CAICT 中国信通院

电信网络诈骗治理与人工智能应用白皮书（2019 年）

中国信息通信研究院安全研究所

2019 年 12 月

CAICT 中国信通院

版权声明

CAICT 中国信通院

前 言

近年来，我国电信网络诈骗活动猖獗，已成为影响人民群众安全感和幸福感的一大社会公害。根据党中央、国务院对防范治理电信网络诈骗工作有关指示要求以及工业和信息化部系列工作部署，信息通信行业相关单位和企业在工信部网安局指导下，紧跟反诈新形势新要求，不断创新思路举措，将人工智能等新技术广泛应用于防范治理技术能力建设中，取得了阶段性的明显效果。以“众智护网”2019 年度防范治理电信网络诈骗创新示范项目评选为例，从入选项目情况看，基础电信企业、互联网企业、安全领域专业技术厂商等信息通信行业企业单位均开展了基于人工智能的技术防范体系研究、开发和建设工作，涵盖了网络侧、业务侧、用户侧等多方面的电信网络诈骗治理需求，并取得了良好的治理成效和社会效益。

技术创新是一把双刃剑，人工智能技术在不断促进防范治理技术发展和进步的同时，也开始被诈骗分子所利用，带来了一定程度的风险隐患。特别是随着基于人工智能的“深度伪造”、群聊群控等诈骗手法的传播和应用，这些风险被进一步集聚、放大，引起了社会各界的关注。

本白皮书结合当前电信网络诈骗防范治理工作实践情况和人工智能技术的发展应用，系统梳理人工智能在治理工作中的积极影响及技术实践应用，同时剖析人工智能不当使用为治理工作带来的风险挑战，在总结国内外电信网络诈骗治理现状的基础上，深入分析当前人

工智能背景下治理工作的趋势走向和问题短板并研究提出对应的措施建议。

目录

一、人工智能在电信网络诈骗治理中的积极影响及技术实践

（一）人工智能在电信网络诈骗治理中的积极影响

随着云计算、大数据、移动互联网的不断发展，人工智能进入到了一个全新的发展阶段：基础算力不断增强，核心算法不断突破，应用场景不断丰富，成为引领创新发展的战略性技术。人工智能技术的蓬勃发展带来了强大的产业拉动效应，不断加速传统产业的数字化、智能化进程，驱动产业优化升级和生产力快速提升，在推进社会进步、经济发展、人民生活质量提高等方面产生了重大而深远的影响。

同时，人工智能技术的发展也为电信网络诈骗治理工作带来了积极影响：随着以大数据分析、机器学习、模式识别、知识图谱为代表的人工智能技术的部署应用，电信网络诈骗技术防范系统的识别准确度更高、监测拦截实时性更强、防护覆盖面更大，为治理工作的不断推进提供了强大动力，有效降低了电信网络诈骗带来的风险与危害。

1、防范识别准确度提高

对诈骗信息和行为的判定识别是防范治理技术工作的基础和前提。相比于人工判定方式，人工智能在诈骗识别方面的应用有效地提高了判定的准确度及可靠性。首先，通过人工智能技术可以对海量历史及实时数据进行多维度分析，挖掘不同数据间的内在联系，使得发现隐蔽诈骗线索和行为的能力快速提升。其次，通过对已有诈骗事件和数据的不断迭代学习，人工智能技术可以及时全面掌握各类诈骗活动的行为特征，从而准确识别具备相同和相似特特征的疑似诈骗信息

或行为。

2、监测拦截实时性加强

对电信网络诈骗行为进行实时预警拦截是防范治理工作的一个重要方面。利用人工智能技术通过对电信网、互联网、金融支付等各方数据的实时分析和深度计算，能够极大提升诈骗信息预警拦截的效率和及时性。从监测拦截角度看，现有人工智能技术可以实时监控疑似手机黑卡的诈骗流通轨迹，快速研判涉诈行为，实施对手机黑卡或涉诈号码的秒级溯源和分钟级关停等处置措施。从预警信息推送角度看，依托人工智能的强大算法和计算能力，可以针对疑似涉诈网站、APP 的活动行为进行实时记录分析，并根据相关模型对疑似高危受害用户进行实时的预警提醒。

3、反诈防护覆盖面变大

随着基于人工智能的防范治理能力不断提升，反诈技术的防护范围得到极大扩展。从数据分析角度看，人工智能技术能够对海量电信和网络数据进行关联分析并提取有效信息，对疑似受害人群和疑似诈骗团伙的搜索查找覆盖范围大大增加。从通讯联络角度看，基于人工智能的防范治理技术覆盖到了网络侧、业务侧和用户终端侧等信息通信过程的各个环节，极大地提高了反诈安全防护的用户覆盖范围。从诈骗信息监测角度看，当前人工智能技术能够有效识别诈骗文本、图片、音视频等各类诈骗信息传播手段，内容监测范围得到明显拓展。

（二）人工智能在电信网络诈骗治理中的技术应用

1、基于大数据分析的技术应用

基于大数据分析的电信网络诈骗防范治理技术应用以数据挖掘分析结果为驱动，整个过程包括“数据采集、数据处理、数据挖掘”等多个环节。

在数据采集和处理层面，主要有三种数据来源：在企业自有系统中沉淀的数据、在网上采集爬取的数据和从第三方购买的数据。这些数据经过智能化处理清洗后为后续开展数据分析和挖掘，识别电信网络诈骗行为，构建完备的技术防范体系奠定了数据基础。

在数据挖掘层面，利用大数据的挖掘能力可以发现诈骗行为的典型规律，精准识别诈骗分子和诈骗行为，进而对电信网络诈骗进行准确预警。

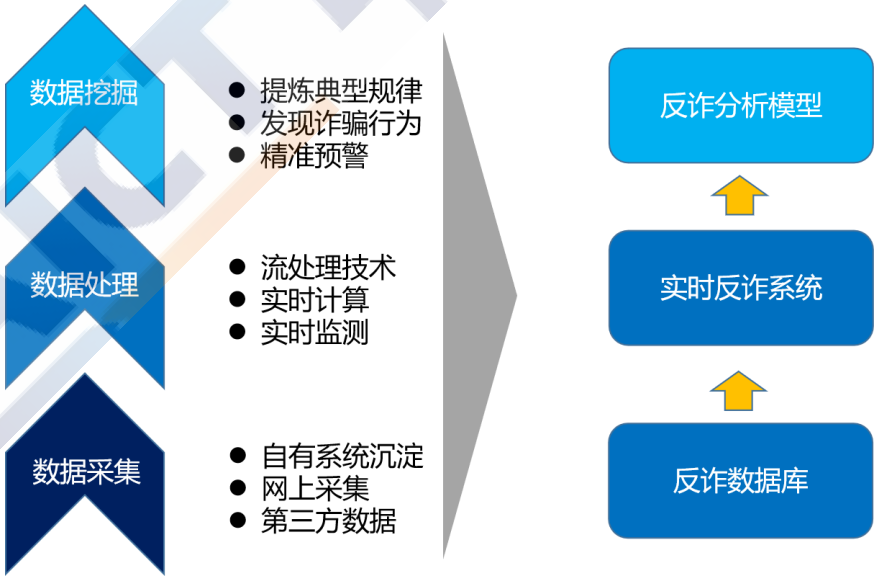

图 1 基于大数据分析的技术应用

2、基于机器学习算法的技术应用

基于机器学习算法的电信网络诈骗防范治理技术应用可以分为

分类和聚类两种应用形式。分类算法通过已知的诈骗样本、案例数据进行模型训练，在此基础上对新的行为事件进行涉诈风险分析预测。聚类算法通过全局分析和高维空间聚类，在无诈骗样本数据的情况下找出数据中隐含的共同特征，从而完成大规模关联诈骗团伙的自动发现。

