

基于语料库的中美主流媒体“人工智能”话语对比研究

王园园

天津外国语大学 天津 300204

【摘要】：本研究运用语料库方法，对中美主流媒体 2018—2023 年中人工智能话语进行系统性对比。基于 NOW 新闻语料与 BCC 新闻语料，构建规模约 15 万词中美媒体可比语料库，借助 AntConc4.2.1 软件，以人工智能/Artificial Intelligence 为节点词，在 L5-R5 搭配窗下提取 MI≥3 的高频搭配词，从议题框架、类联接、语义韵三方面展开分析。研究发现，美媒聚焦科技巨头竞争、风险伦理与规制，中媒侧重国家战略、产业赋能与社会治理；美媒呈现复杂偏负的语义韵，中媒呈显著积极的语义韵；在类联接模式上，美方多用被动警示结构，中媒多用主动发展结构。本研究揭示了中美媒体人工智能话语建构差异，反映两国科技路径、媒体生态及国家战略的不同，为跨文化科技传播提供了实证参考。

【关键词】：语料库语言学；话语分析；人工智能；中美媒体；对比研究

DOI:10.12417/3041-0630.26.07.077

1 引言

《新一代人工智能发展规划》指出，人工智能已成为国际竞争焦点与战略性技术，中美作为关键引领者备受瞩目^[1]。新闻媒体通过议程设置与框架建构塑造公众认知，承担着重要的议程设置与框架建构功能影响社会认知，其报道并非客观现实的简单呈现，而是带有社会文化与意识形态倾向的建构过程，系统对比中美媒体 AI 话语对理解该技术跨文化传播与公众接纳具有重要现实意义。现有研究多为质性分析，缺乏大规模语料库量化对比。本研究采用语料库辅助话语研究范式，依托双语新闻语料库，对比分析中美媒体 AI 话语的议题框架、语言特征及深层动因。

2 文献综述

2.1 基于语料库的话语研究

传统批评话语分析将话语视作社会实践，认为其兼具反映与建构社会现实的功能，权力和意识形态通过话语实践隐蔽再生产^[2-4]，且侧重少量文本的定性分析，易受研究者主观性影响。由此，语料库辅助话语研究被引入，实现量化与质性分析的互补结合^[5]。语料库工具可识别文本数据中隐形词汇语法模式，这正是话语模式和意识形态的语言表征^[6]。学者通过英国媒体语料库，量化地揭示媒体话语如何将群体与负面概念联系起来，建构排斥消极性的社会表征^[5]。其中语义韵能揭示节点词在习惯共现语境中的情感语义氛围，表明该词不易被直觉感知的褒贬色彩^[5]。研究发现，中性词因高频与负面词语搭配，形成了负面语义韵^[6]。

2.2 媒体话语的语料库研究

媒体话语通过词汇、句法的选择来构建议题，是 CADS 的

核心研究对象。研究主要聚焦：一是对特定社会群体的表征研究。通过历时研究系统揭示了媒体话语将穆斯林与负面框架关联的话语模式^[7]。二是对重大社会事件的框架建构研究。媒体倾向于使用隐喻来框架化经济危机，简化复杂议题的同时，隐蔽其危机的结构性社会根源；气候变化、移民危机、公共卫生事件等议题的研究，验证了语料库方法的强大效能；三是跨文化的对比研究，学者依托可比语料库，证实两国媒体话语策略的显著不同反映各自国内的政治生态^[8]。

2.3 “人工智能”相关话语研究

现有相关研究以定性分析为主，表明媒体报道常在“乌托邦”与“反托邦”两极之间摇摆^[9]。部分研究引入语料库方法对历时语料分析发现，媒体焦点已从逻辑推理能力转向其伦理隐患，公众对 AI 的焦虑情绪显著上升，随着 AI 进入跨文化传播视阈，媒体建构的作用愈发凸显^[9]。但当前 AI 话语研究多局限单一英文语境，缺乏深度词汇语法层面的库语言系统对比^[10]。因此，本研究基于语料库辅助批评话语分析范式，对中美主流媒体的“人工智能”话语进行系统性的对比分析。

