数美 NEXTDATA | 扬帆出海

2026

泛娱乐出海
内容合规白皮书

2026 Content Compliance Whitepaper for Global Digital Entertainment

全球泛娱乐出海市场最新发展趋势

欧美、东南亚、中东等热门区域的法规全貌

音视频/社交、游戏等主流行业的合规挑战与风控指南

数美科技AI业务风控系列白皮书

AI 风控专家

版权说明

本册主要采用文献综述、桌面调研、行业访谈等调研方法，所涉及的图片、数据、参考文献、新闻报道均采集于网络公开信息，并已标注来源。

本册以分析泛娱乐行业的企业在出海全球中各个区域的发展现状、合规风险，为泛娱乐类应用业务风控能力搭建提供行动建议，受公开资料和研究方法限制，本册内容和观点仅供读者参考和行业了解。

本册为《2026 泛娱乐出海内容合规白皮书》第一版，如有修订，请关注"扬帆出海"和"数美科技"公众号获取新版。

前言
FOREWORD

AI 时代，合规即是“智创未来”的基石

中国泛娱乐产业出海浪潮澎湃向前，内容合规已成为全球化征程中不可或缺的基石。伴随泛娱乐出海从“规模扩张”向“精细化运营”进阶，内容合规能力的高低直接决定着企业出海的“生存底线”与“发展上限”。从早期出海探索到如今全球化布局，中国泛娱乐企业在内容合规领域历经数载深耕，从被动合规走向主动构建合规体系，这一进程与全球监管环境的迭代深度交织。

近年来，全球内容合规监管呈现**“趋严化、差异化、技术化”**三大特征：欧盟《数字服务法案》（DSA）、美国《消费者告知法案》、东南亚多国数据本地化与内容审核新规等相继落地，各国对音视频、社交、游戏等泛娱乐内容的监管细则不断细化；不同地区在文化价值观、宗教禁忌、数据隐私等方面的合规要求差异显著，给企业全球化运营带来挑战；同时，以 AI 内容审核、隐私计算为代表的技术，正成为破解合规难题的关键抓手。据行业研究显示，2025 年全球内容合规技术市场规模同比增长超 40%，企业在合规技术上的投入占比持续提升。

泛娱乐出海内容合规领域正呈现三大发展趋势：**合规主体从头部企业向全行业延伸**，中小企业因合规能力不足面临更高出海风险；**合规策略从“事后整改”向“技术驱动的全流程合规”升级**，AI 审核、数据脱敏等技术深度融入内容生产、分发、运营全链路；**合规覆盖从单一品类向多元场景拓展**，从社交、直播到短剧、AIGC 内容，合规需求贯穿泛娱乐全业态。

在实践层面，内容合规的重要性愈发凸显：某国际社交平台因未满足欧盟内容审核合规要求面临巨额罚单；而依托智能合规技术构建全链路审核体系的企业，不仅成功进入高门槛市场，更凭借合规信任度实现用户增长。数美科技作为全球领先的智能内容与数据安全服务商，在 AI 内容审核、反欺诈、数据合规等领域积累了深厚技术壁垒；扬帆出海则是泛娱乐出海生态的核心连接器，拥有丰富的全球市场资源与行业洞察。

为助力泛娱乐从业者精准把握全球内容合规脉搏，**数美科技联合扬帆出海重磅推出《2026 泛娱乐出海内容合规白皮书》**。本白皮书深度解读全球五大出海区域的内容合规细则，剖析前沿合规技术的应用场景，挖掘合规驱动下的出海新机会点，旨在为企业筑牢合规防线、开辟全球化增长新路径提供专业指南。

CONTENTS

目录

01

全球泛娱乐出海的市场现状与趋势

Global Digital Entertainment:
Market Overview and Emerging Trends

全球泛娱乐市场版图重塑

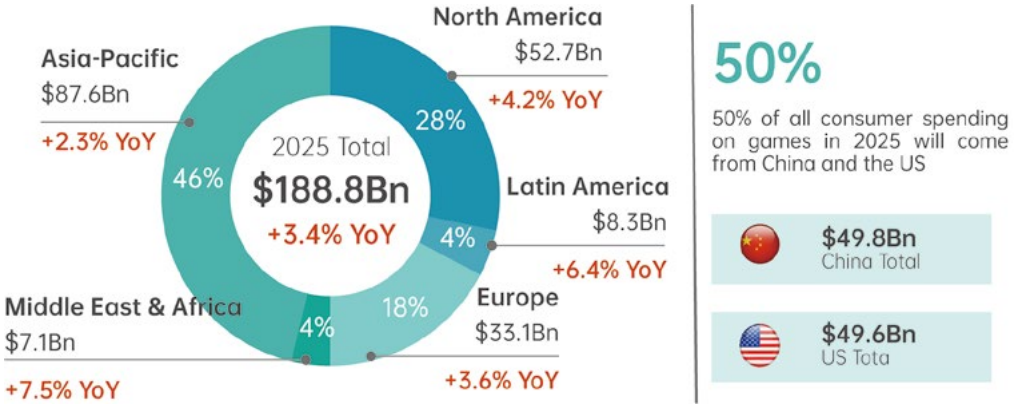
2025-2026 年全球泛娱乐出海整体规模与增速预测

进入 2025 年后，全球泛娱乐（Entertainment & Media, E&M）格局呈现**"增量由质引导、结构由数驱动"**的明显特征。总体上，PwC 将 2024 年行业规模确认在约 2.9 万亿美元，预计未来五年仍将增长但增速放缓（2025–2029 年 CAGR 约 3.7%），这意味着短期内行业更多依赖结构性优化与效率提升来驱动营收增长，而非单纯依靠高速体量扩张。

图源： PWC《2025-2029 年全球娱乐与媒体展望》

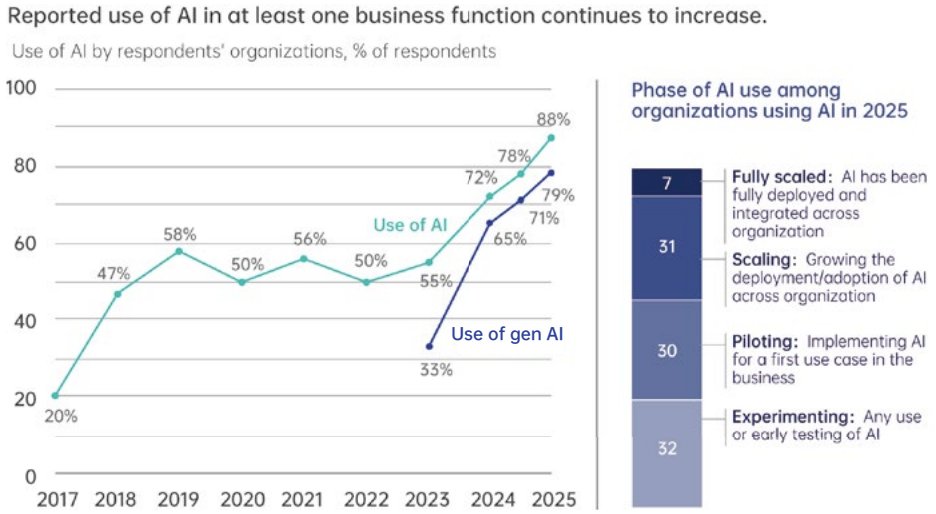
在细分类别上，广告与游戏成为短期内最直接的增量来源。广告端持续向数字化、程序化与 CTV（Connected TV）倾斜：**2024 年数字广告已在整体广告中占据高位**，比重持续抬升，程序化买量与基于大数据 / 模型的定向投放在 2025–2026 年将进一步放大货币化效率，这也对平台的技术能力与数据治理提出更高要求。

游戏保持为消费端最稳定且具弹性的增长引擎。Newzoo 预测 2025 年全球游戏收入约 **188.8 亿美元**，仍将实现小幅增长，并在 2026 年继续受新一批大作、主机 / 主机世代更替与为核心用户服务的高价发行所推动。值得注意的是，2025–2026 年游戏产业增长节奏将更依赖产品节奏与硬件生命周期，而非单一子品类独挑大梁。



The Global Market by Region - Revenue 2025 Forecast © Newzoo Global Games Market Report 2025 Newzoo《全球游戏市场报告》

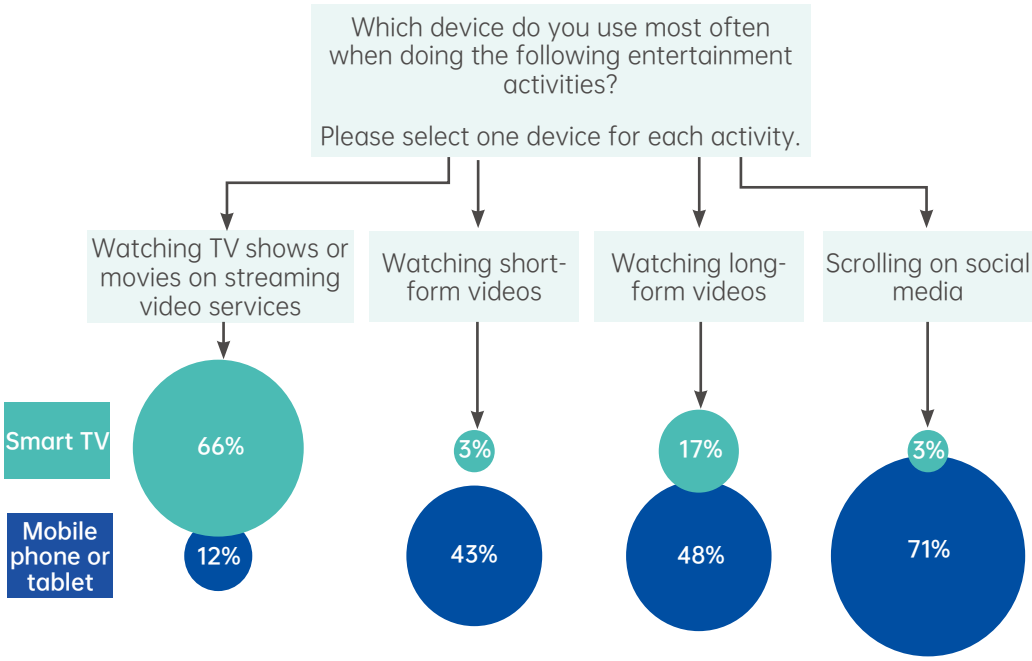
另一股深刻而系统性的变化来自 AIGC 对内容生产与分发链路的渗透。根据 McKinsey2025 年调研，企业在媒体与娱乐领域正从试点向规模化落地过渡。这意味着**短视频、图像、剧本原型等制作环节的自动化与快速迭代变得可行**，从而显著压缩单条内容的边际成本并提升小团队的产能。但同时，AIGC 带来的合规、版权与质量控制问题也在 2025–2026 年成为企业普遍需解决的运营命题。



McKinsey《2025 年人工智能的发展现状：智能体、创新与转型》

图注：McKinsey 调研结果显示，在企业层面，大多数企业仍处于试验或试点阶段，约三分之一的受访者表示其所在公司已开始扩大人工智能项目的规模。

结合上述两点，2025–2026 年全球泛娱乐的商业逻辑正在经历"放大效率 + 精准变现"的双重重构。一方面，AI 与程序化广告提高了内容触达与货币化的边际效益；另一方面，内容同质化、用户注意力碎片化与成熟站点的用户获取成本上升，迫使企业把更多资源投入到创意差异化、社区运营与本地化适配中。Deloitte 的趋势观测也指出，社交平台与创作者生态在吸引用户时间上持续对传统长视频 / 线性媒体形成挤压，内容提供方必须在产品形态与分发渠道上做出更灵活的策略调整。



Deloitte《2025 年数字媒体趋势：社交平台正在成为媒体和娱乐领域的主导力量》

图注：Deloitte 调查结果显示，从不同世代的受访者来看，他们的偏好正在从付费电视转向流媒体视频服务、社交视频平台和游戏。

对中国出海企业而言，2025–2026 年的几个直接影响值得注意：其一，变现端结构性机会（尤其是广告与跨平台分发）可在短期内弥补订阅或付费增速的放缓，但这需要更成熟的数据能力与合规框架以应对隐私与监管压力（如数据本地化与广告监管）。其二，游戏与泛娱乐产品的本地化成本在上升，不仅是语言本地化，还包括**叙事、玩法及社区治理的深度调整**。其三，AIGC 的工具红利为小团队和中小企业提供了更低的创作门槛，但要把"量产"转为"高质量变现"，仍需建立品牌信任与付费路径。

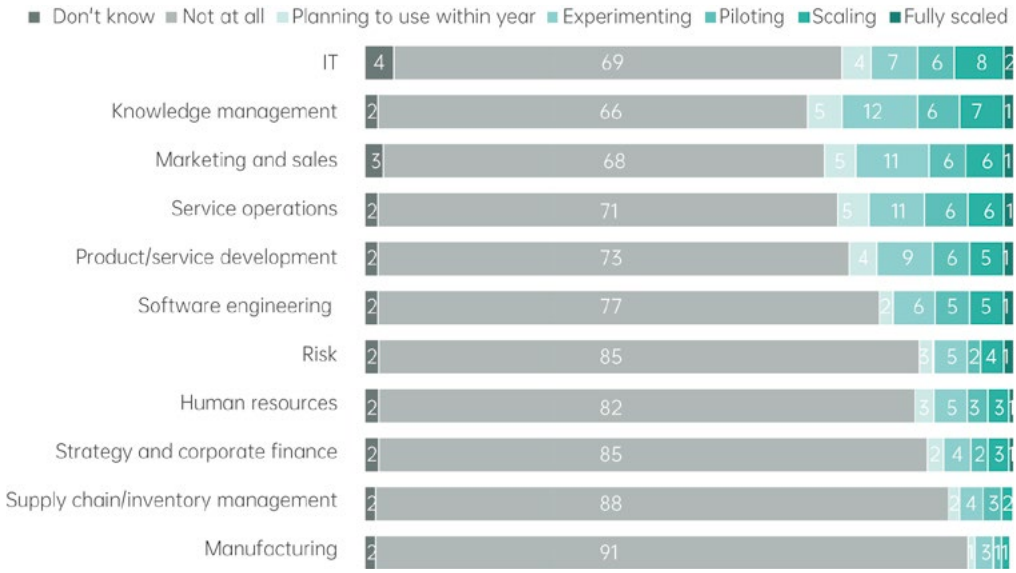
综合来看，2025–2026 年既是机遇窗口，也是优胜劣汰的试金石：能在**产品创新、数据治理与本地化运营**上实现协同的企业，更有可能在新一轮结构性转型中占据优势。

AIGC 对内容生产和消费模式带来的颠覆性影响

进入 2025 年，生成式人工智能（AIGC）已从概念与试验阶段迈向商业化落地。

No more than 10 percent of respondents report scaling AI agents in any individual function

Phase of AI agent use at respondents' organizations, by business function, % of respondents (n = 1,933)



图注：McKinsey 调研结果显示：在任何特定的业务职能部门中，只有不超过 10% 的受访者表示他们的组织正在扩展部署人工智能代理。

首先，AIGC 对生产端的直接效应是"**产能倍增 + 迭代提速**"。工具从文本到图像、音频与视频的生成能力在 2024–2025 年间快速成熟，越来越多的内容团队把 AIGC 嵌入脚本创作、素材合成、配音与多语言本地化流程中。世界经济论坛报告《Artificial Intelligence in Media, Entertainment and Sport》显示，企业正把 AIGC 用于创建短视频、广告片头、营销素材模板以及游戏关卡原型，这既降低了单位内容成本，也允许小团队以较低预算做出"量产级"内容池。

其次，在内容消费与分发层面，AIGC 与程序化广告 / 算法推荐的耦合正在改变流量与变现的节奏。程序化投放在 2025 年占据数字展示广告较高比重，当创意可以由模型按用户细分自动生成并被实时 AB 测试时，广告投放的效能与 ROI 有望提升，但这也进一步把平台的技术能力与数据治理放在核心位置。

同时，行业将面临质量分层与同质化风险并存的挑战。AIGC 降低了创作门槛，但大量低成本生成内容可能造成同质化潮流，使得用户注意力稀释、平台内容噪声增加，从而抬高差异化创意与品牌内容的价值溢价。调研结果显示，创作者对 AIGC 的态度是"既依赖又警惕"：工具能提升效率，但原创性、文化贴合度与情感深度仍是用户付费与长期粘性的关键。

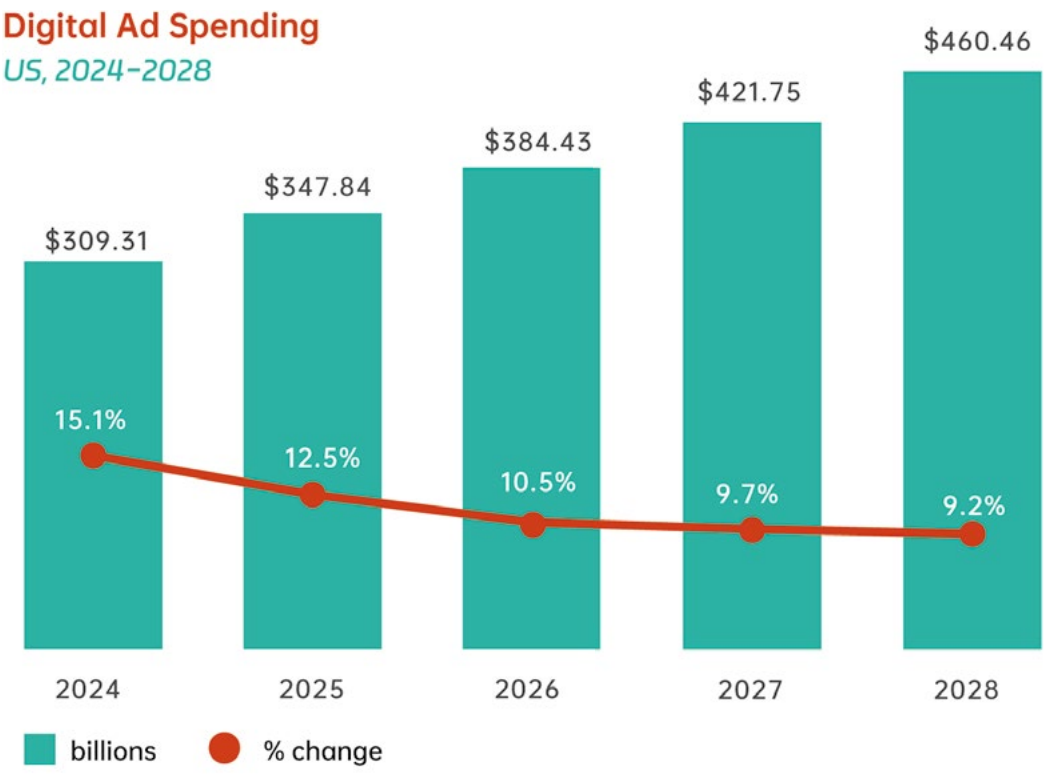
在 AI 创作层面上，内容可信度与平台治理成为消费端的核心议题。AIGC 带来的深度伪造（deepfake）、虚假信息与版权紊乱，易侵蚀用户对平台内容的信任，催生对生成内容溯源标识、内容鉴别工具与治理规则的需求。

最后，商业模式在短期内出现两类明显分化：一类是以 AIGC 为效率工具、将释放的产能用于放大广告与流量变现的轻量化内容商业体；另一类是以 AIGC 为创作放大器但强调原创性、IP 与社区运营的中高端内容生产方。前者受益于快速规模化与低边际成本进入点，后者则通过品牌力、版权与社区粘性来抵御同质化风险。对平台与投资人而言，判断依据从"单纯流量"逐步移向"变现深度与长期用户价值"。

出海企业面临的共同挑战：增长瓶颈、文化冲突、监管收紧

在跨境扩张过程中，企业面临的挑战不仅仅局限于市场竞争或运营模式的调整，更加复杂和多维的难题接踵而至。增长瓶颈、文化冲突以及监管收紧，成为出海企业在全球市场中不断遇到的核心问题。

在众多挑战中，增长瓶颈无疑是出海企业最为直观且紧迫的问题之一。随着全球市场逐渐趋于饱和，尤其是在成熟市场（如欧美、日韩等），用户获取的成本（CAC）急剧上升。根据 eMarketer 的数据，2024 年全球广告市场增速大幅放缓，许多成熟市场的广告成本呈现上涨趋势。对于出海企业而言，获取新用户的成本日益增高，且在这些市场上，竞争已经异常激烈，大型平台和本土企业占据了主要市场份额，致使新进者面临巨大的市场进入障碍。



Note: includes advertising that appears on desktop and laptop computers as well as mobile phones, tablets, and other internet-connected devices, and includes all the various formats of advertising on those platforms
Source: MARKETER Forecast. November 2024

此外，随着互联网广告价格的上涨，企业必须寻求更高效的方式来降低获客成本（CAC）并提高用户留存率。在广告的变现路径上，出海企业面临的挑战不仅仅是吸引新用户，更需要找到可持续的收入模式。例如，许多平台通过订阅、会员制和内购等多种方式来实现收入多元化，但在成熟市场中，用户的付费意愿趋于饱和，企业很难依赖单一的变现路径。在这一背景下，如何在已有市场中不断提高用户价值与生命周期（LTV），成为企业必须面对的重要任务。

除了增长瓶颈，文化冲突无疑是出海企业面临的另一大挑战。特别是在不同区域之间，文化差异更加显著。欧美市场的消费者更倾向于寻求个性化、自由度较高的产品和服务，而亚洲市场则可能更加注重产品的社交性和集体性。这使得很多企业在进入新市场时，需根据本地文化的需求进行相应的本地化策略调整。尽管全球化趋势在加强，但文化的多样性依然不可忽视。

随着全球对数据隐私和网络安全的监管要求日益严格，监管收紧成为出海企业在国际扩张过程中必须面对的一大障碍。除了欧美地区外，其他多个新兴市场近期推出的法规也加大了出海企业在合规方面的压力。

如韩国，自 2025 年 10 月起，外国游戏公司若要在韩国运营，必须设立本地代表，确保能够有效回应监管要求。此举是韩国政府为加强对外国游戏公司的监管力度而采取的措施，未设本地代表的公司将面临处罚。

在墨西哥，2025 年 3 月 21 日生效的《联邦私人持有 - 个人资料保护法》对数据保护框架进行了大幅修改，要求私营部门在处理个人数据时必须更加严格地遵循隐私保护规定。该法律强化了跨境数据流动的监管，要求数据处理方为数据传输提供充分的透明度，并强化了用户的隐私权利。

在印度，2025 年的《数字个人数据保护规则（草案）》对数据处理提出了新的要求，明确规定了数据受托方的通知义务、用户的知情权和数据删除权等权利保障。草案还规定了跨境数据流动的监管要求，要求数据接收方必须具备充分的安全保护措施。随着全球数据隐私保护法规不断收紧，出海企业在全球扩张过程中必须加强合规管理，确保在各个市场中遵循当地的法律要求，以避免合规风险和高额罚款。

随着全球数据隐私保护法规不断收紧，出海企业在全球扩张过程中必须加强合规管理，确保在各个市场中遵循当地的法律要求，以避免合规风险和高额罚款。

核心出海区域发展概览

中国泛娱乐企业出海已从战略备选升级为产业发展的必然路径，在全球数字消费升级与技术革新的驱动下，呈现出规模化、深度本地化的发展趋势。在这一进程中，不同市场展现出差异化的挑战：北美与欧洲市场用户付费意愿强烈，但对数据隐私、内容合规和 AIGC 监管提出极高要求；东南亚、中东等地区蕴含增长红利，却需严格把握宗教与文化敏感性；南美市场潜力巨大，但面临经济波动、支付渠道不完善和语言多样性等现实瓶颈。面对这一全球图景，企业需在合规适配、文化融合和本地运营上构建多层次策略，方能在泛娱乐出海的深水区行稳致远。

北美与欧洲： 市场成熟度、付费能力与高标准的隐私 / 内容治理要求

北美和欧洲作为全球数字经济的核心区域，其泛娱乐市场历经数十年发展，已形成高度成熟的生态体系。从两地应用下载榜 TOP10 来看，视频类平台占据主导地位，其中 TikTok 表现尤为突出，在 2025 年前三季度累计下载量突破 1 亿。榜单中还包括 Instagram、YouTube、Netflix 等本土老牌应用，以及近年来迅速崛起的中国短剧出海平台 ReelShort 和 DramaBox。这一优质生态的形成，主要源于以下几个因素：

2025 年 1-9 月 北美 iOS 与 Google 双平台 娱乐应用下载榜

排名	应用	热门分类	下载量
1	TikTok 内容形式 > 用户生... 109 新加坡 | TikTok Ltd.	娱乐 > 短视频	5438.44万
2	Instagram 内容形式 > 用户生... 82 美国 | Instagram, Inc.	娱乐 > 短视频	4627.01万
3	ReelShort 内容形式 > 短剧 25 美国 | NewLeaf Publishing	娱乐 > OTT	3945.54万
4	Spotify 内容形式 > 原创内容 56 瑞典 | Spotify	娱乐 > 音乐与音频	3417.25万
5	DramaBox STORYMATRIX PTE. LTD.	娱乐 > OTT	3332.88万
6	Netflix 内容形式 > 独家授权 52 美国 | Netflix, Inc.	娱乐 > OTT	2975.40万
7	Amazon Prime Video 内容形式 > 独家授权 62 美国 | AMZN Mobile LLC	娱乐 > OTT	2451.95万
8	HBO Max 内容形式 > 独家授权 41 美国 | WarnerMedia Global Digital Services, LLC	娱乐 > OTT	2409.66万
9	Tubi 内容形式 > 原创内容 36 美国 | Tubi, Inc	娱乐 > OTT	2373.88万
10	YouTube 内容形式 > 用户生... 70 Google	娱乐 > 视频分享	1985.40万

