贝泰妮集团人工智能安全管理政策

前言

云南贝泰妮生物科技集团股份有限公司(以下简称“本公司”）及其各级子、分公司或项目组(以下合称“本集团”。泛指本集团某一单位或所有单位时，通称“集团公司”或“各集团公司”，但不包括本集团参股的各类型合营联营企业)高度重视人工智能（以下简称“AI”）技术应用中的信息安全与隐私保护。为规范 AI 系统的开发、部署、运营与管理，保障数据安全、模型可靠性与系统稳定性，防范 AI 应用过程中可能出现的各类风险，特制定本政策。

政策适用范围

本政策适用于公司全体员工（包括公司及其各级子、分公司、项目组的正式员工、兼职员工、临时工、高级管理人员及董事），涵盖集团内所有涉及人工智能系统需求分析、设计开发、测试部署、运行维护、迭代优化及下线销毁全生命周期的相关工作与人员。

公司鼓励为集团提供 AI 技术服务、数据支持、系统开发等合作的供应商、合作伙伴等第三方共同遵守本政策明确的相关要求，共同维护人工智能应用的安全生态，集团将对第三方相关安全行为进行监督与管理。

人工智能安全管理原则

- 清晰透明，公平可靠：公司在 AI 应用中采用透明标识机制，确保用户能够识别 AI 生成内容。具体措施包括在 AI 生成的内容中添加明确的标识，或通过用户界面提示 AI 输出状态。集团将定期检查标识机制的有效性，确保用户体验清晰明确。
- 风险管理，预防为主：对 AI 系统全生命周期进行全面风险评估，识别潜在安全威胁与风险点，按风险严重程度和发生概率分级分类，针对高风险环节制定针对性防范措施，实现风险的提前预判、动态监测和及时处置，实现网络安全保障。
- 合规为本，底线思维：确保 AI 系统的设计、开发、部署和运营符合国家及地方人工智能、信息安全、数据隐私保护等相关法律法规及行业标准（如 ISO27001 等），坚守数据安全、隐私保护、算法公平等合规底线。
- 数据安全，隐私至上：严格遵循集团信息安全与隐私保护的合法、正当、必要原则，规范 AI 应用中的数据采集、存储、处理、传输、使用等环节，强化敏感数据保护，防止数据泄露、篡改、滥用，保障个人信息隐私权益。
- 权责明确，全程追溯：遵循 “权限最小、按需授权、用户唯一、职责分离” 原则，明确 AI 系统全生命周期各环节的管理责任人

和执行责任人，实现 AI 系统操作、数据访问、算法调整等行为的全程可追溯。

- 透明可解释，公平可靠：保障 AI 决策过程的透明度和可解释性，避免算法“黑箱操作”，对 AI 模型进行偏见检测和鲁棒性测试，确保模型输出结果公平、公正、可靠，杜绝算法歧视。
- 持续改进，动态优化：紧跟人工智能技术发展和安全威胁变化，定期对本政策进行回顾与修订，持续优化 AI 安全管理措施，提升集团人工智能安全防护能力。

人工智能安全管理流程

（一）风险评估与规划

- 需求分析：明确 AI 系统的应用场景与功能需求，初步识别风险。
- 风险识别与分级：采用定性与定量相结合的方法，评估安全威胁，并根据严重程度与发生概率对风险进行分级。
- 制定应对方案：针对不同级别风险，制定相应的防范与缓解措施，如数据加密、访问控制、模型验证等。

（二）开发与测试阶段管理

- 安全设计：在系统设计阶段嵌入安全模块，遵循最小权限原则。
- 安全测试：定期进行应用安全测试，防范 SQL 注入、跨站脚本等常见漏洞。

- 多轮测试：通过单元测试、集成测试等方式，验证系统抗风险能力。
- 模型验证：对 AI 模型进行偏见检测与鲁棒性测试，确保输出结果公平、可靠。

（三）部署与运行阶段管理

- 环境隔离：将 AI 系统部署于隔离的云环境或专用服务器，防范外部攻击。
- 实时监控：部署日志审计与异常检测系统，实现 24 小时运行状态监控。
- 应急响应：制定安全事件应急预案，明确响应流程与责任人。
- 定期审计：定期开展安全审计，检查系统漏洞与违规操作。

（四）数据安全管理

- 数据分级：根据敏感程度对数据进行分级，实施差异化保护。
- 加密存储：对敏感数据进行加密存储，防止泄露。
- 访问控制：实施严格的权限管理，仅授权人员可访问核心数据。
- 脱敏处理：在模型训练或数据共享时，对个人身份信息进行脱敏处理。

（五）持续改进

- 反馈机制：建立用户反馈渠道，收集使用过程中的安全问题。
- 技术更新：定期更新安全工具与补丁，应对新型威胁。

- 复盘总结：安全事件发生后组织复盘，总结经验并优化政策。

第三方安全管理

- 对为集团提供 AI 技术开发、数据服务、系统部署等合作的第三方进行严格的安全资质审查，评估其信息安全和 AI 安全防护能力，将安全要求纳入合作协议。
- 规范第三方对集团 AI 系统、数据资源的访问权限，为第三方分配专用访问账号，设置访问范围和操作时限，监控第三方的访问和操作行为。
- 明确第三方在数据使用、模型开发、信息保密等方面的安全责任，若因第三方行为造成集团信息泄露、系统被攻击等安全事件，集团将依法追究其相关责任。

监督与考核

（一）安全监督

- 监督职责：由集团信息安全管理小组负责制度制定与执行监督。
- 定期会议：定期召开会议，评估安全状况并调整管理策略。

（二）绩效考核

- 目标设定：设定年度安全目标（如零数据泄露、漏洞修复率≥99%等）。

- 评估标准：依据事件发生率、响应时间、整改效果等指标进行考核。
- 奖惩措施：对安全表现优异的团队给予奖励，对违反制度的个人进行处罚。

回顾与修订

本政策由集团信息安全委员会审批通过，信息安全委员会的最高决策人为集团总裁。信息安全委员会将定期监督政策的实施情况，并确保其有效性和全面性。公司将定期对本政策的执行情况进行回顾和评估，结合国家人工智能相关法律法规的更新、监管要求的变化、业务发展的需要、AI 技术的革新及安全威胁的新趋势，及时对本政策进行修订与完善。