

南昌大学人工智能安全与可信数据实验室

在总体国家安全观指导下，为积极应对人工智能技术快速迭代带来的算法安全风险、数据滥用隐患及生成式 AI 治理等难题，南昌大学充分发挥多学科优势，依托公共政策与管理学院国家安全学、公共管理、数据资源与数据智能等学科，联合法学、人工智能与计算机等相关学科，开展交叉学科研究，成立人工智能安全与可信数据实验室。

实验室服务省委网信办、省工信厅、省公安厅、省国安厅等主管部门，围绕总体国家安全观重点领域中的非传统安全问题（人工智能），聚焦人工智能带来的深度伪造、隐私保护、算法透明与公平等核心问题开展技术研究；针对算法责任认定、数据权属划分与伦理评估体系建设等法律伦理问题进行合作探索；为政府人工智能安全与数据治理提供政策咨询，服务国家总体安全战略，助力构建安全、可控、可信的人工智能创新生态。