汉坤法律评述
2023 年 2 月 21 日

HANKUN
汉 坤 律 师 事 务 所
Han Kun Law Offices

北京 | 上海 | 深圳 | 海口 | 武汉 | 香港

探析生成式 AI 系列（一） — ChatGPT 新浪潮下的算法监管

作者：卢在光 | 李潜 | 刘佳艺 | 李茹娜

2022 年 11 月 30 日，一款由 Open AI 公司开发的人工智能聊天机器人 ChatGPT 正式上线，上线后仅用 5 天就产生了百万用户，而短短 2 个月后其用户规模更是突破了 1 亿大关，不仅迅速引起了美国科技和投资界的注意，也在中国引发广泛讨论。受相关概念影响，A 股多家 ChatGPT 概念股随之应声跟涨。伴随着 ChatGPT 人工智能问答技术热度，多位市场人士预测，今年国内的 AIGC 领域或许会借着这股春风，迎来飞速发展的新契机。

AIGC（AI-Generated Content，即人工智能生成内容）是人工智能理解技术的一个发展方向，作为 Open AI 推出的第 3.5 代模型，ChatGPT 是一款基于互联网可用数据进行训练的文本生成深度学习模型，ChatGPT 可以通过从数百万个网站收集信息，以对话式、人性化的方式生成独特的答案，为用户提供信息查询、日常聊天、文章撰写、程序编写等多种服务。

根据调研机构 CB Insights 的调查报告，2022 年 AIGC 领域初创公司完成 110 笔创投交易，融资总额超 26 亿美元。尽管在技术上实现较大突破，但目前生成式 AI 赛道在内容质量、数据来源、算法技术、行政监管、道德伦理、价值偏见等方面均面临着局限与挑战。本文作为“探析生成式 AI”系列的首篇，拟对 AIGC 和 ChatGPT 的底层基础设施之一 — 算法技术（尤其是生成合成类算法技术）在中国法项下的监管态势进行简要分析与总结，以供读者参考。

一、鼓励、引导与监管并存 — 我国算法监管机制渐趋成熟

自党的十八大以来，我国数字经济以加速度态势发展，信息基础设施与信息技术产业体系建设成效显著。就人工智能而言，国务院、工业和信息化部（“**工信部**”）、科学技术部等主管部门相继出台了《新一代人工智能发展规划》《促进新一代人工智能产业发展三年行动计划（2018-2020 年）》《国家新一代人工智能创新发展试验区建设工作指引》《关于加快场景创新以人工智能高水平应用促进经济高质量发展的指导意见》《关于支持建设新一代人工智能示范应用场景的通知》等一系列政策文件支持人工智能产业发展。党的二十大报告也强调我国要加快发展数字经济，推动战略性新兴产业融合集群发展，构建新一代信息技术、人工智能等一批新的增长引擎。但任何新兴产业的健康发展都离不开鼓励、引导和监管的共同作用。伴随着数字经济快速发展，“数字守门人”的监管概念也逐渐付诸实践，近几年全国人大常委会在年度立法工作计划均提出要加快人工智能、大数据、云计算等领域的立法步伐。

在 2021 年之前，算法监管相关规定还主要散见于有关人工智能、电子商务、反不正当竞争、个人信息保护等相关方面立法项下，包括《电子商务法》《反不正当竞争法》《APP 违法违规收集使用个人信息行为认

定方法》《网络信息内容生态治理规定》等。2021 年 9 月 17 日，国家互联网信息办公室（“**国家网信办**”）等九部门联合发布《关于加强互联网信息服务算法综合治理的指导意见》，提出将“利用三年左右时间，逐步建立治理机制健全、监管体系完善、算法生态规范的算法安全综合治理格局”的目标。随后，国家网信办联合工信部、公安部与国家市场监督管理总局在 2021 年 12 月 31 日发布了《互联网信息服务算法推荐管理规定》（“**《算法推荐管理规定》**”），并于 2022 年 3 月 1 日起正式实施。在 2022 年 4 月，各级网信主管部门在国家网信办的牵头下开展“清朗·2022 年算法综合治理”专项行动，从组织自查自纠、开展现场检查、督促算法备案、压实主体责任和限期问题整改五个方面推动《算法推荐管理规定》落地以及算法综合治理工作的常态化和规范化。至此，我国在算法领域的监管进入新篇章。