3 研究问题

(1) 中美主流媒体在建构“人工智能”话语时，分别突出哪些核心议题框架？

(2) 中美主流媒体在词汇语法选择及潜在的评价态度上呈现出何种显著异同？

(3) 中美媒体的 AI 话语建构模式差异背后蕴含何种社会文化与意识形态动因？

3.1 研究步骤

本研究选取 2018 年 1 月至 12 月 NOW 语料库与 BCC 语料库中新闻部分及公开新闻数据源，以“Artificial Intelligence/人工智能”为检索词，分别构建规模 15 万词的中美媒体可比语料库，数据清洗与分词处理参考相关本土建库规范^[11]。

3.1.1 数据分析工具与方法

使用软件 AntConc(V.4.2.1)进行数据分析：以节点词为中心，在 L5-R5 跨度内提取高频搭配词，采用互信息值作为衡量搭配紧密度的主要统计指标^[12]，筛选出 MI 值≥3 的前 20 位搭配词进行语义范畴归类，将其划分为不同的主题框架；其次分析类联接特征；最后进行语义韵分析，结合真实语境判定语义韵的语义褒贬色彩^[13]。

3.2 研究结果与讨论

3.2.1 高频搭配词与语义范畴对比

提取节点词的高频搭配词并进行语义范畴归类：

表 1 中美媒体 AI 话语高频搭配词对比 (MI 值≥3)

排名	美国媒体语料库搭配词	中国媒体语料库搭配词
1	Google's	发展
2	race	技术
3	risks	产业
4	ethical	应用
5	companies	国家
6	human	赋能
7	bias	战略
8	threat	创新
9	Microsoft	企业
10	battle	治理
11	weapons	融合
12	researchers	城市
13	data	机器人
14	jobs	安全
15	regulation	平台

表 2 搭配词的语义范畴归类对比

语义范畴	美国语料库代表词	中国媒体语料代表词
商业/竞争	Google's,race,companies,Microsoft,battle,jobs	产业, 企业, 平台
风险/伦理	Risks,ethical,bias,threat,weapons,regulation	安全
国家/政策	regulation	发展, 国家, 战略, 治理
技术/应用	human,researchers,data	技术, 应用, 赋能, 创新, 融合, 机器人

3.2.2 议题框架的深度解构

(1) 美国媒体市场竞争与风险规制的双螺旋叙事：美媒 AI 话语由两条主线构成：一是市场竞争框架，将 AI 塑造为成以科技巨头为主角的商业竞赛，如例 (1) The deal reflects Microsoft's ambition to lead in the new wave of artificial intelligence.例 2) Inside the multibillion-dollar battle between tech giants to create a true artificial intelligence.例 3) The global race for dominance in artificial intelligence has become a central point of tension between the U.S.and China.例 1)和 2)将 AI 的发展人格化为企业行为，而例 3)进一步将提升至地缘政治层面，此框架下 AI 是资本追逐的赛道；二是风险规制框架，此框架更为突出，反映了美国对 AI 潜在负面影响的忧虑，如例 4)Experts are increasingly concerned about the existential risks posed by advanced artificial intelligence.将 AI 从技术议题转化社会议题。

(2) 中国媒体国家擘画与产业赋能的宏大叙事：中媒 AI 话语呈现出统一的自上而下的宏大叙事，核心为国家战略引领下的产业发展，一是国家擘画，明确 AI 发展的顶层战略定位，如例 7)我国将人工智能提升到国家战略高度，系统布局，积极推动其健康发展，将 AI 视为技术进步，及关乎民族复兴和国家强大的进程；二是产业赋能框架，强调 AI 服务实体经济的功能，如例 9)人工智能正在为千行百业赋能，催生了无人驾驶、智能医疗等新业态，将 AI 描绘成能够激活实体经济、推动社会转型的强大引擎。

(3) 类联接模式对比：美媒中“Artificial Intelligence”常作受事主语，其高频共现动词多为 pose,raise,face 等，形成“AI+be+负面名词/形容词”的警示类联接，如例 11)The rapid deployment of Artificial Intelligence poses unforeseen risks to privacy and raises serious ethical questions that need to be regulated；在行动上偏好将 AI 置于防御性“防范”结构中，模式为“V(mitigate/regulate/address)+N(risk/impact/threat)+of+AI”，如例 12) Lawmakers are struggling to take immediate action to mitigate

the risks of AI and address its potential impact on the workforce.

