中国平安 PING AN
金融 · 科技

专业 让生活更简单

平安集团
AI伦理治理政策声明

2022年1月

本政策适用于中国平安保险（集团）股份有限公司
旗下各成员公司及集团职能部门

随着人脸识别、语音识别、自动驾驶等技术日益成为大众关注的焦点，人工智能 (AI) 与社会、人类、企业正在快速融合。AI 技术的突破与发展，未来可以用于解决一些社会的长期问题，例如低能耗的产品生产、低成本的医疗服务、高质量的教育资源等，实现社会公平的可持续发展。而人工智能技术作为社会与企业的强大加速器，也带来了前所未有的治理与伦理挑战。人工智能行业对技术应用的约束力度及伦理相关概念与准则的模糊，导致监管难以落实，也为企业伦理的实施带来了更大的挑战。

平安致力于成为国际领先的个人金融生活服务集团，坚持“科技引领金融，金融服务生活”的理念，并通过提供符合伦理道德审查的科技与金融服务在金融、医疗、汽车、智慧城市四大领域，目前已布局了 ABC 三大技术，即人工智能、区块链、云计算。

在人工智能的管理上，平安遵循“以人为本、人类自治、安全可控、公平公正、公开透明”的五大伦理原则，现阶段平安主要从健全伦理体系与加强问题监控两方面入手：对外积极谏言，推动国际、国家行业标准的制定，并与同业企业及高校建立联系加强沟通；对内相关伦理委员会、监管会也在筹备建立中，同时针对 AI 伦理六大问题，在实际的项目应用中探索优化管理方式。

我们承诺对人工智能的开发和应用进行科学管控，在实现企业商业价值的同时，履行社会责任，实现企业的可持续发展。

行业六大 AI 伦理问题

通过调研与实际的应用，我们发现 AI 相关的伦理问题可以归纳为应用边界、人身安全、数据隐私、公平公正、责任认定与社会福祉六方面：

- **应用边界问题**
  指 AI 技术在应用时的边界仍不明确，在一些领域中的应用会引起争议；
- **人身安全问题**
  指人工智能在应用过程当中存在一些危害人类及社会安全的隐患，如使用无人机进行走私将会危害社会治安；
- **数据隐私问题**
  指在数据及技术模型发展过程中，对数据的收集、传播与使用通常未经个人授权许可，可能涉及侵犯个人隐私；
- **公平公正问题**
  通常是针对机器学习可能存在的种族、宗教、性别等方面偏见与歧视及 AI 应用产品分配不均等问题;

• 责任认定问题

人工智能应用中产品造成的问题，责任主体如何认定问题也应该引起重点关注，如无人驾驶汽车事故的责任主体是汽车厂商还是汽车的主人；

• 社会福祉问题

目前如何确保人工智能的应用把以人为本作为核心宗旨，并为社会提供福利仍没有明确的规定及准则。

行业五大伦理原则与承诺

平安针对以上 AI 伦理问题，结合目前各国、各行业已发布的相关规定，从自身的业务出发，归纳了平安应对 AI 伦理需遵循的五大原则，并从数据、算法、产品设计等层面全面把控 AI 伦理问题。

