Open Journal of Legal Science 法学, 2024, 12(5), 3028-3034
Published Online May 2024 in Hans. https://www.hanspub.org/journal/ojls
https://doi.org/10.12677/ojls.2024.125430

Hans汉斯

论ChatGPT背景下个人信息风险的行政规制

蒋显耀

宁波大学法学院，浙江 宁波

收稿日期：2024年3月20日；录用日期：2024年4月11日；发布日期：2024年5月22日

摘 要

本文以OpenAI隐私政策为研究对象，认为生成式人工智能ChatGPT在信息处理上存在过度收集、过度利用、信息泄露风险三个方面问题。借鉴欧盟制定的《通用数据保护条例》，针对政府在个人信息权利损害赔偿、个人信息自决和安全监管三个方面的不足，提出四个针对性的行政监管策略：完善个人信息保护的监督机制、加强相关部门的监督能力、设置个人信息评估与行政咨询机构。

关键词

ChatGPT，个人信息风险，行政规制

On the Administrative Regulation of Personal Information Risk under the Background of ChatGPT

Xianyao Jiang

School of Law, Ningbo University, Ningbo Zhejiang

Received: Mar. 20th, 2024; accepted: Apr. 11th, 2024; published: May 22nd, 2024

Abstract

This article takes the OpenAI privacy policy as the research object and believes that the generative artificial intelligence ChatGPT has three problems in information processing: excessive collection, excessive utilization, and information leakage risks. Drawing on the General Data Protection Regulation formulated by the European Union, four targeted administrative supervision strategies are proposed to address the government's deficiencies in three aspects: personal information rights damage compensation, personal information self-determination, and security supervision: improving the supervision mechanism for personal information protection, strengthening the supervi-

文章引用: 蒋显耀. 论 ChatGPT 背景下个人信息风险的行政规制[J]. 法学, 2024, 12(5): 3028-3034.
DOI: 10.12677/ojls.2024.125430

sion capabilities of relevant departments, and setting up personal information consulting and administrative consulting agencies.

Keywords

ChatGPT, Personal Information Risk, Administrative Regulation

1. 引言

2022 年 11 月，美国 OpenAI 公司发布人工智能聊天机器人 ChatGPT，其使用自然语言处理技术与搜索引擎集成技术，建构大型语言和强化学习微调训练模型，连接大量语料库，通过预训练方法处理大模型序列数据，使其拥有与人类相当的语言理解能力和超越人类文本生成、知识学习能力。依托 ChatGPT 服务平台既能以极其低廉的经济成本与时间成本获取所需信息，又能借此将其发散到平台以供所有用户提取。然而，以海量语料库数据为基础的生成式人工智能在为社会带来积极变革和影响的同时，也引发了诸多个人信息安全风险。

2. ChatGPT 背景下个人信息面临的风险

从 2023 年 OpenAI 发布最新的隐私政策来看，ChatGPT 处理个人信息不局限在各自的注册用户，也包括非用户，其对个人信息的收集、处理和存储从政策到实操都充斥着对信息主体的权益侵害风险。从原宗旨来讲，隐私政策的内容即为 ChatGPT 在收集、使用、存储个人信息方面的行为规则。因此我们以下以 ChatGPT 开发公司 OpenAI 的隐私政策为基础来探究其行为产生的个人信息安全隐患[1]。

2.1. 信息数据过度收集风险

当下，个人信息数据已然是一种战略资源，并且随着技术不断创新发展，甚至成为一项具有公共资源属性的社会财富，导致不规范的收集、使用、出售和共享个人信息现象逐渐加剧，对个人的信息安全形成挑战。据此，《中华人民共和国网络安全法》、《中华人民共和国个人信息保护法》、《生成式人工智能服务管理暂行办法》明确规定了收集和处理个人信息方面基本准则，强调“知情同意”的基本行为准则。当然，为了对个人信息领域的绝对保护而直接禁止人工智能发展，这不仅剥夺了人民对便捷、智能美好生活的追求，也可能因此削弱我国在相关领域的国际竞争力，但生成式人工智能 ChatGPT 问世后，以其独有的不可替代的经济市场价值对知情同意规则发起冲击不容忽视。

