陕西省科学技术进步奖提名书

(2023年度)

**一、项目基本情况**

|  |  |
| --- | --- |
| 项目名称 | 社交媒体信息传播分析及智能引导关键技术 |
| 主要完成人 | 孙鹤立，何亮，秦兵，王迎雪，孙玉柱，张鹏，朱琳琳，何晖，林晖，金昊 |
| 主要完成单位 | 西安交通大学，哈尔滨工业大学，中国电子科技集团有限公司电子科学研究院 |

**二、提名意见**（适用于部门、机构提名）

|  |  |  |  |
| --- | --- | --- | --- |
| 提 名 者 | 陕西省教育厅 | 提名等级 | □一等奖☑二等奖及以上□三等奖及以上 |
| 提名意见：该项目针对社交媒体中信息的传播路径和影响力难以预测和控制，信息的情感倾向和传播效果难以评估和管理等挑战，研究了媒体信息传播结构及阻断方法，基于跨模态的共享-私有信息的自动情感预测及智能引导技术，基于情感引导的可控内容生成技术，解决了关键节点和传播社团同时检测、传播网络结构嵌入表示、传播社团检测联合优化、自动情感预测及智能引导、文本可控生成等多个难题。项目发表高水平论文多篇，取得专利多项，项目成果已应用于湖北省公安厅机场公安局、西安康奈网络科技有限公司、人民网股份有限公司、腾讯科技(深圳)有限公司、陕西省互联网信息办公室、陕西省科学技术情报研究院、银川一体化联合作战平台、达州立体化治安防控体系、合肥立体化治安体系等单位，取得重大社会经济效益。研究成果选题准确，研究起点高，创新性强，有重要的学术和工程应用价值，成果材料齐全、规范，无知识产权纠纷，人员排序无争议，符合陕西省科学技术进步奖提名条件，特提名为2023年度陕西省科学技术进步 二等及以上 奖项。 **说明：省科学技术奖一、二、三等奖项目，实行按等级标准提名、独立评审表决的机制。提名单者应严格依据省科学技术奖的标准条件，说明提名项目的贡献程度及等级建议。“仅提名一等奖”评审落选项目不再降格参评二等奖，“提名二等奖及以上”的评审落选项目不再降格参评三等奖。提名项目正式提交后，提名等级建议本年度不得变更。** |

**二、提名意见**（适用于专家提名）

|  |  |
| --- | --- |
| 姓 名 |  |
| 专家类型 | □国家最高科学技术奖获得者 □中国科学院院士 □中国工程院院士 □国家科学技术奖获奖项目第一完成人（需注明获奖等次） □省最高科学技术奖获奖人（或xxxx年省科学技术最高成就奖、xxxx年基础研究重大贡献奖获奖人） □Xxxx年省科学技术奖第一完成人（需注明获奖等次） | 提名等级 | □一等奖□二等奖及以上□三等奖及以上 |
| 责任专家 | □是 □否 |
| 提名意见： |
| **说明：省科学技术奖一、二、三等奖项目，实行按等级标准提名、独立评审表决的机制。提名单者应严格依据省科学技术奖的标准条件，说明提名项目的贡献程度及等级建议。“仅提名一等奖”评审落选项目不再降格参评二等奖，“提名二等奖及以上”的评审落选项目不再降格参评三等奖。提名项目正式提交后，提名等级建议本年度不得变更。** |

**三、项目简介**

|  |
| --- |
| 针对社交媒体中信息的传播路径和影响力难以预测和控制，信息的情感倾向和传播效果难以评估和管理等挑战，基于信息传播、社交网络分析、情感计算、内容生成的理论和技术，主要完成了以下三个方面的创新性工作：**创新点1：**基于传播网络结构特征，提出面向动态、异构传播网络的核心节点及社团检测，并对特殊节点进行分析和阻断，进而对信息传播实现有效的管控。**创新点2：**针对社交媒体的多样化信息，构造跨模态（文本、图像、音频）信息的交互。将具有较高关注权重的模态特征划分为共享特征，将具有较高预测损失的模态特征划分为私有特征。以文本模态为中心，另外两种模态的共享语义增强情感分析的鲁棒性，私有语义与文本语义互补，提供不同视角，提高情感预测的准确性。**创新点3：**基于多模态信息，利用带有情感导向的指示信息，并引入一个轻量级的调节器来调节控制器在不同译码位置的偏置信号，生成带有情感导向的可控内容。 |

