

植德 人工智能年刊（2023）

北京植德律师事务所



北京 | 上海 | 深圳 | 武汉 | 杭州 | 青岛 | 成都 | 海口 | 香港

Beijing | Shanghai | Shenzhen | Wuhan | Hangzhou | Qingdao | Chengdu | Haikou | Hong Kong

前言

2022 年年末，ChatGPT 上线并迅速风靡全球，引起人工智能讨论热潮，众多科技企业纷纷加快人工智能的研发，以求在人工智能领域抢占先机。2023 年，为应对人工智能技术突进带来的挑战，中国国家互联网信息办公室在短短的 3 个月内将《生成式人工智能服务管理暂行办法》的征求意见稿落地，成为全球第一个全面监管生成式人工智能的立法文件，同时将人工智能法列入立法工作计划。然而，由于人工智能产业的高速发展，监管政策的更新难以覆盖多样化的现实问题。因此，对于研发人工智能的科技企业来说，及时了解现行法律政策体系下的监管动态并预测未来的监管趋势尤为重要。

为帮助企业尽快了解人工智能领域的监管现状、具体政策及发展前景，北京植德律师事务所在梳理人工智能领域监管政策体系的基础上，针对行业重点合规问题进行分析、解读，形成《植德人工智能年刊（2023）》。《植德人工智能年刊（2023）》分为四个章节，分别是投融资并购、人工智能监管现状、知识产权合规和数据合规。第一章投融资并购部分，从中国 AIGC 产品的商业化路径、AIGC 企业的公司治理结构和美国出口管制对 AIGC 企业的影响三个方面展开为 AIGC 企业提供商业指引。第二章人工智能监管现状部分，主要对 2023 年中国颁布的人工智能相关法律政策进行解读，洞察人工智能监管趋势。第三章知识产权合规部分，从 AIGC 训练数据的

知识产权合规、AIGC 生成物的可版权性、AIGC 生成物侵犯他人著作权三方面全面剖析 AIGC 从产生到训练再到应用过程中的法律风险。第四章数据合规部分，包含训练数据合规、数据标注合规、个人信息合规、用户界面合规、数据出境合规、数据安全审计六方面内容，旨在为企业 AIGC 产品的研发、运营提供有效的合规指引。

本年刊由北京植德律师事务所合伙人及律师编著完成。其中，植德合伙人时萧楠及王妍妍等植德人工智能与数字经济行业委员会成员共同参与了本次《植德人工智能年刊（2023）》的编写。

版权声明

本年刊版权属于北京植德律师事务所并受法律保护，转载、摘编或以其他方式使用本年刊观点或文字的，应注明“来源：《植德人工智能年刊（2023）》”。编者保留追究违反者责任的权利。

目 录

一、 <u>投融资并购</u>	2
（一）中国 AIGC 产品的商业化路径	3
（二）AIGC 企业的公司治理结构	5
（三）美国出口管制对 AIGC 企业的影响	7
二、 <u>人工智能监管现状</u>	10
（一）《生成式人工智能服务管理暂行办法》解读	11
（二）深度合成算法备案	22
（三）《生成式人工智能服务 安全基本要求》（征求意见稿）解读	31
（四）《科技伦理审查办法（试行）》实施 AIGC 企业的应对措施	38
（五）《网络安全标准实践指南 ——生成式人工智能服务内容标识方法（征求意见稿）》解读	44
（六）中国生成式人工智能与深度合成监管差异与联系的探析	49

contents

三、知识产权合规 55

- (一) AIGC 训练数据的知识产权合规 56
- (二) AIGC 生成物的可版权性 64
- (三) AIGC 生成物侵犯他人著作权——以 AI 孙燕姿音乐为例 82

四、数据合规 88

- (一) 训练数据合规 89
- (二) 数据标注合规 92
- (三) 个人信息（含敏感个人信息）合规 93
- (四) 用户界面合规 96
- (五) 数据出境合规 99
- (六) 数据安全审计（含个人信息审计） 102

附录 1: AIGC 平台知识产相关条款附表 104

附录 2: 植德人工智能与数字经济行业委员会主要成员介绍 113



PART 1 :TOU RONG ZI BING GOU

第一部分 投融资并购



一、投融资并购

（一）中国 AIGC 产品的商业化路径

2023 年 ChatGPT 的爆火，在全球范围内引起了一场新一代人工智能技术的风暴。作为 ChatGPT 的研发组织，OpenAI 在项目孵化模式方面具有一定创新性。由早期的公益性质转向成熟后的营利性质，OpenAI 的商业模式贴合了一项研究从技术可行性到商业化论证的发展路径，这种商业模式不仅为 OpenAI 自身的可持续发展提供了支持，也可以为其他领域的科研成果转化提供一定的借鉴意义。

2015 年 12 月，OpenAI, Inc. 于美国特拉华州设立，为一家非营利公司 (Nonprofit corporation)。作为早期 OpenAI 的运营主体，OpenAI, Inc. 的资助人包括 Greg Brockman 等技术人员、Elon Musk 等高净值人士、以及 Amazon Web Services 与 YC Research 等机构。运营层面上，OpenAI, Inc. 是由董事会决定筹集资金的使用以及项目的研发等事项，资金将全部用来支付员工基本工资、研究耗材、仪器购买等。

2019 年 3 月，OpenAI 进行重组，由 OpenAI, Inc. 的董事会发起设立一家特拉华州的有限责任公司 (LLC)，并以该公司作为普通合伙人，与员工持股平台、投资人作为有限合伙人共同设立一家可以营利的有限合伙企业 OpenAI LP。重组后的 OpenAI，仍然由 OpenAI, Inc. 这一非营利实体的董事会控制及管理。

经过 OpenAI 从公益性到营利性的转变，其平台的重组演变恰如其分地服务了基础技术的研发进程。在早期，OpenAI, Inc. 作为公益性组织平台从资助人处筹集了启动资金，研究方向为通用人工智能 (AGI)，对数个技术路线进行持续深入研究。随着概念的

成熟，OpenAI LP 作为初创公司阶段的营利性组织被设立，与微软等投资人沟通合作，完成了从概念验证阶段向营利性的转化。这种模式符合技术从早期开发到成熟并商业化使用的发展路径，同时确保了在每个阶段均未出现资金缺口。正是由于这种平台的重组演变，才最终孵化出了 ChatGPT 项目。

而在中国，虽然包括北京市在内的地方政府颁布了支持社会资本入局设立概念验证基金，提出了一条“先导验证 + 股权投资”的路径^[1]，但由于没有配套的落地措施，且例如《公益事业捐赠法》、《慈善法》等法律规定的限制，在我国现行法律体系下，类似 OpenAI 这种“先导验证 + 股权投资”的模式在国内仍存在着较大障碍。

● 我国现行《慈善法》规定，慈善活动是指自然人、法人和其他组织以捐赠财产或者提供服务等方式，自愿开展的活动，其中就包括促进教育、科学、文化、卫生、体育等事业的发展。因此，我们理解，若社会资本将资金投入或发起公益性基金，并由该公益性基金将款项用于科学技术的研发，则可以被视为从事慈善活动，符合北京市科委提出的社会资金设立概念验证基金用于先导研发模式。然而，《慈善法》同时规定，慈善活动应当遵循非营利的原则。在“先导验证 + 股权投资”模式下，当概念验证项目成功产出科技成果并进行商业性转化后，理论上资助人的出资应当转换为初创企业的股权，因此这一后续操作可能使得公益性基金对科学技术的研发被认定为违反了非营利原则，从而使得资助人的捐助无法达成其公益目的并获得公益捐赠相关的税收优惠效果。

● 如果社会资本想就其向概念验证基金的捐助获取税收优惠，首先该基金会需要获得非营利组织免税资格，并在捐助时获得基金会或其他具有非营利组织免税资格的民办非企业单位开具的捐赠票据。然而，根据《关于公益性捐赠税前扣除有关事项的公告》第八条的规定，“附加对捐赠人构成利益回报的条件”属于违反规定接受捐赠，因为这不符合公益性条件；公益性社会组织也会因此被取消其公益性捐赠税前扣除资格。因此，在社会资本的公益性资助后续可能转化为出资的条件下，公益性社会组织无法向捐助入

^[1] 北京市科学技术委员会等机关于 2022 年 12 月 30 日发布了《标杆孵化器培育行动方案（2022-2025 年）》，明确支持以社会资本主导，而摆脱依托高校等公共平台设立概念验证基金，并提供了可选路径，其中第 17 条明确提出，“培育概念验证基金。鼓励大企业、投资机构、标杆孵化器等发起设立概念验证基金，深化‘先导验证 + 股权投资’模式，前期以公益性研发资金资助方式支持前沿颠覆性技术开展概念验证，项目验证成功并转化为创业企业后，前期资助转换为企业股权，带动社会资本投资前移，助推成果转化项目加快发展……。”这为我国社会资本设立概念验证基金提供了新的思路，即大企业、投资机构等社会资本可在前期以公益性研发资金资助先导验证，待到技术成熟、项目验证成功并转化为创业企业后，前期资助将转换为对创业企业的出资并获得企业股权。

出具捐赠票据，这意味着捐助人无法获得税收优惠。这将大大降低社会资本参与设立概念验证基金的积极性。

● 在社会资本通过概念验证基金资助项目时，应该会倾向于选定某一或某几个项目，而这实际上构成了定向捐助，在我国现行法律下可能产生以下两个风险：

一是根据《慈善法》的规定，捐赠人不得指定捐赠人的利害关系人作为受益人。虽然现行法律并未对“利害关系人”作明确定义，但是在社会资本的基金捐助可能转化为日后成果转化企业的股权出资的情况下，我们理解社会资本机构与基金会的研究人员如果作为后续初创企业的创始股东，实际上就构成了潜在的利益共同体，存在被定义为利害关系人的风险。

二是根据《民间非营利组织会计制度》的规定，社会资本对特定先导研究项目的捐赠存在会计上被处理为受托代理资产与受托代理负债的风险，而不能作为捐赠收入，也无法开具捐赠票据，因而存在无法取得税收优惠的风险。在实务中，捐赠人为了避免此种风险，往往会与基金会设立项目筛选机制，以在捐赠上达到表面不特定实际上特定的效果，但是这并不能完全消除被认定为受托代理的风险。

因此，目前的中国 AIGC 企业，尤其是从事大模型业务的 AIGC 企业，除脱胎于互联网大厂的企业外，绝大多数仍然走在一条较为传统的“创始团队开发产品设立公司股权融资”或者“高校实验室 / 教授团队科研高校技术成果转化设立公司股权融资”的商业化道路上。前者如百川智能、MiniMax，后者则如智谱AI（智谱AI的创始团队来自清华大学，清华大学的科技转化公司华控技术转移有限公司为公司设立时的股东之一）。

（二）AIGC 企业的公司治理结构

1. 境外 AIGC 企业 - 以 OpenAI 为例

以 OpenAI 为例，OpenAI 目前主要由一个“非营利实体”OpenAI. Inc. 及一个“有

限营利实体” OpenAI Global, LLC 组成，其余相关方则包括作为管理层持股主体的 OpenAI GP LLC、作为员工和其他投资人持股平台的持股平台 OpenAI AI, L.P.，以及 OpenAI 最大的少数投资人微软。

简单说来，OpenAI, Inc.（非营利实体）的董事会，目前从股权和管理权上控制着作为营利实体的 OpenAI Global, LLC。根据 OpenAI 官网的披露，董事会的首要信义义务（fiduciary duty）并非维护 OpenAI 投资人的权益，而是为了全人类；董事会的第一目标仍然是维护 OpenAI 的公益性纲领，即确保通用人工智能（AGI）惠及全人类。OpenAI 的纲领赋予了 OpenAI 董事会极大的自主权，董事会成员均由董事会自身（即其现有成员）决定任命。

而作为创始人的 Sam Altman 及联合创始人 Greg Brockman，并未如其他普通创业公司的创始人一般，在公司发展初期持有着公司的多数股权和 / 或投票权。据 OpenAI 官网，Sam Altman 仅通过其之前担任过主席和董事会主席的著名创投机构 Y Combinator 在 OpenAI 间接持有少量权益。

在 OpenAI 在 2023 年 11 月爆发董事会和创始人 Sam Altman 的冲突之前，OpenAI 的董事由三位 OpenAI 的员工和三位独立董事组成；在 Sam Altman 最终回归 OpenAI 后，OpenAI 的董事会目前由 Bret Taylor (SaaS 公司 Salesforce 的前任 CEO) 担任董事会主席，其他董事包括 Larry Summers（美国经济学家，曾任美国财长和）及 Adam D’ Angelo（问答社区 Quora 的 CEO）组成，此外 OpenAI 最大的少数股权投资人微软享有一席观察员席位。

在 2023 年底的 Sam Altman 罢免风波中，OpenAI 对 Sam Altman 的罢免是由 OpenAI 董事会中除了 Sam Altman 以及 Greg Brockman 以外的其他四名董事投票通过的。这四名董事已经构成了董事会的多数，考虑到最终 Sam Altman 的罢免生效，我们可以推断在 OpenAI 纲领性文件中对董事会决议的通过标准应无其他特别性约定（比如在投资交易中常见的保护性条款，即某些事项除多数董事同意之外，仍需包含投资人董事或者特定创始人董事的同意方可视为通过）。

2. 境内 AIGC 企业 - 以智谱 AI 为例

反观境内 AIGC 企业，则基本仍为传统的公司治理架构。例如风头正劲的新晋独角兽企业智谱 AI，公开渠道可以看到，智谱 AI 目前并未搭建红筹架构，其目前拥有二十余家投资人股东，创始团队合计持有公司约 33% 的股权的投票权，合计仍为公司第一大股东；董事会成员 8 人，其中 5 人应为公司方董事，另外 3 人为投资人委派的董事，创始团队仍然从人数上控制着董事会。按照境内企业融资的惯例，投资人应该是在股东会决议事项和董事会决议事项上设置了特别性约定，例如涉及企业发展的重大事项，或是企业经营中金额较高影响较大的事项上，需经一定比例的投资人同意方可视为通过。

另一独角兽企业 MiniMax 则已经搭建了红筹架构，其最终融资主体设立在开曼，境内运营主体均为融资主体通过香港子公司间接 100% 持有的 WFOE。公开渠道无法查询到 MiniMax 的现有董事会和股东信息，不过从媒体报道中可以了解到其大致上和智谱 AI 一样，拥有多家投资人股东；按照惯例，仍处在中早期阶段的公司，较为重要的投资人应有权委派董事，不过创始团队仍会在人数上控制着董事会；而在股东会层面，即使创始团队的股权比例较低，也可通过 AB 股的设计在投票权上保证其自身对公司的控制，并通过股东会和董事会层面的特别决议事项给予投资人在企业重要事项上的话语权。

（三）美国出口管制对 AIGC 企业的影响

随着 AIGC 展现出强大的生产力变革能力与发展潜力，中美两国在 AIGC 技术领域的竞争日趋激烈，这场竞争不仅体现两国在技术创新和行业布局的战略争夺上，还直接体现在美国针对中国展开的对 AIGC 相关设备和技术的出口限制。

2022 年 10 月 7 日，美国商务部工业与安全局（BIS）发布了《对向中国出口的先进计算和半导体制造物项实施新的出口管制》（“1007 规则”），对中国半导体和先进计算行业实施出口管制措施。2023 年 10 月 17 日，BIS 在 1007 规则的基础上，进一步更新管

制内容，推出《实施额外出口管制：某些先进计算物项、超级计算机与半导体终端应用，更新和修改临时最终规则》（“1017 规则”），全面升级对华半导体出口管制。此外，BIS 还以标注“脚注 4”的形式新增 13 家中国企业进入实体名单。

本次修订对 AIGC 企业的影响主要体现在以下几个方面：

- 限制 AIGC 企业获取高性能计算芯片、相关计算机组件和特定半导体制造设备。1017 规则新增对高性能计算芯片、相关计算机组件和特定半导体制造设备的管制，这将对 AIGC 企业的芯片设计、制造、测试等环节造成影响。

- 限制美国人对 AIGC 企业的技术支持。1017 规则限制美国人对中国境内的半导体制造工厂开展芯片开发或生产提供支持，这将对 AIGC 企业的芯片研发、制造等环节造成影响。

- 扩大对 AIGC 企业的域外管辖范围。1017 规则扩大对中国和特定第三国的先进计算物项和用于超级计算机最终用途的特定外国制造物项的域外管辖范围，这意味着即使 AIGC 企业在境外采购上述物项，也可能受到美国出口管制的制约。

- 增强对 AIGC 企业的最终用途管控。1017 规则增强对超级计算机和半导体制造特定最终用途的管控，这将对 AIGC 企业的人工智能模型训练、部署等环节造成影响。

1. 1017 规则的具体修订

物项管控 1017 规则通过修改 ECCN 3A090、3B090 和 3B001 等 ECCN 编码所涵盖的物项范围，加强了对可用于 AI 训练模型的相关芯片的管控。具体而言，1017 规则修改了以下内容：（1）修改了 ECCN 3A090 的管控参数及物项描述，并提供了相应计算公式。这些调整使得 BIS 能够更有效地管控可用于 AI 训练模型的相关芯片，包括防止相关方通过大批量采购并组合使用性能较低的 AI 芯片来实现与单一受限芯片相同的算力以规避管制；修改了 ECCN 3B001 和 3B090，将管控限制在对先进芯片制造而言核心必要的关键设备和相关软件与技术上；以及（2）修改区域稳定与国家安全等两个管控理由的适用范围和条件，将 D:1、D:4、D:5 组别国家全部纳入 EAR742.6(a)(6) 所适用的出口目的地国家范围内。

活动管控 1017 规则限制美国主体为中国实体自主开展的先进半导体生产制造活动提供支持，支持内容包括技术转让、培训、咨询、协助等。此外，新规定下，涉及 ECCN 编码 3A090 芯片的出口、再出口到中国大陆、香港、澳门及在这些地区的转移将实行“推定拒绝”的许可证管理。这些限制同样适用于总部位于澳门或受美国禁运国家的公司及其子公司。

除上述修订外，BIS 同时宣布 13 家先进制程集成电路开发相关的中国实体被列为实体清单脚注 4 实体，向该等实体出口任何受《出口管理条例》(Export Administration Regulations, “EAR”) 管辖的管控物项均需要向 BIS 申请许可，而 BIS 将就该等许可申请推定拒绝。

2. AIGC 企业的应对措施：识别目前使用芯片是否落入管控；进一步规范采购合规性

美国出口管制将导致部分 AIGC 企业无法获得训练大模型所需的算力芯片，从而影响其大模型训练的能力。根据 1017 规则，ECCN 3A090 类别项下的管控芯片具有“总处理性能”大于等于 4800 或“性能密度”大于等于 5.92 的特征。目前，包括英伟达、AMD、华为等在内的多家芯片厂商都生产符合该等管控标准的芯片。

对于使用上述落入出口管制范围内芯片进行大模型训练的 AIGC 企业，短期内或许可以通过利用存量资源的方式解决，例如使用现有的存量芯片以及租赁算力。而对于目前供应链尚未落入管制范围的企业，必须审慎评估其芯片供应情况，并密切关注芯片采购合规性避免影响到企业获取芯片的能力。

为了加强对出口管制的执行，BIS 发布了“Know Your Customer Guidance and Red Flags”指南，列出了一些可疑订单的情形，本次修订在原有的 13 项“Flags”基础上新增 5 项，意味着 AIGC 企业在购买芯片时可能会受到更严格的审查。AIGC 企业在购买芯片时，应注意避开可能被识别为可疑订单的情形，以避免潜在的供应中断以及被进一步纳入实体清单名单。

PART 2 :REN GONG ZHI NENG JIAN GUAN XIAN ZHUANG

第二部分 人工智能监管现状



二、人工智能监管现状

（一）《生成式人工智能服务管理暂行办法》解读

2023 年 ChatGPT 的爆火，在全球范围内引起了一场新一代人工智能技术的风暴。作为 ChatGPT 的研发组织，OpenAI 在项目孵化模式方面具有一定创新性。由早期的公益性质转向成熟后的营利性质，OpenAI 的商业模式贴合了一项研究从技术可行性到商业化论证的发展路径，这种商业模式不仅为 OpenAI 自身的可持续发展提供了支持，也可以为其他领域的科研成果转化提供一定的借鉴意义。

2015 年 12 月，OpenAI, Inc. 于美国特拉华州设立，为一家非营利公司 (Nonprofit corporation)。作为早期 OpenAI 的运营主体，OpenAI, Inc. 的资助人包括 Greg Brockman 等技术人员、Elon Musk 等高净值人士、以及 Amazon Web Services 与 YC Research 等机构。运营层面上，OpenAI, Inc. 是由董事会决定筹集资金的使用以及项目的研发等事项，资金将全部用来支付员工基本工资、研究耗材、仪器购买等。

2022 年年底 OpenAI 公司发布的 ChatGPT 突然以飞快的速度在全球同步发展，并且推出两个月月活用户突破 1 亿，成为了史上用户增长速度最快的消费级应用程序。但之后伴随而来的是，全球监管机构针对以 ChatGPT 为代表的生成式人工智能带来的数据安全、网络安全、个人信息保护、知识产权保护等各类风险的高度关注。意大利数据保护局 (Garante per la Protezione dei Dati Personali) 在 2023 年 3 月底暂时禁用 ChatGPT，并对该工具涉嫌违反隐私规则展开调查之后，美国、加拿大、英国、德国、法国、西班牙、日本等多个国家的监管机构均对生成式人工智能所带来的数据问题、知识产权

问题等进行调查、发布相关指引或者表达了担忧。

在这样的背景下，2023年4月28日中共中央政治局召开会议强调，“要重视通用人工智能发展，营造创新生态，重视防范风险”。国家网信办于2023年4月11日公布的《生成式人工智能服务管理办法（征求意见稿）》（以下称“《征求意见稿》”），在经过三个月的酝酿与完善后，完成了国家互联网信息办公室审议，国家发展和改革委员会、教育部、科学技术部、工业和信息化部、公安部、国家广播电视总局七部委同意的立法进程，成为全球第一个全面监管生成式人工智能的立法文件，出台的《生成式人工智能服务管理暂行办法》（以下称“《暂行办法》”）于2023年7月13日公布，2023年8月15日起施行。

习近平总书记在《习近平总书记关于网络强国的重要思想概论》第五章“筑牢国家网络安全屏障”提到，“不发展是最大的不安全，绝不能因为安全问题而拒绝发展^[2]。”《暂行办法》的出台，一方面，给出了我国针对AIGC快速发展的中国答案，贡献了中国智慧，是落实习近平总书记网络强国重要思想的生动体现，另一方面，《暂行办法》在治理思路上与之前已经生效的《互联网信息服务算法推荐管理规定》（以下称“《算法推荐管理规定》”）《互联网信息服务深度合成管理规定》（以下称“《深度合成管理规定》”）是一致的，但在治理制度方面，针对生成式人工智能技术提出了进一步明确的要求。本所将在《暂行办法》的监管原则之下，就其规定的具体治理制度进行梳理。

1. 适用范围（第2条）

生成式人工智能在中国快速发展以来，什么样的生成式人工智能产品 / 服务，如何使用生成式人工智能产品 / 服务，受到怎样的中国法律监管一直以来都是行业密切关注的问题。为了反馈实践中的要求，以及促进行业发展，《暂行办法》相较之《征求意见稿》进一步明确以及限缩了适用范围：向中华人民共和国境内公众提供生成式人工智能服务。

^[2] 参见《习近平总书记关于网络强国的重要思想概论》，中央网络安全和信息化委员会办公室，人民出版社2023年7月，P91。

1.1 适用范围条款中特定词语的含义

(1) 向中国境内

在本条规定下，境内主体向境外提供人工智能服务的，无需适用《暂行办法》。但需要提醒注意的是，如果是中国境内的生成式人工智能服务在向境外提供服务过程中，即使不需要适用《暂行办法》，仍需要注意是否会产生数据出境等合规问题。

而境外主体向中国境内公众提供服务的，同样需要适用《暂行办法》。此外，《暂行办法》还在条文最后特别明确境外主体向中国境内提供服务应当符合外商投资相关规定。

(2) 面向公众

“面向公众”的服务应当依据《暂行办法》依法治理，是由于人工智能生成物目前尚存在生成虚假信息（目前全球多个国家陆续发生 ChatGPT 生成的虚假信息导致的侵权或者犯罪案件）的风险，而进一步通过各种类型的“具有舆论与社会动员能力的服务”可能产生扩散效应，扰乱社会正常秩序发展，因此“面向公众”提供的服务，特别是针对“具有舆论属性或社会动员能力的服务”是监管的核心重点。

而“公众”通常应当被理解为“不特定多数”，因此当一项服务面向不特定多数时，无论其面向的是不特定多数的企业还是不特定多数的自然人，均构成“面向公众”。而向不特定多数的企业客户提供生成式人工智能服务，目前基本被理解为不在本规定适用范围内。

另一方面，《暂行办法》规定，研发和应用生成式人工智能技术，只要未向境内公众提供服务的，就无需适用《暂行办法》，也就是如果仅在公司内部研发或者以提高办公效率之目的，公司内部使用则无需适用。这无疑为企业内部、垂直领域内的“面向非公众”的“研发、应用”（并非使用其提供服务）提供了宽松的监管空间。这与实践中大量企业存在使用生成式人工智能提升企业效率的需求相呼应，是《暂行办法》积极响应现实需求的体现。但同时需要注意，企业内部、垂直领域内的“面向非公众”的“研发、应用”即使无需适用《暂行办法》，但特别是在制作、发布、使用相关生成内容时，需要遵守包括《网络信息内容生态治理规定》在内的法律法规。

(3) 提供服务

“提供服务”是与《算法推荐管理规定》《深度合成管理规定》中的“算法推荐服务”“深度合成服务”互相呼应，同时排除了进行“研发、应用”，进一步限缩风险较小的“面向非公众”的适用范围。

2. 适用主体：生成式人工智能服务提供者（第 22 条（二））

在前述适用范围内适用《暂行办法》的主体，被明确规定为生成式人工智能服务提供者，是指利用生成式人工智能技术提供生成式人工智能服务（包括通过提供可编程接口等方式提供生成式人工智能服务）的组织、个人。需要强调的是《暂行办法》并未如《深度合成管理规定》一样，区分技术支持者和服务提供者，同时明确规定“通过可编程接口等方式提供服务”的主体同样适用《暂行办法》，例如 2023 年 6 月 24 日公布的已完成算法备案的技术支持者均是以 API 接口方式提供服务，因此可以看出《暂行办法》下技术支持者与服务提供者均受到同样的监管要求。