通过机器学习两种算法的互相结合，可以有效提升发现识别诈骗行为和团伙的技术能力。以涉诈互联网社交账户识别发现为例，根据诈骗行为在多维空间向量上距离相近的特征，通过构建以登录时间、浏览器类型、IP 地址、GPS 地址、昵称修改等为特征的多维空间向量，利用聚类算法可以将疑似诈骗行为或账户聚为一组并抽取该群组的共性信息生成训练数据。基于聚类算法生产的训练数据，分类算法能够在此基础进行模型训练并进一步发现共性样本群组之外的诈骗行为和账户。两种算法相辅相成，为诈骗风险预警提供高效的检测和研判能力。

3、基于模式识别的技术应用

基于模式识别的电信网络诈骗防范治理技术聚焦已知诈骗行为的样本数据特征，通过分析归纳得到诈骗行为的多维度特征属性并形成涉诈资源模板库，结合自然语言处理、生物特征识别及大数据挖掘分析等技术，对目标对象进行相似度交叉比对分析，研判得出目标对象的涉诈风险，在诈骗电话、诈骗网站的判定识别领域有广泛应用。

以诈骗网址检测识别为例，在提取目标网址的标题、关键词及页面标签元素等多种特征属性的基础上，通过计算目标网址与诈骗资源

模板库中的网址样本之间的特征距离，判断两者之间的相似度。一般来讲，两者特征距离越近说明相似程度越大，目标网址涉嫌诈骗的可能性就越大。

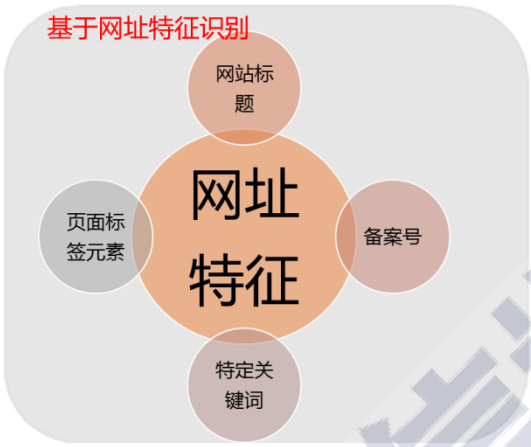

图 2 诈骗网址监测识别

4、基于知识图谱的技术应用

知识图谱是一种基于图的数据结构，可以看作是由数据绘制出来的一张知识图。在防范治理应用中，知识图谱技术能够聚合关联多种数据源，针对监测目标分析识别其脉络、趋势以及特征，在关键诈骗信息搜索、账号涉诈风险评估、诈骗团伙研判、异常行为分析等方面具有重要应用。

以银行卡全周期异常行为分析为例，通过知识图谱技术对全周期内的银行卡关联数据进行分析，并以图的方式进行数据融合及可视化，从而找到银行卡异常行为的内在关联，提升对诈骗资金流的打击效率。相比于人工核验，其效率提升 10 余倍，准确率约为 99.5%。

图 3 基于知识图谱的银行卡异常行为分析

（三）人工智能在电信网络诈骗治理中的实践应用

结合电信网络诈骗治理实践，目前人工智能技术在诈骗电话检测、恶意网址拦截、诈骗文本识别、诈骗风险预警等场景下已经有深入的应用及探索。

1、诈骗电话检测

电话诈骗是电信网络诈骗中的突出类型，诈骗数量较多，诈骗内容花样层出不穷，社会影响恶劣。及时尽早地感知、发现诈骗电话，是减少此类电信网络诈骗活动损失必不可少的方法。由于诈骗分子掌握大量用于诈骗的号码，传统的基于号码的阻断策略容易被绕过，及早发现诈骗电话的难度在不断增大。人工智能时代，利用大数据分析、模式识别等技术可以对海量通信数据进行预处理和数据融合分析研判，从而进行全天候全方位诈骗电话检测，能够有效增强针对电话诈骗的防御和反制能力。

以极限元人工智能语音识别、语音关键词检索技术为例，其推出了诈骗电话检测技术解决方案，可以有效检测识别邮包快递、社保卡、信用卡等十余种诈骗场景及类型。通过人工智能语音关键词检索技术对疑似诈骗电话进行匹配识别，不仅提高了检索识别效率，还能有效

避免误判增加识别准确率。

2、恶意网址拦截

恶意网址是指恶意种植木马、病毒等恶意程序的网站，是目前诈骗分子用于窃取用户个人信息、实施网络勒索诈骗的重要方式。人工智能技术用于恶意网址监测拦截主要基于机器学习的分类和聚类方法。

分类方法主要是收集样本数据进行归一化处理后构造分类器识别是否恶意：首先是根据已经标记的 URL 数据集提取静态特征（主机、URL、网页信息等）和动态特征（浏览器行为，网页跳转关系等），对提取特征进行归一化处理后通过算法构造分类器来识别恶意网站。

聚类方法略有不同：首先在网页采集的 URL 数据集中提取连接关系、URL 特征、网页文本信息等特征，再根据聚类算法模型将 URL 数据集划分为若干聚类，相似度较高的 URL 数据会在同一聚类中，反之则归为不同聚类，最后对已标记的聚类结果识别待测 URL，判断其是否为恶意网页。

以腾讯安全云库为例，其利用云端监测技术，基于机器学习算法和大数据平台，进行云端全网海量网址特征比对检测、页面相似度分析，同时配备人工审核团队审核，建立实时更新的黑白样本库。在不同的应用场景下，基于海量数据的样本库不断调整和适配算法，最终选取严格保证准确率的机器学习模型，在保证及时性的同时，也能达到较好的识别准确率。

3、诈骗信息审核过滤

诈骗信息包括诈骗文本、诈骗图片（头像）、诈骗音视频等多种内容。其中，以诈骗短信为代表的诈骗文本信息是出现最早、较为突出的诈骗类型，由于成本低、目标用户范围广，颇受诈骗分子青睐。依靠社交网络平台传播的诈骗图片（头像）、伪造音视频等诈骗信息形式则更为复杂，通过合成虚假头像、语音或视频能够伪装成受骗人群的亲人好友，并进而以多种理由话术逐步实施诈骗。人工智能在诈骗信息识别中具有重要作用，通过自然语言处理，计算机视觉，和语音识别等深度学习技术，可以实时分析并判断消息的可靠性和敏感程度，及早识别出涉诈或伪造内容，将对公众有危害的信息及早扼杀在摇篮里。

以 360 应龙综合反诈平台为例，其运用模型算法、人工智能等前沿技术，对疑似诈骗的信息、图片、账号进行识别、分类、拦截。高峰时期，在国内单个省份一天能有效识别拦截虚假涉诈信息近 70 万条。

