LATHAM&WATKINS
瑞生国际律师事务所

客户通讯
评论

瑞生国际律师事务所隐私及网络业务组

2023年8月23日 | 第3152号

Read this Client Alert in English

中国人工智能监管新规

中国出台新规旨在防范人工智能的相关风险，且规定了从事人工智能相关业务的实体的合规义务。

要点：

- 中华人民共和国（中国）通过推出并实施如下一系列法规，在人工智能监管方面领先于其他司法管辖区：
  - 互联网信息服务算法推荐管理规定（算法推荐规定），于2023年3月1日生效；
  - 互联网信息服务深度合成管理规定（深度合成规定），于2023年1月10日生效；
  - 生成式人工智能服务管理暂行办法（生成式AI办法），于2023年7月13日公布并于2023年8月15日生效；及
  - 科技伦理审查办法（试行）（征求意见稿）（伦理审查办法（草案）），于2023年4月14日公开征求意见，意见反馈截止时间为2023年5月3日。
- 算法推荐规定重点关注使用算法推荐技术（其中包括生成式和合成式算法）在中国境内提供互联网信息服务。深度合成管理规定重点关注使用深度合成技术（即生成式人工智能技术的其中一个子集）在中国境内提供互联网信息服务。另一方面，伦理审查办法（草案）则重点关注（其中包括）中国境内人工智能技术的研究和开发。生成式AI办法对开发并使用所有生成式人工智能技术在中国境内提供服务进行更广泛的监管。
- 该等法规规定了服务提供者、技术支持者和使用者，以及包括网络平台在内的一些其他实体的义务，其最终目的都是为了应对人工智能生成内容的相关风险并维护中国国家及社会安全。

本客户通讯讨论的内容包括：这些法规涵盖了哪些技术和实体，规定了哪些义务，以及明确了对违规行为的哪些处罚措施。同时，本客户通讯也将这些法规与欧盟的相应法规进行比较，并讨论了其对从事人工智能相关业务的实体带来的潜在影响。

背景

经济合作与发展组织（经合组织）将人工智能系统定义为，能够针对人类定义的一组既定目标作出影响真实或虚拟环境的预测、建议或决策的机器系统。生成式人工智能是人工智能的一个子集，利用深度学习技术基于输入数据生成输出内容（如图像、音乐或文本）。ChatGPT自2022年11月公开发布以来便成为引起大众关注的互动式人工智能语言模型，其正是这类技术的一个典型例子。中国方面，百度推出了旗下的生成式人工智能产品“文心一言”（Ernie Bot），其竞争对手亦已经或者预计会推出其他类似的工具。这类技术所涉风险（如个人数据泄露、违法信息传播及虚假新闻/内容等）也对现行的法律制度带来冲击和挑战，并引起了社会的广泛关注。多个司法管辖区已出台或正在出台新的法律法规以应对这些挑战。

与其他司法管辖区相同，在中国，数据保护、网络安全、不正当竞争以及电子商务方面的现有法律，可适用于人工智能的应用。但是，国家互联网信息办公室（网信办）是世界上首个推出专门规范人工智能的新法规的监管部门：

- **算法推荐规定**：中国推出的首部全国范围的人工智能专门法规是算法推荐规定，于2022年3月1日生效。其规范了使用算法推荐技术在中国境内提供网络服务的行为。

- **深度合成规定**：中国推出的第二个重要的人工智能专门法规是深度合成规定。网信办、工业和信息化部（工信部）及公安部于2022年11月25日联合发布了深度合成规定，于2023年1月10日生效。该规定还随附了一份公告及常见问答。深度合成规定的一项要求是向网信办提交相关算法。2023年6月23日，网信办公布了第一批在中国境内备案的深度合成算法（即人工智能算法备案清单），进一步明确了算法备案义务。

- **生成式AI办法**：2023年7月13日，网信办、国家发展与改革委员会（发改委）、教育部、科学技术部（科技部）、工信部以及公安部联合发布了生成式AI办法，该办法将于2023年8月15日生效，旨在规范更广泛的生成式人工智能技术。该办法还随附了一份公告及常见问答。

- **伦理审查办法（草案）**：2023年4月14日，科技部公布了伦理审查办法（草案），以征求意见，截止2023年5月3日，该办法草案重点关注具有伦理风险的科技活动（比如人工智能技术的研发）的伦理审查。

另外，且值得注意的是，根据中国国务院发布的2023年度立法工作计划，一个人工智能法律草案将在2023年被提交中国的全国人民代表大会常委会（中国的立法机构）审议。但是，人工智能法律制定和生效的预计时间尚不明确。

今后，虽然仍有待制定综合性人工智能法律，算法推荐规定、深度合成规定、伦理审查办法（草案）和生成式AI办法（统称人工智能法规）将成为中国规管人工智能相关服务及产品（包括生成式人工智能及人工智能生成的内容（AIGC））的主要法律。

适用范围

关于适用的行为的范围：

- **算法推荐规定**适用于任何使用算法推荐技术在中国境内提供互联网信息服务的行为；

- **深度合成规定**适用于任何使用深度合成技术在中国境内提供互联网信息服务的行为；

- **生成式AI办法**适用于使用生成式人工智能技术在中国境内向公众提供服务的行为，但并不包括尚未用于在中国境内向公众提供服务的生成式人工智能技术的研究和开发；及
- **伦理审查办法**（草案）一旦实施，将适用于任何涉及人类、实验室动物或者其他涉及伦理风险的科学技术活动。

关于适用的地域范围，人工智能法规并不限定其适用于中国个人及实体。外国个人或实体如涉及使用人工智能技术在中国境内提供服务或研发人工智能技术，亦有可能受到规管。但是，如果在中国境内尚未向公众提供相关服务，则生成式AI办法不适用于中国境内有关人工智能技术的研发。

受到规管的技术有哪些?