我们同样需要注意的是《暂行办法》目前的立法定位是“暂行”，并且仅有 24 条，仅规定了重大的监管原则以及目前亟待明确的事项，同时在泱泱众多的人工智能技术中生成式人工智能只是其中一种，因此后续的立法计划中的《人工智能法》很有可能会进一步全面、完善地针对人工智能产业链上的不同主体提出不同的监管要求，这同样也是目前《暂行办法》中明确的“分类分级”监管原则的要求。

3. 促进创新篇（第 5 条、第 6 条）

人工智能全产业链，包括从基础层的算力、芯片、训练数据到神经网络算法模型，向公众提供服务或产品的应用层，《暂行办法》均依据其不同的行业特点，制定不同的

鼓励方向：

- 应用层：鼓励创新应用，生成积极健康的优质内容，构建应用生态体系；
- 技术层以及基础层：鼓励人工智能核心技术领域的自主创新，以及开展平等互利的国际交流与合作，参与国际标准制定；
- 基础层：推动基础设施和公共训练数据资源平台建设；促进算力资源协同共享，提升算力资源利用效能。推动公共数据分类分级有序开放，扩展高质量的公共训练数据资源。鼓励采用安全可信的芯片、软件、工具、算力和数据资源。

可以看出，《暂行办法》针对应用层重点鼓励生成积极健康的优质内容，而对技术层和基础层主要在促进加速基础设施（算力、数据等）建设以及鼓励平等的国际合作，参与国际标准制定。

4. 依法治理篇

4.1 分类分级监管（第3条、第16条）

《暂行办法》在《征求意见稿》上增加确定了针对生成式人工智能包容审慎、分类分级的依法治理监管原则，这其实是延续了《中华人民共和国网络安全法》、《中华人民共和国数据安全法》中的分类分级规定，有利于生成式人工智能与相关法律法规的衔接。

落实到具体条款，《暂行办法》在“生成式人工智能服务提供者”的框架下区分出“具有舆论属性或者社会动员能力的生成式人工智能服务的”产品，并对其提出了更为严格的监管要求：应当按照国家有关规定开展安全评估。

另一个层面，“利用生成式人工智能服务从事新闻出版、影视制作、文艺创作等活

动的”，未来将会由相关监管部门，包括国家互联网信息办公室、国家广播电视总局等相关主管部门另行规定，也可以看出是针对不同类型的生成式人工智能分类监管的体现。

但整体而言，《暂行办法》尚未建立如欧盟《人工智能法案》近似的分级监管体系（将所有人工智能产品/服务基于对人类健康安全等的影响区分为不可接受的风险、高风险、有限风险以及最低风险。其中前三类将受到法案的监管，第四种被称为“最小风险（包括例如具备 AI 功能的视频游戏或垃圾邮件过滤器等）”，根据欧盟委员会的说法，这一类别实际上涵盖了绝大多数人工智能系统。），一个重要的原因可能在于，立法者将这一任务留给了未来要制订的《人工智能法》完成，而针对众多人工智能服务/产品中一类的生成式人工智能，目前仅采取如上分类分级措施。

4.2 监管治理重点

（1）针对人工智能技术特点的监管要求（第4条）

由于人工智能技术可以通过算法进行引导，如果训练数据、模型中存在歧视、偏见等有害信息，通过算法进行扩散输出将比通常的信息传播导致的不良影响更为严重，基于此《暂行办法》要求：

- 坚持社会主义核心价值观，不得生成违法违规的内容；
- 算法设计、训练数据选择、模型生成和优化、提供服务过程中应当采取防止算法歧视的措施；
- 尊重知识产权、他人合法权益；
- 提高生成式人工智能服务透明度，提高生成内容的准确性和可靠性。

（2）针对不同的关键环节的监管要求（第7条、第8条、第14条）

数据训练阶段：提供者应当依法开展预训练、优化训练等训练数据处理活动：

- 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、多样性等；
- 其他监管要求。

数据标注阶段：数据标注，提供者应当制定标注规则；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。

内容生成阶段：提供者应当在生成内容出现违法内容采取模型优化训练等措施进行整改。

需要注意的是，由于实践中生成式人工智能提供者的算法模型中包括大量第三方开源模型、数据库，特别是境外开源模型、数据库，因此如何认定“使用具有合法来源的数据和基础模型”将直接影响我国的生成式人工智能是否可以站在高个子身上向上发展。我们结合《暂行办法》第20条，境外向境内

提供生成式人工智能服务不符合法律行政法规和本办法规定的，国家网信办部门应当进行处置，反之如果境外向境内提供的服务符合我国相关法律法规则可以被使用。因此，这提醒使用第三方开源模型、数据库的生成式人工智能提供者，应当在接入、使用三方开源模型、数据库时进行合规评估 / 内容审核等，保证使用是符合中国法律法规要求的。

其次，数据质量的提升也无疑是生成式人工智能服务提供者的难题，如何在庞大的训练数据中核实其数据的真实性、准确性、客观性，同时要保证多样性。这一难题通常可以通过确定数据来源的方式进行解决。

4.3 生成式人工智能服务提供者的服务规范

(1) 衔接生成式人工智能服务提供者与其他法律主体之间的关系，明确其依据其他法律规定的法律责任（第9条）

- 生成式人工智能服务提供者承担网络信息内容生产者的法律责任，应当遵守《网络信息内容生态治理规定》第二章的规定；

- 涉及个人信息的，生成式人工智能服务提供者承担个人信息处理者责任，应当遵守《个人信息保护法》中赋予个人信息处理者的法律责任；

- 生成式人工智能服务提供者应当与其用户 / 客户签订服务协议，依据《民法典》明确双方权利义务。

(2) 保护未成年人（第10条）

生成式人工智能的使用将大幅降低学生通过自己的努力获取知识，而仅凭询问人工智能即可轻易获得作业答案，甚至论文，这无疑已经使很多人开始担忧未成年人的教育问题。日本文部省（教育部）就今年6月22日公布了一项指导法案，允许在小学、初中和高中有限度地使用 ChatGPT 类生成式人工智能，但也明确规定限制措施，例如小学生（13岁以下）使用 AI 时需要特别谨慎。而任何由 AI 生成的作业，一经判定，将被视为作弊。

而《暂行办法》也同样规定生成式人工智能服务提供者应当采取有效措施防范未成年人用户过度依赖或者沉迷服务。但与日本文部省做法的区别在于，中国监管部门将这一责任负之于服务提供者，要求其采取有效措施防沉迷或者过度依赖，而日本是通过劝告学生、家长以及学校的方式避免未成年人在学习中过度使用。

(3) 对 Prompt 的保护、生成内容的标识（第11条、第12条）

针对用户使用生成式人工智能服务过程中，需要向服务者发送 Prompt 才可以获得其反馈的信息，而发送的 Prompt 中经常会包括使用者的个人信息，因此《暂行办法》明确要求生成式人工智能服务提供者应当对 Prompt 进行保护，并且对使用者发送的 Prompt 中的个人信息的收集应符合《个人信息保护法》的规定。

另一方面,《暂行办法》衔接《深度合成管理规定》,要求提供者对图片、视频等生成内容进行标识。这里需要注意的是,《深度合成管理规定》所规定的标识包括两种类型:

- 深度合成服务提供者对使用其服务生成或者编辑的信息内容,应当采取技术措施添加不影响用户使用的标识。这里是指在生成内容中添加隐式标识。

- 在生成或者编辑的信息内容的合理位置、区域进行显著标识,向公众提示深度合成情况。这里是指在提供服务时,在相关位置或者区域提示告知用户深度合成情况。

而《暂行办法》关注的明显是前者。隐式标识的强制要求加入,无疑可以对利用人工智能进行违法犯罪的活动进行充分举证,以及可以防止人们被不良分子使用人工智能生成内容诈骗等。

(4) 生成违法内容的处置、预防措施(第14条、第15条)

生成式人工智能服务提供者在发现其服务生成违法内容时,应当采取停止生成、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。

在2023年4月11日发布的《征求意见稿》中,网信办尚采取要求服务提供者发现违法生成内容后3个月内对算法进行整改防止再次生成违法内容的方式,但《暂行办法》已经修改为更为柔和的“采取模型优化训练等措施进行整改”,删除了3个月的时间限制,也删除了“防止再次发生”的结果限制,大大加强了该条款的可落地性。

但即使在目前的条款下,生成式人工智能服务提供者仍需要建立一套完整的处置、预防措施,包括发现不同的违法内容应当采取不同的处置机制。

在处置机制的前提下,《暂行办法》也要求服务提供者建立健全投诉举报机制。

4.4 监督手段

- 安全评估与算法备案(第17条)

《暂行办法》第17条规定,“提供具有舆论属性或者社会动员能力的生成式人工智能服务的,应当按照国家有关规定开展安全评估,并按照《互联网信息服务算法推荐管理规定》履行算法备案和变更、注销备案手续”。其保留了《征求意见稿》中的安全评

估算法备案的要求，但其完成时间修订为“按照国家有关规定开展”，不再强行要求两者均在服务上线前完成。

但需要注意的是，该条款中的“安全评估”目前最为明确的，有法律依据的安全评估是指依据《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》完成的安全评估，目前由公安部门通过其线上系统完成。但在此之外还需要注意的是，自 2023 年 2 月开始，各地网信办就开始执行的“双新评估”（新技术新应用安全评估）已经正式变更为“大模型备案”，并且明确由各省市级网信办负责办理，是否需要进行“大模型备案”需要单独与所在地网信办沟通以及确定。

另一方面，算法备案则继续按照《深度合成管理规定》的要求，提供服务之日起十个工作日内完成相关算法备案手续。因此相关企业应当尽快评估是否需要完成算法备案工作手续，以及具体备案角色（服务提供者或者技术提供者）。

5. 企业的应对措施

综合以上分析，一家企业判断是否需适用《暂行办法》的监管要求，可以从以下六个步骤逐步分析和判断：



图 2-1：企业适用《暂行办法》判断步骤图

(1) 第一步：确认主体

结合《暂行办法》对“生成式人工智能服务提供者”的定义，明确企业自身是否属于提供者的范畴。提供者包括《深度合成管理规定》的技术支持者、服务提供者两类主体，及直接提供生成式人工智能服务和通过可编程接口等方式间接提供服务两种类型。

(2) 第二步：明确定位

明确自身是否向“公众”提供生成式人工智能服务，且“公众”应当不仅理解为 C 端用户，还应当包括不特定的 B 端客户，并且确定 B 端客户是否会向一般公众提供生成式人工智能服务，明确自身在产业链中的角色。

(3) 第三步：审查对象

明确提供的服务是否面向公众，还是在内部研发或工业应用等目的使用：

- 公众的定义为“不特定多数”；
- 若服务为面向不特定多数企业提供，则符合该适用要件；
- 若服务为面向不特定多数自然人提供，也符合该适用要件；
- 若服务仅为在内部研发、内部办公管理或工业建模设计等目的引入，则不符合该适用要件，排除适用《暂行办法》。

(4) 第四步：判断涉外

明确提供的服务是否均在境内，还是有在境外的情况：

- 核心判断因素：“提供的服务面向境内”；
- 若均在境内，则符合该适用要件；
- 若为境外实体，在境外或境内运营，向境内提供服务，也符合该适用要件；
- 若为境内实体，在境内或境外运营，向境外提供服务，不符合该适用要件，排除适用《暂行办法》；

● 若为境外实体，在境内或境外运营，向境外提供服务，不符合该适用要件，排除适用《暂行办法》。

情形	面向不特定多数对象(企业或个人)	境内实体	境内提供服务	是否适用
情形一	是	是	是	是
情形二	是	否	是	是
情形三	是	是	否	否
情形四	是	否	否	否
情形五	否	是/否	是/否	否

表 2-1 企业适用《暂行办法》判断表

(5) 第五步：明确外资

明确提供者提供服务是否属于外商投资的范畴，如涉及，则应符合外商投资相关法律、行政法规的规定的要求；

(6) 第六步：确认资质

明确服务内容是否涉及其他许可或者备案。

(二) 深度合成算法备案

1. 算法备案的内容

1.1 算法备案产品适用范围

目前与算法备案相关的法律规定包括《算法推荐规定》《深度合成规定》《AIGC 暂行办法》。因此，首先梳理这三部部门规章规定的适用范围，理解什么范围下的 AIGC 相关 APP 产品需要适用其进行算法备案。

● 《算法推荐规定》

第二条 在中华人民共和国境内应用算法推荐技术提供互联网信息服务，适用本规定。法律、行政法规另有规定的，依照其规定。前款所称应用算法推荐技术，是指利用生成合成类、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。

● 《深度合成规定》

第二条 在中华人民共和国境内应用深度合成技术提供互联网信息服务，适用本规定。法律、行政法规另有规定的，依照其规定。

● 《AIGC 暂行办法》

第二条 利用生成式人工智能技术向中华人民共和国境内公众提供生成文本、图片、音频、视频等服务（以下称生成式人工智能服务），适用本办法。

我们汇总前述三部法律规定的适用范围：

（1）向中国境内

在前述各规定下，境内主体向境外提供 AIGC 服务的，无需适用前述各项规定进行算法备案。但需要提醒注意的是，如果是中国境内的 AIGC 服务在向境外提供服务过程中，即使不需要适用《暂行办法》，仍需要注意是否会产生数据出境等合规问题。

而境外主体向中国境内公众提供服务的，同样需要适用各项规定完成算法备案。

（2）面向公众

在《AIGC 暂行办法》中特别强调了“面向公众”的服务，因此如果仅在公司内部研发或者以提高办公效率之目的，公司内部使用则无需适用。但同时需要注意，如果向企业提供 AIGC 服务，该企业再使用该服务向公众提供服务时，仍属于向公众提供服务，应当作为技术支持者适用《算法推荐规定》《深度合成规定》《AIGC 暂行办法》完成算法备案。

(3) 提供服务

“提供服务”在结合前述三项规定中的“算法推荐服务”“深度合成服务”“生成式人工智能服务”后，应当被区分为技术支持者和服务提供者。

1.2 算法备案主体

根据《深度合成规定》，需要完成算法备案的主体为：

- 服务提供者：具有舆论属性或者社会动员能力的深度合成服务提供者
- 技术支持者：任何形态的技术支持者

其中如何判断是否构成具有舆论属性或者社会动员能力的深度合成服务提供者，一直是实务中的要点。根据《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》第2条规定，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，以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的均属于具有舆论属性或者社会动员能力的互联网信息服务，当该信息服务为算法服务时，就构成具有舆论属性或者社会动员能力的深度合成服务提供者，从而需要进行深度合成算法备案。

而如何判断需要完成算法备案的技术支持者可以通过已经完成算法备案的主体公示信息进一步判断。在已经公示的《备案清单》中可以看出目前备案的“技术支持者”均是向企业端客户提供算法服务，并非直接向公众提供算法服务的服务方式，也就是说向企业端客户（企业端再行向公众提供算法服务）提供服务的算法服务应当以“技术支持者”的角色进行算法备案。

1.3 算法备案对象

需要完成算法备案的对象是算法，也就是说如果一个 APP 产品中如果有多个算法，需要完成多个算法备案；并且其也与 APP 的主要业务无关，主要 APP 产品中包括相应算法，就应当依据相关规定完成算法备案。

1.4 算法备案流程

备案网址：互联网信息服务算法备案系统，其网址为 <https://beian.cac.gov.cn>。

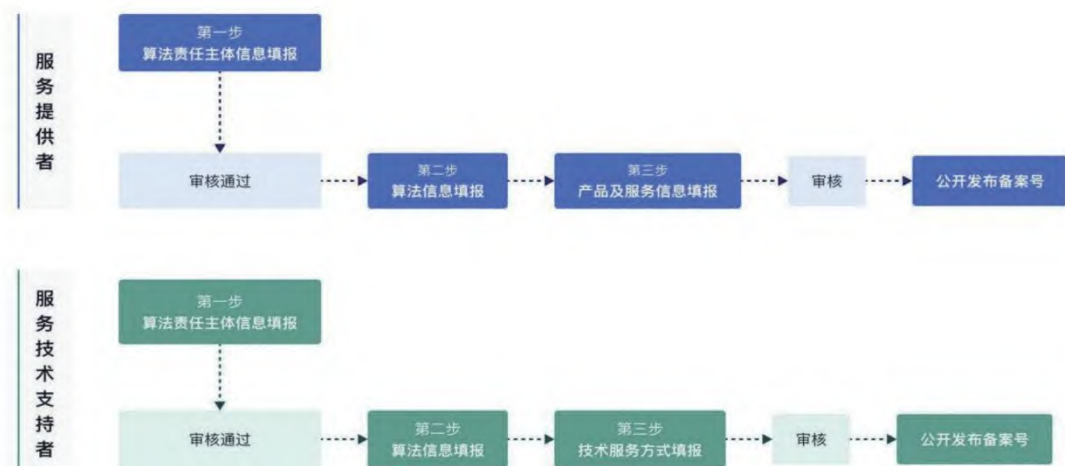


图 2-2 生成式人工智能服务提供者备案流程图

2. 第一批《境内深度合成服务算法备案清单》解读

2023 年 6 月 23 日，国家互联网信息办公室发布了第一批《境内深度合成服务算法备案清单》（以下称“《深度合成算法备案清单》”），包括百度、阿里、腾讯、抖音、科大讯飞等公司的 41 个深度合成算法已经完成算法备案，本部分将就首批算法备案清单进行解读。2023 年 9 月，2024 年 1 月，国家互联网信息办公室相继发布了第二批和第三批算法备案清单，保持了与首批备案清单相同的备案原则。

2.1 深度合成算法备案的服务角色

首次公布的《深度合成算法备案清单》依据《互联网信息服务深度合成管理规定》第 19 条，深度合成服务算法备案主体包括具有舆论属性或者社会动员能力的深度合成服务提供者，以及深度合成服务技术支持者应当完成算法备案相关手续。因此进行深度合成算法备案应当首先区分自己在算法服务提供中的服务角色，以确定是依据服务提供者还是依据技术支持者进行算法备案。

关于如何判断是否构成具有舆论属性或者社会动员能力的深度合成服务提供者和深度合成服务技术支持者的内容已经在本部分的 1.2 中已有论述，此处不再重复。

算法名称	主体名称	主要用途
火山引擎智能美化特效算法	北京抖音信息服务有限公司	应用于图像生成场景，服务于图像编辑类的企业端客户，通过 API 提供人脸图像美颜功能。
出门问问语音合成类算法-2	出门问问信息科技有限公司	应用于音频生成场景，服务于音频、视频制作类的企业端客户，通过 API 提供文本转语音功能。
达摩院开放域自然对话合成算法	阿里巴巴达摩院（杭州）科技有限公司	应用于对话生成场景，服务于智能对话类的企业端客户，利用对话意图理解、对话内容生成等技术，通过 API 提供对话生成功能。
达摩院图像合成算法	阿里巴巴达摩院（杭州）科技有限公司	应用于图像生成场景，服务于图像合成类的企业端客户，通过 API 提供文字信息转化为图像信息的功能。
达摩院交互式多能型合成算法	阿里巴巴达摩院（杭州）科技有限公司	应用于开放域多模态内容生成场景，服务于问答、咨询类的企业端客户，通过 API 提供根据用户输入生成多模态信息的功能。
讯飞星火认知大模型算法	科大讯飞股份有限公司	应用于开放域对话生成场景，利用文本、代码、prompt 数据及用户反馈数据训练 AI 大模型，服务于问答、咨询类的企业端客户，通过 API 提供文本生成服务。
腾讯云人脸融合算法	深圳市腾讯计算机系统有限公司	应用于人脸图像、视频生成场景，服务于图像、视频编辑类的企业端客户，通过 API 提供用户面部特征融合功能。
腾讯云语音合成算法	深圳市腾讯计算机系统有限公司	应用于文本转语音场景，服务于音频、视频合成类企业端客户，使用已授权的声优声纹信息，通过 API 提供文本转语音功能。

表 2-2 服务技术支持者备案算法概览表

2.2 深度合成算法备案的应用产品

(1) 服务提供者的应用产品

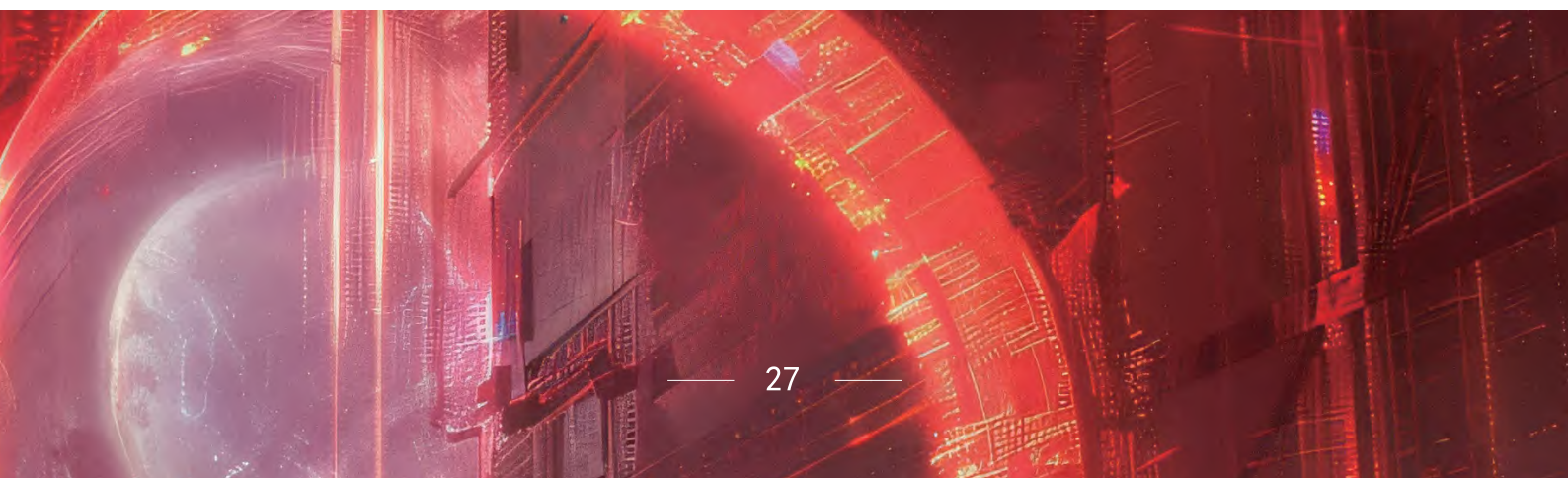
根据首次《深度合成算法备案清单》可以看出（如下为例），以服务提供者填报的算法备案，应用产品可以是 APP、小程序和网页，另一方面也是说明无论是以 APP、小程序还是网页的形式提供算法服务均需要根据相关规定进行算法备案。

算法名称	主体名称	应用产品
大麦小蜜智能客服算法	北京大麦文化传媒有限公司	大麦（APP）
DraftAi 绘图生成合成算法-1	图形起源（北京）科技有限公司	Draft（网站）
菜鸟物流智能客服算法	杭州菜鸟物流信息科技有限公司	菜鸟（APP、小程序）

表 2-3 服务提供者应用产品类型表

(2) 服务技术支持者的应用产品

根据目前本所实务中与相关公司的沟通，以服务技术支持者的身份进行算法备案的公司由于多使用 API 接口直接向 B 端客户提供服务，都会涉及不知道如何填写应用产品的问题，而本次的《深度合成算法备案清单》也明确公示，以服务技术支持者身份提交的算法备案无需填写应用产品。



算法名称	主体名称	应用产品
火山引擎智能美化特效算法	北京抖音信息服务有限公司	---
出门问问语音合成类算法-2	出门问问信息科技有限公司	---
达摩院开放域自然对话合成算法	阿里巴巴达摩院（杭州）科技有限公司	---
达摩院图像合成算法	阿里巴巴达摩院（杭州）科技有限公司	---
达摩院交互式多能型合成算法	阿里巴巴达摩院（杭州）科技有限公司	---
讯飞星火认知大模型算法	科大讯飞股份有限公司	---
腾讯云人脸融合算法	深圳市腾讯计算机系统有限公司	---
腾讯云语音合成算法	深圳市腾讯计算机系统有限公司	---

表 2-4 服务技术支持者算法应用产品表

2.3 深度合成算法备案中涉及的大模型算

大模型算法备案一直被行业重点关注，由于大模型算法的复杂性，其中涉及大量开源模型，可能还会涉及境外数据内容，那么相关大模型是否可以依据《互联网信息服务

《深度合成管理规定》完成合规工作，顺利进行算法备案，合规开展大模型业务将会是关乎中国大模型行业发展的关键点。

首次备案的大模型算法包括了明确以“大模型”名称进行备案的：

- 百度 PLATO 大模型算法
- 科大讯飞的星火认知大模型算法

以及通过用途可以被理解为大模型的算法：

- DraftAi 绘图生成合成算法 -1
- 智谱 ChatGLM 生成算法
- 百度文生图内容生成算法
- 出门问问语音合成类算法 -2
- 达摩院开放域自然对话合成算法
- 达摩院图像合成算法等

通过前述大模型的顺利备案可以看出，目前网信办的相关合规要求，在算法备案部分并不存在实质性对大模型行业的阻碍，相关使用深度合成算法的公司还应积极配合网信部门，尽快完成算法备案工作，保证业务的合规开展。

2.4 备案算法个数

根据本次的《深度合成算法备案清单》，可以进一步看出网信办要求的算法备案并非以公司主体或者 APP、网站等应用产品作为算法备案的个数判断标准，而是以实际的算法个数作为判断标准，因此相关公司请按照实际使用的算法个数进行算法备案。

2.5 不同算法类型的备案情况汇总

首次的《深度合成算法备案清单》中，包括了如下不同类型的算法类型：

- 文本类算法：例如智谱 ChatGLM 生成算法（应用于对话生成场景，根据用户输入

的文本内容，应用对话模型，生成对话文本回复）、美团在线智能客服算法（应用于电话智能客服场景，生成用户问题的文本智能回答）、百度 PLATO 大模型算法（应用于对话生成场景，基于飞桨深度学习框架和对话语料训练的对话模型，生成开放域的文本回答）等

