

狂飙的 ChatGPT，合规“缰绳”何在？

作者：潘永建 | 朱晓阳 | 王雪莹 | 吴若蕸

一. 引言

近期，ChatGPT 的爆火让人工智能的话题在全球范围内再次引发高度关注。据其研发公司 OpenAI 介绍，“ChatGPT 是一种对话式 AI，可运用自然语言连续对话、回答问题并挑战不正确的假设。”作为人工智能技术驱动的自然语言处理工具，它的亮点在于通过学习和理解人类的语言以及聊天的上下文进行互动、交流对话，甚至能完成诸如撰写邮件、视频脚本、文案、翻译、代码等任务。

但 OpenAI 远不是唯一的入局者。ChatGPT 火热之际，微软已宣布将正式接入 ChatGPT 技术，以支持必应搜索和 Edge 浏览器；而 Google、百度等传统搜索引擎类企业也火速入局类 ChatGPT 领域，并计划推出聊天机器人 Bard、文心一言等竞品产品，未来的竞争是否将更加激烈也犹未可知。而随着 ChatGPT、DALL-E 等利用深度合成技术的产品逐一问世，面对未来该领域的广阔发展前景，其也引发了一系列的合规与隐私安全方面的担忧。

ChatGPT 类产品虽然非常复杂，但其原理却并非新鲜事物。ChatGPT 中 GPT 的含义为 Generative Pre-trained Transformer，简而言之是通过预先输入海量的文本内容，使得 ChatGPT 能够学习人类语言的特征和结构，对输出结果进行校正和优化，从而实现人机之间的自然、流畅交流。而其背后所依赖的核心技术，便是深度合成技术，记忆力好的朋友可能还记得几年前闹得沸沸扬扬的换脸技术 deepfake 也正是基于深度合成技术。

.....
如您需要了解我们的出版物，
请联系：

Publication@llinkslaw.com

深度合成技术,是指利用深度学习、虚拟现实等生成类算法制作文本、图像、音频、视频、虚拟场景等网络信息的技术。我国是最先立法规范深度合成技术的国家之一:国家网信办、工信部和公安部于2022年11月25日联合发布了部门规章——《互联网信息服务深度合成管理规定》(“《深度合成规定》”),该规定已于2023年1月10日起正式实施。除《深度合成新规》外,我国其他网络安全与数据合规领域的法律,如《网络安全法》《数据安全法》《个人信息保护法》等均可能适用于 ChatGPT 类产品,因此该等产品在中国法律下可能面临诸多合规风险和挑战。

由此,我们希望通过本文,探讨 ChatGPT 类产品在中国的数据合规与个人信息保护的体系下存在的合规风险,供读者借鉴思考。

二. ChatGPT 们的潜在法律风险

1. 数据合规风险

1.1 收集训练数据阶段:数据来源的合规风险

我们首先关注的,是 ChatGPT 类产品的训练数据库数据来源的合法合规性问题。根据《深度合成新规》第十四条规定,深度合成服务提供者和技术支持者应当加强训练数据管理,采取必要措施保障训练数据安全;训练数据包含个人信息的,应当遵守个人信息保护的有关规定;其中还特别强调了“深度合成服务提供者和技术支持者提供人脸、人声等生物识别信息编辑功能的,应当提示深度合成服务使用者依法告知被编辑的个人,并取得其单独同意。”

而目前来看,该类产品收集的训练数据量绝对不容小觑,以 ChatGPT 为例,从 GPT-3(ChatGPT 前代)公开发表的论文可知,其用于干预训练的文本已达到了 45TB。而在中文数据上,据查询,全球最大的语料库是 WuDaoCorpora,约有 3TB 的中文语料¹。

