EY安永
Shape the future with confidence
聚信心 塑未来

Ernst & Young (China) Advisory Limited
Beijing Branch Office
Level 5, Ernst & Young Tower
Oriental Plaza, 1 East Chang An Avenue
Dongcheng District
Beijing, China 100738

安永（中国）企业咨询有限公司
北京分公司
中国北京市东城区东长安街 1 号
东方广场安永大楼 5 层
邮政编码: 100738

Tel 电话: +86 10 5815 3000
Fax 传真: +86 10 8518 8298
ey.com

人工智能安全合规服务指南

1. 人工智能安全合规服务简介

在 AI 技术重塑全球商业格局的今天，企业面临模型安全、数据隐私、算法偏见、合规准入等系统性风险。安永推出覆盖 AI 全生命周期的安全合规服务，为企业提供从 AI 模型训练到商业落地的全周期安全护航。

本服务由安永咨询提供，安永咨询是国际知名专业服务机构，在科技咨询、安全咨询及人工智能安全管控领域展现出全方位专业优势，具备全球化资源及行业领导力优势。

2. 服务内容

2.1. 人工智能风险评估：精准洞察企业 AI 风险现状

AI 应用场景调研

全面梳理企业 AI 系统部署现状，深入分析业务场景与技术架构，明确 AI 应用的覆盖范围与关键节点。

数据安全合规评估

选取代表性 AI 产品，检查数据采集、存储、使用及共享的合规性，确保数据全生命周期的安全性与合法性。

AI 监管法规对标分析

对标国内外先进 AI 法规（如《生成式 AI 服务管理办法》《数据安全法》等），识别企业 AI 应用与法规要求的差距，提供合规改进建议。

动态风险库构建

通过穿透式风险识别，建立动态风险库，量化与定级潜在风险，明确处置优先级，确保风险应对的及时性与有效性。

2.2. AI 安全技术测试：攻防实战验证与漏洞修复

多模式安全测试

采用红蓝对抗、智能扫描与人工渗透相结合的方式，结合 OWASP LLM Top 10 框架，识别数据投毒、提示注入、模型窃取等多类安全风险。

定制化测试方案

设计贴合企业需求的测试标准与方案，评估 AI 技术架构的合理性与先进性，提供标准化测试报告。

修复实施计划

针对发现的安全漏洞，制定具体修复计划，完善 AI 系统安全防护体系，确保系统稳定性与可靠性。

2.3. AI 合规管理体系咨询：构建纵深防御治理架构

AI 治理框架设计

制定符合企业需求的 AI 安全管理政策、流程及组织架构，确保 AI 应用的合规性与可持续性。

AI 管理体系建设

结合 ISO 42001 等行业优秀实践，通过领导力支持、风险评估、体系设计、绩效评估和持续改进等环节，帮助企业构建全面、系统的 AI 安全合规管理体系，实现从风险识别到持续运营的全生命周期管理，为企业 AI 应用的可持续发展提供强有力的技术支持与合规保障

AI 伦理审查机制建设

帮助企业设立 AI 伦理委员会，制定算法伦理审查流程，评估算法偏见、歧视性决策等伦理问题，提升企业社会责任感。

跨部门协作机制

明确各部门在 AI 应用中的职责，促进跨部门协作，形成制度、流程与人员的闭环管理。

2.4. AI 可信持续运营：赋能企业品牌与技术创新

实时监控与管理

引入先进的 AI 管理工具与平台，实现对 AI 系统的实时监控与动态管理，确保系统运行的稳定性和安全性。

工程化能力嵌入

将 AI 安全工作嵌入企业研发与业务流程，形成稳定的 AI 安全合规工程化能力，提升企业整体竞争力。

定期审计与优化

通过定期审计确保系统合规运行，并结合业务需求进行针对性优化，持续提升 AI 系统的可信性与可靠性。

2.5. 算法备案协助：确保监管合规

备案材料准备

指导企业整理算法说明、安全自评估报告等备案材料，确保材料的完整性和合规性。

备案流程支持

协助企业完成算法备案流程，确保 AI 系统符合监管要求，降低法律风险。

3. 交付件

3.1 核心交付件

- **AI 风险评估**：《AI 应用现状分析报告》、《AI 安全风险评估报告》
- **AI 安全测试**：《AI 应用渗透测试报告》、《AI 安全修复建议》
- **AI 体系建设**：《AI 安全管理架构设计》、《AI 安全管理手册》、《AI 管理体系策略和配套工具》、《AI 合规培训》
- **AI 可信运营**：《AI 风险报表》、《事件响应与预案》、安全托管支撑
- **备案协助**：《算法备案材料》

3.2 按需交付件

基于企业现有的人工智能相关业务和风险情况，按需定制开发相关的安全合规管理工具和实施方案。