图源：点点数据

2025 年 1-9 月 欧洲 iOS 与 Google 双平台 娱乐应用下载榜

排名	应用	热门分类	下载量
1	TikTok 内容形式 > 用户生... 109 新加坡 | TikTok Ltd.	娱乐 > 短视频	5000.90万
2	Instagram 内容形式 > 用户生... 82 美国 | Instagram, Inc.	娱乐 > 短视频	3944.35万
3	Spotify 内容形式 > 原创内容 56 瑞典 | Spotify	娱乐 > 音乐与音频	2568.40万
4	Netflix 内容形式 > 独家授权 52 美国 | Netflix, Inc.	娱乐 > OTT	1938.88万
5	Amazon Prime Video 内容形式 > 独家授权 62 美国 | AMZN Mobile LLC	娱乐 > OTT	1733.08万
6	VK Видео 俄罗斯 | V Kontakte OOO	娱乐 > OTT	1695.03万
7	ReelShort 内容形式 > 短剧 25 美国 | NewLeaf Publishing	娱乐 > OTT	1563.13万
8	YouTube 内容形式 > 用户生... 70 Google	娱乐 > 视频分享	1510.83万
9	Disney 内容形式 > 独家授权 42 美国 | Disney	娱乐 > OTT	1392.90万
10	Twitch 内容形式 > 用户生... 50 美国 | Twitch Interactive, Inc.	娱乐 > 在线直播	1302.67万

图源：点点数据

首先，市场基础设施完善，用户成熟度高。这些地区的智能手机普及率、网络覆盖质量及速度均位居世界前列。更重要的是，用户对于数字娱乐内容的需求早已超越"有无"阶段，进入了追求"品质"和"体验"的深层次阶段。用户习惯于为优质的内容以及独特的增值服务付费。这种成熟的消费观念，直接转化为了强大的付费能力。根据《2025 至 2030 手机客户端行业产业运行态势及投资规划深度研究报告》显示，2025 年北美用户的 ARPU 值预计达到 58.3 美元，显著高于全球平均水平 32.1 美元，欧洲市场发展相对平稳，2025 年市场规模约 5860 亿美元，隐私保护法规的强化促使开发者更加注重数据合规。

其次，"订阅制"付费习惯深入人心。得益于 Netflix、Spotify、Apple Music 等全球流媒体巨头的长期市场教育，北美和欧洲用户已经普遍接受了"为服务持续付费"的模式，这种习惯为出海企业提供了稳定且可预测的收入流。

例如：低成本短剧《The Divorced Billionaire Heiress》以不到 20 万美元的成本，在北美创下 3500 万美元的票房奇迹，收益超 170 倍；语聊应用 YOHO、K 歌应用 StarMaker 等中国泛娱乐出海产品也通过提供高质量的语音聊天室和虚拟礼物系统，也在欧美市场获得了可观的付费用户。

短剧《The Divorced Billionaire Heiress》

然而，通往高价值市场的道路并非坦途。北美和欧洲拥有十分严格的数据隐私保护和内容监管法规，对出海企业提出了极高的合规要求。在数据隐私方面，欧盟的《通用数据保护条例》（GDPR）对用户数据的收集、处理、授权、跨境传输和泄露通知等都制定了极其细致和严苛的规定，违规处罚金额可高达全球营业额的 4% 或 2000 万欧元。在北美，美国加州的《消费者隐私法案》（CCPA/CPRA）也赋予了用户对其个人数据的强大控制权。用户对应用的隐私安全极为看重，任何涉嫌过度收集数据或泄露隐私的行为，都会迅速导致用户流失、品牌声誉受损乃至法律诉讼。

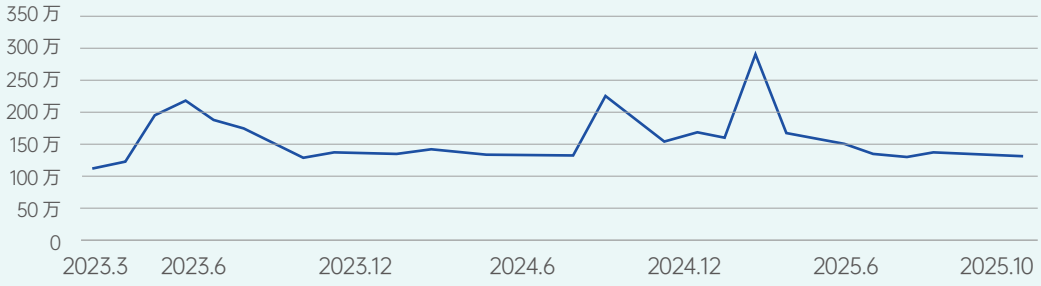
在内容治理方面，欧盟的《数字服务法案》（DSA）对在线平台的内容审核、风险管理和透明报告义务提出了系统性要求。这意味着，泛娱乐应用必须建立高效的内容审核机制，严厉打击仇恨言论、虚假信息、非法内容等，并为用户提供畅通的举报和申诉渠道。任何在内容治理上的松懈，都可能引发巨额罚款和在应用商店下架的风险。

东南亚： 人口红利、移动互联网渗透率与多元化宗教 / 文化敏感性

在全球化战略布局中，东南亚地区正成为中国泛娱乐企业出海的“新热土”。这一区域不仅拥有庞大的人口基数与年轻的用户结构，还展现出强烈的数字娱乐需求与活跃的移动互联网生态。以印尼、菲律宾等为代表的国家，35 岁以下人口占比超过 60%，年轻群体对短视频、社交、AIGC

等新兴娱乐形式接受度极高，形成了显著的"人口红利"与"内容消费红利"。例如，由中国团队开发的地图社交应用 Jagat 迅速风靡越南，不到三年时间，下载量已超 2000 万，显示出本地市场对创新型社交产品的强烈渴求。

2023.3—2025.10 Jagat 全球月下载量



图源：点点数据

东南亚地区在文化层面与中国具有一定亲近性，从今年前三季度东南亚地区下载榜就可以看出，国产短剧平台 DramaBox、Melolo、DramaWave，以及视频平台 BiliBili 都排在靠前位置，华语文化元素在影视、文学、游戏中具有一定接受基础。例如，中国短剧通过翻译配音后在印尼、马来西亚等地上线，用户对其"强情节、快节奏"的叙事模式表现出较高接受度，这与北美市场更偏好本土制作的真人短剧形成对比。

2025 年 1-9 月 东南亚 iOS 与 Google 双平台 娱乐应用下载榜

排名	应用	热门分类	下载量
1	TikTok 内容形式 > 用户生... 109 新加坡 | TikTok Ltd.	娱乐 > 短视频	4392.43万
2	TikTok Lite 内容形式 > 用户生... 109 新加坡 | TikTok Pte. Ltd.	娱乐 > 短视频	4376.67万
3	Instagram 内容形式 > 用户生... 82 美国 | Instagram, Inc.	娱乐 > 短视频	4167.02万
4	ReelShort 内容形式 > 短剧 25 美国 | NewLeaf Publishing	娱乐 > OTT	3975.52万
5	DramaBox STORYMATRIX PTE. LTD.	娱乐 > OTT	3908.70万
6	Spotify 内容形式 > 原创内容 56 瑞典 | Spotify	娱乐 > 音乐与音频	2770.11万
7	Instagram Lite 内容形式 > 用户生... 82 美国 | Instagram	娱乐 > 短视频	2372.93万
8	Netflix 内容形式 > 独家授权 52 美国 | Netflix, Inc.	娱乐 > OTT	2357.59万
9	Lark Player Biao Xie		2334.58万
10	GoodShort 内容形式 > 短剧 25 新加坡 | SINGAPORE NEW READING TECHNOLOGY PTE. L...	娱乐 > OTT	2252.15万

图源：点点数据

然而，东南亚也是一个宗教氛围浓厚、文化多元复杂的区域，尤其是伊斯兰教、佛教等在当地社会生活中具有深远影响力。宗教敏感性成为内容创作与运营中不可逾越的红线，一旦触碰可能引发用户抵制甚至法律风险。

在内容监管方面，多个国家已出台针对性法规。例如，印尼通过《电子信息与交易法》（UU ITE）明确禁止传播亵渎宗教、破坏宗教和谐的内容；马来西亚的《多媒体与通讯法案》也对"侮辱宗教"行为设定了严厉处罚。

> ***例如，****在印尼，一位知名网红曾公开发表"希望创建个人宗教，并让粉丝成为信徒"的言论，这一偏激观点在以伊斯兰教为主体的印尼社会引发强烈舆论反弹。更严重的是，2018 年 7 月 3 日，印尼政府以平台存在"亵渎神灵、色情及其他不当内容"为由，对 TikTok 实施全面封禁。尽管 TikTok 在 3 天内完成整改并恢复服务，但这一事件充分显示了宗教敏感问题可能引发的运营风险。*

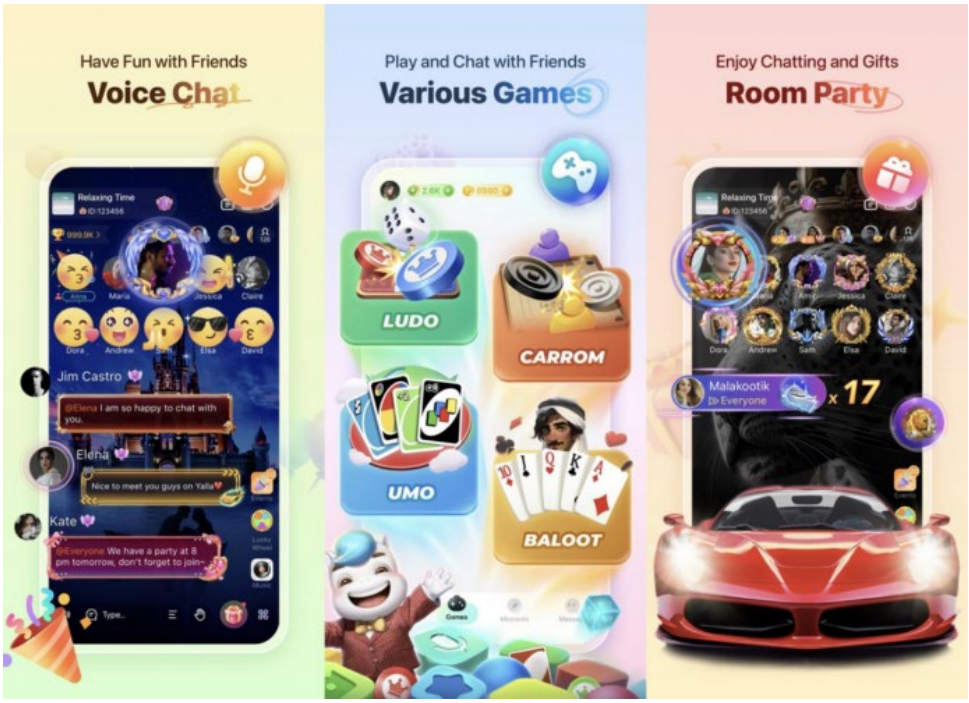
此外，即便在非宗教主题的泛娱乐产品中，文化冲突也常在用户互动中浮现。社交应用中的头像、表情包或虚拟礼物若使用宗教符号或特定动物形象，可能被用户视为不敬；直播主播的言行若被认定违背宗教伦理，也极易引发举报甚至舆论风波。例如，2023 年，印尼女网红作为穆斯林，为博流量发布自己吃猪肉的视频遭到了举报。

这些看似细微之处，恰恰是企业本土化运营中的"暗礁"。

中东与北非（MENA）： 强劲的消费能力、文化 / 宗教的严格监管

中东北非地区（MENA）正成为全球泛娱乐出海的重要市场，其中沙特阿拉伯、阿联酋等海湾国家表现尤为突出，具有人均收入高、互联网渗透率高、年轻人口占比大的特点，为数字娱乐产业提供了优质的发展土壤。据统计，2025 年初、中东的互联网普及率已超过三分之二，阿联酋、沙特等国更是超过 96%，且用户付费意愿强烈。

在这样一个高消费潜力、高娱乐化的市场，独特的阿拉伯文化背景既带来机遇也带来挑战。在机遇方面，伊斯兰教规对社交活动的限制催生了特定的数字娱乐需求。语聊应用因其匿名性和不依赖视觉呈现的特点，在当地获得快速发展。中国出海应用 Yalla、MICO 凭借游戏、PK 等强互动内容以及当地用户强劲的打赏能力，在 MENA 地区获得超过千万月活跃用户。



语聊应用 -Yalla

2025 年前三季度，中东地区社交应用下载榜上，除 WhatsApp、Facebook、X 等国际主流平台外，中国出海产品 Weplay 同样占据一席之地，通过提供纸牌、推理、绘画等轻量级的小游戏，帮助陌生人快速破冰。

2025 年 1-9 月 中东 iOS 与 Google 双平台 社交应用下载榜

排名	应用	热门分类	下载量
1	Telegram Messenger	社交聊天 > 即时通讯	1493.13万
2	Instagram	社交聊天 > 媒体共享网络	1339.52万
3	WhatsApp Messenger	社交聊天 > 即时通讯	1301.81万
4	Snapchat	社交聊天 > 媒体共享网络	1289.50万
5	Facebook	社交聊天 > 社交网络	920.30万
6	Getcontact	社交聊天 > 即时通讯	813.13万
7	Pinterest	社交聊天 > 媒体共享网络	812.56万
8	Threads	社交聊天 > 博客论坛	729.31万
9	X	社交聊天 > 博客论坛	683.89万
10	WePlay(ويبلاي)	社交聊天 > 其他社交应用	682.82万

图源：点点数据

然而，宗教文化也带来严格的合规要求。所有娱乐内容必须符合伊斯兰教法，严禁出现亵渎宗教、批判王室、宣扬 LGBTQ+ 内容或展示不当两性关系的情节。沙特视听媒体总局（GCAM）明确要求所有媒体内容不得违背伊斯兰价值观。2022 年，流媒体巨头 Netflix 就曾一部动画片中因出现的同性拥抱场景，被沙特电视台批评其向儿童"宣扬性偏差"，并要求平台删除"被认为违反伊斯兰价值观"的内容。

在商业化方面，当地用户对虚拟礼物、订阅制等付费模式接受度高，但支付渠道需要与当地系统对接。同时，内容策划需注意斋月等宗教节期的特殊性，运营策略需作相应调整。

总体而言，MENA 市场虽然潜力巨大，但需要出海企业在文化理解、合规投入和本地化运营方面做好充分准备。

南美：市场潜力、经济波动与语言（西班牙语 / 葡萄牙语）的复杂性

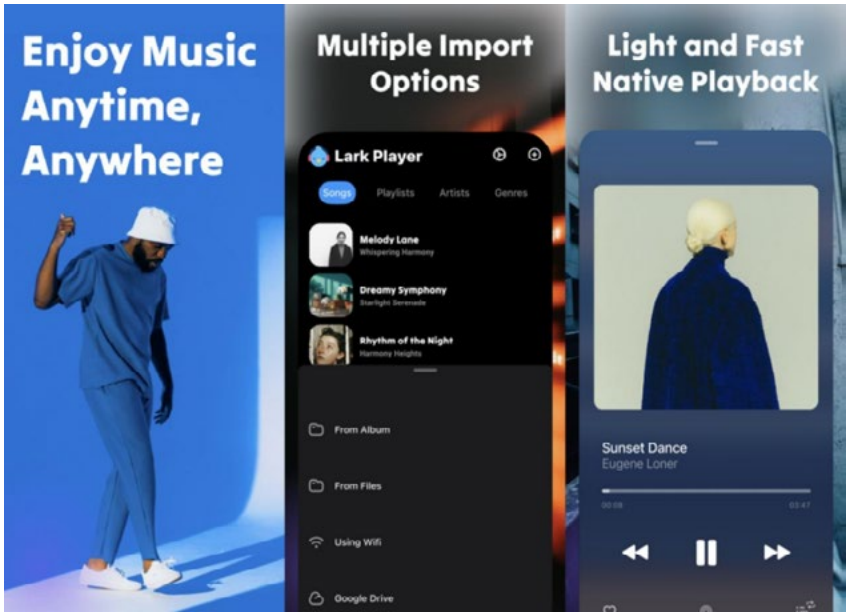
南美地区正以其庞大的人口基数、高涨的娱乐需求及快速增长的数字化进程，成为中国泛娱乐出海版图中一片充满潜力的新兴市场。巴西、阿根廷互联网渗透率已均超过 85%，年轻人口占比高，社交与娱乐应用使用时长持续增长。

2025 年 1-9 月 南美 iOS 与 Google 双平台 娱乐应用下载榜

排名	应用	热门分类	下载量
1	TikTok 内容形式 > 用户生... 109 新加坡 | TikTok Ltd.	娱乐 > 短视频	4392.43万
2	TikTok Lite 内容形式 > 用户生... 109 新加坡 | TikTok Pte. Ltd.	娱乐 > 短视频	4376.67万
3	Instagram 内容形式 > 用户生... 82 美国 | Instagram, Inc.	娱乐 > 短视频	4167.02万
4	ReelShort 内容形式 > 短剧 25 美国 | NewLeaf Publishing	娱乐 > OTT	3975.52万
5	DramaBox STORYMATRIX PTE. LTD.	娱乐 > OTT	3908.70万
6	Spotify 内容形式 > 原创内容 56 瑞典 | Spotify	娱乐 > 音乐与音频	2770.11万
7	Instagram Lite 内容形式 > 用户生... 82 美国 | Instagram	娱乐 > 短视频	2372.93万
8	Netflix 内容形式 > 独家授权 52 美国 | Netflix, Inc.	娱乐 > OTT	2357.59万
9	Lark Player Biao Xie		2334.58万
10	GoodShort 内容形式 > 短剧 25 新加坡 | SINGAPORE NEW READING TECHNOLOGY PTE. L...	娱乐 > OTT	2252.15万

图源：点点数据

南美深厚的音乐文化传统为特定娱乐形态提供了天然土壤，特别是巴西、哥伦比亚等国民众对音乐有着极高热情，户外聚会、公放音乐是常见的社交方式。中国出海企业敏锐捕捉到这一需求，例如音乐播放器应用 Lark Player 通过将音量输出比普通播放器提高 200%—300%，精准契合了当地用户"大声公放"的习惯，成为本地化创新的成功案例。也进入了 2025 年前三季度南美娱乐下载榜排名 Top10 之内。



音乐播放器 -Lark Player

短剧赛道在南美同样蕴含巨大机会。基于肥皂剧收视基础，当地用户对狗血剧情接受度高。不过，目前成功打入市场的中国短剧产品仍属少数，核心挑战在于本地化深度不足。短剧出海不仅需完成西班牙语、葡萄牙语等语言的精准翻译，更需在题材、价值观、演员选择上与本地团队深度合作。直接移植一些传统的东亚题材往往会遭遇水土不服，而融合拉美本地社会、文化议题（如毒枭、黑帮、贫民窟、足球等）的内容更易引发共鸣。

然而，**南美市场的潜力与挑战并存**。其经济波动性较高，阿根廷、委内瑞拉等国通胀问题长期存在，影响用户稳定付费能力，且支付基础设施亦不均衡。语言多样性进一步增加运营难度，除西班牙语、葡萄牙语外，印第安语、瓜拉尼语等小众语言群体仍需覆盖。此外，当地华人从业者远少于北美，导致市场洞察、资源对接和文化理解成本较高，提升了初创团队的进入门槛。

综合而言，**南美市场为泛娱乐出海提供了可观的空间**，但成功关键在于"深度本地化"与"韧性运营"的双重能力。出海企业需在内容创作上真正融入本地文化语境，同时在商业模型中灵活适配经济波动与支付习惯，方能在这一高潜力但高门槛的市场中建立可持续的竞争力。

2026 泛娱乐出海三大合规焦点

随着全球化的深入发展，泛娱乐产业的出海已成为众多企业寻求增长的重要战略。然而，在拓展海外市场的进程中，企业面临着日益复杂的合规挑战。2026 年，价值观合规、数据 / 隐私合规以及 AIGC 合规将成为泛娱乐出海的三大关键焦点，这些焦点问题不仅关系到企业在海外市场的合法运营，更直接影响着企业的品牌形象、用户体验和市场竞争力。

价值观合规：跨文化敏感性、政治正确与仇恨言论的识别

跨文化敏感性

在全球化的泛娱乐市场中，不同国家和地区有着独特的文化背景、宗教信仰、社会习俗和价值观念。泛娱乐内容，如游戏中的角色设定、剧情故事，影视作品中的情节和形象等，如果忽视了这些跨文化差异，很容易引发文化冲突和用户反感。在某些宗教文化中，特定的形象或行为是被严格禁止或视为不尊重的。若游戏或影视作品中不小心涉及到这些元素，可能会引起当地用户的强烈不满，甚至引发抵制行为，如“龙”这一形象，在东方文化中是权力和祥瑞的象征，但在西方文化中，“龙”却常与邪恶、破坏相关联，因此在产品中需要格外注意对此类形象的设计方式。此外，中东地区受伊斯兰教影响，要避免涉及猪肉、十字架等元素；东南亚国家多元宗教并存，需格外注意宗教节日和习俗。

再如，不同文化对于颜色、数字、手势等也有着不同的象征意义。在中东地区，受伊斯兰文化影响，绿色被视为神圣的颜色，代表吉祥、安宁；而黄色在部分中东国家则象征死亡和不吉利；数字 5 因与伊斯兰教中的“五功”相关，有积极的寓意，而数字 13 和数字 4 则常被认为不吉利。针对这些文化敏感要素，如果在泛娱乐内容中不加以注意，可能会传递出错误的信息。

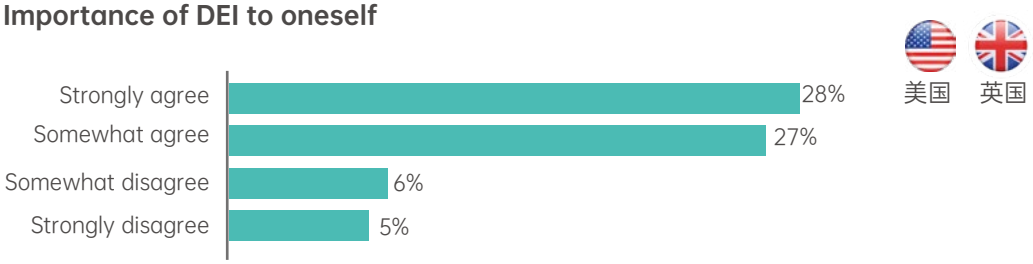
以《原神》为例，该游戏在中东版本中额外增加了附档和低温化身没的绿色系服饰，并且在推出限时活动时也避开了可能引发歧义的数字组合，设计了五人小队挑战，并推出"绿色能量徽章"等道具，同时将游戏的关键操作按键调整至屏幕右侧（中东用户多为右手持机，左手被视为'不洁'）。这一案例说明，颜色、数字等看似细微的文化符号，实则是泛娱乐产品出海时必须直面的"隐性规则"，唯有精准识别并适配，才能跨越文化鸿沟，实现产品的本地化落地。此外，还需要注意的是，文化是动态变化的，企业需要持续关注和更新相关知识。并且考虑到企业内部的创作团队可能来自不同的文化背景，如何确保团队成员在创作过程中充分考虑到目标市场的文化差异，也是一个需要解决的问题。

政治正确与仇恨言论的识别

不同文化背景的用户对"政治正确"的敏感点存在显著差异。在欧美市场，公众对种族平等（如避免黑人角色被刻板化为'罪犯'或'运动员'、亚裔角色被固化成'数学天才'或'外卖员'）、性别平等（如女性角色不能仅被塑造成"花瓶"或"辅助者"）、LGBTQ+ 群体可见性（如游戏中需包含非二元性别或同性伴侣角色）、残疾人群体代表性（如视障 / 听障角色需有合理叙事功能而非单纯"励志符号"）的关注度极高。

中东与亚洲市场虽对"政治正确"的表述方式与欧美存在差异，但对宗教尊重（如伊斯兰文化中避免描绘先知形象、禁用酒精 / 猪相关元素）、性别规范（如部分国家要求女性角色着装符合传统习俗）同样敏感。例如，沙特阿拉伯要求游戏内女性角色必须佩戴头巾（除非剧情明确设定为非穆斯林角色），否则可能面临下架；印度市场对"牛"等宗教象征物的呈现极为谨慎，相关情节可能引发大规模抗议。

因此，推出一款面向全球市场的产品需要遵循 DEI（Diversity, Equity, and Inclusion，即多元、平等、包容）相关策略，以游戏产品为例，最基础的，就是在角色创建时提供更多的发型、体型选择以及 NPC 外貌和年龄的多样性等。根据 Newzoo 对美国和英国玩家进行的一项调查显示，针对"DEI 对我来说重要吗"这一问题填写重要的玩家数量占样本总数的 55%。



图源：Newzoo Gamer Sentiment Study 2022：DEI

但，并非所有关于 DEI 的反馈都需要执行，正确理解 DEI 和表达方式很重要，如果盲目按照玩家发聩执行所有内容，就会令其变成某种为了正确而正确的标签性产物，从现实角度看是极其不合理的。如《黑神话：悟空》也曾面临"性别歧视"的指控，同样受到这类"政治正确"影响的还有《消逝的光芒 2》《战神 4》以及《刺客信条》等系列作品，只不过，后者们无奈接受了"政治正确"的调控，创造出了一系列令人难以评价的女主角，而《黑神话：悟空》则拒绝向 Sweet Baby Inc（加拿大游戏资讯公司，有名的政治正确团队）妥协，拒绝支付 700 万美元起步的整改费，坚决保护自己创作的角色，依然在游戏上线后收获了惊人的效果。

