

外审评议与作者回复

一、外审评议

根据文档内容，我对文档提出以下意见。

1. 文章内容丰富，系统全面地阐述了生成式人工智能在证券公司衍生品业务中的应用，包括多个业务环节的具体应用方式，以及不同阶段的系统搭建方式。整体内容充实，论点清晰。

2. 文章比较务实地提出了短期方案和终局方案，充分考虑到了金融行业内部严监管控风险的现状。尤其是这些短期方案有很强的落地意义。可以给行业内计划做相同方向的工作者提供系统的有价值的建议，可称得上是一本简明指南。

3. 文章提供了多个具体案例，如银河证券利用百度智能云 AI 技术提高交易询报价效率等，有效支持了文档的论点。这些例子佐证了作者对趋势的判断。

4. 文章不足之处在于对 GenAI 在金融公司落地的建议中，提炼升华的部分不足，建议进一步提炼，以追求更强的泛化性。

5. 本文语言表达流畅，使用了规范的学术写作语言；本文引用了多篇相关文献，反映了当前该领域的研究现状。有两处 typo：①引言第一段第 8 行，“积极【相应】党”；②第 8 页段落“弱者监督强者悖论”的第 4 行末尾，“人类难以【对齐】的”。

综上所述，我对这篇文章的质量表示认可，建议予以采纳。

2024 年 3 月 28 日

二、作者回复

感谢外审专家的意见和建议。

对于第四条意见，已经对第四部分“发展方向和建议”下面第一段进行修订，引入证监会刚发布的《关于加强证券公司和公募基金监管加快推进建设一流投资银行和投资机构的意见（试行）》相关观点，进一步提炼并升华了 GenAI 应用在证券行业和机构的意义。同时进一步补充了金融监管的科技赋能作用。

对于第五条意见，笔误已经完全按照专家老师的意见完成修订，并再次通读文章避免出现其他错误。

作者：张橙逸

2024 年 3 月 31 日

未来已来：生成式人工智能如何落地证券公司？

——从场外衍生品业务场景谈起（附录）

附录 1 生成式人工智能（GenAI）的关键技术

GenAI 主要包括两个关键技术：深度学习和自然语言处理。

一、深度学习（Deep Learning, DL）

DL 属于机器学习（Machine Learning, ML）领域，它基于神经网络，最终目标是让机器能够像人一样具有分析学习能力，能够识别文字、图像和声音等数据。深度学习可以自动对数据进行筛选，自动提取数据的高维特征，简单原理如图 1。

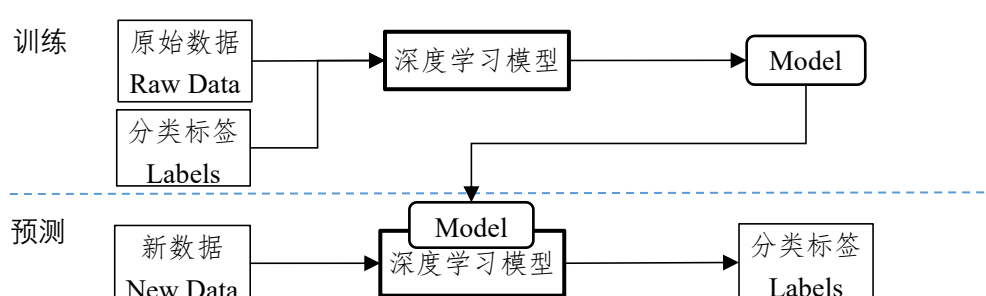


图 1 深度学习原理

（一）生成对抗网络模型

生成对抗网络（Generative Adversarial Networks, GAN,）是深度学习领域里的一个重要模型，旨在找出给定观测数据内部的统计规律，并且能够基于所得到的概率分布模型，产生全新的、与观测数据类似的数据。具体的，GAN 有两个神经网络，一个是判定器，一个是生成器。生成器的目的是通过调整参数使其生成的样本是真实的，而判定器想要最大化判定器的输出，两个神经网络呈现对抗的状态（Adversarial Training，图 2）。

