汉坤法律评述
2025 年 9 月 13 日
HANKUN
汉 坤 律 师 事 务 所
Han Kun Law Offices
北京 | 上海 | 深圳 | 香港 | 海口 | 武汉 | 新加坡 | 纽约 | 硅谷

# 企业使用者视角下：生成式 AI 从“能用”到“可控”的合规落地路径

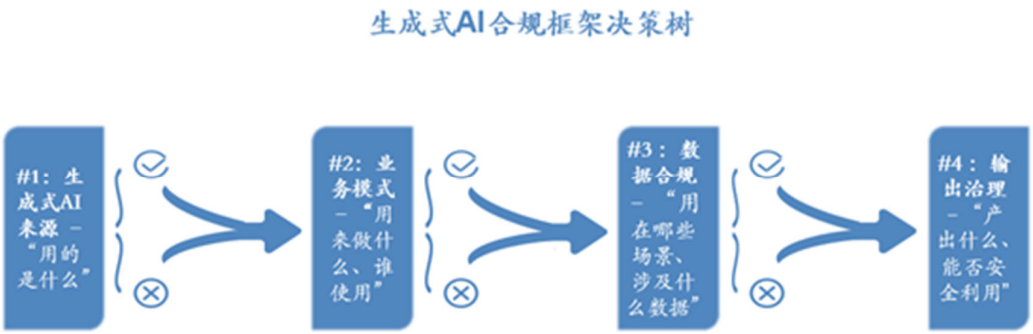
**作者：李珺**[1]

近年来，生成式人工智能（AI）技术已经融入到企业的各个业务场景，新技术的广泛应用也带来了新的法律挑战。我国已陆续出台了《互联网信息服务算法推荐管理规定》、《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》等专项规定，初步构建了人工智能领域的监管框架。

然而，面对众多存在交叉与重叠的合规要求，**许多企业作为生成式 AI 技术的使用者 — 尤其是在尝试新技术或新业务模式时 — 往往感到困惑，难以准确把握合规边界与责任范围**。这在一定程度上导致企业对生成式 AI 的应用持观望态度，从而拖慢了创新与市场响应。

## 一、结构化的合规思路

面对这些挑战，企业不妨可以引入一种**“决策树”式的合规分析框架**，化繁为简，**精准聚焦于关键环节**，避免“撒网式”的治理思路。其核心推进逻辑如下：




**步骤一：来源属性**

**判断生成式 AI 的来源，是否来源于中国境外，是后续合规工作的起点** — 比如会不会涉及数据跨境、

[1] 实习生朱永晗对本文的写作亦有贡献。

有没有受制于特殊的使用场景等。目前，国内法规对生成式 AI 的来源还没有非常明确的认定标准，尤其是从“地理”角度做区分时，实践中很容易出现混淆。由于技术本身复杂、商业安排往往不透明，很多时候并不能简单用“非黑即白”的方式来划分 AI 的来源。

为夯实来源认定，避免在后续合规中陷入被动，**建议企业构建一套结合商业安排、技术架构以及交付路径的来源评估机制，形成完整闭环的判断逻辑**；同时，需关注技术架构迁移或引入第三方组件等动态变化，持续开展评估。其中，在认定生成式 AI 来源属于中国境内时，前提之一是确保该上游技术提供方具备履行中国法规定的相关责任的能力与义务。

**步骤二：业务模式**

明确了技术来源，下一步就需要审视我们如何使用它。**对于源自境外的生成式 AI，其在境内的合规落地需结合具体应用场景，并以是否具有“公众属性”作为判断主要合规义务的起点。**总体而言，将生成式 AI 用于内部运营与对外提供产品或服务时，所适用的合规要求存在显著差异：相比仅限于内部使用的封闭场景，将生成式 AI 嵌入面向不特定公众（尤其是 C 端用户）的服务，通常需满足算法备案以及安全评估等要求。对于源自境外且技术存在“黑盒”情形的生成式 AI，目前尚缺乏成熟的备案与评估合规路径，建议企业依据具体业务场景与技术实现方式，开展进一步的合规可行性评估。