其一，“明确告知”规则的消解。有研究估计，ChatGPT 的自然语言模型训练和运行模型需要庞大的算力，训练耗能相当于汽车往返地月[2]。可见，ChatGPT 模型的技术前提就是既有的庞大语料数据库(简称 LLM)，不过到目前为止，OpenAI 公司并没有公布其信息数据的来源。有人认为：“目前 ChatGPT 属于算法黑箱，相关训练数据库是否获得授权存在疑问。”[3]尽管如此，我们通过对 ChatGPT 模型技术基石-LLM 的架构技术可以大致厘清其数据来源：第一，LLM 是通过爬虫技术和 cookie 技术收集网络用户产生的公开信息，并且这种收集是无目的、全面的；第二，LLM 的训练是无监督地自主进行，无需人工介入和标注。在数据获取过程中，ChatGPT 自己也承认尽管已经尽可能地减少侵犯权利的分享，也不能

保证所有使用的数据都经过了原始创作者的许可[4]。LLM 技术特征就决定了如此庞大的信息收集必然会漠视“知情同意规则”；另一方面，获取数据不受人类干预和控制，那么获取非法数据不可避免，训练过程不受人类监督，那么其学习过程本身不可避免侵犯个人信息安全、数据安全等。

其二，“取得同意”规则的消解。根据 OpenAI 最新公布的隐私政策，其收集注册用户的信息包括创建账户信息以及适用 ChatGPT 过程产生的数据信息，大致可以分为三类：内容信息，指的是用户与 ChatGPT 对话过程产生的内容；社交信息，是指它是指用户在其他社交媒体网站上使用基于 ChatGPT 提供的 API 创建的 ChatGPT 页面时产生的信息；技术信息，指的是用户使用 ChatGPT 过程中自动向其发送包括国家、日期、IP 地址等数据。在内容数据上涉及到个人对话的隐私，甚至具有敏感性内容，社交信息与技术信息中也没有对涉及到公共安全信息进行排除，而全部予以择取并默认用户同意。但事实上基于“不同意即走”的用户协议，要使用其提供的智能服务就必须接受其提供的“霸王条款”，明显违反了个人自由意愿。

2.2. 个人信息数据过度利用风险

根据 OpenAI 最新公布的隐私政策，OpenAI 表示，ChatGPT 其对用户个人信息进行处置的目的主要为三个方面：为了更好提供服务；对系统进行管理、维护、改进和分析；开发新的计划和服务以及保护用户和第三方权利、隐私、安全。ChatGPT 对个人信息的使用行为的本身就具有极强的隐蔽性，作为普通的用户是没有能力和精力区追踪 ChatGPT 究竟是如何处理自己的信息，而隐私政策中“提供更好的服务”“开发”等模糊化的词汇就更进一步增加了用户对个人信息的监控。其次，OpenAI 在隐私政策中提到的关于使用个人信息的方式主要是：通过个人信息汇总分析用户的一般行为和特征并将此与第三方共享此类汇总信息；出于研究或统计的目的对个人信息进行匿名化或去识别化处理后无限期地适用这些信息，并且不进行通知。可见，ChatGPT 收集个人信息目的就是通过学习分析出个人行为的特征，虽然其对个人信息进行了匿名化和去识别化处理，但至今未有对“去识别化、匿名化”的效果评价机制，以及利用个人信息汇总提取用户行为一般特征的影响评价机制。

2.3. 信息存储与共享的泄露风险

首先，ChatGPT 并非是本土部署。根据 OpenAI 发布的隐私政策规定，ChatGPT 的国际用户产生的个人信息及相关数据会自动发送到美国的 OpenAI 公司服务器。因此，如此庞大的本土人群信息为他国掌握，不仅仅威胁到个人信息安全，也无形中产生对国家安全的威胁。

其次，隐私政策中提到，向供应商和服务提供商共享、业务转让过程中向交易方披露、对关联公司披露的情况下，不对披露信息涉及的个人进行通知。OpenAI 在隐私政策的格式条款中规定了不另行通知用户的情形，但并没有对因为这些行为导致信息泄露进行救济的规定，也没有规定自身应当对此承担的责任。一方面，OpenAI 通过隐私政策这种格式条款来规避责任并将这种风险转嫁给用户，另一方面正是因为其不承担责任而可能对信息数据的安全维护降低了关注度。并且，如今类似于 OpenAI 的人工智能开发公司大规模数据泄露并不罕见。