算法推荐技术

算法推荐规定中定义的“算法推荐技术”一词，包括以下算法：

- 生成及合成（例如AIGC相关产品和服务）；
- 个性化推送（例如根据网络购物应用的消费习惯作出产品推荐）；
- 分类和选择（例如社交媒体应用的热点搜索或选择）；
- 检索和过滤（例如搜索引擎应用自动识别和排除非法词汇）；及
- 调度相关决策（例如打车平台和订餐应用）。

深度合成

深度合成规定所定义的“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”。更具体来说，此类技术的用例包括：

- 篇章生成、文本风格转换、问答对话等生成或者编辑**文本内容**（例如，类似于ChatGPT的应用）；
- 文本转语音、语音转换、语音属性编辑等生成或者编辑**语音内容**；
- 音乐生成、场景声编辑等生成或者编辑**非语音内容**；
- 人脸生成、人脸替换、人物属性编辑、人脸/姿态操控等生成或者编辑**图像、视频内容中生物特征**（例如，深度伪造技术）；
- 图像生成、图像增强、图像修复等生成或者编辑**图像、视频内容中非生物特征**（例如，类似于DALL-E 2的应用）；及
- **三维重建、数字仿真**等生成或者编辑数字人物的技术。

生成式人工智能

生成式AI办法所定义的“生成式人工智能技术”，是指“有能力生成文本、图片、声音、视频等内容的模型和相关技术”。该办法并未具体说明此类技术的任何用例或示例，但表面上此词所指的较“合成技术”的范围

更广，因为任何内容生成技术将被纳入规制范围内，而有关内容可以是文本、图片、声音、视频和在互联网以外的其他信息。

具有伦理风险的科技活动

伦理审查办法（草案）广泛适用于“具有伦理风险的科技活动”并且其中特别列明：研发具有舆论属性或社会动员能力的算法模型、应用以及系统，属于具有高伦理风险的科技活动。

受到规管的实体有哪些?

算法推荐规定主要规定了服务提供者的义务。

深度合成规定对服务提供者、技术支持者、使用者以及任何涉及应用深度合成技术的其他实体（包括网络应用程序分发平台）规定了全面的义务。

生成式AI办法主要规定了服务提供者（即，通过生成式人工智能技术提供生成式人工智能服务的组织机构和个人（包括通过应用编程接口（API）向直接服务提供者开放生成式人工智能技术的技术支持者））的义务。

伦理审查办法（草案）规定了任何从事相关科研活动的高校、科研机构、医疗和卫生机构以及企业的伦理审查义务。

以下定义有助于理解所涉及的相关实体:

- “技术支持者”可以是指开发基础深度合成或其他生成式人工智能技术并提供技术支持的公司，如OpenAI和百度等。人工智能算法备案清单进一步阐明，“技术支持者”主要指通过应用程序编程接口（API）向企业客户（即B2B）提供生成式或合成类算法服务的主体（更多详情请参见下文）。
- “服务提供者”是指提供深度合成或其他生成式人工智能服务的公司（方法例如通过将有关服务整合到公司旗下的B2C互联网信息服务内）。
- 深度合成服务使用者包括利用深度合成服务制作、复制、发布、传播信息的组织、个人。
- “网络应用程序分发平台”是指App Store、Google Play store、Oppo app store和华为应用市场（Huawei AppGallery）等的互联网应用程序分发商店。

规定了哪些义务?

受人工智能法规规管的实体一般需要确保履行如下义务（如适用）：

序号	义务人	义务
A. 在推出新的人工智能产品、服务或应用程序时或之前		
算法的备案（无论生成式或非生成式人工智能）		
1.	所有算法推荐服务提供者（包括全部生成式人工智能服务提供者和技术支持者）	**算法备案** 根据算法推荐规定第24条的规定，义务人应在提供服务之日后10个工作日内向网信办或其地方网信部门进行有关算法（AIGC情况下，主要为生成式或合成类算法）的备案。 值得注意的是，算法推荐规定、深度合成规定和生成式AI办法均规定，只有在产品、服务或应用程序具有“舆论属性或社会动员能力”时才会产生备案义务。我们认为，这项规定旨在澄清关于是否触发备案的不清晰之处，但是鉴于如下所述该条件的范围广泛，算法备案将可能成为推出任何人工智能产品、服务或应用程序的强制性前置条件。 *舆论属性或社会动员能力* 人工智能法规并未对“舆论属性或社会动员能力”的构成给出清晰的定义或标准。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）于2018年由网信办及公安部联合公布，是对此事宜的唯一适用指引。根据安全评估规定，下述互联网信息服务应被视为具有舆论属性或社会动员能力： • 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；及 • 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。 所涉范围非常宽泛，可涵盖几乎所有互联网信息服务。在实践中，互联网公司通常会主动履行算法备案手续，但监管当局也可主动要求备案。 *备案* 算法备案应包含服务提供者或技术支持者（以适用者为准）的名称、服务或技术支持形式、算法类型、应用领域/产品类型（例如，应用程序、网站及小程序），以及算法自评估报告。备案手续完成后，应及时在相关网站或应用程序上展示备案编号和公示信息链接。 以下关于算法备案的不清晰之处，人工智能算法备案清单中做了进一步澄清：