4、受害者发现预警

及时发现识别高危受害者并进行快速预警提醒是防范阻断电信网络诈骗活动，切实保障人民群众财产安全的重要手段。人工智能的发展为解决高危受害人发现及诈骗信息预警提供了可能，通过对电信、网络数据的融合分析，能够有效、及时地发现诈骗活动并进行预警。

以阿里钱盾反诈公益平台为例，通过对涉及诈骗的黑号码、黑卡

等黑名单库的共享共建，利用人工智能模型算法分析处理用户的异动情况，不断更新迭代预警模型，进而形成自动化的诈骗行为和受害用户的发现预警的能力，通过对疑似受害用户的及时预警提醒，大大地降低了群众资金损失。

5、诈骗团伙研判分析

当前电信网络诈骗呈现出团伙化、组织化特点，且诈骗分子往往分布在不同地区，乃至不同国家，相互勾连配合实施诈骗。检测识别某一诈骗帐号或某一诈骗分子往往效果不佳，识别诈骗团伙，进行全面的防范治理才是更加高效有用的治理方式。人工智能技术的应用，可以大大提高诈骗团伙识别能力，助力反电信网络诈骗工作推进。

以腾讯“反诈大脑”为例，通过人工审核清洗出互联网侧精准的电信网络诈骗举报数据，经由诈骗团伙识别模型进行团伙识别、聚类分析，将离散的单点举报聚类成可疑诈骗团伙，输出诈骗团伙作案设备、嫌疑人、受害人等相关线索，可以有效提高电信网络诈骗防范治理工作效率。

二、人工智能给电信网络诈骗治理带来的风险与挑战

人工智能在推动防范治理工作的不断前进的同时，也开始被不法分子用来实施诈骗，当电信网络诈骗用上人工智能，往往会扩大诈骗的危害广度和深度，给防范治理工作带来一定的风险与挑战。

（一）电信网络诈骗实施的四个主要环节

一般来讲，电信网络诈骗活动可以归结为精准信息获取、诈骗脚本设计、通讯联络诱导、资金支付转移四个关键环节。**在精准信息获取环节**，诈骗分子主要非法窃取或购买社会上各行各业泄漏的个人信息，包括身份证信息、电话号码、家庭地址，以及网络账号和密码、银行账号和密码等信息。目前看，个人信息泄露是精准诈骗的根源。**在诈骗脚本设计环节**，诈骗分子模拟真实的经济社会活动场景，精心设计各种诈骗脚本，如近期高发的代办信用卡、兼职刷单、冒充网购客服、冒充公检法等诈骗案件。**在通讯联络诱导环节**，诈骗分子通过电话、短信、互联网等通讯渠道联络受害人，利用之前设计的诈骗脚本与获取的受害人个人信息，骗取受害人信任进而实施诈骗。**在资金支付转移环节**，诈骗分子引导受害人通过银行转账、网上支付等方式向其指定账户转款，再经由预先设计的诈骗分赃销赃渠道快速从指定账户中转移受害人资金。

（二）在“精准信息获取”环节

在“精准信息获取”环节，人工智能加剧个人隐私泄露。

一是借助人工智能技术更容易通过网络攻击破坏系统并窃取数据。随着机器学习算法研究的推进，智能软件技术快速发展，使信息窃取更为便捷。从近几年多起侵犯公民个人信息案件看，犯罪团伙往往以人工智能软件为犯罪工具，通过“撞库”等方式，非法获取在网站后台存储的用户注册信息。在人工智能的帮助下，智能恶意软件攻

击效率更高，针对性更强，可轻易破坏受害者的防御系统，获取系统中的个人信息，例如聊天记录、家庭关系、个人习惯、音频视频等。

二是利用人工智能技术可获取大量用户生物特征信息。近期走红网络的“ZAO”、DeepFakes 和 DeepNude 等换脸 APP，用户借助 AI 技术只需要一张正脸照，就可以替换为影视作品或者小视频中的人物，生成以自己为主角的视频片段。但同时关于换脸的安全性及隐私性问题，很快引起了社会的广泛讨论和关注。用户如果想要下载或分享换脸视频，则需要在摄像头前进行张嘴、眨眼、转头等动作进行验证，这一过程中搜集了用户的面部特征等核心个人信息。生物特征信息的泄露是永久、不可逆的，也就意味着对信息滥用者来说，存在着“一次窃取，永久有效”的“超便捷性”和“超高性价比”。

三是利用人工智能技术“晒密撞库”精准提取个人重要信息。2018 年 10 月，苹果 APP Store 爆出大规模的“免密”盗刷事件，主要原因是不法分子进行“晒密撞库”导致的个人信息泄露。所谓“晒密撞库”就是不法分子在窃取网站数据库后，通过验证的方式筛选账号密码等对应的有价值数据。利用图像识别技术的“打码平台”，提供图片验证码自动识别服务，为批量识别、提高“晒密”效率、突破验证码安全体系起到了关键作用，该平台识别精准度极高，验证码识别正确率达 95%以上，实现了批量晒密撞库的功能。

四是获取个人信息的途径更加多样化，方式更加隐蔽。剑桥大学的一项最新研究表明，利用人工智能技术可以通过“偷听”获取个人信息。如当用户轻敲手机和平板电脑的屏幕时会产生声波，这些声波

通过设备的内置麦克风收集、恢复，就可“听到”敲击屏幕的声音，结合人工智能算法，听到的声音与键盘位置关联，就可实现对用户信息的“偷听”。另外，当前智能家居市场兴起，涌现出大量被声音控制的产品和控制产品的软件，均存在窃取隐私的风险。不法分子利用人工智能技术，将智能手机与智能家居变成一张无孔不入的网，将用户个人隐私网罗其中。

（三）在“诈骗脚本设计”环节

在“诈骗脚本设计”环节，利用人工智能技术设计定制化脚本、精准选取受害人。

一是针对特定人群产生定制化脚本。人工智能技术能够越来越容易精确的模仿一个人，越来越多的语音交互程序会让特定人的声纹更容易被收集建立。智能音箱越来越熟悉主人声音和身份账号背后的关联。对于人工智能，最大的威胁，并不是替代人类的工作岗位，而是经过大量的数据输入和深度学习之后，计算机可以准确判断一个人的喜好、状态甚至模拟人类做出决策。2019 年 12 月 12 日，新华社首个智能化编辑部正式建成并投入使用，开启了一场新闻生产与传播的智慧革命，“双十一”期间，阿里巴巴智能设计平台“鲁班”自主设计超过四亿张海报，给用户展现“千人千面”的购物界面。通过对特定人群的行为特征的训练学习，人工智能系统同样可以生产出千万个定制化的诈骗脚本。

二是利用人工智能技术精准筛选受害人。随着人工智能应用的发

展，公民个人信息在采集、存储和处理的方式上发生了很大变化。大量的个人信息被采集下来，组成一个多维度智能数据库，这些信息被深度地整合分析，挖掘出更有价值的信息。通过分析公众发布在网上的各类信息，诈骗分子会根据所要实施的骗术对人群进行筛选，从而选出目标人群。例如实施情感诈骗时，可以筛选出经常发布感情信息的人群；实施金融诈骗时，可以筛选出经常搜集投资信息的人群。