首先,ChatGPT 收集的训练数据难以避免对个人信息的收集,尤其是公开渠道的个人信息。根据《个人信息保护法》(“《个保法》”)的相关规定,个人信息处理者需取得个人信息主体的同意或具备其他合法性基础,方可处理其个人信息。然而,在 OpenAI 公布的隐私政策中²,只描述了个人访问 OpenAI 及其关联公司的网站或使用其相关产品和服务时对个人信息的处理,并未提及其收集训练数据环节的个人信息的保护政策。而在实操中,数据训练环节想要取得个人同意几乎不可行。如果个人信息的数据源仅限于公开渠道,那么 ChatGPT 类产品公司可能可以援引“在合理的范围内处理个人自行公开或者其他已经合法公开的个人信息”或兜底条款来论证收集使用个人信息的合法性,但如何证明训练 AI 属于“合理的范围内”仍将是一大合规难题。

¹ 《ChatGPT 的发展历程、能力来源和复现它的关键之处》, https://blog.csdn.net/qq_35082030/article/details/128988965

² Open AI Privacy Policy, <https://openai.com/privacy/>

另一方面, ChatGPT 类产品收集个人信息的方式应注意是否违反“最小范围”原则。《个保法》对于企业收集个人信息提出了“限于实现处理目的的最小范围”的要求。与 GDPR 数据最小化规则的规定类似, 这要求 ChatGPT 类产品所收集的数据仅限于最低限度、且仅提供给此项产品及其服务使用³。

OpenAI 的隐私政策指出, 所有收集的数据都将仅用于合同指定目的。尽管如此, 该隐私政策并未说明其用于合同指定目的的数据是否包括存储在其模型中的数据。同时, 收集如此海量的数据的同时, 其中所涉的个人信息是否真正仅用于 ChatGPT 也未可知。

此外, GPT 模型的预训练模型挖掘大量受著作权法保护的第三方文字作品内容投入数据库作为训练素材, 因而 ChatGPT 类产品的训练数据库需要注意避免侵害第三方的著作权。以复制使用文本数据为例, 将触及著作权侵权、挖掘行为授权、二次创作许可等问题。ChatGPT 类产品可能会主张对上述版权作品构成合理使用, 但目前中国对合理使用仍采取封闭模式规制, 并未承认文本挖掘构成合理使用, 因此文本挖掘可能侵犯文字作品的复制权、改编权、署名权等著作权⁴。此外, 从二次创作许可来看, 如果其生成物是在保留他人作品的基础上进行创造性变动, 可能会侵犯他人的改编权; 如果对数据进行个性化选择和编排, 则可能会侵犯他人的汇编权⁵。

最后, 即使数据中不涉及个人信息, 或者作为知识产权被保护的作品, 其本身也是一种竞争性财产权益。《中华人民共和国反不正当竞争法》的最新修订草案征求意见稿中指出, 经营者不得以不正当的方式获取或者使用其他经营者的商业数据, 如违反约定或者合理、正当的数据抓取协议, 获取和使用他人的商业数据(见 18 条规定)。在淘宝诉美景一案中⁶, 杭州市中级人民法院也认为, 由于网络数据产品不同于网络原始数据, 其数据内容经过网络运营者大量的智力劳动成果投入, “网络运营者对于其开发的数据产品应享有独立的财产性权益”。由此, 针对 ChatGPT 类产品的训练数据, 虽然互联网上存在大量可以自由访问的信息, 但从诸如具有禁止第三方爬取数据条款的网站收集海量数据, 该等数据很有可能被认定为竞争性财产权益, 因此如何避免对该等数据的收集违反《反不正当竞争法》的相关规定也将是 ChatGPT 类产品的一大课题。

1.2 使用阶段: 信息安全风险及个人信息主体行权问题

而从使用的角度而言, 用户在 ChatGPT 类产品的界面中输入相关请求时, 可能会披露个人相关信息、企业的业务数据等, 甚至可能涉及公司的商业秘密等机密信息。而 ChatGPT 使用的训练方法表明, 用户使用过程中的输入和交互信息可能会用于其持续迭代训练。