最后，隐私政策中提到，国际用户产生的所有信息会自动发送到 ChatGPT 在美国的设施和设备进行存储。其中包含的个人信息既包括用户的名称、使用数据等一般数据，也包括行踪轨迹(IP 地址)、金融账号等个人敏感信息。并且，如此庞大数量的国内个人信息被用于处于美国控制下的 ChatGPT 学习分析以提炼出国内大部门人群的一般行为特征和行为倾向等，因此产生的侵权纠纷必然是涉外纠纷。如此，不仅个人信息权益遭受侵害时难以救济，而且导致的经济利益损害也难以救济，更甚于在发生战争时用于分析国民行为倾向，并应用与战争中，可能威胁到国家整体安危。因此，对于数据的存储可能产生的

风险，我们应当给予更多的关注。

3. ChatGPT 背景下个人信息安全救济困境

3.1. 个人信息权益救济困境

为保护在人工智能迅猛发展时代的个人信息安全，2016 年通过的《中华人民共和国网络安全法》、2020 年通过的《中华人民共和国民法典》、2021 年 6 月通过的《中华人民共和国数据安全法》、2021 年 8 月通过的《中华人民共和国个人信息保护法》、2023 年推出的《生成式人工智能服务管理暂行办法》，一系列法律规范从民事、行政和刑事责任等不同侧面构建出个人信息保护的法律框架。但是，我们不得不面对一个尴尬的事实：刑事责任、行政责任不断强化，个人信息安全的环境却不断恶化[5]。“天下之事，不难于立法，而难于法之必行。”究其原因，主要是法律实施上的原因：

首先，损害赔偿面临损害不能确定的难题。在个人信息侵权的有些情况下，侵权行为导致的财产和身体损失是比较明显和确定的[6]。例如，个人信息泄露可能导致账户被盗窃和财产损失这样的案例。但是，在更多普遍的侵害个人信息案例中，损害并不明显。有的情形下，侵害带来的是骚扰。例如，很多广告、诈骗信息频繁地推送给用户，造成时间成本和心理负担外，还可能使得个人信息容易暴露。有的情形下，侵害带来的是风险。例如，收集个人信息的企业可能存储用户的敏感信息，其中涉及的大量数据可能存在失窃和非法传输的风险。用户暂时未发现任何实际损失，但长期来看，这种风险却可能导致他们遭受更大的经济和精神损失。此外，侵害可能带来纯粹的焦虑感。例如，一些企业可能进行完全合法而合规的个性化推荐，推荐内容与收集到的个人信息高度吻合，但这种过度的个性化反而增加了用户对于被“标签化”的不安感和恐惧，使得他们变得非常谨慎和怀疑，进而影响个人的获取信息的自由权力。

其次，个人信息侵权损害赔偿数额难以计量。《个人信息保护法》规定个人信息权益遭受侵害后可以请求损害赔偿，赔偿责任按照个人损害或处理者获益来确定，不能确定的根据“实际情况”确定赔偿；《网络安全法》、《数据安全法》都同样规定，违反本法给他人造成损害的，依法承担民事责任；《民法典》第 1182 条规定等，一系列法律都明确规定了信息处理者、网络运营者、数据处理者等的侵权损害赔偿责任，但是并没有具体地就赔偿责任的性质属于精神损害赔偿还是财产损害赔偿进行确定，也没有就具体的赔偿数额进行规定。

因此，人工智能时代下的个人信息权益损害早已不是单个主体的问题，权益损害的赔偿如何确定的问题就凸显出来了。因为 ChatGPT 对个人信息侵权行为总是表现为对不确定的大规模、大范围的单个个人信息权益的侵犯，所以很难确定个人损害赔偿的数额[7]。

再次，个人信息权益维护热情低迷。以生成式人工智能为代表的数字时代下，对个人信息的滥用通常是一种“大规模的微型损害”，即涉及大量的网络用户，但个体损害可能是轻微的[8]。大量案件中，用户都无法感受到即时性的伤害，或者因为侵权的隐秘性根本未能发觉自己权益受到侵害，又或者即使发现了也选择忽视。因此，个人信息权利的救济成本，直接影响到权利被侵害人是否原意进行权利维护，以及保护个人信息投入经济效益[9]。在个人权利意识较为突出的欧洲国家，因个人信息受到损害而提起诉讼的案件也非常少。