序号	义务人	义务
		• *谁应作为技术支持者提出申请*：通过 API 向企业客户提供生成式或合成类算法服务（即 B2B）（而非直接服务公众）的任何实体。 • *哪些种类的服务会受到规管*：使用算法推荐技术的任何形式的服务，包括应用程序、网站和小程序。 • *大语言模型（LLM）算法是否受到规管*：是，百度（作为服务提供商）和科大讯飞（作为技术支持）已分别提交了LLM算法。 技术支持者应当参照服务提供者履行备案手续。但是，生成式AI办法尚未明确，如因提供API而被视为服务提供者的技术支持者，已经完成算法备案，则直接服务提供者是否还需要办理相关备案手续。在人工智能算法备案清单中，服务提供者和技术支持者的备案没有重叠，迄今为止，在41个备案算法中，仅8个算法是由技术支持者备案的。这表明，自算法推荐规定生效以来，直接服务提供者在算法备案过程中发挥了更加积极主动的作用。 *算法推荐规定、深度合成规定和生成式AI办法均规定了这一义务。*
产品、服务、应用程序等（无论生成式或非生成式人工智能）的安全评估		
2.	所有算法推荐服务提供者（包括全部生成式人工智能服务提供者和技术支持者）	**安全评估** 义务人在推出产品、服务、应用程序、功能或工具之前，应根据安全评估规定进行安全评估，并向所在地地市级网信部门和公安机关提交安全评估报告。 算法推荐规定、深度合成规定、生成式AI办法和安全评估规定均规定，只有在特定条件下才会触发安全评估。该等条件包括： • 上线的新产品、服务、应用程序或功能具有“舆论属性或社会动员能力”或导致“舆论属性或社会动员能力”发生重大变化（见上文的分析）； • 用户规模显著增加，导致“舆论属性或者社会动员能力”发生重大变化； • 发生违法有害信息传播扩散，表明已有安全措施无法有效防控网络安全风险； • （如为深度合成技术）模型、模板等的工具用于生成或者编辑人脸、人声等生物识别信息，或可能涉及国家安全、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；或 • 地市级以上网信部门或者公安机关书面通知需要进行安全评估的其他情形。 但是，鉴于“舆论属性或社会动员能力”的条件范围广泛，根据安全评估规定开展的安全评估，将很可能成为推出任何人工智能产品、服务或应用程序的强制性前置条件。

序号	义务人	义务
		安全评估 这项安全评估是不同于网信办在个人信息保护法（PIPL）下要求就个人信息出境进行的安全评估，以及（构成算法备案一部分的）算法自评估报告，且与上述各项分开的。 根据安全评估规定，服务提供者可以自行实施安全评估，也可以委托第三方安全评估机构实施。安全评估重点评估下列内容： • 确定与所提供服务相适应的安全管理负责人、信息审核人员或者建立安全管理机构的情况； • 用户真实身份核验以及注册信息留存措施； • 对用户的账号、操作时间、操作类型、网络源地址和目标地址、网络源端口、客户端硬件特征等日志信息，以及用户发布信息记录的留存措施； • 用户账号和通讯群组名称、昵称、简介、备注、标识，信息发布、转发、评论和通讯群组等服务功能中违法有害信息的防范处置和有关记录保存措施； • 个人信息保护以及防范违法有害信息传播扩散、社会动员功能失控风险的技术措施；及 • 服务提供者是否已建立(i) 投诉、举报制度，公布投诉、举报方式等信息，及时受理并处理有关投诉和举报的情况；及(ii) 为相关互联网信息服务及国家安全或公安机关履行其监督管理职责提供技术、数据支持和协助的工作机制的情况。 安全评估报告应基于安全评估制定，由服务提供者通过全国互联网安全管理服务平台提交所在地地市级以上网信部门和公安机关。该报告应包括：(i) 服务功能、服务范围、软硬件设施、部署位置等基本情况和相关证照获取情况；(ii) 安全管理制度和技术措施落实情况及风险防控效果；(iii) 安全评估结论；及(iv) 其他应当说明的相关情况。 *算法推荐规定、深度合成规定及生成式AI办法均规定了该项义务。*
应用程序分发平台进行程序性审查（主要是生成式人工智能）		
3.	网络应用程序分发平台（AIGC的情况）	**应用商店进行核实验证** 在网络应用程序分发平台上推出应用程序之前，必须核实适用法律法规所规定的安全评估和备案手续均已完成。虽然这不属于应用商店提起的实质性审查，但却似乎增加了另一层仅适用于应用程序的程序性审查。尽管网站无需完成该第三方审核，但其仍须遵守算法备案和安全评估要求。 *仅深度合成规定规定了这项义务。*

序号	义务人	义务
人工智能有关研发（无论生成式及非生成式人工智能）的伦理审查		
4.	从事研发活动的人工智能服务提供者及技术支持者	**伦理审查** 根据伦理审查办法（草案）第5条，任何高校、科研机构、医疗卫生机构和企业从事包括人工智能在内的某些“伦理敏感领域”科技活动的，都必须成立科技伦理审查委员会。 伦理审查办法（草案）附录列出了高风险科技活动清单，其中包括具有舆论社会动员能力和社会意识引导能力的算法模型、应用程序及系统的研发。根据伦理审查办法（草案）第 32 条的规定，这些高风险活动应接受 (i) 伦理审查委员会的初步审查；(ii) 当地或相关行业主管部门的专家复核。 虽然“伦理敏感”的含义尚不明确，但鉴于人工智能相关研发已被明确列为上述高风险科技活动，成立伦理审查委员会并对人工智能相关研发进行伦理审查很可能成为人工智能服务提供者和技术支持者的强制性（而非有条件）前置条件。 登记 义务人应在科技部建立的国家科技伦理管理信息登记平台上，分别 (i) 在伦理审查委员会成立后 30 天内（伦理审查办法（草案）第 44 条）登记伦理审查委员会；(ii) 在高风险科技活动（包括伦理审查委员会的初审和地方或相关行业主管部门的专家复审）经伦理审查核准后 30 天内（伦理审查办法（草案）第 45 条）登记该高风险科技活动。登记内容应包括：(i) 伦理审查委员会及其组成、章程、工作制度等相关信息；(ii) 科技活动实施方案、初审和专家复核结果等相关信息。相关内容发生变化时，应及时更新登记内容。 *该义务源于伦理审查办法（草案），尚不具有法律约束力。*
B. 在人工智能产品、服务或应用程序的日常营运期间		
一般义务		
1.	所有算法推荐服务提供者（包括全部生成式人工智能服务提供者和技术支持者）	**遵守法律、社会公德及伦理** 根据算法推荐规定第4条，提供算法推荐服务，应当遵守法律法规，尊重社会公德和伦理，遵守商业道德和职业道德，遵循公正公平、公开透明、科学合理和诚实信用的原则。 与之类似，根据深度合成规定第4条，提供深度合成服务，应当遵守法律法规，尊重社会公德和伦理道德，坚持正确政治方向、舆论导向、价值取向，促进深度合成服务向上向善。