因此总结而言，政治正确与仇恨言论的识别，本质上是泛娱乐企业对全球多元文化的"尊重能力"与"治理能力"的综合考验。在合规和保护自身产品价值的前提下，如何将"政治正确"的价值观融入内容创作的核心逻辑，将"仇恨言论"治理转化为用户信任的建立过程，将会成为打造兼具商业价值与社会责任的全球化泛娱乐产品的关键。

数据 / 隐私合规：GDPR、CCPA 等区域性隐私法的落地与执行

区域性隐私法的核心要求：以用户为中心的权利体系与责任边界

在数字经济与全球化深度融合的 2026 年，数据已成为泛娱乐产业（涵盖游戏、影视、动漫、网络文学等）的核心生产要素与用户连接纽带。随着全球用户对个人隐私保护的认知大幅提升，各国针对数据收集、存储、使用与共享的监管力度持续强化，区域性隐私法如欧盟《通用数据保护条例》（GDPR）、美国加州《加州消费者隐私法案》（CCPA）及其后续扩展法规（如 CPRA）、加拿大《个人信息保护和电子文件法》（PIPEDA）、新兴的巴西《通用数据保护法》（LGPD）、印度《个人数据保护法案》、南非《个人信息保护法》（POPIA）等，已从"推荐性指引"演变为泛娱乐企业出海的"强制性合规门槛"。这些法规虽在具体条款上存在差异，但核心均指向对用户个人数据隐私权的严格保护，要求企业在数据全生命周期管理中承担明确的法律责任，并且可以预见的是，全球隐私法规还将继续演进，AI 驱动的监管或许将成为焦点，预计还会有更多国家出台类似 GDPR 的框架，这些将成为泛娱乐企业全球化运营中不可忽视的挑战。

以最具代表性的 GDPR 与 CCPA 为例，其核心是通过构建"用户赋权 + 企业问责"的双重机制，确保个人数据的合法合规使用。GDPR 作为全球隐私保护的'黄金标准"，明确规定了数据控制者（如泛娱乐企业）与处理者（如第三方数据分析服务商）的多项义务：

> *其一，合法性基础要求企业在收集用户个人数据（包括但不限于姓名、邮箱、IP 地址、设备标识符、游戏内行为数据、消费记录等）时，必须基于用户的"明确同意"或法定的其他合法依据（如合同履行必需），且同意需以清晰、易懂的方式获取，不得通过默认勾选或模糊表述诱导用户；*
>
> *其二，用户权利保障赋予用户完整的控制权——用户有权随时访问、更正、删除其个人数据（即"被遗忘权"），有权限制企业对其数据的使用，有权拒绝企业基于商业目的的数据分析或精准营销，并有权获取数据被共享至第三方的具体信息；*
>
> *其三，数据最小化与目的限定原则要求企业仅收集实现特定功能所必需的最少数据，且数据用途不得超出最初告知用户的范围（例如，若游戏仅因登录验证需要收集手机号，则不得将其用于广告推送）；*
>
> *其四，安全保障义务要求企业采取"适当的技术和组织措施"（如加密存储、访问权限控制、定期安全审计）防止数据泄露、篡改或丢失，若发生数据安全事件，需在 72 小时内向监管机构报告，并视情况通知受影响的用户。*

此外值得一提的是，自 2025 年开始，GDPR 的执行更注重 AI 应用，比如在社媒算法中使用个人数据时，必须评估隐私影响。

CCPA 及其扩展法案 CPRA 则聚焦于加州消费者的隐私保护，核心在于赋予用户对个人数据的"知情权"与"选择权"。根据规定，泛娱乐企业若在加州开展业务并收集消费者个人数据（包括通过游戏内购、广告互动、账号注册等场景），需明确告知用户收集的数据类型（如姓名、联系方式、支付信息、设备信息、浏览历史）、使用目的（如个性化推荐、广告分发）及数据共享对象（如第三方广告商）；用户有权要求企业删除其个人信息，有权拒绝企业出售其个人数据（特别是针对 16 岁以下未成年人的数据，默认禁止出售），并有权获得企业数据收集行为的清晰说明。与 GDPR 相比，CCPA 更侧重于规范企业的数据商业化行为（如数据交易与精准广告），但其执法力度同样严格——加州隐私保护局（CPPA）可对违规企业处以高额罚款（最高可达每起违规事件 7500 美元 / 用户）。

除欧美地区外，其他新兴市场也在加速完善隐私法规体系：巴西 LGPD 借鉴了 GDPR 的核心框架，要求企业设立数据保护官（DPO）、建立数据跨境传输的合法性基础；印度《个人数据保护法案》则将个人数据分为"关键个人数据"（仅限境内存储）、"敏感个人数据"（如健康、财务信

息需严格限制跨境传输）与"一般个人数据"，并针对不同类别数据设定差异化的处理规则。这些法规的共同特点是：从被动合规转向主动隐私设计，以用户为中心构建权利体系，以企业为责任主体明确义务边界，且监管范围覆盖所有处理用户数据的主体（无论是否位于当地），持续加强对跨境数据转移的审查。

泛娱乐企业的合规挑战：全球化运营与区域差异的复杂博弈

对于泛娱乐企业而言，区域性隐私法的落地执行带来了多维度的挑战，核心矛盾在于"全球统一服务模式"与"区域差异化合规要求"的冲突。2025 年，随着中美欧数据摩擦加剧，这些冲突愈发凸显。

首先，数据处理的复杂性显著增加。泛娱乐产品的用户覆盖全球，同一款游戏或应用可能同时收集来自欧盟、美国、东南亚、中东等不同地区用户的数据。例如，一款全球发行的 MMORPG 游戏中，欧盟用户的 IP 地址、游戏内社交关系链、充值记录属于 GDPR 保护范畴；加州用户的设备传感器数据（如 GPS 定位用于个性化活动推荐）、广告 ID（用于精准广告投放）需符合 CCPA 要求；巴西用户的个人身份信息（如 CPF 税号）则受 LGPD 约束。企业需要针对不同区域的数据类型、收集场景与使用目的，分别制定合规策略——这意味着同一功能模块（如用户登录系统）可能需要部署多套数据收集逻辑，根据用户所在地区动态调整权限请求与信息披露内容。

其次，用户权利的响应难度升级。GDPR 与 CCPA 均赋予用户"数据可携权"（如导出个人数据副本）与"被遗忘权"（如要求彻底删除账户及相关数据），而泛娱乐企业的用户数据通常分散在多个系统中：游戏服务器存储基础账号信息与行为日志，第三方分析工具记录用户活跃度与留存数据，广告平台持有用户画像标签，云端存储保存游戏存档与虚拟物品交易记录。当用户行使权利时，企业需跨系统整合数据、验证用户身份真实性，并在规定时限内（GDPR 要求 72 小时内响应数据访问请求，CCPA 要求 45 天内响应删除请求）完成操作。

再者，第三方合作的合规风险传导加剧。泛娱乐企业的运营高度依赖第三方服务提供商——从游戏引擎、数据分析工具、广告投放平台，到云服务、支付渠道。根据 GDPR 与 CCPA 的"连带责任"原则，若第三方合作伙伴在数据处理过程中违规，企业作为数据控制者仍需承担主要责任。

1）AI 与新兴技术隐患

目前生成式 AI 已经在社媒广告中应用广泛，但如果涉及未经授权的个人数据，则可能被认定为违规。GDPR 要求在高风险 AI 使用场景下进行 DPIA（数据保护影响评估），泛娱乐企业若忽

视这一点，可能受到监管机构调查。

2）儿童与敏感数据保护

美国 COPPA（美国儿童在线隐私保护法）要求 13 岁以下用户数据必须经家长同意，欧盟也对未成年用户信息设置额外保护。特别是泛娱乐产品需要格外注意。

最后，监管执法的严厉性与不确定性并存。欧盟数据保护机构（如爱尔兰数据保护委员会、法国国家信息与自由委员会）对跨国企业的监管力度持续加强，罚款金额屡创新高——2025 年某国际社交平台因违规处理儿童用户数据被 GDPR 处罚 12 亿欧元；美国加州隐私保护局则重点打击隐蔽数据商业化"行为（如通过游戏内活动变相收集未成年人信息用于广告定向）。更复杂的是，部分新兴市场（如东南亚、中东）虽已出台隐私法规，但具体执行细则仍在完善中，企业可能因对"当地监管尺度"的误判而面临合规风险（如某中东国家要求游戏内用户数据必须存储在本国境内，但未明确"存储期限"与"跨境传输例外"，导致企业合规方案反复调整）。

此外，在中国泛娱乐企业出海时，还需注意中国本土的《个人信息保护法》（PIPL）与海外法规的衔接。虽然 PIPL 虽与 GDPR 相似，但跨境数据出口还需经国家安全评估，进一步增加了双重合规风险。

合规策略与实践：从被动应对到主动治理

为了应对区域性隐私法的挑战，泛娱乐企业需构建"全流程覆盖、多主体协同"的合规体系。

首先，可以建立隐私合规团队，聘请熟悉 GDPR 和 CCPA 等法规的法律顾问，定期审计数据流程。2025 年起，推荐采用"Privacy by Design"原则，从营销策划阶段嵌入隐私考虑。

其次，在选择合作伙伴时，也要将隐私合规能力纳入供应商评估的核心指标，要求第三方提供 GDPR/CCPA 合规认证、定期提交数据处理报告，并在合作协议中明确约定数据保护责任。

第三，优化用户同意机制。部署先进的同意管理平台，提供易懂的隐私政策和撤回选项，避免隐形数据收集。

第四，加强数据安全与本地化。对于跨境传输，使用欧盟认可的 SCC 或绑定性公司规则（BCR）。泛娱乐企业可选择在目标市场设立数据中心，同时，实施数据匿名化技术，减少个人标识符使用。

第五，企业需对用户数据进行精细化分类，并针对不同区域法规要求，明确每类数据的"收集必要性"。例如，在欧盟市场，若产品可通过匿名 ID 实现登录（无需绑定手机号或邮箱），则优

先采用匿名化方案；在加州市场，若广告推送可通过上下文关联而非用户画像实现，则减少对敏感标签（如年龄、性别）的收集。同时，通过"Privacy by Design"理念，在产品开发初期嵌入数据合规要求——例如，游戏注册页面仅展示必需填写的字段（如用户名），可选字段（如生日、性别）需单独征得用户同意，并明确说明用途。

最后，前瞻布局新兴趋势，动态适应监管变化。随着隐私监管趋严，零方数据（用户主动分享的偏好信息，如问卷填写的兴趣标签）与上下文广告（基于当前页面内容而非用户画像的精准推送）将成为减少对个人数据依赖的重要方向，企业可提前探索此类模式的可行性。对于 AI 技术的应用，需严格执行"数据保护影响评估（DPIA）"，确保算法不歧视或泄露隐私。此外，企业需建立常态化的全球监管动态跟踪机制，及时了解全球监管动态。

AIGC 合规： AI 生成内容的版权、虚假信息（Deepfake）治理

在 AIGC 技术深度渗透泛娱乐产业的 2026 年，AI 生成内容已成为内容生产的核心工具之一。然而，其'无明确创作主体'"数据来源复杂'"生成逻辑不可解释'的特性，也带来了两大核心合规挑战：AI 生成内容的版权归属模糊性与虚假信息的传播风险。

传统版权法的核心逻辑是"保护人类创作者的独创性表达"，而 AIGC 的"非人类创作"属性直接冲击了这一基础。当前全球尚未形成统一的 AI 生成内容版权认定规则，但主流司法辖区（如欧美、日韩）的监管趋势已逐渐清晰：若 AI 生成内容缺乏人类创作者的"实质性智力投入"（如仅通过简单指令生成通用模板），则该内容通常无法获得版权保护；反之，若人类对生成过程进行了深度干预（如设定具体情节框架、调整艺术风格参数、筛选最终输出结果），则可能基于"人类创造性贡献"主张部分权利。

这一规则在泛娱乐场景中引发了多重争议。首先是版权归属的模糊性：举例来说，某游戏公司使用 AI 生成了一组奇幻风格的角色立绘，若策划团队仅输入"精灵战士"等宽泛关键词，而 AI 自主完成了细节设计，则该立绘可能因"缺乏人类独创性表达"无法登记版权；但若美术设计师基于 AI 生成的初稿，进一步修改了面部表情、配色方案等，最终成果则可能被认定为"人类主导的衍生创作"，从而获得版权保护。其次是训练数据的合法性问题：AI 模型的训练通常依赖海量现有作品，若训练数据中包含受版权保护的内容，且生成内容与原作品存在实质性相似，则可能构成"间接侵权"——即使企业未直接复制原作品，但因使用了包含侵权数据的模型，仍需承担法律责任。

此外，更复杂的挑战来自跨国版权规则的差异：欧盟部分国家要求 AI 生成内容必须明确标注是否包含受版权保护数据的训练痕迹"，否则可能被认定为"虚假署名"；美国版权局则明确拒绝为"完

全由 AI 生成且无人类干预"的内容颁发版权证书，但允许人类对 AI 辅助创作的内容申请有限保护；日本则倾向于通过行业自律规范，要求企业披露 AI 训练数据的来源及使用范围。这些差异使得泛娱乐企业在出海时，需针对目标市场的版权规则调整内容生产策略——例如，在欧盟市场发布 AI 生成的音乐专辑时，需附上"训练数据未包含受版权保护作品"的合规声明；在北美市场推广 AI 辅助设计的游戏皮肤时，则需保留人类设计师调整过程的记录，以备版权审查。

Deepfake 技术在泛娱乐领域的应用具有两面性：

> *一方面，它可以低成本生成高质量的虚拟角色动画、影视特效或游戏过场动画，提升内容生产效率；*
>
> *另一方面，若被滥用，则可能引发严重的虚假信息传播问题，甚至演变为声誉危机与法律纠纷。*

从法律层面看，各国对 Deepfake 的监管正趋于严格：欧盟《数字服务法案》（DSA）要求平台对"可能造成公众误解的合成内容"添加明显标识（如"本视频包含 AI 生成内容"），并建立快速删除机制；美国部分州（如加州、德州）已立法明确，未经授权使用他人肖像或声音生成 Deepfake 内容属于违法行为，受害者可提起民事赔偿；日本则将针对公众人物的 Deepfake 诽谤内容纳入《刑法》中的"名誉毁损罪"范畴。泛娱乐企业若未能有效管控 Deepfake 风险，可能面临用户诉讼、平台下架甚至监管处罚。

在下一章节，将针对不同区域的内容合规法规详细阐述区域性的合规差异和治理要求。

02

全球内容合规的法规差异与治理要求

Global Expansion of Digital Entertainment in the AI Era:
Compliance Challenges and Vertical-Specific Solutions

东南亚：文化多元与宗教敏感交织

东南亚作为全球经济快速发展的区域之一，人工智能行业也展现出蓬勃的发展势头。东南亚各国在政治、文化、法律制度等方面具有鲜明的区域特性，因此，在生成式人工智能的应用和发展过程中，合规问题显得尤为重要。不同国家的法律框架、AI 合规相关的法规、伦理标准以及政府监管政策等方面都需要特别关注和遵守，以确保技术创新与法律规范的有机结合。

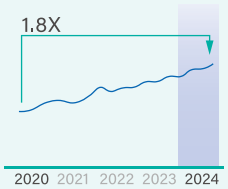
东南亚地区经济快速发展，各国在科技创新方面也展现出巨大的潜力。然而，由于各国在政治、经济和文化上的差异，生成式 AI 的法律监管和合规要求也各不相同。这使得东南亚的 AI 行业在合规方面面临独特的挑战和要求。针对出海东南亚的六个热门国家——**新加坡、越南、泰国、马来西亚、印度尼西亚以及菲律宾**，围绕内容合规重点以及相关监管政策进行介绍与解读，为企业出海东南亚时提供必要的合规参考。

据统计，东南亚 6.75 亿人口的互联网渗透率已达 75%。《2024 年东南亚数字经济报告》预测，到 2030 年，东南亚地区数字经济规模将攀升至 2 万亿美元，其中 AI 市场规模有望突破 5800 **亿美元**。与此同时，东南亚国家积极向外资企业释放政策红利，在税收减免、财政补贴、投资准入及数据治理便利化等领域提供有力支持。

报告同时显示，2020 至 2024 年，**东南亚人工智能相关搜索量增长 11 倍**。其中**新加坡、菲律宾、马来西亚更跻身全球 AI 搜索量前十国家**，这一数据进一步印证东南亚在人工智能领域的领先地位，使其成为企业出海的优选市场。

The increasing digital fluency of SEA users opens up new avenues for growth as businesses adapt

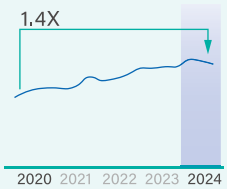
Category exploration



More room for competition as consumers explore options

The increasing prevalence of brand-agnostic searches for services and products intensifies competition, compelling businesses to actively seek ways to differentiate their offerings.

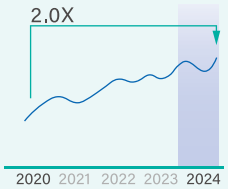
Use of local languages



Brands can win by speaking the language of their consumers

SEA users consistently demonstrate a preference for using local languages. Businesses that invest in locallyrelevant experiences and offers can unlock stronger growth, increasing the likelihood of usage, repeat purchases, and platform loyalty.

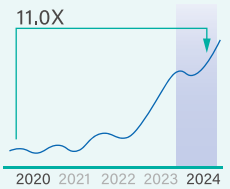
Search specificity



Consumers search for relevance with increasingly specific queries

Searches are becoming longer, signalling demand for information and services that are more tailored to their needs. To improve conversion rate and deepen engagement, brands must adapt their strategies to become more relevant.

AI interest



The potential of AI comes to light as consumers seek to learn more

The surge in searches relating to AI indicates growing consumer interest in exploring and adopting AI-powered tools and services.

来源：《2024 年东南亚数字经济报告》

- **文化多元复杂的东南亚地区，在内容审查上有哪些重点关注的领域？**
- **企业出海东南亚又有哪些必须注意的内容红线？**

以下将针对新加坡、越南、泰国、马来西亚、印度尼西亚以及菲律宾六国内容合规要点进行解读。

新加坡

AI 发展及相关监管政策

新加坡在其国家人工智能战略（NAIS）中宣布了一个宏伟的目标，就是要走在开发和部署具有可扩展性和影响力的人工智能解决方案的前沿，**成为全球人工智能解决方案的开发、测试、部署和扩展的中心**。该战略提出了五个"生态系统促进因素"，旨在推动人工智能的应用，其中一个就是为人工智能创造一个"进步和可信"的环境，即在创新和社会风险之间找到平衡点的环境。为了建立这样一个"进步和可信"的环境，新加坡目前对人工智能的监管采用了一种较为宽松和自主的方式。

2019 年 1 月，新加坡个人数据保护委员会**发布了亚洲首个《人工智能治理模型框架》**，并于 2020 年对其进行了更新，框架**奠定了"透明性""以人为中心"两项基本治理原则**，形成了新加坡人工智能治理框架的雏形。

面对生成式人工智能放大传统风险和激发新兴风险的问题，新加坡政府认为有必要更新早期的模型治理框架，以全面解决新出现的问题。因此，2023 年 6 月，新加坡通讯与信息部下的资讯通信媒体发展管理局（IMDA）与多家全球领先科技企业发起建立"AI Verify"基金会，专门推进 AI 治理研究与工具开发工作。

2024 年 1 月，基金会在讨论文件《生成式人工智能：对信任和治理的影响》的基础上，整合多方反馈意见，发布**《生成式人工智能治理的模型框架草案》**，并于 5 月正式发布**《生成式人工智能治理模型框架》**，细化 9 个治理维度内容，以期打造更可信的人工智能环境，为全球人工智能治理提供参考。

Fostering a Trusted AI Ecosystem

1. Accountability
Putting in place the right incentive structure for different players in the AI system development life cycle to be responsible to end-users

2. Data
Ensuring data quality and addressing potentially contentious training data in a pragmatic way, as data is core to model development

3. Trusted Development and Deployment
Enhancing transparency around baseline safety and hygiene measures based on industry best practices in development, evaluation and disclosure

4. Incident Reporting
Implementing an incident management system for timely notification, remediation and continuous improvements, as no AI systems is foolproof

5. Testing and Assurance
Providing external validation and added trust though third-party testing, and developing common AI testing standards for consistency

6. Security
Addressing new threat vectors that arise through generative AI models

7. Content Provenance
Transparency about where content comes from as useful signals for end-users

8. Safety and Alignment R&D
Accelerating R&D through global cooperation among AI safety Institutes to improve model alignment with human intention and values

9. AI for Public Good
Responsible AI includes harnessing AI to benefit the public by democratising access, improving public sector adoption, upskilling workers and developing AI systems sustainably

图源：新加坡《生成式人工智能治理模型框架》

企业自我监管

依据《生成式人工智能模型管理框架》，企业在开发和应用生成式人工智能技术时，需建立全面的内容管理和审核体系。

政府监管与执法

新加坡政府相关部门依据《人工智能治理模型框架》和《组织实施和自我评估指南》等文件，积极履行监管职责，建立起全方位、多层次的监管体系。**政府部门制定了详细的生成式人工智能产品和服务的审核标准与流程**，要求企业在产品上市前提交内容审核报告、风险评估报告和合规承诺书等材料。同时，政府定期组织专业人员对市场上的生成式人工智能应用进行抽查和评估，重点检查内容的合法性、公正性、道德性以及准确性，及时发现并处理存在内容实质错误与 AI 幻觉问题的产品和服务。

其他互联网监管政策

新加坡的互联网立法体系建制起步较早，是世界上较早通过立法规范互联网使用的国家之一。在 AI 监管体系之前，新加坡已经制定了相当完善的互联网监管法规体系，并取得了良好的执行效果。现行的适用于管制互联网内容的法规包括:《互联网操作规则》、《防止网络假信息和网络操纵法案》、《维护宗教融合法》、《行业内容操作准则》等。

《互联网运行准则》

《互联网运行准则》是新加坡管制互联网内容最主要的法律，其规定网络上禁止出现与公共利益、公共道德、公共秩序、公共安全、国家和谐相违背以及任何违反现行新加坡法律的内容。

除了以上绝对禁止的内容以外，如果含有某些因素，该内容也有可能被判定为违规内容：

1. *以故意挑逗的方式展示裸体或性器官；*
2. *性暴力或强迫 / 非自愿性行为、16 岁以下未成年性行为；*
3. *鼓励同性恋，描绘或鼓励乱伦、恋童癖、恋兽癖、恋尸癖等；*
4. *大肆渲染极端暴力、残酷暴行；*
5. *在种族和宗教之间制造仇恨。*

值得注意的是，为了减轻互联网内容提供商的内容审核负担，《准则》也规定了一些豁免情况：

1. *对基于网络服务而建立的私人讨论区（如聊天室），只要保证不设定"禁止内容"范围内的话题，即可免责；*
2. *对基于网络服务而建立的公共展示区（如 BBS 系统），只要在日常编辑、检查的过程中关闭了含有"禁止内容"的链接，即可免责；*
3. *网络内容提供商必须按照 MDA 的要求，关闭含有"禁止内容"的网页。*

《防止网络虚假信息和网络操纵法案》

新加坡在 2019 年通过了《防止网络虚假信息和网络操纵法案》，其核心目的在于"防止虚假谎言在新加坡的传播与交流，禁止资助、推广和支持虚假谎言的在线传播渠道，监测、控制和防止虚假账户和自动程序的滥用，监管具有政治目的的付费内容，突出可靠的信息来源。"

该法案也延续了新加坡政府一向的监管思路，重点管控两类特殊内容，一类是可能影响选举、有特殊政治目的的内容，另一类则是危害国家安全、种族宗教和谐的内容。

《维护宗教融合法》

新加坡是一个多种族、多宗教信仰的国家，种族宗教和谐对新加坡的社会稳定而言至关重要。1990 年制定并实施至今的《维护宗教融合法》是新加坡最主要的宗教管理法律，它明确了新加坡的多元宗教政策，通过立法维持宗教和谐。

《维护宗教融合法》规定，禁止个人煽动不同宗教团体之间的敌意、仇恨，不得侮辱宗教或伤害他人的宗教情感。2019 年，新加坡政府还推出了修正案，将网络空间也纳入宗教言论的管理范围。

《刑法》——"种族和谐"

新加坡是一个多种族的国家，华人占 74% 左右，其余为马来人、印度人和其他种族。新加坡独立至今共发生两次大规模种族暴动，对于种族冲突、歧视问题尤其敏感，1997 年开始更是将 7 月 21 日定为"种族和谐日"。

新加坡将维护宗教和谐写入了刑法，严厉打击煽动种族仇恨的行为。刑法规定，禁止故意伤害他人的种族情感、煽动种族仇恨、破坏种族和谐和社会安宁。

《远程赌博法》

随着网络赌博逐渐兴起，为防止网络赌博成为非法活动的资金渠道，新加坡开始对网络赌博进行管控。2014 年，新加坡颁布了《远程赌博法》，其第 8 条规定，"禁止任何形式的以互联网、无线电、电话或者可以与之通信的电子设备所进行的在线赌博，如果违反此法律，将面临被判入狱 6 个月或罚款 5000 美元。"

AI 发展及相关监管政策

越南于 2021 年 1 月发布《越南至 2030 年人工智能研究、发展与应用国家战略》，提出到 2030 年将越南打造为东盟领先的人工智能创新中心，并推出了《数字技术产业法（草案）》《个人数据保护法（草案）》《关于加强半导体芯片、人工智能和云计算领域高素质人力资源培训政府令》等一揽子政策，为 AI 发展明确了路径。