此后，国家网信办与工信部、公安部在历经近一年的征求意见后于 2022 年 11 月 25 日发布了《互联网信息服务深度合成管理规定》（“**《深度合成管理规定》**”），并于 2023 年 1 月 10 日正式实施。《深度合成管理规定》在《算法推荐管理规定》等相关规定的基础上，专门对近两年以突飞猛进之势加速发展的深度合成技术进一步作出了针对性合规指引。随着上述规定相继出台，我国对算法技术的引导与监管逐步强化，并形成了“各级网信与电信、公安、市场监管等主管部门协同监管+行业协会强化政策指导与自律管理+落实企业端主体责任”的整体监管机制。

二、算法监管体系的重点 — 算法备案与算法安全评估

根据《算法推荐管理规定》《深度合成管理规定》，“应用算法推荐技术”指利用生成合成类、个性化推送类、排序精选类、检索过滤类、调度决策类等算法技术向用户提供信息。其中，“深度合成技术”是指利用深度学习、虚拟现实等生成合成类算法制作文本、图像、音频、视频、虚拟场景等网络信息的技术。而最近热度颇高的 ChatGPT 人工智能问答技术主要对应生成合成类算法。因此，未来拟从事 ChatGPT 相关业务的企业需遵守应用算法推荐技术的监管要求。

根据《算法推荐管理规定》《深度合成管理规定》以及其他法规的相关规定，我国对应用算法推荐技术的监管已涵盖了算法服务规范与安全义务（包括算法机制机理定期审核与评估、科技伦理审查、用户注册、内容管理与审核、生成合成信息显著标识义务、数据安全和个人信息保护、网络日志留存、反电信网络诈骗、算法备案与安全评估、安全事件应急处置与技术保障措施、算法推荐服务规则、反垄断与反不正当竞争等）与用户权益保护义务（包括算法透明性、保障用户的算法知情权/选择权、禁止算法歧视/过度推荐/非法操纵/诱导用户沉迷和过度消费、劳动者/未成年人/老年人/消费者权益保护机制等）等方面。我们在此不再逐一展开分析，以下主要以问答形式对广受关注的算法备案与算法安全评估中的合规热点问题进行梳理和分析，并对监管实践作出一些经验分享。

（一）算法备案

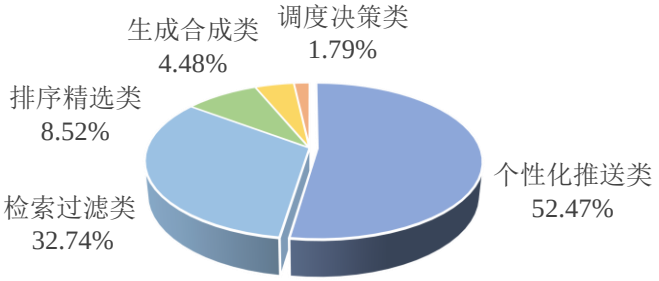
1. 哪些企业需要进行算法备案？

在中国境内提供算法推荐服务的具有舆论属性或者社会动员能力的算法推荐服务提供者（包括深度合成服务提供者、深度合成服务技术支持者）。需要注意的是，算法备案实行属地管辖，也即境外主体在中国境内提供算法推荐服务且达到备案要求的也需要履行算法备案义务。

根据互联网信息服务算法备案系统公示的境内互联网信息服务算法备案清单，截至 2023 年 1 月，完成互联网信息服务算法备案的算法共计 223 项，完成备案的算法类型主要分布如下图所示[1]。通过前

[1] 数据来源参见链接：https://beian.cac.gov.cn。

述备案清单，不难发现各主要互联网头部企业已进行了相关算法备案，备案应用产品以业内知名网站、APP、小程序为主，同一主体名下完成备案的算法最高达 13 项。



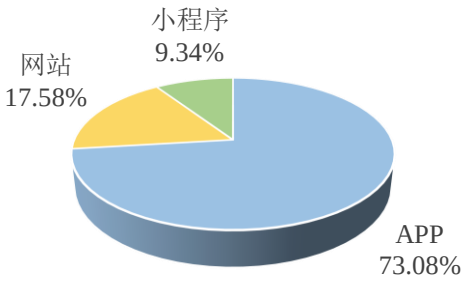
类型	范围
个性化推送类	利用用户属性数据或用户行为数据实现信息个性化分发的算法，常见应用场景比如个性化广告、图文/音视频/资讯/商品/服务推荐、信息流推送等。
检索过滤类	包括检索算法和过滤法，其中检索算法指按照输入条件或检索需求匹配相应网络信息内容的算法，过滤算法指按照给定条件识别并筛选相应网络信息内容的算法，常见应用场景比如搜索引擎/搜索功能、内容检索等。
排序精选类	以客观因素或主观因素为依据，设置、调整网络信息内容排列顺序的算法，常见应用场景比如热搜、榜门榜单/话题榜单/音视频榜单、商品/新闻排行榜/推荐榜等。
生成合成类	自动或辅助生成、编辑文本、图像、语音、视频等网络信息内容的算法，常见应用场景比如语音/文字/图片识别、自动写稿/新闻合成等。
调度决策类	自动或辅助生成供需匹配、供需调节、路径规划等调度决策结果，或提供调度决策依据的算法，常见应用场景比如导航路线规划、外卖配送/网约车订单调度等。

