

人工智能赋能多模态大模型安全产教协同育人 创新实践

宋斌^{1,2} 张志勇^{1,2} 梁晗^{1,2}

¹河南科技大学 信息工程学院（人工智能学院），河南省 洛阳市 471023

²河南科技大学 人工智能创新研究院，河南省 洛阳市 471023

摘要 针对高校人工智能学科人才培养中教学内容滞后、实践资源不足及与产业脱节等问题，本研究以多模态大模型安全为切入点，系统分析现状后构建了“教学科研—技术创新—人才培养”三位一体协同育人模式。本文以河南科技大学与新华三集团联合开发的“雏盾·多模态大模型安全检测系统”为载体，介绍了共建课程、实训平台以及推动技术落地的实践。该模式有效提升了学生工程实践与创新能力，推动了关键技术研发、多行业应用与标准制定，为人工智能赋能教育高质量发展提供了可复制的范例。

关键字 人工智能安全，产教融合，多模态大模型，协同育人，实战能力

Artificial Intelligence Empowers Multimodal Large Model Security: Innovative Practice of Industry Education Collaborative Education

Bin Song^{1,2} Zhiyong Zhang^{1,2} Han Liang^{1,2}

College of Information Engineering and Artificial Intelligence
Henan University of Science and Technology
Luoyang 471023, China
songbin@haust.edu.cn, xidianzzy@126.com, hanliang@haust.edu.cn

Institute of Artificial Intelligence Innovations
Henan University of Science and Technology
Luoyang 471023, China

Abstract—In response to the problems of lagging teaching content, insufficient practical resources, and disconnection from the industry in the cultivation of artificial intelligence talents in universities, this study takes multimodal big model security as the starting point, systematically analyzes the current situation, and constructs a "teaching and research technology innovation talent cultivation" three in one collaborative education model. Using the "Luodun Multimodal Large Model Security Detection System" jointly developed by Henan University of Science and Technology and New H3C Group as a carrier, this article introduces the practice of jointly building courses, practical training platforms, and promoting technology implementation. This model effectively enhances students' practical and innovative abilities in engineering, promotes key technology research and development, multi industry applications, and standard setting, and provides a replicable example for empowering high-quality education development with artificial intelligence.

Keywords—Artificial Intelligence Security, Integration of Industry and Education, Multimodal Large Model, Collaborative Education, Practical Ability

1 引言

随着社会的不断进步和发展，对于人才的要求也日渐提升，不仅需要扎实的基础知识，还要具有较强的创新思维、合作意识、沟通表达等多方面的能力。

对此，学校要重视对学生的全方位培养，使其为新时代的发展贡献力量。对此，要重视与企业之间的合作，落实产教协同的育人模式，为社会发展培养更多优秀的人才^[1]。

人工智能已成为推动科技创新与产业变革的核心动力。国家《新一代人工智能发展规划》和《教育强国建设规划纲要（2024-2035年）》明确提出，加快高端AI人才培养，推动教育与科技、产业深度融合，构建面向未来的多学科融合教育体系^{[2][3]}。人工智能技术

***基金资助：**本文得到河南省高等教育研究项目重点项目“人工智能赋能高校教学模式创新和多学科交叉学术型人才培养机制研究与实践”（No.2025SXHLX064）、河南省高等教育教学改革研究与实践项目“产教融合驱动下高校人工智能安全治理通识教育创新与实践”、河南省科技攻关项目“基于多模态多尺度语义融合信息的情感分析方法研究与应用”（No.262102210089）、河南省校企协同创新项目“基于多模态多尺度语义融合的工业互联网安全感知研究与应用”（No.26AXQXT029）、河南省科技副总（No.HNFZ20250022）的资助。

*通讯作者：宋斌 songbin@haust.edu.cn。

正以前所未有的深度与广度重塑社会经济结构、工作方式以及知识传播模式，其在教育领域的应用孕育出深刻的人才培养革命^[4]。河南科技大学信息工程学院（人工智能学院）与新华三集团合作，围绕多模态大模型安全开展协同育人实践，旨在破解传统教育中理论与实际脱节的问题，推动教育高质量发展。科教融合与产教融合的深度结合是人工智能领域人才培养的核心路径^[5]，为了实现这一目标，需要在教学中充分发挥新质生产力的作用，依靠“产学研”融合之力，推动产学结合、产教相融、协同育人^{[6][7]}。