³ Unveiling the Crucial 5 GDPR Obstacles of ChatGPT That Can't Be Ignored, <https://www.fieldfisher.com/en/insights/unveiling-the-crucial-5-gdpr-obstacles-of-chatgpt>

⁴ 《建议著作权法修法将文本挖掘行为纳入合理使用的情形》, <https://mp.weixin.qq.com/s/VTGb265WQXyh-dqpzbDczw>

⁵ 《ChatGPT 爆火, 带来哪些版权问题?》, https://mp.weixin.qq.com/s/5JfO2dAzNrcOG1pL_s2SBQ

⁶ 淘宝(中国)软件有限公司诉安徽美景信息科技有限公司不正当竞争纠纷案, <https://www.chinacourt.org/article/detail/2019/10/id/4591196.shtml>

尽管 ChatGPT 声称会采取适当的安全保障措施(如匿名化、加密等技术)来提高数据安全性,相关数据的泄露风险还是得到了一定证实。据报道, Amazon 内部的通信工具 Slack 泄露的一条信息显示,亚马逊的公司律师称,公司在 ChatGPT 生成的内容中发现了与公司机密“非常相似”的文本,可能是由于 Amazon 员工在使用 ChatGPT 生成代码和文本时输入了公司的机密数据⁷。而这些信息的外泄不仅可能造成公司商业秘密的直接泄露,也可能被 OpenAI 存储于数据库中并作为训练数据,如遭受黑客攻击,公司将进一步蒙受损失。

另一方面的合规担忧则源自 ChatGPT 是否能满足保障个人信息主体权利的相关要求,尤其是针对个人信息的删除权。虽然其隐私政策中规定了用户对其个人信息的查阅、更正和删除权等,但基于请求该类 AI 系统擦除数据的复杂性⁸,对于深度合成技术提供者能否真正删除个人信息的存在痕迹、满足现有合规要求仍存疑。

2. 内容合规风险

内容方面, ChatGPT 类产品的主要风险则在于可能生成包含违法信息的答复。《深度合成新规》对内容审核提出了较高要求,其中包括禁止深度合成服务提供者、使用者在内的主体利用该服务从事危害国家安全和利益、损害国家形象、侵害社会公共利益、扰乱经济和社会秩序、侵犯他人合法权益等法律、行政法规禁止的活动(如制作、发布、传播假新闻等),且服务提供者应采取技术或者人工方式审核使用者的输入数据和输出结果;另,该新规以及《互联网信息服务管理办法》《网络信息内容生态治理规定》还规定了建立违法和不良信息特征库、违法和不良信息报告与处置、辟谣机制建立、申(投)诉渠道设置及处理等一系列内容合规义务。

因此,落地中国的 ChatGPT 类产品还需关注其是否建立健全了相关内容审核、违法和不良信息报告与处置、辟谣、申(投)诉等一系列机制,特别是审核其输出结果(如问题的答案)是否可能涉及中国法律禁止的内容,如涉嫌危害国家安全、破坏国家统一,或含有暴力恐怖、淫秽色情元素等等。

3. 算法合规问题

《深度合成新规》进一步要求,深度合成服务提供者和技术支持者应当加强技术管理,定期审核、评估、验证生成合成类算法机制机理,具有舆论属性或者社会动员能力的还应按照《互联网信息服务算法推荐管理规定》(“《算法推荐规定》”)的要求履行备案手续等。

ChatGPT 类产品作为基于大规模语言模型的对话系统产品,其应用的生成合成类算法技术毫无疑问属于《算法推荐规定》下的算法推荐技术之一。虽然相关算法合规细则尚未出台,但未来若

⁷ Amazon Warns Employees to Beware of ChatGPT, <https://news.yahoo.com/amazon-warns-employees-beware-chatgpt-174500629.html>

⁸ Does ChatGPT Comply with EU-GDPR Regulations? Investigating the Right to be Forgotten, <https://www.fieldfisher.com/en/insights/does-chatgpt-comply-with-eu-gdpr-regulations-inves>