3.2. 个人信息自决权困境

德国学者施泰姆勒认为信息自决权是“人们有权自由决定周遭世界在何种程度上获知自己的所思所想以及行动的权利”[10]。而具体到以生成式人工智能时代的个人信息保护领域，个人信息自决权集中体现在“告知同意原则”上。虽然告知同意原则在个人信息领域被奉为圭臬，但其内生性矛盾与人工智能

时代的潮流已经有些格格不入。

其一，以 ChatGPT 为标志的强人工智能时代下，制度设计的告知义务难以保障知情权的实现。同大多数网络运营者一样，ChatGPT 也是使用隐私政策的方式对用户告知其个人信息的收集情况。这种分散的方式是无法让用户对自己个人信息被收集使用情况获得全面的了解。其二，“不选即走”的、为规避法律而产生的捆绑式的隐私政策变相迫使个人放弃信息控制权，事实上剥离了用户的“同意权”。ChatGPT 同大多数智能产品提供者一样仅仅为用户提供两种选择的可能性，即或是全盘接受条款内容，或是全盘否定条款内容而退出服务。软件的应用程序要求用户进行注册并同意隐私协议，否则就无法使用该软件，被破同意成为常见现象[11]。此外，网络平台提供的用户协议内容极其庞杂而重点不突出，极大地增加用户的阅读难度。有研究表明，用户仅阅读一年中所使用的网络服务的隐私政策就需要花费 224 个小时，同时还要考虑到用户的教育背景不同对于各条款的理解能力有所偏差所带来的现实性困境[12]。根据 OpenAI2023 年更新的隐私政策文本来看，其词汇数量达到 2900 多个，其使用的语言也非常抽象。实质上使得“合理目的”这一规制的目的落空，并且用户可能根本不能明白同意 COOKIES 追踪会导致什么样的后果，而只能以“勾选同意”为代价继续使用服务。

3.3. 个人信息安全监管困境

首先，存在监管不能的领域。基于 ChatGPT 模型的技术设计，在某些方面对个人信息的非法收集、泄露不具有监管的可能性：一是不能杜绝用户故意输入他人敏感信息以及其他隐私数据。破窗定律认为如果一个地方出现了一个未被修补的破窗玻璃，而没有人去修补它，那么人们就会有一种无视规则的信号，随之而来的是更多的窗户被砸和更严重的犯罪行为。这个定律同样也适用当前的情景，在泄露他人隐私的不良行为没有及时受到制止时，就必然会导致更多的人发生这项不良行为。并且这种现象已经发生，例如据《Economist》报导，近期三星半导体员工疑似因使用 ChatGPT，导致在三起不同事件中泄露公司机密。调查原因皆因员工将公司机密资讯输入 ChatGPT 而导致[13]。由于用户与 ChatGPT 的对话都会上传至 ChatGPT 数据库，因此很多员工在将自己的问题输入 ChatGPT 时，实际上就已经产生了数据泄露。二是不能杜绝 ChatGPT 输出他人的个人信息。事实上，GPT-2 在训练过程中就已经发生过生成式人工智能不当输出用户隐私信息的情况[14]。其实这种现象已经比较普遍。

其次，监管制度的不完善消减个人信息维护力量。一方面，如何确保《个人信息保护法》规定的“个人信息保护负责人”的独立性、以及第 58 条第 1 项中规定的“独立机构”的独立性是存在疑问的。在人工智能迅猛发展的时代个人信息庞大且涉及面广，缺乏对个人信息安全的监管的强有力的、强独立性的机构是难以维护公民切身权益的。另一方面，《个人信息保护法》中规定国家网信部门协调相关部门对个人信息进行保护，该法中并没有明确规定监管人工智能侵犯个人信息的专属权利归于哪一机关，网信部门是否能够统筹协调也是一个疑问，人工智能领域的个人信息具有复杂性，没有专门的监管行政机关就会出现无人监管或者双重监管的情况，不仅导致行政资源浪费、效率低下，也使得公民的权益难以得到及时救济。

