天融信大模型安全评估服务

www.topsec.com.cn

服务概述

大模型安全评估服务划分为两部分：通用安全评估、模型安全评估。通用安全评估主要涉及：服务器、网络设备、安全设备、数据库、docker容器等漏洞扫描和配置核查，以及应用层传统的渗透测试和代码审计。大模型安全评估主要涉及：根据大模型业务场景的安全测试，如语言大模型的提示词注入、API接口权限测试、非授权代码执行、会话权限测试等；两大环节协同形成闭环，全面保障大模型的安全性、可靠性与合规性。



服务优势

全面管控系统与业务风险

针对客户大模型部署环境展开漏洞扫描、渗透测试、代码审计等工作，充分掌握系统风险现状。同时，根据大模型特定业务场景进行专项测试，兼顾应用系统风险与大模型原生安全隐患，完整识别全链路脆弱点，保障大模型的安全性、可靠性与合规性。

测评流程严格合规

测评严格依据《生成式人工智能服务安全基本要求》等国标、团体标准，覆盖内容合规、数据、模型全维度测评项，配套标准化访谈、工具测试、渗透验证方法，输出可落地风险整改方案，既完成漏洞隐患排查，又满足大模型备案、安全报备等监管合规硬性要求。

专业资质与实战攻防技术支撑

拥有CNCERT最高级应急支撑单位证书，构建了业内完备的资质矩阵；依托专业技术实验室，在历年国家级攻防演习、省级攻防演练斩获优秀成绩，持续跟进新型 AI攻击手段，测试能力贴合真实攻防场景。