要实现 ChatGPT 类产品的本土落地，是否能够达到目前相关规定提出的要求仍不可忽视。因此，我们列举了以下合规要求以供参考：

- (1) 审核算法的机制机理，如设置防止用户沉迷、过度消费等机制；
- (2) 应以显著方式告知用户其提供推荐服务的情况，并以适当方式公示其推荐服务的基本原理、目的和主要运行机制等(如利用描述信息、用户画像等进行自动化决策)；
- (3) 应向用户提供不针对其个人特征的选项，如选择或删除用于推荐的用户标签等功能；关闭其推荐算法服务的便捷选项等；
- (4) 如向未成年人提供服务，应当依法履行未成年人网络保护义务，开发适合其使用的模式，不可向其推送可能诱导未成年人模仿不安全行为、违反社会公德、沉迷网络、损害身心健康的信息。

三. 《深度合成新规》重点提示

《深度合成新规》针对深度合成服务提供者、技术支持者、使用者三类主体的各项义务进行了较详细的规定，请参考如下重点提示。当然，对于细节性规则，仍有待监管部门的进一步明确。

义务类型	义务具体内容
深度合成服务提供者	
数据合规	(1)建立管理制度及技术保障措施义务；(2)制定和公开管理规则义务；(3)身份信息认证义务；(4)个人信息保护义务；(5)安全评估义务
内容合规	(1)内容审核义务；(2)建立违法和不良信息特征库义务；(3)违法和不良信息报告与处置义务；(4) 建立辟谣机制义务；(5)申(投)诉渠道设置及处理义务
算法合规	(1)深度合成内容标识义务；(2)算法机制机理审核评估义务；(3)算法备案义务
深度合成服务技术支持者	
数据合规	(1)个人信息保护义务；(2)安全评估义务
算法合规	(1)算法机制机理审核评估义务；(2)算法备案义务
深度合成服务使用者	(1)内容审核义务；(2)配合身份信息认证义务

四. 结语

作为网安与数据合规领域的律师，我们非常乐于见到、尝试并学习类似 ChatGPT 这样的新型产品。基于本文的分析，OpenAI 等 ChatGPT 类产品的开发者的合规之路无疑是任重而道远，而另一方面，ChatGPT 类产品的用户也并不是“百无禁忌”，也必须要在使用该类产品时注意自身的合规风险与义务。如果您对 ChatGPT 类产品的用户的合规风险与义务感兴趣，敬请关注我们的后续文章。

如您希望就相关问题进一步交流，请联系：



潘永建
+86 21 3135 8701
david.pan@llinkslaw.com



朱晓阳
+86 21 3135 8683
nigel.zhu@llinkslaw.com

如您希望就其他问题进一步交流或有其他业务咨询需求，请随时与我们联系: master@llinkslaw.com

上海

上海市银城中路 68 号
时代金融中心 19 楼
T: +86 21 3135 8666
F: +86 21 3135 8600

北京

北京市朝阳区光华东里 8 号
中海广场中楼 30 层
T: +86 10 5081 3888
F: +86 10 5081 3866

深圳

深圳市南山区科苑南路 2666 号
中国华润大厦 18 楼
T: +86 755 3391 7666
F: +86 755 3391 7668

香港

香港中环遮打道 18 号
历山大厦 32 楼 3201 室
T: +852 2592 1978
F: +852 2868 0883

伦敦

1/F, 3 More London Riverside
London SE1 2RE
T: +44 (0)20 3283 4337
D: +44 (0)20 3283 4323



www.llinkslaw.com



Wechat: LlinksLaw

本土化资源 国际化视野

免责声明：

本出版物仅供一般性参考，并无意提供任何法律或其他建议。我们明示不对任何依赖本出版物的任何内容而采取或不采取行动所导致的后果承担责任。我们保留所有对本出版物的权利。

© 通力律师事务所 2023