（四）在“通讯联络诱导”环节

在“通讯联络诱导”环节，人工智能使通讯联络手段更加隐蔽，诈骗信息传播更加精准。

一是机器人的大量使用让沟通人力成本更低。以机器人拨打诈骗电话为例，应用深度学习技术，将接听人问题自主学习进知识库，并随着交互数据的不断累计总结，自动更新维护。同时，基于语音识别、自然语音处理、语音合成三大技术支持，人机交互流畅自然。人工拨打方式每人每天可以打 300-500 个电话，而一个外呼机器人每天最多可打 5000 个电话，人力成本降低 80%，工作效率提升 200%。而聊天机器人软件，一个人可以同时控制着几十个端口，在微信、婚恋网站中假装成各种“美女”，和对方聊天交友，批量“谈恋爱”，不少人就被这些“虚拟美女”诱惑上当受骗。另外，一个充分利用人工智能的 8 人钓鱼团队能发挥 8000 人的水平，人工操作成本被大大的缩减。

二是通讯联络方式隐蔽性更强。利用人工智能技术进行声音合

成、换脸变声，甚至连愤怒、高兴等不同语气情绪都能够做到惟妙惟肖。用伪造的声音或视频与受害人联系，可信度高、迷惑性强，实现以假乱真。2018 年底，河北发生一起微信语音转发诈骗，不法分子利用人工智能技术通过提取语音文件，转发他人语音实施诈骗。据《每日邮报》信息，今年 3 月份在英国曝出诈骗分子利用人工智能语音模仿软件冒充某能源公司高层，骗取子公司 CEO 22 万欧元。基于 AI 的恶意软件还可以搜索攻击对象的电子邮件与其他文件，模仿受害者的写作风格，发送真假难分的钓鱼信件实施诈骗。

三是非法获取号码资源效率更高。不法分子利用人工智能技术主要为突破企业安全策略，进行技术对抗。在人工智能技术的应用初期，不法分子主要通过脚本控制大量的设备进行拟人化操作，防止僵尸账号被企业安全策略打击，比较典型的是通过群控框架及脚本，控制上百甚至上千台手机组成手机墙，保持帐号活跃性，伪装成普通用户帐号。现阶段，不法分子已利用人工智能技术突破企业验证码体系，进行恶意、批量注册，以获取储备号码资源，提供给诈骗分子使用。以打码为例，在人工智能技术应用下，不法分子每秒可破解验证码次数达千余次，且成功率在 90%以上，降低了突破企业验证码安全策略的成本。

四是给信息源头治理和打击带来挑战。从治理实践看，不法分子不断利用新型技术设备实施诈骗，对抗拦截封堵，逃避追查打击。如诈骗分子利用智能群呼网关（如 GoIP、SIMBANK 等），通过远程操控、机卡分离实现诈骗呼叫异地落地，并使用智能化策略隐藏自身网络、

业务特征，设法规避公安部门落查打击和通信行业技术防范策略，对现有诈骗电话大数据预警处置模型产生冲击。从监测数据看，通过国际互联网入境的 GoIP 诈骗电话达到日均 10 万余次。

（五）在“资金支付转移”环节

在“资金支付转移”环节，人工智能的大量使用，给支付环节的使用带来了风险。在“金融科技”迅速兴起的大背景下，国内银行和支付机构纷纷利用人工智能推出新业务，如“智能语音”支付功能、智能理财机器人、人脸支付、“无感支付”、“刷脸”取款等，实现金融服务的智能化、个性化、定制化。人工智能在支付领域的广泛应用，同时也存在安全风险：**一是**信息泄露的风险。应用人工智能必然会面临海量数据采集和处理，这些数据一旦被成功攻击，会暴露用户的个人隐私，也极有可能对客户造成财产损失，甚至是人身安全。**二是**盗刷盗用风险。不法分子通过获取的用户面部特征、虹膜、声纹、指纹等生物特征信息，冒充用户身份盗刷盗用他人账户资金。

三、人工智能背景下国内外电信网络诈骗治理动态

近年来，人工智能技术为防范打击电信网络诈骗工作开辟了以“智”图“治”治理新格局，为促进国家治理体系和治理能力现代化增添了新动能，但同时也逐渐成为诈骗分子实施精准诈骗的“新利器”，为诈骗治理工作带来新挑战新风险。为做好诈骗治理工作，各国政府纷纷布局，瞄准精准诈骗实施主要环节、关键要素，统筹推进

诈骗治理工作有序稳妥进行。

（一）国际主要国家和地区治理经验做法

从国际上看，换脸换声诈骗、自动刷单诈骗等一系列利用人工智能技术的新型诈骗模式进入人们的视线，成为当前电信网络诈骗治理的痛点难点。为防范治理这种新型电信网络诈骗，世界主要国家和地区纷纷从深化个人信息保护、提升技术防范能力、强化企业守法自律、加强警示宣传教育等方面着手，开展推进打击防范治理工作，遏制人工智能诈骗活动的泛滥蔓延。

（1）注重立法先行，深化个人信息保护制度。

一是制定专门性法律法规。人工智能诈骗实施的关键要素是获取受害者的个人信息，因此个人信息保护是人工智能诈骗治理的源头。目前全球已有 126 个国家制定了专门针对个人信息保护的法律。2015 年日本出台《个人信息保护法》，细化了个人信息保护权相关规定，确立了对个人信息保护的一体化监督机制。2018 年，欧盟出台《通用数据保护条例》，建立了用户个人信息访问、修正和删除请求相关机制，明确指出未取得数据主体的同意，不得使用用户画像。随后，印度出台《2018 年个人数据保护法案（草案）》，明确要求数据控制者在使用数据画像时，要开展数据保护影响评估，否则不得处置。2019 年 6 月，美国国会先后提出了两部法律草案《深度伪造责任法案》和《2019 年深度伪造报告法案》，探索立法应对措施，以防范深度伪造技术潜在滥用风险。2019 年 7 月，美国弗吉尼亚州扩大了

正式生效《非同意色情法》（Nonconsensual pornography law）的修正案，将深度伪造内容纳入该修正案。

二是加大违法行为惩处力度。欧盟出台的《通用数据保护条例》是目前公认全球对个人信息泄露惩罚最严格的法律，明确要求在可行的情况下，不得迟于发现数据泄露后的 72 小时通知主管监管部门，未履行通知义务的企业将面临一千万欧元或全球年营业额 2%以上的处罚。此外，美国、澳大利亚、英国等西方主要国家和地区，均在立法中对“数据泄露通知”进行了规定，美国部分州规定，民事处罚分为 500 美金到 5 万美金不等、泄露一次 2500 美金、每天最高 1000 美金三种。2018 年 12 月，美国联邦议员提出《2018 年恶意伪造禁令法案》，规定制作深度伪造内容引发犯罪和侵权行为的个人，以及明知内容为深度伪造还继续分发的平台，可处以罚款和/或长达两年的监禁，如果严重的煽动暴力或扰乱政府或选举，监禁将长达 10 年。