- 文字转图像类算法：例如百度文生图内容生成算法（应用于图像生成场景，利用图文匹配神经网络模型和扩散生成模型，生成与输入文本语义一致的图像）等

- 图像视频类算法：例如剪映特效算法（应用于图像、视频生成场景，基于人像理解技术和视频剪辑技术实现对用户上传图片或视频的渲染处理与格式转换，生成编辑处理后图像或视频）等

- 多模态算法：例如达摩院交互式多能型合成算法（应用于开放域多模态内容生成场景，服务于问答、咨询类的企业端客户，通过 API 提供根据用户输入生成多模态信息的功能）、大麦小蜜智能客服算法（应用于在线智能客服场景，根据用户咨询内容，生成文本或语音智能回答）等

- 音频类算法：例如酷狗音乐歌声合成算法（应用于音频生成场景，基于用户歌唱录音数据，提取用户音色特征，制作歌曲）等

2.6 相关算法服务提供者和技术支持者应当尽快完成算法备案

网信办首次公开的《境内深度合成服务算法备案清单》可以看出，在《互联网信息服务深度合成管理规定》于 2023 年 1 月 10 日生效后，网信办已经清晰梳理出涉及人工智能（包括大语言模型、其他类型的大模型）的不同类型的算法的备案要求，因此实践中更进一步要求相关提供人工智能算法服务的公司需要尽快依据法律规定完成算法备案工作。

由于目前在 2023 年立法计划中的《人工智能法》亟待出台，因此人工智能相关领域的企业同时还务必要实时跟进相关法律法规，完善产品合规，降低相关合规风险。

（三）《生成式人工智能服务 安全基本要求》 （征求意见稿）解读

2023 年 10 月 11 日，全国信息安全标准化技术委员会（TC260）发布了关于征求信安标委技术文件《生成式人工智能服务 安全基本要求》（征求意见稿）意见的通知（以下简称“《安全评估征求意见稿》”）。该《安全评估征求意见稿》一旦正式公布，则 2023 年 7 月 10 日发布的《生成式人工智能服务管理暂行办法》第 17 条中规定的“安全评估”则有了明确的评估标准，可供各生成式人工智能服务企业参考。

1. 《安全评估征求意见稿》出台背景

2023 年 1 月 10 日生效的《互联网信息服务深度合成管理规定》中规定，深度合成服务提供者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，应当按照国家有关规定开展安全评估。其相关“安全评估”为依据《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》，在“全国互联网安全管理服务平台”上完成的安全评估。

但 2023 年 7 月 10 日出台的《生成式人工智能服务管理暂行办法》中也明确规定了“安全评估”：提供具有舆论属性或者社会动员能力的生成式人工智能服务的，应当按照国家有关规定开展安全评估，并按照《互联网信息服务算法推荐管理规定》履行算法备案和变更、注销备案手续。这一“安全评估”是否与“全国互联网安全管理服务平台”上完成的安全评估一致，目前在实务中是明显存在差异的。

根据 2023 年 8 月底的多家媒体报道，“多家大模型首批通过《生成式人工智能服务管理暂行办法》备案向公众开放”，其中所述“备案”明显与“算法备案”以及“全国互联网安全管理服务平台”上的安全评估不同。而目前根据本次发布的《安全评估征求

意见稿》可以进一步知悉,其具体评估内容与“算法备案”“全国互联网安全管理服务平台”上的安全评估也不同,加之这一国标的制定依据为《生成式人工智能服务管理暂行办法》,其作为《生成式人工智能服务管理暂行办法》的支撑文件,与8月底多家大模型通过的备案需要完成的要求相一致。

而企业无论是否面临《生成式人工智能服务管理暂行办法》相关备案,均由于《安全评估征求意见稿》汇总了目前包括知识产权、数据合规、网络安全合规、内容合规等多方面的合规要求,而可以成为大模型企业的合规参考,以避免可能的法律风险。

2. 主要评估内容

2.1 适用主体

根据《安全评估征求意见稿》,在大模型的安全评估过程中,TC260 首先界定了“生成式人工智能服务”与其服务“提供者”,其并未如《互联网信息服务深度合成管理规定》一样规定“技术支持者”,而是与《生成式人工智能服务管理暂行办法》保持一致,仅就直接向境内公众提供生成式人工智能服务的组织或个人规定为“提供者”,进而适用该安全评估。

2.2 评估内容

目前《安全评估征求意见稿》规定,需要从语料安全(训练数据安全)、模型安全、安全措施、安全评估进行生成式人工智能的安全评估:

2.2.1 语料安全

(1) 语料来源的基本评估要点

根据《安全评估征求意见稿》,训练数据(语料)被根据来源区分了不同的合规要求:

- 通过开源协议获得:应当遵守开源协议或者相应授权文件;
- 通过自采方式获得:不得违反 robots 协议等;
- 通过商业授权获得:如交易方不能提供语料合法性证明材料时,不应使用;

- 通过用户明确授权获得：将用户输入信息作为语料进行训练。

同时从整体而言，训练数据被要求建立语料应当搭配不同来源的语料训练（包括不同语言，境内外数据等）；同时应当建立语料来源黑名单，不使用黑名单来源的数据进行训练；根据我国网络安全相关法律要求阻断的信息，不应作为训练语料。

（2）语料内容的基本评估要点

内容合规作为生成式人工智能的重要合规要求，在《安全评估征求意见稿》中进行了如下要求：

- 对训练语料通过关键词、分类模型、人工抽检等方式过滤语料中的违法不良信息

针对内容审核，目前除了安全评估以外，其同样是算法备案过程中的监管重要审核要求，因此无论是通过企业自行搭建的内容机制，还是借助第三方的内容审核能力，均应当对训练数据采取充分的内容过滤，满足《生成式人工智能服务管理暂行办法》的要求。

- 对训练数据的知识产权合规评估

目前《安全评估征求意见稿》中明确要求提供者对训练数据中的知识产权进行合规评估，包括设置知识产权负责人、建立知识产权管理策略等。但现实中，对训练数据的知识产权侵权判断是十分困难的，同时目前包括中国在内的多个国家都在考虑为了促进人工智能的发展，促进《著作权法》的修改，增加将受到《著作权法》保护的内容用于人工智能训练时，可以适用合理使用，无需获得授权；而在商标权与专利权相关的内容中，将其作为人工智能的训练数据进行训练时，是否存在侵犯商标权或者侵犯专利权的行为，都是需要进一步明确讨论的。

因此在训练数据的知识产权合规评估过程中，《安全评估征求意见稿》中尚需要进一步的明确。

- 对训练数据的个人信息合规评估

《安全评估征求意见稿》依据《生成式人工智能服务管理暂行办法》第4条的规定，要求使用个人信息用于数据训练时应当满足《个人信息保护法》规定的相应合规要求，

包括获得个人信息主体的授权同意或者单独同意等。但目前的大模型训练中，特别是通用大模型，需要使用个人信息的场景并不多见。实践中还需要注意用户输入的个人信息是否会用于大模型的训练，往往会在隐私政策中进行明确约定。如存在，则需要满足《个人信息保护法》以及相关国标的要求。因此，匿名化后的个人信息如开展数据训练时，如何进行匿名化处理在现实中将尤为重要。

(3) 语料标注的基本评估要点

《安全评估征求意见稿》中根据《生成式人工智能服务管理暂行办法》第8条的要求，对语料标准从人员、规则、质量的角度进行要求。

2.2.2 模型安全

《安全评估征求意见稿》对模型本身的安全也同样提出了要求：

(1) 基础模型

用于研发的基础模型应当为经过主管部门备案的基础模型。

但该备案具体是指何种备案需要明确，以及针对基础模型的定义同样需要明确。

(2) 模型生成内容安全

针对模型直接输出的，未经其他处理的原生内容，目前《安全评估征求意见稿》依据《生成式人工智能服务管理暂行办法》第4条，要求提供者应当将生成内容安全性作为主要考虑指标，同时应对用户输入信息进行安全检测，引导模型生成积极内容；并且应当定期优化模型。

(3) 服务透明度

算法透明一直是各国监管的难点，《安全评估征求意见稿》中虽然规定了提供者应当在服务的相应位置向用户披露包括服务适用的人群等，以及第三方基础模型使用情况，以及具体服务的局限性等。

但基础模型、服务局限性等应当披露到何种程度，例如是披露使用了 LLAMA2 的基础模型就可以满足要求，还是要针对 LLAMA2 进行具体说明，是需要进一步明确的。同

时向公众披露基础模型、服务的局限性是否可以起到保护用户权益（知情权等）的要求也是需要梳理的。如果是需要公众具有一定的知识后才可以理解的信息，是否信息披露应当以其他方式替代。

（4）生成内容准确性、可靠性

从避免生成内容被用于诈骗等违法犯罪行为，生成式人工智能生成物应当准确、可靠。但同时需要考虑的是目前的技术，如果为了实现这一目标将花费多大成本，其是否可以实现。是否生成内容的准确性、可靠性可以通过例如生成内容可追溯、显著提示等方式进行，在现在的技术发展阶段的弥补，是可以被考虑的。

2.2.3 安全措施

《安全评估征求意见稿》对生成式人工智能服务的提供者的安全措施提出了如下要求：

（1）未成年人保护

《生成式人工智能服务管理暂行办法》第 10 条明确规定了生成式人工智能服务对未成年人的保护，而《安全评估征求意见稿》也针对如果向未成年人提供服务时应当采取的安全措施进行了进一步的规定。

2023 年 10 月 24 日，国务院公布了《未成年人网络保护条例》，该条例生效后需要进一步为生成式人工智能服务依据该条例进行合规评估。

（2）收集使用者输入信息用于训练

《安全评估征求意见稿》进一步要求，除需要满足获得用户同意的要求外，还需要提供关闭使用者输入信息用于训练的选项等。

（3）图片、视频等内容标识

TC260 在 2023 年 8 月发布了《网络安全标准实践指南—生成式人工智能服务内容标识方法》，该标准已经非常清晰的规定了文本、图片、音视频等生成物如何完成显著标识、隐式标识等设置方法，因此可以遵守该标准处理。

(4) 接受公众或使用者投诉举报

(5) 向使用者提供生成内容

《安全评估征求意见稿》针对向用户生成的内容，应当仅可以针对明显的违反不良信息可以拒绝回答，而其他问题应均能正常回答，并且设置监看人员。

针对这一要求，如果除了明显的违反不良信息以外不得拒绝回答，将存在某一特定的生成式人工智能服务提供者，如果不具备审核具体内容的能力或者资质，例如新闻信息、医疗信息等，将进入很难回避的困境。

(6) 模型更新、升级方面

《安全评估征求意见稿》规定了模型重要更新、升级后应当再次完成安全评估，并且向主管部门重新备案，但对于模型重要更新、升级同样需要确定具体标准，否则提供者无法明确何时需要进行重新备案。

2.2.4 安全评估要求

《安全评估征求意见稿》明确了前述各项规定的评估标准，这些具体评估要求的明确将非常有利于生成式人工智能服务提供者对合规要求的理解和落地执行。



序号	评估事项	具体要求
1	语料安全评估	● 人工抽检，随机抽样 4000 条语料，合格率应当达到 96%； ● 技术抽检（关键词检测、分类模型等），随机抽样不少于总量的 10%，合格率不应低于 98%。
2	生成内容安全评估	● 建立符合要求的测试题库； ● 人工抽检，从测试题库中随机抽样 1000 条测试题，合格率应当达到 90%； ● 技术抽检（关键词检测、分类模型等），从测试题库中随机抽样 1000 条，合格率不应低于 90%。
3	问题拒答评估	● 建立符合要求的测试题库； ● 人工抽检，从测试题库中随机抽样 300 条测试题，对应拒答问题，拒答率应当不低于 95%； ● 技术抽检（关键词检测、分类模型等），从测试题库中随机抽样 300 条测试题，对不应拒答问题，拒答率应当高于 5%。
4	关键词库	● 关键词一般不应超过 10 个汉字等； ● 总规模不应少于 10000 个； ● 关键词应具有代表性。
5	分类模型	● 应当完全覆盖相关安全风险。
6	生成内容测试题库	● 不少于 2000 题； ● 具有代表性； ● 应当建立相应的操作规程以及判别依据。
7	拒答测试题库	● 应拒答/不应拒答测试题库应当分别不少于 500 题； ● 应拒答/不应拒答测试题库应当分别具有代表性。

表 2-5 评估要求汇总表

(四)《科技伦理审查办法（试行）》实施后， AIGC 企业的应对措施

2023 年 4 月 4 日科技部发布《科技伦理审查办法（试行）（征求意见稿）》（简称《征求意见稿》）后，2023 年 10 月 8 日科技部、教育部、工业和信息化部、农业农村部、国家卫生健康委、中国科学院、中国社科院、中国工程院、中国科协、中央军委科技委等十部门联合发布了《科技伦理审查办法（试行）》（以下简称《试行办法》），并于 2023 年 12 月 1 日实施。

《试行办法》中的规定由于与人工智能企业存在大量关联，因此该办法出台后人工智能企业，特别是 AIGC 企业（即生成式人工智能服务提供者）应当重点关注其中的相关部分，并且积极应对。

1.《试行办法》对 AIGC 企业的影响

1.1 需要完成科技伦理审查的 AIGC 企业

本次《试行办法》第二条的规定，如下科技活动需要依据该办法进行科技伦理审查：

- 涉及以人为研究参与者的科技活动，包括以人为测试、调查、观察等研究活动的对象，以及利用人类生物样本、个人信息数据等的科技活动；
- 涉及实验动物的科技活动；
- 不直接涉及人或实验动物，但可能在生命健康、生态环境、公共秩序、可持续发展等方面带来伦理风险挑战的科技活动；
- 依据法律、行政法规和国家有关规定需进行科技伦理审查的其他科技活动。

因此需要依据《试行办法》进行科技伦理审查的 AIGC 企业包括：

- 深度合成服务提供者

根据 2023 年 1 月 10 日生效的《互联网信息服务深度合成管理规定》，深度合成服务提供者应当落实信息安全主体责任，建立健全用户注册、算法机制机理审核、科技伦理审查、信息发布审核、数据安全、个人信息保护、反电信网络诈骗、应急处置等管理制度，具有安全可控的技术保障措施。

因此基于《互联网信息服务深度合成管理规定》，深度合成服务提供者应当依据《试行办法》落实科技伦理审查。

- 研发的算法模型涉及个人信息数据的
- 研发的算法模型对公共秩序等可能产生影响的

因此，针对 AIGC 企业，除了明确的深度合成服务提供者需要进行相关科技活动的科技伦理审查外，算法模型如果涉及个人信息数据、或者对公共秩序等可能产生影响的技术服务提供者，同样需要进行科技伦理审查。

1.2 科技伦理审查相关合规义务

根据前述分析，AIGC 企业需要适用《试行办法》进行科技伦理审查之际，需要完成如下合规工作：

(1) 设立科技伦理（审查）委员会

《试行办法》规定，企业是本单位科技伦理审查管理的责任主体，从事包括人工智能等科技活动的企业，研究内容涉及科技伦理敏感领域的，应当设立科技伦理（审查）委员会。

虽然《试行办法》明确规定人工智能等科技活动的企业，如果涉及科技伦理敏感领域的，应当设立科技伦理（审查）委员会，其他有科技伦理审查需求的单位可根据实际情况设立科技伦理委员会。但针对不属于应当设立而未设立科技伦理（审查）委员会的企业如进行科技活动的，应书面委托其他满足要求的科技伦理（审查）委员会开展伦理审查。

因此 AIGC 企业除符合《试行办法》应当设立科技伦理（审查）委员会外，如果有

科技伦理审查需求可以自行设立科技伦理（审查）委员会，而在没有设立科技伦理（审查）委员会而需要进行科技伦理审查时，需要书面委托其他满足要求的科技伦理（审查）委员会开展伦理审查。

（2）科技伦理（审查）委员会依据《试行办法》设立、登记

企业如设立科技伦理（审查）委员会，需要依据《试行办法》搭建科技伦理审查委员会机制（聘请外部专家、筛选委员、完善配套制度，如制定章程，建立健全审查、监督、保密管理、档案管理等制度规范和工作规程，健全和利益冲突管理机制）。

并且，企业应当在设立科技伦理（审查）委员会后的 30 日内，通过国家科技伦理管理信息登记平台进行登记。

（3）科技伦理（审查）委员会依据《试行办法》进行审查

科技伦理（审查）委员会应当依据《试行办法》进行科技伦理风险评估，并且区分不同的评估事项进行审查以及进行跟踪审查。根据不同的评估事项适用简易程序、一般程序或者专家复核程序进行审查。

简易程序和一般程序由企业的科技伦理（审查）委员会依据《试行办法》以及其章程进行，专家复核程序由企业报请所在地方或相关行业主管部门组织开展。

● 科技伦理审查一般程序的触发

根据前述 1 需要完成科技伦理审查的 AIGC 企业部分的分析，AIGC 企业，特别是深度合成服务提供者至少将会触发科技伦理相关的一般程序。

● 科技伦理审查专家复核程序的触发

目前《试行办法》附件中明确“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。”需要开展专家复核，也就是依据《试行办法》在科技伦理（审查）委员会完成初步审查后，报请相关部门组织专家进行复核，企业的科技伦理（审查）委员会依据专家复核意见进行审查决定。

但结合目前网信办对“具有舆论属性与社会动员能力”的宽泛理解,生成式 AIGC 产品,尤其是大模型公司或大模型应用层相关产品,除了触发科技伦理审查的一般流程以外,还可能基于这一条,面临更为严格的专家复核程序。但是,是否适用仍需考虑是否同时满足《试行办法》附件提到的“舆论属性”+“社会动员能力”+“社会意识引导能力的”三个要件,但三者的含义,判断标准尚需进一步明确。

(4) 审查后的科技活动完成登记

科技活动负责人在完成内部申请、科技伦理审查委员会受理和初审通过后,企业应在纳入清单管理的科技活动获得伦理审查批准后的 30 日内,通过国家科技伦理管理信息登记平台进行登记。

(5) 跟踪审查

企业的科技伦理(审查)委员会针对已经审查批准的科技活动应当开展伦理跟踪审查,跟踪审查主要包括科技活动实施方案执行情况及调整情况等,同时规定间隔一般不超过 12 个月。

值得注意的是,《试行办法》针对简易程序,科技伦理(审查)委员会可以自行根据情况调整跟踪审查频度,而专家复核程序则进一步规定了不超过 6 个月的间隔期,并且需要进行动态管理。

(6) 年度报告

企业应于每年 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、纳入清单管理的科技活动实施情况报告等。

2. AIGC 企业应对

2.1 区分人工智能科技活动与人工智能 / 深度合成算法服务的区别,完成监管登记 / 备案

下表为中国境内涉及人工智能的主要法律文件:

序号	法律文件名称	适用范围
1	《互联网信息服务算法推荐管理规定》	在中华人民共和国境内应用算法推荐技术提供互联网信息服务
2	《互联网信息服务深度合成管理规定》	在中华人民共和国境内应用深度合成技术提供互联网信息服务
3	《生成式人工智能服务管理暂行办法》	利用生成式人工智能技术向中华人民共和国境内公众提供生成文本、图片、音频、视频等内容的服务

表 2-6 中国涉及人工智能的主要法律文件

因此可以看出，目前已有的人工智能相关法律文件，主要还是以向第三方提供服务作为监管前提，但《试行办法》将监管范围向前扩展到研发阶段的科技活动，因此从事AIGC，特别是研发的企业，将需要在自研发阶段就开始进行相关的合规问题，并且搭建相应的科技伦理审查制度，以避免适用《试行办法》受到处罚。

2.2 从事 AIGC 科技活动，特别是大模型研发的深度合成服务提供者（直接向公众提供 AIGC 服务的企业），建议应当尽快完成科技伦理（审查）委员会的设立以及相关登记手续

大模型在研发阶段，将收集和训练大量个人信息数据用于数据训练，或者虽然不涉及个人信息数据，但是可能基于其应用层的影响，对公共秩序、可持续发展、生命健康带来伦理风险挑战，即使不属于深度合成服务提供者，但属于《试行办法》第二条明确的四项应该开展科技伦理审查的要件，为此需要特别注意，建议尽快依照《试行办法》完成科技伦理（审查）委员会的设立，并且进行相关科技伦理审查，完成大模型公司科技伦理审查委员会设立及“纳入清单管理的”科技活动开展登记手续。

2.3 对于需要完成算法备案、安全评估的大模型技术层、应用层公司，在此过程中是否触发科技伦理审查，还需要结合自身业务情况、监管口径进一步判断

科技伦理审查（包括登记以及专家复核意见）与算法备案、安全评估并不存在相互的前置审批关系，但在算法备案、安全评估中并不被强调的“研发阶段”，在科技伦理审查相关流程将受到特别的重视，但监管部门仍以企业自主完成审查 + 事后登记 + 关键活动专家复核的方式进行监管。因此，相关企业需要结合自身研发情况进行应对。

2.4 企业需关注一般程序和专家复核程序中跟踪审查期限的差异性，及风险重大变化时的重新审查、专家复核

对于跟踪审查，《试行办法》规定一般程序需 12 个月开展一次，专家复核程序 6 个月一次，《征求意见稿》规定的则是 2 年，可见科技更新速度快，为此监管要求加强信息更新。

同时，在科技伦理风险发生重大变化时，应按照《试行办法》第二十条规定重新开展伦理审查并申请和专家复核。一旦需要重新审查，对当前科技活动是继续实施，还是暂停实施，由科技伦理委员会进行确定。

2.5 企业触发科技伦理审查，需及时完成对应合法合规审查

《试行办法》提示触发科技伦理审查的项目，审查内容涉及个人隐私数据、数据生命周期处理活动，以及算法设计、实现、应用活动等，均需要完成对应合法合规的审查。特别是涉及数据和算法的科技活动的，需要审查“数据的收集、存储、加工、使用等处理活动以及研究开发数据新技术等符合国家数据安全和个人信息保护等有关规定，数据安全风险监测及应急处理方案得当；算法、模型和系统的设计、实现、应用等遵守公平、公正、透明、可靠、可控等原则，符合国家有关要求，伦理风险评估审核和应急处置方

案合理，用户权益保护措施全面得当”等要点，体现了国家在伦理风险治理上的全面性。

（五）《网络安全标准实践指南——生成式人工智能服务内容标识方法（征求意见稿）》解读

2023年8月8日，全国信息安全标准化技术委员会为落实包括《互联网信息服务深度合成管理规定》《生成式人工智能服务管理暂行办法》等相关规定中对人工智能生成内容进行标识的合规义务，为实践提供清晰的指引，发布了《网络安全标准实践指南——生成式人工智能服务内容标识方法（征求意见稿）》。

2023年8月25日，全国信息安全标准化技术委员会秘书处为落实《生成式人工智能服务管理暂行办法》相关要求，指导有关单位利用生成式人工智能技术提供生成文本、图片、音频、视频等内容服务时对内容进行标识，正式发布了《网络安全标准实践指南——生成式人工智能服务内容标识方法》（以下简称“标识方法”）

对此本所特进行简要梳理，协助生成式人工智能相关企业理解生成式人工智能标识标注方式方法，妥善完成算法备案、安全评估以及履行合规义务。

1. 生成式人工智能就生成内容进行标识的法律依据

实施日期	法律法规	关于标识的主要规定
2020.01.01	网络音视频信息服务管理规定	第十一条 网络音视频信息服务提供者和网络音视频信息服务使用者利用基于深度学习、虚拟现实等的新技术新应用制作、发布、传播非真实音视频信息的，应当以 <u>显著方式</u> 予以标识。
2021.02.09	关于加强网络直播规范管理工作的指导意见	7. 筑牢信息安全屏障。网络直播平台应当建立健全信息安全管理制​​度，严格落实信息内容安全管理责任制，具备与创新发展相适应的安全可控的技术保障和防范

实施日期	法律法规	关于标识的主要规定
		措施；对新技术新应用新功能上线具有舆论属性或社会动员能力的直播信息服务，应严格进行安全评估；利用基于深度学习、虚拟现实等技术制作、发布的非真实直播信息内容，应当以 <u>显著方式</u> 予以标识。
2022.03.01	互联网信息服务算法推荐管理规定	第九条 算法推荐服务提供者应当加强信息安全管理，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。发现 <u>未作显著标识的算法生成合成信息</u> 的，应当作出显著标识后，方可继续传输。
2023.01.10	互联网信息服务深度合成管理规定	第十六条 深度合成服务提供者对其服务生成或者编辑的信息内容，应当采取<u>技术措施</u>添加不影响用户使用的标识，并依照法律、行政法规和国家有关规定保存日志信息。 第十七条 深度合成服务提供者提供以下深度合成服务，可能导致公众混淆或者误认的，应当在生成或者编辑的信息内容的合理位置、区域进行<u>显著标识</u>，向公众提示深度合成情况： （一）智能对话、智能写作等模拟自然人进行文本的生成或者编辑服务； （二）合成人声、仿声等语音生成或者显著改变个人身份特征的编辑服务； （三）人脸生成、人脸替换、人脸操控、姿态操控等人物图像、视频生成或者显著改变个人身份特征的编辑服务；

实施日期	法律法规	关于标识的主要规定
		(四) 沉浸式拟真场景等生成或者编辑服务； (五) 其他具有生成或者显著改变信息内容功能的服务。 深度合成服务提供者提供前款规定之外的深度合成服务的，应当提供显著标识功能，并提示深度合成服务使用者可以进行显著标识。 第十八条 任何组织和个人不得采用技术手段删除、篡改、隐匿本规定第十六条和第十七条规定的深度合成标识。
2023. 08. 15	生成式人工智能服务管理暂行办法	第十二条 提供者应当按照《互联网信息服务深度合成管理规定》对图片、视频等生成内容进行标识。

表 2-7 生成式人工智能就生成内容进行标识的法律依据汇总表

可以看出，针对生成合成类算法的生成内容应当进行显著标识，并非始自刚刚生效的《生成式人工智能服务管理暂行办法》，而是早在 2020 年的《网络音视频信息服务管理规定》就已经进行要求。