2 产学研存在的现实问题与研究目标

2.1 现实问题

课堂学习仅是起点，学生须在课外投入更多时间和精力进行自主学习和实践。当前，以大模型为代表的新一代人工智能技术正加速向金融、医疗、制造、交通等各行业渗透，深刻改变产业生态与人才需求格局，对高校人才培养体系在知识结构、技术能力和工程素养等方面提出了更高、更紧迫的要求^[1]。人工智能类课程本身具有知识迭代快、强应用导向的特点^[8]，然而现实中高校人工智能及相关专业的人才培养仍普遍面临多重现实困境：一方面，课程体系以单一学科理论为主，智能技术模块缺失或碎片化^[9]，课程内容更新速度滞后于技术迭代节奏，教材与行业真实问题脱节，导致学生所学难以直接应用于工程实践；另一方面，高校普遍缺乏高可用、高仿真的多模态AI训练与安全攻防实践资源，实验平台功能单一、数据陈旧，难以支撑前沿安全议题的深入探究。此外，传统教学模式以理论传授和封闭式实验为主，缺乏与企业真实需求和技术架构的深度对接，学生工程实现能力、系统思维与跨场景迁移能力明显不足。

尤其在人工智能安全、多模态大模型治理等交叉前沿领域，上述矛盾更为突出。网络空间安全已成为国家战略需求的重要领域，数据安全与隐私保护是该领域的热点方向^[10]，这些领域不仅要求学生掌握坚实的算法基础和系统知识，更需具备在开放、对抗、高维环境中解决实际安全问题的综合素养，包括漏洞挖掘与防御、模型鲁棒性增强、生成内容安全监测、隐私与合规处理等能力。现实中产学研各主体之间差别较大，策略也不尽相同。特别是大模型产业中，产方大多是大型互联网企业，技术壁垒较高、对高端人才的需求大，与传统企业在经营特征上有很大区别^{[12][13]}。

为破解上述难题，河南科技大学信息工程学院（人工智能学院）联合新华三集团，构建产教融合协同育人新机制。鉴于大模型具有启发性内容生成、对话情境理解、序列任务执行和程序语言解析能力，因此依托学院在人工智能领域的科研基础和教学体系，结合新华三集团“百业灵犀”大模型等产业级技术资源，

双方围绕人工智能安全开展系统合作。双方通过共建课程、共研项目、共享平台，推动产业真实需求与高校教学科研深度融合，提升学生工程实践与创新能力，探索“人工智能+教育”融合发展新路径，服务区域智能教育高质量发展战略^[14]。

2.2 研究目标

河南科技大学信息工程学院（人工智能学院）联合新华三集团，围绕人工智能安全领域开展战略合作，构建“教学科研—技术创新—人才培养”三位一体的协同育人模式，联合开发“雏盾·多模态大模型安全检测系统”，共建课程体系、教材资源与实训平台，推动AI前沿技术融入教学，提升学生综合实践能力，打造人工智能安全人才培养新范式。具体目标如下：

（1）创新产教协同育人模式，通过与新华三集团深度合作，构建高校与人工智能企业协同联动的人才培养机制，推动计算机科学与技术及网安密码领域学科建设。

（2）共建教学与实践资源体系，依托“百业灵犀”大模型等企业级AI平台，联合开发多模态安全检测等前沿课程、项目化教材及真实应用场景实践平台，构建贴合产业需求的实践教学体系，提升人才培养的针对性与实效性。

（3）打造典型示范案例，探索并形成一套可落地、可复制的“人工智能+教育”融合模式和实践成果，助力河南省高等教育高质量发展，为全国高校提供可推广的优秀范式。

3 协同育人设计与运行机制

校企双方建立常态化协作机制，明确分工与资源共享方式。共同开发课程与教材，建设AI安全实验室与边缘云安全原型系统，支持漏洞挖掘、攻防演练等高阶实践。学生通过参与实际研发与竞赛活动，显著提升工程实践与创新能力。该协同运行机制的逻辑架构与闭环流程如图1所示，外部“政策指导+产业需求”输入驱动，中层通过“资源建设、机制保障、实践融合、成效反馈”形成D-P-A-C循环运作，最终输出“技术突破、高端人才、行业标准、产业生态”成果，实现服务与赋能产业的目标。

