人工智能企业出海合规风险分析

作者：杨迅 | 郑学易

中国人工智能企业在 2025 年实现现象级突破后，加速拓展海外市场，但面临严峻的法律合规挑战。然而，中国人工智能企业的出海之路面临多重挑战，其中本地合规体系适配是核心重点。深入理解并严格遵守不同国家及地区在数据隐私保护、知识产权管理、人工智能合规监管等领域的规制要求，同时强化本地化文化适配能力，已成为人工智能企业出海的必备前提。

一. 数据合规与隐私保护

人工智能技术的发展高度依赖对海量数据的采集、存储与处理，数据合规和隐私保护因此成为人工智能企业合规管理的重要组成环节。这一要求既体现在模型训练阶段对训练数据集的合法收集与处理，也贯穿于模型落地应用阶段对用户数据的合规性处理。在企业出海场景中，数据隐私保护风险尤为突出，不同国家及地区基于自身数据主权、公民权益保护需求，构建了差异化的法律体系与监管标准。

例如，欧盟《通用数据保护条例》(GDPR)、美国《加州消费者隐私法》(CCPA)及《加州隐私权法案》(CPRA)、巴西《通用数据保护法》(LGPD)等代表性法规，不仅强制要求企业在收集、处理用户数据时履行透明度义务，确保相应个人信息处理活动具备合法性基础，更对跨境数据流动设置了严格的监管壁垒。若企业忽视上述规则，可能面临高额罚款、品牌声誉严重受损、甚至被直接限制市场准入的风险。比如，我国初创人工智能企业 DeepSeek 于 2025 年初因涉嫌违反欧盟《通用数据保护条例》，先后在意大利、法国、荷兰、德国等欧洲国家被监管机构立案调查，并最终遭到多国监管机构的封禁措施。社交媒体 TikTok 更是由于数据跨境传输的违规遭受高额罚款。无论 DeepSeek 还是 TikTok，其都在数据保护和隐私合规方面做出了努力，但是仍然由于对欧盟法律理解和执行的偏差，遭遇法律风险。

那么，中国出海人工智能企业应该如何减小该等风险呢？

如您需要了解我们的出版物，
请联系：

Publication@llinkslaw.com

1. 知情同意机制构建

"公开透明"是全球数据保护立法的共识性原则，企业需确保用户能够清晰了解其个人信息被处理的类型、目的、方式等内容，这既是法律规定的强制性义务，也是建立用户信任、保障业务可持续性的关键基础。实践中，企业主要通过制定隐私政策、个人信息处理规则等告知文本履行透明度义务，对于出海企业而言，尤其需要注意不同国家及地区对隐私政策的内容深度、表述规范存在的差异化要求。例如，韩国《个人信息保护法》要求明确公布受托处理方的具体信息，并明确要求区分"委托处理"与"提供给第三方",禁止使用"共享(share)"等模糊的词汇，其合规标准显著高于中国《个人信息保护法》。

基于此，出海企业在制定隐私政策时，可采用"基准模板---属地附录"的方法论，以中国《个人信息保护法》或欧盟《通用数据保护条例》为基础框架，提炼各国数据收集、存储、使用等共性合规要求并纳入基准模板；再针对目标国的特殊规制，通过增设属地附录的方式补充说明。此外，企业亦应重视不同法域对于"同意"的差异化规定，通过设置具有区别化的同意机制(opt-in 或 opt-out)，以满足各国对同意的法定要求。

2. 数据跨境传输体系设计

对出海企业而言，数据跨境流动既是商业模式全球化的必然选择，但同时也是合规风险高发环节。尤其是对于互联网产品(比如在线 AI 产品)，虽然绝大部分数据可以本地储存，但是大部分产品由中国技术团队远程运维，在这过程中，客户个人信息和其他数据可能被中国运维团队所接触，从而构成数据和个人信息的转移。在这转移过程中，一旦出现合规瑕疵，不仅可能面临监管处罚，更可能影响全球业务布局。