**四、客观评价**

|  |
| --- |
| （限2页。围绕创新性、应用效益和经济社会价值进行客观、真实、准确评价。填写的评价意见要有客观依据，主要包括与国内外相关技术的比较，国家相关部门正式作出的技术检测报告、验收意见、鉴定结论，国内外重要科技奖励，国内外同行在重要学术刊物、学术专著和重要国际学术会议公开发表的学术性评价意见等，可在附件中提供证明材料。非公开资料（如私人信函等）不能作为评价依据。）1、技术评价 1）课题结题验收意见① 2022年3月29日，国家自然科学基金委员会对哈尔滨工业大学承担的国家自然科学基金资助项目“社交媒体中文本情感语义计算理论和方法（项目编号：61632011）” 按有关规定已审核完毕，准予结题。② 2022年5月31日，项目承担单位科大讯飞股份有限公司组织专家对课题“文本生成及复杂语言问题求解关键技术与系统”进行绩效评价。认为课题完成了任务书规定的考核指标，支撑了项目任务目标的完成。 2）检测报告信息产业信息安全测评中心对群体关系网络发现系统进行了第三方测试。经检验本项目的系统达到测试通过准则的要求。 3）科技奖励该项目部分成果获陕西省高等学校科学技术研究优秀成果奖一等奖。2、学术评价1） ACM/IEEE Fellow、伊利诺伊大学芝加哥分校计算机科学系杰出教授刘兵在其发表于《Deep learning for sentiment analysis: A survey》文章中肯定了我们提出的情感特定词嵌入（SSWE）的学习模型，不仅包含语义信息，而且将情感信息都嵌入到学习的词向量中。2）美国国家生物技术信息中心、国家医学图书馆、国立卫生研究院（NIH）研究员陈庆宇在其发表于《BioWordVec, improving biomedical word embeddings with subword information and MeSH》文章中指出我们的情感特定词嵌入方法比传统的独热编码表示更适合于各种自然语言处理任务。3）白宫科技政策办公室数据和民主部主任助理、美国东北大学计算机科学教授Alan Mislove在其发表于《Using millions of emoji occurrences to learn any-domain representations for detecting sentiment, emotion and sarcasm》文章中使用我们提出的八种情感表示创建了一个预训练数据子集，来分析使用多样化表情符号集的效果。 |

**五、应用情况**

|  |
| --- |
| **1．应用情况（限2页）**该项目已应用于湖北省公安厅机场公安局、西安康奈网络科技有限公司、人民网股份有限公司、腾讯科技(深圳)有限公司、陕西省互联网信息办公室、陕西省科学技术情报研究院、银川一体化联合作战平台、达州立体化治安防控体系、合肥立体化治安体系、军科委1XX、虚实融合的XXXX预测技术、云南省XXX大数据中心（二期）、海南XXXX信息化平台一期建设项目工程等转化推广，通过对社交媒体信息风险识别和治理关键技术进行一系列研究，获得了重要的学术成果应用，在应用过程中显著改善了社交媒体的潜在风险，推动网络舆情治理的进一步发展，产生良好的经济效益和社会效益。①团队提出的传播社团结构及核心节点检测算法用于其中的重要的传播结构分析及关重人物定位，并基于以上算法设计重大网络舆情早期检测方法，成功应用于西安康奈网络科技有限公司研发的网络内容生态综合治理平台建设中，为网络舆情提供早期检测与预警、溯源及发展趋势预测等解决方案，为舆情管理提供了有效的决策依据。截至2021年4月已支持平台完成有效预警1,365,768次，曾成功在西安地铁网络舆情事件中，于当日早5点成功预警，为相关单位进行舆情处置方案的提早准备提供了依据。网络内容生态综合治理平台现服务于新疆生产建设兵团互联网信息办公室、陕西省互联网信息办公室、宁夏回族自治区互联网信息办公室、陕西省公安厅、贵州省公安厅、青海省公安厅，取得重大社会效益。西安康奈网络科技有限公司通过对团队研究成果的应用，解决了关键节点和传播社团同时检测、传播网络结构嵌入表示和传播社团检测联合优化等问题，提升了公司产品的竞争力，促进了公司销售业绩的持续增长。②团队提出的网络节点传播影响力度量及关键节点发现模型，成功应用于湖北省公安厅机场公安局，针对业务中需要对网络舆情高风险关键重点人员检测与监控的真实应用场景需求，解决了“舆情传播网络中关键节点及传播社群同步检测”这一关键技术难题，团队研发的传播网络枢纽节点识别、基于密度的传播社群检测技术在湖北省公安厅机场公安局进行了成功部署应用，显著提升了舆情传播过程中重点枢纽节点的识别精度、有效拓展了传播社群的实时检测性能。截至 2022 年12月，湖北省公安厅机场公安局使用基于该模型的平台进行监控及检测，检测准确率提升20%以上，提高了检测效率，节省了大量人力成本。带来间接经济效益达3800万元，对湖北省公安厅机场公安局高效开展新技术预研及产品竞争力突破做出重要贡献。准确、提前定位网络舆情传播中的关键重点人物并加以监控，对控制舆情传播扩散起到了重要作用，带来了极大的社会效益。③团队提出的社交网络机器账号识别及特征可解释性学习技术，成功在银川一体化联合作战平台、达州立体化治安防控体系、合肥立体化治安体系、军科委1XX、虚实融合的XXXX预测技术、云南省XXX大数据中心（二期）、海南XXXX信息化平台一期建设项目工程等转化推广，为中国电科集团在中央企业人工智能领域CYD建设提供了有力支撑，转化推广金额约9亿元。自1XX工程02专项感知认知智能XX方向专家组设立XXX主题以来，项目组全程组织参与了“信息空间XX感知”专题的策划论证，在面向社交网络机器账号的安全风险感知领域大多观点被专家组采纳，论证材料多次向XXX首长、机关和专家委员会汇报，并获得肯定。项目技术成果成功在基于项目成果，构造覆盖六大专题模型库的模型体系，并取得一定的实战效果，共挖掘到潜在重点舆情相关人员1051人，其中高风险潜在重点人员59人，中风险潜在重点人员49人，低风险潜在重点人员848人。同时，技术在推动行业科技进步、改善民生方面做出重大贡献，基于该技术撰写的《美国网络空间治理思路分析一一基于推特删除账号的视角》研究报告，被中央网信办采用，为网络强国建设做出了一定的贡献。 |