3.1 校企协同机制落地实施

项目立足长效协作，与新华三集团共同成立了联合工作组和管理办公室，建立定期联席会议制度和专项沟通渠道，统筹人工智能安全领域技术研发与人才培养的战略规划与实施路径。双方明确校方主导理论教学与前沿研究、企业负责技术赋能与实战资源供给的分工机制，实现人员互聘、资源互通和成果互认，真正形成“共商、共建、共管、共享”的协同治理结

构^[17]。这一机制不仅保障了合作过程的规范性和持续性，也为教学资源建设、技术协同研发等环节提供了坚实的制度基础和组织支持。

3.2 教学资源与实践场景共建落地

在协同机制有效运作的基础上，双方系统推进教学资源与实战环境的一体化建设。依托新华三“百业灵犀”大模型、真实业务数据及企业级开发框架，共同构建覆盖模型安全、数据隐私、对抗防御等方向的

模块化课程群和项目化教材体系，将企业最新工程案例、技术规范与行业标准融入教学大纲和实验设计。同时，校企联合打造高度仿真的AI安全攻防实训平台和边缘计算安全测试床，支持多模态漏洞挖掘、实时威胁响应和系统韧性评估等综合实践，实现企业技术体系向教育场景的完整导入。该环节有效承接协同机制的顶层设计，将制度优势转化为优质教学资源与实战能力赋能平台，为技术研发与教学深度融合提供内容与环境支撑。