以出海欧盟为例，通常，中国出海企业以 SCC(欧盟下的标准合同条款)作为向中国境内提供数据的合法基础。但是，仅有 SCC 是不够的。对于向中国这样未被欧盟充分认可的法域，中国出海企业还需开展影响评估，主要评价 SCC 在数据接收国(如中国)法律下的可执行性，以及如果数据所在国法律不能提供有效保护对应的减小风险的措施。2025 年 5 月 TikTok 被爱尔兰数据保护委员会处以 5.3 亿欧元罚款案揭示了标准合同条款(SCC)的局限性：尽管 TikTok 采用 SCC 作为传输基础，但监管方认为其未能验证中国法律环境不会削弱 SCC 效力，且其隐私政策未明确列出中国作为数据传输目的地及中国境内人员可远程访问存储于美、新服务器的数据。 这一裁决表明，SCC 并非形式上的"护身符"，企业必须开展传输影响评估(TIA)，深入分析第三国法律环境下公权力获取数据的可能性及其对 SCC 有效性的削弱作用。

当前，各国围绕数据本地化的监管模式呈现多元化特征，主要包括三类：一是无强制本地化存储要求，但对跨境传输设置相应的合规前置条件；二是要求境内数据需留存副本，境外处理需符合特定条件；三是严格限制数据出境，仅允许在境内完成存储与处理。在全球数据监管趋严的背景下，出海企业必须把数据合规纳入战略，建立"合规评估---方案设计---流程管控"的数据跨境传输管理体系。首先，通过法律调研，识别目标国在数据跨境传输领域的法律要求；其次，综

合考量业务功能协同需求、法律和合规成本、IT 架构部署的成本和效率、本地化团队与合作方权责分工等因素，设计适配的跨境数据传输方案，例如，搭建本地化数据中心、采用合规跨境传输工具等；最后，针对数据收集、存储、传输等关键环节，制定标准化操作规程，并引入必要的审批和监控机制，确保数据跨境流动过程的合规性。

二. 知识产权合规

以生成式人工智能为代表的人工智能技术革新，对传统创作模式、知识产出路径带来了剧烈变革。人工智能正重塑知识产权的创造、运用、保护与管理体系。对于出海人工智能企业而言，知识产权领域面临的核心风险集中于著作权侵权风险与开源合规风险两大维度。

1. 著作权侵权风险防控

人工智能模型的训练需依赖大规模高质量数据集，当前行业内普遍采用网络爬虫技术等方式从公开网页抓取新闻、学术论文、博客内容、图片素材等信息用于模型训练。尽管上述信息具备公开可访问性，但"公开可访问"不等于"可自由使用"，若企业未经著作权人授权且无合法免责事由，直接将他人作品用于模型训练，可能构成对著作权人复制权、信息网络传播权的侵犯；同时，若人工智能生成内容与原作品在表达形式上存在实质性相似，还可能进一步引发著作权侵权有关争议。此争议表面是现有法律框架滞后于技术发展，深层实为传统资源方与人工智能开发者之间的利益再分配问题：前者主张数据训练的授权许可，后者则寻求低成本的"合理使用"空间。

从全球立法与司法实践来看，目前仅新加坡、日本等少数国家对人工智能数据训练的著作权豁免做出明确规定。但即便在上述国家，豁免范围也存在严格限制：例如，日本文化厅于 2025 年 10 月发布的《关于 AI 与版权的紧急事务官室见解》明确，数据挖掘阶段虽可依据《著作权法》第 30 条之 4"非享受性使用"条款主张豁免，但后续生成的内容若与原作品表达"高度近似"，著作权人仍有权主张停止侵害、损害赔偿，并可向法院申请临时禁令。此外，为应对 OpenAI 旗下 Sora2 人工智能应用对日本动漫产业的冲击，日本内容海外发行协会于 2025 年 10 月正式向 OpenAI 发出书面请求，要求其停止使用协会成员的版权内容用于 Sora2 模型训练，这进一步凸显了人工智能领域著作权争议的复杂性。

美国法院通过"合理使用四要素"分析框架为此类纠纷提供了重要裁判逻辑，其核心在于判断使用行为的"转换性"及是否构成市场替代。 在 Thomson Reuters 诉 Ross Intelligence 案中，法院认定被告复制法律数据库的编辑内容用于训练竞争性 AI 检索工具，因缺乏转换性目的且对原告市场构成直接威胁，不构成合理使用；而在 Bartz 等诉 Anthropic 案及 Kadrey 诉 Meta 案中，法院则认为大语言模型训练构成高度转换性使用，且输出内容未实质性替代原作品，故认定构成合理使用。值得注意的是，美国司法实践普遍接受"完整性复制是 LLM 训练的技术必要"这一技术事实，在第三要素判断中给予宽容，但最终裁决仍取决于第四要素"对原作品市场或价值的影响"的细致评估。