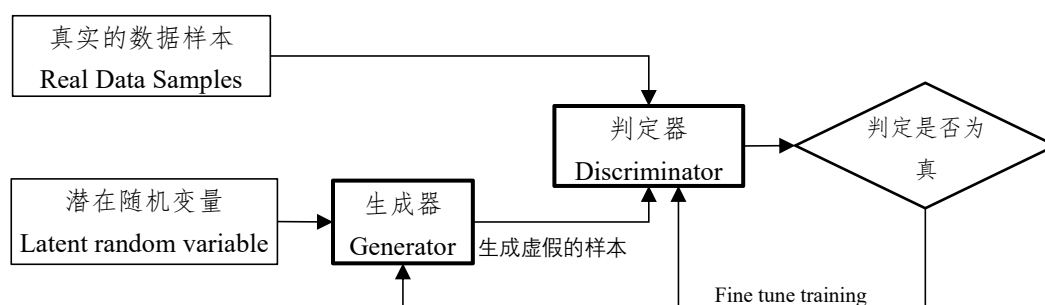


图 2 生成对抗网络原理

（二）大语言模型

大语言模型（Large Language Models, LLMs）是深度学习领域里的另一个重要模型，专门设计用于处理和理解自然语言文本。其通常包含数十亿甚至数万亿个参数，并使用大量的文本数据进行训练，从而能够理解和生成与人类语言类似的文本。大语言模型的核心是神经网络，特别是所谓的“变压器”（Transformer）架构（图 3），这是由谷歌在 2017 年提出的一种注意力机制。这种架构使模型能够同时关注输入序列的所有元素，从而在处理语言时更好地理解上下文。

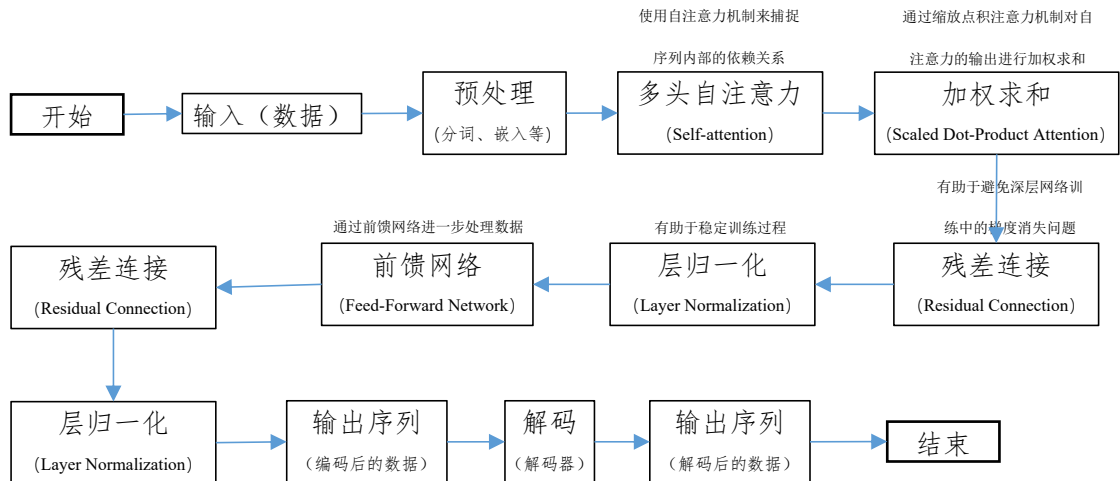


图 3 大语言模型 Transformer 架构运行逻辑

人类偏好对齐（Human Preference Alignment）是大语言学习中的一项重要训练方式，是指人工智能（AI）的行为和决策与人类的价值观、道德标准、偏好和期望相一致的过程。这种对齐涉及确保 AI 系统能够理解和遵循人类的意愿，如果 AI 和人类的价值观并不能对齐，可能会出现：AI 的行为并不符合人类意图的目标；在多种设定目标冲突时做出错误取舍；伤害人类的利益；脱离控制。但同时，由于人类偏好对齐关注的并不是“什么是正确的”，而是专注于与人类的意图（无论是什么）保持一致。这就会产生与人类社会类似的问题，比如歧视、错误认知等。