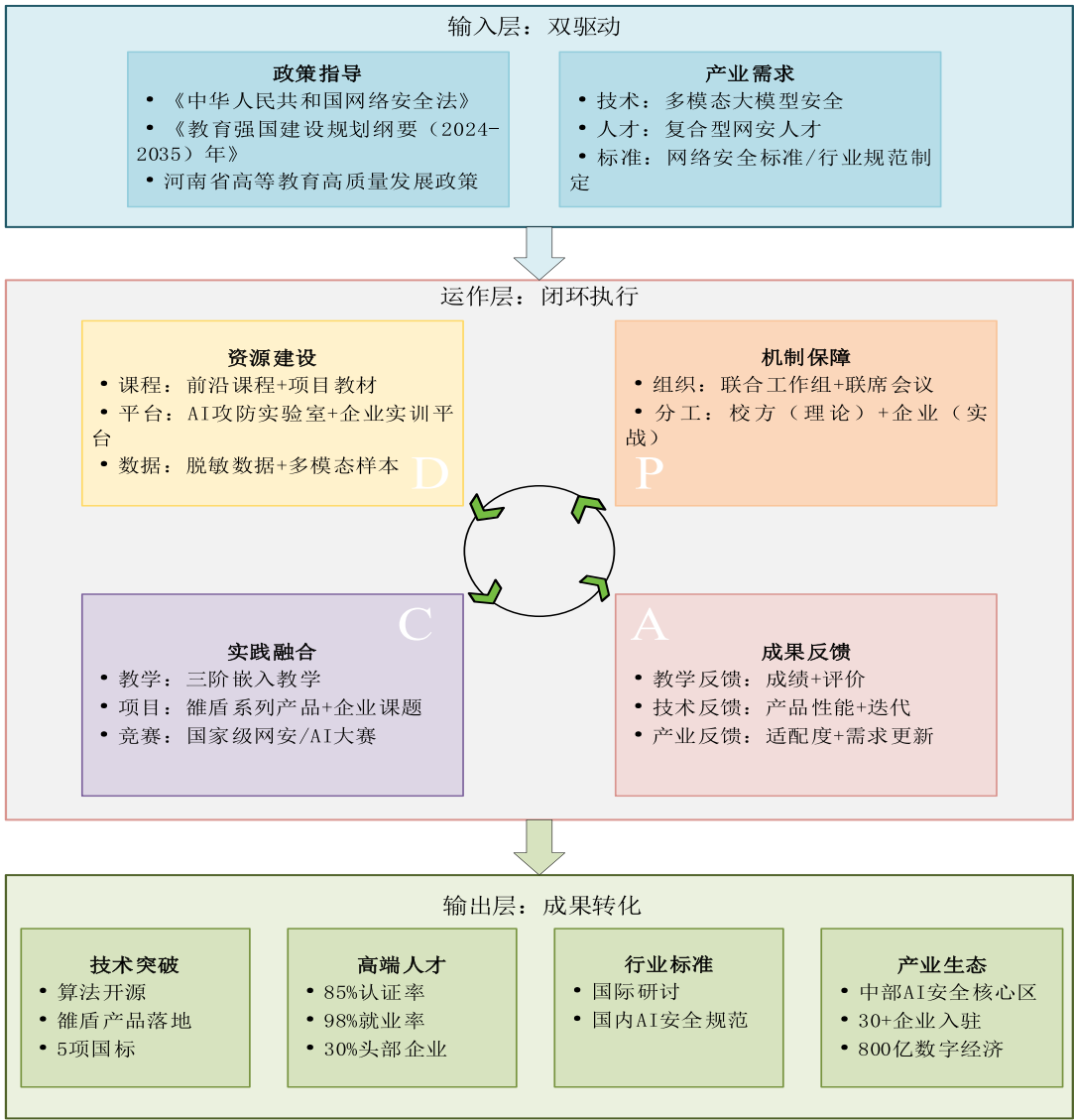


图1 产教协同运行机制闭环流程图

3.3 技术研发与教学实践深度融合推进

雏盾·多模态大模型安全体系架构如图2所示。以前沿技术问题和真实项目需求为纽带，将“雏盾·多模态大模型安全检测系统”的联合研发全面嵌入教学

过程，推进“教、研、用”一体化发展。教师团队与企业工程师共同指导学生团队开展系统设计、数据治理、模型训练、安全性评估及部署运维等全流程研发任务，把企业真实攻关课题转化为毕业设计、综合实训和竞赛项目，形成“任务引领、研中学、学中做”的循环赋能模式。学生通过深度参与技术迭代与应用

推广，不仅巩固了理论认知，更显著提升系统构建、团队协作和工程创新能力。该环节是资源建设成效的实践检验，推动教学体系从知识传授向能力生成转型，实现人才培养与产业技术研发的同频共振。

锥盾·多模态大模型安全体系架构如图2所示，图中呈现核心检测算法（含对抗攻击、模型窃取等8类威胁）、6大攻击防御模块（明确各场景攻击手段与检测指标），并展示“上下文数据经API接入-传入多模态模型-由安全检测模块执行防控”的完整工作流程。

3.4 产教协同成效初步显现

通过前述三个环节的有机联动与持续积累，本项目已形成一批阶段性成果并实现初步应用转化。联合开发的课程与教材已在人工智能、网络安全等相关专业投入使用，教学反馈良好；“锥盾”系统在多行业安全防护中的实际应用，不仅验证了技术方案的可行性，也为师生参与真实研发提供了宝贵场景。实战平台支持下的学生团队在多项国家级竞赛中表现突出，部分成果获行业关注并进入孵化阶段。人才培养与产业应用之间的通道逐步畅通，一批具备创新能力和工程素养的高水平人才进入领域内重要岗位，合作模式与运行机制也为后续持续优化和规模推广奠定了坚实基础。

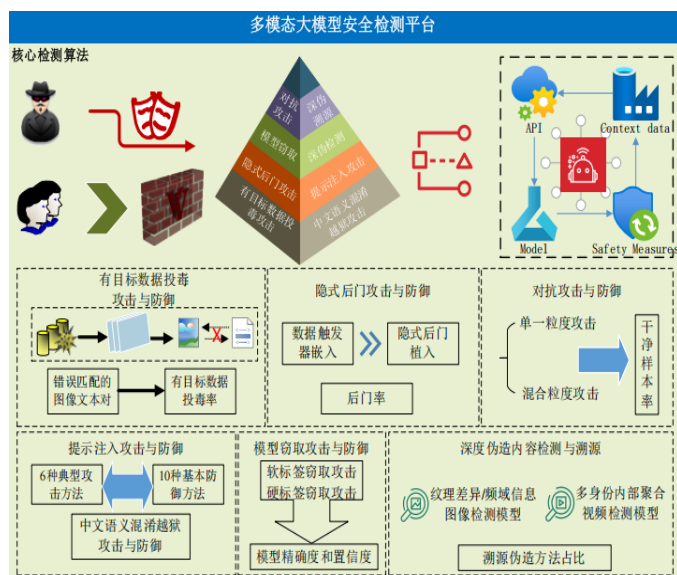


图2 锥盾®·多模态大模型安全体系架构

4 实践成效

通过技术突破、人才培养、产业协同及社会赋能，为人工智能安全领域的教育改革、技术创新与社会治理提供了全维度、可持续的示范方案。

4.1 技术突破与行业引领

河南科技大学与新华三集团基于产教融合战略合作，共同构建“教学科研—技术创新—人才培养”三位一体协同育人模式，在推进人工智能安全领域人才培养的同时，于核心技术突破与行业引领方面取得以下重要成果：