人类自治

应用过程中确保人类知情权，告知可能产生的风险和影响，人类能够干预或监督软件做出的每一个决定；

建立人工智能问责机制，明确研发者、使用者和受用者等的责任，防范利用人工智能进行非法活动。

安全可控

在个人信息的收集、存储、使用等各环节设置边界，建立规范，尊重和保护个人隐私；

确保人工智能系统运作时不会损害用户（人类、资源、环境），提高人工智能鲁棒性及抗干扰性，并形成人工智能安全评估和管控能力。

以人为本

遵循人类的价值观和伦理道德，促进人机和谐，服务人类文明进步

以保障社会安全、尊重人类权益为前提，避免误用，禁止滥用、恶用人工智能技术

公开公正

促进协调发展，推动各行业领域

转型升级，使用人工智能赋能多行业生态圈；

促进共享发展，避免数据与平台垄断，努力消除数字鸿沟，进行开放有序竞争；

通过持续提高技术水平、改善管理方式，在数据获取、算法设计、技术开发、产品研发和应用过程中消除偏见和歧视。

公开透明

对使用人类数据或影响人类或可能产生其他道德上重大影响的模型，应确实公开：数据来源、开发过程和利益相关者的选择与决策；

人工智能算法应具有可解释性，系统使用的判决数据，包括训练数据在必要时刻是可以被访问的；

加强与政府、企业、高校间的交流合作。

* 参考欧盟《人工智能道德准则》、日本 AI 伦理纲要草案、Partnership on AI 宗旨、百度 AI 伦理四大原则、未来生命研究院（FLI）阿西洛马人工智能原则、电气电子工程师学会（IEEE）《人工智能设计的伦理准则》。