二、自然语言处理（Natural Language Processing, NLP）

即研究如何让计算机能够理解、处理、生成和模拟人类语言的能力，从而实现与人类进行自然对话（图 4）。它包括多个子领域，如语言模型、语义分析、机器翻译和情感分析等。自然语言处理技术是大语言模型更广泛的应用，其目标是实现机器和人类更平等、更流畅地沟通交流。

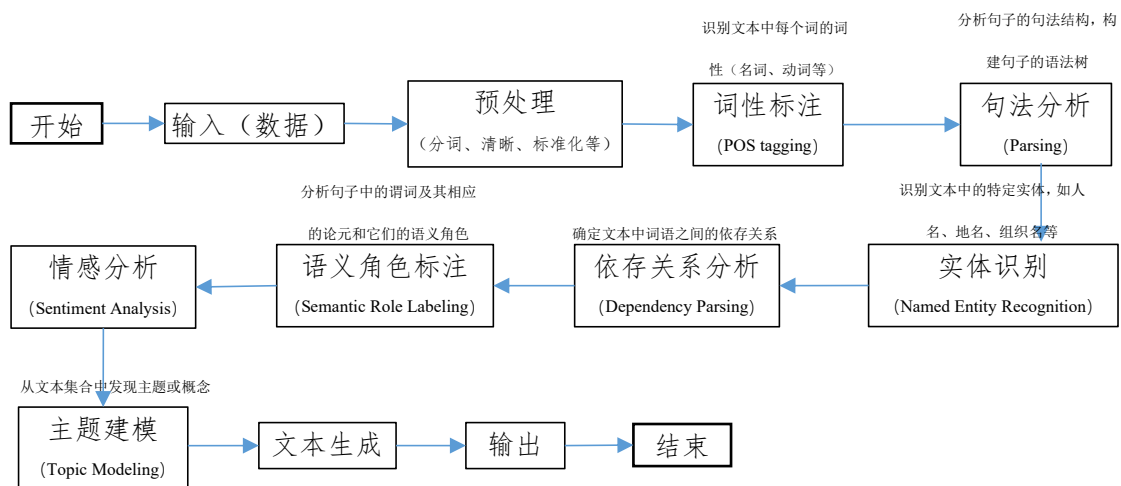


图 4 自然语言处理运行逻辑

附录2 境内通用大模型和金融垂类大模型开发信息表¹
(截至2024年2月)