序号	义务人	义务
		虽然还有反映这些价值的更具体的义务（如下面列出的关于输入和输出的具体义务），但这些是各项服务的总体价值和要求。除了更具体的要求外，这些义务在实践中如何适用还有待观察。 *算法推荐规定和深度合成规定规定了该项义务。*
2.	所有算法推荐服务提供者（包括全部生成式人工智能服务提供者和技术支持者）	**不得使用算法推荐服务进行操纵** 根据算法推荐规定第14条，算法推荐服务提供者不得利用算法虚假注册账号、非法交易账号、操纵用户账号或者虚假点赞、评论、转发，不得利用算法屏蔽信息、过度推荐、操纵榜单或者检索结果排序、控制热搜或者精选等干预信息呈现，实施影响网络舆论或者规避监督管理行为。 *算法推荐规定规定了此项义务。*
3.	所有算法推荐服务提供者（包括全部生成式人工智能服务提供者和技术支持者）	**不得使用算法推荐服务进行不正当竞争** 根据算法推荐规定第15条，算法推荐服务提供者不得利用算法对其他互联网信息服务提供者进行不合理限制，或者妨碍、破坏其合法提供的互联网信息服务正常运行，实施垄断和不正当竞争行为。 与之类似，根据生成式AI办法第4(3)条，义务人不得利用算法、数据、平台等优势，实施垄断和不正当竞争行为。
训练数据和数据标注（主要是生成式人工智能）		
4.	所有生成式人工智能服务提供者和（通过API提供技术的）技术支持者	**训练数据规则** 生成式AI办法第7条规定，义务人应当依法开展预训练、优化训练等训练数据处理活动，遵守以下规定： • 使用具有合法来源的数据和基础模型； • 涉及知识产权的，不得侵害他人依法享有的知识产权； • 涉及个人信息的，应当取得个人同意或者符合法律、行政法规规定的其他情形； • 采取有效措施提高训练数据质量，增强训练数据的真实性、准确性、客观性、多样性； • 中华人民共和国网络安全法、中华人民共和国数据安全法、中华人民共和国个人信息保护法等法律、行政法规。

序号	义务人	义务
		与之类似，深度合成规定第14条要求，深度合成服务提供者和技术支持者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。 我们注意到，生成式AI办法正式版本中的规定“采取有效措施”增强训练数据的真实性、准确性、客观性、多样性，相比生成式AI办法的草案（要求义务人“确保”训练数据的真实性、准确性、客观性、多样性），有所放松。 *深度合成规定和生成式AI办法规定了该项义务。*
5.	所有生成式人工智能服务提供者和（通过API提供技术的）技术支持者	**训练数据标注规则** 根据生成式AI办法第8条，在生成式人工智能技术研发过程中进行数据标注的，义务人应当制定符合本办法要求的清晰、具体、可操作的标注规则；开展数据标注质量评估，抽样核验标注内容的准确性；对标注人员进行必要培训，提升尊法守法意识，监督指导标注人员规范开展标注工作。 “数据标注”一词与人工智能法规下所要求的内容标识（见下文）有所不同。数据标注是在训练模型时用人手将标记或标签添加到相片、视频、文本和音频等原始数据的过程。这些标记说明数据的实体类型，引用该数据的各种属性和特征。这样的话，机器学习模型便可在没有标签的情况下学习识别数据，且对建立准确的模型十分重要，但前提是标注人员具备所需经验和训练有素。 *仅生成式AI办法规定了这项义务。*
6.	所有生成式人工智能服务提供者和（通过API提供技术的）技术支持者	**配合监管部门对训练数据作出说明** 根据生成式AI办法第19条，有关主管部门依据职责对生成式人工智能服务开展监督检查，义务人应当依法予以配合，按要求对训练数据来源、规模、类型、标注规则、算法机制机理等予以说明，并提供必要的技术、数据等支持和协助。 *仅生成式AI办法规定了这项义务。*
核实用户的真实身份（主要是生成式人工智能）		
7.	深度合成服务提供者	**核实用户的真实身份** 根据深度合成规定第9条，深度合成服务提供者应当基于移动电话号码、身份证件号码、统一社会信用代码或者国家网络身份认证公共服务等方式，依法对深度合成服务使用者进行真实身份信息认证，不得向未进行真实身份信息认证的深度合成服务使用者提供信息发布服务。 *仅深度合成规定规定了这项义务。*