2024 年 7 月 2 日，越南发布了《数字技术产业法》（征求意见稿）， 越南国会于 2025 年 6 月 14 日通过了《数字技术产业法》，将于 2026 年 1 月 1 日正式生效。该法涵盖了数字技术产业活动、数字技术产品和服务在内的数字技术产业。《数字技术产业法》（征求意见稿）中，越南第一次在法律法规中 以专章方式明确了人工智能发展的合规框架。 而数字技术产品和服务包括信息技术产品和新技术产品， 但不限于人工智能、大数据、云计算、物联网、区块链、虚拟现实 / 增强现实， 以数字化现实、收集 、存储、传输 、处理信息和数字数据。

越南虽然目前尚未发现具体的监管案例，但从上述法律法规和指导方针中可以看出越南政府对于人工智能产品内容安全的高度关注。越南政府希望通过建立相应的监管框架来确保人工智能技术的健康发展，并平衡商业利益与消费者权益。

其他互联网监管政策

近年来，从行政法令到法律再到行为守则，越南的互联网内容监管法律法规不断完善，但就违规内容来说并没有太大的变化，反政府内容一直是越南内容审查的重中之重。

刑法“反国家宣传罪”

在越南，批评政府及领导人的言论是被严令禁止的。

越南刑法第 117 条规定：禁止“制作、储存、散发、宣传含有煽动宣传反对越南社会主义共和国内容的信息和资料”，这一条也被称为“反国家宣传罪”。违者将面临 5 至 20 年的有期徒刑，实施未遂也会被处以 1 至 5 年的监禁处罚。

越南刑法第 331 条规定，禁止“滥用民主自由权侵犯国家利益和组织、个人的合法权益”。这两条法律经常被用于起诉在社交媒体上发表反政府言论的个人。

2020 年 6 月，河内市公安局对涉嫌煽动反对国家罪的两名被告人进行起诉。据越通社报道，两名被告在社交网络上直播，造谣、诽谤政府和越南共产党。此外，他们还两次焚毁国旗，焚毁党旗，并使用剪刀剪裁并焚烧胡志明主席的图片等。

72 号法令

2013 年起施行的 72 号法令是越南监管互联网服务的主要立法文件，对从事社交网络与网络游戏的企业来说尤为重要，监管机构多引用 72 号令要求平台配合删除违规内容。72 号令规定禁止以下行为：

> *任何利用互联网信息、互联网使用权和网上信息等进行破坏越南政府，危害越南国家安全、社会秩序，破坏民族大团结，鼓励发动战争与恐怖主义行动，煽动民族仇恨、引发民族或宗教矛盾，宣传激发暴力行为，散布淫秽、色情、犯罪、社会丑恶现象，传播迷信、破坏社会美德，泄露国家机密，散布谣言，损害组织信誉及个人荣誉与品德，广告宣传违法商品，散布被禁止的文学艺术作品，冒充组织或个人散布谎言信息，侵害组织和个人合法权益等行为。*

除了规定禁止内容，72 号令还将网络用户在社交网站、微博以及博客页面上能够发布的内容限定在 “个人信息”上，并且要求在越南营运的外国互联网公司将服务器部署在越南境内。2021 年，72 号令进行了新一轮修订，对网络游戏、直播等事项进行了更新。截至目前，修正案尚在审议中。

《网络安全法》

2018 年 6 月，越南在十四届国会第五次会议上以 86% 的高票通过《网络安全法》，并于 2019 年 1 月 1 日起正式实施。该法以国家安全与网络安全为出发点，加强了政府对互联网服务提供商多项运营环节的控制。首先，《网络安全法》对网络上的违规内容作出详细规定，禁止以下行为:

1. *使用网络空间从事煽动颠覆国家政权、欺骗、煽动分裂越南社会主义共和国。*
2. *歪曲历史、否认革命成就、破坏国家统一、宗教歧视、种族歧视、区别对待。*
3. *编造、传播虚假信息扰乱经济秩序和社会秩序。*
4. *卖淫、社会弊端、贩卖人口；淫秽色情信息等。*
5. *促使他人去实施犯罪活动。*
6. *从事网络攻击、网络恐怖袭击、网络间谍、网络犯罪活动等。*
7. *利用保护网络安全活动以侵犯国家主权、利益、扰乱国家安全和社会治安秩序，侵害组织和个人的合法权益或以获得利益。*

根据该法，互联网服务提供商有义务配合监管部门调查请求，**应在收到通知 24 小时内撤下平台上的违禁内容**。其次，《网络安全法》也赋予**越南政府对企业及用户数据更大的管辖权**，其规定在越南境内提供互联网相关服务的国内外企业，需将用户信息数据存储库设在越南境内，相关外国企业需在越南设立办事处。**在越南境内提供互联网相关服务的国内外企业需验证用户注册信息，并应公安部门调查要求提供相关信息**。《网络安全法》还对国家关键信息系统、网络攻击、儿童保护等议题进行了细致规定。

《社交媒体行为守则》

2021 年 6 月，越南信息和通信部发布《社交媒体行为守则》，鼓励个人和组织分享官方或有可靠来源的信息；弘扬符合越南民族道德、文化和传统价值观的行为。《社交网络行为规范》要求组织和个人不使用煽动仇恨、鼓吹暴力、性别或宗教歧视的言论；不发布违法内容、侵犯他人名誉和人格尊严的信息；不使用违反优良风俗习惯的语言或散布虚假信息、歪曲事实的新闻；不发布影响社会治安秩序的违法广告和经营等。但目前尚不清楚这套社媒行为守则具有多大的法律约束力或将如何具体执行。

《个人数据保护法》

2024 年 2 月，越南公共安全部（MPS）启动《个人数据保护法》（PDP 法）制定工作，旨在强化国内数据隐私监管，并推动数字化转型。2025 年 6 月 26 日，越南国会正式通过该法，取代此前作为临时依据的第 13/2023/ND-CP 号法令。新法适用于在越南境内处理个人数据的实体，以及处理越南居民数据的境外机构，无论数据处理行为发生地是否在越南。

泰国

AI 发展及相关监管政策

泰国政府于 2025 年成立国家人工智能委员会，计划两年内投资 154 亿美元用于人工智能基础设施建设，并设立 9 个人工智能卓越中心。数字经济与社会部数据显示，泰国已投入超 60 亿泰铢用于人工智能专项人才培养，并建立 ThaiLLM（泰语开源人工智能基础设施）等国家级项目

泰国目前暂无 AI 和机器学习专项法律，但有意进行监管，据泰国媒体报道泰国近期公布该国首部人工智能法案并公开征求公众意见，目标是建立更清晰透明的法律框架应对高风险人工智能应用程序。

泰国政府在制定相关法律法规和指导方针时，特别强调了生成式人工智能产品的内容安全问题。具体来说：

《AI 伦理原则及指引》将透明性和责任作为 AI 伦理原则之一 ，要求人工智能的研究、设计、开发、服务和应用应该保持透明，能够解释和预测，包括可以追溯各种活动 。《关于使用人工智能系统的商业运营的皇家法令草案》提出，高风险的人工智能系统服务的服务开发者除履行登 记要求外 ，还有责任确保其系统在整个服务期间内具备且遵守适当的风险管理措施，并向公众发布

相关信息 。《泰国促进和支持人工智能创新法案草案》 以及两项子法规草案征求公众意见， 旨在通过提供人工智能监管沙箱 、要求部分最低限度的内容需要向公众披露 ，以确保透明度并防止不公平的歧视性合同条款 ，促进人工智能的发展。

其他互联网监管政策

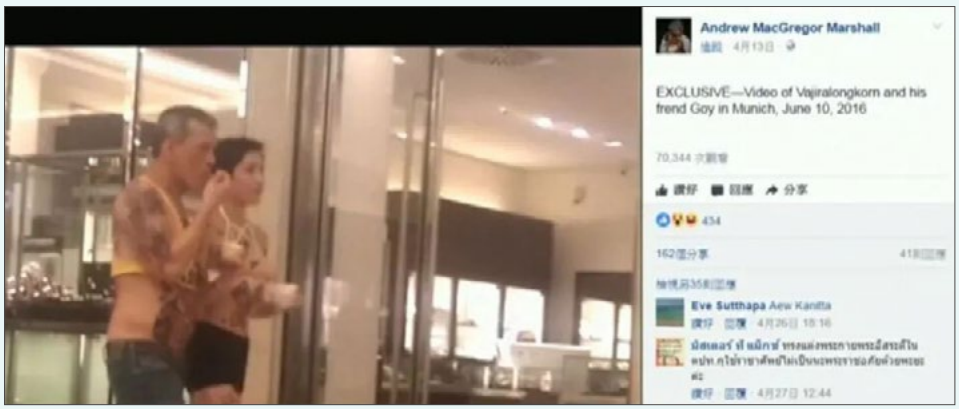
立法与行政法令是泰国政府最主要的两种互联网监管手段。立法虽然时有更新，但相对来说较为固定，在执法中最常援引的当属刑法"冒犯君主罪"、《计算机犯罪法》以及赌博法。行政法令则是监管部门在日常执法中针对某一突出网络现象、问题而制定的应急解决方案，较多针对社交平台、网站上的社会、政治、历史、文化话题。行政法令的出台往往有着较大的不确定性，往往随着某一社会特点话题而来，以平台撤下内容而终。下面我们将主要介绍泰国的互联网监管立法。

泰国刑法 112 条"冒犯君主罪"

在泰国，冒犯王室是一项严重的罪行。泰国刑法第 112 条规定，任何人诽谤、侮辱或威胁国王、王后、王储或摄政王，最高可被判处 15 年徒刑。

除了逮捕发表冒犯君主言论的人，社交媒体平台也因其强大的传播能力成为泰国监管机构的重点关注对象。

- *2020 年，泰国政府向 Facebook 发出警告，扬言要对其采取法律行动，原因是 Facebook 没有撤下被认为是诽谤君主制的内容。*
- *因传播侮辱国王的视频，泰国曾多次关闭视频网站 YouTube。*



NBTC 要求 Facebook 删除皇室不雅视频 来源：Now News

《计算机犯罪法》

2006 年，泰国军方趁总理他信出国访问之际发起政变，全面接管国家政权。此后，他信的支持者及反军政府势力频繁借助互联网和广播发表批评言论、开展宣传以及组织示威活动。为巩固政权，军政府在 2007 年快速通过了《计算机犯罪法》以管控网络反动言论。 自法案实施以来，泰国当局已频繁援引相关条款开展逮捕行动、封禁违法网页。据路透社报道，在 2007 至 2009 年间，泰国当局屏蔽了将近 2 万个被认为有辱国王尊严的网页。其中引用最多的为第 14、15 以及 16 条。

1. *《计算机犯罪法》第 14 条规定，在网络上传播歪曲或虚假信息损害公共利益，传播虚假信息损害国家安全、公众安全、经济安全、公共基础设施，传播淫秽、危害国家安全、暴恐内容，均为违法行为，最高可判处五年有期徒刑，或最高 10 万铢罚款，或两者并罚。*
2. *第 15 条规定，供应商如被发现故意为这类违法行为提供支持，也将面临同样的刑事处罚。*
3. *第 16 条规定，在网络上传播、修改、编辑他人肖像并对他人造成损失的，最高可面临三年有期徒刑或 20 万铢罚款。*

《1935 赌博法》

泰国是一个全面禁赌的国家，禁止赌博的力度堪称位居全球前列。泰国《1935 赌博法》规定，除国家发行的彩票和赛马以外，禁止其他一切形式的赌博活动，就连在公共场所玩扑克牌都可能被警察抓起来。

近年来泰国也开始倡导赌博合法化，但前总理佩通坦宣布赌博合法化仅仅过去半年时间，新任总理阿努廷立刻就给推翻了。

数据安全与隐私保护立法

泰国的数据安全与隐私保护立法还处于初始阶段。2019 年，泰国通过了《个人数据保护法》，这也是泰国国内管辖数据保护的统一法案。虽然《个人数据保护法》早已通过，但由于多方原因，

法案要到 2022 年 6 月才会被全面执行。就个人数据保护而言，《个人数据保护法》提出了多项原则，但具体措施则有待进一步规定。

2019 年泰国政府还通过了《网络安全法》以预防应对网络威胁。

马来西亚

AI 发展及相关监管政策

2024 年 12 月，马来西亚成立国家人工智能政策法规办公室，负责制定人工智能相关的法律法规，并评估技术安全性和可靠性。该办公室还承担教育发展计划制定，以培养人工智能领域专业人才。

同时，马来西亚政府**高度重视生成式人工智能产品的内容安全问题**，并通过一系列法律法规来确保内容的安全性和合规性。目前，虽然马来西亚尚未出台专门针对人工智能治理的立法，但已有的数据保护法律和其他相关政策为人工智能的内容安全提供了基础框架。

《2021-2025 年国家人工智能路线图》明确了负责任 AI 的七项原则，其中包括**透明性、公平性、可靠性和控制、隐私和安全等**，这些都是内容安全的重要组成部分。路线图强调了人工智能系统必须遵守数据收集、使用和存储的隐私法律，以确保个人数据得到妥善保护。这有助于确保生成式人工智能产品在内容生成过程中遵守隐私保护的要求。科技创新部 (MOSTI) 正在制定的人工智能法案预计将解决与人工智能相关的透明度和问责制等问题，进一步加强对内容安全的监管。

2018 年通过的《反假新闻法令》规定，制造、发布或传播虚假新闻是违法行为。生成式人工智能产品在内容生成和传播过程中，应避免制造和传播虚假信息，以免触犯该法。马来西亚政府在制定相关法律法规时，**强调了透明度和责任制**，要求生成式人工智能产品开发者和运营者必须公开其系统的运作和数据处理机制，并对其生成的内容负责。包括：

告知义务

人工智能系统在与用户互动时，必须明确告知用户他们正在与人工智能系统互动，确保用户知情权。当使用人工智能系统生成或修改图像、声音或视频时，可能导致误认为是真实或原物的，开发者有义务向公众披露这些内容是人工合成或被修改过的。

内容责任

生成式人工智能产品开发者需对其生成和传播的内容负责，确保内容不包含虚假信息、不侵犯他人隐私、不具煽动性或威胁性。

其他互联网监管政策

目前，马来西亚已有多部法律法规将互联网监管纳入其中。

《1998 通信与多媒体法令》

《1998 通信与多媒体法令》CMC 是马来西亚最重要的互联网内容监管法律，它整合了原有的《1950 年电讯法令》、《1985 年电讯服务法令》及《1988 年广播法令》，被誉为马来西亚在信息技术立法工作方面最全面的一部，从**经济、技术、社会**三方面对**网络设备、网络服务、应用程序服务和内容应用程序服务提供商**的活动作了调整。

具体到互联网内容监管上，CMA 主要通过设立执照门槛以及内容限制两种手段来对互联网内容进行限制。在执照发放方面，《1998 年通讯与多媒体法令》第 126 条款规定，任何人未取得执照，皆不得从事、拥有或提供网络设备、提供网络服务、或提供应用服务等行为。假使获得执照，业者必须遵守执照上订明之条件 （第 127 条款）。

CMA 禁止未取得执照提供内容应用服务，网络设备服务提供者、网络服务者或应用服务者、内容应用服务提供者若取得之执照为类别执照，须先向 MCMC 注册即可营业。在内容监管上，

CMA 最著名的 211 条款规定，“禁止令人反感的内容”，内容应用服务提供者及其他使用内容应用服务的人不能提供有**伤风化的、猥亵的、不正确的、危险的内容**，或是具有攻击性意图造成使任何人困扰、咒骂、威胁或骚扰等。触犯上述条款的最高刑罚为罚款五万令吉或监禁一年，或两者兼施；若在定罪后还继续犯行，可以每日一千令吉的罚款处罚之。

另一用来检举网民的条款是第 233 条款的“不当使用网络设备或网络服务”，规定禁止使用网络设备或网络服务来传播不良内容。第 233（3）条款阐明，触犯此罪行者，最高刑罚为罚款五万令吉或监禁一年，或两者兼施；若在定罪后还继续犯行，可以每日一千令吉的罚款处罚之。

除了 CMA 这部专门针对互联网领域的法律，任何马来西亚线下被视为违法的行为，在线上仍然适用。

《1948 年煽动叛乱法》

《1948 年煽动叛乱法》将针对**政府、王室、宗教、种族具有“煽动性”倾向的言论定为犯罪，包括“煽动仇恨或蔑视”或激起对政府的不满或引起“不同种族之间的恶意和敌对情绪”的言论**。被裁定犯有煽动罪的人可能会判刑入狱三年，或五千令吉罚款，或两者并罚。自 2013 年以来，至少有 14 人因“煽动叛乱罪”被起诉，其中多为政治家、活动人士、记者。 由于《1948 年煽动叛乱法》对于“煽动叛乱”的定义过于宽泛，会随执政当局的标准而变化，政府可轻易滥用该法来打压异议人士，因此近年来也受到马来西亚国内外的抵制。

《1953 住家公开赌博法》

来源：网络

伊斯兰教为马来西亚的联邦官方宗教，全国有超过 60% 的人口信奉伊斯兰教。根据《马来西亚联邦宪法》规定，马来西亚法律制度实行双司法系统制度，即适用于全国范围的英美法系（普通法）和仅适用于穆斯林生活事务的伊斯兰教法并行。在这两种法律体系中，赌博都是被严格禁止的。

《1953 住家公开赌博法》（Common Gaming Houses Act 1953，CGHA）是政府打击赌博犯罪的主要依据。马来西亚皇家警察是负责打击赌博犯罪的主要执法力量。为遏制赌博犯罪蔓延，马来西亚皇家警察会与 MCMC 合作查处赌博网站。**警方将疑似涉赌网站名单提交至 MCMC，MCMC 在对有关网站进行审核后，对确属违法的涉赌网站域名进行封锁。**

《刑法》禁止同性恋

作为一个以穆斯林居多的国家，受伊斯兰教的影响，**马来西亚的联邦法律与宗教法律均不认可同性恋权利**。《马来西亚刑法》第 377A 条文中禁止违反自然性行为，第 377B 条文则列明男性违反者将会被处以二至二十年徒刑和鞭刑。伊斯兰教法由各州伊斯兰宗教局负责执法，马来西亚 13 个州的伊斯兰教法全都禁止穆斯林"男性"扮成"女性"。因此，涉及 LGBT+ 话题的网站也是 MCMC 的封禁目标之一。

《2010 个人数据保护法案》

马来西亚在 2010 年颁布了**《个人数据保护法》**，该部法律主要规制在商业行为中个人数据的收集和使用行为，并确立了个人数据保护的一般原则。

为了应对数据泄露和网络攻击频发的现状，**2024 年马来西亚实施了《2024 年个人数据保护（修订）法案》**。该法案引入了一系列重大变革，对马来西亚的企业提出了更为严格的合规要求。其中一项显著的改变是要求企业任命数据保护官（DPO），以监督数据保护策略并确保法规的遵守。此外，数据处理者的责任得以明确，跨境数据传输规则也进行了修订，增加了**强制性的数据泄露通知制度**，而对于违规行为的惩罚也加重了。

对于出海企业来说，PDPA 的数据跨境传输规定需要注意。PDPA 第 129 条规定，**禁止将马来西亚境内存储的个人数据传输至境外**。不过，在特定情况下，企业仍可将数据传输至海外，如征得用户同意、取得政府部门批准等。

印度尼西亚

AI 发展及相关监管政策

印尼政府对人工智能的顶层设计呈现出鲜明的**"国家能力主导"**特征。2025 年 8 月发布的《国家人工智能战略》(STRANAS KA) 并非简单的技术路线图，而是**与"2045 黄金印尼"国家愿景深度绑定的发展纲领**。该战略将 AI 定位为实现经济跨越式发展的核心杠杆，设定了三阶段发展路径：**2020-2024 年建立制度基础，2025-2035 年推动广泛应用，2035-2045 年实现全面集成**。这种长周期、分阶段的规划思路，反映出印尼政府避免技术冒进、注重实效的务实态度。

政策架构的四大支柱构成了印尼 AI 生态的基石。在伦理治理方面，通信与信息技术部 (KOMINFO) 发布的《AI 伦理通函》(9/2023) 确立了九大原则，强调**包容性、人本性和透明度，要求高风险 AI 模型必须提供印尼语可解释性文档**。

不过目前，印尼仍然缺乏针对公共和私营部门开发和使用人工智能的明确监管框架。印尼仍然依赖现有立法来**解决与新兴人工智能模型的开发和使用相关的问题**。

其他互联网监管政策

《反色情法》

2008 年，印尼通过了《反色情法》(Anti-Pornography Law)，这一法案一直以来受到较大争议，但获得很多伊斯兰教徒的支持。《反色情法》禁止公民生产、制作、复制、分发、广播、进口、出口、

提供、买卖、租借色情内容。生产、制作、复制、分发、广播淫秽内容的最高可判处 12 年有期徒刑，最高可处罚款 66 万美元；下载淫秽作品的最高会被判处 4 年有期徒刑，22 万美元罚金。

印尼穆斯林游行支持《反色情法》 *来源：Reuters*

《反色情法》将含有性行为、性暴力、性器官、裸体的敏感内容都归为“pornography”，但对尺度的把握并没有明确定义。考虑到印尼将近 87% 的人口为穆斯林，印尼政府对色情内容的管控可能会更为严格。

《部颁条例 5》

2020 年 11 月，印度尼西亚开始施行《部颁条例 5》(Ministerial Regulations 5,MR 5)，进一步加强政府对网络内容与用户数据的控制。

MR5 的监管对象涵盖很广，包括社交媒体及其他类型的内容分享平台、网络购物、搜索引擎、金融服务、数据处理服务、即时讯息、视频、游戏等。

MR5 要求私人电子系统运营商应向有关部门提交注册信息，取得牌照后方能在印尼开展业务，并要求企业在印尼当地派驻主管人员。运营商还应向监管机构提供系统数据接口，以供监管机构可以开展监控等执法行动。如果企业拒绝提供接口，将会面临警告、暂时封禁、完全封禁、撤销牌照等处罚。

MR5 第 13 条要求运营商撤销其平台上的违禁信息或违禁文件。此处的违禁信息泛指违反印度

尼西亚法律法规的任何规定，或会带来"社区焦虑"、"扰乱公共秩序"的信息或内容。运营商应确保他们的平台、服务、网站不会被用于助长违法信息的传播。这一条款要求运营商开展审核工作，对平台内容进行过滤。

《个人数据保护法》

2022 年 10 月 17 日，《印度尼西亚个人数据保护法》（以下简称《个人数据保护法》）正式生效，该法关于个人数据跨境传输的规定与印度尼西亚其他相关法律共同构成了印度尼西亚个人数据跨境传输规则。

根据 PDP 法的规定，个人数据被定义为通过电子或非电子手段，直接或间接与特定个人相关的信息。印尼法律明确界定了个人数据的定义，并为数据主体赋予了多项权利，如知情权、更正权、访问权等。这还包括反对基于完全自动化的决策和数据可移植的权利。

AI 发展及相关监管政策

菲律宾已经认识到人工智能的潜力及其在推动经济增长和创新方面的重要性，政府 、私营部门和学术机构正在努力推动人工智能的研究、开发 和应用。 根据 Statista 的预测，菲律宾的**AI 市场在 2025 年将达到 10.25 亿美元，预计到 2030 年将达到 34.87 亿美元，年均增长率约 28%。**

菲律宾目前还没有人工智能的强制性标准或国家标准，也没有专门针对人工智能的法律法规。菲律宾正在寻求通过立法，成立" 人工智能发展管理局 "，**负责制定国家 AI 战略和框架， 指导企业在菲律宾开发和部署 AI 技术** 。

2021 年发布《国家 AI 战略路线图》（NAISR），2024 年更新纳入生成式 AI 创新，但缺乏配套法律支撑。《菲律宾发展计划（2023-2028）》将 AI 列为**核心增长引擎**，强调人才培训与基础设施投入。

其他互联网监管政策

菲律宾在互联网内容监管上还处于早期阶段，相关法律法规仍不完善。任何刑法所规定的线下违法行为，在网络上仍然受法律监管。

《反儿童色情法》

在联合国儿童基金会及其他组织的呼吁下，2009 年菲律宾政府通过了《反儿童色情法》，儿童性虐待在菲律宾第一次被定为违法犯罪行为。《反儿童色情法》规定，任何形式的网络儿童性剥削均为违法行为。ISP 以及互联网内容平台应在发现儿童性剥削材料的七天内通知菲律宾国家警察及国家调查局，且有义务积极配合相关监管部门的调查请求。ISP 还应配有相应的技术手段，以确保可对儿童性剥削内容进行封禁或过滤，违反可面临 200-300 万比索罚款。

《网络犯罪预防法》

2012 年通过的《网络犯罪预防法》是菲律宾最主要的互联网内容监管法律，其第 4（c）节规定，网络性行为、儿童色情、违规商业传媒、诽谤均为违法行为，违者将面临监禁或最高 25 万比索罚款。《网络犯罪预防法》规范的是个人层面上的网络违法行为，平台责任尚未被纳入管辖。

《反恐怖主义法》

菲律宾深受恐怖主义侵害。2020 年，杜特尔特总统签署颁布《反恐怖主义法》，以强化打击恐怖主义力度。相比之前的《人类安全法》，新反恐法拓宽了"恐怖主义"的定义边界，处罚措施也更为严厉。该法规定，除实施恐怖主义行为属违法之外，任何提议、煽动、密谋及参与恐怖主义计划、训练、准备和宣传工作，以及向恐怖分子提供物质支持或协助招募恐怖组织成员也属违法行为。《反恐怖主义法》并未将平台的审核义务纳入其中，但也有消息显示，菲军方在未来有可能将社交媒体纳入反恐范围。

东南亚内容合规总结

国家	监管政策	合规重点
新加坡	《反歧视法》 《生成式人工智能模型管理框架》	价值观导向问题，如歧视与偏见 社会道德与伦理问题、 内容实质错误与AI幻觉问题
越南	《人工智能开发指南》 《数字技术产业法》（草案） 《人工智能生命周期国家标准》（草案）	人类价值和社会道德、 智能生成物的可识别性
泰国	《AI伦理原则及指引》 《泰国促进和支持人工智能创新法案草案》 《关于使用人工智能系统的商业运营的皇家法令草案》	透明性和责任归属
马来西亚	《2021-2025年国家人工智能路线图》 《反假新闻法令》	透明度和责任制， 包括对用户的告知义务以及要求 对生成及传播的内容负责
印度尼西亚	《信息与电子交易法》《人工智能伦理准则》 《网络与信息空间安全法》《关于人工智能道德指南的通知第9号》 《关于金融科技行业中负责任和可信赖的人工智能的道德指南》	道德和伦理规范、种族主义等
菲律宾	《网络安全法》《数据隐私法案》《人工智能伦理准则》 《反网络性虐待或性剥削儿童和反儿童性虐待或性剥削材料法案》	色情内容、诽谤、网络欺诈 未成年人相关内容

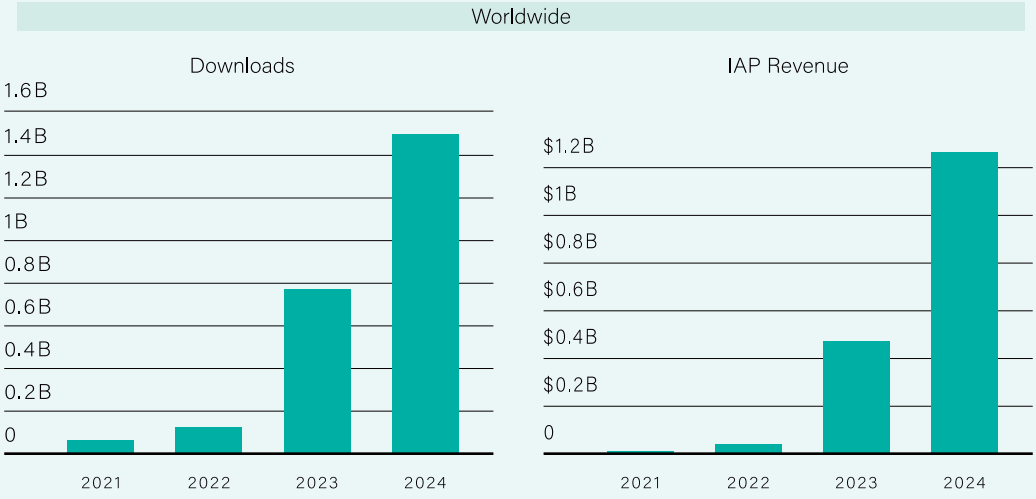
北美：言论自由与社区标准的平衡

美国

据 Sensor Tower 数据显示，截至 2024 年，美国是 AI 应用的绝对头部市场，占全球收入的 45%。美国市场拥有良好的用户付费意愿、成熟的支付习惯，使其成为中国 AI 企业出海战略中不容忽视的一站。

Yearly Trends for Generative AI Apps

The United States is the clear top market for Generative AI apps in 2024, according for 45% if global revenue.