（2）强化技术手段，提升诈骗音视频识别与拦截能力。

一是政府主动出击，研制开发伪造视频拦截工具。2018 年 8 月，美国国防部积极开展新型诈骗防范技术研发，依托媒体取证 （Media Forensics）项目，研发出全球首款“反换脸”AI 刑侦工具，能够更高效更准确的自动检测出被修改过的图片和利用深度伪造技术生成的视频，防止诈骗分子利用合成信息诱导实施诈骗，斩断人工智能诈骗黑产链条关键节点。2019 年 6 月，美国众议员在提交美国众议院审议的情报授权法案中提及要开展深度伪造鉴别技术竞赛，法案要求国家情报总监通过情报高级研究计划局局长实施一项有竞争力的奖

励计划，刺激技术的研究、开发或商业化，以自动检测被机器操纵的媒体。

二是各方积极响应，开发出多项音视频造假检测技术。目前，联合国区域间犯罪与司法研究所人工智能与机器人中心正在研究检测虚假视频的技术。美国赛门铁克公司正在开发一种检测音频造假的技术，通过该技术可以准确检查出利用人工智能技术合成的音频内容。此外，美国的 Facebook 除建立了一个机器学习模型来检测虚假图片和视频外，还与 Partnership on AI、微软以及麻省理工学院、牛津大学、加州大学伯克利分校等美国多所大学合作，发起了“假视频检测挑战”(Deepfake Detection Challenge)活动，推动创建高质量的深度伪造视频数据集，促进检测识别技术的研究与开发，旨在利用人工智能技术更好地甄别深度伪造和合成内容、开发对抗深度伪造技术滥用的方法和工具，实现从源头上遏制人工智能诈骗案件的滋生与蔓延的目的。

（3）强化企业自律，加大对诈骗电话呼叫与转帐环节的管控。

一是政府主导，聚力诈骗资金转帐管理。2019 年 2 月，英国政府联合银行业界、消费者群体等多方力量，发布了《授权推送付款（APP）诈骗自律守则》，规定银行金融机构若未能达到相关要求而导致客户遭受 APP 诈骗，则需对客户损失进行赔偿。目前共有包括巴克莱银行、劳埃德银行、汇丰银行、Metro 银行、苏格兰皇家银行、国民西敏寺银行、桑坦德银行与全英房屋抵押贷款协会在内的 8 家机构签署加入《守则》。

二是企业自发，携手自动呼叫检测拦截。2019 年初，美国 AT&T 和 Comcast 两大通信公司携手，联合开展 SHAKEN/STIR（安全的电话身份重新访问/使用令牌安全处理断言的信息)协议的验证，并于 3 月 21 日宣布完成了在 AT&T 的移动电话网和 Comecast 的 VoIP 电话网络两个不同网络间的认证，实现机器人电话的自动检测与警示。与此同时，AT&T、Verizon 和 T-Mobile 等运营商承诺将在各自网络中采取措施部署该协议。

（4）加强警示教育，提升民众防范意识。

2018 年，以色列国家网络管理局（INCD）发布了“新型网络攻击”警告，提醒公众诈骗分子很可能利用 AI 技术模仿管理层，命令员工进行资金转移等。同年，澳大利亚税务局（ATO）发布了一则警告，提醒公众有不法分子冒充 ATO，并要求被害人使用比特币或其他加密货币支付子虚乌有的税费。2019 年，奋韩网联合韩国国家情报院、金融监督院等机构针对冒充国家公共机关（大使馆、公安局）诈骗以及招人取钱的问题进行联合宣传，警示用户要严防人工智能时代各类仿冒诈骗。

（二）国内信息通信行业主要治理举措与成效

基于人工智能的电信网络诈骗凭借其新颖的诈骗形式、精准的诈骗脚本、逼真的伪造音视频，让人防不胜防，已经严重影响了社会稳定和群众幸福感。为有效遏制基于人工智能的新型电信网络诈骗产生蔓延，及时铲除此类诈骗活动滋生的土壤，我国信息通信行业积极开

展了行业源头治理工作：工业和信息化部作为国务院打击治理电信网络诈骗新型违法犯罪工作部际联席会议工作单位和行业主管部门，全面贯彻党中央、国务院有关决策部署，在工信部网安局指导下，信息通信行业相关企业单位深入落实主体责任，纵深推进防范治理工作，形成了全链条多层次的治理格局。

（1）加强组织部署和任务落实，充分发挥行业主管部门引领作用。

一是巩固深化顶层统筹，扎实推进电信网络诈骗治理迈向新台阶。利用人工智能技术的电信网络诈骗一般呈现涉及面广、危害性大、对抗性强三大特点，加强组织领导与统筹布局是保障治理工作有序推进的坚实基础。工信部高度重视电信网络诈骗防范治理工作，依托防范打击通讯信息诈骗工作领导小组，全面负责统筹相关重点任务；时刻关注基于人工智能的电信网络诈骗新形势以及演进态势，积极研究形成治理新方案新思路；针对当前诈骗治理重点难点，工信部网安局多次组织指导相关专题会议，部署督导人工智能诈骗治理相关工作，确保治理工作有效推进。同时，为保障促进人工智能在安全领域应用，工信部发布《促进新一代人工智能产业发展三年行动计划》，构建完善以法律法规、监管政策、技术标准、监督执法为核心的人工智能安全监管体系。

二是制定出台相关管理制度，强调个人信息保护与实名登记并重。个人信息保护以及电话实名制作为治理人工智能诈骗的重要手段，已成为当前监管部门的治理工作重点。2013 年工信部发布《电

信和互联网用户个人信息保护规定》（工信部令 24 号）以及《电话用户真实身份信息登记规定》（工信部令 25 号），明确了用户个人信息保护以及电话实名登记的相关要求。随后，工信部网安局先后制定出台《工业和信息化部关于进一步做好防范打击通讯信息诈骗相关工作的通知》（工信部网安函〔2015〕601 号文）、《关于进一步防范和打击通讯信息诈骗工作的实施意见》（工信部网安函〔2016〕452 号文）、《工业和信息化部关于纵深推进防范打击通讯信息诈骗治理工作的通知》（工信部网安函〔2018〕157 号文）等文件，明确提出要加强电话用户实名登记，加强用户个人信息保护等。截至目前，基于人工智能的用户实名登记人像比对试点及在线视频实人认证等工作正在逐步深化落实，用户实名制信息准确率超过 95%。2019 年 5 月，工信部网安局制定实施《工业和信息化部办公厅关于进一步做好 2019 年防范治理电信网络诈骗重点工作的通知》（工信厅网安函〔2019〕108 号文），明确指出要坚持关口前移，要及时汇总分析电信网络诈骗样本模板、关键特征等信息，利用大数据技术防范打击电信网络诈骗。

三是强化通报约谈，持续强化责任落实。近年来，越来越多的诈骗分子利用人工智能技术，借助企业管理漏洞，实施电信网络诈骗，强化监督检查、压实企业责任成为行业监管部门事中事后监管的主要手段。截至目前，工信部共计下发电信网络诈骗风险行业和重点地区通报 33 期，累计点名通报问题企业 60 余次。2019 年 7 月，针对媒体公开报道和用户曝光的“ZAO”APP 存在数据安全问题，工信部网

安局对北京陌陌科技有限公司相关负责人进行了问询约谈，要求其依法依规组织开展自查整改，并进一步加强新技术新业务安全评估，积极防范自有业务平台被利用实施人工智能电信网络诈骗。