序号	义务人	义务
针对更改生物识别信息取得单独同意（主要是生成式人工智能）		
8.	深度合成服务提供者和技术支持者	**通知及获得深度合成服务主体的同意** 根据深度合成规定第14条，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。例如，如果个人/公司甲向深度合成服务输入了个人乙的人脸照片以更改个人乙的人脸，个人/公司甲必须取得个人乙的单独同意。 *仅深度合成规定规定了这项义务。*
输入数据（主要是生成式人工智能）		
9.	所有生成式人工智能服务提供者和（通过API提供技术的）技术支持者	**保护用户输入数据的隐私** 根据生成式AI办法第11条，提供者对使用者的输入信息和使用记录应当依法履行保护义务，不得收集非必要个人信息，不得非法留存能够识别使用者身份的输入信息和使用记录，不得非法向他人提供使用者的输入信息和使用记录。 *仅生成式AI办法规定了这项义务。*
输出内容审核（无论生成式及非生成式人工智能）		
10.	所有算法推荐服务提供者（包括全部生成式人工智能服务提供者和技术支持者）及使用者	**呈现符合主流价值导向的信息，倡导社会主义核心价值，以及内容限制** 算法推荐规定第11条要求义务人应当在首页首屏、热搜、精选、榜单类、弹窗等重点环节积极呈现符合主流价值导向的信息。 算法推荐规定第6条也笼统地规定，算法推荐服务提供者应当坚持主流价值导向，积极传播正能量。 与之类似，深度合成规定第4条规定，提供深度合成服务，应当坚持正确政治方向、舆论导向、价值取向，促进深度合成服务向上向善。 算法推荐规定和深度合成规定均规定，义务人不得利用服务： • 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；或 • 制作、复制、发布、传播法律和行政法规禁止的信息。 生成式AI办法第4条进一步规定，义务人： • 坚持社会主义核心价值观；

序号	义务人	义务
		• 不得生成危害国家安全和利益、损害国家形象，宣扬恐怖主义、极端主义，宣扬民族仇恨、民族歧视，暴力、淫秽色情，以及虚假有害信息等法律、行政法规禁止的内容； • 采取有效措施防止产生民族、信仰、国别、地域、性别、年龄、职业、健康等歧视； • 尊重知识产权、商业道德，保守商业秘密，不得利用算法、数据、平台等优势，实施垄断和不正当竞争行为； • 尊重他人合法权益，不得危害他人身心健康，不得侵害他人肖像权、名誉权、荣誉权、隐私权和个人信息权益。 *算法推荐规定、深度合成规定和生成式AI办法均规定了该项义务。*
11.	所有生成式人工智能服务提供者和（通过API提供技术的）技术支持者	**采取有效措施提高AIGC的准确性和可靠性** 根据生成式AI办法第4(5)条，义务人应当基于服务类型特点，采取有效措施，提高生成内容的准确性和可靠性。 我们注意到生成式AI办法正式版本中的该项规定（采取有效措施提高AIGC的准确性和可靠性），相比其草案（其中义务是“确保”AIGC真实准确且不包含任何歧视性内容）有所放松。 *仅生成式AI办法规定了这项义务。*
12.	所有算法推荐服务提供者（包括全部合成服务提供者及使用者）	**特别禁止传播虚假消息** 根据算法推荐规定第13条，算法推荐服务提供者提供互联网新闻信息服务的，应当依法取得互联网新闻信息服务许可，规范开展互联网新闻信息采编发布服务、转载服务和传播平台服务，不得生成合成虚假新闻信息，不得传播非国家规定范围内的单位发布的新闻信息。 根据深度合成规定第6条，深度合成服务提供者和使用者不得利用深度合成服务制作、复制、发布、传播虚假新闻信息。 这些规定特别禁止传播假新闻或并非来自国家授权实体的新闻。但根据人工智能生成条例，真实性和准确性也是必须履行的义务。根据上述关于真实性和准确性的生成式AI办法，这项义务也是必要的。 *算法推荐规定和深度合成规定均规定了此项义务。*
13.	所有算法推荐服务提供者（包括全部合成服务提供者）	**屏蔽非法有害信息** 根据算法推荐规定第9条，义务人应当建立健全用于识别违法和不良信息的特征库，完善入库标准、规则和程序。

序号	义务人	义务
		深度合成规定第10条进一步规定，深度合成服务提供者应当加强深度合成内容管理，采取技术或者人工方式对深度合成服务使用者的输入数据和合成结果进行审核。 算法推荐规定第12条鼓励算法推荐服务提供者综合运用内容去重、打散干预等策略（避免特定推荐结果集过度集中）。 *算法推荐规定和深度合成规定均规定了此项义务。*
14.	所有算法推荐服务提供者（包括全部生成式人工智能服务提供者和技术支持者）	**防范虚假、违法或有害信息的措施** 总体而言，义务人应当采取措施防范和制止传播虚假、违法或有害信息。 根据算法推荐规定，发现违法信息的，应当立即停止传输，采取消除等处置措施，防止信息扩散，保存有关记录，并向网信部门和有关部门报告。发现不良信息的，应当按照网络信息内容生态治理有关规定予以处置。 根据深度合成规定第11条，义务人发现利用深度合成服务制作、复制、发布、传播虚假信息的，应当及时采取辟谣措施，保存有关记录，并向网信部门和有关主管部门报告。 与之类似，生成式AI办法规定，义务人发现违法内容的，应当及时采取停止生成、停止传输、消除等处置措施，采取模型优化训练等措施进行整改，并向有关主管部门报告。义务人发现使用者利用生成式人工智能服务从事违法活动的，应当依法依约采取警示、限制功能、暂停或者终止向其提供服务等处置措施，保存有关记录，并向有关主管部门报告。 此外，生成式AI办法第9条规定，义务人应当依法承担网络信息内容生产者责任。尽管在生成式AI办法中并无“网络信息内容生产者”及其义务的明确定义，但相同概念见于网信办发布的网络信息内容生态治理规定（网络生态治理规定，于2020年3月1日生效）。根据网络生态治理规定，网络信息内容生产者不得制作、复制、发布违法信息；应当采取措施，防范和抵制制作、复制、发布不良信息。 *算法推荐规定、深度合成规定和生成式AI办法规定了该项义务。*
合成内容标识（主要是生成式人工智能）		
15.	所有生成式人工智能服务提供者和（通过API提供技术的）技术支持者	**标识内容** 根据深度合成规定第16条和生成式AI办法第12条，关于使用深度合成服务生成或者编辑的信息内容，必须以不影响用户使用有关服务的方式进行标识。有关服务提供者和技术支持者必须保存相关网络日志信息的记录。 人工智能法规没有说明添加内容标识方法的相关具体要求，而是参照深度合成规定下的标识规定。