境内通用大模型开发信息			
公司	大模型名称	发布时间	应用场景
华为	盘古大模型	2021年4月	物流监控、药物研发、气象预测、铁路巡检、金融等
达摩院	八卦炉	2022年	
腾讯	混元大模型	2022年4月	广告营销、金融风控、交互翻译、数智人客服等
阿里	通义大模型	2022年9月	办公学习助手、电商跨境模态搜索、开放域人机对话、AI辅助设计、法律文书学习、医疗文本理解等
百度	文心大模型	2023年3月	办公助手、商业营销、社交文娱、金融、能源、传媒等
字节跳动	飞书 My AI	2023年4月	总结会议纪要、分析经营数据、文本续写等
月之暗面	Moonshot 大模型	2023年10月	长程对话、长文本阅读、AI智能体，数学运算、内容创作、虚拟数字人
知乎	知海图 AI	2023年4月	问答社区场景，热榜摘要、搜索聚合
昆仑万维	天工大模型	2023年4月	智能问答、聊天互动、文本生成、AI作曲、游戏设计等
麒麟合盛	天燕大模型	2023年4月	文本交互、图像生成、视频创作、语音识别与合成等
360	360智脑	2023年5月	办公助手、AI收缩、数字员工、智能安防、税务等
京东	言犀大模型	2023年7月	在线服务机器人、智能语音外呼、智能语音应答等
网易	玉言	未发布	语音助手、智能客服、互联网收缩、游戏创作、媒体等
蚂蚁	贞仪大模型	未发布	
MiniMax	自研大模型	2021年12月	办公助手、游戏创作、广告营销、交易助手、代码生成
竹间智能	EmotiBrain	2022年6月	办公、营销，为金融、通信、互联网企业提供服务
澜舟科技	孟子 GPT	2023年3月	银行客服、智能投研、行业搜索、AI写作与绘画等
智谱 AI	ChatGLM	2023年3月	代码生成、内容创作、智能问答等
商汤科技	日日新大模型	2023年4月	代码开发、智慧医疗、商业营销、AI数字人、智能汽车等，目前在金融、医疗、珠宝等行业落地
毫末智行	DriveGPT	2023年4月	汽车端，城市 NOH、捷径推荐、智能陪练、场景脱困
出门问问	序列猴子	2023年4月	办公助手、商业营销、社交文娱、金融、能源、传媒等
第四范式	式说	2023年4月	企业级办公助手，数据资源治理、应用统一管理
科大讯飞	星火大模型	2023年5月	办公助手、代码开发、广告营销、客服、企业管理
清华大学	GLM-130B	2022年7月	
复旦大学	MOSS	2023年2月	搜索引擎、文生图、计算器等
上海人工智能实验室	风乌	2023年4月	气象预测
北京智源 AI 研究院	悟道 3.0	2023年6月	文本对话、多语种生成、文图生成、人脸编辑、代码写作等
武汉人工智能研究院 ²	紫东太初	2023年6月	手语教学、法律咨询、交通出行、医疗机器人、医学影像判读、短视频内容审核、音乐理解等

¹ 信息和数据来源于《2024年金融业生成式AI应用报告》《2024年证券行业发展趋势大展望：大模型 大券商 大服务》《中国AI治理的独立思考——生成式人工智能发展与监管白皮书》《2024年人工智能行业专题分析：券商大模型应用进展几何？》。