针对上述风险，出海人工智能企业应当密切跟踪各国及地区在人工智能著作权领域的立法更新与司法判例，及时调整合规策略，特别在内容生成过程中设置相应的过滤机制，排除可能侵犯版权的材料，监控输出内容，确保不包含受版权保护的作品的片段。**DeepSeek 近期因涉嫌使用 OpenAI 模型输出通过"模型蒸馏"技术训练自身系统而面临知识产权争议，正体现了后发企业在训练数据来源合法性方面的典型风险。**

2. 开源合规管理

开源(OpenSource)是指源代码及相关技术文档向公众开放，用户可自由查看、修改、再发布的技术协作模式。传统开源软件的核心是源代码和相关文档共享，而人工智能领域的开源范畴进一步扩展，涵盖模型架构、预训练权重、训练数据集、算法实现方案及使用说明等内容。相较于美国人工智能企业，中国人工智能企业作为后发者，普遍通过开源模式打破闭源生态壁垒、快速积累用户与开发者资源，但开源大模型在开发应用过程中亦面临开源标准、开源许可证内容、下游接入者相关责任等方面的挑战。

为此，在大模型开发阶段，企业需将开源合规纳入全流程管理，积极开展开源组件合规审计，通过专业工具对项目中集成的开源库、框架及其许可证进行系统性排查，明确不同许可证的义务要求(如是否需开源衍生代码、是否允许商业使用)，避免因许可证冲突导致合规风险。同时，在开源大模型发布阶段，企业亦需从负责任人工智能角度出发，构建标准化合规流程，制定并对外提供详细的使用指南、合规文档，以及安全实践指南，详细说明模型的局限性和潜在风险；并通过制定和实施限制性协议，明确模型的授权范围、使用限制、责任分担等内容。

三. 人工智能监管合规

全球主要国家及地区均在加速推进人工智能治理框架构建，形成了差异化的立法路径与监管模式。总体来看，人工智能监管已从原则倡导进入规则落地阶段，各国立法围绕透明度、问责性、风险管理、人类监督等核心原则形成普遍共识，在监管模式、实施路径方面则存在差异。针对上述监管环境，人工智能出海企业需重点关注以下两方面合规要求:

1. 限制性或禁止性用途管控

多个国家及地区的监管机构基于国家安全、公共利益、人权保障等考量，对特定人工智能系统或应用场景实施限制性或禁止性措施。以欧盟《人工智能法案》(AI Act)为例，其将人工智能系统划分为禁止的人工智能系统、高风险人工智能系统、有限风险人工智能系统和最小风险人工智能系统四个级别，实施差异化监管: 其中禁止的人工智能系统因对基本权利及欧盟价值观构成不可接受风险，被明确禁止在欧盟境内部署与使用，具体包括: 用于操纵人类行为的 AI 系统(如通过心理暗示影响决策)、基于社会特征的歧视性评分系统、用于预测个人刑事犯罪风险的评估系统等。

除 AI 专门立法外，欧盟《数字服务法》(DSA)对超大型在线平台(VLOPs)的系统性风险管理要求亦构成重要合规维度。 2024 年 10 月欧盟委员会对 Temu 的调查表明，平台企业需依据 DSA 第 34 条开展真实、有效的系统性风险评估，识别其商业模式、推荐算法、成瘾性设计(如游戏化奖励计划)等功能对消费者保护、产品合规等带来的真实风险，而非依赖过时或通用研究报告。Temu 因使用 2021 年市场报告评估 2023 年才上线的业务，且未分析其联盟营销计划对非法产品传播风险的促成作用，被认定违反勤勉评估义务，这揭示出 DSA 风险评估必须具有勤勉性、针对性与数据驱动性。

据此，企业在进入海外市场前，应当开展针对性的合规筛查，通过梳理目标国监管清单，明确自身产品是否落入禁止性或限制性范畴，若相关人工智能应用落入高风险限制性使用范畴，需按照监管要求完成相应合规义务，包括但不限于技术文档记录、影响评估、算法审计等，确保满足目标国的合规要求。

2. 透明度及标识义务履行

随着生成式人工智能在社交传播、内容创作、客户服务等场景的广泛应用，虚假信息传播、身份欺诈、深度伪造等风险日益凸显。在此背景下，"人工智能活动标识"成为全球监管共识，中国、欧盟、美国等主要经济体均将透明度义务纳入立法：要求人工智能服务提供者通过明确标识，区分人工智能生成内容与人类创作内容、人工智能交互行为与人类交互行为，帮助用户识别伪造信息、增强警惕、降低误信和受骗的风险。