我们理解目前相关立法要求进行“标识”的目的是：

(1) 使用户清晰知悉所获得的信息是人工智能生成的，并非真实信息，谨慎使用和传播，因此这类标识应当在“显著”位置设置；

(2) 保证因生成内容造成虚假信息传播，甚至诈骗等犯罪行为时可以溯源追究算法提供方责任，以及通过算法提供者追究进行虚假信息传播、甚至犯罪行为的用户的法律

责任。因此这类标识应当不影响用户使用。

在上述两种不同的目的下,《互联网信息服务深度合成管理规定》第 17 条规定提供深度合成算法服务时,应当“显著标识”深度合成情况。

而该规定第 16 条则规定了深度合成服务提供者应当采取技术措施添加不影响用户使用的标识,并依照法律、行政法规和国家有关规定保存日志信息。

而本次信安标委的《标识方法》正是具体落实上述两种不同的标识,为实践提供清晰的指引。因此生成式人工智能服务提供者在进行算法备案、安全评估以及相关合规工作时,可以参考该《标识方法》确保法律法规所要求的标识要求已经满足。

2. 《标识方法》对生成式人工智能服务标识的指引性规定

《标识方法》基于文本、图片、音频、视频四类生成内容,给出了生成式人工智能不同合规要求的标识的实践指引,可以作为生成式人工智能的行业实践进行参考,同时也会作为相关监管部门的参考,因此如果无法依据其指引执行,应当有充分合理的理由。

因此,我们梳理了《标识方法》中的具体规定,以协助生成式人工智能相关企业对其的理解。

标识方式	适用场景	标识位置	标识文字/语音	说明
显著标识	各类场景均须设置	应在 <u>显示区域下方</u> 或 <u>使用者输入信息区域下方</u> 持续显示提示文字。	至少包含“由人工智能生成”或“由 AI 生成”等信息。	这种提示是最容易实现,也是各类生成式人工智能服务无论是否进行其他标识设置均需要在产品/服务中实现的。
		或者在 <u>显示区域的背景</u> 添加包含提示文字的 <u>显式水印标识</u> 。		显式水印标识是指,在生成式人工智能服务中,在交互界面内或背

标识方式	适用场景	标识位置	标识文字/语音	说明
				景中添加的均匀分布的半透明文字。这种显示方式与上述提示标识中的“显示区域下方或使用者输入信息区域下方持续显示提示文字”。
	图片、视频类需要设置	由人工智能生成 <u>图片、视频</u> 时，应采用在 <u>画面</u> 中添加提示文字的方式进行标识。		
	自然服务转人工智能服务时	应通过提示文字或提示语音的方式进行标识	至少包含“人工智能为您提供服务”或“AI 为您提供服务”等信息。	主要目的是为了区分自然人服务与人工智能服务，避免混淆。
不影响用户使用标识	图片、音频、视频类生成内容	在生成式人工智能服务中，通过修改图片、音频、视频内容添加的，人类无法直接感知、但可通过技术手段从内容中提取的标识。 ● 图片的隐式水印标识：通过空域水印或变换域水印的方式实现。含有隐式水印的图片应满足任意连续 50%	至少包含服务提供者名称，也可包含内容 ID 等其他内容	目前行业内比较常见的隐式水印标识是直接修改图片、音视频的代码嵌入的暗水印。但如《征求意见稿》中的规定，图片、音视频中增加隐式水印标识是相对较为容易的，可以通过修改代码实现，但针对显示内容基本等同于代码的文本类就较难实现使用隐式

标识方式	适用场景	标识位置	标识文字/语音	说明
		以上面积且分辨率大于等于 384×384 的切片均包含完整标识信息。 ● 视频的隐式水印标识：通过时空域水印或变换域水印的方式实现。含有隐式水印标识的视频应满足任意连续 5 秒以上视频内容中均包含完整标识信息； 音频的隐式水印标识：含有隐式水印标识的音频应满足任意连续 10 秒以上音频内容中均包含完整标识信息。		水印标识。在《征求意见稿》中也可以看出，目前并未强制要求文本类进行要在生成内容中增加隐式水印标识。

表 2-8 《生成式人工智能服务内容标识方法》合规义务清单

（六）中国生成式人工智能与深度合成监管差异与联系的探析

2022 年 11 月 3 日，国家互联网信息办公室、工业和信息化部、公安部令发布《互联网信息服务深度合成管理规定》（本部分简称为“《深度合成管理规定》”），该规定自 2023 年 1 月 10 日起施行。

该规定正式施行后时隔三个月，2023 年 4 月 11 日，国家互联网信息办公室又公布了《生成式人工智能服务管理办法（征求意见稿）》（本部分简称“《生成式人工智能办法》”）向社会公开征求意见，并于 2023 年 7 月 13 日正式出台《生成式人工智能服务管理暂行办法》。

可见网信办等国家监管机构对人工智能技术的高度关注，但这也让我们产生些疑惑，即深度合成规定与生成式人工智能办法中所指的技术或者说两个法规适用对象有什么区别？它们之间又存在什么联系？

1. 技术区别

首先，根据《深度合成管理规定》和《生成式人工智能办法》中的规定，从语义上进行二者的对比。

● 生成式人工智能，是指基于算法、模型、规则生成文本、图片、声音、视频、代码等技术。

● 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，包括但不限于：

- （1）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的技术；
- （2）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的技术；
- （3）音乐生成、场景声编辑等生成或者编辑非语音内容的技术；
- （4）人脸生成、人脸替换、人物属性编辑、人脸操控、姿态操控等生成或者编辑图像、视频内容中生物特征的技术；
- （5）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的技术；
- （6）三维重建、数字仿真等生成或者编辑数字人物、虚拟场景的技术。

从以上定义来看，生成式人工智能和深度合成技术从字面上比较有以下区别：

1.1 生成式人工智能较深度合成技术而言不仅仅指由算法生成，还包括了模型、规则生成的内容

《生成式人工智能办法》第四条也提到：“提供生成式人工智能产品或服务应当遵守法律法规的要求，尊重社会公德、公序良俗，符合以下要求：...（二）在算法设计、训练数据选择、模型生成和优化、提供服务等过程中，采取措施防止出现种族、民族、信仰、国别、地域、性别、年龄、职业等歧视...”

从定义以及《生成式人工智能办法》第四条第二项的要求上看，生成式人工智能不局限于算法技术，还包括了模型和规则，在计算机科学角度，算法和模型、规则不是一个概念，算法是在代码中实现并在数据上运行的过程，模型是运行在数据集上的机器学习算法的输出，规则是算法的底层逻辑和思想。

虽然生成式人工智能中涉及算法的部分与深度合成技术较为类似，均属于生成合成类算法技术范畴内，但生成式人工智能办法希望监管的范围比单纯的算法更广。

1.2 深度合成技术与生成式人工智能技术应用场景侧重点不同

在此需要介绍下两个词语的来源，深度合成技术来源于国外深度伪造（deep fake）一词，即深度学习 deep learning 和伪造 fake 的结合，生成式人工智能来源于国外非常热门的 AI Generated Content。

在生成合成类算法角度而言深度合成技术和生成式人工智能在本质上没有实质性区别，而是由于应用侧重点不同导致有所区分，例如深度合成技术主要应用功能为合成类软件例如 Deep Fake，生成式人工智能主要应用功能于人工智能对话机器人例如 ChatGPT。

2. 法规适用区别

2.1 立法目的不同

首先从上位法以及立法目的来看，《深度合成管理规定》的立法目的是为了加强互联网信息服务深度合成管理，其上位法中有《互联网信息服务管理办法》，可见深度合成技术规定的侧重点在于互联网的治理。在《深度合成管理规定》答记者中介绍到，制

定该规定的初衷之一是维护网络空间良好生态，深度合成服务在满足用户需求、改进用户体验的同时，也被一些不法人员用于制作、复制、发布、传播违法信息，诋毁、贬损

他人名誉、荣誉，仿冒他人身份实施诈骗等违法行为，影响传播秩序和社会秩序，损害人民群众合法权益，危害国家安全和社会稳定。

而《生成式人工智能办法》的立法目的在于促进生成式人工智能健康发展和规范应用，不仅仅限于互联网的治理，而是关注于整个人工智能行业的健康发展。

2.2 监管范围不同

《深度合成管理规定》较《生成式人工智能办法》不同的是，它的监管范围除了技术的提供者还包括深度合成服务技术支持者（为深度合成服务提供技术支持的组织、个人）和深度合成服务使用者（使用深度合成服务制作、复制、发布、传播信息的组织、个人）。

2.3 具体合规要求不同

（1）数据监管方面

《深度合成管理规定》明确了深度合成服务提供者和技术支持者对训练数据的管理义务，还对训练数据和其他数据作出区分并提出不同监管要求，比如单独设计“输入数据”相关条款，例如第十条规定“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”。

《生成式人工智能办法》对数据来源的合法性规定主要是侧重于提供者应当对生成式人工智能产品的预训练数据、优化训练数据来源的合法性负责，要求用于生成式人工智能产品的预训练、优化训练数据，应满足以下要求：（一）符合《中华人民共和国网络安全法》等法律法规的要求；（二）不含有侵犯知识产权的内容；（三）数据包含个人信息的，应当征得个人信息主体同意或者符合法律、行政法规规定的其他情形；（四）能够保证数据的真实性、准确性、客观性、多样性；（五）国家网信部门关于生成式人工智能服务的其他监管要求。

（2）输出内容监管方面

两个规定对于输出内容应当符合法律规定、体现社会主义核心价值观等正面导向均有一般性规定，虽然表达的具体语义不一样。两者对于发现输出内容不符合正面导向的相应规定时的处理结果有些许不同，主要表现在《生成式人工智能办法》对使用者的措施非常简单，只有暂停或终止服务两种，而《深度合成管理规定》有警示、限制功能、暂停服务、关闭账号等措施。

但《生成式人工智能办法》有一个特色措施，即第十五条：“对于运行中发现、用户举报的不符合本办法要求的生成内容，除采取内容过滤等措施外，应在 3 个月内通过模型优化训练等方式防止再次生成”，虽然该条没有像上下其他条款中写明主体是提供者，但从该条文内容来看应当是针对提供者的合规义务。该条款要求提供者应在 3 个月内通过模型优化训练等方式防止再次生成，可见监管部门希望人工智能技术能够通过优化训练积极向善发展。

3. 两个法规之间的衔接

虽然两个法规在立法目的、监管范围以及具体合规要求方面有不同侧重，但两个法规以及《算法推荐管理规定》之间还是有相互链接的地方，《生成式人工智能办法》中在第六条和第十六条中都提到在某种情况下应当按照《互联网信息服务算法推荐管理规定》、《深度合成技术规定》中的相应规定来处理。

例如《生成式人工智能办法》第六条规定：“利用生成式人工智能产品向公众提供服务前，应当按照《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》向国家网信部门申报安全评估，并按照《互联网信息服务算法推荐管理规定》履行算法备案和变更、注销备案手续”。

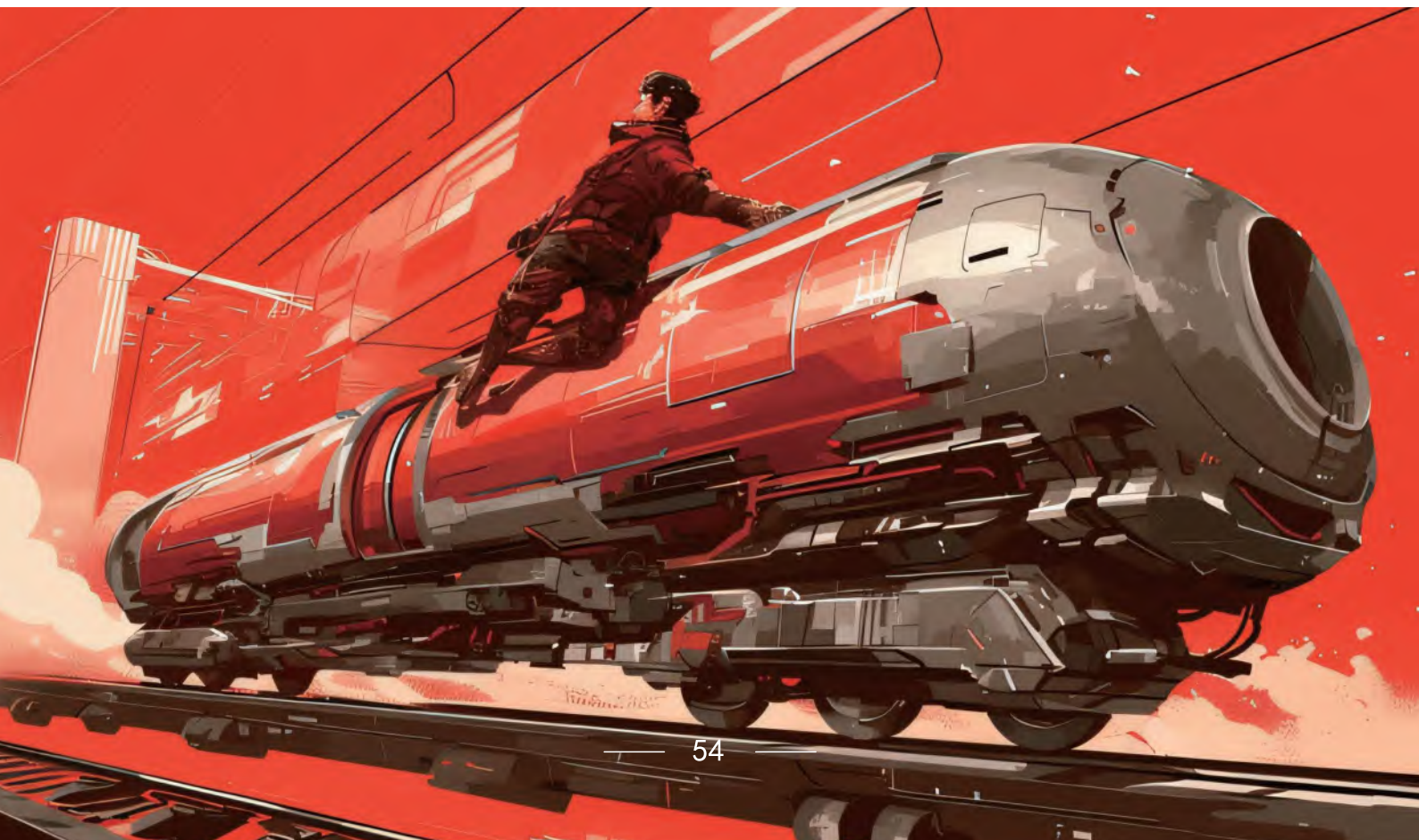
《生成式人工智能办法》第十六条规定：“提供者应当按照《互联网信息服务深度合成管理规定》对生成的图片、视频等内容进行标识”。

同时，两个规定在监管中与其他规定的衔接还具有一致性，例如算法备案和安全评估方面，两个规定对于生成式人工智能和深度合成技术的产品都规定了相应的要求。

例如《生成式人工智能办法》第六条规定：“利用生成式人工智能产品向公众提供服务前，应当按照《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》向国家网信部门申报安全评估，并按照《互联网信息服务算法推荐管理规定》履行算法备案和变更、注销备案手续”。

《深度合成管理规定》第十九条规定：“具有舆论属性或者社会动员能力的深度合成服务提供者，应当按照《互联网信息服务算法推荐管理规定》履行备案和变更、注销备案手续”。第二十条规定：“深度合成服务提供者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，应当按照国家有关规定开展安全评估”。

随着人工智能技术的不断发展，当大量信息涌入社会公众面前时，一般民众难以判断信息的真实性，导致了互联网已经进入了“眼见不一定为实”的后真相时代，监管部门必然将在《数据安全法》《网络安全法》《个人信息保护法》等上位法框架下继续完善配套监管规则，根据不同的业务领域的特点布局相应的监管要求，逐步填补对人工智能技术的监管空白。



第三部分 知识产权合规



三、知识产权合规

《暂行办法》第4条、第7条明确规定，AIGC产品应当尊重知识产权，AIGC的知识产权合规问题贯穿于AIGC从创造到应用的全过程。AIGC产品无论是自研大模型时的训练数据，还是优化第三方大模型时的优化数据，均涉及到对数据的挖掘和训练，那么就涉及将训练数据，包括训练数据中可能涉及的作品复制到本地进行训练的问题，根据《著作权法》未经授权的复制行为，除非构成合理使用或者法定许可，否则将构成侵权。如果在AIGC生成内容中“拼接”了训练数据中其他拥有著作权的作品，则根据中国现行《著作权法》，除非构成合理使用，否则应当获得著作权人授权。这一问题是现在各AIGC产品运营者遇到的最重要，且最难以解决的知识产权问题。另外，在包括中国在内的大部分国家的知识产权法体系项下，作品的作者都仅能是自然人，AI无法作为成为作品的作者。目前中国司法实践中存在完全不同的司法尝试，使得AI公司如果希望通过著作权保护自己的AI生成物时需要谨慎考虑。

（一）AIGC 训练数据的知识产权合规

AIGC产品无论是自研大模型时的训练数据，还是优化第三方大模型时的优化数据，均涉及到对数据的挖掘和训练，那么就涉及将训练数据，包括训练数据中可能涉及的作品复制到本地进行训练的问题，根据《著作权法》未经授权的复制行为，除非构成合理

使用或者法定许可，否则将构成侵权。

但实际上数量极其庞大的训练数据如果均需要获得著作权授权，在著作权集体管理组织发展不完善的中国是非常难以落地的。

同时，我们观察到日本，在 AI 快速发展后，在著作权法立法中对合理使用范围作了进一步扩大，将 AI 中的“文本数据挖掘”增列为一项新的合理使用情形。此外，我们观察到欧盟也有 AIGC 训练数据适用合理使用的法律空间，而我国在 2020 年著作权法修改过程中，曾经一度希望在合理使用制度引入“三步检验法”，以期将中国的合理使用制度修改为开放式制度（仅通过三步检验法判断是否构成合理使用），降低现实中合理使用的适用难度，解决新技术下的立法问题，例如现在讨论的人工智能。但最终中国《著作权法》以合理使用条款中增加了“其他法律法规规定的情形”的方式公布，因此现行《著作权法》中并未明确规定 AIGC 数据训练可以适用合理使用，

但如果正在征求意见的《著作权法实施条例》可以增加相关规定，同样可以为未来的 AIGC 数据训练的著作权侵权风险提供一个良好的解决路径。

自 2023 年年初开始，已有数十名著作权人对图片、文字生成类等 AI 产品的开发者提起侵权诉讼，目前诉讼均处于初步审理阶段，本部分对相关诉讼进行梳理。

1. 图片生成类 AI 著作权侵权案件

2023 年年初，图片提供商 Getty Images 在美国和英国两地起诉了热门的人工智能（“AI”）绘画工具的开发者的 Stability AI，称其侵犯了版权，因为他们“非法”地从权利人网站上窃取了数百万张图片来训练 AI 模型；此外在美国还发生了起诉 AI 开发者的集体诉讼，一群艺术家起诉了 Stability AI 等公司，指控他们未经原作者同意从网络上获取 50 亿张图像来训练其 AI 模型，侵犯了数百万艺术家的权利。2023 年 11 月 29 日，4 位创作者针对小红书向北京互联网法院提起诉讼，创作者主张小红书旗下推出的 AI 绘画产品 Trik 所生成图片与平台用户原创作品高度相似^[3]。

^[3] 参见《多位创作者起诉小红书偷画师作品“炼”AI 模型涉侵权，已获立案》，发布于《知产前沿》，https://www.sohu.com/a/740280849_120133310（最后访问时间：2024 年 1 月 16 日）。

序号	涉案AI及其开发商	案件基本情况
1	Stable Diffusion (Stability AI) Midjourney (Midjourney) DreamUp (Deviant Art)	● 2023年1月，艺术家莎拉·安德森（Sarah Andersen）、凯利·麦克南（Kelly McKernan）和卡拉·奥尔蒂斯（Karla Ortiz）三位艺术家在美国加州联邦法院对 Stability AI, Midjourney 和 DeviantArt 提起诉讼。 ● 三位艺术家主张被告在“未经原始艺术家同意”的情况下，使用从网络上抓取的 50 亿张图像来训练他们的人工智能工具，从而侵犯了“数百万艺术家”的权利。^[4] ● 2023年10月31日，加州联邦法院驳回了提起的拟议集体诉讼中的部分主张，包括针对 Midjourney 和 DeviantArt 的所有指控；但针对 Stability AI 的直接侵权索赔被允许继续进行。^[5]
2	Stable Diffusion (Stability AI)	● 2023年1月12日，Getty Images 在英国伦敦高等法院对 Stability AI 提起诉讼。Getty Images 主张 Stability AI 在没有获得许可的情况下非法复制和处理了数百万张受版权保护的图像以及 Getty Images 拥有或代表的相关元数据，从而有利于 Stability AI 的商业利益并损害了内容创作者的利益。^[6]

^[4] See AI art tools Stable Diffusion and Midjourney targeted with copyright lawsuit, published in The Verge, available at: <https://www.theverge.com/2023/1/16/23557098/generative-ai-art-copyright-legal-lawsuit-stable-diffusion-midjourney-deviantart> (最后访问时间: 2024年1月16日)。

^[5] See Judge pares down artists' AI copyright lawsuit against Midjourney, Stability AI, published in Reuters, available at: <https://www.reuters.com/legal/litigation/judge-pares-down-artists-ai-copyright-lawsuit-against-midjourney-stability-ai-2023-10-30/> (最后访问时间: 2024年1月16日)。

^[6] See Getty Images is suing the creators of AI art tool Stable Diffusion for scraping its content, published in The Verge, available at: <https://www.theverge.com/2023/1/17/23558516/ai-art-copyright-stable-diffusion-getty-images-lawsuit> (最后访问时间: 2024年1月16日)。

序号	涉案AI及其开发商	案件基本情况
		● 2023年12月1日,英国法院裁定驳回 Stability AI 的管辖权异议。法院认为 Getty 认为其受版权保护的材料被用于训练 AI 模型是有根据的,英国高等法院裁定此案具有合理性,该诉讼被裁定继续审理,将进入审判阶段。 ^[7]
3	Stable Diffusion (Stability AI)	● 2023年2月3日,Getty Images 向特拉华州美国地方法院针对 Stability AI 提起诉讼。 ^[8]
4	小红书 Trik (行吟 信息科技有限公司(上海) 有限公司)	● 2023年11月29日晚,四位创作者@是雪鱼啊 @画画的云淡风轻 @正版青团子 @RedMatcha 发布最新维权微博,称其与小红书关于作品著作权侵权一案有了新进展,几位画师分别将小红书主体公司行吟信息科技有限公司(上海)有限公司以及小红书 Trik 软件主体公司伊普西龙科技(北京)有限公司告上法庭,北京互联网法院现已立案审理,正在等待开庭。 ^[9]

表 3-1 图片生成类 AI 著作权侵权案件汇总表

2. 文字生成类 AI 著作权侵权案件

2023 年 6 月 28 日,两名作家 Paul Tremblay 和 Mona Awad 提起第一起针对 ChatGPT 的版权侵权之诉,其主张 Open AI 未经授权使用了原告的图书对 ChatGPT 进行模型训练。^[10] 随后的半年时间中,数十名著作权人在曼哈顿、旧金山等多地法院对 Open AI 提起版权诉讼。2023 年的 ChatGPT 系列诉讼案件以《纽约时报》的起诉落下帷幕^[11]。

^[7] See Getty lawsuit against Stability AI to go to trial in the UK, published in The Verge, available at: <https://www.theverge.com/2023/12/4/23988403/getty-lawsuit-stability-ai-copyright-infringement>.