序号	义务人	义务
		深度合成规定和生成式AI办法规定了该项义务。
16.	深度合成服务提供者	**标识合成内容** 根据深度合成规定第17条，可能导致公众混淆或者误认的深度合成内容，义务人必须在合理位置、区域进行显著标识，显示该信息为深度合成内容。该内容包括模拟真人风格进行的智能对话或智能写作（例如，ChatGPT）、仿声、人脸图像合成、人脸操控、沉浸式拟真场景等生成或者编辑服务。 就其他深度合成内容而言，义务人必须提示使用者可以进行显著标识，显示相关信息为合成内容。 根据深度合成规定第18条，个人和组织机构均不得采用技术手段删除、篡改、隐匿上述标识。 *该项义务主要由深度合成规定规定。算法推荐规定也规定，任何由算法生成的合成内容应当显著标识。*
保护使用者的权利（无论生成式或非生成式人工智能）		
17.	所有算法推荐服务提供者	**使用标签和选择** 根据算法推荐规定第11条，义务人应当加强用户模型和用户标签的管理，完善用户模型和用户标签管理中的兴趣记录规则，不得将违法有害信息作为用户兴趣关键词记录，或者制作成用户标签作为推荐信息内容的依据。 此外，根据该规定的第12条，算法推荐服务提供者还应通过建立健全人工干预机制来加强生态系统管理，并允许用户自主选择内容。 *算法推荐规定规定了此项义务。*
18.	所有算法推荐服务提供者	**用户权利** 根据算法推荐规定第17条，义务人应当保护用户的以下权利： • **不被针对（通过个人特征）的权利**：义务人应当向用户提供不针对其个人特征的选项； • **选择退出的权利**：义务人应当向用户提供便捷的关闭算法推荐服务的选项；及 • **删除个人特征的权利**：义务人应当使用户能够选择或者删除用于算法推荐服务的针对其个人特征的用户标签。 此外，根据算法推荐规定第21条，用户享有以下权利：

序号	义务人	义务
		• **不被实施差别待遇的权利**：义务人向消费者销售商品或者提供服务的，不得根据消费者的偏好、交易习惯等特征，利用算法在交易价格等交易条件上实施不合理的差别待遇等违法行为。 根据生成式AI办法第11条，义务人应当依法及时受理和处理个人关于查阅、复制、更正、补充、删除其个人信息等的请求。 *算法推荐规定和生成式AI办法规定了该项义务。*
19.	所有生成式人工智能服务提供者和（通过API提供技术的）技术支持者	**使用条款** 根据生成式AI办法第9条，义务人当与生成式人工智能服务使用者签订服务协议，明确双方权利义务。 *生成式AI办法规定了该项义务。*
20.	所有算法推荐服务提供者（包括全部生成式人工智能服务提供者和技术支持者）	**披露规定** 根据算法推荐规定第16条，义务人应当以显著方式告知用户其提供算法推荐服务的情况，并以适当方式公示算法推荐服务的基本原理、目的意图和主要运行机制等。 应用算法对用户权益造成重大影响的，应当依法予以说明并承担相应责任。 根据算法推荐规定第7条和第12条，义务人制定并公开算法推荐服务相关规则，并且鼓励义务人优化检索、排序、选择、推送、展示等规则的透明度和可解释性，避免对用户产生不良影响，预防和减少争议纠纷。 根据生成式AI办法第10条和第18条，义务人应当明确并公开其服务的适用人群、场合、用途，指导使用者科学理性认识和依法使用生成式人工智能技术。 *算法推荐规定和生成式AI办法均规定了该项义务。*
21.	所有算法推荐服务提供者	**不得诱导用户沉迷或过度消费** 根据算法推荐规定第8条，义务人不得设置诱导用户沉迷、过度消费等违反法律法规或者违背伦理道德的算法模型。 针对未成年人，也特别要求履行此项义务（见下文）。 *算法推荐规定规定了该项义务。*
22.	所有算法推荐服务提供者（包括全部生成式人工智能服务提供者和技术支持者）	**保护未成年人、老年人和劳动者** 根据算法推荐规定第18条，义务人应当依法履行未成年人网络保护义务，并