（1）核心算法国际领先：锥盾系统提出的多模态云数据安全方案及大模型对抗防御技术，经中国自动化学会鉴定为“整体技术达到国际先进水平，其中社会情境分析方法和大模型对抗防御技术达到国际领先水平”。核心算法已在GitHub开源，推动全球技术协作。该系统通过河南省电子信息产品质量检验技术研究院鉴定测试（报告编号：R20250949号），具备正式工程应用资质，且相关技术成果入选全国产教融合与科教融汇典型案例。

（2）标准制定实现突破：牵头制定《网络安全技术生成式人工智能服务安全基本要求》等5项国家标准，参与国际标准研讨，结束我国在多模态安全检测领域“无标可依”局面。

（3）多场景应用成效显著：产品成功应用于舆情治理、工业互联网、智慧城轨等10余个场景，验证了技术的稳定性与通用性。相关核心成果来自河南省重点研发专项“多模态大模型安全检测与对抗防御关键技术研发及示范化应用”，项目由河南科技大学联合新华三集团共同承担。

锥盾系统核心技术与行业引领方面的主要突破与影响如表1所示。

表1 锥盾系统核心技术突破与影响一览表

技术突破维度	具体内容	影响
核心算法	多模态云数据安全方案、大模型对抗防御技术、社会情境分析方法	国际先进，部分国际领先（由中国自动化学会鉴定）
标准制定	牵头制定《网络安全技术生成式人工智能服务安全基本要求》等5项国家标准	结束“无标可依”局面，参与国际研讨
应用场景	互联网内容治理、工业互联网边缘安全、智慧城轨调度等10余个场景	验证了技术的稳定性与通用性