² 和中国科学院自动化研究所和华为联合研发。

达观数据	曹植	2023 年 7 月	申报材料自动生成、金融报告智能写作，金融、政府、医药、能源等领域
鹏城实验室	鹏城·脑海	未发布	
网易有道	子曰	2023 年 5 月	AI 虚拟人口语教练、中文作业批改，AI 学习机
学而思	MathGPT	未发布	数学领域的解题和讲题
理想汽车	MindGPT	2023 年 6 月	智能汽车场景，声纹识别、出行规划、AI 计算等
Vivo	BlueLM	2023 年 11 月	应用在手机应用中，进行基础手机指令操作、实时语音助手、查询信息以及一些基于手机端的办公应用
境内金融垂直类大模型开发信息			
公司	大模型名称	发布时间	应用场景
海通证券	泛海言道		综合应用前沿大模型技术，成功落地智能问答、智能研报、智能研发等业务场景
国金证券	国金证券大模型		LangChain 和 ChatGLM2 构建大模型，能够通过知识库检索大模型能够准确回答证券相关的专业问题
	国金代码大模型		aiXcoder 与国金证券共同构建代码大模型，提高了开发团队编码整体生产力和创造力
	国金 AI 办公助手		采用外部 API 和内部自研大模型相结合的模式，实现智能化的办公处理和管理，提高公司的办公效率
银河证券	百度智能云金融智能场外交易发现平台		基于百度智能云开元解决方案金融行业应用大模型底座，率先实现大模型在衍生品交易场景的应用落地
广发证券	广发证券大模型		涵盖了大模型问答&生成服务（ChatRAG）、代码生成服务（ChatCode）、文生图服务（ChatImage）和大模型 API 基础服务（LLM-API）
	金钥匙智能犇犇		智能客服
国泰君安	君极		率先提供一站式、全链路、专业化的公司级交易平台
	灵犀布道		以通用大模型为基石，辅以开源大模型集群的协同，对投顾、投研、投行、运营、风控、协作、运维等核心业务场景进行全面智能升级
	国芯证道		基于国产 AI 芯片的大模型解决方案，成功在国产芯片和全栈信创环境下进行完整的金融大模型工程，创新性提出 Early Exit、Token Skipping、权重共享、低秩分解、注意力优化等多项技术
申万宏源	SWHY-Matrix		为机构金融客户及专业投资者提供全方位一体化服务的财富管理平台，包含了极速交易服务、产品募集及运营服务、专业投研服务、金融创新及衍生品
中信证券	Bond Copilot		全方位支持债券承揽、承做、承销三大环节，具备自动撰写、研报摘要生成、债券法规问答、产品的定价分析、发行结构分析、投资人画像分析、市场数据分析、多维数据问答等功能
东吴证券	东吴秀财 GPT		基于 1.8 万亿通用语料、4000 亿金融语料在 40 台 H800 上训练出 130 亿参数证券垂类大模型，支持智能搜索、智能编排、专业回话、数据分析四大中台应用

兴业证券	兴业证券大模型		员工的智能助手，辅助海量的信息自助整理汇总
农业银行	ChatABC	2023 年 3 月	全方位的金融知识理解和智能问答应用
奇富科技	奇富 GPT	2023 年 6 月	
度小满	轩辕大模型	2023 年 6 月	
恒生电子	LightGPT	2023 年 6 月	为金融机构的投顾、客服、运营、合规、投研、交易等业务系统注入 AI 能力
蚂蚁集团	AntFinGLM	2023 年 9 月	在“研判观点提取”“金融意图理解”等众多领域达到行业专家水平
腾讯云	腾讯云金融行业大模型	2023 年 9 月	智能客服、智能咨询、辅助分析和决策服务
同花顺	HithinkGPT+AI iFinD	2024 年 1 月	协助投资者进行标的选择、标的诊断、资讯解读、投资教学等全方位投顾服务
东方财富	妙想大模型	2024 年 1 月	提升用户各场景使用体验和服务能力，未来或创造新的场景以挖掘客户全生命周期需求，进一步巩固流量优势并提高客户转化率

附录3 境内外涉及人工智能的监管和立法进程

一、境外监管和立法进程

（一）联合国

2023年10月26日，联合国秘书长古特雷斯（António Guterres）在纽约举行新闻发布会，成立新的人工智能咨询机构（Artificial Intelligence Advisory Body），负责人工智能的风险、机遇和国际治理。

（二）美国

2016年10月，美国发布《为人工智能的未来做好准备》（Preparing for the Future of AI）和《国家人工智能研发战略规划》（The National AI Research and Development Strategic Plan），提出7大人工智能研发策略。

2016年12月，美国白宫发布《人工智能、自动化与经济》（AI, Automation, and Economy），倡导各方开发、训练AI促进转型。

2018年1月，美国国防部发布该年《国防战略报告》（National Defense Strategy）强调AI对美国国家安全具有重大战略意义。

2018年5月，美国白宫成立人工智能特别委员会（Select Committee on Artificial Intelligence）。

2019年2月，美国总统行政令《维护美国在人工智能领域的领导地位》（Executive Order on Maintaining American Leadership in AI）。