例如，欧盟《人工智能法案》第 50 条规定，"提供商应确保，旨在与自然人直接互动的人工智能系统在设计和开发时，能让相关自然人知晓自己正在与人工智能系统互动……人工智能系统(包括通用人工智能系统)的提供者，若其系统生成合成音频、图像、视频或文本内容，应确保该人工智能系统的输出以机器可读取的格式进行标记，且可被检测为人工生成或篡改的内容。"此外，美国加州《人工智能透明度法案》(California AI Transparency Act)也要求人工智能服务提供者在生成内容时明确披露其"人工智能生成"的性质。

此外，欧盟 DSA 进一步强化了透明度要求，这些要求适用于大型或超大型的人工智能平台。DSA 第 27 条不仅要求平台以"简单易懂的语言"阐述推荐系统主要参数，还规定超大型平台必须提供至少一项不基于用户画像的推荐选项。这种"颗粒度"要求反映出欧盟监管对技术可解释性与用户控制权的极致追求，与我国企业追求效率优先、功能集成的产品思维存在根本性文化差异。对于出海人工智能企业而言，这表明仅遵守国内《人工智能生成合成内容标识办法》尚不足够。部分域外地区(如欧盟)对于"机器可读格式"和"互操作性"的强调意味着企业需从产品设计阶段入手，构建适配多区域监管要求的标识系统，确保既符合中国合规标准，又满足海外市场监管规定。

3. 文化合规与本地化

受训练数据的地域局限性与历史偏见的影响，人工智能应用的底层模型往往存在文化盲区，对不同国家及地区的宗教教义、历史叙事、社会习俗、政治敏感点缺乏深度理解与适配。这种文化适配不足可能导致企业在海外运营中触发舆情危机甚至法律风险。实践中，海外对虚拟角色生成内容的文化敏感度高，政治正确的讲究较多，相关产品稍有不慎就可能面临因内容违规被下架的风险。例如，在宗教信仰严格的国家和地区，人工智能生成的图片内容需要考虑人物衣着的覆盖区域和面积等因素。

更深层次的文化合规挑战源于监管价值观的根本差异。欧盟秉承二战以来个人权利至上的传统，高度重视用户自主权、透明度与弱势群体保护，而我国企业普遍受效率优先导向影响，习惯于将"方便性"置于"隐私"之前。Temu 案揭示，在中国习以为常的商业模式(如游戏化奖励、无限滚动)在欧盟可能因被认定为"黑暗模式"或成瘾性设计而遭到严格审查。

这种商业逻辑背后是文化价值观的差异，要求企业不能仅进行形式合规，而需理解并内化欧洲价值观，在产品设计中嵌入"尊重人性尊严"的理念。因此，人工智能企业需要关注文化的差异性，避免触发深埋的"文化地雷"，企业可采取以下策略:

- 一是构建本地化数据集，通过采集目标国文化相关数据，开展模型迁移学习，提升模型对本地文化的理解能力;
- 二是制定文化适配性内容过滤规则，结合目标国文化禁忌，定制系统提示词，从源头减少不当内容生成;
- 三是启用专项语义过滤模块，针对宗教、民族、历史等敏感领域，对人工智能输出内容进行二次审查，确保符合本地文化认知。

如您希望就相关问题进一步交流，请联系：

杨 迅
+86 21 3135 8799
xun.yang@llinkslaw.com

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

上海	北京	深圳
上海市银城中路 68 号 时代金融中心 19 楼 T: +86 21 3135 8666 F: +86 21 3135 8600	北京市朝阳区光华东里 8 号 中海广场中楼 30 层 T: +86 10 5081 3888 F: +86 10 5081 3866	深圳市南山区科苑南路 2666 号 中国华润大厦 18 楼 T: +86 755 3391 7666 F: +86 755 3391 7668
香港	伦敦	
香港中环遮打道 18 号 历山大厦 32 楼 3201 室 T: +852 2592 1978 F: +852 2868 0883	1/F, 3 More London Riverside London SE1 2RE T: +44 (0)20 3283 4337 D: +44 (0)20 3283 4323	

www.llinkslaw.com

Wechat: LlinksLaw

本土化资源 国际化视野