四是建立健全技术手段，巩固升级诈骗治理联动防线。善于把人工智能技术运用到预防打击电信网络诈骗工作中，是推动治理体系和治理能力现代化的重要举措，是确保与诈骗分子较量时占得先机、赢得主动的关键。为进一步提升电信网络诈骗发现、预警和处置能力，在工信部网安局指导下，地方通信管理局、中国信息通信研究院、国家互联网应急中心等单位，建设了全国诈骗电话防范系统，初步构建了部、省两级反诈骗治理技术体系，并持续推进基于通话行为和语义内容的深度学习诈骗电话识别、基于文本分析的诈骗短信识别以及基于无监督机器学习算法的涉诈网站识别等技术的实际应用落地。

五是进一步加强宣传教育，提升用户风险防范意识。在人工智能技术的推动作用下，电信网络诈骗呈现成本低、花样新、波及广、迷惑性大等特点，增强用户识骗、防骗意识和能力变得十分必要。近年来，工信部紧跟诈骗新形势新特点，利用大数据技术分析研判当前诈骗趋势，聚焦人工智能诈骗重点地区、易受害人群，先后组织“号码安全伴你我，全民防骗公益行”、“反诈校园行”等活动 10 余次，在北京、上海、广东等多地校园内引发热烈讨论，有效提升了高校师生的防范意识。2019 年，工信部通过工信微报、人民邮电报等渠道编发系列诈骗风险预警提示 40 余篇，持续宣传电信网络诈骗政策、工作动态、诈骗案例等信息。同时，指导三家基础电信企业利用短彩

信、微信公众号等，发送风险提醒信息十亿余条，及时向公众发布人工智能诈骗经典案例与防范之策，有效防范诈骗分子利用人工智能等新技术实施电信网络诈骗。

（2）企业充分发挥技术优势，加强新型防控技术研发。

一是基础电信企业积极落实主体责任，不断推进技术能力升级。基础电信企业作为电信网络诈骗治理工作的核心责任主体，瞄准“认识到位、执行到位、责任到位、成效到位”的工作目标，在精准管理上用力、在精准打击上发力、在精准服务上着力。在组织管理方面，三家基础电信企业严格落实行业管理部门的相关部署要求，聚集涉及人工智能的诈骗治理工作，积极推进电话实名制、重点业务管理等源头治理；在防范打击方面，先后建设了重点业务管理系统、短信提醒系统、诈骗电话核查处置系统等，汇总分析诈骗样本模板、关键特征等精准信息，利用大数据等技术手段，不断加大对涉及人工智能的电信网络诈骗等违法信息的识别、预警、拦截与处置能力，实现精准打击。在服务引导方面，三家基础电信企业利用人工智能技术分析诈骗新手法、新套路，筛选易受害人群，组织多次定点、定向宣传教育活动，提升民众反诈意识。

二是互联网企业主动作为，加大对人工智能诈骗手法的识别能力研发。近年来，腾讯公司开展人工智能技术在黑色链条中的应用分析研究，为政府部门提供反诈动机分析。百度公司聚焦于诈骗网址安全性识别研究，对高风险链接进行标记、分类乃至屏蔽，将治理电信网络诈骗工作从被动变为主动。科大讯飞股份有限公司建设了人工智

能类诈骗防控技术手段，能够从语言层面自动理解并识别对方意图，实现早期预警，并通过早期预警，对容易受骗的受害人群体进行标记。北京得意音通技术有限责任公司致力于声纹识别技术、语音识别技术、情感识别技术、语言理解技术等打击人工智能诈骗技术研究。

（3）科研院校充分发挥创新主体作用，积极支撑做好行业治理工作。

为支撑做好信息通信行业防范打击电信网络诈骗工作，中国信息通信研究院（以下简称“中国信通院”）、中国互联网协会等行业组织和研究机构凭借自身优势，积极发挥支撑单位创新主体作用，全力支撑防范打击电信网络诈骗相关工作。

一是充分发挥中心作用，体系化推进治理工作。在工信部网络安全管理局的指导下，中国信通院成立了“中国信息通信研究院电信网络诈骗治理支撑与服务中心”，下设秘书处、政策研究组、监管支撑组、技术服务组、专家组等四个二级常设工作小组和 1 个外部专家组，为精准打击基于人工智能的电信网络诈骗提供科研支撑，聚焦监管职责与工作重点，从分析诈骗形势、制定技术标准、建设技术手段、厘清治理思路、研提治理建议、支撑监管工作等方面着手，深入推进诈骗治理支撑工作。目前已支撑工信部开展监管政策制定、监督检查、成效评估、通报举报、责任落实等相关工作。

二是凭借技术优势，智能化推进治理工作。在电话诈骗治理的高位挤压以及人工智能技术的推动作用下，电信网络诈骗向互联网领域，向自动化精准化方向拓展延伸，创新技术管理模式，提升技术治

理能力，是防范治理电信网络诈骗的一柄利刃。在工信部网安局指导下，目前中国信通院正在推动互联网反诈部级系统建设，实现了与部分省系统的对接联动以及数据共享，将诈骗治理工作从电信网向互联网领域拓展延伸，进一步提升对基于人工智能技术的诈骗信息的拦截与处置能力。与此同时，中国信通院会同三大基础电信企业、奇虎、腾讯、电话邦、泰迪熊等 6 家互联网标记企业建立了涵盖号码举报、标记、回收等全环节的数据信息共享机制，在终端用户和电信企业间实现了基于人工智能的诈骗电话的标记信息共享与联动处置。

三是调动各方力量，精准化推进治理工作。人工智能技术使得电信网络诈骗呈现具有低成本、取证难、收益高等特点，促使犯罪活动案件高发频发，为着力推进人工智能时代电信网络诈骗治理工作，相关科研院校积极作为，大力支撑推进开展治理工作。2019 年，中国互联网协会联合中国信通院成功举办“2019 年防范治理电信网络诈骗论坛”，组织 11 家重点互联网企业签订“防范治理电信网络诈骗责任书”，遴选 30 项具有行业示范价值的创新实践应用项目，形成并印发《防范治理电信网络诈骗创新实践示范项目应用汇编》，推动基于人工智能的技术防范体系研究，促进诈骗治理工作实现以“智”提“质”。同时，中国互联网协会积极发挥行业自律平台作用，将涉诈 APP、公众号、网站等纳入用户投诉举报渠道，扩大基于人工智能的诈骗信息投诉受理渠道，完善电信网络诈骗用户投诉受理处置机制。

四、人工智能背景下治理工作面临的难点与问题

在工信部的统筹部署和全行业的共同努力下，信息通信行业电信网络诈骗源头治理和综合治理不断向前推进，已经取得了阶段性的明显效果。但是面对人工智能时代电信网络诈骗的高隐蔽性、高危害性等特点，我国防范治理工作仍然存在着一定的难点与问题，法律法规尚需完善、行业管理范围仍需延伸、技术反制能力亟待提高、社会宣传引导和协同治理的力度和程度还有进一步强化和提升空间。