1. **主要知识产权和标准规范等目录（限10条）**

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 序号 | 知识产权类 别 | 知识产权具体名称 | 国家（地区） | 授权号 | 授权日期 | 证书编号 | 权利人 | 发明人 |
| 1 | 发明专利 | 社交网络事件推荐方法、系统、设备及存储介质 | 中国 | ZL202110264628.2 | 2021.07.09 | 5822111 | 西安交通大学 | 孙鹤立; 姜晓龙; 何亮 |
| 2 | 发明专利 | 一种面向事件社交网络的规划方法及系统 | 中国 | ZL202110876077.5 | 2021.11.26 | 5637054 | 西安交通大学 | 孙鹤立; 郏靖宇; 姜晓龙; 何亮; 贾晓琳 |
| 3 | 发明专利 | 一种建立社交组的方法和装置 | 中国 | ZL201410242171.5 | 2014.09.17 | 2619432 | 西安交通大学; 腾讯科技（深圳）有限公司 | 孙鹤立; 陈川; 岳亚丁; 贺鹏; 管刚; 黄健斌; 吕泽; 钟翔 |
| 4 | 发明专利 | 一种用于舆情解析的多模态数据特征提取与关联方法 | 中国 | ZL202011299703.0 | 2022.09.16 |  | 中国电子科技集团公司电子科学研究院；天津大学  | 金昊；杨悦；宋丹；刘安安；李阳阳；王亚珅 |
| 5 | 发明专利 | 一种基于图神经网络的机器账号识别及伪装对抗方法 | 中国 | ZL202110442722.2 | 2022.08.16 |  | 中国电子科技集团公司电子科学研究院；深圳市网联安瑞网络科技有限公司 | 李阳阳；刘弋锋；武文翰；金昊；郭庆浪；杨智钦；彭浩；杨阳朝；石珺  |
| 6 | 发明专利 | 一种反溯源安全受控入网通道资源优化配置算法 | 中国 | ZL202011228961.X | 2021.03.09 | 5550019 | 中国电子科技集团公司电子科学研究院; 西北大学 | 金昊; 尹小燕; 金博; 罗茜; 韩青洁; 王亚珅; 李阳阳 |
| 7 | 论文 | Platform-Oriented Event Time Allocation（Extended Abstract） | 中国 | 10.1109/ICDE53745.2022.00139 | 2022.5 | 2022 IEEE 38th International Conference on Data Engineering  | 西安交通大学 | 孙鹤立；王宁；郏靖宇；黄健斌；熊辉；何亮；刘兴旺；张珊；乔少杰；赵继中 |
| 8 | 论文 | Sentiment Embeddings with Applications to Sentiment Analysis. IEEE Transactions on Knowledge and Data Engineering | 中国 | 10.1109/TKDE.2015.2489653 | 2015.10 | IEEE Transactions on Knowledge and Data EngineeringLinguistics: ACL 2022 | 哈尔滨工业大学 | Duyu Tang, Furu Wei, Bing Qin, Nan Yang, Ting Liu, and Ming Zhou |
| 9 | 论文 | Learning Semantic Representations of Users and Products for Document Level Sentiment Classification | 中国 | 10.3115/v1/P15-1098 | 2015.7 | Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing | 哈尔滨工业大学 | Duyu Tang, Bing Qin, Ting Liu |
| 10 | 论文 |  Learning Sentiment-Specific Word Embedding for Twitter Sentiment Classification | 中国 | 10.3115/v1/P14-1146 | 2014.06 | Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics | 哈尔滨工业大学 | Duyu Tang, Furu Wei‡, Nan Yang, Ming Zhou, Ting Liu†, Bing Qin |

