2026 第二届大学生人工智能安全竞赛
——开放式自主命题作品赛参赛指南

竞赛的目的是为培养、选拔、推荐优秀人工智能安全人才创造条件，促进高等学校网络空间安全和人工智能专业课程体系、教学内容和方法的改革，培养学生的创新意识与团队合作精神，普及人工智能安全知识，增强学生人工智能安全意识。

本指南为学生、指导教师和高校如何参与本次人工智能安全竞赛开放式自主命题作品赛提供具体指导。

一、学生参赛及报名

1.报名截止日期内具有正式学籍的全日制在校本科生、专科生均可报名参赛（硕士生和博士生不能参赛）。评审时，如发现参赛队员不符合参赛规定，将取消参赛队伍的参赛或获奖资格。

2.每支参赛队不超过 4 名学生（包括 1 名组长），每支参赛队限指定 1 名指导教师，每名学生限参加 1 支参赛队，各高校参赛队数不限，不可跨校组队。

3.报名步骤：

(1)各高校收到通知后，确定一位老师作为本校唯一联络老师，并填写“高校联络教师登记表”。盖章，扫描后于 2026 年 6 月 10 日前将电子版和盖章纸质版扫描件发至组委会秘书处邮箱 xieyi233@sjtu.edu.cn。

(2)学生收到通知后，联系本校负责老师，进行集体报名。本次竞赛不接收个人报名。

(3)报名时间为2026年6月1日—7月1日。

4.在线提交作品：2026年7月2日—8月2日。高校联络教师将本校所有作品上传至https://ai-contest.sjtu.edu.cn/。

5.参赛名单公布日期：2026年8月3日。

6.联络信息：

作品赛联络群：281820424。

高校教师联络群：474997879。

二、参赛作品

1.参赛作品要体现一定的创新性和实用性。

2.参赛作品可以是软件、硬件等。参赛作品的内容以人工智能安全技术与应用设计为主要内容。参赛队自主命题，自主设计。

3.参赛作品内容包含四个方面之一：一是AI内生安全（主要是AI自身的安全，如隐私泄露、模型窃取、版权保护）；二是AI应用安全（主要是由于AI的应用引发的安全，如数据投毒、深度伪造、对抗样本生成等）；三是AI赋能安全（主要是将AI技术应用于传统的安全技术中来提升其性能或效率，例如将AI技术应用于恶意代码检测、漏洞挖掘、渗透测试、密码设计与分析系统等）；四是AI衍生安全（主要是AI Agent权限管理、API投毒、skill工具篡改等）。

4.参赛队自主命题，自主设计。本次竞赛采用开放式，不限定竞赛场所，参赛队利用课余时间，在规定时间内完成作品的设计、调试及设计文档撰写。所有参赛题目须得到组委会认可后方

能参赛。如果参赛队伍所报题目及内容违反赛事精神和章程，组委会有权要求参赛队伍进行修改。本赛事只接受防御性的题目，不接受任何具有攻击性质或与国家有关法律、法规相违背的题目。

5.参赛队的参赛作品应该是参赛队员独立设计、开发完成的原创性作品，严禁抄袭、剽窃、一稿多投等行为。凡发现此类行为，将取消参赛队伍的参赛资格，并追究相关指导教师和高校的责任。

6.凡已公开发布并已获得商业价值的产品不得参赛；凡有知识产权纠纷的作品不得参赛；与企业合作即将对外发布的产品不得参赛；凡在其他公开竞赛（不含校内比赛）中获奖的作品不得参赛。