来源：State of Mobile 2025

然而，同其他行业一样，文化、制度与法律环境的巨大差异，意味着曾在国内行之有效的商业逻辑，到了美国可能不得不进行"重构"。从**知识产权、数据隐私、内容合规**等，每一个环节都可能影响 AI 企业在美国市场的长期可持续发展。因此，若想在美国扎根，第一步就必须打下合规的地基。

AI 发展及相关监管政策

美国形成了联邦与州相协同的双轨规范架构，通过分散式立法、行政指引与司法判例共同构建动态人工智能法律治理体系。

联邦层面法律与政策

- 2021 年 1 月，美国正式颁布了《2020 年国家人工智能倡议法案》。该法案从国家层面推动可信人工智能系统在公共和私营部门的开发与应用，旨在进一步巩固美国在人工智能研发领域的地位。
- 2022 年，美国政府发布《人工智能权利法案蓝图》，以政策文件的形式确立了人工智能技术应用的五项原则，为公私主体划设技术伦理边界，确保人工智能技术的开发和应用符合社会需求，并保护公民权利。
- 2025 年，美国总统特朗普签署《消除美国人工智能领导地位的障碍》行政令，优先考虑以监管最小化路径巩固美国在全球人工智能技术竞争中的优势，解除对私营部门开发和部署的多重限制，以加速人工智能技术的创新突破，但此举可能削弱对技术的伦理约束与风险防控效能。
- 2025 年 7 月 23 日，白宫发布了长达 28 页的《赢得 AI 竞赛：美国 AI 行动计划》（Winning the AI Race: U.S. AI Action Plan）（以下简称《行动计划》），标志着美国总统特朗普 AI 领域总体规划思路逐渐成型，也标志着美国国家人工智能战略的重大转向。

特朗普在"赢得人工智能竞赛"峰会上发表讲话　　*图源：路透社*

在此行动计划中，与 AI 治理相关核心内容：通过最大程度地扫除监管障碍，为美国私营部门的 AI 创新"松绑"和"加油"；其次，确保 AI 技术的发展方向和内容产出，符合本届政府的价值观和政治议程，即所谓的"美国价值观" 。为实现这一双重目标，计划部署了三套强有力的政策工具。

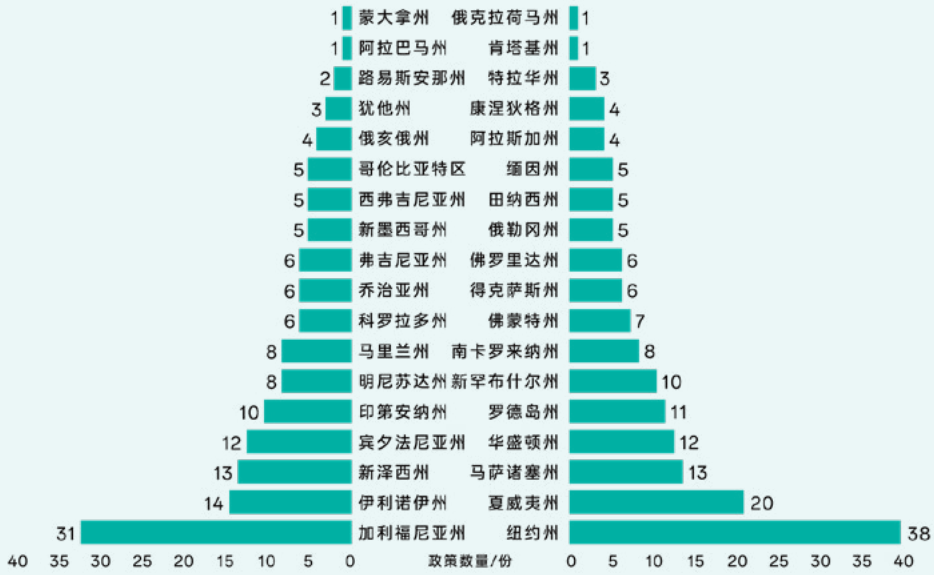
1. ***"监管大扫除"***：*这是最直接的"松绑"措施。由 OMB 统筹，要求联邦机构 60 天内审查并修订或废除阻碍 AI 开发部署的法规; 白宫科技政策办公室（OSTP）向产业界征询信息，设快速通道让企业直报监管痛点，实现精准解障。*
2. ***风险框架"瘦身"***：*计划明确指示国家标准与技术研究院（NIST）修订《AI 风险管理框架》，移除或淡化错误信息、多样性公平性包容性（DEI）、气候变化等社会伦理考量，将风险管理焦点收窄至技术安全与性能风险，回应行业对合规成本及 "政治正确" 限制创新的不满。*
3. ***联邦采购的"思想门槛"***：*通过行政令更新采购指南，要求政府部门采购的前沿大语言模型需 "客观且无意识形态偏见"，以 "真实性" 和 "意识形态中立" 为原则，明确规避 DEI 等 "意识形态教条"，实质是借联邦采购力量对 AI 模型进行意识形态审查。*

这反映了美国 AI 治理思路的根本性转变。它从拜登政府时期强调"过程合规"（如进行风险评

估、影响报告）的模式，转向了更注重"结果控制"的模式。一方面通过去监管简化过程，另一方面通过采购标准直接控制最终 AI 产品的"思想"产出。这种从"如何做 AI "到" AI 应该说什么"的转变，代表了一种更直接、更具干预性的治理思路。

美国州级 AI 立法实践

截至 2024 年，在美国 50 个行政区中，以加利福尼亚州、纽约州和夏威夷州为代表的 30 个州正式通过了有关人工智能监管的政策。



美国各州 AI 监管政策数量图　图源：《科学学研究》论文

在州级立法实践中，犹他州与科罗拉多州率先实现突破。2024 年颁布的《犹他州人工智能政策法案》规定了使用生成式人工智能的披露制度，通过人工智能政策办公室统筹制定标准，并设立技术沙盒实验室推动公私协同研发。

同年颁布的《科罗拉多州人工智能法案》则聚焦高风险人工智能系统的算法歧视治理，明确将就业、信贷、教育等关键决策场景纳入监管范畴，在涉及与人工智能系统交互的消费者保护领域建立包含影响评估、数据溯源、算法解释的透明度义务框架。至此，人工智能州级法律治理呈现出三大转型态势：从软性指南转向硬性约束、从普适原则转向场景规制、从分散治理转向机构化运作。

美国的双轨法律治理框架强调"弱监管"原则下的行业自律，各州监管力度与重点差异显著——既通过弹性监管为本土企业创造制度红利，又借助州际竞争筛选最优治理方案，但立法权分散也引发了标准互认的困境与跨国规制冲突。如何在激发地方创新动能的同时把控整体风险，逐渐成为美国人工智能法律治理体系优化的核心议题。

美国具体内容监管要求

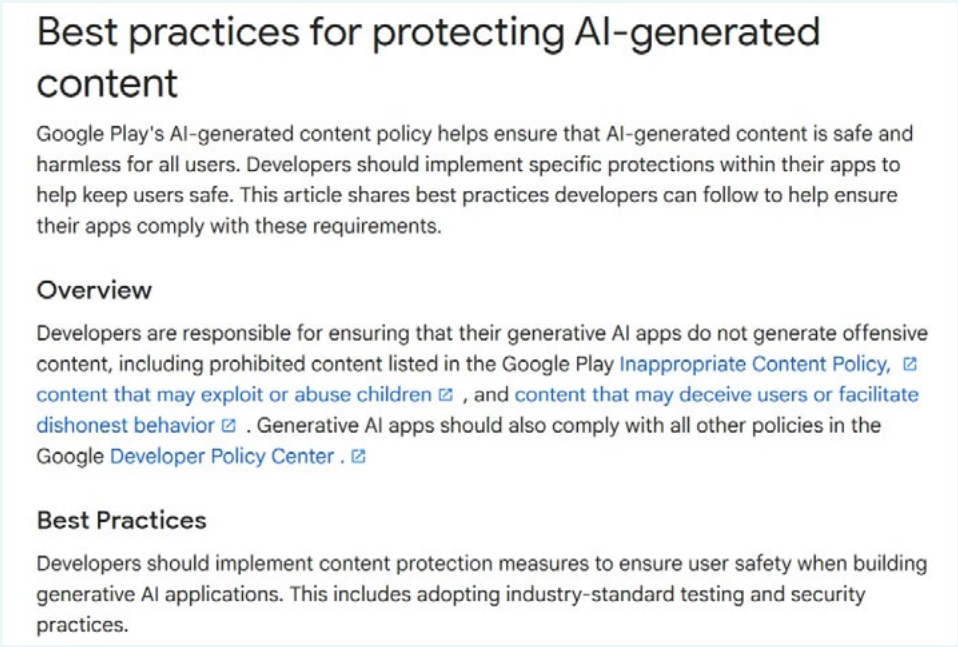
美国的内容监管体系呈现出联邦与州分权、多领域法律交叉、言论自由与安全治理动态平衡的显著特征。

美国政府在网络内容监管上较为宽松，其著名的第 230 条法规规定，互联网服务的运营商无需对平台上的第三方言论承担法律责任，这几乎等同于是网络平台的豁免令。虽然美国近年来就废除 230 条款一事展开多轮辩论，但截至目前，互联网平台无需为平台内容负责的立法现状仍未改变。然而，政府立法监管的缺位并不代表美国网络世界毫无限制。

__其一__，美国政府虽然囿于对自由言论的保护，无法通过立法约束平台，但美国国会也在不断向企业施压，要求企业采取措施解决网络内容乱象。

__其二__，在美国活跃的社交平台基本上都属于全球头部企业，虽然美国本土无立法限制，但其他地方的立法要求，如欧盟，也会倒逼美国企业对网络内容进行限制。

__其三__，海外的主流应用商店 Google Play 和 App Store 均为上架 APP 设置了内容审核条款，头部行业自发设置的准入门槛成为美国互联网企业不得不遵守的硬性要求。2025 年 7 月 10 日，Google Play 发布政策公告，新增了 "AI 生成内容政策"，针对 AI 生成内容主要规定了 AI 生成内容的定义、生成式 AI 的免责使用、开发者责任三大板块。应用平台接纳 AI 生成内容并逐步建立起对 AI 生成内容的管理已是大势所趋。



Best practices for protecting AI-generated content

Google Play's AI-generated content policy helps ensure that AI-generated content is safe and harmless for all users. Developers should implement specific protections within their apps to help keep users safe. This article shares best practices developers can follow to help ensure their apps comply with these requirements.

Overview

Developers are responsible for ensuring that their generative AI apps do not generate offensive content, including prohibited content listed in the Google Play Inappropriate Content Policy, content that may exploit or abuse children , and content that may deceive users or facilitate dishonest behavior . Generative AI apps should also comply with all other policies in the Google Developer Policy Center .

Best Practices

Developers should implement content protection measures to ensure user safety when building generative AI applications. This includes adopting industry-standard testing and security practices.

图源：Google Play

接下来，针对美国网络内容监管的重点领域进行分析。包括未成年人、色情内容、恐怖主义、网络赌博以及仇恨言论。

未成年人

20 世纪 90 年代，由于未成年人容易接触到互联网上大量的色情内容，引发民众与政府的担忧与抵制，由此开始了美国第一波互联网管制浪潮。此后，美国联邦多次尝试通过立法建立一套强制内容审核系统，但由于《美国宪法第一修正案》**对言论自由的强力保护，几次提案都以失败告终**。几次比较著名的立法提案包括 1996 年《通讯规范法案》（Communications Decency Act, CDA）、1998 年《儿童在线保护法案》（Child Online Protection Act, COPA）、 2000 年《儿童互联网保护法案》（Child Internet Protection Act）以及 2000 年《儿童在线隐私保护法案》（Child Online Privacy Protection Act, COPPA）。

2024 年 7 月 31 日，美国参议院通过了《儿童和青少年在线隐私保护法案》（简称 COPPA 2.0）和《儿童在线安全法案》（KOSA），两项新法案的出台将禁止对未成年人进行定向广告投放和未

经同意的数据收集，同时赋予家长和孩子删除其社交媒体平台个人信息的权利，此外还将明确社交媒体公司在未成年人使用其产品时的"注意义务"，重点关注互联网平台的合规设计和公司监管义务。2025 年 1 月，美国联邦贸易委员会（FTC）修订《儿童在线隐私保护法》（COPPA 2.0），以加强儿童信息保护，赋予家长更多控制权。

此前美国佛罗里达州一女子控告谷歌公司和 Character.AI 平台，指控后者的人工智能（AI）聊天机器人"教唆"其 14 岁儿子自杀。在美国高度重视未成年人保护的背景下，引发各界关注和讨论。

色情内容

在美国，网络色情是一个非常有争议的话题。自 20 世纪 90 年代以来，公众、立法者以及法院就如何控制网络色情内容争论不休。反对者希望从道德的角度出发对网络色情内容进行管制，支持者则认为网络色情内容同样受《宪法第一修正案》保护，是公民言论自由的一部分。

单从立法层面而言，在美国通过网络浏览色情内容并不违法，但儿童色情是坚决不能触碰的红线。1990 年，最高法院裁定儿童色情内容不受宪法修正案保护，**禁止制作、持有、传播或意图传播任何以未成年人（不满 18 周岁）为对象的淫秽视觉材料**。2025 年修订的《儿童性虐待材料法》，明确将 AI 生成的儿童色情内容纳入"视觉材料" 范畴，持有或传播此类内容最高可判 10 年监禁，累犯或涉及真实儿童性虐待的案件可加重至终身监禁。

2025 年生效的《删除法案》（TAKE IT DOWN Act）进一步要求社交媒体平台在 48 小时内删除用户举报的儿童深度伪造影像，否则将失去《通讯规范法案》第 230 条的免责保护，并可能面临刑事处罚。

恐怖主义

随着社交媒体的蓬勃发展，越来越多的恐怖分子和暴力极端分子也开始利用社交媒体开展宣传、招募、运营等。由于绝大多数活跃的互联网平台都是美国企业，美国互联网巨头也日益受到来自政府、股东、公民社会团体、媒体、民众等多方压力，要求平台采取措施对有害内容进行管制。近年来，多方呼吁美国政府通过立法直接对网络恐怖主义内容进行监管，但从实际情况来看，美国政府的动作还是较为克制，而社交平台自身成为网络恐怖主义内容管控的中坚力量。

经过长时间的讨论与尝试，美国现在已形成一套政府与行业共同参与的复杂的私下调解机制。美国政府自身不直接介入平台上的内容管制，而主要将这一责任"外包"给企业。一方面，美国政府虽然不会直接制定法律来监管网络上的恐怖主义内容，但它会通过立法权向社交媒体企业施压。另一方面，美国政府也通过提供专家指导、公私合作的方式，鼓励企业采取审核手段对抗平台上的恐怖主义内容。

在公众和美国政府的施压之下，美国各大社交媒体平台开始加快反恐步伐。事实上，美国互联网巨头最先感受到的压力来自欧洲政府。在 2015 年和 2016 年欧洲发生多次恐怖袭击事件后，Facebook、Twitter 和 Google 为回应欧洲各国政府的施压，开始移除其平台上的恐怖主义内容。

网络赌博

在过去很长一段时间里，美国联邦政府一直将网络赌博列为非法行为。2006 年，美国更是通过了《取缔非法互联网赌博法案》，要求美国的银行和信用卡公司不得为互联网博彩网站提供信用卡、支票、电子支付等转账支付，违者将受处罚。

转变发生在 2011 年，这一年联邦裁定，只有体彩赌博在网络上是被禁止的，其他形式的网络赌博不受法律约束，对网络赌博的立法权被交到了各州政府手上。截至 2025 年，美国共有 48 个州宣布网络赌博合法，仅夏威夷州和犹他州通过立法明确禁止所有形式的网络赌博。

虽然法律并不禁止赌博，但各社交平台都对线上赌博广告做出了限制。以谷歌为例，谷歌仅允许国营实体投放彩票广告，仅允许在在线赌博合法的州投放赛马、体育博彩、在线赌场的广告。同时，广告客户不得定位到未满 21 周岁的用户或获得许可的州以外的用户，必须在登录页或广告素材中添加针对成瘾性和强迫性赌博的危险警告和相关帮助信息。

仇恨言论

管制网络仇恨言论最先是由欧盟提出。2015 年，欧洲难民危机一再升级，网络仇恨言论开始成为欧盟立法机构与网络巨头争论的焦点。2016 年 5 月，欧盟委员会宣布与 Facebook、Twitter、YouTube 和 Microsoft 就控制网络仇恨言论开展合作，四家公司承诺"接到举报后 24 小时内屏蔽和删除相关仇恨言论"。欧盟各国也相继立法对网络仇恨内容进行管制。

但与欧盟通过立法管控仇恨言论不同，由于美国宪法对言论自由的高度保护，再加上《通信端正法》对互联网企业的豁免条款（第 230 条规定，互联网平台无需为其平台上的任何内容负责），美国政府以技术手段或立法过滤网络内容的情况较少出现，企业自律与行业合作成为互联网公司对抗仇恨言论的主要途径。

以 Facebook 为例，Facebook 在其社区守则中将"仇恨言论"定义为"针对他们受保护的特征，而非观念或习俗直接发起的直接言论攻击，这些特征包括：民族、种族、原国籍、残疾、宗教信仰、种姓、性取向、性别、性别认同，以及严重疾病"，并明确禁止在平台上发布仇恨言论。

美国内容合规总结

1. *美国的内容监管体系呈现出联邦与州分权、多领域法律交叉、言论自由与安全治理动态平衡的显著特征。*
2. *AI 监管从强调“过程合规”，转向了更注重“结果控制”的模式，AI 技术的发展方向和内容产出，符合政府的价值观和政治议程，即所谓的“美国价值观”。*
3. *未成年人相关内容是出海美国不可逾越的合规红线，即便是 AI 生成的涉未成年人内容，亦需严格遵循这一刚性准则。*
4. *海外应用商店中 Google Play 率先新增了 “AI 生成内容政策”，应用平台接纳 AI 生成内容并逐步建立起对 AI 生成内容的管理已是大势所趋。*

加拿大

加拿大作为科技创新的先锋国家，正通过完善政策框架和治理体系，在全球 AI 发展中扮演着领导角色。自 2017 年首次发布国家级人工智能战略以来，加拿大通过一系列政策和立法，奠定了 AI 领域的治理基础。与此同时，加拿大在国际合作、伦理规范、标准化建设以及高影响力人工智能系统的法律约束方面均展现出了卓越的前瞻性和引领力。

加拿大虽非全球 AI 应用收入的 “头部玩家”，但凭借高用户付费能力、完善的数字基础设施及政府对 AI 产业及高等教育的强力扶持，成为北美市场中具有增长潜力的“优质赛道”。

政策红利与产业导向

2017 年，加拿大政府与加拿大高等研究院（ Canadian Institute for Advanced Research CIFAR）联合发布《泛加拿大人工智能战略》（PCAIS），这一文件使得加拿大成为全球首个设立国家级人工智能战略方针的国家，这一战略通过推动人工智能技术的商业化应用、建立人工智能的标准化以及培养支持顶尖人工智能人才三大支柱，为加拿大奠定了成为全球人工智能行业领导者的基础。

在 PCAIS 实施的第一阶段期间（2017 年 -2021 年），加拿大政府不仅通过大规模资助推动了蒙特利尔、多伦多和埃德蒙顿三大人工智能研究中心的建设，也通过吸引全球顶尖学者强化了 AI 领域的研究实力。与此同时，该战略推动了多项人工智能技术向实际商业应用的转化，与行业伙伴合作的研究项目不断增加，为构建国际人工智能生态体系打下了坚实的产业基础。在 2022 年开始的第二阶段实施期间，2022 年启动的第二阶段则更加注重人才培养、研究前沿和商业化应用之间的联系，同时强化了对人工智能伦理规范的支持，并致力于通过标准化建设加强国际合作。

2024 年 11 月 12 日，加拿大创新、科学与工业部部长弗朗索瓦 - 菲利普·尚帕涅宣布成立加拿大人工智能安全研究所（Canadian Artificial Intelligence Safety Institute，简称 CAISI），以增强加拿大应对 AI 安全风险的能力，进一步巩固加拿大在安全、负责任地开发和采用 AI 技术方面的领先地位。这是 2024 年加拿大联邦预算中宣布的 24 亿加元人工智能投资计划的一部分，该计划旨在帮助研究人员和企业负责任地开发和采用 AI 技术。

AI 监管政策：联邦主导、省域补充的 "风险分级" 框架

加拿大采用"联邦 - 省" 双层监管体系，联邦政府负责制定全国性 AI 治理原则与核心法律，各省根据本地产业特点出台补充规则，整体呈现"强伦理约束、弱意识形态干预、重风险适配"的特征，与美国 2025 年《赢得 AI 竞赛》计划的 "去监管化" 思路形成鲜明对比。

加拿大的AI及数据管理规定和指引

名称	主要内容	发布时间	效力层级
《全球人工智能标准化现状及对加拿大启示的白皮书》	对分析全球人工智能标准化的发展现状进行分析，涵盖国际组织(如ISO、IEC、IEEE)及主要国家在标准制定中的关键举措和重点领域，包括**数据管理、算法透明性、伦理规范及风险管理**等方面，同时提出了加拿大在国际标准化进程中提升竞争力和引领作用的战略建议。	2023年3月	行业指引
《生成式人工智能行为守则》	该文件旨在推动生成性AI的负责任使用，涵盖伦理原则、用户保护、问责机制和社会影响等方面。该准则强调透明性和公平性，建议**开发者对AI系统采取有效的安全措施**，明确AI系统生成内容的侵权责任，并鼓励行业和公众参与对AI系统监督，以确保技术的发展符合社会价值。	2023年3月	行业指引
《AI和数据法案》(AIDA)	该法案着重于**确保AI系统的安全性和非歧视性**，并使企业在其控制下的AI活动中承担责任。AIDA的目标是建立一个适应新技术发展的法律框架，以应对AI系统可能带来的潜在风险，并为加拿大政府监管AI提供法律依据。	2022年	法律法规(还未通过)
《泛加拿大人工智能战略》(PCAIS)	加拿大于2017年启动的全球首个国家级人工智能战略，旨在通过支持研究、培养人才和推动技术商业化，巩固加拿大在人工智能领域的全球领导地位。战略由三大支柱构成:**人才与研究**(通过CIFAR和三大AI研究机构吸引顶尖学者并推进研究)，**商业化**(支持人工智能技术的实际应用和产业转化)，**以及标准**和**政策制定**(推动AI技术的规范化和国际合作)。	2017年3月	国家战略