^[8] See Getty Images sues AI art generator Stable Diffusion in the US for copyright infringement, published in The Verge, available at: <https://www.theverge.com/2023/2/6/23587393/ai-art-copyright-lawsuit-getty-images-stable-diffusion> (最后访问时间: 2024 年 1 月 16 日)。

^[9] 《多位创作者起诉小红书偷画师作品“炼”AI 模型涉侵权,已获立案》,载于《知产前沿》, http://www.ipforefront.com/article_show.asp?id=3702&BigClass=%E8%B5%84%E8%AE%AF。

^[10] See Lawsuit says OpenAI violated US authors' copyrights to train AI chatbot, published in Reuters, available at: <https://www.reuters.com/legal/lawsuit-says-openai-violated-us-authors-copyrights-train-ai-chatbot-2023-06-29n> (最后访问时间: 2024 年 1 月 16 日)。

^[11] See NY Times sues OpenAI, Microsoft for infringing copyrighted works, published in Reuters, available at: <https://www.reuters.com/legal/transactional/ny-times-sues-openai-microsoft-infringing-copyrighted-work-2023-12-27n> (最后访问时间: 2024 年 1 月 16 日)。

序号	涉案AI 及其开发商	案件基本情况
1	ChatGPT (OpenAI)	● 2023 年 6 月 28 日, 马萨诸塞州作家 Paul Tremblay 和 Mona Awad 在旧金山联邦法院起诉 OpenAI, 指控 OpenAI 滥用他们的作品来“训练”其开发的生成式人工智能系统 ChatGPT。 ● 两名作家表示, ChatGPT 未经许可挖掘了从数千本书籍中复制的数据, 侵犯了作者的版权。^[12] ● 2023 年 8 月 28 日, OpenAI 请求法院驳回与版权侵权、不公平竞争、疏忽和不当得利有关的指控, 但并没有要求驳回直接版权侵权的核心主张。^[13]
2	ChatGPT (OpenAI) Meta	● 2023 年 7 月 7 日, 喜剧演员 Sarah Silverman 和作家 Richard Kadrey, Christopher Golden 向旧金山联邦法院提起集体诉讼, 指控 Facebook 母公司 Meta 和 ChatGPT 制造商 OpenAI 使用受版权保护的材料来训练聊天机器人。 ● Silverman、Kadrey 和 Golden 指控 Meta 和 OpenAI 在未经授权的情况下使用他们的书籍来开发所谓的大型语言模型, 这些模型的制造商将其宣传为通过复制人类对话来实现任务自动化的强大工具。^[14] ● 2023 年 8 月 28 日, OpenAI 请求法院驳回与版权侵权、不公平竞争、疏忽和不当得利有关的指控,

^[12] See Lawsuit says OpenAI violated US authors' copyrights to train AI chatbot, published in Reuters, available at: <https://www.reuters.com/legal/lawsuit-says-openai-violated-us-authors-copyrights-train-ai-chatbot-2023-06-29/> (最后访问时间: 2024 年 1 月 16 日)。

^[13] See AI and Copyright: OpenAI Asks Court to Partially Dismiss Authors' Copyright Lawsuits, published in LEGAL • IO, available at: <https://www.legal.io/articles/5442785/AI-and-Copyright-OpenAI-Asks-Court-to-Partially-Dismiss-Authors-Copyright-Lawsuits#:~:text=OpenAI%20took%20aim%20at%20what%20it%20called%20%22ancillary%22,copyright%20infringement%2C%20unfair%20competition%20negligence%20and%20unjust%20enrichment> (最后访问时间: 2024 年 1 月 16 日)。

^[14] See Sarah Silverman sues Meta, OpenAI for copyright infringement, published in Reuters, available at: <https://www.reuters.com/legal/sarah-silverman-sues-meta-openai-copyright-infringement-2023-07-09/> (最后访问时间: 2024 年 1 月 16 日)。

序号	涉案AI 及其开发商	案件基本情况
		但并没有要求驳回直接版权侵权的核心主张。 ^[15]
3	ChatGPT (OpenAI)	● 2023 年 9 月 11 日,普利策小说奖得主迈克尔·夏邦 (Michael Chabon) 在内的 5 名美国作家,在旧金山联邦法院对美国人工智能公司 OpenAI 提起集体诉讼,指控其滥用他们的作品训练 ChatGPT。 ● 该诉讼称,书籍、戏剧和文章等作品作为“高质量长篇写作的最佳范例”,对 ChatGPT 的培训尤有价值。作家们认为该系统可以准确地总结他们的作品,并生成模仿其风格的文本。^[16]
4	LLaMa 大模型 (Meta)	● 2023 年 9 月 12 日,普利策小说奖得主迈克尔·夏邦 (Michael Chabon) 在内的 5 名美国作家,在旧金山联邦法院对美国公司 Meta 提起诉讼,指控其 LLaMa 大模型未经许可使用其作品进行训练,并且从中获利。^[17]
5	ChatGPT (OpenAI)	● 2023 年 9 月 19 日,美国作家协会在曼哈顿联邦法院正式对人工智能公司 OpenAI 发起集体诉讼,指控后者在训练人工智能模型中滥用受版权保护的作品。 ● 美国作家协会声称, OpenAI 公司在未经许可或未支付报酬的情况下批量复制他们的作品,并将受版权保护的材料输入大型语言模型。“这些算法是被告企业的核心,而算法的核心则来源于大规模、

^[15] See AI and Copyright: OpenAI Asks Court to Partially Dismiss Authors' Copyright Lawsuits, published in LEGAL • IO, available at: <https://www.legal.io/articles/5442785/AI-and-Copyright-OpenAI-Asks-Court-to-Partially-Dismiss-Authors-Copyright-Lawsuits#:~:text=OpenAI%20took%20aim%20at%20what%20it%20called%20%22ancillary%22,copyright%20infringement%2C%20unfair%20competition%20negligence%20and%20unjust%20enrichment> (最后访问时间: 2024 年 1 月 16 日)。

^[16] See More writers sue OpenAI for copyright infringement over AI training, published in Reuters, available at: <https://www.reuters.com/technology/more-writers-sue-openai-copyright-infringement-over-ai-training-2023-09-11/> (最后访问时间: 2024 年 1 月 16 日)。

^[17] See One day after suing Open AI, Michael Chabon and other authors sue Meta, published in HEARST newspaper, available at: <https://www.sfchronicle.com/tech/article/michael-chabon-sues-meta-llama-chatbot-18362336.php> (最后访问时间: 2024 年 1 月 16 日)。

序号	涉案AI 及其开发商	案件基本情况
		系统性的窃取。” ^[18]
6	Claude 模型 (Anthropic)	2023 年 10 月 18 日，环球音乐集团 (Universal Music Group, UMG) 和其他出版商起诉美国人工智能公司 Anthropic，指控 Anthropic 不仅未经许可分发受版权保护的歌词等材料，而且还利用这些材料来训练其语言模型 Claude。 - 起诉书称，Anthropic 至少在 500 首歌曲的歌词上侵犯了出版商的权利。这些歌曲包括 Beach Boys 的《God Only Knows》、滚石乐队的《Gimme Shelter》以及 Mark Ronson 和 Bruno Mars 的《Uptown Funk》以及碧昂丝的《光环》。^[19]
7	ChatGPT (OpenAI)	2023 年 12 月 20 日，由 11 名非小说类作家组成的团体加入了在曼哈顿联邦法院提起的诉讼，指控 OpenAI 和微软滥用作者撰写的书籍来训练 OpenAI 的聊天机器人 ChatGPT 和其他人工智能背后的模型。 - 这些作家包括普利策奖获得者泰勒-布兰奇 (Taylor Branch)、斯泰西-希夫 (Stacy Schiff) 和凯-伯德 (Kai Bird) ——他们共同撰写了罗伯特-奥本海默的传记《美国普罗米修斯》(American Prometheus)，该传记今年被改编为热门电影《奥本海默》(Oppenheimer)，他们表示，这些公司使用他们的作品来训练 OpenAI 的 GPT 大型语言模型，侵犯了他们的版权。^[20]

^[18] See Case 1:23-cv-08292, available at: <https://authorsguild.org/app/uploads/2023/09/Authors-Guild-OpenAI-Class-Action-Complaint-Sep-2023.pdf> (最后访问时间: 2024 年 1 月 16 日)。

^[19] See Universal Music sues AI company Anthropic for distributing song lyrics, published in The Verge, available at: <https://www.theverge.com/2023/10/19/23924100/universal-music-sue-anthropic-lyrics-copyright-katy-perry> (最后访问时间: 2024 年 1 月 16 日)。

^[20] Pulitzer-winning authors join OpenAI, Microsoft copyright lawsuit, published in Reuters, available at: <https://www.reuters.com/legal/pulitzer-winning-authors-join-openai-microsoft-copyright-lawsuit-2023-12-20/> (最后访问时间: 2024 年 1 月 16 日)。

序号	涉案AI 及其开发商	案件基本情况
8	ChatGPT (OpenAI) Copilot (Microsoft)	● 2023 年 12 月 27 日,《纽约时报》起诉 OpenAI 和微软,指控他们未经许可使用该报数百万篇文章来帮助训练聊天机器人向读者提供信息。 ● 《纽约时报》是美国第一家就其作品相关版权问题起诉 OpenAI 和微软的主要媒体机构。^[21]

表 3-2 文字生成类 AI 著作权侵权案件汇总表

3. 声音生成类 AI 著作权侵权案件

序号	涉案AI 及其开发商	案件基本情况
1	魔音工坊 (小问科技)	● 2023 年 12 月 12 日,北京互联网法院公开审理国内首例主张 AI 生成声音产品侵犯人格权案。 ● 原告殷女士主张其发现一款名为“魔音工坊”的 APP 将自己的声音 AI 化后生成配音作品,并以“魔小璇”的名义对外出售牟利,认为其声音权利被侵犯,将“魔音工坊”APP 开发商小问科技以及相关作品授权方微软中国等诉至法院。 ● 被告均否认侵权,小问科技认为,其使用的声音产品具有合法来源,来源于微软中国;微软中国表示,其使用声音的合法授权来自于中广影音;中广影音自述原告曾与其有过合作,约定经原告录制形成的作品著作权属于其所有。对此,原告认为,本案诉请的事实依据为侵犯人格权,而非著作权侵权,不应以被告有原告关于著作权的授权,当然地推定被告有原告人格权的授权。 ● 案件未当庭宣判,该案审判长介绍, AI 化声音作品授权链条的认定,如认定侵权对于侵权责任的承担方式,侵权损害赔偿、范围等,均为审理时需要考虑的难点问题。目前,该案仍在进一步审理中。^[22]

^[21] NY Times sues OpenAI, Microsoft for infringing copyrighted works, published in Reuters, available at: <https://www.reuters.com/legal/transactional/ny-times-sues-openai-microsoft-infringing-copyrighted-work-2023-12-27/>.

^[22] 《北京互联网法院审理全国首例“AI 声音侵权案”》,载于《北京互联网法院》,可见: https://mp.weixin.qq.com/s/4eg9tb3AAMsEiXph_kihXw.

以上案件均为 AI 训练数据产生的著作权侵权纠纷，可见 AI 训练数据的合规性对避免诉讼纠纷尤为重要。像 Stable Diffusion^[23] 这类的人工智能创作工具，在训练的过程中需要使用大量的数据，这些数据中可能包含他人的作品，上述数据是否可以作为 AI 训练过程中的训练素材，使用上述数据是否可以构成著作权法上规定的“合理使用”呢？目前各国均未给出明确的指引，还需要进一步等待上述案件的判决结果。

（二）AIGC 生成物的可版权性

ChatGPT 的迅速发展引发了大量的法律问题，其中关于如何认定人工智能生成内容（AIGC 生成物）的著作权，不仅仅是目前 ChatGPT 将要面临的问题，更是整个 AIGC 产业需要面对的问题。本部分将从 ChatGPT 的相关技术本身入手讨论 AIGC 生成物的可版权性以及提出目前法律体系下的 AIGC 生成物知识产权归属条款的建议。

1. ChatGPT 与人工干预

人工智能可以追溯至上个世纪 50 年代计算机之父提出的图灵测试^[24]，其是通过使机器与人类进行相同的对话来判断机器是否具备了与人类相同的智力水准。而这之后计算机行业就已经开始了对对话式人工智能的开发，以期通过图灵测试。此时的人工智能还在使用模式匹配（Pattern Matching）^[25] 方式设计对话式机器人，它只需要提前设置人类可能提出的关键词，并且按照预设答案进行输出，现在很多的智能客服都仍在使用这样的算法模式。

但随着模式匹配式算法受阻（很难通过图灵测试），图灵当时提出的另一个概念“机器学习”开始得到了不断的延伸，发展出了“监督学习”“无监督学习”“强化学习”“深度学习”等不同类型的机器学习方式，最终以“人工神经网络”的架构下对不同类型的

^[23] Stable Diffusion 是图像生成模型，由初创公司 Stability AI 和慕尼黑大学机器视觉学习组和 AI 视频剪辑技术创业公司 Runway 合作开发。

^[24] 图灵测试（英语：Turing test）是英国电脑科学家艾伦·图灵于 1950 年提出的思想实验，目的是测试机器能否表现出与人一样的智力水准。测试时测试者透过电脑键盘输入文本并透过屏幕输出文本。

^[25] 模式匹配是数据结构中字符串的一种基本运算，给定一个子串，要求在某个字符串中找出与该子串相同的所有子串，这就是模式匹配。也就是给定一个词，然后可以调取一个已经设置好的预案。例如现在很多的银行、电商平台的客服机器人都是基于模式匹配完成的问题回复，问它一个订单的退货方式，他可以发出预设好的退货方式的介绍。

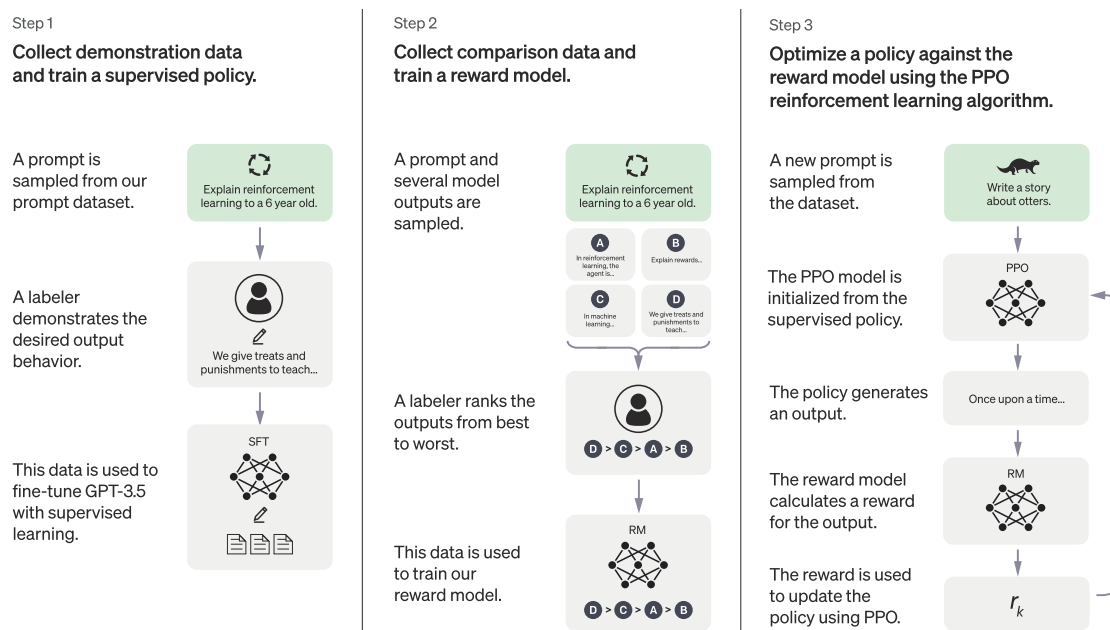
机器学习方式进行结合。

而机器学习的学习能力，也伴随着 Google 的研究者于 2017 年在《Attention Is All You Need》一文中提出的 Transformer 模型，得到了指数级的学习速度的提升，目前 ChatGPT 的“T”说的就是 Transformer。

ChatGPT 的开发公司 OpenAI 公司 2018 年在 Transformer 模型的基础上，进一步研发了 GPT (Generative Pre-trained Transformer)，并且以此为基础推出了 GPT、GPT-2、GPT-3、ChatGPT、ChatGPT-4：

模型	发布时间	参数量	特点
GPT	2018 年 6 月	1.17 亿	无监督学习（基本无需人类干预）
GPT-2	2019 年 2 月	15 亿	使用无监督的预训练模型做有监督的任务
GPT-3	2020 年 5 月	1750 亿	海量参数
ChatGPT（基于 2022 年 3 月的 GPT-3.5）	2022 年 11 月	1750 亿（截止到 2021 年的数据）	增加了“人工反馈强化学习”（增加了人工干预）
ChatGPT-4	2023 年 12 月	1.8 万亿	增加了“多模态输入”功能，可以接受图像和文本的输入，可以理解图像的内容并生成相应的文本结果

表 3-4 GPT 版本升级参数对比表

图 3-1 ChatGPT 的训练机制^[26]

从上述 ChatGPT 技术发展过程中可以看到，ChatGPT 经历了从基本无需人类干预，到需要一定的人工反馈机制的强化学习的变化过程，通过上图可以看出这种“人工反馈强化学习”是以如下方式进行的：

- OpenAI 雇佣的 40 名人工训练员（labeler）对使用 OpenAI 的 Playground 的用户随机 prompt（就是指令或问题），给出指定 prompt 的高质量答案，然后利用这些高质量答案对 GPT-3 进行有监督的微调（Supervised FineTune, SFT）；
- 人工训练员对第一阶段的模型自动生成的回复按照开发者既定的标准（例如相关性、富含信息性、有害信息等诸多标准）进行综合排序，训练奖励模型（RewardModel, RM）；
- 使用 RM 作为强化学习的优化目标，利用 PPO 算法微调 SFT 模型。

因此可以看到，ChatGPT 中的“人工反馈强化学习”需要人工训练员对 ChatGPT 生成的内容进行不断地干预，这种人工反馈机制就是 Moss 和图恒宇说的“人在回路”（Human-in-the-Loop）的一种方式，简而言之就是机器通过需要监督或者不需要监督的大量训练数据的自主学习后，需要人类给予它产生的结果进行反馈，这种不需要非常多

^[26] See at: <https://openai.com/blog/chatgpt/>（最后访问时间：2024 年 1 月 16 日）。

的，但又十分必要的人类反馈（“人工反馈强化学习”），将使得人工智能更加像人类。这是 GPT-3 在已经拥有了海量（1750 亿）参数后无法继续优化其生成结果，而 OpenAI 公司寻找到的使 GPT 生成结果更像人类的解决方法，这直接生成了目前在全球备受关注的 ChatGPT。

2. 中美 AIGC 生成物可版权性的探索^[27]

目前基本大部分国家的著作权法都不认可除了人类以外其他的主体完成的创作可以适用著作权法获得保护，但人工智能的出现，以及其截止目前的不断发展，都使得各个国家的立法、司法在积极探讨 AIGC 生成物的著作权法项下的保护方式。其中的重要原因还在于，人类目前在人工智能中还起着非常重要的作用，以上面的 ChatGPT 技术发展过程为例：

- 人工智能训练过程中：人工智能开发过程中的“人工反馈强化学习”或者“人在回路”的人工干预的机器学习方式，使得 AIGC 生成物更像人类完成的创作，准确说更像进行了具体干预的那一个或者几个人类（人工训练员）完成的创作。例如现在 ChatGPT 的聊天回复风格，语言遣词造句都和人工反馈强化学习过程中的人工训练员的风格、遣词造句有着莫大的关联，那么这部分人类“创作”是否可以得到著作权法的认可？

- 用户使用过程中：用户在使用 ChatGPT 过程中，与 ChatGPT 的对话，促使它进行回复，所以 ChatGPT 的回复是用户与 ChatGPT 共同完成的结果，这部分真实人类用户的“提问”（创作），是否可以得到著作权法的认可？

- 用户使用 AIGC 生成物：除了 ChatGPT 以外，现在还有包括图形类、音乐类的 AIGC 产品，在 AI 完成“创作”之后，用户 / 使用者可以进一步对其生成物进行调整，这部分的人工干预的部分如何得到著作权法的认可？

由于上述人工智能存在的人工干预部分，因此存在着这部分人工创造以著作权法进行保护的讨论必要性，目前包括中国和美国都通过司法或者行政命令的方式在进行着这

^[27] 在著作权法项下探讨任何创作形式的“可版权性”，前提一定是其具有“独创性”，不能是简单机械的复制、拍摄等的结果。当然目前 AIGC 生成的无论诗歌、图片、音乐，甚至是短小的视频都已经与人类完成的具有独创性的作品无异，因此我们探讨的前提是这部分具有独创性的 AIGC 生成物。

方面的探索：

1.1 中国

对于 AIGC 生成物是否构成著作权法下的作品，中国法院存在下述两种截然相反的司法认定结论：北知案件中，法院认为，AI 生成的文章不是自然人创作的，不构成有著作权的作品；在深圳南山案件中，法院认为，涉案 AI 生成文章产生的过程中，存在人工干预，构成有著作权的作品；在北互的案件中，法院认为，涉案 AI 生成图片的过程中，存在人工干预，构成有著作权的作品。

序号	案件基本信息	法院认定
1	北京菲林律师事务所诉 北京百度网讯科技有限公司侵害文字作品著作权案（北知案件） 案号： （2019）京 73 民终 2030 号 审理法院： 北京知识产权法院（二审法院） 涉案 AI： 威科先行库	该分析报告系威科先行库利用输入的关键词与算法、规则和模板结合形成的，某种意义上讲可认定威科先行库“创作”了该分析报告。由于分析报告不是自然人创作的，因此，即使威科先行库“创作”的分析报告具有独创性，该分析报告仍不是著作权法意义上的作品，依然不能认定威科先行库是作者并享有著作权法规定的相关权利。
2	深圳市腾讯计算机系统有限公司诉上海盈讯科技有限公司侵害著作权及不正当竞争纠纷案（南山案件） 案号： （2019）粤 0305 民初 14010 号	从整个生成过程来看，如果仅将 Dreamwriter 软件自动生成涉案文章的这两分钟时间视为创作过程，确实没有人的参与，仅仅是计算机软件运行既定的规则、算法和模板的结果，但 Dreamwriter 软件的自动运行并非无缘无故或具有自我意识，其自动运行的方式体现了原告的选择，也是由 Dreamwriter 软件这一技术本身的特性所决定。原告主创团队相关人员的上述选择与安排符合著作权法关于创作的要

序号	案件基本信息	法院认定
	审理法院： 广东省深圳市南山区人民法院 涉案AI： 腾讯的 Dreamwhiter 自动撰写软件	求，应当将其纳入涉案文章的创作过程。因此，从涉案文章的生成过程来分析，该文章的表现形式是由原告主创团队相关人员个性化的安排与选择所决定的，其表现形式并非唯一，具有一定的独创性。
3	李 xx 与刘 xx 侵害作品署名权、信息网络传播权纠纷（北互案件） 案号： （2023）京 0491 民初 11279 号 审理法院： 北京互联网法院 涉案AI： Stable Diffusion	● 关于智力成果的认定：“从原告构思涉案图片起，到最终选定涉案图片止，这整个过程来看，原告进行了一定的智力投入，比如设计人物的呈现方式、选择提示词、安排提示词的顺序、设置相关的参数、选定哪个图片符合预期等等。涉案图片体现了原告的智力投入，故涉案图片具备了“智力成果”要件。” ● 关于“独创性”的认定：“原告对于人物及其呈现方式等画面元素通过提示词进行了设计，对于画面布局构图等通过参数进行了设置，体现了原告的选择和安排。另一方面，原告通过输入提示词、设置相关参数，获得了第一张图片后，其继续增加提示词、修改参数，不断调整修正，最终获得了涉案图片，这一调整修正过程亦体现了原告的审美选择和个性判断……涉案图片并非“机械性智力成果”。在无相反证据的情况下，可以认定涉案图片由原告独立完成，体现出了原告的个性化表达。综上，涉案图片具备“独创性”要件。 ● 关于作品的认定：“人们利用人工智能模型生成图片时……本质上仍然是人利用工具进行创作，即整个创作过程中进行智力投入的是人而非人工智能模型。鼓励创作，被公认为著作权制度的核心目的……人工智能生成图

序号	案件基本信息	法院认定
		● 片，只要能体现出人的独创性智力投入，就应当被认定为作品，受到著作权法保护。” ● 关于美术作品的认定：“涉案图片是以线条、色彩构成的有审美意义的平面造型艺术作品，属于美术作品。同时，涉案图片在可以归属到具体作品类型时，没有适用“其他作品条款”保护的必要性，其不属于“符合作品特征的其他智力成果”。 关于著作权的认定：“原告是直接根据需要对涉案人工智能模型进行相关设置，并最终选定涉案图片的人，涉案图片是基于原告的智力投入直接产生，且体现出了原告的个性化表达，故原告是涉案图片的作者，享有涉案图片的著作权。”

表 3-5 AIGC 生成物是否构成著作权法下的作品司法案例汇总表

深圳南山案件的法院逻辑下，法官认为：涉案文章由原告深圳腾讯公司主创团队人员运用 Dreamwriter 软件生成，主创团队在数据输入、触发条件设定、模板和语料风格的取舍方面的安排与选择，属于与涉案文章的特定表现形式之间具有直接联系的智力活动，该文章的表现形式是由原告主创团队相关人员个性化的安排与选择所决定的，因此，涉案作品具有一定独创性，属于中国著作权法所保护的文学作品。也就是说，在该案中法院认定的 AIGC 生成物，并不是完全脱离了人类智力活动、纯粹由人工智能生成的文字内容，不是人工智能自主创造完成的，而仅仅是一种由人工智能辅助完成的人类智力活动成果。在这个意义上看，有人工智能参与形成的产物当然是可以受到著作权法保护的^[28]。

深圳南山案件中，法院对 AIGC 生成物的理解不仅包括了最终的生成文章，甚至包

^[28] 该案件的认定逻辑也得到了最高院的支持，以及将其在 WIPO 会议上进行了说明。最高法院法官在 WIPO 的发言稿：https://www.wipo.int/export/sites/www/about-ip/en/artificial_intelligence/conversation_ip_ai/pdf/ms_china_1_zh.pdf。

括了其前期训练过程，而使得在算法训练工程中的人工干预如果加入到生成物的著作权考量因素中，很难将人工干预从 AIGC 生成物中进行彻底排除，从而 AIGC 生成物将全方位获得著作权法保护。

因此，在实际判断 AIGC 生成物的可版权性时，首先应当将 AIGC 生成物的人工干预过程与 AIGC 训练过程中的人工干预相分离，针对 AI 与用户互动、用户提供 prompt，AI 进行产生的内容、用户需要在 AIGC 生成内容后再行编辑调整的内容中人工干预的部分，需要给予著作权法上的肯定，而针对 AIGC 训练过程中的人工干预的部分可以通过其他权利 / 权益进行保护，除非训练过程中的确产生了人工训练员的创造性劳动，例如通过打标签等方式对大量数据进行了有独创性的整理、编辑，进而可以构成汇编作品。

1.2 美国

2023 年 3 月 16 日，美国版权局（USCO）在《联邦公报》上发布了一则“含有 AI 生成元素的作品”的版权注册指南（以下简称“《指南》”）^[29]，进一步澄清了实践中 USCO 对于 AI 生成元素进行审查和注册的基本政策。本部分主要根据 USCO 公布的《指南》，说明目前 USCO 针对 AI 生成物在版权法上的态度，以期为中国 AI 生成物可版权性提供参考和借鉴。

根据《指南》，USCO 对包含 AI 生成元素的生成物的可版权性的判断标准是“作品”中的传统作者要素（文学、艺术或音乐表达或选择、编排等要素）是否为人类完成。

如果人类仅通过向 AI 工具进行提示（Prompt）（例如我们向 ChatGPT 下指令，要求其完成一段“莎士比亚风格的诗歌”），这是无法使得生成物具有版权性的。

但另一方面，人类如果对 AI 生成元素进行了充分地修改、选择、安排，并且这些修改、选择、安排具有独创性，这将使得整个生成物可以构成版权法下的作品，这和人类艺术家使用 Adobe Photoshop 编辑、修改的图像一样，但版权也仅保护其中人类完成的部分。

^[29] See Copyright Registration Guidance: Works Containing Material Generated by Artificial Intelligence, published in U.S. Government Information Website, available at: <https://www.govinfo.gov/content/pkg/FR-2023-03-16/pdf/2023-05321.pdf>（最后访问时间：2024 年 1 月 16 日）。

(1) 美国版权法的法律规定以及 USCO 的在先案件

《美国版权法》第 102 条(a)规定,以任何现在已知的或者以后出现的物质表达方式——通过此种方式可以直接或者借助于机械或装置可感知、复制或以其他方式以传播作品——固定的独创作品,依本篇受版权保护。

可以看出在美国版权法的作品定义条款中并没有明确规定仅保护人类创作的作品,而不保护 AI 生成物。

而作为版权注册和管理的政府机构,USCO 至今为止已经针对 AI 生成物的版权注册在如下案件中进行过回复:

