

国轩高科负责任人工智能政策

一、概述

国轩高科股份有限公司（以下简称“公司”）致力于推动人工智能技术在产品研发、生产制造、运营管理及客户服务等领域的安全、可靠和负责任应用，通过建立人工智能治理框架，规范人工智能技术开发、部署和使用过程中的管理要求，促进技术创新与风险管控协同发展，充分发挥人工智能对经营效率提升、产品创新及可持续发展的赋能作用。本政策确立了指导性原则，旨在确保我们的人工智能工作在伦理、合法、安全、隐私及问责等方面均遵循最高标准。

二、适用范围

本政策适用于公司及下属子公司人工智能相关活动，包括人工智能技术的开发、测试、部署、应用、维护及退出管理，以及与人工智能相关的数据处理、模型管理、第三方合作和业务应用场景管理。同时，公司倡导并鼓励合作伙伴（包括但不限于供应商、承包商等）参照执行本政策。

三、管理架构

公司将人工智能治理纳入整体信息安全与风险管理体系，并建立有跨部门协同管理机制。由数字化中心、数智产品院、数字工程分院等技术团队及业务部门共同参与人工智能系统的开发、部署、运行及监督管理。对于人工智能应用过程中产生的异常结果、数据问题、安全事件或业务风险，公司按照事件等级开展调查分析、责任认定及整改优化，确保人工智能应用全过程可追溯、可问责。公司

定期向高级管理层汇报人工智能应用进展、重大风险及管理成效，并根据业务发展情况持续完善人工智能技术相关管理机制。

四、管理原则

公司参考国内外相关法律法规，以及经济合作与发展组织（OECD）、联合国教科文组织（UNESCO）、电气与电子工程师协会（IEEE）等组织对于“负责任 AI”的定义及要求，结合公司管理实践，明确负责任 AI 管理原则。

1. 保障数据安全与隐私保护

公司在人工智能开发和应用过程中严格遵守数据安全、网络安全及个人信息保护相关要求。公司对训练数据、业务数据及用户信息实施分类分级管理，通过权限控制、脱敏处理、数据隔离、加密保护及访问审计等措施，保障数据在采集、存储、传输、使用和销毁过程中的安全性与合规性。

2. 坚持公平、透明、可解释

公司关注人工智能模型可能产生的偏差风险，通过数据预处理、模型训练优化及专家复核等措施持续提升模型公平性和准确性。对于人工智能生成内容及人工智能驱动的分析结果，公司要求进行明确标识，并向相关用户提供必要的模型逻辑说明、应用场景说明及结果解释，提升人工智能应用透明度。

3. 坚持人机协同与审慎应用

对涉及产品安全、客户服务、设备运行及其他重大业务决策的人工智能应用，公司坚持“人工参与决策”原则，保留人工审核、干预和否决权限。公司同时明确人工智能系统的适用范围、能力边界和禁止事项，防止超出授权场景使用人工

智能技术。公司禁止开发、部署或使用具有操纵用户行为、利用脆弱群体、社会评分、未经授权生物识别监控、违法内容生成或其他违反法律法规及商业伦理要求的人工智能系统。

4. 考虑环境和社会影响

我们将考虑人工智能系统对环境和社会的影响。我们将从可持续性、资源保护和生态稳定性等方面对人工智能系统进行评估。在整个 AI 系统生命周期中，我们还将考虑政府与社会诚信、人权和法治等价值观。公司将根据具体情况，采用与技术水平相适应的机制和保障措施。通过采取负责任的人工智能发展方式，我们旨在减轻环境和社会风险，并走在可持续创新的前沿。

五、管理措施

1. 人工智能系统开发与运行管理

公司建立人工智能模型持续监测机制，定期评估模型准确性、稳定性及适用性。当发现模型性能下降、业务环境变化或数据特征发生偏移时，及时开展模型优化、参数调整或重新训练，确保模型持续满足业务需求。

2. 人工智能安全防护

公司根据业务需求对人工智能模型、训练环境及算力资源实施权限管理和分级授权，限制敏感功能、关键数据及核心模型的访问范围，并通过网络隔离、入侵检测、堡垒机管理等措施提升人工智能系统安全防护能力。

3. 推动绿色和高效人工智能

公司关注人工智能技术应用过程中的资源消耗与环境影响，在满足业务需求和技术效果的前提下，持续推动人工智能模型、算力资源及基础设施的高效利用，

提升人工智能应用的资源效率和可持续发展价值。

公司优先采用计算效率更高、资源消耗更低的模型架构和技术方案，结合实际应用场景持续优化模型规模、计算复杂度及算力配置。在人工智能应用建设过程中，公司积极探索以小模型替代部分高资源消耗的大模型应用场景，通过模型优化、轻量化部署及资源调度等方式，在保障业务效果的同时降低算力资源消耗，提高人工智能系统运行效率。

公司在人工智能基础设施建设和采购过程中，综合考虑设备能效水平、资源利用效率及运行稳定性，优先选用高效能算力设备和技术平台，并持续探索提升算力利用效率和降低能源消耗的可行路径。

公司持续评估人工智能技术在产品研发、生产制造、设备管理及产品运营等场景中的应用成效，通过分析生产效率提升、设备故障率下降、产品质量改善、能源利用效率优化及资源消耗降低等指标，评估人工智能应用对经营绩效和可持续发展目标的贡献，并推动相关成果持续优化与推广。

六、能力建设

公司定期开展人工智能安全、数据保护及负责任人工智能应用培训，持续提升员工对人工智能风险、伦理要求及安全规范的理解与应用能力。

七、问题反馈与持续改进

公司建立面向客户、合作伙伴、员工及其他相关方的沟通反馈机制。对于人工智能系统产生的分析结果、预警信息或业务决策建议，相关方可通过客户服务渠道、项目沟通机制、投诉举报渠道（邮箱：info_sec_opt@gotion.com.cn）等

提出异议或申诉。公司将组织相关业务及技术团队开展核查、复审和反馈，并根据实际情况采取纠正或优化措施。

公司将结合业务发展情况，逐步评估人工智能治理体系与国际标准的对标需求，持续完善人工智能管理机制，并探索引入第三方评估或认证机制。

八、政策实施与修订

本政策由公司战略与 ESG 委员会批准通过，自发布之日起执行。公司承诺根据最新法律法规及时进行评审修订，确保其适宜性、充分性和有效性。本政策以中文和英文版本制备，若中英文版本不一致，以中文版本为准。

本政策将通过内部信息网络全面传达至公司每位员工，并通过培训强化全员意识与执行，确保政策内容有效落地。