4.2 人才培养质量显著提升

在深化产教融合、协同育人的过程中，河南科技大学与新华三集团以产业需求和前沿技术为导向，重构人才培养体系，全面推动学生综合素质与实战能力

提升，人才培养质量取得显著成效，教学改革与实践取得突破性成效，具体表现在以下方面：

（1）课程教学改革成效

校企共建多模态大模型安全核心课程体系，完成8门前沿课程开发，将产业真实案例、技术标准全面融入教学，教学质量大幅提升：

①实践教学课时占比从25%提升至55%，产业真实案例融入数量从12个增至46个；

②学生课程满意度从72.3%提升至94.6%，课程考核优秀率从18.5%提升至 36.8%；

③企业认证课程实现100%覆盖，教学内容与产业需求匹配度显著提高。

图3为课程教学改革核心指标对比柱状图。由图可知，经产教协同课程改革后，各项指标均实现跨越式增长，充分验证课程教学改革的有效性 with 实用性。

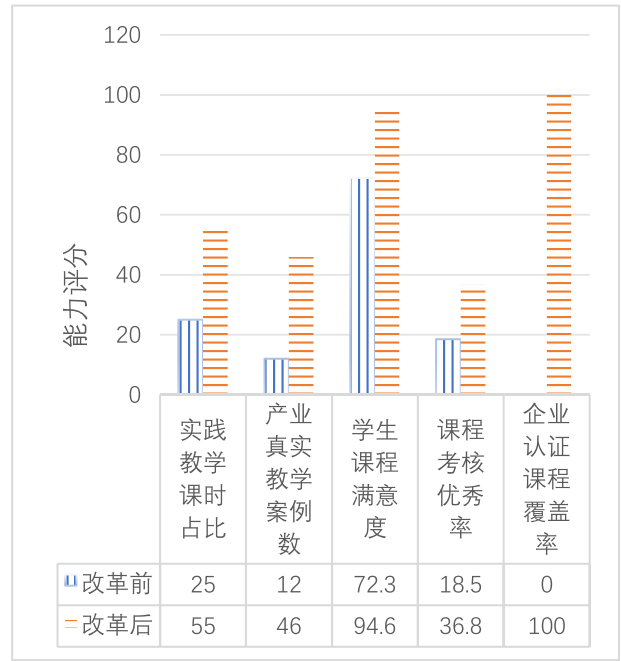


图3 课程教学改革核心指标对比柱状图

（2）实践教学改革成效

依托“锥盾·多模态大模型安全实训平台”开展全流程实战教学，实践教学体系全面升级，实施成效统计如表2所示：

①平台累计服务人工智能、网络安全专业学生1260人次，实训项目完成38个；

②学生必修实践参与率100%，实践报告优秀率达41.2%；

③人均实战训练时长146学时，218名学生深度参与锥盾系统研发全流程。

④研究生 100%参与系统研发，覆盖数据清洗、模型训练、对抗测试、安全部署全环节。

表 2 实施成效统计表

实践教学指标	统计数量	覆盖范围
实训平台服务学生人数	1260 人次	人工智能、网络安全全专业
完成产业级实训项目数	38 个	多模态安全、对抗防御全方向
学生必修实践参与率	100%	全体在读学生
实践报告优秀率	41.2%	参与实训学生
人均实战训练时长	146 学时	深度参与实训学生
深度参与系统研发学生数	218 名	本硕贯通培养学生
研究生科研参与率	100%	人工智能、网安专业研究生

经教学改革与产教实践，学生核心能力得到全方位提升，如图4所示，工程实践、创新研发、安全攻防、团队协作能力均实现大幅增长。

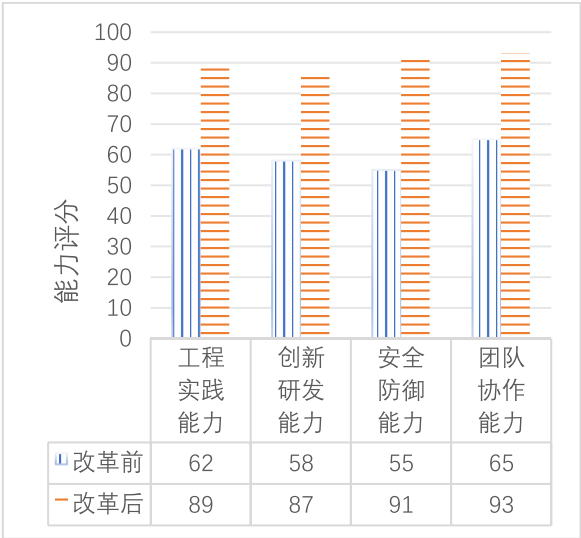


图4 教学改革前后学生核心能力提升对比柱状图

（3）学生综合能力与就业成效

①课程体系精准对接产业需求：开设“多模态安全检测技术”“大模型对抗防御”等前沿课程，引入新华三、360集团认证体系，2025届毕业生认证覆盖率达85%，就业率98%，30%进入行业头部企业。

②学生实战能力突出：学生团队在“中国研究生网络安全创新大赛”中获全国三等奖，项目“基于CTTA框架的大模型对抗攻击系统”获企业孵化支持并成功应用于金融风控，欺诈识别效率提升40%；“DeepCensor：多模态大模型安全检测系统”获“华为杯”第七届中国研究生人工智能创新大赛一等奖；多模态大模型安全检测系统V1.0.0获2025年网络安全优秀创新成果大赛优胜奖。

③科研教学融合发展：学生参与雏盾系统研发，相关成果发表于中科院一区期刊，实现“教学—科研—产业”良性循环。

产教融合模式下学生在认证、就业、竞赛及科研等方面的突出成效如表3所示。

表 3 产教融合模式下人才培养成效数据表

培养成效维度	具体指标	成果
课程与认证体系	前沿课程开设数量、企业认证引入	开设“多模态安全检测技术”等课程，引入新华三、360 认证体系
学生认证与就业	毕业生认证覆盖率、整体就业率、头部企业就业率	2025 届认证覆盖率 85%，就业率 98%，30%进入行业头部企业
学生竞赛与创新	国家级竞赛获奖等级、项目孵化与应用效果	获“中国研究生人工智能创新大赛”全国一等奖、“中国研究生网络安全创新大赛”全国三等奖；“基于CTTA 框架的对抗攻击系统”应用于金融风控，欺诈识别效率提升40%
科研产出	学生参与研发并发表的高水平论文级别	成果发表于中科院一区TOP 期刊

4.3 产业协同与区域经济赋能

基于产教深度融合的创新机制，河南科技大学与新华三集团的合作不仅提升了人才培养质量，更在推动企业技术创新与区域产业生态构建方面取得显著成效，具体表现在以下方面：

1、企业技术升级与市场拓展：新华三通过雏盾系统技术赋能，2025年营收增长25%，新增客户12家，技术竞争力跻身国内前三。

2、区域生态构建与规模效应：推动郑州成为中部人工智能安全产业核心区，吸引30余家企业入驻，形成“研发—育人—应用”闭环生态，预计带动区域数字经济规模突破800亿元。