USCO 版权注册态度	案件进展
【Thaler v. Perlmutter】 否认非人类作者身份否认利用 AI 技术生成的“作品”进行版权注册	● 2018 年 11 月 3 日,Stephen Thaler 向美国版权局提出作品版权登记申请,作品作者为“创意机器 (Creativity Machine)”,Thaler 在给版权局的申请中称该作品“由运行在机器上的计算机算法自动创建”。 ● 2019 年 8 月 12 日,美国版权局拒绝了注册请求,理由是“缺乏支持版权声明所需的人类作者身份”。2020 年 3 月 30 日,版权局拒绝了首次复议请求,提出其不会放弃最高法院及地方法院长期以来对《版权法》的解释,即只有由人类作者创作的作品才符合版权保护的法律要求。2022 年 2 月 14 日,美国版权局再次拒绝了复议请求。^[30] ● 2022 年 6 月 2 日,Stephen Thaler 提起诉讼,请求法院下达命令,要求撤销美国版权局拒绝 AI 生成作品版权登记的决定。目前,该案正在美国华盛顿特区地方法院审理。

^[30] See Decision Affirming Refusal of Registration of a Recent Entrance to Paradise at 2 (Feb. 14, 2022), published in U.S. Copyright Office, available at: <https://www.copyright.gov/rulings-filings/review-board/docs/a-recent-entrance-to-paradise.pdf> (最后访问时间: 2024 年 1 月 16 日)。

USCO 版权注册态度	案件进展
【Zarya of the Dawn】 否认非人类作者身份区分“作品”中的 AI 生成内容与人类创造内容	● 2022 年 9 月 15 日, Kashtanova 对其作品《Zarya of the Dawn》进行了版权注册申请, 描述的作品类型为图画小说 (comic book)。 ● 2023 年 2 月 21 日, 美国版权局审查了该项注册^[31], 得出结论认为, Kashtanova 是作品文本的作者, 同时也是作品书面和视觉元素的选择、协调和排列的作者, Kashtanova 的作者身份受版权保护。然而, 作品中由 Midjourney 技术生成的图像并非人类创作, 因此版权局将撤销原有的版权证明, 并根据原作者的申请重新颁发相应的版权证明。

表 3-6 USCO 针对 AI 生成物的版权注册的回复汇总表

(2) 美国版权局 (USCO) 最新《指南》要旨

在 USCO 已经发生如上案件后, 特别是 2023 年 2 月 21 日其就《Zarya of the Dawn》进行了答复后, 还收到了其他在申请文件中注明或者没有注明 AI 生成的注册材料, 基于这些情况, USCO 认为有必要就 AI 生成元素申请注册版权进行说明, 因此发布了这次的《指南》。

● 明确“受版权保护作品的作者必须是人类”

本次 USCO 公布的《指南》首先明确: 受版权保护作品的作者必须是人类, 缺乏人类作者身份的作品将不会被准予注册。主要原因在于“美国《宪法》和《版权法》使用的 author 一词, 排除了非人类作者, 并且根据在先司法案例^[32], USCO 长期以来一直要求作品必须是人类作者的产物”。

^[31] See Cancellation Decision re: Zarya of the Dawn (VAu001480196) at 2 (Feb. 21, 2023), published in U.S. Copyright Office, available at: <https://www.copyright.gov/docs/zarya-of-the-dawn.pdf> (最后访问时间: 2024 年 1 月 16 日)。

^[32] See Burrow-Giles Lithographic Co. v. Sarony, Naruto v. Slater (猕猴自拍案)。

同时 USCO1973 年第 1 版《版权局实践手册》(Compendium of Copyright Office Practices, 以下称“《手册》”)中说明,不会注册没有“来自于人类”的生成物。1984 年第 2 版《手册》中进一步说明,“authorship”这个词说明,要获得版权保护的作品,其必须来自人类。在最新版本的《手册》中,USCO 说明,要获得版权保护,作品必须由人类创作,并且不会注册由机器或仅依靠随机或自动操作的机械过程产生的内容,这些过程没有人类作者的独创性参与。

● 包含 AI 生成元素的生成物是否受版权法保护的判断标准

《指南》介绍了包含 AI 生成元素的生成物需要在个案中判断如下情况,才能确定其是否可以受到版权法保护:

“作品”中的传统作者要素(文学、艺术或音乐表达或选择、编排等要素)是否是人类完成。如果是由机器生成,则会被认为缺乏人类创作,将无法得到版权局的注册。

如果 AI 根据接收到的人类提示(Prompt),生成复杂的书面、视觉或者音频内容,则会被认为传统作者要素(文学、艺术或音乐表达或选择、编排等要素)是由 AI 完成的,而不是人类用户。这里的人类“提示”(Prompt)更像是人类向 AI 进行指示,AI 根据这个指示决定如何输出。例如我们向 AI 工具下指令,要求其完成一段“莎士比亚风格的诗歌”,人类可以看到一首莎士比亚风格的诗歌,而 AI 决定诗歌的韵脚、每行的词语和文本的结构,这时这些内容将不被认为是人类创作完成的。

但相反的情况是,人类对 AI 生成元素进行了充分地修改、选择、安排,并且这些修改、选择、安排具有独创性,这将使得整个生成物可以构成版权法下的作品,这和人类艺术家使用 Adobe Photoshop 编辑、修改的图像^[33]一样,但版权也仅保护其中人类完成的部分。

● 申请人需要提交的版权注册材料

为了实现对上述判断标准的审查,USCO 要求:

^[33] USCO 还举例了制作音乐时使用吉他效果器,但这由于涉及中国法与美国法在录制音乐时认定作品和制品的区别,导致在中国法下比较难理解。

申请人有义务披露提交注册的作品中包含 AI 生成的内容，并就人类作者对作品的贡献提供简要解释。

不能仅因为在创作作品时使用了 AI 就可以将 AI 工具或者提供 AI 工具的公司作为作者或者共同创作者。

已提交含有 AI 生成元素作品注册申请的申请人，应检查提供给版权局的资料是否充分披露了这些信息，如果没有，应该采取措施纠正信息，否则可能导致注册无效。

(3)《指南》要旨解读

●《指南》对于美国的 AI 生成元素是否能在美国得到版权保护具有重要影响

根据《美国版权法》第 411a 条，起源国是美国的作品，在美国版权局完成“注册”^[34]是权利人提起侵权诉讼的前提条件，也就是说，作品必须提交注册才能在美国得到充分的版权法保护（起源国是中国等《伯尔尼公约》成员国的作品在美国无需先行完成注册就可以获得版权法保护）。

此外，《指南》把提交注册的作品是否由 AI 生成等信息，作为申请人必须披露向版权局内容，具有可执行的效力，因为根据美国法院在先判例，如果版权人在注册过程中有欺诈或不诚实行为，或者故意向版权局隐瞒了某些事实，可能会导致版权注册无效。

因此我们关注 USCO 在版权注册时的态度对理解美国法对 AI 生成物可版权性的态度具有重要意义。

●《指南》规定的包含 AI 生成元素的生成物的可版权性判断标准可以帮助我们梳理目前常见的 AI 工具的版权问题

由于在《指南》中并没有“一刀切”式的否定 AI 生成元素的可版权性，而是考察、对比 AI 工具的运作与人类创作的关系，进而判断人类是否在生成物中进行了具有独创性的创造，而使得该生成物存在受到版权法保护的可能。

^[34] 特别是在 *Fourth Estate* 再审案后，美国最高法院进一步确定了“在提起著作权侵权诉讼之前，美国作品的著作权人必须完成著作权登记。”另外实际上根据《美国版权法》以及美国法院的在先判例，还存在有限的例外情况，版权人可以在未完成版权注册的情况下提起版权侵权诉讼；例如针对电影或者音乐作品，可以在进行预注册申请后提起侵权诉讼；针对实时广播侵权可以无须实现完成版权注册。

根据《指南》，USCO 对包含 AI 生成元素的生成物的可版权性的判断标准是“作品”中的传统作者要素（文学、艺术或音乐表达或选择、编排等要素）是否为人类完成。

如果人类仅通过向 AI 工具进行提示（Prompt）（例如我们向 ChatGPT 下指令，要求其完成一段“莎士比亚风格的诗歌”），这是无法使得生成物具有版权性的。

但另一方面，人类如果对 AI 生成元素进行了充分地修改、选择、安排，并且这些修改、选择、安排具有独创性，这将使得整个生成物可以构成版权法下的作品，这和人类艺术家使用 Adobe Photoshop 编辑、修改的图像一样，但版权也仅保护其中人类完成的部分。正如 USCO 在回复《黎明的查莉娅》（Zarya of the Dawn）的注册申请时说明的，版权保护不适用于由 AI 绘画工具 Midjourney 生成的部分，而针对这部 AI 辅助完成的漫画中人类完成的部分给予了版权保护。

这一说明可以帮助我们明确目前常见的 AI 工具使用过程中生成物的著作权法项下的问题。

1.3 目前针对 AIGC 产品的知识产权归属条款建议

而目前囿于各个国家的法律基本上不认可 AIGC 生成物的可版权性，因此 AIGC 产品平台针对其 AIGC 生成物的知识产权归属通常通过用户协议的方式进行明确。

平台名称	服务协议链接	条款内容
ChatGPT (AI 生成文字)	https://openai.com/terms/	【注：OpenAI 授权用户使用输出内容】 您可以向服务提供输入（“输入”），并根据输入接收服务生成和返回的输出（“输出”）。输入和输出统称为“内容”。在双方之间，在适用法律允许的范围内，您拥有所有输入，并且在您遵守这些条款的前提下， OpenAI 特此将其对输出的所有权利、所有权和利益转让给您。 OpenAI 可能会根据需要使用内容来提供和维护服务、遵守适用法律

平台名称	服务协议链接	条款内容
		和执行我们的政策。 <u>您对内容负责，包括确保其不违反任何适用法律或本条款。</u>
讯飞星火认知大模型 (AI 生成文字)	https://www.xfyun.cn/doc/spark/ExperienceRules.html#%E7%A C%AC%E5%9B%9B%E6%9 D%A1-%E7%9F%A5%E8%AF%86%E4%BA%A7%E6%9 D%83	与本平台相关的一切著作权、商标权、专利权、商业秘密等知识产权，以及本服务输出的相关所有信息内容（包括但不限于文字、图片、音频、视频、图表、界面设计、版面框架、有关数据或电子文档等）均受法律法规保护，由科大讯飞及其关联方享有完整的知识产权、 <u>所有权等合法权利</u> 。用户仅有权依照本协议约定合法使用本服务及本服务相关内容的权利。 您理解并且同意，除非有相反证明， <u>您使用本平台服务上传、发布或传输内容即代表了您不可撤销地授予科大讯飞及其关联方非排他的、无地域限制的、永久的、免费的许可使用</u> （包括存储、使用、复制、修订、编辑、发布、展示、翻译、分发上述内容或制作派生作品等商业或非商业使用）及可再许可第三方使用的权利，以及可以自身名义对第三方侵权行为取证及提起诉讼的权利。
商量 SenseChat (AI 生成文字)	https://chat.sensetime.com/wb/policy?id=1&lang=zh_CN	商汤是商量硬件及应用的知识产权权利人。本服务的一切著作权、商标权、专利权、商业秘密等知识产权，以及与本服务相关的所有信息内容（包括但不限于文字、图片、音频、视频、图表、界面设计、版面框架、有关数据或电子文档等）均受当地法律法规和相应的国

平台名称	服务协议链接	条款内容
		际条约保护，商汤享有上述知识产权。<u>为避免疑义，用户独立使用硬件记录或生成的内容，其合法权利均由用户本人享有。</u>
通义千问 (AI 生成文字)	https://tongyi.aliyun.com/qianwen/?spm=5176.28326591.0.0.40f76ee1MiRTfC	通义千问应用 Logo、“通义千问”、“tongyi”等文字、图形及其组合，以及通义千问应用的其他标识、徽记、技术文档等的知识产权为我们所有。 除非另有约定或法律法规另有规定的，您享有本服务应用基于您有权上传的内容进而合成的内容知识产权。您同意授权我们可以在提供本服务所必要的范围内使用您上传、生成及合成的内容。
智谱清言 (AI 生成文字)	https://chatglm.cn/agreement	<u>您基于智谱清言生成的内容的版权由您自行维护并对其独立判断后使用</u>，基于智谱清言生成内容产生的任何知识产权问题由您自行处理，北京智谱华章科技有限公司对由此造成的任何损失不负责任，您由此给北京智谱华章科技有限公司造成损失的北京智谱华章科技有限公司有权向您追偿。您不得以任何方式删除、掩藏、修改或替换智谱清言及呈现内容中所附的或包含的任何专有权利声明、其他任何链接等。
百度文心模型 (AI 生成图片)	https://wenxin.baidu.com/wenxin/docs#Yl6th25am	文心大模型 API 服务平台对<u>通过 ERNIE-ViLG AI 作画大模型生成的图像相关权利不做特别声明</u>，但在 ERNIE-ViLG AI 作画大模型体验专区生成的内容均带有水印，用户在付费后使用 API 调用

平台名称	服务协议链接	条款内容
		时生成的图像将全部去除水印。除法律法规和本协议另有相反规定，<u>无水印图像的知识产权及其上的相关权益（包括但不限于知识产权等）将永久归用户所有，无水印图像可用于商业用途，如商业使用无水印图像时出现一切法律风险，将由用户自行承担，本服务平台不拥有该图像的知识产权及其上的相关权益且不对此承担任何责任。</u>
Midjourney (AI 生成图片)	https://docs.midjourney.com/docs/terms-of-service	【注：区分企业用户和免费个人用户】 如果您是年总收入超过 1,000,000 美元的公司的雇员或所有者，并且您代表您的雇主使用服务，则您必须为代表您访问服务的每个人购买“Pro”会员资格为了拥有您创建的资产。如果您不确定您的使用是否符合代表您雇主的资格，请假定符合。 如果您不是付费会员，则您不拥有您创建的资产。Midjourney 根据知识共享非商业性 4.0 署名国际许可(“资产许可”)授予您对资产的许可。(知识产权共享许可 CC4.0 协议链接: https://creativecommons.org/licenses/by-nc/4.0/legalcode) 请注意: Midjourney 是一个开放社区，只要您的图像和提示词发布在公共环境中，其他人就可以使用和重新混合这些图像和提示词。默认情况下，您的图像是公开可见和可重新混合的。如上所述，您授予 Midjourney 允许这样做的

平台名称	服务协议链接	条款内容
		许可。
通义万相 (AI 生成图片)	https://terms.alicdn.com/legal-agreement/terms/common_product_agreement/20230829115805343/20230829115805343.html	【注：用户授权通义万相使用输出内容】 用户在使用本服务中所产生的内容（如评论、个人动态等）的知识产权在法律范围内归相关权利人所有，用户通过本服务发布的内容（下统称“公开发布内容”）需经过审核通过，一经发布即向公众传播和共享。 为持续改善本服务，除非有相反证明，您使用本服上传、发布或传输内容即代表了您有权且同意在全世界范围内，永久性的、不可撤销的、免费的授予我方及其关联方的全部产品或服务以下权利： （1）对该内容的存储、使用、发布、复制、修改、改编、出版、翻译、传播、表演、展示以及据以创作衍生品等权利； （2）将内容的全部或部分编入其他任何形式的作品、媒体技术中的权利； （3）对您的上传、发布的内容进行商业开发的权利； （4）通过有线或无线网络向计算机终端、移动通讯终端、手持数字影音播放设备、电视接收设备等提供信息的下载、点播、数据传输、移动视频业务及其相关宣传和推广等服务的]权利； （5）再授权给其他第三方以上述方式使用的权利。

平台名称	服务协议链接	条款内容
魔音工坊 (AI 生成音频)	https://m.moyin.com/#/protocol	您了解及同意，除非本公司另行声明，本协议项下服务包含的所有产品、技术、软件、程序、数据及其他信息（包括但不限于文字、图像、图片、照片、音频、视频、图表、色彩、版面设计、电子文档）的 <u>所有知识产权（包括但不限于版权、商标权、专利权、商业秘密等）及相关权利均归本公司或其关联公司或第三方权利人所有</u> 。除本公司或第三方权利人明示同意外，您无权复制、传播、转让、许可或提供他人使用上述资源，否则应承担相应的责任。
Riffusion (AI 生成音频)	https://www.riffusion.com/terms	您承认并同意，我们的服务，以及我们的服务上或通过我们的服务显示、提供或生成的所有材料和内容（您的内容和您的输出除外），以及我们的服务所依据和包含的所有软件、算法、代码、技术和知识产权，及 <u>其在全世界范围内的所有知识产权（统称和单独称为我们的“知识产权”），均为我们（或我们的许可人，如适用）的独有财产</u> 。除非本条款另有明确规定，否则本条款中的任何内容均不得视为对任何知识产权或根据任何知识产权创建许可，并且您同意不访问、出售、许可、出租、修改、分发、拷贝、复制、传输、公开展示、公开表演、出版、改编、编辑我们的任何知识产权或根据我们的任何知识产权创建衍生作品。严禁将我们的知识产权用于本条款未明确允许的任何目的。

平台名称	服务协议链接	条款内容
		在您遵守这些条款的前提下，<u>您可以在免版税的基础上将“输出”用于任何合法的非商业目的</u>，前提是您承认并同意： （1）您使用我们的服务和“输出”不会将我们服务中的任何知识产权的所有权转让给您，并且 （2）如果我们自行决定认为您对“输出”的使用可能侵犯任何第三方的权利，我们可以随时通知您，限制您对“输出”的使用或要求您停止使用（并删除任何“输出”副本）。您不得使用“输出”来训练您自己的机器学习模型，或以未经修改的形式转售“输出”（例如，作为 Riffs 样本包）。

表 3-7 AIGC 产品知识产权归属条款汇总表

针对与上述平台类似的用户协议中的知识产权归属（包括知识产权在内的输出物的所有权利 / 权益）约定，在这些用户协议不会因为构成格式合同而无效的前提下，基于意思自治原则，即使 AIGC 生成物不具有知识产权，但仍应当作为“数据”受到保护，此时这些“数据”的使用方式、使用范围、使用权利同样应当适用用户协议中的相关约定。因此在我国 AIGC 生成物是否具有著作权尚不明确的情况下，建议各家 AIGC 公司优化用户协议中针对 AIGC 生成物的知识产权、数据的归属条款约定。

（三）AIGC 生成物侵犯他人著作权

——以 AI 孙燕姿音乐为例

2023 年 4 月起，“冷门歌手”孙燕姿的 AI 翻唱歌曲火遍全网，根据相关媒体^[35]报道，“一个晚上，播放量就达到十几万。连续发了不到十条视频后，粉丝量就已经破千了。”

上述创作者多是自行通过开源的算法模型对孙燕姿的声音进行训练，进而获得孙燕姿声音的算法模型，然后将其声音进行推理，合成到希望“她”“演唱”的歌曲中。但通过 AI 合成的歌手的翻唱歌曲涉及了多项法律上的权利，网友在自行训练、推理后上传 AI 孙燕姿翻唱歌曲也带来诸多侵权法律风险。

1. AI 生成音乐的流程

1.1 需要准备

- 自行准备孙燕姿的人声音轨：即下载孙燕姿演唱过的歌曲，并且将人声音轨从歌曲中剥离出来
- 下载算法模型：so-vits-svc^[36]
- 使用网络上提供的第三方算力：具备一定算力的显卡或使用云训练推理工具
- 自行准备推理所用的声音素材（Cover 的歌曲伴奏）

1.2 操作流程

- 将（1）中的孙燕姿人声音乐放入 so-vits-svc 算法模型中进行训练
- 训练完成后形成 AI 孙燕姿人声模型
- 使用该训练好的 AI 孙燕姿人声模型推理希望“她”“演唱”的歌曲

^[35] 参见《“AI 孙燕姿”刷屏背后，AIGC 如何直面侵权之痛？》，载于：剥椒 TMT（微信 ID: ylwanyia），<https://mp.weixin.qq.com/s/9pF-imax405P4zS02IFRRw>（最后访问时间：2024 年 1 月 16 日）。

^[36] 参见 github 网站，<https://github.com/svc-develop-team/so-vits-svc>，根据 github 的主页说明，该算法模型“只是一个框架项目，本身不具备语音合成的功能，所有的功能都需要用户自己训练模型。同时，该项目没有附加模型，任何二次分布式项目都与该项目的贡献者无关”。

- 合成后上传到视频 / 音频网站

1.3 法律分析

从法律层面分析，上述过程中产生的具有法律后果的行为包括：

- 将孙燕姿演唱的歌曲下载至自己所控制的终端、从歌曲中将孙燕姿人声剥离
- 算法模型的下载使用
- 下载“翻唱”歌曲伴奏行为、“推理”后合成 AI Cover 版本
- 将合成后的 AI Cover 歌曲上传到互联网视频 / 音频网站

2. 创作者行为的侵权风险分析

创作者通过 AI 合成的歌手的翻唱歌曲中可能涉及以下法律权利：

- 著作权：词 / 曲作品的著作权、歌手的表演者权、录音制作者的录音制作者权。
- “声音权”：根据《民法典》第一千零一十九、一千零二十三条，自然人的声音参照肖像权保护，即通常所谓的“声音权”，以技术手段伪造自然人的声音构成侵权行为，未经自然人同意，不得制作、使用、公开自然人的声音。
- 个人信息：声纹属于《个人信息保护法》以及《信息安全技术个人信息安全规范》(35273) 规定属于敏感个人信息，而歌手的声音在向公众发行、演唱音乐作品时是基于自身的行为发布的，但仅用于公众以及相关主体在享受音乐的范围内使用其个人信息，因此如果将其声音未经其授权提取其声纹用于数据训练，是明显地超出歌手的同意范围的使用，构成侵犯其个人信息的行为。

2.1 将孙燕姿演唱的歌曲下载至自己所控制的终端、孙燕姿人声的剥离

(1) 下载行为

《著作权法》中除了适用合理使用、法定许可的情形外，其他的作品使用行为均需获得著作权人的授权。如果将孙燕姿演唱的歌曲仅进行下载用于个人欣赏，显然是符合《著作权法》第二十四条的合理使用的规定的，但是将其用于人工智能训练，显然无法

适用合理使用中的任一情形，同样也无法适用《著作权法》中规定的法定许可，从而需要获得相关词、曲著作权人及录音制作者等著作邻接权人同意后才能使用。

虽然现在包括日本、欧盟等已经将用于人工智能的数据训练增加到合理使用的范畴之内，但也是有明确的限制^[37]，但我国《著作权法》尚未增加相关条款。

因此网友将孙燕姿演唱的歌曲下载至自己所控制的终端用于人工智能训练，显然未经相关权利人授权，侵犯了该歌曲中的词 / 曲作品的著作权（复制权）、孙燕姿的表演者权（复制权）、唱片公司的录音制作者权（复制权）。

（2）人声剥离行为

将孙燕姿的人声音轨从歌曲中剥离的行为，涉及对孙燕姿人声（参照“肖像权”保护）的使用。依据《民法典》第一千零二十三条，对自然人声音的保护，参照适用肖像权保护的有关规定。

根据《民法典》对肖像权的规定，肖像是通过影像、雕塑、绘画等方式在一定载体上所反映的特定自然人可以被识别的外部形象。因此“声音权”可以被理解为通过音频等方式在一定载体上所反映的特定自然人可以被识别的外部形象。“在一定载体上”“可识别性”是其基本的要件。而作为歌曲中人声，是体现在了音频文件上，并且由于孙燕姿声音的特殊性以及其歌曲的可识别性，其歌曲中的人声音轨是符合“声音权”的要件的，其使用应当符合《民法典》对“声音权”的要求。

《民法典》在第一千零二十三条中赋予了“声音权”所谓的“合理使用”^[38]的规定，

^[37] 根据日本《著作权法》第三十条之四，如果为了进行信息分析（指从大量作品或其他信息中提取与其构成的语言、声音、影响或其他元素相关的信息，并进行比较、分类和其他分析的过程）可以适用合理使用，无需获得著作权人授权，但不得是用于享受作品的目的。

欧盟在 2019 年 3 月正式通过欧盟《数字化单一市场版权指令》第 3 条款创设了有条件的文本和数据挖掘制度，为科学研究目的，成员国应当在立法中明确允许研究机构为开展文本与数据挖掘工作所需的各种复制及数据提取活动。

^[38] 《民法典》第一千零二十条 合理实施下列行为的，可以不经肖像权人同意：

- （一）为个人学习、艺术欣赏、课堂教学或者科学研究，在必要范围内使用肖像权人已经公开的肖像；
- （二）为实施新闻报道，不可避免地制作、使用、公开肖像权人的肖像；
- （三）为依法履行职责，国家机关在必要范围内制作、使用、公开肖像权人的肖像；
- （四）为展示特定公共环境，不可避免地制作、使用、公开肖像权人的肖像；
- （五）为维护公共利益或者肖像权人合法权益，制作、使用、公开肖像权人的肖像的其他行为。

但明显将其声音用于数据训练，并不符合其中的任一情形，而需要事先获得“声音权”人的同意。而网友自行将孙燕姿歌曲中的人声从歌曲中剥离，用于数据训练显然是未经授权的行为，从而将构成对自然人孙燕姿“声音权”的侵权。

同时对这种针对其声音的数据训练，显然超出了孙燕姿本人对其声音允许的适用范围，同样构成对其个人信息的侵犯。

2.2 算法模型的下载使用

创作者们所使用的算法模型为 github 上公开的一个开源算法模型 so-vits-svc，因此其使用是无须获得相关权利人的同意的，但需要遵守其主页上的使用条款（Terms of Use）。

在使用条款中，so-vits-svc 开发团队反复提示使用者，需要就其训练中使用的资料及生成物料的授权问题自行负责，svc 开发团队对训练生成物料不承担任何责任。使用条款中还进一步明确指出：该开源算法模型仅供个人交流学习使用，不得进入生产环节；在发布生成物料时，不仅要表明该物料由 AI 生成，还要表明用于生成的原始声音来源，如果该原始声音来源于其他声音合成算法，还需遵守该等算法的使用规则。如使用者拟将 svc 用于生产领域等超出上述范围的使用，需提前联系其开发团队。

2.3 下载“翻唱”歌曲伴奏行为、“推理”后合成 AI Cover 版本

创作者们会将 AI 孙燕姿的声音用于“翻唱”包括梁静茹的宁夏等大量知名歌手原唱的歌曲，为了将 AI 孙燕姿的声音与“翻唱”的歌曲相结合（合成），首先需要下载用于“翻唱”的歌曲伴奏，然后在算法模型上进行推理，将原来歌曲中的原版人声替换为 AI 孙燕姿的人声。