（一）在法律法规方面，一是刑事打击困难，针对电信网络诈骗持续高发的形势，公安机关在总结诈骗案件态势时指出：“目前诈骗犯罪成本低、风险低、收益高。犯罪分子流窜程度加剧、地域性犯罪突出、职业化趋势明显，传统犯罪与互联网犯罪高度融合，团伙构成、作案手段更加复杂，隐蔽性更强，侦查打击的难度更大。”随着人工智能技术在电信网络诈骗领域的不断应用，不法分子实施诈骗的精准性、迷惑性、隐蔽性大大加强，直接导致公安和检察机关的侦查打击难度进一步加大，人工智诈骗治理工作所面临的“侦查破案难、电子证据调取难、认定处理难”等一系列问题更加突出。**二是**法律适用难，电信网络诈骗已经呈现出明显的链条化、平台化作案特征。例如上游黑产提供各种基于人工智能技术的专用设备工具、算法，为下游电信网络诈骗源源不断地提供技术支持，危害巨大。但是在这个产业链中，不法分子利用互联网以提供服务的方式分工配合，使得诈骗案件发生时，位于上游的不法分子完全可以以自己不知情来规避法律风险。这种法律依据的不明确及其导致的上游犯罪节点定责难，是当

前电信网络诈骗治理不能回避一个重要问题。

（二）在行业管理方面，一是目前以举报通报、现场检查、年度考核等多种形式构建组成的电信网络诈骗企业责任落实督导体系在电信行业已经比较完善和细化，但是在人工智能技术和业务快速发展的互联网等相关领域的督导执行力度仍相对较弱。根据行业和技术特点，将已有责任落实督导体系向互联网行业合理延伸拓展是当前行业管理需要研究的重要问题。**二是**数据隐私泄露已经成为“精准诈骗”的源头之一。虽然现有的用户个人信息保护、数据安全等行业管理基础性制度已经具备，但在具体的企业责任落实方面，特别是对人工智能应用在数据采集、传输、存储、使用、共享等各环节的隐私保护、数据安全等方面的监督落实力度仍需加强。面对人工智能对个人隐私保护和数据安全带来的挑战，督促行业切实加强数据安全保护是当前需要重点解决的问题之一。**三是**对当前涉诈风险较大的基于人工智能的群呼（GoIP、SIMBANK）网关类设备的监管政策和管理方式仍需确定。当前 GOIP 类设备可以从各种渠道轻易获得，生产、销售、使用环节监管尚未形成合力。就目前情况来看，这类设备广泛应用于网络黑灰产，对网络违法犯罪起到了推波助澜作用，因此对此类设备的监管方式需要进一步研究确定。

（三）在技术利用方面，一是人工智能技术的不当应用大幅提高了不法分子获取个人隐私信息，实施精准诈骗的效率。人工智能技术和算法在个人隐私信息数据的爬取、清洗和分析等环节更加高效，成本更加低廉。不法分子可以通过技术的关联分析挖掘，在公共

数据、匿名化数据中推导出精准的个人隐私信息。**二是**人工智能技术被不当利用的门槛不断降低。在软件和算法方面，当前大量被不当利用的人工智能技术均以开源方式对外公开发布，不法分子可以通过互联网开源社区直接获取源代码并进行修改使用；此外，以“语音合成”、“图像识别”等业务为例，有大量人工智能技术被封装成数据接口供使用，不法分子可以不了解底层复杂算法直接调用接口。在硬件设备方面，GoIP 设备以其远程化、智能化的特点，成为不法分子实施诈骗、逃避追查的重要方式。**三是**基于人工智能的仿冒和识别技术发展不均衡。以 ZAO、appFace 等应用为代表，基于人工智能深度伪造的换脸换声工具及应用日趋完善，肉眼识别仿冒视频极为困难，最新技术甚至可以在不需要事先准备各种角度的人脸照片的前提下，实现从视频到视频的直接变脸。相较于深度伪造应用的快速发展，仿冒识别技术研发难度更高，且受限于仿冒数据集样本不足等多种原因，其发展速度相对较慢。**四是**针对人工智能为防范治理工作带来的的新风险，已有技术平台和策略应对能力尚显不足，在对电信网络诈骗实现事前有效防范和事中事后精准溯源等方面还需深入探索。

（四）在宣传引导方面，一是关于电信网络诈骗风险提示等的相关宣传内容仍需增加。目前面向大众的反诈宣传提醒以较为传统的电信网络诈骗手法为主，基于人工智能的诈骗的方式方法和应对策略等相关宣传内容较少，一定程度上造成社会普通民众尤其是学生、老人等重点人群对此类诈骗的特点认知和防范意识不足。**二是**关于个人信息安全保护的宣传力度尚需加强。当前很多用户对人工智能时代个

人隐私泄露的危害仍然认识不够，面对一些应用的霸王隐私条款，往往选择以隐私换取效率和方便。与此同时，学生，农村留守老人等重点群体的个人隐私保护意识较为淡薄，仍存在为了蝇头小利出卖个人信息甚至身份证件供不法分子使用的不良现象。

（五）在协同治理方面，一是随着电信网络诈骗治理逐步进入“深水区”，过往基于单一数据源建立的相关技术防范系统，已不能形成良好的协同效应，治理效果逐步减弱。相关部门与企业、行业组织虽然建立了高效的会商和联动处置机制，但在跨行业、跨部门数据融合共享方面尚不充分，各治理主体的反诈系统仍相对分散，难以对各方数据进行及时的综合分析，一定程度上制约了整体治理效能的持续提升。**二是**部分企业自律意识不足，在用户不知情的情况下，过度收集或使用个人信息。从 App 专项治理工作组通报相关的情况看，包括一些主流 APP 在内的大量应用均存在违规收集使用用户个人信息的问题。**三是**跨国协同合作仍需加强。人工智能技术的应用使诈骗分子实施远程精准诈骗的能力越来越强。为增强隐蔽性，降低风险，大量诈骗团伙越来越倾向于聚集在境外实施诈骗。跨国、跨境电信网络诈骗的识别、拦截、追踪和定责必须在国际协同合作的框架下解决，在已有基础上进一步推动国际各方协同治理已经成为我国电信网络诈骗治理需要解决的重要问题。

五、人工智能背景下电信网络诈骗治理的建议措施

随着防范治理工作的不断向前推进，电信网络诈骗攻防对抗逐步

升级，诈骗手法和方式不断向智能化、精准化、链条化方向发展，新问题新情况不断出现，反诈工作的艰巨性、持久性、复杂性仍然没有改变。面对人工智能时代下的电信网络诈骗的治理问题，需坚持发展与治理并重的思路，划定法律红线，加强行业监管，大力推进技术反制手段建设和社会宣传引导等工作，促进各方协同共治，全面提升我国人工智能时代的电信网络诈骗源头治理和综合治理能力，有效保障我国数字经济和智能社会的健康稳步发展，维护人民群众财产安全和切身利益。

（一）明晰治理思路，坚持发展与安全并举

一是坚持促进发展和依法管理相统一，既大力培养人工智能等新技术在发展数字经济、改善社会民生、治理电信网络诈骗等方面的落地及应用，又积极利用法律法规、行业监管等多种方式引导人工智能新技术新应用的规范发展，保障个人信息安全，降低技术滥用风险，在发展中逐步探索解决人工智能背景下电信网络诈骗的治理问题。**二是**坚持安全可控和开放创新并重，深化人工智能背景下电信网络诈骗治理研究，实时掌握人工智能新技术所存在的诈骗风险，推动增加人工智能技术透明度，开展技术安全审查，加强基于人工智能的技术防范能力开发和建设；立足于开放环境维护网络安全，搭建国家级电信网络诈骗协同治理创新平台，通过开展国际交流合作和推进社会共同治理等方式促进治理模式创新，在治理过程中促进人工智能等新技术健康发展。