平安的 AI 应用

平安大力发展科技力量，将 AI 技术聚焦五大领域，赋能于集团其他生态领域，平安均致力于提供符合伦理道德审查的人工智能应用。

AI+ 金融

金融数据隐私

严控用户信息及业务经营数据的保密及使用规范。

保证公平

避免数据缺陷导致的歧视、不公平；

保证使用 AI、大数据技术开展产品设计与营销过程中的公平。

AI+ 医疗

医疗数据隐私

严控医疗数据的保密性及使用规范。

明确权责

确保医疗科技产品的权责划分明确。

维持公平

在 AI 辅助决策中，充分考虑及规避不公平问题。

AI+ 政务

政务数据隐私

严控政务信息的保密与使用规范。

保障决策公平

充分考虑 AI 辅助政务决策的公平性。

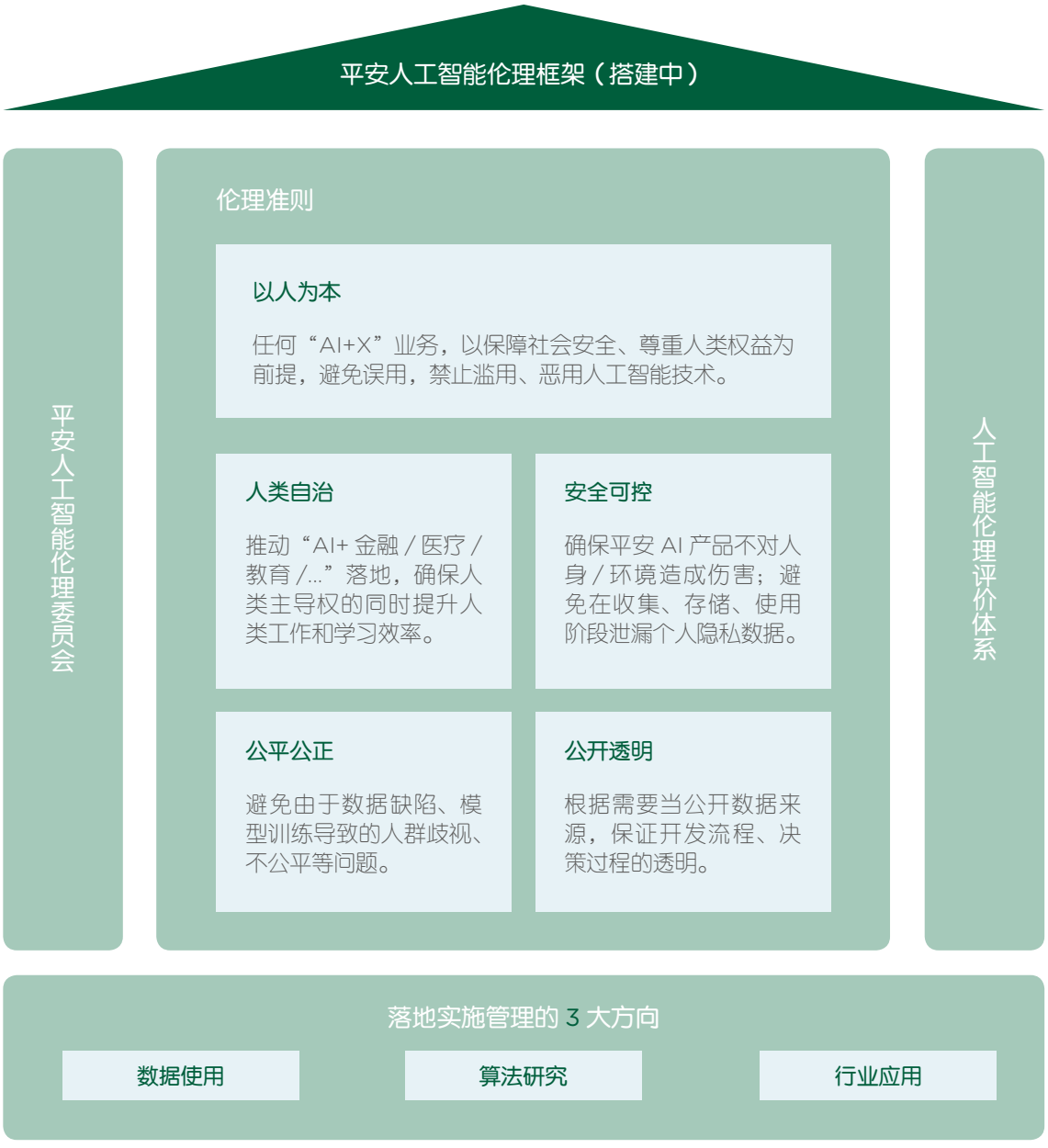
AI+X

X 代表交通、房产法律等其他领域

结合各领域特点严控 AI 伦理管理。

平安的 AI 治理组织管理

平安集团所有科技类管理均在董事会执行委员会下属科技发展委员会的职能范围内，面对未来潜在的伦理隐患，集团正在搭建全面的伦理目标及治理框架。信息安全与数据隐私风险是平安集团全面风险管理之中“五大风险”之一，我们对于该风险给予严密关注和管控，制定了详细风险管理措施。更多信息详见《信息安全管理政策》以及《隐私保护声明》。



平安集团在科技发展委员会架构内已成立由人工智能、大数据、科技管理等领域专家组成的人工智能伦理委员会，专门负责组织进行人工智能及大数据相关业务伦理领域政策及技术研究、治理宣导、管理统筹及定期检视等工作，下一步将逐步完善集团人工智能伦理体系并推动落地。

平安的 AI 伦理治理目标

技术是 AI 的核心，平安从数据、算法、应用三方面制定了伦理目标，并进行问题的监控。

数据使用

个人敏感信息审慎处理

对个人敏感信息的收集处理需经过主体明确同意，同时显示对个人信息的自动化处理，对其进行加密储存或采取更为严格的访问控制等安全保护措施。

隐私数据充分保护

对个人信息的使用不超出收集个人信息时所声明的范围，严格控制共享与开放使用，遵守集团数据治理与信息安全需求。

算法研究

技术透明

在不伤害算法所有者利益的情况下，公开其人工智能系统中使用的源代码和数据，避免“技术黑箱”。

算法可靠

在一定时间内、一定条件下算法可以无故障地实现特定的功能，并对非法数据的输入做出适当的应对，避免输出具有伦理风险的结果。

决策可解释

算法所有者或使用者应尽可能地对算法的过程和特定的决策提供解释，维护算法消费者的知情权，避免和解决算法决策的错误性和歧视性。

运行结果可验证

在一定条件下可以复现算法运行产生的结果。

行业应用

服务人类

应用人工智能技术的目的不应违背人类伦理道德的基本方向，在使用过程中不作恶。

公平无偏

使用完备且相对中立客观的数据训练模型，力保人工智能的算法及应用不具有某些偏见或歧视。

平安的 AI 伦理治理应对措施

平安积极参与人工智能全球治理，加强内部与外部的组织协调，构建多维度、多层次的人工智能治理体系，推动人工智能的下一步发展，实现经济利益与道德伦理之间的综合平衡。

内部管理

成立平安 AI 伦理委员会

研究委员会

- AI 伦理问题研究
- 政府监管 \ 行业标准和政策研究
- AI 伦理技术解决方案研究

管理委员会

- AI 业务伦理审核
- AI 伦理事件调查
- AI 伦理危机公关

制定平安 AI 伦理管理体系

- 确认 AI 业务伦理原则
- 明确 AI 业务边界
- 制定违规惩处规则
- 定制 AI 伦理事件应对机制
- 细化 AI 业务的具体标准、要求

落实伦理评估标准

- 全方位评估，加强管理制度执行检测。
- PDCA：plan-do-check-action，持续改进，有效管理。
- 注重架构及系统伦理设计、研究可解释方法加以采用