针对**应用生成式 AI 向 B 端客户提供服务的情形，需基于业务模式，综合考察服务的受众范围、提供方式、以及潜在传播链路等要素，以准确界定其是否具备公众属性**，并据此识别对应的合规责任。

从更广泛的监管框架看，除 AI 的专门性规定外，还需要关注行业监管要求和准入规定。例如，如果企业属外资主体，则需关注外资准入限制，例如增值电信业务经营许可证（如 ICP 证）等资质要求。因此，生成式 AI 的应用合规不仅是人工智能治理体系中的关键环节，也是企业整体合规战略中的重要决策节点。

**步骤三：数据合规**

无论在内部运营还是对外服务中使用生成式 AI，都可能涉及个人信息处理。**不同应用场景下的风险存在差异。**例如，在内容生成类应用中，生成内容可能包含来自训练数据或知识库的个人信息；而在多模态处理场景中，生物特征信息（如声音、面部识别数据）则存在滥用的隐患。

因此，可以结合自身 AI 应用的具体场景，识别其中的高风险环节，并在遵循告知、合法性基础、最小必要、加密传输等基本合规要求的基础上，**结合不同场景情境的特殊风险，在输入、上下文管理、存储、决策、输出及监测等关键环节，制定差异化的管理策略**，将合规控制要求嵌入实际业务流程中。

对于重要数据，由于其在技术应用的多环节中被频繁处理与调用，数据泄露的潜在风险也相应增高。对此，**一个可行的办法是建立与业务模式相匹配的数据分级分类与授权使用制度，框定重要数据的使用边界，实行场景化授权管理**。同时，应系统开展数据映射工作，协同技术、采购等相关团队，从制度规范、技术隔离、权限管理及供应商管理等多个维度全面落实合规要求。

**步骤四：输出治理**

在应用生成式 AI 的过程中，企业可能面临由生成内容引发的多重风险，包括因事实错误导致的误导性宣传、涉及著作权或商标权的知识产权争议、不正当竞争问题，以及商业秘密与敏感信息泄露等法律与合规隐患。为系统管控此类风险，**我们需要为生成内容建立一套覆盖全生命周期的输出管理机制**，将 AI 输出错误所带来的法律和商业影响控制在可接受范围内。企业可以预先设定规则，明确内容生成的“禁区”与“高压线”，并通过实时监控与干预，对不同风险等级的内容采取差异化的治理策略，将合规要求嵌入业务全流

程。

在生成内容的下游使用方面，自 2025 年 9 月 1 日起，《人工智能生成合成内容标识办法》开始施行。尽管该办法未强制要求对仅限于内部使用的生成内容进行标识，但为更好地防范因后续泄露或误用可能带来的风险，企业可考虑通过自研或采购第三方工具，**逐步建立生成内容的全程留痕管理机制，并依据内容敏感性构建分级治理体系**。建议对语音克隆、数字人等高敏感类内容予以重点关注并实施标注，配套建立与风险水平相适应的溯源管理措施。对于拟对外发布的生成内容，均建议在发布前履行标识核查程序；若发现未标识情形，完成补充标识后再行传播。

## 二、合规支撑体系

最终，建议企业构建一套**流程驱动、可持续演进**的生成式 AI 合规制度体系，以确保创新应用与风险管控的动态平衡。该体系应涵盖以下核心治理维度：**来源属性与供应链管理、业务模式的可行性评估与界定、数据与个人信息合规义务的映射与控制，以及生成内容的全面风险识别与闭环管理**，将合规要求融入企业的日常管理与运营流程。

## *特别声明*

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

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

**李珺**

电话： +86 21 6080 0981
Email： jun.li@hankunlaw.com