4.4 社会效益与安全能力提升

雏盾系统服务于中央网信办“清朗·打击网络谣言”专项行动，2025年累计检测虚假信息120万条，阻断传播路径85万次，有效净化了网络生态。系统V2.0.0已完成计算机软件著作权登记，为技术规模化推广奠定了法律基础。

5 结 语

本研究通过河南科技大学与新华三集团的深度产教融合，系统构建了“教学—研发—应用”一体化的人工智能安全人才培养新范式，有效破解了传统教育中理论与实践脱节、人才培养与产业需求错位的突出问题。项目以“雏盾”系统为纽带，将真实企业项目、前沿技术研发与多层次教学实践紧密结合，不仅显著提升了学生的工程实战能力和创新素养，也推动了人工智能安全技术在多行业的创新应用与标准引领。该实践形成了可复制、可推广的协同育人机制与资源建设模式，为新时代高等教育深度融合产业发展、实现高质量人才培养与高效能治理提供了重要路径参考与典型范例，对构建具有中国特色的产教融合教育体系具有积极的示范意义。

参 考 文 献

- [1] 钟雪丽. 产教协同创新创业育人模式研究[J]. 现代职业教育, 2021(4):38-39.
- [2] 中华人民共和国教育部. 积极推动人工智能赋能教育强国建设[EB/OL]. (2025-04-02) [2025-09-08]. http://www.moe.gov.cn/jyb_xwfb/s5148/202504/t20250402_1185864.html.
- [3] 中国政府网. 国务院印发《教育强国建设规划纲要（2024—2035年）》[EB/OL]. (2025-01-19) [2025-09-08]. https://www.gov.cn/zhengce/202501/content_6999913.htm.
- [4] 邹雄, 刘宇航, 刘栓, 等. 人工智能技术对高校人才培养的影响[J]. 计算机技术与教育学报, 2024, 12(4):79-84.
- [5] 郭俊. 科教融合视域下人工智能类课程教学案例探析[J]. 计算机技术与教育学报, 2023, 11(2):46-50.
- [6] 黄璐, 任航, 曹晓丽, 等. 面向有组织产学研协同创新的合作主题挖掘[J]. 科学学研究, 2025, 43(3):548-559. DOI:10.3969/j.issn.1003-2053.2025.03.010.
- [7] 包云岗. 基于开源模式的产学研协同新机制探索与实践[J]. 中国科学院院刊, 2025, 40(3):465-470.
- [8] 董帅, 庄宇, 李悦乔. 大模型赋能的人工智能导论实践教学改革的[J]. 计算机技术与教育学报, 2024, 12(5):109-114.
- [9] 覃福钊, 李晶, 杜晓辉, 等. “新工科”背景下人工智能教学的改革探索[J]. 计算机技术与教育学报, 2025, 13(6):244-247.
- [10] 殷红建, 郝银凤, 马宇翔. 人工智能赋能网络空间安全方向研究生教学改革的首探[J]. 计算机技术与教育学报, 2025, 13(2):74-78.

- [11] 向函, 赵云, 夏峰. 交互式项目开发在程序设计类课程教学中的应用实践[J]. 计算机教育, 2025(8):35-39.
- [12] 郭园方, 余梓彤, 刘艾杉, 等. 多模态大模型安全研究进展[J]. 中国图象图形学报, 2025, 30(06):2051-2081.
- [13] 沈映春, 潘淑苓. 大模型产业产学研机制研究[J]. 中国科技论坛, 2024(10):92-103, 116.
- [14] 于秋月, 刘坚. “人工智能+教育”的伦理分析及主体重塑[J]. 学术探索, 2024(1):148-156.
- [15] 巫娜. “生成式人工智能+教育”的伪主体间性及其风险[J]. 现代大学教育, 2024, 40(4):27-37.
- [16] 周杰, 刘珍, 童卫丰. 生成式人工智能增强职业教育适应性学习的价值逻辑与实践探赜[J]. 中国职业技术教育, 2025(15):51-59.
- [17] 项导, 鲍蓉, 胡局新. 学科竞赛和产教融合驱动的人工智能应用型人才培养探索 [J]. 计算机技术与教育学报, 2024, 12 (3):124-128.