对外交流

近两年我国对 AI 伦理问题关注加强，国内相关协会、组织开始兴起，平安也积极参与其中。平安已加入国家人工智能标准化总体组、专家咨询组并建立联系；同时也参加新一代人工智能治理专业委员会前期意见收集。

加入 AI 伦理协会（规划中）

- 国家人工智能标准化总体组、专家咨询组（已加入）。
- 新一代人工智能治理专业委员会（已参与标准讨论）。
- 中国人工智能学会伦理专业委员会。
- 中国人工智能产业发展联盟。
- 人工智能的社会、伦理与未来研究专业委员会。
- Partnership on AI

制定平安 AI 伦理管理体系

- 参与国家行业标准建立，确立平安行业地位。
- 加强政府互动，谏言产业发展规划。
- 与国际标准接轨，维护平安国际声誉。

落实伦理评估标准

- 与高校密切合作，产学研结合共同实现 AI 伦理目标。
- 与同行企业深入交流，通过合力组织协会等方式共同助力 AI 技术应用健康发展。

平安相信 AI 技术是精确可控的，只要人们采取积极主动的态度及处理措施，未来 AI 技术的应用将会使社会变得更加安全、公正和美好。

--- 陈心颖，平安集团联席 CEO。

附：平安集团数据治理声明与政策

人工智能技术的发展是以大数据为基础的，大数据的广泛应用使得之前离散分布的数据之间产生了关联，在带给人类更多便利的同时也带来数据安全的风险。从目前曝光的各类安全事件来看，数据泄露特别是个人敏感信息 / 个人隐私的泄露最为严重。正是出于对个人数据的保护，世界各国都出台了相关法律法规和管理制度来约束企业正确、安全的使用个人数据，也对各类违法行为进行严厉打击。我国先后颁布了《中华人民共和国网络安全法》、《信息安全技术 - 个人信息安全规范》、《信息安全技术 - 大数据安全管理指南》、《信息安全技术 - 数据出境安全评估指南》、《信息安全技术 - 个人信息去标识化指南》、《个人信息和重要数据出境安全评估办法》等相关国家法律法规和安全标准，要求企业必须严格执行信息安全、数据安全保护措施。

平安集团作为国内首屈一指的金融集团，每日产生的海量真实数据涉及数亿人口的金融、医疗、交通等与生活息息相关的方方面面，如何在充分做好数据利用的同时提供全面的数据安全保护是一项艰巨的挑战。平安集团目前已经建立了以客户数据保护为核心的数据安全治理模型，该模型从数据安全策略、数据安全管理和数据安全运营三个维度全方位进行客户数据安全防护。

数据安全策略方面，集团高级管理层以国家法规政策、监管合规要求为基线批准制定了数据安全战略方针，重点保护客户隐私数据，并提出了数据安全零容忍的管理要求。

数据安全管理方面，集团成立了专门的数据安全管理团队，团队成员均为具有多年的数据安全防护经验的资深专家。结合集团的战略发展目标，制定了平安集团数据安全保护规范体系、风险及应急管理体系，同时以数据安全为基础，定期进行信息安全和数据管理合规审计，对集团各子公司管理高层设立了专门的考核标准，自上而下彻底做到数据安全管理全覆盖，实现数据安全人人有责，人人防护。

数据安全运营方面，从数据管理者、数据所有者以及数据使用者三种角色出发，按照分级分类、采集、传输、存储、处理、交换以及销毁的数据全生命周期的制定客户数据安全管理流程，最大程度规范数据操作流程，确保数据安全，保护用户个人隐私信息。为了确保数据安全治理模型的落地执行，集团信息安全团队、系统研发团队、基础架构团队与数据安全管理团队通力配合，共同推进数据安全保护的管控措施落地，从业务系统、主机安全、网络防护等方面增强了数据安全的管控技术。