4. 个人信息安全风险的行政规制

欧盟在个人信息保护方面的发展起步较早，并通过一系列法律制度来确保对个人信息的保护。2015 年欧洲议会颁布生效的《一般数据保护条例》，并被称为史上最严格的条例。该条例不仅重视个人数据主体权利保护，还规范数据控制者的个人数据处理行为，以平衡个人和社会利益。《一般数据保护条例》在全球都产生了广泛的影响和借鉴作用，多数国家在信息数据保护方面的立法不可避免第借鉴了欧洲经验。故此，我们可以通过分析欧盟《一般数据保护条例》，从相关个人信息、数据安全风险应对经验中

得到启示。

4.1. 完善个人信息保护监督机制

《中华人民共和国个人信息保护法》中引入了欧盟《一般数据保护条例》的两大制度，即“数据保护专员”和“独立监管机构”(在个人信息保护法中表现为“个人信息保护负责人”) [15]。根据《个人信息保护法》规定，当个人信息的处理量达到一定规模时，个人信息处理者应指定“个人信息保护负责人”，并由其负责对个人信息处理活动与采取的保护措施进行监督。此外，规模较大的个人信息处理者还应成立独立的个人信息保护监督机构，以加强对个人信息的管理和保护。但是其独立性、公正性依然饱受猜忌。针对“独立性”的问题，欧盟《一般数据保护条例》第 38 条明确规定数据控制着和处理者不得对数据专员发出任何有关执行数据专员职务的指示，数据专员不因执行自身职务而被解雇或遭受处罚。并且要求数据控制者与处理者负担保障数据专员顺利执行职务的义务。此外，还对数据专员的重要职务范围进行列举式的规定。第 52 条专门规定了独立监管机构的独立性问题，明确独立监督机构成员只听从监管机构的指示，独立监管机构的监督活动不受任何组织和个人的影响。在我国《个人信息保护法》中，就对个人信息处理者的监管制度中，就仅仅涉及到监管机构、监管专员进行规定，而对其独立性和监管效能却置之不理，这显然难以起到监管作用。因此，在个人信息保护的监管制度建构中，我们应当对欧盟关于信息数据保护监管机制进行深入研究，而不能停留在简单的制度模式的模仿，以此完善我国个人信息保护监督机制。

4.2. 加强职能部门对个人信息保护的监督能力

为了保护个人信息安全，需要加强公民的信息安全意识和维权意识，普及相关知识，增强保护个人信息的常识和动力。同时，信息从业者需要提高行业自律意识，约束自身行为，推进行业规范的完善和发展。LinkdIn 服务器被黑客攻击，超过 1 亿名用户的个人信息资料被窃取。这表明大多数社交网络用户缺乏相关专业知识，很容易成为信息安全隐患的受害者。政府应当加强事前监管和事后监督，对侵权审查工作人员进行基本业务素质的考核和奖惩标准制定，确保个人信息保护侵权审查各环节衔接有序。此外，在财政投入方面，应该增强大数据服务商的安全技术能力并提高智能设备水平，从数据收集、分析、应急处置直至最终的数据使用全程严格监督和审查。在研究上，需要建立严格的数据市场准入制度，以流程和交易方式规范数据获取、分发和使用，从源头把控数据的安全性。政府还应该加强事后监管，对侵权行为进行严厉处罚[9]。

4.3. 设置个人信息评估机构与行政咨询机制

网络的权力难以规制，依赖第三方的认证与标记能够为互联网用户提供直接的信赖判断，并成为信赖判断的直接依据，可以减少搜寻成本和内心忧虑。有调查表明，人们更愿意在可信任的网站分享个人信息，因为它们能够严格执行个人信息保护政策或者允诺有限使用个人信息[16]。事实上，对于此类机制，欧盟早在《一般数据保护条例》中就有所规定[17]。《一般数据保护条例》第 35 条规定当数据处理行为带来高度风险时，数据控制者应当制作“数据处理对个人数据保护影响的评估”。并且设置事先咨询机制。该条例的第 36 条规定，当数据控制者制作的影响评估表明在控制者缺乏减轻风险的措施会导致高风险时，数据控制者应当在处理前向监管机构咨询。系列相关规定从源头上消除数据控制者侵犯个人信息权益的机会，更好地衔接企业、行业自律与行政部门监督机制。而我国相关的个人信息保护立法、数据安全立法并没有较好的设置影响评估机制、咨询机制，因此在我国企业在个人信息数据处理上，一般是不进行事前的影响评估，也没有通过向行政部门咨询来进行企业合规的渠道。我国应当借鉴欧盟经验，