7.本次竞赛不支持论文参赛。

三、初赛

1.2026 第二届大学生人工智能安全竞赛分初赛和决赛。凡取得参赛资格的参赛队均自动进入初赛。

2.初赛作品提交截止时间为 2026 年 8 月 2 日。各参赛队应在 2026 年 7 月 2 日—8 月 2 日期间完成参赛作品并网上提交，以参加初赛。

3.各参赛队应在截止时间前提交参赛作品相关材料。提交材料包括作品报告、作品原创性声明、源程序、可执行程序及其他作品相关材料。

4.作品报告和作品原创性声明应使用竞赛官网发布的统一模板填写。其中，作品报告应完整说明作品的系统设计方案、功

能实现、测试方案、测试结果、创新特色及其他必要内容；作品原创性声明须由所有参赛队员手写签名，并按要求加盖学校教务处或教务部公章。

5.提交方式：各参赛队将参赛作品及相关材料交由高校联络教师，由高校联络教师统一通过竞赛官网提交。

6.本次竞赛的组委会将在全国范围内组织专家对参赛队伍提交的作品进行网络评审，初赛评审时间为2026年8月5日—8月15日。依据网络评审结果，由专家组评审并最终确定进入决赛名单。进入决赛的参赛队伍由专家组根据参赛队伍总数及参赛作品质量确定。

7.评审方式：评审专家审阅作品报告，依据评审规则对参赛作品进行打分，并给出评审意见。每一件作品将至少由2至3位专家进行评审。

8.专家评审的主要内容包括：作品的原创性与创新性、作品完成程度、作品的性能、作品的应用价值、相关文档的规范性等。

四、决赛

1.组委会将在2026年8月16日公示进入作品赛决赛的名单。

2.在获得决赛资格后，各参赛队伍可以继续对参赛作品进行完善和修改。

3.线下决赛会评安排为2026年8月20日报到，8月21—22日决赛。获得决赛资格的参赛队伍应在规定时间内参加决赛。决赛分为作品演示和答辩两个环节。

4.作品演示

参赛队自行携带作品、文档及设备，到决赛地点进行演示。决赛时，承办方提供因特网接入环境。各参赛队伍须自带测试设备，如对作品的演示环境有特殊要求，请提前与组委会秘书处协商(注意：本次竞赛演示时，原则上不能以视频方式演示；只能现场操作演示。如有特殊演示要求，需要提前申请并获得组委会秘书处同意)。

5.答辩

答辩时间为 15 分钟（10 分钟讲解，5 分钟提问），包括 PPT 陈述、演示、测试与专家提问，专家会现场检查源代码。

6.评审专家对每个竞赛作品实行分项打分，集体讨论，综合评定，最终确定参赛作品的获奖等级。

7.决赛时，由竞赛承办方统一提供附近宾馆、饭店等相关信息，食宿及相关费用由参赛学校自理。

五、获奖

1.本届竞赛设一等奖、二等奖和三等奖。获奖比例由竞赛专家委员会开会决定。

2.竞赛颁发统一的获奖证书，所有获奖队伍及名单将以多种方式公布，并报送相关高校，作为高校评定奖学金、推荐研究生等的参考。

3.获奖队伍将获邀参加 2026 年 8 月 23 日进行的“2026 第二届大学生人工智能安全竞赛颁奖”。

六、指导教师

1.指导教师必须是参赛队伍所在高校在职教师。

2.指导教师可以指导学生选题和设计方案的论证，但具体的硬件制作、软件编程、系统调试、作品文档撰写必须由参赛学生独立完成。

3.指导教师负责把握所指导学生参赛作品的原创性，并确保其不具攻击性，以及不与国家法律、法规相违背。

七、参赛高校

1.各高校在收到参赛通知后，指定1位教师作为联络人（联络人须为高校领队），负责本校竞赛相关事宜，并在竞赛网站上下载“高校联络教师登记表”，将该教师信息填写完后，发至组委会秘书处邮箱：xieyi233@sjtu.edu.cn（含电子版和盖鲜章纸质版的扫描件）。

2.各高校负责本校范围内的竞赛组织、选拔等工作，并对本校范围内参赛队伍及指导教师的真实性负责。

3.本次竞赛将对在竞赛组织工作中表现出色和作出贡献的高校给予奖励。

4.各高校应从培养和选拔创新人才的角度出发，对获奖学生在奖学金评定等方面予以优先考虑。

5.禁止参赛高校弄虚作假。对违反国家有关法律、法规以及大赛章程的行为，组委会将取消相关奖项，并依照有关规定进行处罚。