*. 1898, 134: 278 - 283.
- [7] 翟晓梅,邱仁宗. 生命伦理学导论[M]. 北京:清华大学出版社, 2020: 3. Zhai X M, Qiu R Z. *Bioethics: an Introduction* [M]. Beijing: Tsinghua University Press, 2020: 3.
- [8] 加扎尼加,伊夫里,曼根. 认知神经科学:关于心智的生物学[M]. 北京:中国轻工业出版社, 2011: 3. Gazzaniga M S, Ivry R B, Mangun G R. *Cognitive Neuroscience: The Biology of the Mind* [M]. Beijing: China Light Industry Press, 2011: 3.
- [9] Anderson M, Anderson S L. Machine ethics: Creating an ethical intelligent agent [J]. *AI Magazine*, 2007, 28(4): 15.
- [10] Backus A R, Bosch S E, Ekman M, et al. Mnemonic convergence in the human hippocampus [J]. *Nature Communications*, 2016, 7(1): 11991.
- [11] Cassinadri G, Ienca M. Non - voluntary BCI explantation: Assessing possible neurorights violations in light of contrasting mental ontologies [J]. *Journal of Medical Ethics*, 2024: jme - 2023 - 109830.
- [12] Friedrich O, Racine E, Steinert S, et al. An analysis of the impact of brain - computer interfaces on autonomy [J]. *Neuroethics*, 2021, 14(1): 17 - 29.
- [13] Schermer M. The mind and the machine. On the conceptual and moral implications of brain - machine interaction [J]. *Nanoethics*, 2009, 3(3): 217 - 230.
- [14] Gilbert F, Viaña J N M, Ineichen C. Deflating the “DBS causes personality changes” bubble [J]. *Neuroethics*, 2021, 14(1): 1 - 17.
- [15] Leuthardt E C, Moran D W, Mullen T R. Defining surgical terminology and risk for brain computer interface technologies [J]. *Frontiers in Neuroscience*, 2021, 15: 599549.
- [16] Martínez Beltrán E T, Quiles Pérez M, López Bernal S, et al. Noise - based cyberattacks generating fake P300 waves in brain - computer interfaces [J]. *Cluster Computing*, 2022, 25(1): 33 - 48.
- [17] Naufel S, Klein E. Brain - computer interface (BCI) researcher perspectives on neural data ownership and privacy [J]. *Journal of Neural Engineering*, 2020, 17(1): 016039.
- [18] Bublitz C, Wolkenstein A, Jox R J, et al. Legal liabilities of BCI - users: Responsibility gaps at the intersection of mind and machine? [J]. *International Journal of Law and Psychiatry*, 2019, 65: 101399.
- [19] Nierula B, Spanlang B, Martini M, et al. Agency and responsibility over virtual movements controlled through different paradigms of brain - computer interface [J]. *Journal of Physiology*, 2021, 599(9): 2419 - 2434.
- [20] 罗尔斯. 正义论[M]. 北京:中国社会科学出版社,

- 1988;56 – 61. Rawls J. A Theory of Justice [M]. Beijing: China Social Sciences Press, 1988; 56 – 61.
- [21] 森. 正义的理念[M]. 北京:中国人民大学出版社, 2012; 276 – 278. Sen A. The Idea of Justice [M]. Beijing: China Renmin University Press, 2012; 276 – 278.
- [22] Chandler J A, Van der Loos K I, Boehnke S, et al. Brain computer interfaces and communication disabilities: Ethical, legal, and social aspects of decoding speech from the brain [J]. *Frontiers in Human Neuroscience*, 2022, 16: 841035.
- [23] Mattout J. Brain – computer interfaces: A neuroscience paradigm of social interaction? A matter of perspective [J]. *Frontiers in Human Neuroscience*, 2012, 6:114.
- [24] 哈贝马斯. 交往行为理论[M]. 上海:上海人民出版社, 2018; 125 – 129. Habermas J. The Theory of Communicative Action [M]. Shanghai: Shanghai People's Publishing House, 2018; 125 – 129.
- [25] Nijboer F, Clausen J, Allison B Z, et al. The Asilomar Survey: Stakeholders' opinions on ethical issues related to brain – computer interfacing [J]. *Neuroethics*, 2013, 6(3): 541 – 578.
- [26] Kögel J, Jox R J, Friedrich O. What is it like to use a BCI? – Insights from an interview study with brain – computer interface users [J]. *BMC Medical Ethics*, 2020, 21(1):2.
- [27] 于雪. 人工智能伦理治理的实验方法论探究[J]. *科学学研究*, 2024, 42(9): 1793 – 1799, 1807. Yu X. Experimental methodology of ethical governance in artificial intelligence [J]. *Studies in Science of Science*, 2024, 42(9): 1793 – 1799, 1807.
- [28] Lyreskog D M, Zohny H, Mann S P, et al. Decentralising the self – ethical considerations in utilizing decentralised web technology for direct brain interfaces [J]. *Science and Engineering Ethics*, 2024, 30(4): 28.
- [29] Zwitter A, Hazenberg J. Decentralized network governance: Blockchain technology and the future of regulation [J]. *Frontiers in Blockchain*, 2020, 3: 12.

Ethical research on brain – computer interface: Interdisciplinarity and risk evolution

WANG Gao – feng¹, YAO Ming – yue¹, GAN Ye – tong²

(1. Department of Philosophy of Science and Technology, University of Science and Technology of China, Hefei 230026, China;

2. Institute for Advanced Studies in Humanities and Social Sciences, Guangdong

University of Technology, Guangzhou 510006, China)

Abstract: With breakthrough advances and accelerated industrial expansion in brain – computer interface (BCI) technology, its ethical implications have increasingly garnered academic attention. Based on scientometrics and combined with interdisciplinary and risk – evolution analytical perspectives, this study examines the landscape, knowledge base, and research hotspots in international research on the ethics of brain – computer interface: (1) Publications in research on the ethics of brain – computer interface demonstrate a three – stage evolutionary trajectory, with fluctuations closely linked to technological cycles, social events, and policy shifts, culminating in 2024 as a window of explosive growth. (2) Bioethics, cognitive neuroscience, and machine ethics form a modular interdisciplinary citation network spanning normative values, technical pathways, and human – machine collaboration frameworks, thereby supporting a multidimensional theoretical foundation. (3) Research hotspots have formed clusters in the domains of subjectivity, technology, and society, highlighting the dynamic characteristics of risk evolution. (4) With the deepening of brain – computer integration, issues concerning the neural rights of integrated subjects and social acceptance urgently require attention. Based on the assessment of the current situation, future research should focus on fostering interdisciplinary empirical collaboration, driving industrial optimization and upgrading, and promoting decentralized network governance, so as to contribute to the improvement and development of global research on the ethics of brain – computer interface.

Key words: brain – computer interface; technology ethics; interdisciplinarity; risk evolution; scientometrics