有必要通过建立个人信息数据影响评估机制和行政咨询机制来进一步完善以个人信息数据为产业核心要素的企业的合规，在最前端消除对个人信息安全的损害风险。

5. 结语

在憧憬 ChatGPT 的丰富的未来适用场景式，人们理应更加重视其暗含着的个人信息安全风险。以 ChatGPT 为代表的人工智能技术迭代更新远远超过人类的预估，但法律也不应当简单地全盘否定、过于保守。现阶段，ChatGPT 尚且处于市场化初期，仅仅根据其不确定性信息安全风险而限制其发展并非最佳抉择。因此，一个合乎逻辑的行政规制路径应当基于现有的相关法律条款进行解释和延伸使用，对现有的风险和权益困境进行行政立法规制，促使该项技术真正为人类所用才是我们的最终目的。

参考文献

[1] OpenAI. 隐私政策(Privacy Policy) [EB/OL]. https://openai.com/policies/privacy-policy, 2024-04-10.

[2] 朱光辉, 王喜文. ChatGPT 的运行模式、关键技术及未来图景[J]. 新疆师范大学学报(哲学社会科学版), 2023, 44(4): 113-122.

[3] 李昀锴. ChatGPT 内容商业使用的法律风险及应对[EB/OL]. https://mp.weixin.qq.com/s/8fzvmnyhEbIwWVTAV-m8WWA, 2024-04-17.

[4] 从立先, 李泳霖. 聊天机器人生成内容的版权风险及其治理——以 ChatGPT 的应用场景为视角[J]. 中国出版, 2023(5): 16-21.

[5] 周林兴, 韩永继. 大数据环境下个人信息治理研究[J]. 情报科学, 2021, 39(3): 11-18.

[6] 丁晓东. 从个体救济到公共治理: 论侵害个人信息的司法应对[J]. 国家检察官学院学报, 2022, 30(5): 103-120.

[7] 张建文, 时诚. 个人信息的新型侵权形态及其救济[J]. 法学杂志, 2021, 42(4): 39-52.

[8] 迈尔·舍恩伯格. 删除: 大数据取舍之道[M]. 袁杰, 译, 杭州: 浙江人民出版社, 2013: 56-76.

[9] 金泓序, 何畏. 大数据时代个人信息保护的挑战与对策研究[J]. 情报科学, 2022, 40(6): 132-140.

[10] 万方. 隐私政策中的告知同意原则及其异化[J]. 法律科学(西北政法大学学报), 2019, 37(2): 61-68.

[11] 郭雪慧. 人工智能时代的个人信息安全挑战与应对[J]. 浙江大学学报(人文社会科学版), 2021, 51(5): 157-169.

[12] 范为. 大数据时代个人信息保护的路径重构[J]. 环球法律评论, 2016, 38(5): 92-115.

[13] 张勇毅. ChatGPT 一周“紧急刹车”: 算力不足、隐私数据泄露[R]. https://www.thepaper.cn/newsDetail_forward_22644353, 2024-04-10.

[14] 李振林, 潘鑫媛. 生成式人工智能背景下数据安全的刑法保护困境与应对——以 ChatGPT 为视角的展开[J]. 犯罪研究, 2023(2): 25-33.

[15] 数据法律资讯公众号. 欧盟《一般数据保护条例》第四节: 数据保护专员[EB/OL]. https://www.dingxiang-inc.com/event/gdpr/gdpr.pdf, 2024-03-22.

[16] 调查: 68%的消费者愿意分享个人信息, 以满足个性化需求[EB/OL]. https://www.sohu.com/a/337742468_115514, 2024-04-11.

[17] 数据法律资讯公众号. 欧盟《一般数据保护条例》[EB/OL]. https://www.dingxiang-inc.com/event/gdpr/gdpr.pdf, 2024-03-18.