1）联邦层面

随着 AI 技术应用的迅速扩展，加拿大在 2022 年推出了《数字宪章实施法案》（C-27 法案）。这部法案是加拿大政府为应对快速变化的数字经济和技术挑战而提出的重要法律框架，其中包括几项关键子法案，分别针对不同领域的问题进行规制。

《人工智能与数据法案》（AIDA）

作为《数字宪章实施法案》的关键组成部分，AIDA 是加拿大首部专门针对 AI 的专项立法。AIDA 的主要目标是规范涉及人工智能系统的国际和省际贸易活动，确保责任主体对 AI 技术的开发和使用行为负责，同时保护公众免受高风险 AI 系统可能造成的潜在危害。法案定义了"人工智能系统"为利用算法（如遗传算法、神经网络和机器学习）自动或部分自动处理数据并生成内容、做出决策或预测的技术。

AIDA 法案特别关注高影响力 AI 系统，责任主体需评估其系统是否属于高影响力类别，并采取措施识别、评估和缓解可能的危害或偏见，确保其运行符合公众利益。AIDA 还明确禁止使用非法获得的个人信息进行 AI 开发，并对可能造成严重社会危害的 AI 系统实施禁令。这些规定旨在强化人工智能技术使用过程中的安全性与合规性，并赋予政府部门强有力的监管和执行权力。AIDA 法案虽尚未通过，但联邦政府仍将其作为 AI 监管核心框架推进。

《消费者隐私保护法案》（CPPA）

作为 C-27 法案的另一重要组成部分，意图取代现行的《个人信息保护与电子文档法案》（PIPEDA），从而进一步加强了数据隐私保护力度。通过赋予数据主体删除权和数据可携权，CPPA 确保了用户对其数据的更大掌控权，同时要求企业在 AI 开发中优先考虑隐私保护（Privacy by Design），并在处理数据时保持透明性。

虽然 CPPA 并非专门针对人工智能（AI）系统和技术的立法，但其通过间接方式规范了 AI 技术的使用。该法案的核心在于加强对个人数据的保护，而 AI 系统的发展和运行通常依赖于对大量数据的处理，这种对隐私保护的高标准规范，间接迫使 AI 技术开发者和使用者遵守更为严格的数据管理和伦理标准。

截至目前，该 C-27 法案虽尚未正式通过，但该法案的提出标志着加拿大在人工智能监管领域迈出了重要一步，并为未来的政策制定奠定了基础。

2）省域层面：产业适配性监管实践

除了联邦层面的政策和法律，加拿大的省级政府也开始探索 AI 监管的地方实践。例如，魁北克省更新了隐私保护法律，其间接影响了 AI 技术在数据使用方面的合规性。更新后的法案要求在数据处理和跨境转移时严格遵循隐私设计原则，同时确保数据主体对信息处理有充分的知情权。

安大略省则启动了数字健康解决方案创新计划（IDHS），由安大略创新中心（OCI）主导，这一计划覆盖了 13 个项目，目标是通过数字化创新提升医疗服务质量。与此同时，安大略省还在制定可信人工智能框架，以指导 AI 在政府服务中的使用。此外，在医疗和教育等领域，安大略省正在试点 AI 技术的应用，旨在提高医疗诊断的效率和准确性，以及优化教育资源分配。

省级的地方性实践不仅丰富了加拿大 AI 监管的多样性，也为联邦层面的立法提供了宝贵的试验和经验支持。

3）加拿大在国际治理与全球合作体系中的角色

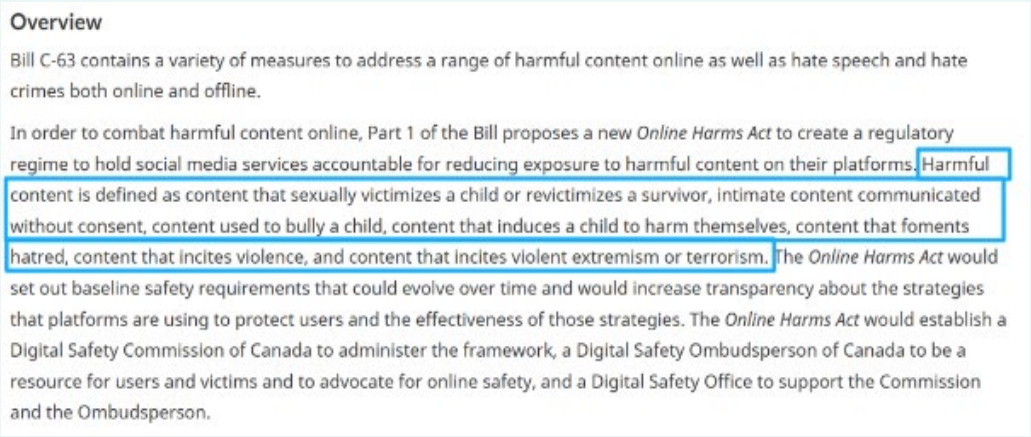
2020 年，加拿大与其他 G7 成员国及欧盟共同发起了《全球人工智能伙伴关系倡议》（Global Partnership on Artificial Intelligence, GPAI）。这一多边倡议致力于推动人工智能的负责任开发与应用，通过覆盖人权、隐私保护和可持续发展等领域的研究，构建多利益相关方参与的平台。作为 GPAI 的创始成员之一，加拿大在这一倡议中发挥了主导作用，不仅推动了人工智能技术的伦理规范，还进一步增强了其在国际人工智能治理中的话语权。

2023 年，加拿大创新、科学与工业部发布了《生成式人工智能自愿行为守则》（Voluntary Code of Conduct on Advanced Generative AI Systems）。该守则围绕问责、安全、公平性、透明度、人类监督与监控以及系统稳健性六大原则，为生成式人工智能技术（如 ChatGPT 和 MidJourney）的开发者和管理者提供了指导。尽管该守则为自愿性质，但它通过风险评估、信息透明化等手段，填补了正式立法前的监管空白，并为未来法律的实施提供了实践依据。

加拿大具体内容监管政策

加拿大对 AI 生成内容及网络公开传播内容的监管，以“保护弱势群体（未成年人）、维护社会公共利益、尊重多元文化”为核心，通过“法律强制 + 平台责任 + 行业自律”三重机制实现，重点监管 7 类有害内容：

- 儿童性剥削材料
- 非自愿私密内容（含深度伪造）
- 欺凌儿童
- 诱导儿童自伤
- 煽动仇恨
- 煽动暴力
- 恐怖主义
- 暴力极端主义



Overview

Bill C-63 contains a variety of measures to address a range of harmful content online as well as hate speech and hate crimes both online and offline.

In order to combat harmful content online, Part 1 of the Bill proposes a new *Online Harms Act* to create a regulatory regime to hold social media services accountable for reducing exposure to harmful content on their platforms. Harmful content is defined as content that sexually victimizes a child or revictimizes a survivor, intimate content communicated without consent, content used to bully a child, content that induces a child to harm themselves, content that foments hatred, content that incites violence, and content that incites violent extremism or terrorism. The *Online Harms Act* would set out baseline safety requirements that could evolve over time and would increase transparency about the strategies that platforms are using to protect users and the effectiveness of those strategies. The *Online Harms Act* would establish a Digital Safety Commission of Canada to administer the framework, a Digital Safety Ombudsperson of Canada to be a resource for users and victims and to advocate for online safety, and a Digital Safety Office to support the Commission and the Ombudsperson.

图源：加拿大司法部官网

未成年人保护 -《刑法》、《互联网儿童色情强制报告法》等

加拿大将未成年人网络保护视为"国家优先事项"，相关法律与政策远超美国早期监管力度，核心内容包括：

> ***严格禁止内容：****强制平台移除涉及儿童性剥削、教唆儿童自残、欺凌未成年人、未经同意的儿童私密内容等危害内容（含 AI 生成）。*
>
> ***24 小时删除义务：****各大社交媒体公司必须迅速删除性侵害儿童的内容以及未经同意传播的私密内容，平台需在用户举报后 24 小时内删除，并接受监督和审查，否则最高可被处以其全球总收入 6% 的罚款。*
>
> ***年龄验证与家长控制：****要求面向 13 岁以下用户的平台实施年龄验证（如身份证 / 生物识别），并提供家长控制台，允许父母查看孩子的 AI 交互记录 。*

仇恨言论与有害内容：《刑法》修订

《刑法》新增仇恨"犯罪罪名：明确基于种族、宗教、性别等可识别群体的仇恨言论"为独立罪名，最高可判终身监禁（如煽动种族灭绝）。

> ***平台主动审核义务:****要求平台对仇恨言论、恐怖主义内容（含 AI 生成）实施实时监测。*
>
> ***72 小时响应机制：****非儿童类有害内容（如成人仇恨言论），平台需在 72 小时内响应举报，否则 CRTC （加拿大广播电视和电信委员会）可介入调查。*

本土内容保护：《在线流媒体法案》

> ***加拿大内容配额：****外国数字媒体平台需将在加拿大境内营收的 5% 用于支持本土内容制作。*
>
> ***豁免与处罚：****未达标平台将面临罚款，小型平台（如独立播客）可申请豁免。*
>
> ***禁止内容独家性：****禁止平台对加拿大本土电视节目实施独家播放（如仅在某 ISP 提供），确保全民可及。*

平台通用义务：年龄验证、举报机制

年龄验证：*所有涉及未成年人的平台（如 AI 教育、社交 APP）需要使用可信年龄验证（如 Equifax 认证、信用卡验证），禁止 13 岁以下儿童使用成人功能，目前只是行业规范，无联邦立法强制执行。*

一键举报与响应：*平台需在显著位置设置"举报有害内容" 按钮，举报后需向用户反馈处理进度。*

算法透明度：*针对推荐算法（如 TikTok 加拿大版），需向 CRTC 备案推荐逻辑，不过目前尚未实施强制备案。*

加拿大内容合规总结：

1. 监管框架：实行“联邦主导 + 省域补充” 双层模式，联邦定全国 AI 治理核心规则，魁北克、安大略等省按产业特点出补充细则（如隐私法更新、可信 AI 框架）。
2. 联邦战略与立法：2017 年推全球首个国家级 AI 战略 (PCAIS)，2022 年提 C-27 法案（含 AI 专项法 AIDA、隐私法 CPPA），2024 年设 CAISI 并纳入 24 亿加元 AI 投资。
3. 7 类危害内容监管：禁儿童性剥削材料等 7 类内容（含 AI 生成），儿童类违规内容需 24 小时删除，仇恨言论入《刑法》。
4. 平台义务：涉未成年人平台需可信年龄验证，推荐算法需向 CRTC 备案（暂未强制）。
5. 本土保护与国际参与：《在线流媒体法》要求国外平台将加拿大境内营收的 5% 投入到本土内容，2020 年参与发起 GPAI，2023 年发布生成式 AI 自愿行为守则。

南美：以巴西为主导的隐私保护与内容审核挑战

巴西

巴西作为拉美最大经济体与人口大国（人口超 2.16 亿，居全球第六），既是区域经济增长的核心引擎，也是新兴市场中 AI 产业发展的关键枢纽。

2025 年《巴西人工智能法》的通过与 2025 年中巴 AI 合作备忘录的签署，标志着巴西在平衡技术创新与风险管控上迈出关键步伐。本文基于巴西联邦议会、国家数据保护局（ANPD）等官方渠道的政策文本，解读其 AI 战略及法规政策、内容监管要点等，为出海企业提供可落地的政策指引与风险应对方案。

巴西 2024 年 GDP 总量达 11.7 万亿雷亚尔（约合 2.02 万亿美元），人均 GDP 超 1 万美元，叠加 87.5% 的城镇化率、81% 的互联网渗透率、82% 的智能手机普及率及 73% 的社交媒体渗透率，形成了极具吸引力的数字经济土壤 —— 尤其对寻求新兴市场的中国 AI 与互联网企业而言，巴西被业内称为“全球最后一片出海热土”（东南亚市场趋于饱和、非洲基建尚未完善，巴西的市场规模与数字基础形成独特优势）。

从产业导向看，巴西正通过密集的战略布局与立法行动构建 AI 监管体系：2025 年《巴西人工智能法案》正式通过，2025 年中巴签署 AI 合作备忘录，标志其在“技术创新”与“风险管控”的平衡上迈出关键步伐。但需注意，尽管拉美整体市场环境开放，外国企业（尤其是互联网内容平台与 AI 应用）仍需重点关注内容合规与监管响应 —— 此前 X 平台、WhatsApp 因未配合司法要求遭临时封禁的案例，凸显了本地合规的必要性。

国家战略布局

基础战略奠基：2021 年《巴西人工智能战略》（EBIA）确立伦理原则、研发投资、人才培育等六大目标，构建 AI 发展基本框架，截至 2025 年已推动 127 个州级 AI 创新项目落地。

专项投资计划：2024 年推出《巴西人工智能投资计划（PBIA 2024-2028）》，四年投入 230 亿雷亚尔（约 42 亿美元），重点建设国家级 "主权云"、葡语大语言模型（覆盖农业、医疗专业术语库）及 28 个州的算力基础设施升级。

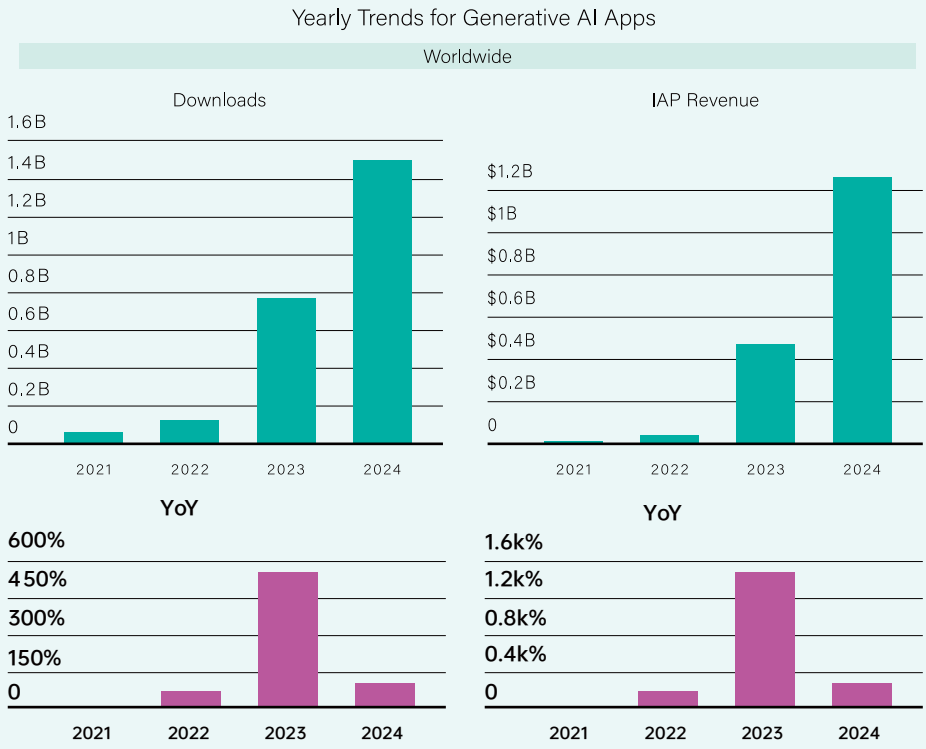
国际合作深化：2025 年 6 月，巴西总统卢拉与中国国家主席习近平签署 AI 合作谅解备忘录，明确在智能制造、农业科技、人才培训、开源技术共享四大领域开展技术对接，设立中巴 AI 联合实验室推动合规标准互认。

发展现状及潜力

从投资与产业发展来看，人工智能正成为巴西经济增长的重要驱动力。根据 Statista 数据库预测，无论在乐观、中等还是保守情景下，人工智能市场对巴西国内生产总值的影响都将是显著的。2025 年巴西人工智能项目投资预计将超过 130 亿雷亚尔（约 24 亿美元），主要覆盖农业、金融、制造业和公共服务等关键领域。

巴西地理统计局（IBGE）最新发布的《创新半年度调查报告》（Pintec）显示，过去两年间，巴西工业企业人工智能（AI）应用数量从 1619 家跃升至 4261 家，增幅高达 163%。研究指出，生成式人工智能的普及是推动增长的关键因素。自 ChatGPT 问世并迅速扩散后，巴西工业企业在行政管理、商业销售、生产流程优化等环节广泛应用 AI。

另据 Sensor Tower 数据显示， 2024 年巴西已成为全球第七大生成式人工智能应用市场，这一排名远超其在"全球移动消费总市场" 中的位次（未进入前十），凸显出巴西在生成式 AI 领域的独特增长潜力。



The genre is also quite popular in Brazil, which ranks #7 for Generative AI (and outside the top 10 markets by overall mobile consumer spend). Generative AI revenue is lagging a bit in China Mainland, which accounted for less than 2% of the genre's IAP revenue in 2024.

图源：State of Mobile 2025

巴西人工智能政策：创新发展与风险管控并重

《巴西人工智能战略》（以下简称"EBIA"），是巴西首个专门针对人工智能的联邦战略。EBIA 确立了六大核心目标：**促进负责任人工智能的伦理原则发展**、推动人工智能研发的持续投资、消除人工智能创新障碍、培养和教育人工智能生态系统专业人才、在国际环境中刺激巴西人工智能创新发展以及促进政府、私营部门、产业和研究中心在人工智能发展方面的合作环境。

2024 年 7 月，巴西推出了《巴西人工智能投资计划（PBIA 2024-2028）》，主题为"人工智能惠民"。该计划在 EBIA 基础上进行了升级，**提出了 31 项即时影响行动和 54 项结构性行动**，涵盖基础设施和人工智能发展、培训和能力建设、人工智能改善公共服务、人工智能业务创新以及人工智能监管和治理流程五大支柱。计划的结构性行动包括升级 Santos Dumont 超级计算机，将其打造成全球五大最强超算之一。

2024 年 12 月 10 日，巴西参议院批准了第 2338/2023 号法案（PL 2338/2023），即《巴西人工智能法案》，2025 年 2 月众议院已审议通过。该法案采用**基于风险和权利的监管方法，旨在进一步加强巴西合乎伦理和负责任的人工智能使用**，并通过制定国际标准和最佳实践来提升其在全球人工智能治理中的引领作用。该法案与联合国 2030 年可持续发展议程和联合国教科文组织《人工智能伦理问题建议书》的要求保持一致。法案的核心内容包括提升透明度和信息权利保障、明确问责和损害赔偿机制以及强调环境保护要求等。

巴西互联网监管法规体系

巴西政府**整体上更强调开放自由的互联网氛围，对内容审核持谨慎态度**，目前并不明确要求 APP 具有事前审核能力，更强调事后处置。因此在内容审核的监管上，出现频率最高的是法院：法院在收到起诉后会要求 APP 删除相关内容、提交数据等，如果 APP 不配合或未能及时回应，**法院将会要求巴西国家电信局、联邦警察等执法机构封禁该 APP**。此前，X 平台曾因未能及时回应法院判决而遭到封禁。

4 新闻背景

环球时报

第6324期 2024年9月2日 星期一

大法官发出通牒 马斯克拒不遵守

巴西下令关闭X平台

本报特约记者 李一 王迅

公告

X 平台在巴西被禁　来源：环球时报

监管体系

目前，巴西专门监管互联网领域的顶层法律有两部：《巴西通用数据保护法》以及《巴西互联网法》。

1）《巴西通用数据保护法》（LGPD）

2020 年 LGPD 正式生效，是巴西首部综合性数据保护法律，其制定大量借鉴了 GDPR，对个人数据的收集、使用、处理和存储规则做了详细规定。

与 GDPR 一样，LGPD 允许数据跨境传输，但需要满足以下条件：1）充分性认定；2）使用标准合同条款或数据保护机构批准的其他机制。

未能遵守 LGPD 的企业最高可面临营收 2% 的罚款，单次罚款最高 1290 万美元。

2024 年巴西数据保护局（ANPD）暂停执行美国 Meta 公司的最新隐私政策，禁止 Meta 根据巴西用户发布的帖子训练生成式人工智能（AI）模型。

DIÁRIO OFICIAL DA UNIÃO

Publicado em: 02/07/2024 | Edição: 125 | Seção: 1 | Página: 180

Órgão: Ministério da Justiça e Segurança Pública/Autoridade Nacional de Proteção de Dados/Conselho Diretor

DESPACHO DECISÓRIO Nº 20/2024/PR/ANPD

Processo nº 00261.004509/2024-36

Interessado: META PLATFORMS INC - FACEBOOK SERVIÇOS ONLINE DO BRASIL

O CONSELHO DIRETOR DA AUTORIDADE NACIONAL DE PROTEÇÃO DE DADOS PESSOAIS, no uso de suas atribuições legais e regulamentares, com base no art. 45 da Lei nº 9.784/1999; nos arts. 52, III, e 54, da LGPD; no art. 26, IV, do Decreto nº 10.474/2020; e nos arts. 7º, IV e 55 do Regimento Interno da ANPD, nos termos do Voto 11/2024/DIR-MW/CD, cujas razões acolhe e integra à presente decisão, conforme autoriza o § 1º do art. 50 da Lei nº 9.784/1999, considerando a deliberação tomada no Processo nº 00261.004509/2024-36, resolve proferir medida preventiva para determinar à Meta Platforms INC - Facebook Serviços Online do Brasil, até ulterior decisão desta Autoridade, a imediata suspensão no Brasil: (i) da vigência da nova política de privacidade da empresa, no que toca à parte relativa ao uso de dados pessoais para fins de treinamento de sistemas de IA generativa; e (ii) do tratamento de dados pessoais dos

图源：巴西数据保护局（ANPD）网站通报

2）《巴西互联网法》以及 2016 年修正案

《巴西互联网法》又称《网络民法》，有巴西"网络宪法"之称，对用户、企业和公共机构在巴西境内使用互联网的原则、权利与义务给出了详细规定。

《巴西互联网法》对 APP 运营者在数据存储、内容审核以及配合政府审查等方面的义务如下：

数据存储

APP 的登录记录应保存 6 个月，且应确保存储环境加密、安全、可控。在特殊情况下，监管机构有可能要求保存时间超过 6 个月。

内容审核

为了确保言论自由，巴西在处理网络违规内容时，主要使用国外较为通行的"Notice and Take Down"原则。《巴西互联网法》规定 APP 运营者无需为第三方发布的违规内容负责，也无需前置审核措施。但是，运营者有义务配合监管机构的要求及时删除违规内容，如果在收到法院通知后拒绝删除相关内容，则会面临处罚。

在收到法院的删除通知后，如果运营商拥有违法内容发布者的联系方式，运营商有义务与其沟通删除原因。

当有人在平台上私自散布其他人的裸体、性行为照片、影像等内容，造成侵犯他人隐私的，运营商应在收到当事人删除请求后立刻删除相关内容，否则也将承担相应的法律责任。

配合政府审查

按照《互联网法》及 2016 年的修订案，巴西监管机构有权要求 APP 运营者提供用户注册信息：任职信息（filiation information）、住址（address）、个人信息（姓名、婚姻状况 understood as name, surname, marital status and occupation of the user）。监管机构不应要求 APP 运营者批量提供用户注册信息，应明确所需用户信息以及希望其提供的内容。

如果 APP 运营者在注册时并不收集此类信息，则可免除配合义务。

WhatsApp 在巴西多次被禁：*WhatsApp 在巴西曾多次因法律纠纷被法院封禁，其中 2015 年 12 月至 2016 年 7 月这八个月的时间里，巴西当局就曾三次下令封禁 WhatsApp，原因是 WhatsApp 拒绝向执法部门提供有关犯罪案件相关的用户数据。这三次封禁令都在短暂生效后解除。*

来源：网络

内容合规要点

巴西整体较为开放的社会环境，互联网内容在巴西并不会受到非常严格的监管。但近年来，随着社交媒体上各类消息泛滥、真假难辨，尤其是有关总统选举、种族歧视的内容往往会引起较大的社会动荡，因此近年来，巴西政府也开始针对这些内容进行了一些治理动作。

1）种族歧视

种族歧视在巴西是一种非常严重的犯罪行为。根据巴西法律规定，任何煽动和实施基于种族、肤色、民族、宗教或国籍歧视与偏见的种族主义行为都属于犯罪，将面临 1 至 3 年监禁及罚款，且不允许保释。

据新闻报道，巴西研究人员曾于 2015 年开发一款应用程序，用于监测社交媒体上包含宣扬暴力、仇恨、种族主义等元素的不当言论。该软件通过对社交网络聊天对话中涉及性暴力、种族主义和针对有色人种、土著居民、移民群体、变性人等歧视性关键词的自动搜索，允许监管者对不

良用户进行识别，从而为当局者提供管理处罚依据。目前尚不清楚该应用程序是否仍在被使用。

2）假新闻

巴西目前正在推动制定《假新闻法案》，截至目前该法案尚未通过。不过值得注意的是，《假新闻法案（草案）》除了要求平台加大对假新闻的处理力度外，还要求外国企业在巴西当地设立法人、定期公开内容移除报告等。出海巴西企业应及时跟踪《假新闻法案》的进展。

3）儿童

虽然巴西对色情内容并不排斥，甚至会有专门放映色情影片的电影院，但儿童色情是绝对不能逾越的红线。根据巴西的《儿童和青少年法规》，儿童色情全链条的参与者都会面临严厉的处罚，在当地运营的 APP 应尤为注意对儿童色情内容的甄别。巴西国会众议院 8 月 20 日通过一项法案，对社交媒体平台保护未成年人提出更多要求，包括实行更严格的用户年龄验证机制以及绑定 16 岁以下未成年用户及其父母账号。

4）AI 标识义务

所有 AI 生成内容需添加符合巴西国家标准的不可见数字水印，包含生成时间、系统标识；选举广告需在显著位置标注 “AI 生成” 。

巴西内容合规总结：

1. *AI 监管法案：第 2338/2023 号法案（2025 年通过），要求 AI 透明、明确问责，对齐国际伦理*
2. *数据与互联网合规：LGPD 管数据（最高罚营收 2%）；《互联网法》要求存登录记录 6 个月，配合删除内容及调查，拒配合可封 APP*
3. *内容红线：种族歧视犯罪（1-3 年监禁）、儿童色情零容忍（需年龄验证）、AI 内容加水印、假新闻草案要求外企设本地法人*
4. *案例警示：X 平台、WhatsApp 因拒配合遭封禁，需设本地代表、及时响应司法要求*

欧洲：全球最严的数字内容治理高地

欧洲作为全球数字经济高地与监管标杆，既是科技企业出海的核心市场，也是合规门槛最高的区域之一。其覆盖 27 个欧盟成员国及英国、挪威等非欧盟国家，总人口超 7.4 亿，经济总量占全球 20% 以上，数字经济规模达 8.5 万亿欧元，是全球第二大数字市场。对于中国出海企业而言，欧洲市场的高消费力、高互联网渗透率与成熟的产业生态极具吸引力，但同时，以《人工智能法案》《GDPR》为核心的监管体系，构建了全球最严苛的合规框架，合规已成为进入欧洲市场的"敲门砖"。

2025 年，欧洲人工智能市场规模预计突破 1500 亿欧元，较 2023 年实现年均复合增长率超 28%，成为全球增长最快的区域市场之一。IDC 预测，2024-2028 年欧洲 AI 支出 CAGR 将达 30.3%，2028 年总支出将达 1446 亿美元，其中生成式 AI 占比将从当前的 1/4 提升至 1/3。从市场结构来看，西欧（德、法、英等）凭借 97% 的互联网普及率和 85% 的网购渗透率，成为成熟核心市场；东欧则以 86% 的互联网使用率和 57% 的网购渗透率，展现出强劲增长潜力。

14 3月 2025

European AI Spending to Reach $144 Billion by 2028, Reflecting 30% CAGR, According to IDC

MILANO, March 17, 2025 — According to a new forecast from the International Data Corporation (IDC) *Worldwide AI and Generative AI Spending Guide*, European spending in artificial intelligence will reach $144.6 billion in 2028, based on a compound annual growth rate (CAGR) of 30.3% over the 2024-2028 forecast period.