（二）完善法律法规，加大执法与惩戒力度

一是在立法层面，结合当前防范治理电信网络诈骗的实际需求，从源头治理的角度出发，推进人工智能、个人信息保护、数据安全等专项立法工作。在法律上对数据非法采集、个人敏感信息非法交易、深度伪造等涉及人工智能的电信网络诈骗的违法违规行为进行规制，明确不同参与主体在电信网络诈骗案件中相关主体民事责任、刑事责任、行政责任，为电信网络诈骗治理提供基本法律依据。**二是**在执法层面，加快制定基于人工智能实施的电信网络诈骗事件的调查取证方法和规范指引，进一步明确政府相关部门与企业、行业组织联动处置的流程规范。加强安全执法，特别是对数据过度采集、技术资源滥用、侵犯个人隐私、恶意网络攻击、伪造仿冒他人等行为加大执法惩戒力度，促进相关法律和规章有效落地执行。

（三）强化行业监管，推进源头与综合治理

一是加强监督检查。依照国家法律法规和行业规章制度，针对人工智能数据过度采集、技术资源滥用、深度伪造仿冒等涉诈安全风险，通过现场检查、技术监测、社会公众举报监督等多种方式实施督导检查，及时发现和防范涉诈安全隐患，通过约谈、责令整改、行政处罚、通报、公开曝光等工作机制，督促相关主体切实落实安全责任。**二是**强化人工智能数据安全保护。研究出台电信和互联网企业网络数据安全保护指导意见，重点明确人工智能应用在数据采集、传输、存储、使用、共享等各环节的数据安全的范围边界、责任主体和具体要求，

督促企业健全数据安全管理制度，加强安全防护技术手段建设，对情节严重或影响恶劣的数据泄露事件加大处罚力度。**三是**开展人工智能涉诈风险技术安全审查和检测评估。依托行业组织或者第三方机构，构建人工智能涉诈安全风险审查和检测评估平台，制定人工智能产品、应用和服务的涉诈风险审查检测方法和评估指标体系，通过审查-测试-验证-改进的方式强化人工智能产品及应用的数据安全与隐私保护，降低涉诈风险。**四是**对涉诈高风险技术或业务实施专门管理，防止技术滥用。对于有合理应用场景但涉诈安全风险较高的人工智能技术、设备或应用，明确规定其使用主体及适用领域范围，对规定范围之外的使用的予以及时清理，涉及相关企业单位或个人的予以警告处罚。对于没有合理使用场景的涉诈人工智能设备或业务予以坚决打击。**五是**建立健全行业信用体系，充分发挥信息通信市场信誉管理机制对电信和互联网企业的监督约束作用，对监督检查过程中所发现问题拒不整改或整改不力的企业纳入违法不良记录信息库，对于情节严重的，列入电信业务经营不良名单或电信业务经营失信名单，并在全行业实现信息共享。

（四）加快技术研发，提高识别与反制能力

一是加强电信网络诈骗技术反制的理论研究和技术研发。以基金引导和政策鼓励等方式，推动产业界和学术界加快突破深度伪造仿冒研判、智能群呼设备识别等以人工智能为手段的电信网络诈骗反制关键技术。**二是**利用以大数据分析和智能预警算法为基础的人工智能技

术进行电信网络诈骗治理，推动人工智能技术在已有反诈系统中的正向应用。通过将人工智能和大数据分析能力集成进已有的反诈系统中，进一步提高现有系统的预警和分析能力，加强系统研判的及时性和准确性。**三是**加快制定与完善深度伪造、智能群呼设备等高风险人工智能技术的使用标准和规范，通过统一技术标准，为识别涉诈虚假视频声音及通信终端类型提供统一接口和模式。**四是**积极探索将生物识别技术应用于金融支付风险防控领域。通过加强生物识别技术与风险防控结合的理论研究与实践应用，推动解决生物识别技术的准确度及成功率、个人生物信息数据安全等方面的问题，促进生物识别技术在金融支付领域的全面普及，以减少发生在“资金支付转移”环节的欺诈事件。

（五）创新宣传方式，增强防范与安全意识

一是拓展宣传思路，针对电信网络诈骗出现的新特点、新手法，由政府部门组织企业、行业研究机构等主体，充分利用传统媒体及短彩信、微信、微博、短视频、手机 APP 等多种渠道编发最新人工智能电信网络诈骗风险提示和应对方法，并及时向社会公众发布；**二是**组织推动地方行业主管部门及企业深入学校、乡村、社区实地开展防范电信网络诈骗宣传活动，针对老年人、青少年等重点群体，有的放矢的开展面对面宣传工作；**三是**进一步加强电信网络治理创新示范项目的推广应用，鼓励行业针对积极探索利用人工智能等新技术解决电信网络诈骗治理的重点难点问题，对取得较好治理成效和社会效益的项

目进行一定的奖励并推动其在行业内的推广应用。**四是**开展人工智能电信网络诈骗攻防技术大赛，引导企业技术人员、高校学生等群体积极参与电信网络诈骗反制技术的研究，对其中的优秀项目予以资金鼓励和创业扶持，推动其在产业界部署落地。

（六）促进协同治理，深化跨行业与跨国合作

一是强化顶层设计，构建工信、公安、金融等主要相关行业或部门深度协同参与的全链路立体化电信网络诈骗综合治理体系，加强电信、互联网、资金、公安等各方数据汇聚分析，并在此基础上对电信网络诈骗进行综合研判，进而在信息获取、脚本编制、通信联络、转账汇款等电信网络诈骗的关键环节建立完善跨行业、跨部门的联合防范治理机制，共同提升基于人工智能的电信网络诈骗的发现识别、打击治理的能力。**二是**进一步畅通举报渠道，建立国家统一的电信网络诈骗举报窗口，完善电话、短信、微信、手机 APP 等多种用户举报方式，通过有奖举报，评选先进个人等多种方式激励广大用户参与防范治理电信网络诈骗工作，推进全民共同治理。**三是**加强行业自律制度建设，组织针对深度伪造等诈骗风险较大的人工智能业务制定专门性的行业自律公约，指导企业在不明确技术泛滥后果的前提下，合理地释放技术成果。**四是**加强并拓宽同美国、欧盟以及东南亚相关国家等重点国家和地区之间的在人工智能、电信网络诈骗防范治理等领域的合作，推动在国际框架下制定防范治理标准和规则，建立国际层面的电信网络诈骗预警、防范及联合惩处机制。

致　谢

感谢工业和信息化部网络安全管理局在白皮书研究和撰写过程中的全面指导。

感谢腾讯守护者计划、腾讯安全战略研究中心和腾讯110，积极贡献其在电信网络诈骗防范治理和人工智能技术对抗方面的研究积累和实战能力，参与白皮书的研究发起和内容编撰。

感谢三六零安全科技股份有限公司、阿里巴巴（中国）有限公司、蚂蚁金服集团等单位的大力支持。

CAICT 中国信通院

中国信息通信研究院

地址：北京市海淀区花园北路 52 号

邮政编码：100191

联系电话：010-62305900

传真：010-62300264

网址：www.caict.ac.cn