序号	义务人	义务
		通过开发适合未成年人使用的模式、提供适合未成年人特点的服务等方式，便利未成年人获取有益身心健康的信息。义务人不得： • 向未成年人推送可能引发未成年人模仿不安全行为和违反社会公德行为、诱导未成年人不良嗜好等可能影响未成年人身心健康的信息； • 利用算法推荐服务诱导未成年人沉迷网络。 根据生成式AI办法第10条，义务人应当采取有效措施防范未成年人用户过度依赖或者沉迷生成式人工智能服务。 根据算法推荐规定第19条，算法推荐服务提供者向老年人提供服务的，应当保障老年人依法享有的权益，充分考虑老年人出行、就医、消费、办事等需求，按照国家有关规定提供智能化适老服务，依法开展涉电信网络诈骗信息的监测、识别和处置，便利老年人安全使用算法推荐服务。 此外，根据算法推荐规定第20条，算法推荐服务提供者向劳动者提供工作调度服务的，应当保护劳动者取得劳动报酬、休息休假等合法权益，建立完善平台订单分配、报酬构成及支付、工作时间、奖惩等相关算法。 *算法推荐规定和生成式AI办法均规定了该项义务。*
算法的定期审核（无论生成式或非生成式人工智能）		
23.	所有算法推荐服务提供者（包括全部深度合成服务提供者和技术支持者）	**算法的定期审核** 服务提供者和技术支持者必须定期审核、评估、验证算法机制机理、模型、数据和应用结果。（算法推荐规定第8条和深度合成规定第15条） *算法推荐规定和深度合成规定均规定了该项义务。*
24.	从事研发活动的人工智能服务提供者和技术支持者	**定期伦理审查** 根据伦理审查办法（草案）第46条，义务人应于每年3月31日前，向国家科技伦理管理信息登记平台提交上一年度科技伦理（审查）委员会工作报告，以及纳入清单管理的科技活动实施情况报告。 此外，根据伦理审查办法（草案）第39条，义务人应对高风险科技活动进行跟踪审查，跟踪审查频度在该等活动被批准后不少于每年两次。科技伦理风险发生重大变化的，应重新开展伦理审查和专家复核，获得批准后方可继续实施 *该义务源于伦理审查办法（草案），尚不具有法律约束力。*

序号	义务人	义务
25.	所有生成式人工智能服务提供者和（通过API提供技术的）技术支持者	**其他监管部门的监督和核查** 根据生成式AI办法第16条，生成式人工智能服务受多个相关监管部门的监管，包括网络、发展改革、教育、科技、工业和信息化、公安、广播电视、新闻出版等部门。 相关监管部门应当按其各自职能对生成式人工智能服务进行监督和核查。但是，生成式AI办法在这方面并没有作出任何更详细的规定，包括如何开展监督和核查以及在此过程中需要审查哪些内容。 *生成式AI办法规定了该项义务。*
数据安全管理系统以及技术措施（无论生成式或非生成式人工智能）		
26.	所有算法推荐服务提供者（包括全部生成式人工智能服务提供者和技术支持者）	**遵守数据安全和个人信息保护法律** 提供人工智能相关服务和产品，应当遵守法律法规（算法推荐规定第4条；深度合成规定；及生成式AI办法）。在AIGC的情况下，生成式人工智能服务提供者和技术支持者必须确保遵守中国的适用数据安全和个人信息保护法律法规，包括网络安全法、数据安全法、个人信息保护法和科学技术进步法。 *算法推荐规定、深度合成规定和生成式AI办法规定了该项义务。*
27.	所有算法推荐服务提供者（包括全部深度合成服务提供者）	**管理制度和技术措施** 一般而言，人工智能服务提供者和技术支持者必须建立健全用户注册、基础算法和技术的原则审查和伦理审查、信息发布审查、内容审核、数据安全及个人信息保护、反电信网络诈骗、安全评估监测，以及安全事件应急处置及数据泄露等管理制度和技术措施（算法推荐规定第7条；及深度合成规定）。 *算法推荐规定和深度合成规定规定了该项义务。*
28.	网络应用程序分发平台（AIGC的情况）	**网络应用程序分发平台的管理措施** 根据深度合成规定第13条，应用程序分发平台必须落实上架审核、日常管理、应急处置等安全管理措施。对违反适用法律法规的应用程序，相关平台必须对所涉应用程序采取不予上架、警示、暂停服务或者下架等措施。 *仅深度合成规定规定了此项义务。*
辟谣及投诉处理机制（无论生成式或非生成式人工智能）		
29.	深度合成服务提供者	**辟谣机制** 深度合成规定第11条规定，服务提供者应当建立健全辟谣机制，发现利用任何虚假信息的，应当及时采取辟谣措施，保存有关记录，并向网信部门和有关主管部门报告。

序号	义务人	义务
		仅深度合成规定规定了此项义务。
30.	所有算法推荐服务提供者（包括全部生成式人工智能服务提供者和技术支持者）	**建立用户投诉处理机制** 一般而言，义务人应当建立投诉处理机制。 根据算法推荐规定第22条，义务人应当设置便捷有效的用户申诉和公众投诉、举报入口，明确处理流程和反馈时限，及时受理、处理并反馈处理结果。 根据深度合成规定第12条，义务人应设置便捷的公众及用户投诉渠道，公布投诉处理流程和反馈时限，及时受理、处理和反馈处理结果。 与之类似，生成式AI办法第15条规定，义务人应当建立健全投诉、举报机制，设置便捷的投诉、举报入口，公布处理流程和反馈时限，及时受理、处理公众投诉举报并反馈处理结果。 *算法推荐规定、深度合成规定和生成式AI办法规定了该项义务。*

处罚

算法推荐规定

算法推荐规定第31条规定，服务提供者违反该规定的，法律、行政法规有规定的，依照其规定；法律、行政法规没有规定的，由网信部门和电信、公安、市场监管等有关部门依据职责给予警告、通报批评，责令限期改正；拒不改正或者情节严重的，责令暂停信息更新，并处一万元以上十万元以下罚款。

深度合成规定

深度合成规定本身没有明确列明不合规行为的处罚。相反，第22条规定，深度合成服务提供者或技术支持者违反该规定的，应依照适用法律法规的规定相应处罚。此外，该规定也没有注明使用者或网络平台不遵守法律法规的后果。

尽管如此，该规定赋予网信、电信及公安等部门监督深度合成业务的合规情况并对之进行检查。网信部门和其他有关主管部门认为深度合成服务存在较大信息安全风险的，可以责令深度合成服务提供者和技术支持者采取暂停信息更新、用户账号注册或者其他相关服务等措施，要求整改违规行为。