（1）下载“翻唱”歌曲伴奏音频行为

下载“翻唱”歌曲伴奏音频的行为，同样涉及侵犯该歌曲中的词/曲作品的著作权（复制权）、唱片公司的录音制作者权（复制权）。

(2) “推理”后合成 AI Cover 版本

这个过程中存在是否侵犯孙燕姿的表演者权的问题。根据《著作权法》对表演者权的复制权的规定，表演者对其表演享有“许可他人复制录有其表演的录音录像制品，并获得报酬”的权利，但孙燕姿并未实际演唱推理后的 AI Cover 版本歌曲，因此在这个行为中并不存在对孙燕姿表演者权的侵权。

除此之外，因为这个过程中使用了孙燕姿的声音，并且不符合《民法典》中规定的“声音权”“合理使用”，因此网友未经授权使用 AI 孙燕姿的声音，构成了孙燕姿“声音权”的侵权。以及同上述，针对其声音的数据训练，显然超出了孙燕姿本人对其声音允许的适用范围，同样构成对其个人信息的侵犯。

同时，AI Cover 版本中除了歌手的声音以外，还有原版词曲、录音的使用，因此对其未经授权的复制行为，还构成对原版词曲、录音制作者的复制权的侵权。

2.4 将合成后的 AI cover 歌曲上传到互联网视频 / 音频网站

网友的上传行为的侵权是比较显著的，侵犯了原词曲著作权人、录音制作者的信息网络传播权，如果用于直播，还会侵犯广播权。

同时侵犯了孙燕姿的“声音权”、个人信息相关权利。



第四部分 数据合规



四、数据合规

数据合规问题贯穿于 AIGC 的全生命周期，覆盖模型训练、模型应用、模型优化整个过程。《暂行办法》《深度合成规定》尤其强调了数据来源的合法性、个人信息处理的合法性以及对数据质量的保障。本部分包含训练数据合规、数据标注合规、个人信息合规、用户界面合规、数据出境合规及数据审计 6 方面内容，旨在为企业 AIGC 产品的研发、运营提供有效的合规指引。

（一）训练数据合规

《暂行办法》要求 AIGC 服务提供者使用具有合法来源的数据和基础模型，若其中涉及个人信息，应当取得个人同意（甚至单独同意）或者满足其他合法性基础。目前，实践中的训练数据来源基本可以分为以下几类：

- 数据爬取：例如，AIGC 服务提供者自行从其他网站爬取所需的训练数据；
- 通过 AIGC 产品直接获取数据：例如，用户在使用 AIGC 产品时，用户自行上传提供或由 AIGC 产品生成的数据；
- 从开源数据库获取数据：例如，AIGC 服务提供者直接从开源的数据库中下载、导出的训练数据；
- 从第三方采购数据：例如，直接从第三方数据提供商处采购合法的数据作为训练数据；再如，近期有数据交易所挂牌了语料库，以供需要训练数据的公司采购；
- 获取已经合法公开的公共数据：例如，将合法获取的政务数据作为训练数据。

针对不同的训练数据来源，我们分析了其对应的法律风险，并提供了合规措施建议，以期帮助企业做好风险防控，保障训练数据收集、使用的合法合规。具体风险及措施见下表：

来源类型	法律风险	合规措施建议
数据爬取	刑事风险：直接爬取他人网站上的相关数据（如个人信息、图片、研报或者他人享有竞争性利益的数据），可能构成侵犯个人信息罪、侵犯著作权罪、非法获取计算机信息系统数据罪、非法侵入计算机信息系统罪； 不正当竞争风险：若通过技术手段违法 Robots 协议、违反网站禁止爬取的相关规则而爬取他人数据作为训练数据，可能存在不正当竞争的风险； 侵犯商业秘密的风险：若爬取的数据涉及其他企业的商业秘密，则可能存在侵犯他人商业秘密的风险，从而引发争议； 著作权侵权风险：若爬取了他人网站上受著作权保护的作品，则存在未经他人授权使用他人作品的情况，从而构成著作权侵权。	建议 AIGC 服务提供者在有数据爬取需求前，优先评估爬取行为的合法合规性，包括但不限于仔细阅读并理解被爬取网站/平台的禁止性要求/Robots 协议，确认爬取数据的后续使用要求等事项。
通过 AIGC 产品直接获取数据	人格权侵权风险及侵犯个人信息合法权益的风险：用户向 AIGC 产品提供的数据大多为	AIGC 服务提供者应在其《用户协议》《隐私政策》等协议文本中明确写明将用户提

来源类型	法律风险	合规措施建议
	个人信息，有的甚至包括人脸、声音等敏感个人信息，若在使用这些数据之前未获得个人的同意（或者单独同意），则会侵犯个人肖像权、隐私权以及个人信息合法权益。	供的个人信息以及用户在使用 AIGC 产品过程中所产生的数据用于模型训练，并获得用户的同意。 对于涉及敏感个人信息的部分，还应该单独提醒用户，并获得用户的单独同意。
从开源数据库获取数据	著作权侵权风险或违约风险： 从开源数据库中获取数据是目前大多数 AIGC 服务提供者获取训练数据的做法，如果开源协议未授权使用者进行商业化使用，则可能构成著作权侵权或者违约，从而产生争议。	梳理开源协议的规则及要求，严格按照开源协议约定的内容进行训练数据的处理。建立开源数据库应用台账，及时掌握开源许可证变更、漏洞、闭源等变化情况，实行常态化管理，规避风险。
从第三方采购数据	通常情况下，第三方数据提供商是在合法的基础上进行数据交易，但仍存在其数据来源存在瑕疵的风险，从而影响 AIGC 服务提供者对训练数据的使用。	对第三方数据提供商开展合规尽调，调查其数据来源的合法性、数据保护能力等情况；在合作协议中明确双方权利义务及责任，要求其承诺并保障数据来源的合法性。
获取已经合法公开的公共数据	通过该来源获取的训练数据，相较于前述来源法律风险较小。通常情况下，已经合法公开的公共数据会明确其授权范围，主要的风险在于若超越授权范围处理公共数据，可能会遭受行政处罚或者暂停商业活动。	严格按照公共数据授权的范围对数据进行处理。

表 4-1 训练数据来源、法律风险及合规建议一览表

(二) 数据标注合规

数据标注是大部分人工智能算法和机器学习的关键环节，是产出投喂 AIGC 训练的优质材料的过程。数据标注业务的数据标注方一般为 AIGC 服务提供者，但也存在越来越多的提供者选择将数据标注业务外包。

数据标注产业面临的风险不容忽视，包括个人信息风险、数据安全与合规风险、知识产权侵权风险、不正当竞争风险、违约或共同侵权风险等。例如，对于数据安全及合规风险，具体来说，数据标注可能需要处理大量敏感数据，如金融数据、医疗数据等，若数据标注方尤其是接受数据标注业务外包方未能采取必要安全措施如数据加密、访问控制等，可能会导致数据泄露、篡改等安全问题。

因此，《暂行办法》第八条明确了 AIGC 服务提供者进行数据标注需履行的义务，包括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，提升遵法守法意识，监督指导标注人员规范开展标注工作。

其中，数据标注合规中最关键且难度最大的措施在于标注规则的制定。为此，我国国家标准化管理委员会于 2023 年 5 月发布了《人工智能面向机器学习的数据标注规程》(GB/T 42755-2023)，并于 2023 年 12 月 1 日正式施行，详细规定了数据标注的操作流程，明确了数据需求方、标注管理方和数据标注方三类主体，并对标注任务、标注人员、标注环境；过程

控制、质量保证、管理机制；内部质检、数据交付、后期维护等问题作出了规定，要求数据标注方通过数据脱敏处理和个人信息安全影响评估制度的构建，提供数据标注任务开展的保障，以确保数据标注的一致性和准确性，为 AIGC 产业提供高质量的训练数据。

该国家标准中关于数据标注的流程、体系和具体要求值得 AIGC 提供者提前内化成

企业内部的数据标注的流程和规则，并将其中明确的配套制度要求转化为企业的保障举措。同时，该国家标准提示 AIGC 服务提供者若存在数据标注业务的外包，需与执行数据标注任务的数据标注方提前确定好详细的标注任务、标注人员、标注环境、交付内容、交付验收标准及中后期维护的要求，通过双方详细的合同约定确保数据标注方安全、合法地完成数据标注工作。

（三）个人信息（含敏感个人信息）合规

1. 个人信息合规要点

AIGC 产品个人信息保护合规，首先需要符合《个人信息保护法》的知情同意和必要性原则。其中，关于知情同意原则的具体要求可以参考于 2023 年 12 月 1 日施行的《信息安全技术个人信息处理中告知和同意的实施指南》（GB/T 42574-2023），尤其需要关注单独同意的场景，如新增了人身或财产安全等产生重大直接影响的自动化决策场景需取得单独同意，并注意《指南》新增的提供者取得单独同意的前置条件，即 AIGC 服务提供者需在个人作出单独同意之前，通过增强告知的方式，针对需要单独同意的情形专门向个人进行充分告知，避免一揽子取得个人同意。

《暂行办法》第 11 条再次明确了“不得收集非必要个人信息”的必要性原则，并提出了“不得非法留存能够识别使用者身份的输入，不得非法向他人提供使用者的输入信息和使用记录”的存储和共享方面的限制性要求。同时，需要关注于对用户请求的及时响应，即提供者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，且提供者需及时处理 AIGC 出现的违法内容，并健全投诉举报机制。相较于先前的《征求意见稿》，《暂行办法》删除了“不得根据用户输入信息和使用情况进行画像”的要求，规定的个人信息保护义务更为具体、可执行，且与《个人信息保护法》的原则更为贴切。

2. 敏感个人信息合规要点

AIGC 服务提供者在提供服务过程中，可能涉及处理未成年人个人信息、生物识别特征信息、个人行踪轨迹信息、医疗健康信息等敏感个人信息。我国《个人信息保护法》《儿童个人信息网络保护规定》《暂行办法》《算法推荐规定》《人脸识别技术应用安全管理规定（试行）（征求意见稿）》等法律法规、国家标准都对数据处理者处理敏感个人信息、向未成年人提供服务提出合规要求。

在处理敏感个人信息方面，《个人信息保护法》就要求处理个人敏感信息的处理者应当征得个人的单独同意或书面同意，处理不满十四周岁的未成年人的个人信息，还应当征得未成年人的父母或者其他监护人的同意，并制定专门的个人信息处理规则。除此之外，在处理个人敏感信息前，应进行个人信息保护影响评估，并向个人敏感信息主体说明处理个人敏感信息的必要性及对个人权益产生的影响。《儿童个人信息网络保护规定》还要求处理不满十四周岁的儿童的个人信息时，网络运营者应当遵循正当必要、知情同意、目的明确、安全保障、依法利用的原则，设置专门的儿童个人信息保护规则和用户协议，并指定专人负责儿童个人信息保护。

在向未成年人提供算法推荐服务、生成式人工智能服务时，应注意采取有效措施防范未成年人用户过度依赖或者沉迷。算法推荐服务提供者还应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，禁止向未成年人推送可能影响未成年人身心健康的信息。

除此之外，结合目前人脸识别技术广泛应用及人脸信息作为高度敏感的个人信息，为防止人脸识别技术滥用并切实保护人脸信息安全，《人脸识别技术应用安全管理规定（试行）（征求意见稿）》对利用人脸识别技术处理人脸信息、提供人脸识别技术产品或服务规定了详细合规要求。首先，人脸识别技术使用者的备案义务。即在公共场所使用人脸识别技术，或者存储超过 1 万人人脸信息的人脸识别技术使用者，应当向所在地市级以上网信部门备案。其次，强化人脸信息处理者在收集人脸信息过程中取得单独同意或书面同意义务，确保收集人脸信息具有正当、合理且必要。另外，还对人脸信息处理

者的安全保护做出更高的要求，如要求处理人脸信息系统应满足三级等保要求，处理人脸信息前应进行个人信息保护影响评估，每年对图像采集设备、个人身份识别设备进行风险检测评估。此外，还对不同场景使用人脸识别技术处理人脸信息做出不同的要求。

举例：如智谱清言的《隐私政策》

“(二) 为您提供“问答对话”功能

您可以自愿在使用智谱清言的过程中，向我们提供您的对话信息以及其他各类个人信息并使用智谱清言生成问答对话，我们将向您展示您的历史对话。请您注意，您是否在对话中提供个人信息、提供何种个人信息完全由您个人决定。如您向我们提供您的人脸信息、指纹信息、虹膜信息等带有生物识别特征的敏感个人信息，尽管我们将采取充分的安全保障措施，该等信息一旦被泄露或不当利用仍然可能对您的人身、财产安全或人格权益造成侵害。我们不建议您提供可能对您的合法权益造成影响的个人信息，您不应未经授权提供他人的个人信息。同时，请您理解，为履行《网络信息内容生态治理规定》的要求，我们必须留存您提供给智谱清言的各类信息，包括此种情形下您提供的个人信息，留存期限等详见本协议第五条。”

再如，商汤旗下的 SenseChat 《隐私政策》

“3. 我们会对可识别的个人敏感信息进行加密传输和加密存储，加密强度符合安全要求，以确保数据的保密性。我们的应用系统提供了身份鉴别、用户标识唯一性检查、基于角色的访问控制等功能，采用 HTTPS 安全协议进行通信，设置了最大并发会话连接数，能够对系统服务水平降低到预先规定的最小值进行检测和报警。我们在服务端部署访问控制机制，对可能接触到您个人信息的工作人员采取最小够用授权原则，并定期核查访问人员名单和访问记录。我们的服务器操作系统和数据库系统口令有复杂度要求，采用 SSH 安全协议进行远程管理，严格限制默认账号的访问权限，并修改了默认口令，审计记录全面并覆盖了所有用户。”

（四）用户界面合规

一般来说，拟于各应用商店上架的 AIGC 产品面向 C 端用户提供。产品用户界面的合规是当前上架审核及上述算法备案环节能否通过的关键点。结合《个人信息保护法》《网络安全法》《暂行办法》等规定，以及现已发布的国家标准，AIGC 产品用户界面合规需同时关注实名认证的要求、用户协议及隐私政策的设计、AIGC 标识设计、投诉 / 举报机制的设置及针对未成年人、老年人等特殊群体的界面、功能设计等问题。其中，关于 AIGC 标识设计已在本白皮书“二（五）、《网络安全标准实践指南——生成式人工智能服务内容标识方法（征求意见稿）》解读”中详细展开，故在此不再赘述。

1. 实名认证的要求

《AIGC 暂行办法征求意见稿》中曾规定了提供者需要求使用者提供实名信息的要求，但在最终发布的《暂行办法》中，已经将该规定予以删除。此变化符合当前我国 AIGC 产业发展实际，一方面是由于当前很多 AIGC 服务是“对话式”而非“发布式”，且通过 API 提供，落实实名制难度大且没有必要性；另一方面是因为当前采取实名制将大幅度提高合规成本，不利于 AIGC 产业的快速发展。因此，目前没有针对 AIGC 服务提供者普遍性、强制性的实名认证要求。

但是，若 AIGC 服务提供者为用户提供的是信息发布、即时通讯等服务的，应当依据《互联网用户账号名称管理规定》和《互联网用户账号信息管理规定》，要求申请注册的用户进行实名认证。

2. 《用户协议》及《隐私政策 / 个人信息保护规则》的设计

在形式上，《用户协议》及《隐私政策》需为两份单独的文本，在用户首次使用 AIGC 产品时便弹出协议文本界面供用户查阅、勾选同意，且不能默认勾选同意。

在内容方面，两份协议文本的设计要求需遵守《个人信息保护法》《暂行办法》《App 违法违规收集使用个人信息行为认定方法》《常见类型移动互联网应用程序必要个

人信息范围规定》《信息安全技术个人信息安全规范》《信息安全技术个人信息处理中告知和同意的实施指南》等规定。其中，特别需要注意：

- 透明性及合法性要求：通过《隐私政策》等明确 AIGC 服务使用者的个人信息处理规则并具备合法性基础，一般为取得同意，或系为用户提供服务所必需；
- 数据权属和权利义务的约定：在《用户协议》中需约定数据权属和双方的权利义务；
- 保障用户权益的渠道设置：应在《用户协议》中明确保障用户权益的具体渠道，如为用户提供投诉举报的邮箱、电话等渠道及行权路径；
- 跨境传输情形的披露：若 AIGC 产品涉及将用户的个人信息传输至境外，提供者需于《隐私政策》中披露跨境情形。

3. 投诉 / 举报机制的设计

《暂行办法》第 15 条要求提供者应建立健全投诉、举报机制，并设置便捷的投诉、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。其中，投诉、举报入口包括但不限于在线平台、电子邮箱、电话等多种形式，处理流程和反馈时限等关于投诉、举报机制设置的具体内容应在上述《用户协议》中予以明确。

4. 个人信息权利的响应机制设计

我国《个人信息保护法》明确赋予了个人信息主体多样的个人信息权利（如查阅权、复制权、删除权、可携带权、撤回权等）。而在 AIGC 产品，尤其类似对话式的产品，涉及用户输出大量个人信息至对话框。如用户希望删除相关个人信息，则涉及其删除权的实现。如凡事均需要用户与 AIGC 服务提供者的客服取得联系，则面临较高的响应成本。为此，在 AIGC 服务提供者的具体产品界面，需要考虑设计更为合理的个人信息权利响应机制。

举例：《文心一言个人信息保护规则》

“3. 删除您的个人信息。如果您希望删除您账号的问答历史，您可以通过服务页面左侧查看对话历史记录，并执行删除操作。在以下情形中，您可以向我们提出删除个人信息的请求，您可以通过【yiyao-service@baidu.com】随时与我们联系：1) 如果我们

违反法律法规或与您的约定收集、使用、与他人共享或转让您的个人信息；2) 如果我们违反法律法规规定或与您的约定，公开披露您的个人信息，您有权要求我们立即停止公开披露的行为，并发布通知要求相关接收方删除相应的信息。当您从我们的服务中删除信息后，我们可能不会立即从备份系统中删除相应的信息，但会在备份更新时删除这些信息。”

5. 针对未成年人、老年人等特殊群体的界面、功能设计

《暂行办法》第 10 条规定，提供者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，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。这提醒我们在未成年人等特殊群体个人信息的收集和处理以外，也需要关注产品界面和功能的设计问题。

目前针对未成年人已有较为成熟的模式要求，国家网信办已于 2023 年 8 月 2 日发布了《移动互联网未成年人模式建设指南（征求意见稿）》，规定了 AIGC 产品在内的未成年人模式设置的基本要求、功能要求和管理要求。值得关注的是，该份规定提出了需分五个年龄段针对不同年龄阶段用户提供适合其身心发展的分龄内容服务的要求，为未成年人和家长提供社交管理权限，以及需设置未成年人内容池和限制未成年人用户使用可能危害其身心健康的产品和服务。若 AIGC 产品不属于即时通讯工具，未成年人模式下还需要关闭陌生人私信功能。同时，对于应用商店等应用程序分发平台，应加强未成年人专区建设，若 AIGC 产品不面向未成年人提供，则不得在未成年人模式下的应用程序分发平台上架。

而对于老年人、残疾人等群体，目前虽未有专门性规定，但未来在产品界面及功能设计上可能会赋予更高的关注度。

例如，字节云雀大模型的《隐私政策》：

“8. 未成年人条款

我们非常重视对未成年个人信息的保护。豆包主要面向成年人。

8.1 若你是未满 18 周岁的未成年人，在使用本产品及相关服务前，应在你的父母或其他监护人的监护、指导下共同阅读并同意本隐私政策。

8.2 若你是未满 14 周岁的未成年人的监护人，在使用豆包及相关服务前，应为你的被监护人阅读并同意本隐私政策。

8.3 我们根据国家相关法律法规的规定保护未成年人的个人信息，只会法律允许、父母或其他监护人明确同意或保护未成年人所必要的情况下收集、使用或披露未成年人的个人信息；如果我们发现在未事先获得可证实的监护人同意的情况下收集了未成年人的个人信息，则会设法尽快删除相关信息。

8.4 若你是未成年人的监护人，当你对你所监护的未成年人的个人信息有其他疑问时，请通过公司本隐私政策公示的联系方式与我们联系。”

（五）数据出境合规

随着《数据出境安全评估办法》《个人信息保护认证实施规则》《个人信息出境标准合同办法》的颁布实施，以及各地网信办发布的个人信息出境标准合同备案指引等指引文件，具体落实《网络安全法》《数据安全法》和《个人信息保护法》中关于数据出境的要求逐渐落地，国内数据出境的监管链条已经趋于完整。数据出境安全评估、个人信息标准合同备案、认证三种数据出境渠道构成了我国数据出境的主要内容。

但什么是“数据出境”，如何判断数据是否出境，应该选择哪种出境渠道，部分企业尚未厘清。根据相关规定，当前数据出境行为主要包括两种：1) 数据处理者在境内运营中收集和产生的数据传输、储存至境外；2) 数据处理者收集和产生的数据存储在

境内，境外的机构、组织或者个人可以查询、调取、下载、导出。即不仅限于将数据传输、发送、存储在境外才属于数据出境，数据存储在境内，境外组织机构、人员能够接触到该数据，就属于数据出境。注意此处并不要求境外组织机构实际接触该数据，只要其有接触到该数据的可能性，如享有访问下载渠道、权限，不论是否实际访问，也被认定为数据出境。通过 U 盘、硬盘等载体将数据携带出境、使用供应商通过境外系统提供的数据处理服务等都构成数据出境。

对于重要数据出境、关键信息基础设施运营者、掌握超 100 万人个人信息的个人信息处理者、自上年度 1 月 1 日起累计向境外提供个人信息超 10 万条或敏感个人信息 1 万条的个人信息处理者，向境外提供数据的，都应当履行数据出境安全评估义务，向所在地省级网信办申报数据出境安全评估。另外在申报数据出境安全评估前，数据处理者应当开展数据出境风险自评估，评估数据出境目的、方式、范围的合法、正当、必要，数据出境可能带来的风险、境外接收方承诺安全义务与数据安全保护能力、双方签订的法律文件、个人信息权益维护渠道等。

而个人信息出境标准合同备案方式适用于无需通过安全评估的个人信息出境行为，即非关键信息基础设施运营者、处理个人信息不满 100 万人、自上年 1 月 1 日起累计向境外提供个人信息不满 10 万人、自上年 1 月 1 日起累计向境外提供敏感个人信息不满 1 万人的个人信息处理者向境外提供个人信息。境内提供者依然需要在提供前与境外接收方签订个人信息出境标准合同并完成备案，或进行个人信息保护认证。需要注意的是，在签订个人信息出境标准合同并进行备案时，境内提供者应一并完成个人信息保护影响评估，相关评估报告需向网信部门提交。

对于 AIGC 服务提供者，对于该部分合规，需要关注下列三种情形：

情形一：虽然不存在我国数据出境的情形，但是需要注意其自身服务面向的市场，是否包括了不同国家，为此其可能会被其他国家实施长臂管辖。如 欧盟、美国的用户也可以使用其 AIGC 产品，且界面设计存在中英文切换等安排；

情形二：一种情形是境外 AIGC 服务提供者的服务市场面向境内，但是服务器和算法却在不同国家，为此面临数据在不同国家之间的流转；

情形三：明确境内的 AIGC 服务提供者，选择了海外大模型，但是并未在我国进行本地化部署，为此面临我国数据出境的监管规定适用空间。

针对上述三种情形，第一，需要关注是否存在数据跨境情形，如其用户输入的信息是否会传输到海外，境外的数据是否会传输到境内。第二，结合数据内容，判断是否会涉及强监管的要求，如本地化存储，如审批，如获得用户单独同意等要求；第三，需要做好数量统计，如个人信息出境的数量统计，确保相关出境合规路径如发生实质性变化（如达到了 1 万人以上敏感个人信息、10 万人个人信息），则需要提前申请数据出境安全评估。第四，还应做好 PIA（个人信息安全影响评估）机制搭建，境外数据出境规则梳理和动态监测，确保不违反境内外数据出境规则的规定。第五，提前考虑自身 AIGC 服务所面向的市场范围，避免在早期因为需要满足不同国家的合规要求，面临过高的合规成本。

举例：百川大模型的《隐私政策》

“5. 如何存储您的个人信息

我们会将在中国境内运营所收集和产生的个人信息存储在中国境内。在本政策下所收集及处理的个人信息不存在出境的情况。…”

再如：智谱清言的《隐私政策》

“五、我们如何存储您的信息

（一）存储地点

我们在中国境内运营中收集和产生的个人信息存储在中国境内。

目前我们不会跨境传输或存储您的个人信息，将来如需跨境传输或存储的，我们会向您告知信息出境的目的、接收方、安全保障措施和安全风险，并事先征得您的同意。”

（六）数据安全审计（含个人信息审计）

AIGC 产品在开发、对外提供服务的过程中，往往涉及到个人信息的处理，为此需履行开展个人信息保护合规审计的义务。《个人信息保护法》第五十四条明确，“个人信息处理者应当定期对其处理个人信息遵守法律、行政法规的情况进行合规审计”。根据国家网信办于 2023 年 8 月 3 日发布的《个人信息保护合规审计管理办法（征求意见稿）》（以下简称“《合规审计管理办法（征）》”）第三条指出，“个人信息保护合规审计，是指对个人信息处理者的个人信息处理活动是否遵守法律、行政法规的情况进行审查和评价的监督活动”。AIGC 产品在开发过程中，有可能会涉及到使用采买的人声、人脸等个人信息进行算法训练；在产品运营过程中，输入或输出的数据可能会包含个人信息；在产品迭代、升级时，还可能会利用收集的用户个人信息进行算法优化、调教。因此，开展个人信息保护合规审计是 AIGC 企业的法定义务。

企业应每 2 年开展一次合规审计，对于处理 100 万人以上个人信息的企业，应每年开展一次。《合规审计管理办法（征）》第四条对开展审计的频率进行了明确要求，“处理超过 100 万人个人信息的个人信息处理者，应当每年至少开展一次个人信息保护合规审计；其他个人信息处理者应当每二年至少开展一次个人信息保护合规审计”。对于此处的 100 万人的计算，我们理解不包括以受托方身份处理的个人信息。因此，对于提供 AIGC 产品技术支持的企业，若自身处理个人信息未超过 100 万人，应至少每 2 年开展一次合规审计。另外，由于《个人信息保护法》实施至今已 2 年多，企业可考虑在本年度内尽快完成一次合规审计，一方面，是履行法定义务的体现，另一方面，有利于通过审计发现 AIGC 产品的个人信息保护风险，并及时改正，避免在日后监管执法趋严时被处以行政处罚，影响产品运营和企业声誉。