2020年5月，《生成式人工智能网络安全法案》（GenAI Networking Security Act）发布。

2020年11月，《人工智能监管原则草案》（Guidance for Regulation of AI Applications）发布。

2021年1月，《2020年国家人工智能倡议法》（The National AI Initiative ACT of 2020）要求科学技术政策办公室（OSTP）成立国家人工智能计划办公室和国家人工智能咨询委员会。

2021年5月，发布《人工智能能力与透明度法案》（AI Capabilities and Transparency ACT of 2021）和《军用人工智能法案》（AI for the Military ACT of 2021）。

2021年11月，美国授权商务部国家标准与技术研究院（National Institute of Standards and Technology, NIST）开发人工智能风险管理框架（AI RMF）。

2022年2月，发布《算法问责法草案》（Algorithmic Accountability ACT of 2022）。

2022年10月，美国白宫科技政策办公室发布《人工智能权利法案蓝图》（Blueprint for an AI Bill of Rights，非法规，无法强制执行，只是提供指导），提出五项原则：安全有效的系统、算法歧视保护、数据隐私保护、通知和解释、替代方案。

2022年12月，发布《2022推进美国人工智能法案》（Advancing American AI ACT）。

2023年1月，发布《人工智能风险管理框架（第一版）》（AI Risk Management Framework）。

2023年3月，美国NIST成立可信与负责AI资源中心，推进AI RMF的执行。

2023年6月，美国NIST宣布将成立新的AI公共工作组，应对生成式AI挑战。

2023年10月，美国白宫发布名为《安全、可靠、可信地开发和人工智能》（Executive Order on the Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence）的行政命令。

2024年1月，美国商品期货交易委员会（CFTC）发布金融市场的AI应对征求意见稿。

（二）英国

2023年3月，英国发布《促进创新的人工智能监管方法》（A pro-innovation approach to AI regulation）白皮书，列出了监管机构在应对AI相关风险时应遵守的五项“注重价值观的跨部门”原则：安全性、保障性和稳健性；适当的透明度和可解释性；公平性；问责制和治理；

可竞争性和补救。

（三）欧洲

2018年4月，发布《欧洲人工智能战略》（European Strategy for AI）。

2018年5月，欧盟《通用数据保护条例》（General Data Protection Regulation）施行。

2018年12月，发布《人工智能协调计划》（Coordinated plan on AI）。

2019年4月，发布《可信赖人工智能伦理准则》（Ethics guidelines for trustworthy AI）。

2022年11月，欧盟《数字服务法》（Digital Service ACT）和《数字市场法》（Digital Markets ACT）生效。

2023年6月，欧盟《人工智能法案》（AI Act）在欧洲议会通过。

2023年底，欧盟《人工智能法案》（AI Act）正式生效成为法律。

该法案将AI风险分为4个等级：**禁用AI系统**：因“侵犯基本权利”而被明确禁用的系统，例如公共场所的实时生物识别系统。此外，《人工智能法案》严禁任何蓄意利用成人或儿童的弱点造成伤害的系统。**高风险AI系统**：用于或属于受欧盟产品安全立法约束的产品或《人工智能法案》附件III所列产品的系统。高风险AI系统受《人工智能法案》全面监管并须满足多项要求。**有限风险AI系统**：可与人类直接交互的系统。此类系统必须满足透明度要求并告知用户他们正在使用AI。**低风险或最小风险AI系统**：简单的AI系统，例如垃圾邮件过滤器或AI视频游戏。此类系统不受任何限制。

（四）澳洲

2019年11月，澳大利亚工业、科学和资源部发布《人工智能伦理框架》（Artificial Intelligence Ethics Framework），该框架是指导负责任的AI解决方案的自愿原则。该等原则旨在确保AI对人类有益且用于预期目的，以及AI系统负责人对系统的影响负责。

（五）亚洲

1. 中国香港

2021年8月，香港个人资料隐私专员公署（PCPD）发布《开发及使用人工智能道德标准指引》。内容包括人工智能的数据管理价值和伦理原则，并提供人工智能战略治理实践指导，以帮助组织制定适当的人工智能战略和管理模式，进行风险评估，并制定相关的监督安排等。