生成式AI办法

根据生成式AI办法第21条：

- 服务提供者（包括通过API提供技术的技术支持者）违反本规定的，应由网信部门和其他有关主管部门按照网络安全法、数据安全法、个人信息保护法和科学技术进步法等法律法规的规定予以处罚；及
- 法律、行政法规没有规定的，由有关主管部门依据职责予以警告、通报批评，责令限期改正；拒不改正或者情节严重的，责令暂停提供相关服务。

具体而言，对于来自中国境外但向境内个人提供的生成式人工智能服务，如果违反了中国相关法律法规，网信办可要求相关部门采取技术措施或其他必要措施。

伦理审查办法

与深度合成规定类似，伦理审查办法（草案）本身也没有明确规定任何违规处罚。相反，第 48 条仅规定，如果任何从事相关科技活动的实体违反该办法的规定，可根据其他适用的法律法规进行相关处罚。

2022年3月，网信办联合其他部门开展专项行动，对算法推荐服务的合规情况进行检查，并对违反算法推荐规定的行为采取执法行动。由于深度合成规定于今年生效，生成式AI办法最近才生效，伦理审查办法（草案）尚未生效，我们没有发现根据这些法规采取的执法行动。

与欧盟的发展进程比对

2021年，欧盟委员会提出了一项范围广泛的新法规，以协调统一适用于欧洲内部市场的人工智能系统规则（欧盟人工智能法提案）。欧盟人工智能法提案正在进行密集的立法辩论。

背景

欧洲理事会于2022年12月6日通过了欧盟人工智能法的谈判版本。经过几个月的谈判，欧洲议会于 2023 年 6 月批准了其谈判版本，提出了对欧盟人工智能法提案的实质性修订。除其他外，欧洲议会提出的修正案还包括针对一般基础模型的额外义务，以及针对生成式人工智能系统中使用的基础模型的特定义务（例如要求公布用于人工智能训练的受版权保护数据的摘要）。欧盟立法者目前正进入立法程序的最后阶段，他们将在此期间就欧盟人工智能法的最终形式进行谈判；预计欧盟人工智能法提案在成为最终文本之前将进行大量修改。一旦文本定稿（预计在 2023 年底或 2024 年初），欧盟人工智能法将在欧盟官方公报上公布一个月后生效，预计目前的提案文本包含欧盟人工智能法的大部分义务，将在两年后生效，即最迟在 2026 年年中生效。

欧盟人工智能法提案

欧盟人工智能法提案提供了按照基于风险的方法区分人工智能实践的框架。某些威胁到人们安全、生计和权利的特定人工智能实践被认为风险水平不可接受，并被禁止。针对高风险人工智能，提出了登记、风险评估、人工监督和其他义务，而针对有限风险人工智能，相关主体须履行拟议的透明度义务。欧盟人工智能法提案规定，罚款最高可达 3000 万欧元或全球年度总营业额的 6%，以两者中较高者为准。相比之下，中国的人工智能法是一套部门规章，侧重于人工智能系统的不同方面。中国目前还没有专门的人工智能法律，但我们注意到，起草这样一部法律已被列入国务院 2023 年立法工作计划。

加强反虚假信息行为守则

2022年6月16日，欧盟公布了加强<反虚假信息行为守则>（欧盟守则），以减低网络虚假信息产生的影响。以规制范围而言，与欧盟守则相比，深度合成规定对范围更广的实体施加了更广泛的义务。欧盟守则要求签署方（主要为大型科技公司）落实约定的承诺与措施，以打击深度伪造和虚假信息的情况。同时，深度合成规定适用于所有涉及深度合成服务的参与者，包括服务提供者、技术支持者、使用者和网络平台等。此外，欧盟守则主要侧重于规制错误信息、虚假信息，以及通用数据保护条例（个人信息保护）、经修订的数字服务法（适用于网络服务和平台的透明度义务和若干用户保护措施）及欧盟人工智能法（全面监管人工智能系统）中涉及或将要涉及的其他领域。同时，深度合成规定从不同的角度（包括网络安全、数据管理、个人信息保护、算法审计和备案、实名验证及内容审核等）提供一个全方位的义务框架。

影响

人工智能法规标志着中国致力于监管人工智能相关业务，重点关注AIGC（尤其是应对与深度伪造和其他合成内容有关的风险）方面向前迈进一步。由于该等法规涵盖范围广泛，合规义务增多，所以，从事人工智能业务的个人和实体（特别是服务提供者和技术支持者）应当尽早进行必要的自我评估，以评定合规情况。同时，编制一份全面的合规核对表也会对此有帮助。

除此之外，生成式AI办法的某些规定可能在实践中为服务提供者带来挑战。服务提供者和技术支持者应当密切关注这方面的进一步立法发展情况和监管指引。

如对本客户通讯有任何疑问，请联系下列作者之一或您通常咨询的瑞生律师：

徐辉（Hui Xu）
hui.xu@lw.com
+86.10.5965.7006
北京

Kieran Donovan
kieran.donovan@lw.com
+852.2912.2701
香港

Esther C. Franks
esther.franks@lw.com
+65.6437.5345
新加坡

李晓霖（Bianca Lee）
bianca.lee@lw.com
+852.2912.2500
香港

黄奕瑜（Michelle Wong）
michelle.wong@lw.com
+852.2912.2729
香港

本客户通讯在瑞生北京办事处的李芷莹和瑞生香港办事处的许琪协助下编制。

阁下可能感兴趣的其他文章：

中国发布个人信息出境标准合同草案并明确数据出境机制

中国出台有关在海外上市互联网平台运营者的网络安全审查新规

中国首次出台综合性个人信息保护法律

中国新数据安全法：新规要点

中国发布关于保护关键信息基础设施的新条例