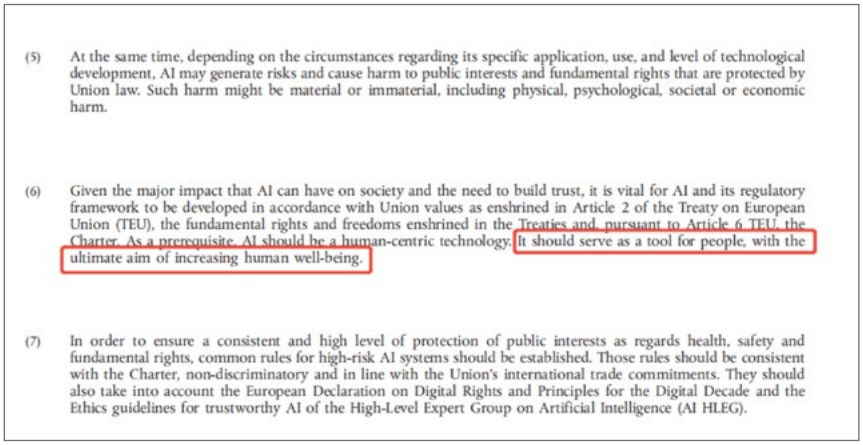
Generative AI (GenAI) is gaining momentum in Europe, currently accounting for nearly a fourth of the total AI market. By 2028, its share is expected to grow to one-third. According to IDC's *2024 Industry IT & Communications Survey*, 87% of European companies plan to allocate up to 30% of their total AI budget to GenAI solutions. "Generative AI is increasingly requiring more coordinated efforts to be deployed company wide," says Carla La Croce, research manager, Data and Analytics at IDC. "GenAI is a top priority for C-suite leaders and opens opportunities to develop more integrated internal strategies among C-suites."

图源：IDC 官方网站

中国企业出海欧洲的核心赛道集中在 AI 技术应用、跨境电商、数字内容等领域。2024 年，中国 AI 企业在欧洲的市场份额已达 8%，主要聚焦制造业智能化、跨境电商 AI 服务等场景，但合规问题仍是主要挑战。欧洲监管的核心逻辑是"权利优先、风险前置"，在保障用户隐私、基本权利的前提下推动技术创新，这要求出海企业必须将合规融入产品设计、运营全流程。

AI 发展战略及监管政策：欧盟框架为核心，各国补充适配

欧洲 AI 监管以"欧盟统一框架 + 成员国补充规则"为核心特征，欧盟层面通过《人工智能法案》确立全球首个全面 AI 监管标准，将"人本中心"（Human-centric）确立为 AI 发展的伦理底线。AI 被视为"服务人类、增进福祉的工具"，而非独立主体。欧盟各成员国在遵循统一要求的基础上，结合本国国情制定专项法规，形成"一体多元"的监管格局。



(5) At the same time, depending on the circumstances regarding its specific application, use, and level of technological development, AI may generate risks and cause harm to public interests and fundamental rights that are protected by Union law. Such harm might be material or immaterial, including physical, psychological, societal or economic harm.

(6) Given the major impact that AI can have on society and the need to build trust, it is vital for AI and its regulatory framework to be developed in accordance with Union values as enshrined in Article 2 of the Treaty on European Union (TEU), the fundamental rights and freedoms enshrined in the Treaties and, pursuant to Article 6 TEU, the Charter. As a prerequisite, AI should be a human-centric technology. It should serve as a tool for people, with the ultimate aim of increasing human well-being.

(7) In order to ensure a consistent and high level of protection of public interests as regards health, safety and fundamental rights, common rules for high-risk AI systems should be established. Those rules should be consistent with the Charter, non-discriminatory and in line with the Union's international trade commitments. They should also take into account the European Declaration on Digital Rights and Principles for the Digital Decade and the Ethics guidelines for trustworthy AI of the High-Level Expert Group on Artificial Intelligence (AI HLEG).

图源：欧盟《人工智能法案》截图

欧盟层面：《人工智能法案》主导的风险分级监管体系

2024 年 8 月 1 日，欧盟《人工智能法案》（AI Act）正式生效，成为全球首部全面监管 AI 的综合性法规，标志着欧洲 AI 监管进入法治化阶段。其核心框架围绕"风险分级"构建，覆盖 AI 全生命周期，且具有强大的域外效力——无论企业是否位于欧盟境内，只要向欧盟用户提供 AI 服务或产品，均需遵守该法案。

Official Journal of the European Union

EN
L series

2024/1689

12.7.2024

REGULATION (EU) 2024/1689 OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL

of 13 June 2024

laying down harmonised rules on artificial intelligence and amending Regulations (EC) No 300/2008, (EU) No 167/2013, (EU) No 168/2013, (EU) 2018/858, (EU) 2018/1139 and (EU) 2019/2144 and Directives 2014/90/EU, (EU) 2016/797 and (EU) 2020/1828 (Artificial Intelligence Act)

(Text with EEA relevance)

THE EUROPEAN PARLIAMENT AND THE COUNCIL OF THE EUROPEAN UNION,

Having regard to the Treaty on the Functioning of the European Union, and in particular Articles 16 and 114 thereof,

Having regard to the proposal from the European Commission,

After transmission of the draft legislative act to the national parliaments,

Having regard to the opinion of the European Economic and Social Committee [1],

Having regard to the opinion of the European Central Bank [2],

Having regard to the opinion of the Committee of the Regions [3],

图源：欧盟《人工智能法案》截图

1）核心监管逻辑：四级风险分类

- **不可接受风险：**明确禁止潜意识操纵、公共场所实时远程生物识别、工作场所情绪识别等 AI 应用，违者最高可处全球年营业额 7% 或 4000 万欧元罚款（以较高者为准）。
- **高风险 AI：**涵盖医疗诊断、教育评分、招聘筛选、生物识别等领域，需在欧盟数据库注册、通过合规评估并加贴"CE"标志，提供者需留存 10 年技术文件，违规罚款最高为全球年营业额 4% 或 2000 万欧元。
- **有限风险：**如聊天机器人等，需向用户明确标注 AI 身份，保障可追溯性，违规罚款最高为全球年营业额 1% 或 750 万欧元。
- **低风险：**如 AI 游戏、垃圾邮件过滤器等，无强制义务，鼓励企业自愿遵循行业行为准则。

2）通用 AI 与基础模型的特别监管

针对 GPT、Stable Diffusion 等基础模型，法案要求提供者履行注册、风险评估、质量管理体系建设等义务，需披露训练数据来源并保障版权合规。欧盟委员会于 2025 年启动通用 AI 行为准则咨询，计划 4 月前完成草案制定，进一步细化透明度与风险管控要求。

3）实施节奏与监管机制

法案条款分阶段生效，通用 AI 相关规定自 2025 年 8 月起实施，其余条款于 2026 年全面落地。欧盟设立专门的 AI 监管办公室，协同欧洲数据保护委员会（EDPB）及成员国监管机构开展执法，建立"主导监管机构"制度，简化跨境 AI 服务的监管协调。

最新报道显示，在大型科技公司和美国政府的巨大压力下，欧盟委员会正考虑暂停执行《人工智能法案》部分条款。此外，欧盟建议将违反 AI 透明度的罚款执行推迟至 2027 年 8 月。

成员国层面：欧盟框架下的专项补充法规

欧盟成员国需将《人工智能法案》转化为国内法，但可基于本国需求制定更细化的规则，部分非欧盟国家（如英国）则构建了独立监管体系：

1）欧盟成员国补充规则

- **德国：**在遵循《人工智能法案》基础上，通过《联邦数据保护法》强化 AI 数据处理的隐私保护，要求公共部门 AI 应用需进行额外的伦理审查。其 2018 年生效的《社交媒体管理法》对 AI 生成的仇恨言论、煽动性内容设定严格管控，200 万用户以上的平台需在 24 小时内删除明确非法内容，违者最高罚款 5000 万欧元。
- **法国：**发布《国家 AI 战略 2025》，聚焦医疗、能源等重点领域的 AI 创新，要求政府采购 AI 系统优先考虑"法国制造"或欧盟合作项目。同时通过《数字服务法》国内实施细则，强化对 AI 虚假信息的监管，选举期间禁止 AI 生成的政治广告投放。

- **西班牙：**制定《AI 伦理与治理条例》，要求 AI 企业设立专门的合规部门，配备至少 1 名欧盟认证的 AI 伦理专家，重点监管金融、医疗领域的 AI 算法公平性。

2）非欧盟国家独立体系：以英国为例

英国脱欧后未加入欧盟《人工智能法案》体系，2024 年通过《AI（监管）法案》，采用"风险匹配 + 监管沙盒"模式，平衡创新与安全。其核心差异在于：不设统一的 AI 监管机构，由信息专员办公室（ICO）、金融行为监管局（FCA）等按行业分工监管；设立为期 2 年的监管沙盒，允许企业在受控环境中测试高风险 AI 应用；罚款上限为全球年营业额 4%，低于欧盟《人工智能法案》的 7%。

其他核心监管政策：GDPR 为基，数字双法护航

(Legislative acts)

REGULATIONS

REGULATION (EU) 2016/679 OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL
of 27 April 2016
on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation)
(Text with EEA relevance)

THE EUROPEAN PARLIAMENT AND THE COUNCIL OF THE EUROPEAN UNION,

Having regard to the Treaty on the Functioning of the European Union, and in particular Article 16 thereof,

Having regard to the proposal from the European Commission,

After transmission of the draft legislative act to the national parliaments,

Having regard to the opinion of the European Economic and Social Committee [1],

图源：欧盟 GDPR 截图

除 AI 专项监管外，欧洲构建了以《通用数据保护条例》（GDPR）为基础，《数字服务法》（DSA）、《数字市场法》（DMA）为核心的数字监管体系，覆盖数据合规、内容审核、市场竞争等全维度，与 AI 监管形成协同。

数据合规核心：GDPR 及 2025 跨境执行新规

GDPR 自 2018 年生效以来，成为全球数据保护的“黄金标准”，其核心原则包括数据最小化、目的限制、知情同意、本地存储与跨境传输授权等，适用于所有处理欧盟公民个人数据的企业。

1）2025 年关键更新：跨境执行机制强化

2025年7月，欧盟理事会与欧洲议会达成临时协议，针对GDPR跨境案件执行效率不足的问题，推出五项核心优化措施：

统一投诉受理标准，消除成员国程序差异导致的处理延迟；

明确投诉人与被调查企业的程序权利，保障双方陈述权；

设定调查时限：普通跨境案件 15 个月内完成，复杂案件可延长至 27 个月，简化程序案件 12 个月内完成；

引入“早期解决机制”，企业主动纠正违法行为可快速结案；

建立简化合作程序，减少成员国监管机构间的磋商环节。

2）核心合规要求与处罚标准

- ***数据本地化：****核心个人数据（如健康、生物识别数据）需存储在欧盟境内服务器，跨境传输需通过“充分性认定”或签订标准合同条款（SCCs）；*
- ***用户权利保障：****用户享有数据访问、更正、删除、被遗忘权，企业需在 72 小时内响应用户请求；*
- ***罚款力度：****最高可处全球年营业额 4% 或 2000 万欧元罚款（以较高者为准），2024 年 Meta、谷歌等企业因跨境数据传输违规均被处以超 10 亿欧元罚款。*

网络安全监管核心：NIS2 指令

2024 年 10 月 17 日，欧盟《网络安全法》（NIS2 指令）正式生效，取代 2016 年版 NIS 指令，成为欧盟网络安全监管的基础性法规，覆盖所有"关键基础设施运营商"（如能源、金融、医疗）及"数字服务提供商"（如云计算、电商平台、社交媒体），域外效力与 GDPR 一致。

> ***核心义务：****企业需建立网络安全管理体系，制定 incident 响应预案，每年开展风险评估并向监管机构报备；发生数据泄露或网络攻击后，需在 24 小时内通知监管机构，严重事件需同步通知欧盟网络安全局（ENISA）；*
>
> ***分级监管：****按行业重要性划分"核心服务商"（如银行、大型云服务商）和"一般服务商"，核心服务商需满足更严格的安全标准，配备专职 CISO（首席信息安全官）；*
>
> ***处罚标准：****违规企业最高可处全球年营业额 4% 或 2000 万欧元罚款（以较高者为准），核心服务商违规将被暂停欧盟境内服务资质。*

内容合规核心：DSA 与 DMA 构建的全链条监管

2023-2024 年，欧盟《数字服务法》（DSA）与《数字市场法》（DMA）先后生效，前者聚焦非法内容治理，后者规范平台垄断行为，共同构成数字内容合规的"双保险"。

(Legislative acts)

REGULATIONS

REGULATION (EU) 2022/2065 OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL
of 19 October 2022
on a Single Market For Digital Services and amending Directive 2000/31/EC (Digital Services Act)

(Text with EEA relevance)

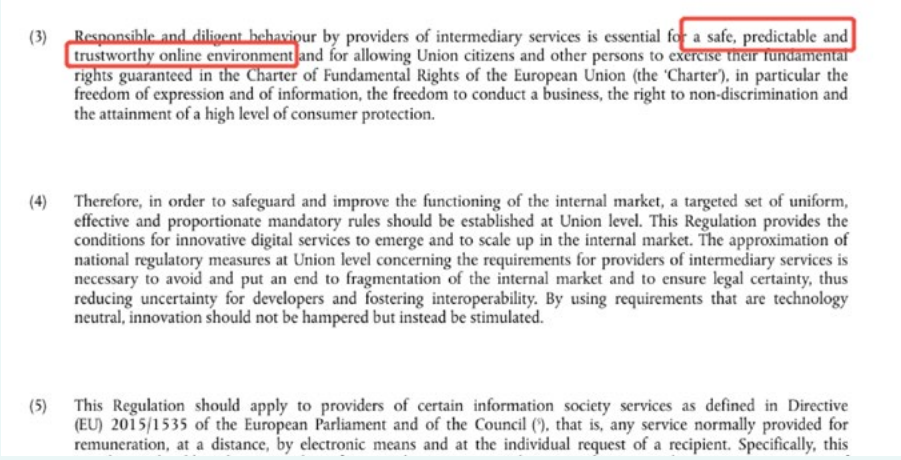
THE EUROPEAN PARLIAMENT AND THE COUNCIL OF THE EUROPEAN UNION,

Having regard to the Treaty on the Functioning of the European Union, and in particular Article 114 thereof,

Having regard to the proposal from the European Commission,

图源：欧盟 DSA 截图

1）《数字服务法》（DSA）：非法内容的"零容忍"规则



(3) Responsible and diligent behaviour by providers of intermediary services is essential for a safe, predictable and trustworthy online environment and for allowing Union citizens and other persons to exercise their fundamental rights guaranteed in the Charter of Fundamental Rights of the European Union (the 'Charter'), in particular the freedom of expression and of information, the freedom to conduct a business, the right to non-discrimination and the attainment of a high level of consumer protection.

(4) Therefore, in order to safeguard and improve the functioning of the internal market, a targeted set of uniform, effective and proportionate mandatory rules should be established at Union level. This Regulation provides the conditions for innovative digital services to emerge and to scale up in the internal market. The approximation of national regulatory measures at Union level concerning the requirements for providers of intermediary services is necessary to avoid and put an end to fragmentation of the internal market and to ensure legal certainty, thus reducing uncertainty for developers and fostering interoperability. By using requirements that are technology neutral, innovation should not be hampered but instead be stimulated.

(5) This Regulation should apply to providers of certain information society services as defined in Directive (EU) 2015/1535 of the European Parliament and of the Council (9), that is, any service normally provided for remuneration, at a distance, by electronic means and at the individual request of a recipient. Specifically, this

图源：欧盟 DSA 截图

DSA 适用于所有向欧盟用户提供服务的数字平台（包括社交平台、电商、搜索引擎等），按平台规模分级设定合规义务：

- ***普通平台：****需设立投诉处理系统，在用户举报后 7 日内处理争议内容，每年发布透明度报告；*
- ***超大型平台：****需开展年度风险评估、自费接受独立审计、向监管机构和研究人员共享数据，建立合规职能部门；*
- ***核心处罚：****非法内容未及时删除（明确违法内容需 24 小时内处理）、虚假信息传播等违规行为，最高可处全球年营业额 6% 的罚款。*

2）《数字市场法》（DMA）：平台垄断的"反垄断利剑"

DMA 针对"守门人平台"（年营业额 ≥75 亿欧元或市值 ≥750 亿欧元，月活用户 ≥4500 万）设定严格义务，禁止自我优待、限制用户卸载应用、滥用数据投放定向广告等行为，违规罚款最高为全球年营业额 20%，多次违规可被拆分。

3）电子隐私监管：ePrivacy 指令

作为 GDPR 在电子通信领域的补充，欧盟《电子隐私指令》（ePrivacy Directive 2002/58/EC）聚焦通信内容隐私与营销规范，虽正修订为《电子隐私条例》（ePrivacy Regulation），但当前指令仍为执法依据，核心要求包括：

> **Cookie 合规：**网站使用非必要 Cookie（如广告追踪）需获取用户"明确同意"，禁止默认勾选同意，需提供便捷的撤回同意选项；
>
> **通信保密：**禁止未经授权监听、记录电子通信（如邮件、即时消息），通信服务商需保障数据传输加密；
>
> **营销限制：**发送商业短信、邮件需事先获取用户同意，内容需标注发件人信息及退订方式，违规罚款最高达 2000 万欧元。

4）数字版权监管：《数字单一市场版权指令》

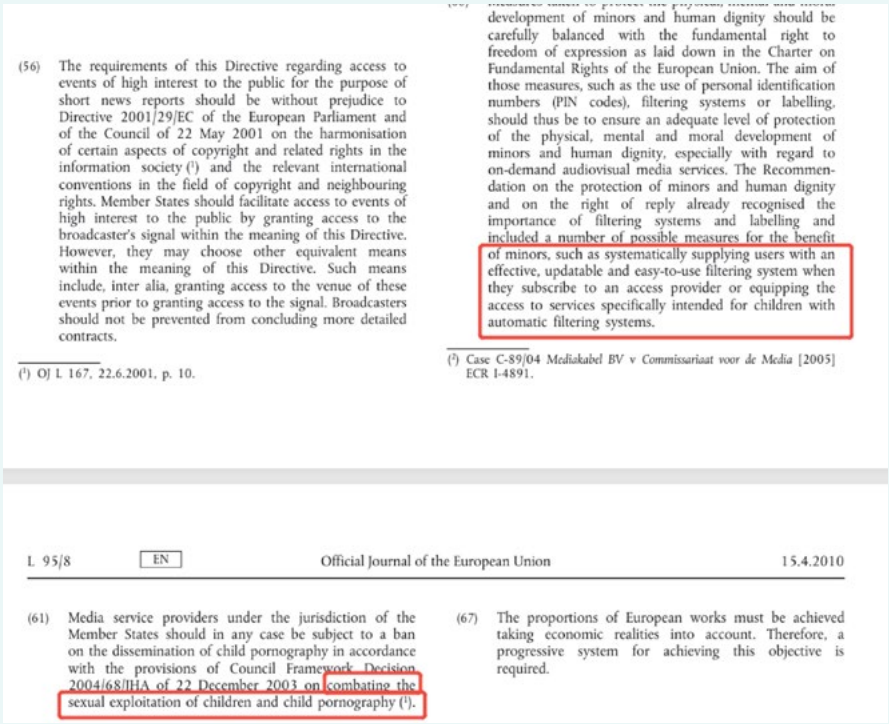
2019 年生效的《数字单一市场版权指令》（Copyright Directive）是欧盟版权监管的核心法规，规定了平台版权过滤义务和用户合理使用权益，重点要求：

> ***平台义务：****视频、音乐、文字内容平台需建立版权过滤系统，主动识别并移除侵权内容，未履行义务需承担侵权赔偿责任；*
>
> *网络内容共享服务提供商（如 YouTube）如果要向公众传播作品或为公众提供作品，则需要得到作者的事先授权（例如，可以通过许可协议获得）。该规则称为授权规则。此外，根据《版权指令》的规定，网络内容共享服务提供商对任何侵犯第三方版权的用户生成以及用户上传的内容负有主要责任，因为其行为构成向公众传播。*
>
> ***用户权益：****允许用户出于评论、讽刺等目的合理引用受版权内容，保障残疾人获取无障碍格式内容的权利。*

。

5）《视听媒体服务指令》（AVMSD）：视听内容的全场景规范

AVMSD 是欧盟监管视听媒体内容的核心法规，聚焦在禁止 / 限制内容、未成年人保护、广告规范及视频共享平台（VSP）的内容治理等领域，与 DSA 形成互补。



(56) The requirements of this Directive regarding access to events of high interest to the public for the purpose of short news reports should be without prejudice to Directive 2001/29/EC of the European Parliament and of the Council of 22 May 2001 on the harmonisation of certain aspects of copyright and related rights in the information society (1) and the relevant international conventions in the field of copyright and neighbouring rights. Member States should facilitate access to events of high interest to the public by granting access to the broadcaster's signal within the meaning of this Directive. However, they may choose other equivalent means within the meaning of this Directive. Such means include, inter alia, granting access to the venue of these events prior to granting access to the signal. Broadcasters should not be prevented from concluding more detailed contracts.

(1) OJ L 167, 22.6.2001, p. 10.

development of minors and human dignity should be carefully balanced with the fundamental right to freedom of expression as laid down in the Charter on Fundamental Rights of the European Union. The aim of those measures, such as the use of personal identification numbers (PIN codes), filtering systems or labelling, should thus be to ensure an adequate level of protection of the physical, mental and moral development of minors and human dignity, especially with regard to on-demand audiovisual media services. The Recommendation on the protection of minors and human dignity and on the right of reply already recognised the importance of filtering systems and labelling and included a number of possible measures for the benefit of minors, such as systematically supplying users with an effective, updatable and easy-to-use filtering system when they subscribe to an access provider or equipping the access to services specifically intended for children with automatic filtering systems.

(2) Case C-89/04 *Mediakabel BV v Commissariaat voor de Media* [2005] ECR I-4891.

L 95/8 EN Official Journal of the European Union 15.4.2010

(61) Media service providers under the jurisdiction of the Member States should in any case be subject to a ban on the dissemination of child pornography in accordance with the provisions of Council Framework Decision 2004/68/JHA of 22 December 2003 on combating the sexual exploitation of children and child pornography (1).

(67) The proportions of European works must be achieved taking economic realities into account. Therefore, a progressive system for achieving this objective is required.