2023年8月，香港资讯科技总监办公室发布《人工智能道德框架》（Ethical Artificial Intelligence Framework）。

2. 中国台湾

2023年8月，台湾金融监督管理委员会发布《金融业运用AI核心原则与相关推动政策》草案。

3. 日本

2021年7月，日本经济产业省发布《实施人工智能原则的治理指南》（Governance Guidelines for AI Principles），对AI应用的潜在影响进行了考量，并就最大程度减少负面影响提供了指导。相关原则包括以人为中心，教育应用，隐私保护，安全保障，公平竞争，公正、问责和透明以及创新。

4. 新加坡

2021年1月，新加坡金融管理局领导的Veritas联盟发布了五份白皮书，阐述了在人工智能和数据分析（AIDA）公平、道德、负责和透明（FEAT）原则的评估方法。

5. 韩国

2023年2月，韩国国会通过《促进人工智能产业和建立可信人工智能框架法案》（Act on Fostering the AI Industry and Establishing a Foundation for Trustworthy AI），内容包括禁止任何人在未经政府批准的情况下开发AI以及对被认为足以影响人类生活的“高风险”AI进行分类。

6.越南

2022年6月，越南新的《保险业务法》（Law on Insurance Business）获得通过，其中政府鼓励保险公司使用包括AI在内的技术来销售创新型保险产品和服务。

7.印度尼西亚

2022年9月，印尼政府发布《国家人工智能战略》（National Strategy for AI），旨在推动印尼社会各部门实现自动化，其中提出针对AI、机器人技术和印尼科技企业进行投资。

8.马来西亚

2021年，马来西亚科技创新部发布《2021—2025年国家人工智能路线图》（2021-2025 National AI Roadmap），阐明国家促进AI发展的六项战略以及负责任AI的七项原则。

9.泰国

2024年1月，泰国发布《国家人工智能战略和行动计划》（National AI Strategy and Action Plan）。

10.菲律宾

2023年5月，菲律宾成立“人工智能发展管理局”，负责制定国家AI战略和框架。

二、境内监管和立法进程

2017年6月，《中华人民共和国网络安全法》施行。

2017年7月，《新一代人工智能发展规划的通知》（国发〔2017〕35号）印发。

2019年12月，网信办发布《网络信息内容生态治理规定》（国家互联网信息办公室令第5号）。

2000年9月，《互联网信息服务管理办法》（国务院令第292号）施行。

2021年1月，《中华人民共和国个人信息保护法》施行。

2021年9月，《中华人民共和国数据安全法》施行。

2022年3月，《互联网信息服务算法推荐管理规定》（国家网信办、工业和信息化部、公安部、市场监督管理总局联合令第9号）施行。

2023年1月，《互联网信息服务深度合成管理规定》（国家网信办、工业和信息化部、公安部联合令第12号）施行。

2023年2月，国家科技部在国务院新闻办公室举行的新闻发布会上表示将把人工智能作为战略性新兴产业，作为新增长引擎，继续给予大力支持。

2023年4月，《信息安全技术人工智能计算平台安全框架》（全国信息安全标准化技术委员会秘书处）征求意见。

2023年4月，中共中央政治局会议特别指出，要重视通用人工智能发展，营造创新生态，重视防范风险。

2023年6月，《关于发布深度合成服务算法备案信息的公告》（国家网信办）发布。

2023年8月，《生成式人工智能服务管理暂行办法》（国家网信办、国家发改委、教育部、科技部、工业和信息化部、公安部和广播电视总局第15号令）施行。

2023年10月，国家网信办提出《全球人工智能治理倡议》。

2023年10月，工业和信息化部、网信办、教育部、国家卫健委、中国人民银行、国务院国资委印发《算力基础设施高质量发展行动计划》（工信部联通信〔2023〕180号）。