2. 如何判断是否“具有舆论属性或者社会动员能力”？

根据我们与主管部门的近期咨询与沟通以及相关项目经验，目前关于“具有舆论属性或社会动员能力”尚无进一步细化的评判标准，主要判断依据仍为国家网信办与公安部于 2018 年 11 月发布的《具有舆论属性或社会动员能力的互联网信息服务安全评估规定》（“**《安全评估规定》**”）。根据《安全评估规定》，具有舆论属性或社会动员能力的互联网信息服务包括：（一）开办**论坛、博客、微博客、聊天室、通讯群组、公众账号、短视频、网络直播、信息分享、小程序等**信息服务**或者附设相应功能**；（二）开办**提供公众舆论表达渠道或者具有发动社会公众从事特定活动能力**的其他互联网信息服务。从字面规定来看，该等定义范围较为宽泛，大部分互联网信息服务均可能落入该定义范围。根据我们的了解，目前算法备案的监管实践主要是企业对照相关规定与自身实际情况自主判断、依法依规备案即可，但不排除主管部门主动联系和要求相关企业进行算法备案的可能。若相关企业不确定是否需要进行算法备案，建议企业结合自身使用的算法技术类型、提供的服务类型/内容类别、用户规模、算法推荐技术处理的数据重要程度、对用户行为的干预程度等因素，参考境内互联网信息服务算法备案清单项下已备案相同或近似类型或业务的企业向主管网信部门进行进一步咨询确认以作出判断。

3. 使用同一算法的不同产品类型可否一起备案？

使用同一算法的不同产品类型可以一起备案，目前算法备案适用的产品类型包括 APP、网站、小程序，截至 2023 年 1 月完成算法备案的产品类型主要分布如下图所示[2]。算法备案以算法为核心、以算法类型为基础进行分类管理，备案主体就不同算法类型项下的不同算法分别备案。也就是说，若同一应用产品涉及不同算法，则需要就各算法分别进行备案；而若同一算法应用于多种产品类型（比如 APP、网站、小程序），则可同时备案（在备案时添加相应产品信息即可）。根据境内互联网信息服务算法备案公示清单，可以看到一些知名互联网企业已经就其 APP 所使用的个性化推送类、检索过滤类、排序精选类、生成合成类和/或调度决策类算法分别进行了备案，并对所涉个性化推送类、检索过滤类和/或生成合成类项下不同算法分别进行了备案。



4. 算法披露义务与算法信息公示是否存在风险？

很多企业可能会对算法披露义务存在商业与技术秘密保护方面的担忧，但实际上算法披露义务并不要求披露核心商业与技术秘密。通过目前互联网信息服务算法备案系统的公示记录，也可以发现公示内容主要包括算法名称、算法类型、算法基本原理、算法运行机制、算法应用场景、算法目的意图和其他算法公示情况（选填）等基础属性信息，而不会涉及备案主体的落实算法安全主体责任基本情况、算法安全自评估报告以及详细属性信息（如算法数据、算法模型、算法策略、算法风险和防范机制等）。

5. 备案后义务有哪些？

- 公示信息链接：完成备案的算法推荐服务提供者应当在其对外提供服务的网站、应用程序等的显著位置标明其备案编号并提供公示信息链接。
- 变更备案：算法推荐服务提供者的备案信息发生变更的，应当在变更之日起 10 个工作日内办理变更手续。
- 注销备案：算法推荐服务提供者终止服务的，应当在终止服务之日起 20 个工作日内办理注销备案手续，并作出妥善安排。

（二）算法推荐服务（包括深度合成类服务）安全评估

1. 哪些企业需要进行算法推荐服务安全评估？

在中国境内提供算法推荐服务的具有舆论属性或者社会动员能力（相应评判标准分析参见以上算法备案部分）的算法推荐服务提供者均需进行安全评估。就深度合成技术而言，包括（1）在中国境内