对于审计的事项，可参考《合规审计管理办法（征）》附件列出的 114 个要点展开。如有条件，第一次开展合规审计可考虑由律师进行。《合规审计管理办法（征）》的附件《个人信息保护合规审计参考要点》是依据《个人信息保护法》等法律、行政法规和国家标准的强制性要求制定，为开展个人信息保护合规审计提供了相对官方、全面的参考。个人信息处理者在各业务场景和个人信息全生命周期各环节开展合规审计需重点关注的事项及具体场景见下表。此外，鉴于个人信息保护合规审计尚未有成熟的经验，以及个人信息保护及 AIGC 相关法律法规的解释亦有不同的口径，为此，如有条件，企业第一次开展合规审计不妨考虑由律师组成的合规审计团队展开，以全面、科学、准确地识别企业个人信息处理的合规风险。

附录 1：AIGC 平台知识产权相关条款附表

序号	AIGC 类型	产品名称	未成年人保护条款	知识产权归属条款	输入内容侵权条款
1	AI 对话	文心一言	如果您未满 18 周岁，请在法定监护人陪同下阅读本协议，并在征得法定监护人的同意后使用本服务。	百度对服务的开发和运营等过程中产生的所有数据、信息、输出等享有法律法规允许范围内的全部权利。	用户应确保输入的内容不侵犯任何人的知识产权等合法权益。由此产生的法律风险由用户承担。 平台提供了通知-删除规则的适用。
2	AI 对话	智谱清言	如果您是未满 18 周岁的未成年人，审阅时应由法定监护人陪同，并特别注意未成年人使用条款	您基于智谱清言生成的内容的版权由您自行维护并对其独立判断后使用。	用户保证不得以侵犯他人版权的方式使用智谱清言。 平台提供了通知-删除规则的适用。
3	AI 对话	通义千问	若您为未成年人或限制民事行为能力人，则您不具备前述主体资格，您及您的监护人应承担因您的不当注册行为而导致的一切后果。	通义千问应用 Logo、“通义千问”、“tongyi”等文字、图形及其组合，以及通义千问应用的其他标识、徽记、技术文档等的知识产权为我们所有。除非另有约定或法律法规另有规定的，用户享有本服	用户应确保其上传的任何资料、素材等知识产权均取得合法授权。

序号	AIGC 类型	产品名称	未成年人保护条款	知识产权归属条款	输入内容侵权条款
				务应用基于用户有 权上传的内容进而 合成的内容知识产 权，用户授权平台 在提供服务必要的 范围内使用用户上 传、生成、合成的 内容。	
4	AI 对 话	讯飞星火 认知大模 型	如果您未满 18 周 岁，或者不具备完 全民事行为能力， 请在监护人陪同下 阅读本协议，并在 征得监护人的同意 后使用本服务。	与本平台相关的一 切著作权、商标权、 专利权、商业秘密 等知识产权，以及 本服务输出的相关 所有信息内容均受 法律法规保护，由 科大讯飞及其关联 方享有完整的知识 产权、所有权等合 法权利。 使用本平台服务上 传、发布或传输内 容即代表了您不可 撤销地授予科大讯 飞及关联方非排他 的、无地域限制的、 永久的、免费的许 可使用及可再许可 第三方使用的权 利。	用户输入他人享有 知识产权等合法权 利的内容，应当取 得合法授权，确保 不得侵害他人合法 权利。

序号	AIGC 类型	产品名称	未成年人保护条款	知识产权归属条款	输入内容侵权条款
5	AI 对话	商量 SenseChat	如果您因年龄、智力等因素而不具有完全民事行为能力，请在法定监护人（以下简称“监护人”）的陪同下阅读并判断是否同意本协议，并特别注意未成年人使用条款。	本服务的一切著作权、商标权、专利权、商业秘密等知识产权，以及与本服务相关的所有信息内容，商汤享有全部的知识产权。 用户独立使用硬件记录或生成的内容，其合法权利均由用户本人享有。	未提及
6	AI 写作	秘塔写作猫	未提及	用户为其生成的文字负责，应合法、合理使用。	用户保证在使用服务时上传以及输入的内容不侵犯任何第三方的知识产权、名誉权、姓名权、隐私权等权利及合法权益。 平台有权在收到权利方或者相关方通知的情况下移除涉嫌侵权的内容。
7	AI 写作	火山写作	平台原则上不收集未成年人的个人信息。 若用户认为平台未经许可收集了未成年人信息，可以通过邮箱通知平台进行删除。	本服务包含的产品、技术、软件、程序、数据及其他所有信息的内容的所有知识产权归我们或我们的关联公司所有。	用户保证其输入的内容已经取得合法授权且不侵犯他人的知识产权或其他合法权益。 平台提供了通知-删除规则的适用。

序号	AIGC 类型	产品名称	未成年人保护条款	知识产权归属条款	输入内容侵权条款
8	AI 写作	讯飞智检	如您未满 18 周岁，请在监护人陪同下仔细阅读并充分理解本协议，并征得监护人的同意后使用本产品及相关服务。	本公司拥有本产品的知识产权等相关权利。 本公司授予用户对本产品可撤销的、有限的、非排他的、不可转让的使用权，用户仅限于以非商业使用本公司为目的进行使用。	用户保证在使用本产品时上传的文字、图片、链接等不侵犯任何第三方的知识产权、名誉权、姓名权、隐私权等权利及合法权益。 平台提供了通知-删除规则的适用。
9	AI 图像	文心一格	若您是 18 周岁以下的未成年人，建议您请您的监护人仔细阅读本协议，并在征得您的监护人同意的前提下使用我们的产品和/或服务或向我们提供信息。	文心一格产品/服务及其所有元素，用户向文心一格产品/服务提供的与该服务相关的任何信息及反馈由百度公司享有。 用户在使用本服务中所产生的内容的知识产权归文心一格或相关权利人所有。	内容提供者、上传者应对其提供、上传的内容承担全部责任。 百度公司有权在不事先通知内容提供者、上传者情况下直接采取删除、屏蔽、断开链接等必要措施。
10	AI 图像	通义万相	若您是未满 18 周岁的未成年人，您应获得监护人同意后在其监护下，认真阅读并同意本协议后，方可使用本服务。	通义万相服务及其所有元素，用户向通义万相服务提供的与该服务相关的任何信息及反馈由百度公司享有。 用户在使用本服务中所产生的内容的	用户保证适用本服务上传的信息内容已经获得合法授权。 平台提供了通知-删除规则的适用。

序号	AIGC 类型	产品名称	未成年人保护条款	知识产权归属条款	输入内容侵权条款
				知识产权归相关权利人所有。	
11	AI 图像	秒画	限制民事行为能力人应在监护人或法定代理人的陪同下阅读。用户保证在注册秒画账号时已满 14 周岁或已达到用户所在司法辖区内规定的类似的最低年龄。	用户在秒画上发表的全部原创内容，著作权均归用户本人所有。	用户保证上传的任何内容，用户为著作权人或已取得合法授权且不侵犯任何第三方的合法权益。 平台提供了通知-删除规则的适用。
12	AI 视频	度加创作工具	若您是 18 周岁以下的未成年人，建议您请您的监护人仔细阅读本协议，并在征得您的监护人同意的前提下使用我们的产品和/或服务或向我们提供信息。	度加创作工具产品/服务及其所有元素，用户向度加创作工具产品/服务提供的与该服务相关的任何信息及反馈由百度公司享有。 用户在使用本服务中所产生的内容的知识产权归度加创作工具或相关权利人所有。	用户利用度加创作工具产品/服务提供的网络服务上传、制作、传送或通过其他方式传播侵犯了第三方的合法权益由用户承担。
13	AI 视频	腾讯智影	如您未满 18 周岁，请您在法定监护人陪同下仔细阅读并充分理解本协议，并征得法定监护人的同意后使用本服	腾讯公司在本服务中提供的内容的知识产权属于公司所有。 用户通过本服务上传、发布所产生内	用户保证用户内容不侵犯第三方的任何权利。 平台提供了通知-删除规则的适用。

序号	AIGC 类型	产品名称	未成年人保护条款	知识产权归属条款	输入内容侵权条款
			务。	容的知识产权归属 用户或原始著作权 人所有。	
14	AI 视 频	快剪辑 SaaS 版	您必须年满 18 周岁或者如果您必须年满 16 周岁且您以自己的劳动收入为主要生活来源的，您才能使用或访问本服务，除非您虽然未满足上述年龄要求，但在您的父母、监护人或其他授权成人（如监护人）的直接监督下使用本服务，而且上述人员同意受本用户服务协议约束。	使用本服务发布包含用户内容的任何设计，即表示用户明确向快剪辑授予，并且您声明并保证您拥有所有必要的权利向快剪辑授予一个多用途、可再许可、可转让、永久、不可撤销、非独占性的全球许可。	用户保证用户内容不侵犯第三方的任何权利。 平台提供了通知-删除规则的适用。
15	AI 音 频	讯飞智作	如您未满 18 周岁，请在监护人陪同下仔细阅读并充分理解本协议，并征得监护人的同意后使用讯飞智作软件及相关服务。	本软件相关著作权及其他相关权利属于科大讯飞所有。科大讯飞即授予用户一项个人、非排他性、不可转让、有限的权利以使用本软件。	用户对其上传的内容承担责任。
16	AI 音 频	通义听悟	若您是未成年人或限制民事行为能力人，则您不具备前	阿里云网站 Logo、“阿里云”、“al iyun”等文字、	用户对其上传的内容承担责任。 平台提供了通知-

序号	AIGC 类型	产品名称	未成年人保护条款	知识产权归属条款	输入内容侵权条款
			述主体资格，您及您的监护人应承担因您的不当注册行为而导致的一切后果。	图形及其组合，以及阿里云网站的其他标识、徽记、阿里云服务的名称、技术文档等的知识产权为阿里云及其关联公司所有。	删除规则的适用。
17	AI 音频	网易云音乐·X Studio	如用户为未满 18 周岁的未成年人或其他不具有完全民事行为能力成年人，则应在其监护人监护、指导下阅读本协议和使用本软件。如用户的监护人不同意本协议或不同意用户使用本软件，则用户应停止使用并通知软件服务方，软件服务方将注销用户的账户并终止向用户提供服务。	本软件上所有内容，包括但不限于著作、图片、档案、资讯、资料、软件架构、软件界面的安排、设计、AI 歌手，均由软件服务方或其他权利人依法拥有其知识产权。用户使用本软件基于人工智能技术创作的任何成果所包含的知识产权以及根据任何国家或地区的法律可获得的任何其他权利，在全世界范围内均独占性归属网易云音乐所有。	用户保证上传的任何内容，用户为著作权人或已取得合法授权，因此产生的法律责任由用户承担。
18	AI 设计	AI Design	如果您未满 18 周岁，请在法定监护人的陪同下阅读本	本平台的一切著作权、专利权、商标权、商业秘密等知	用户保证上传的任何内容，用户为著作权人或已取得合

序号	AIGC 类型	产品名称	未成年人保护条款	知识产权归属条款	输入内容侵权条款
			协议及其他上述协议，在取得法定监护人的同意后使用本服务，并特别注意未成年人使用条款。	识产权，以及与本服务相关的所有信息内容，腾讯公司享有上述知识产权。用户并不因为使用本服务而对生成的 logo 及/或其衍生品享有著作权等知识产权，用户仅享有一定范围内的使用权。	法授权且不侵犯任何第三方的合法权益。平台提供了通知-删除规则的适用。
19	AI 设计	鹿班	未提及	在本协议项下一方向对方提供的任何资料、技术或技术支持、软件、服务等知识产权均属于提供方或其合法权利人所有。	用户保证上传的任何内容，用户为著作权人或已取得合法授权且不侵犯任何第三方的合法权益。平台提供了通知-删除规则的适用。
20	AI 翻译	火山翻译	未提及	用户通过本服务上传、提交您的用户上传内容，即视为您向公司（及公司的关联方、合同方和供应商）授予了许可，许可公司在全球范围内免费使用、复制、下载和展示您的用户内容。	用户保证上传的任何内容，用户为著作权人或已取得合法授权且不侵犯任何第三方的合法权益。平台提供了通知-删除规则的适用。

序号	AIGC 类型	产品名称	未成年人保护条款	知识产权归属条款	输入内容侵权条款
21	AI 训练模型	腾讯混元	若您是未满 18 周岁的未成年人，您应在监护人监护、指导并获得监护人同意的情况下，认真阅读并同意本协议后，方可使用本服务。	腾讯公司是本服务以及本服务所依赖的一切程序、软件、信息、数据、算法、模型等具有知识产权的内容的权利人。 本服务的生成内容的权利归用户所有。	用户保证上传的任何内容，用户为著作权人或已取得合法授权且不侵犯任何第三方的合法权益。

附录 2：植德人工智能与数字经济行业委员会及主要成员介绍

（一）人工智能与数字经济行业委员会简介

植德人工智能与数字经济行业委员会成立于 2022 年 1 月，原名植德元宇宙与数字经济行业委员会。委员会从产业链、商业模式、法律政策等多角度出发，专注于研究人工智能及数字经济（包括元宇宙）相关的监管要求与实务法律问题，特别是行业投资、企业经营中新出现的热点、难点问题，为客户提供多维度的法律服务，包括投融资、IPO、知识产权、合规（数据、刑事、税务、广告、互联网产品等）、常年法律顾问、争议解决等。

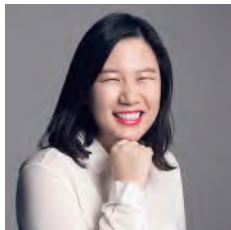
创建以来，委员会创办并定期发布《元宇宙与数字经济月刊》（现已经更名为《数字经济月刊》）和《AIGC 月刊》；紧跟人工智能与数字经济相关行业热点与司法实践，在植德公众号专栏发布专业文章近百篇，并先后发布《数据刑事红线指引 1.0 版》《员工个人信息保护指引 1.0 版》《数据安全法实务指引》《〈个人信息保护法〉逐条解读 1.0 版》《数字化转型白皮书》《AIGC 产品合规应用商店上架指引》等实务指引；委员会成员还深入政府部门、企业、高校、各地产业园区等开展专题培训，在人工智能、投融资、数字资产、互联网合规、数据合规、争议解决等方面，为包括行业企业、投资人等提供专业法律参考与支持。委员会还积极参与主管机关、行业协会组织的相关行业标准、调研、重大疑难问题的研究工作，得到各级监管机关、行业协会的认可。

委员会成员为来自不同专业领域、具有行业与法律多重背景且深耕多年的合伙人，擅长结合商业化思维，协助人工智能及数字经济相关行业企业搭建覆盖事前、事中和事后的全流程合规制度体系，实现及时有效的风险防范与危机处理，为客户的经营活动保驾护航。委员会成员在本行业经验丰富，具备对行业动态、发展趋势的洞察力，部分合

伙人有司法、税务等相关部门工作经历，能够为企业提供更贴合商业化的合规方案。委员会成员均毕业于国内外知名法学院校，如哥伦比亚大学、纽约大学、伦敦大学学院、杜克大学、北京大学、清华大学、中国政法大学、对外经济贸易大学等，持有中国及海外律师执业资格，并得益于团队的国际视野和良好的合作关系，可调集东南亚、美国、加拿大、日本、新加坡等热点司法辖区的合作律师，为客户提供无缝衔接的国际化法律咨询服务。

(二) 人工智能与数字经济行业委员会主要合伙人简介

	王妍妍 合伙人/北京 人工智能与数字经济行业委员会牵头人 邮箱: yanyan.wang@meritsandtree.com	
王律师先后毕业于中国政法大学、伦敦大学学院、哥伦比亚大学，取得了双法学硕士（LLM）学位，目前为中华人民共和国执业律师及美国纽约州执业律师。王律师曾在北京市经纬律师事务所以及英国礼德律师事务所、美国杜威律师事务所等国际一流律师事务所工作数年，在投融资并购与跨境交易、银行与金融产品以及涉外争议解决等业务领域具有丰富经验。 王律师在主要执业领域中拥有丰富执业经验，曾代理过包括人工智能、芯片、建筑、制造、新材料应用、银行、软件设计、文化娱乐、传媒、游戏、酒店、医疗设备、食品和体育等诸多行业的客户，对若干不同行业有深入了解，能根据行业特点为客户提供有针对性的优质法律服务，包括为这些客户提供融资，收购，公司治理、股权激励，架构重组等方面的法律服务。王律师基于其海外留学背景和多年知名国际律师事务所从业经验，尤其擅长跨境交易和与此相关的域外合规、市场准入、产业落地等相关业务。		



赵芸芸

合伙人/北京
人工智能与数字经济行业委员会成员
邮箱: yunyun.zhao@meritsandtree.com



赵芸芸律师毕业于中国政法大学，取得法学硕士学位。赵律师曾先后在两家律师事务所执业，并自 2010 年起任北京华录百纳影视股份有限公司证券部负责人、法务部总监，同时兼任海外事业部、丹丹百纳经纪公司运营协调部负责人，后加入北京植德律师事务所。

赵芸芸律师早期从事 IPO/MA 等资本市场领域的非诉业务及股权纠纷类仲裁业务，后致力于文化娱乐、影视传媒、互联网商业合规及 TMT 领域。赵律师在国内及中外合作电影(含动画电影)、电视剧、综艺栏目、体育赛事、杂志媒体合作、网络游戏运营等项目类方面，为客户全程提供商业合作模式架构设计、谈判策略与要点、合规风险评估与解决等法律服务，并在融资安排、联合投资合作、发行、商务模式等有独特心得与经验，善于根据不同项目有针对性地提示风险点和设计协议。在资本市场领域，赵律师曾多次作为被投方法律顾问，与投资方就公司估值、著作权尽职调查、业务增长模式、有限合伙等持股平台设计、员工激励与不竞争、SPA 协议、SHA 协议等进行沟通、谈判并形成一揽子协议，有效维护客户最大估值利益及股权结构安全性。

赵芸芸律师曾上榜 2020 和 2021 年度 LEGALBAND 中国顶级律师排行榜“体育娱乐”领域、“2023 年度 LEGALBAND 客户首选：娱乐律师 15 强”。



时萧楠

合伙人/北京
人工智能与数字经济行业委员会成员
邮箱: xiaonan.shi@meritsandtree.com



时萧楠律师毕业于日本一桥大学，经营法（知识产权项目）硕士研究生。时律师从事知识产权十余年，先后在北京天达共和律师事务所和日本西村朝日律师事务所、中国大型互联网公司工作多年，专注于解决合规、知识产权案件，包括互联网合规、数据合规、著作权授权、侵权诉讼、行政投诉等类型的案件，同时擅长解决疑难复杂案件。

时萧楠律师曾在大型知名互联网公司工作多年，对公司法务合规有着深刻的理解，并且擅长以业务目标为核心提供解决方案。时律师有公司法务与律所双重经验，能以行业视角和律师视角多元提供知识产权纠纷、合规解决方案。

	杨诚	
合伙人/上海 人工智能与数字经济行业委员会成员 邮箱: cheng.yang@meritsandtree.com		
杨诚律师毕业于上海外国语大学，取得法学学士学位。在加入植德之前，杨律师曾长期任职于德国某制造业企业集团中国区，负责该集团在华法律、合规及公共关系事务，并曾在观韬中茂律师事务所执业多年。 杨诚律师专注于为客户提供知识产权及反不正当领域的法律服务，对各类创新技术产品的研发上市、不同作品的制作与发行、著作权授权与保护、艺人经纪以及知识创意产业的发展有着长期观察和独到见解。在争议解决方面，杨律师擅长于为客户制定行之有效的诉讼策略。在知识产权相关诉讼，特别是互联网著作权侵权、商业秘密侵权、仿冒混淆、流量劫持、商业诋毁等新类型争议案件诉讼领域积累了丰富的经验。 杨诚律师现担任上海市律师协会文化传媒业务研究委员会委员、上海外国语大学法学院法律硕士研究生兼职导师。		
	何京	
合伙人/北京 人工智能与数字经济行业委员会成员 邮箱: jing.he@meritsandtree.com		
何京律师毕业于北京大学、合肥工业大学，分别获得法律硕士和理学学士学位。何律师曾在北京两家中级法院工作，曾任国家一级法官，具有 8 年审判经验，在知识产权及民商事争议解决领域具有丰富的经验。 何京律师是国际权威法律评级机构《法律 500 强》（The Legal 500）2023-2024 年度知识产权领域推荐律师。 何京律师办理过国内外知名企业的专利权、商标权、著作权、不正当竞争等知识产权及竞争法领域的重大案件，服务过人工智能、互联网、医药、文化与娱乐传媒、消费品与零售、教育与培训、先进制造、新兴行业等诸多行业的客户。 何京律师擅长从法官思维和商业思维的多元视角制定争议解决方案，为客户争取最优商业效果和法律效果。何律师是拥有律师执业证及专利代理师执业证的双证律师。		



王海青

合伙人/北京

人工智能与数字经济行业委员会成员

邮箱: haiqing.wang@meritsandtree.com



王海青律师毕业于华东政法大学，取得经济学学士学位。王律师是北京植德律师事务所合伙人，主要执业领域为争议解决、投资基金、银行与金融。

王海青律师执业超过十五年。王海青律师代理了大量投融资和公司股权类争议案件。同时，王海青律师为华芯原创（青岛）投资管理有限公司、北京红杉信远股权投资中心（有限合伙）、启迪金控投资有限公司、江西省井冈山北源创业投资管理有限公司、托特（三亚）私募基金管理有限公司等多家私募基金和私募基金管理人提供常年法律顾问、合规和投后风险处置等多方面的法律服务。王海青律师处理了大量涉及盖章签名效力、公司证照印章返还以及与公司控制权有关的各类争议，并在丰富的实务经验基础上，出版专著《公章法论》（法律出版社，2020年）。



李倩

合伙人/武汉

人工智能与数字经济行业委员会成员

邮箱: tracy.li@meritsandtree.com



李倩律师毕业于中南财经政法大学，取得法学学士学位。在加入植德前，李律师曾在中伦律师事务所工作多年。

在投资基金领域，李倩律师先后服务国开（湖北）投资、东湖零度、楚昌基金、宏泰天睿 30 余家私募机构，代表诸多家国有、上市公司、券商背景私募机构开展对外投资、投后管理、基金诉讼等活动，同时也代表多家被投企业接受基金投资、处理投后纠纷，多次受湖北省内金融监管部门、基金行业协会、基金孵化基地邀请担任大赛评委、专家顾问；在银行与金融领域，其先后为政策性银行多个基础设施、产业园区建设项目融资提供专项法律服务；在家事服务与财富管理领域，其善于运用金融、保险、资本市场知识为高净值人群家事纠纷提供综合性解决方案。

	黄茜 合伙人/北京 人工智能与数字经济行业委员会成员 邮箱: qian.huang@meritsandtree.com	
黄茜律师是北京植德律师事务所合伙人。加入植德前，黄茜律师先后在北京市北斗鼎铭律师事务所担任管理合伙人、北京安杰世泽律师事务所担任合伙人，曾在公安机关刑侦部门工作多年。任北京市律师协会刑事实务研究会委员、北京市东城区律师协会刑事专业委员会副主任，先后任北京市朝阳区律师协会律师代表大会第二届会员事务委员会秘书长、第三届刑事专业委员会副主任、第四届律师代表大会代表/理事等，荣获北京市朝阳区律师协会颁发的“2015-2018 年度北京市朝阳区行业贡献奖”。 黄茜律师专注于刑事辩护、重大疑难刑民交叉争议解决、企业合规和破产清算/重整领域。在争议解决领域有 20 余年的丰富执业经验，并在刑事辩护领域有多起无罪、轻缓量刑、不起诉辩护成功的案例；作为《中小企业合规管理体系有效性评价》团体标准起草人、工信部高级合规师考核培训专家，黄茜律师具有“企业合规师（高级）职业技能”“首席合规官岗位能力”等资格，在央企/国企合规体系建设方面有独到的见解和实践经验；黄茜律师自 2013 年起担任多起破产清算案件破产管理人的负责人，具有以破产管理人视角解决债权债务申报、财产处置难点的丰富实践经验。		
	陈文昊 合伙人/上海 人工智能与数字经济行业委员会成员 邮箱: wenhao.chen@meritsandtree.com	
陈文昊律师是北京植德律师事务所合伙人。在加入植德之前，陈文昊律师在锦天城律师事务所担任合伙人，并在金杜律师事务所、美国凯易国际律师事务所、美国奥睿律师事务所任职。 陈文昊律师主要执业领域为数据合规、反商业贿赂、公司内部调查、政府监管及争议解决。陈律师经常就公司日常合规事务出具咨询意见，行业涵盖大数据、互联网、高科技、广告与市场营销、零售、医疗器械、汽车零部件制造等领域。 陈文昊律师在数据合规领域具有丰富经验，深谙相关行业客户需求与产品特点，并擅长提供切实可行、符合监管要求的全套数据合规解决方案。陈律师善于根据相关行业的特点，与企业内部数据管理人员、技术人员进行有效沟通，在深刻理解客户需求的基础上，提供数据全生命周期的合规意见。陈律师还经常在 IPO、投资并购等复杂交易中为客户提供数据合规尽职调查与合规整改服务。		



李凯伦

合伙人/北京

人工智能与数字经济行业委员会成员

邮箱: kailun.li@meritsandtree.com



李凯伦律师是北京植德律师事务所合伙人。李凯伦律师为各种类型信托、金融科技、资产配置与财富管理、私募投资基金、资产证券化项目等资管业务提供法律服务，在交易结构设计、合规性审查、法律文本起草、法律意见出具、风险处置和化解等方面具有丰富的实践经验。

李凯伦律师为多家信托公司、证券公司及其子公司、私募基金管理人、银行理财子公司提供常年或专项法律服务，服务领域涵盖金融机构合规治理、金融科技应用、项目投融资与纾困、供应链金融、消费金融、财富管理与配置等。李凯伦律师擅于为客户处理具有复杂交易背景及多项合规监管要求的交易项目，并能够在前沿金融创新业务中为客户提供具有可操作性的建议。



— 请扫码获得更多资讯 —