**七、主要完成人情况表**

|  |  |  |  |
| --- | --- | --- | --- |
| 姓 名 | 孙鹤立 | 排 名 | 1 |
| 行政职务 | 副院长 |
| 技术职称 | 教授 |
| 工作单位 | 西安交通大学 |
| 完成单位 | 西安交通大学 |
| 对本项目主要学术贡献：(1)负责项目的整体规划、实施; (2)提出了传播结构分析及阻断方法，情感分析方法及引导内容自动生成方法等，支撑创新点1，2，3; |

|  |  |  |  |
| --- | --- | --- | --- |
| 姓 名 | 何亮 | 排 名 | 2 |
| 行政职务 | 无 |
| 技术职称 | 讲师 |
| 工作单位 | 西安交通大学 |
| 完成单位 | 西安交通大学 |
| 对本项目主要学术贡献：(1)参与项目的整体规划及实施;(2)提出了传播结构分析及阻断方法及引导内容自动生成方法等，支撑创新点1，3; |

|  |  |  |  |
| --- | --- | --- | --- |
| 姓 名 | 秦兵 | 排 名 | 3 |
| 行政职务 | 社会计算与信息检索研究中心主任 |
| 技术职称 | 教授 |
| 工作单位 | 哈尔滨工业大学 |
| 完成单位 | 哈尔滨工业大学 |
| (1)负责项目的整体规划、实施; (2)提出了情感分析方法及引导内容自动生成方法等，支撑创新点2，3; |

|  |  |  |  |
| --- | --- | --- | --- |
| 姓 名 | 王迎雪 | 排 名 | 4 |
| 行政职务 | 赋能中心主任 |
| 技术职称 | 高级工程师 |
| 工作单位 | 中国电子科技集团有限公司电子科学研究院 |
| 完成单位 | 中国电子科技集团有限公司电子科学研究院 |
| 对本项目主要学术贡献：(1)参与项目的整体规划及实施;(2)提出了传播结构分析及阻断方法等，支撑创新点1;(3)项目研究成果的推广应用; |

|  |  |  |  |
| --- | --- | --- | --- |
| 姓 名 | 孙玉柱 | 排 名 | 5 |
| 行政职务 | 无 |
| 技术职称 | 工程师 |
| 工作单位 | 西安交通大学 |
| 完成单位 | 西安交通大学 |
| 对本项目主要学术贡献：(1)参与项目的整体规划及实施;(2) 提出了引导内容自动生成方法等，支撑创新点3; |

|  |  |  |  |
| --- | --- | --- | --- |
| 姓 名 | 张鹏 | 排 名 | 6 |
| 行政职务 | 副处长 |
| 技术职称 | 工程师 |
| 工作单位 | 西安交通大学 |
| 完成单位 | 西安交通大学 |
| 对本项目主要学术贡献：(1)参与项目的实施，支撑创新点1;(2)项目研究成果的推广应用; |

|  |  |  |  |
| --- | --- | --- | --- |
| 姓 名 | 朱琳琳 | 排 名 | 7 |
| 行政职务 | 无 |
| 技术职称 | 博士研究生 |
| 工作单位 | 西安交通大学 |
| 完成单位 | 西安交通大学 |
| 对本项目主要学术贡献：(1)参与项目的实施;(2)提出了引导内容自动生成方法等，支撑创新点3； |