[2] 数据来源参见链接：https://beian.cac.gov.cn。

开发上线具有舆论属性或者社会动员能力的新产品、新应用、新功能的深度合成服务提供者；（2）在中国境内提供具有（i）生成或者编辑人脸、人声等生物识别信息的；（ii）生成或者编辑可能涉及国家安全、国家形象、国家利益和社会公共利益的特殊物体、场景等非生物识别信息的功能的模型、模板等工具的深度合成服务提供者和技术支持者。如同算法备案，算法推荐服务安全评估要求亦实行属地管辖原则。

2. 算法推荐服务安全评估与其他网络安全与数据合规领域相关法律法规项下的安全评估之间存在何种关系？

根据我们向主管部门的近期咨询，《算法推荐管理规定》《深度合成管理规定》项下的算法推荐服务安全评估与其他网络安全与数据合规领域相关法律法规项下的安全评估以及算法备案过程中备案主体需提交的算法安全自评估报告是相互独立的监管环节，但目前就算法推荐服务的安全评估尚未有进一步的实操规定，也尚不存在通过主管部门进行此项安全评估的渠道。因此，为满足《算法推荐管理规定》《深度合成管理规定》项下算法推荐服务安全评估要求，企业可考虑参照《算法推荐管理规定》《深度合成管理规定》项下相关要求定期自行或者委托专业机构开展安全评估并留存相关支撑性资料。

3. 算法推荐服务安全评估的要求是什么？

由于《算法推荐管理规定》项下算法安全评估实操细则尚不明确，通过对上市案例的检索，我们注意到也存在部分拟发行人实际上按照《安全评估规定》通过全国互联网安全管理服务平台向主管网信部门和公安机关提交安全评估报告，以满足《算法推荐管理规定》项下算法推荐服务安全评估要求。但如前所述，我们理解该两项安全评估应该属于相互独立的监管要求。

如同算法备案，目前算法推荐服务安全评估的监管实践亦主要是企业对照相关规定与自身实际情况自主判断、依法依规进行安全评估，但不排除主管部门主动联系和要求相关企业进行算法推荐服务安全评估的可能。因此，若相关企业对是否需要进行算法推荐服务/深度合成服务安全评估以及如何进行该等安全评估存在疑问，建议进一步咨询主管部门进行具体确认。

另外，需要注意的是，为了强化深度合成技术监管，《深度合成管理规定》还进一步要求“互联网应用商店等应用程序分发平台”应当落实上架审核、日常管理、应急处置等安全管理责任，**核验深度合成类应用程序的安全评估、备案等情况；对违反国家有关规定的，应当及时采取不予上架、警示、暂停服务或者下架等处置措施**，国家对深度合成技术实施常态化监管的态度也可从中窥知一二。

三、算法合规的上市审核关注要点

随着《算法推荐管理规定》等监管新规的出台，算法推荐服务合规也逐渐成为上市监管机构对主营业务涉及算法推荐服务的发行人的问询关注点之一。经梳理近期科创板及创业板上市申报案例，结合上市监管部门的审核关注要点及发行人的应对方案，我们注意到上市监管机构一般会要求发行人说明信息推送、交易推荐、提供算法推荐服务、用户权益保护等内容是否符合《算法推荐管理规定》等相关规定，发行人需对其算法安全义务、算法服务规范、用户权益保护、算法备案与安全评估等方面的合规情况进行论述，如发行人历史上存在算法推荐服务不合规记录，相关整改情况也将被上市监管机构问询关注。

四、结语

随着 ChatGPT 的横空出世，AIGC 行业正在以超乎想象的速度发展，加上深度合成技术因其自身属性可能存在或带来的各种法律风险，深度合成类算法推荐服务与应用程序必将成为各级网信部门的监管重点

之一。自《算法推荐管理规定》于 2021 年底生效以及互联网信息服务算法备案系统自 2022 年 3 月 1 日上线以来，在不到一年的时间里，相当数量的互联网企业逐步完成各类算法备案，算法备案监管规定的落地与执行情况已渐趋成熟。尽管目前算法推荐服务安全评估相关实操要求还有待主管部门出台进一步实施细则或指引，考虑到网信部门监管口径与态度趋严，建议相关企业也提高对算法推荐服务合规（尤其是算法备案与安全评估）的重视程度。

特别声明

汉坤律师事务所编写《汉坤法律评述》的目的仅为帮助客户及时了解中国或其他相关司法管辖区法律及实务的最新动态和发展，仅供参考，不应被视为任何意义上的法律意见或法律依据。

如您对本期《汉坤法律评述》内容有任何问题或建议，请与汉坤律师事务所以下人员联系：

卢在光

电话： +86 10 8525 4623
Email： zaiguang.lu@hankunlaw.com