中媒中,人工智能常作施事主语,搭配发展、推动、赋能等积极动词,形成“AI+正向动词/名词”的主动类联接,如例13)人工智能正在加快推动传统产业转型升级,作为核心驱动力全面赋能实体经济的高质量发展;在行动上高度聚焦“应用融合”结构,模式为“以+AI+为核心/引擎”或“AI+V(赋能/深度融合)+N(实体经济/制造业/千行百业)”,强调其工具性与建设性,如例14)我国应以人工智能为核心引擎,持续推动AI与先进制造业的深度融合,加速赋能千行百业。

4 语义韵的显著差异

美媒语境中,呈现出中性偏负的复杂语义韵,常在“While...but...”句式结构中,如例15)While artificial intelligence promises to revolutionize medicine,it also brings unprecedented privacy risks.凸显审视防范态度;中媒语境中,呈现积极推进式语义韵,绝大多数搭配正面动词和形容词,如例16)人工智能作为新一轮产业变革的核心驱动力,正在释放巨大的增长潜能,即便是讨论风险也多置于正面语境,视为能达成的积极目标,而非内在威胁。

参考文献:

- [1] 熊澄宇.全球人工智能发展现状与趋势[J].人民论坛·学术前沿,2018,(13):6-15.
- [2] Fairclough N.Critical discourse analysis:The critical study of language[M].London:Longman,1995.
- [3] Partington A.Corpora and discourse,a most congruous beast[M]//Partington A.Corpora and Discourse.Bern:Peter Lang,2004:11-20..
- [4] Baker P.Using corpora in discourse analysis[M].London:Continuum,2006.
- [5] Sinclair J.Corpora,concordance,collocation[M].Oxford:Oxford University Press,1991.
- [6] Stubbs M.Words and phrases:Corpus studies of lexical semantics[M].Oxford:Blackwell,2001.
- [7] Baker P,Gabrielatos C,McEnery T.Discourse analysis and media attitudes:The representation of Islam in the British press[M].Cambridge:Cambridge University Press,2013.
- [8] 刘涛.中美媒体对“一带一路”倡议的话语建构对比研究—基于语料库的分析[J].国际新闻界,2018,40(6):115-136.
- [9] Fast E,Horvitz E.Long-term trends in the public perception of artificial intelligence[C]//Proceedings of the AAAI Conference on Artificial Intelligence,2017:963-969.
- [10] 姜飞,王润泽.人工智能时代的传播范式变革[J].新闻与写作,2019,(2):18-23.
- [11] 胡开宝.基于语料库的当代中国媒体英语研究[M].上海:上海交通大学出版社,2010.
- [12] 卫乃兴.基于语料库的词汇语义学研究[M].上海:上海外语教育出版社,2011.
- [13] 卫乃兴.语义韵研究的一般方法[J].外语教学与研究,2002,34(4):300-307.
- [14] 李开复,王咏刚.人工智能[M].北京:文化发展出版社,2017.

5 讨论

中美主流媒体的人工智能话语建构的差异,本质是两国政经体制、媒体生态与国家战略的结构性差异:一是叙事主体与驱动机制差异,美媒以科技巨头为核心叙事主体,彰显市场主导的科技生态;中媒聚焦国家与产业,体现了我国国家引导型科技创新路径,凸显AI作为国家核心竞争力和产业升级关键的战略定位^[14];二是媒体角色与话语功能差异,美媒侧重批判性监督资本与技术权力,语义韵偏负面;中媒则以政策阐释、凝聚共识为核心,语义韵呈正面;三是战略定位与社会预期差异,美国作为技术领导者,对AI社会伦理风险存防御性焦虑;中国作为技术赶超者,立足技术跨越与产业升级,将风险视为可控的治理问题。

6 结论

研究验证了多维语料库分析模型在跨文化科技话语研究中的阐释力,丰富了相关研究案例,为理解中美在AI领域的认知差异提供实证了支撑,也为优化我国跨文化科技传播策略提供切实参考。