|  |  |  |  |
| --- | --- | --- | --- |
| 姓 名 | 何晖 | 排 名 | 8 |
| 行政职务 | 无 |
| 技术职称 | 高级工程师 |
| 工作单位 | 西安交通大学 |
| 完成单位 | 西安交通大学 |
| 对本项目主要学术贡献：(1)参与项目的实施;(2)提出了传播结构分析及阻断方法等，支持创新点1. |

|  |  |  |  |
| --- | --- | --- | --- |
| 姓 名 | 林晖 | 排 名 | 9 |
| 行政职务 | 社会治理事业部主任 |
| 技术职称 | 高级工程师 |
| 工作单位 | 中国电子科技集团有限公司电子科学研究院 |
| 完成单位 | 中国电子科技集团有限公司电子科学研究院 |
| 对本项目主要学术贡献：(1)参与项目的实施;(2)项目研究成果的推广应用; |

|  |  |  |  |
| --- | --- | --- | --- |
| 姓 名 | 金昊 | 排 名 | 10 |
| 行政职务 | 无 |
| 技术职称 | 高级工程师 |
| 工作单位 | 中国电子科技集团有限公司电子科学研究院 |
| 完成单位 | 中国电子科技集团有限公司电子科学研究院 |
| 对本项目主要学术贡献：(1)参与项目的实施;(2)项目研究成果的推广应用; |

**八、主要完成单位情况表**

|  |  |
| --- | --- |
| 单位名称 | 西安交通大学 |
| 对本项目主要学术贡献： 针对本项目研究内容，西安交通大学研究了社交媒体信息传播结构分析及传播阻断方法、内容自动生成方法，解决了关键节点和传播社团同时检测、传播网络结构嵌入表示、传播社团检测联合优化，网络评论内容自动生成等多个难题。 |

|  |  |
| --- | --- |
| 单位名称 | 哈尔滨工业大学 |
| 对本项目主要学术贡献： 针对本项目的研究内容，哈尔滨工业大学研究了情感分析方法及引导内容自动生成方法，解决了情感预测准确性低和以及无法精准引导的问题。 |

|  |  |
| --- | --- |
| 单位名称 | 中国电子科技集团有限公司电子科学研究院 |
| 对本项目主要学术贡献： 中国电子科技集团有限公司电子科学研究院研究了社交网络机器账号识别及特征可解释性学习技术，解决了社交网络机器账号的自动识别和特征分析问题。 |

**完成人合作关系说明**

本项目完成人从2011年开始，共同开展合作研究。项目第一、第二、第五、第六、第七、第八完成人均在第一完成人课题组学习或工作，在各级基金和学校的支持下，成立了数据智能与社会治理团队。何亮、何晖为早期成员，孙玉柱毕业后加入。团队研究方向一致，共同承担科研项目，共同发表学术论文，共同取得知识产权，构成本项目的列表研究成果，具体情况如下：

(1)各完成人在项目执行过程中，共同发表了项目中列出的系列论文、申请并获批了相关发明专利，并先后共同承担了若干国家重点研发计划，省级重点项目等；

(2)本项目第一完成人、第三完成人均为中文信息学会情感计算专委会成员，共同承担科研项目；

(3)本项目第一完成人、第四完成人，共同承担科研项目3项，合作学术论文1篇，申报发明专利1项；

(4)本项目第六完成人、第七完成人在第一完成人教授指导下攻读博士学位，分别合作获得论文成果1项，申报发明专利2项；

(5)本项目第四、第九、第十完成人，共同承担科研项目多项，合作发明专利多项。

**完成人合作关系情况汇总表**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| 序号 | 合作方式 | 合作者/项目排名 | 合作起始时间 | 合作完成时间 | 合作成果 | 证明材料 |
| 1 | 科研项目 | 孙鹤立，何亮，何晖 | 2013年 | 2021年 | 共同完成课题3项，完成论文3篇 |  |
| 2 | 科研项目 | 秦兵、孙鹤立 | 2019年 | 2022年 | 共同承担科研项目1项 |  |
| 3 | 科研项目 | 孙鹤立王迎雪林晖金昊 | 2018年 | 2022年 | 共同承担科研项目3项，合作学术论文1篇，申报发明专利1项； |  |
| 4 | 科研项目 | 孙鹤立，孙玉柱， | 2019 年 | 2020 年 | 完成论文2篇 |  |
| 5 | 科研项目 | 孙鹤立张鹏朱琳琳 | 2020年 | 2022年 | 指导攻读博士学位，分别合作获得论文成果1项，申报发明专利2项 |  |
| 不限条目 |  |  |  |  |  |  |