2024年1月，工业和信息化部就《国家人工智能产业综合标准化体系建设指南》公开征求意见。

2024年2月，《人民日报》刊发中央金融委员会办公室和中央金融工作委员会联合拟写的《锚定建设金融强国目标 扎实推动金融高质量发展（深入学习贯彻习近平新时代中国特色社会主义思想）》主题文章，其中提出“数字金融把握机遇，加快金融数字化、智能化转型，提高金融服务便利性和竞争力”。

2024 年 2 月，国务院国资委召开“AI 赋能 产业焕新”中央企业人工智能专题推进会，强调中央企业要把发展人工智能放在全局工作中统筹谋划，深入推进产业焕新，加快布局和发展人工智能产业。

2024 年 3 月，《2024 年国务院政府工作报告》中，将“深入推进数字经济创新发展”列入 2024 年政府工作任务之一，并进一步明确要“深化大数据、人工智能等研发应用，开展‘人工智能+’行动，打造具有国际竞争力的数字产业集群”。

参考文献

- [1]Department of Industry, Science and Resources.Australia’s artificial intelligence ethics framework[EB/OL].<https://www.industry.gov.au/publications/australias-artificial-intelligence-ethics-framework>.
- [2]European Commission.Proposal for a regulation of the European parliament and of the council laying down harmonized rules on artificial intelligence (artificial intelligence act) and amending certain union legislative acts[EB/OL].<https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex%3A52021PC0206>.
- [3]Ministry of Economy, Trade, and Industry.Governance guidelines for implementation of ai principles ver. 1.1[EB/OL].https://www.meti.go.jp/shingikai/mono_info_service/ai_shakai_jisso/pdf/20210709_8.pdf.
- [4]Ministry of Science, Technology and Innovation.Malaysia national artificial intelligence roadmap[EB/OL].<https://www.mosti.gov.my/en/dasar/>.
- [5]Monetary Authority of Singapore.Veritas document 3 FEAT principles assessment methodology[EB/OL].<https://www.mas.gov.sg/-/media/MAS-Media-Library/news/media-releases/2022/Veritas-Documents-3---FEAT-Principles-Assessment-Methodology.pdf>.
- [6]National Assembly of Vietnam.Insurance business law[Z].2022-06.
- [7]OECD.AI strategies and policies in Indonesia[EB/OL].<https://oecd.ai/en/dashboards/countries/Indonesia>.
- [8]OECD.AI strategies and policies in Thailand[EB/OL].<https://oecd.ai/en/dashboards/policy-initiatives/http:%2F%2Fai.oecd.org%2F2021-data-policyInitiatives-27299>.
- [9]Republic of the Philippines House of Representatives.An act establishing a regulatory framework for a robust, reliable, and trustworthy development, application, and use of artificial intelligence (AI) systems, creating the Philippine council on artificial intelligence, delineating the roles of various government agencies, defining and penalizing certain prohibited acts[Z].2023-05.
- [10]The White House.Executive order on the safe, secure, and trustworthy development and use of artificial intelligence[EB/OL].<https://www.whitehouse.gov/briefing-room/presidential-actions/2023/10/30/executive-order-on-the-safe-secure-and-trustworthy-development-and-use-of-artificial-intelligence/>.
- [11]United Nations.Secretary-general's advisory body members - artificial intelligence[EB/OL].https://www.un.org/sg/en/content/sg/personnel-appointments/2023-10-26/secretary-generals-advisory-body-members-artificial-intelligence?_gl=1*1crkyh1*_ga*MjA1NTE4Mjc0My4xNzA4NjY3MjQ0*_ga_S5EKZKSB78*MTcwODY2NzI0My4xLjEuMTcwODY3MDE4NC4zNC4wLjA.*_ga_TK9BQL5X7Z*MTcwODY3MDE1My4yLjEuMTcwODY3MDE4NC4wLjAuMA.