图源：欧盟 AVMSD 关于未成年人保护条款截选

- **适用范围：**涵盖"线性服务"（如直播电视）和"点播服务"（如 Netflix、短视频平台），只要向欧盟用户提供服务，无论企业注册地是否在欧盟均需合规；
- **核心内容标准：**禁止传播仇恨言论、煽动暴力、歧视性内容，儿童节目需设置家长控制功能，避免未成年人接触暴力、色情内容；广告时长需符合比例要求（线性服务每小时广告不超过 12 分钟），禁止误导性广告；
- **处罚机制：**违规企业将被限制在欧盟境内提供服务，最高可处服务收入 10% 的罚款，多次违规将被吊销欧盟境内运营资质。

欧洲内容监管核心要点汇总

1. 监管逻辑：以“用户权利保障”为核心，坚持“风险前置、分级管控”，AI 与数据合规深度绑定，形成全生命周期监管；
2. 欧盟与各国关系：欧盟框架为统一底线，成员国可制定更严格的补充规则，非欧盟国家（如英国）需单独适配其独立体系；
3. 核心合规红线：禁止 AI 歧视、未成年人有害内容、仇恨言论（Hate Speech）与虚假信息（Disinformation）传播；严格遵守数据本地化、跨境传输及 Cookie 同意规则；落实网络安全 incident 上报、版权保护及电子商务信息披露义务；
4. 处罚力度：罚款上限普遍达全球年营业额 4%-7%，超大型平台面临额外的审计与拆分风险，合规成本极高。

中东北非：宗教教法主导下的文化禁忌与内容净化合规体系

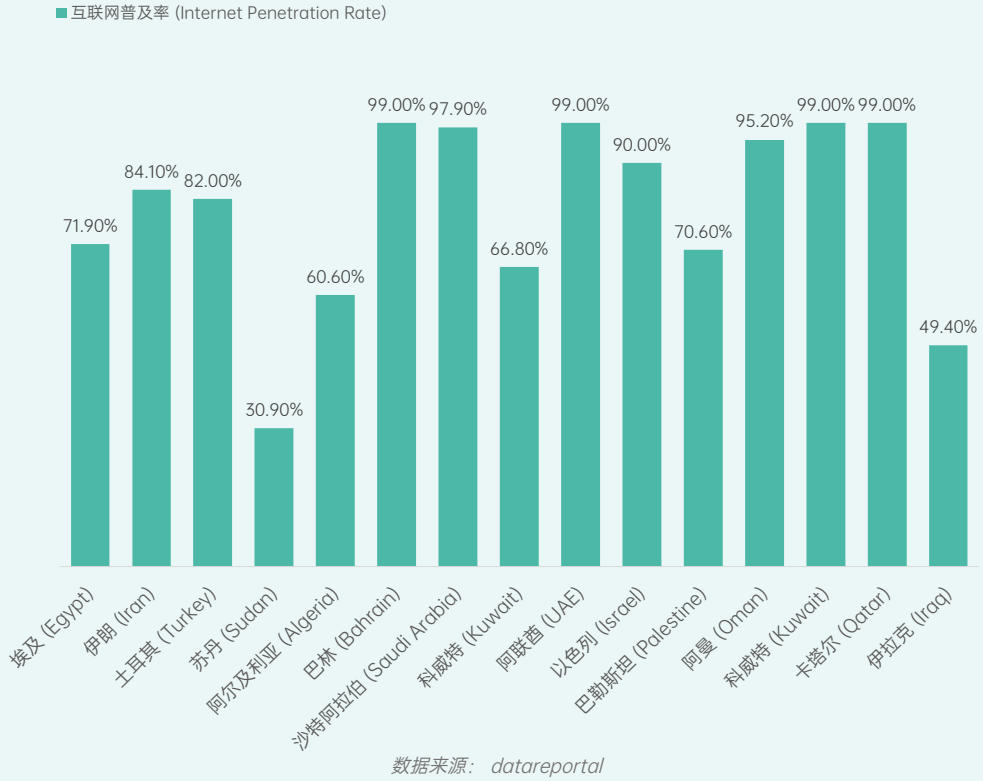
中东横跨欧亚非，22 个经济体覆盖约 1500 万平方公里，战略地位关键，更是我国“一带一路”倡议的重要合作区域。 区域内各国发展差异显著，本次聚焦四国核心样本：阿联酋与沙特作为 GCC（海湾合作委员会）核心成员国，是中东 AI 投资与基建“双引擎”，2024 年 AI 市场合计占区域 65%，政策落地性强，是科技出海核心目的地。土耳其坐拥欧亚地缘枢纽与 8500 万消费人口，全球游戏开发者的聚焦中心，是游戏、跨境电商的关键布局点，监管政策辐射中东欧及中亚；埃及作为中东第一人口大国与非盟核心，是“中东 - 非洲”双市场联动的天然跳板，政策动向反映非 GCC 国家监管共性。

阿联酋、沙特阿拉伯、土耳其、埃及这四个国家覆盖不同发展水平、地缘属性与监管模式，为企业出海提供全面合规参考 .

2025 年，中东地区正经历一场前所未有的 AI 革命。根据中研普华研究院报告，截至 2025 年，中东人工智能市场规模已突破 200 亿美元，预计到 2030 年将以 35% 的年均复合增长率扩张至 1000 亿美元。这一增速不仅远超全球平均水平，更将使中东成为全球 AI 市场增长的核心引擎之一。

中东各国政府通过制定国家战略、加大投资力度、推动国际合作等措施，为人工智能行业的发展提供了有力支持。阿联酋"面向未来 50 年国家发展战略"明确将 AI 列为经济多元化的核心支柱，沙特"2030 愿景"则计划在 AI 领域吸引 200 亿美元投资，并培训 2 万名专业人才。政策红利下，中东互联网普及率已达 75%，其中海湾六国的平均互联网渗透率超 90%， 其中阿联酋和沙特尤为靠前，分别达 99%、98%。作为最早推出商用 5G、世界最大的数据市场之一，中东地区拥有领先世界水平的互联网连接速度，为 AI 应用提供了海量终端设备支持。

部分中东国家互联网普及率



数据来源： datareportal

阿联酋

AI 发展战略与监管政策

阿联酋在 AI 领域的布局堪称全球标杆。2017 年推出的《2031 年阿联酋人工智能战略》是全球首个国家级 AI 战略，明确将 AI 对 GDP 贡献率提升至 12% 的目标。2025 年初发布的《2025—2027 年政府数字化战略》进一步计划投入 130 亿迪拉姆（约合 253 亿元人民币）推动 AI 在政府服务中的应用。

监管框架呈现创新与规范并重的特点。2024 年成立的人工智能和先进技术委员会（AIATC）负责制定 AI 研究、基础设施和投资的战略。区块链与 AI 委员会推出的整合框架要求所有 AI 模型必须在区块链注册表中登记，通过智能合约自动执行伦理准则，开创了 " 代码即监管 " 的新模式。

数据治理方面，阿联酋严格实施数据本地化政策。根据《2025—2027 年政府数字化战略》，所有 AI 训练数据必须存储在境内服务器，这一要求促使 G42 等本土企业与国际巨头合作建设符合规范的数据中心。高风险 AI 需通过 ISO 42001 认证，商业 AI 服务需取得媒体监管办公室（MRO）专项许可，核心用户数据需本地存储，跨境传输需完成 DPIA 评估。8 月 TDRA 推出“数字欺诈猎人”工具，用户可上传可疑信息快速核验，分析过程不收集个人数据，严格保护隐私。值得注意的是，阿布扎比国际金融中心（ADGM）为 AI 企业提供特殊许可，允许在自贸区范围内进行数据跨境流动试点。

伦理治理走在全球前列。阿联酋是首个采用联合国教科文组织 AI 伦理建议的中东国家，《人工智能伦理宪章》规定了公平、透明、问责等核心原则。迪拜卫生局推出的医疗 AI 政策要求所有诊断系统必须保留人工复核机制，算法决策需提供可解释的依据文档。

其他互联网监管政策

阿联酋的宗教系伊斯兰教，整个社会氛围及政府都偏向保守。例如，伊斯兰教禁止色情内容，因此阿联酋政府对色情内容的监管较为严格。根据 TDRA（阿联酋电信和数字政府监管局）的统计数据，2025 年二季度共封禁 7289 个违规网站，绝大多数是因为"毒品、赌博、知识产权、钓鱼欺诈、色情裸露"内容遭到封禁。阿联酋警方也曾逮捕在社交媒体发裸照的用户，称"（迪拜警方）严禁此类行为，它们不被社会所接受，与阿联酋的价值伦理相悖"。

禁止内容类别 Prohibited Content Category	被屏蔽网站数量 Number of Blocked websites	占总屏蔽网站的百分比 Percentage of total blocked websites
Bypassing Blocked Content (绕过被屏蔽内容)	6	0.08
Pornography nudity and vice (色情、裸体及不良行为)	217	2.98
Impersonation fraud and phishing (冒充、欺诈和网络钓鱼)	1930	26.48
Insult slander and defamation (侮辱、诽谤和中伤)	2	0.03
Invasion of Privacy (侵犯隐私)	1	0.01
Supporting criminal acts and skills (支持犯罪行为和技能)	0	0
Drugs (毒品)	2204	30.24
Medical and pharmaceutical practices in violation of the laws (违反法律的医疗和药品行为)	13	0.18
Infringement of intellectual property rights (侵犯知识产权)	2544	34.9
Discrimination racism and contempt of religion (歧视、种族主义和蔑视宗教)	0	0
Viruses and malicious programs (病毒和恶意程序)	0	0
Promotion of or trading in prohibited commodities and services (推广或交易被禁商品和服务)	3	0.04
Illegal communication services (非法通信服务)	110	1.51
Gambling (赌博)	2204	30.24
Terrorism (恐怖主义)	0	0
Prohibited TLD (被禁顶级域名)	0	0
Illegal Activities (非法活动)	154	2.11
Upon order from judicial authorities or in accordance with the law (根据司法当局命令或依照法律)	0	0
Offences against the UAE and public order (危害阿联酋和公共秩序的犯罪)	3	0.04
Total (总计)	7289	100%

2025 年二季度 TDRA 封禁网站统计　来源：TDRA 官网

阿联酋政府的执法行为其实有法可依。21 世纪以来，阿联酋政府就通过立法加大了对媒体领域的监管，包括互联网内容，并由上述提到的 MRO 和 TDRA 负责监管执法。

阿联酋的互联网内容监管法律可大致分为两类：一类是阿联酋的几部顶层法律，如《刑法》、《打击恐怖主义罪行法》、《反歧视法》，这类法律对阿联酋境内的所有犯罪行为进行监管，包括线上。

> *2015 年颁布的《反歧视法》规定，禁止任何形式（包括口头与书面文字、书籍、手册或在线媒体）的反宗教、侮辱先知、亵渎圣地的行为，违者将面临 6 个月至 10 年的监禁处罚，或处以 5 万至 200 万迪拉姆的罚款，或两者并罚。*
>
> *《阿联酋刑法》规定，禁止利用宗教引起骚乱或破坏国家统一，禁止诽谤、侮辱、冒犯伊斯兰教义以及宣扬其他宗教。*

另一类则是专门的互联网内容监管政策法规，比较重要的有《网络犯罪法》、《媒体内容标准》以及《互联网接入管理政策》这三部，它们以上述阿联酋顶层法律为立法基础，细致地规定了互联网平台的内容审核责任、违规内容分类、许可证申领等。

实际上，这三部法规对违规内容的分类基本一致，包括但不限于：

> *绕过封禁的内容；色情、裸体和恶习；冒充、欺诈和网络钓鱼；侮辱、诽谤和诽谤；侵犯隐私；侵犯阿联酋和公共秩序的内容；支持犯罪行为和为其提供技术的；毒品；违反法律的医疗和制药行为；侵犯知识产权；歧视、种族主义和蔑视宗教；病毒和恶意程序；促进或交易违禁商品和服务；非法通信服务；赌博；恐怖主义；封禁的顶级域名；非法活动；根据司法当局的命令或按照法律被禁止的内容。*

但是这三部法律的执法机构以及监管对象都有所差异。

《互联网接入管理政策》

《互联网接入管理政策》主要由 TDRA 负责监管执法，主要的监管对象是电信服务供应商。如前所述，TDRA 会要求当地电信服务供应商对网络上的违禁内容进行拦截。

《媒体内容标准》

《媒体内容标准》由 MRO 制定并负责执法，游戏、电影、社交平台、音视频等媒体内容都属于 MRO 的监管对象。MRO 负责评判这些平台、产品或出版物的内容是否符合《媒体内容标准》，

并有权决定它们是否可以在阿联酋正常使用。5 月 29 日起实施修订版《媒体内容标准》，盈利性内容创作者（含网红）需取得官方许可（费用起价 1.5 万迪拉姆），AI 生成内容需经中央 AI 分析系统预筛查，单次违规最高罚款 100 万迪拉姆，重复违规翻倍至 200 万迪拉姆或永久关停。

2022 年 6 月，MRO 宣布，皮克斯动画电影《光年正传》将不会在阿联酋的电影院放映，原因是因为这部电影"违反了该国的媒体内容标准"，外界猜测可能是因为影片中包含两个女性同性恋角色的亲吻画面。

宣布不会上映《光年正传》 *来源：MRO 推特*

《网络犯罪法》

2002 年，阿联酋通过了《网络犯罪法》，2014 年及 2021 年分别对其进行修订。《网络犯罪法》的监管范围更为广泛，除了违规内容，其他线上违法犯罪行为也受其管辖，其主要执法机构为警方。

新《网络犯罪法》规定禁止通过网络开展赌博、传播色情、教唆或诱拐他人卖淫等危害公共道德规范的行为。新法还对在网络上进行诽谤、亵渎宗教、人口贩卖、仇恨言论、种族主义、分裂主义、国家安全、公共秩序、武器贩卖、恐怖主义、麻醉品、洗钱等违禁行为作出详细规定。

> *例如，该法第 29 条规定："在网站或任何计算机网络或信息技术手段上发布信息、新闻、言论或谣言，意图讽刺或损害国家或其任何机构或其总统、副总统、酋长国的任何统治者、他们的王储或酋长国的副统治者、国旗、国家和平的声誉、威望或地位、其标志、国歌或其任何符号，属违法行为。"*

沙特阿拉伯

AI 发展战略与监管政策

沙特的 AI 雄心与其石油资本深度绑定。" 愿景 2030" 计划明确将 AI 作为经济多元化的核心引擎，目标是到 2030 年成为全球 AI 领域前 15 名领导者。2025 年 2 月，沙特通信和信息技术部宣布投入 109 亿美元用于 AI 基础设施建设和初创企业发展，这一数字相当于其当年非石油 GDP 的 3.2%。

2025 年 7 月推出 " 国家 AI 指数 "，主权基金投资 OpenAI 等企业。2024 年全球人工智能指数位列第 14 位，" 政府 AI 战略 " 指标全球第一，主导成立联合国科教文组织下属国际人工智能研究与伦理中心（ICAIRE）。

监管框架呈现原则先行的特点。沙特数据与人工智能局（SDAIA）2022 年发布的《人工智能道德准则》提出公平、隐私与安全、人类中心等七项核心原则，要求所有 AI 系统在全生命周期贯彻这些准则。 2025 年 4 月由沙特阿拉伯通信、空间与技术委员会（CST）推出《全球人工智能

中心法》草案，旨在建立数据主权中心并寻求国际意见。这一草案不仅揭示了沙特在数据治理方面的雄心壮志，更预示着全球数据治理格局的重大调整。2025 年 6 月发布的《AI 相关作品版权原则》进一步明确：完全由 AI 生成的内容不享有版权保护，只有包含足够人类创造性输入的作品才能获得保护。

其他互联网监管政策

沙特一直以来都秉持着严格的内容审查方针，并通过立法与执法来强化媒体内容监管。随着媒介形式的不断发展演进，沙特的立法也在不断完善。

上世纪 90 年代初公布的《基本法》仅对媒体内容做了非常宽泛的规定；到 21 世纪初，针对传统媒体的专项法律出台，纸质媒体的监管开始收紧；2008 年，《反网络犯罪法》出台，标志着互联网正式成为沙特政府的监管目标；到 2017 年制定《视听媒体法》，将新兴的社交媒体、游戏等用于休闲娱乐的互联网产品纳入监管范围。

《沙特阿拉伯政府基本法》

沙特为政教合一的君主制国家，无成文宪法，1992 年公布的《基本法》是沙特阿拉伯一部类似于宪法的法律，其规定《古兰经》和先知穆罕默德的圣训是国家执法的依据，且《基本法》不能凌驾于伊斯兰教法之上。

在沙特，任何西方商业法律或制度都是从伊斯兰教法的角度进行调整和解释的。伊斯兰教法规范很大程度上被视为沙特人的社会规范，并被严格遵守，如禁止饮酒、同性恋行为、赌博、色情等。

《基本法》也对媒体内容进行了规定，其第三十九条写道，“大众公共传媒及其它各类媒体应当使用文明用语、遵守国家法律，促进国民教育、维护国家统一。禁止一切煽动叛乱和分裂、危害国家安全、破坏国家与公众关系、侵犯他人人格尊严与权利的行为和言论。有关出版及言论的相关事宜由法律另行规定。”

这一条也被视为沙特政府开展互联网审查最主要的法律依据。

《新闻与出版法》

2000 年沙特政府制定《新闻与出版法》，针对出版物的流通及其内容展开全面监管，包括印刷品、书店、外媒、纸质新闻、电视广播。该法规中有一些规定对自由言论做出了限制。

> *例如，所有形式的媒体必须在信息部获取牌照后方能正常运营；言论自由不得超出伊斯兰教法和法律所规定的界限；禁止出版侮辱沙特政府体制的印刷内容，等等。违者可面临最高 6 万里亚尔的罚款或暂时关停机构两个月，严重可永久关停。*

为了更好地适应社交媒体的发展，2011 年，沙特政府为重新修订《新闻与出版法》而制定了《电子出版执行规定》，将社交媒体内容也纳入监管范围。

《反网络犯罪法》

2007 年沙特开始施行《网络犯罪法》，其列举了沙特地区可能导致监禁和罚款的网络犯罪，其中有关网络违规内容的违法行为，尤其是涉及恐怖主义和国内外安全、国家经济的行为，将面临最高等级的处罚。

该法律第六条明确规定，禁止通过信息网络或计算机生产、准备、传播或储存以下内容，违者将面临最高五年的监禁，或者最高 300 万里亚尔的罚款，或两项并罚：

> - *侵犯公共秩序、宗教价值观、公共道德和隐私产生负面影响的内容；*
> - *在信息网络或计算机中搭建或宣传人口贩卖网站；*
> - *筹备、出版以及推广违背公共道德的色情、赌博网站；*
> - *在信息网络或计算机中搭建或宣传有关毒品、精神药品买卖的网站。*

而据该法律第七条规定，涉及恐怖主义，以及危害国内外安全或国家经济的网络行为将会面临最高 10 年监禁，或者最高 500 万里亚尔的罚款，或两项并罚。

《视听媒体法》

2017 年，《视听媒体法》出台，沙特政府又进一步加强了对所有媒体行业的监管，包括传统媒体和社交媒体，其规定：

> *开展任何音视频活动都应向 GCAM 申请运营执照，包括媒体内容生产制作、广告、影院、广播、IPTV、游戏等。*
>
> *音视频活动运营者有义务配合监管部门执法，如保存所有媒体内容记录 90 天，并按要求向 GCAM 提供这些记录数据。*
>
> *音视频活动运营者需要遵守当地的"内容方针"，如应尊重阿拉、《古兰经》、先知及相关人物；禁止歪曲伊斯兰教义；对国王及王子表示敬意；避免扰乱公共秩序、国家安全和公共利益；禁止讨论可能引起冲突、分裂、仇恨的话题，或挑起暴力，威胁安全等。*
>
> *禁止讨论可能伤害与他国外交关系的话题，以及可能煽动恐怖主义、威胁安全的话题。*
>
> *禁止传播有碍公共道德的内容，或其他裸体、不雅着装、挑起性冲动、使用粗俗语言的内容。*

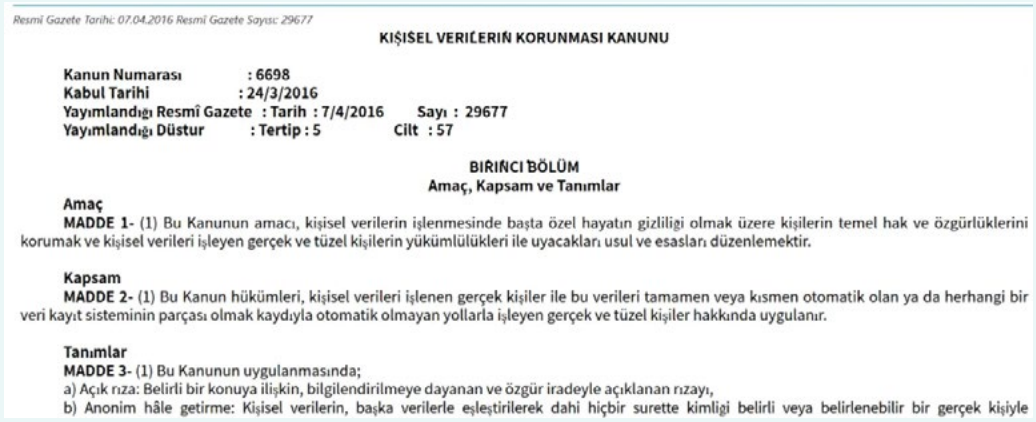
同时，沙特的数字身份体系建设成效显著，"Absher" 平台已整合生物识别技术，提供从出生登记到商业注册的全生命周期服务。2025 年新增的 AI 监控模块可自动识别异常交易和可疑行为，使金融欺诈率同比大幅下降。这种 " 数字治理 +AI 赋能 " 的模式正成为沙特互联网监管的鲜明特色。

土耳其

AI 发展战略与监管政策

土耳其的 AI 战略呈现务实追赶特征。2025 年启动的 " 国家人工智能转型计划 " 虽然没有设定沙特、阿联酋那样的宏大目标，但聚焦医疗、农业等民生领域的 AI 应用，计划到 2030 年培养 5 万名 AI 专业人才，建立 20 个国家级 AI 研究中心。这种 " 小而美 " 的发展路径使其在有限资源下取得了显著进展。

监管框架仍在完善过程中。土耳其目前尚未出台专门的 AI 法律，土耳其 2025 年 1 月成立议会研究委员会，5 月完成 AI 监管研究，但未提交正式《AI 法案》草案，目前主要依托《个人数据保护法》和《信息和通信技术法》进行监管。2025 年 4 月实施的《远程通信商品流通监管条例》要求所有 AI 产品必须通过土耳其技术法规认证，安全警示标识必须使用土耳其语，这一规定增加了外国 AI 企业的合规成本。



Resmî Gazete Tarihi: 07.04.2016 Resmî Gazete Sayısı: 29677

KİŞİSEL VERİLERİN KORUNMASI KANUNU

Kanun Numarası : 6698
Kabul Tarihi : 24/3/2016
Yayımlandığı Resmî Gazete : Tarih : 7/4/2016 Sayı : 29677
Yayımlandığı Düstur : Tertip : 5 Cilt : 57

BİRİNCİ BÖLÜM
Amaç, Kapsam ve Tanımlar

Amaç
MADDE 1- (1) Bu Kanunun amacı, kişisel verilerin işlenmesinde başta özel hayatın gizliliği olmak üzere kişilerin temel hak ve özgürlüklerini korumak ve kişisel verileri işleyen gerçek ve tüzel kişilerin yükümlülükleri ile uyacakları usul ve esasları düzenlemektir.

Kapsam
MADDE 2- (1) Bu Kanun hükümleri, kişisel verileri işlenen gerçek kişiler ile bu verileri tamamen veya kısmen otomatik olan ya da herhangi bir veri kayıt sisteminin parçası olmak kaydıyla otomatik olmayan yollarla işleyen gerçek ve tüzel kişiler hakkında uygulanır.

Tanımlar
MADDE 3- (1) Bu Kanunun uygulanmasında;
a) Açık rıza: Belirli bir konuya ilişkin, bilgilendirilmeye dayanan ve özgür iradeyle açıklanan rızayı,
b) Anonim hâle getirme: Kişisel verilerin, başka verilerle eşleştirilerek dahi hiçbir surette kimliği belirli veya belirlenebilir bir gerçek kişiyle

图源：土耳其《个人数据保护法》截图

数据本地化要求严格。根据 2024 年修订的《个人数据保护法》，处理超过 100 万用户数据的 AI 系统必须在土耳其境内设立数据中心。违反这一规定的企业将面临最高 4% 全球营业额的罚款，谷歌、Meta 等公司已因此调整了其中东数据存储策略。

伦理治理注重文化适应性。土耳其科学技术研究理事会（TÜBİTAK）2025 年发布的《AI 伦理指南》特别强调算法需尊重当地文化价值观，禁止在产品推荐、内容过滤等场景中引入宗教或政治偏见。这一要求反映了土耳其作为欧亚文化交汇点的特殊考量。

其他互联网监管政策

土耳其的互联网监管以严格著称。2025 年，土耳其信息和通信技术管理局（BTK）共封禁 311,091 个网站，其中 82% 的禁令由 BTK 主席直接下令。值得注意的是，2025 年 9 月，土耳其宣布封禁 Tango、Azar 等 29 款社交应用，理由是涉及 " 不当内容和未成年人保护问题 "，这一举措影响了约 500 万土耳其用户。

图源：土耳其媒体 WebTekno

土耳其主要通过《5651 号互联网法》对线上内容进行监管，其他的例如《刑法》以及《反恐法》则对普遍的违法行为进行了规定，也适用于线上。

《5651 号互联网法》

《5651 号互联网法》（简称《互联网法》）是土耳其首部对互联网内容进行监管的法律，也是土耳其目前管理互联网内容最主要的法律。

《互联网法》于 2007 年首次通过实行，并于 2014 年及 2020 年经过两次较大的修订。《互联网法》旨在保护儿童及家庭免受网络有害内容的侵扰，包括教唆自杀、性骚扰儿童、教唆吸毒、提供有害健康的药物、淫秽、卖淫、赌博等，并对互联网内容提供商、第三方平台、接口提供商的义务与责任都做了较为详细的规定。

1) 互联网内容提供商的责任

内容提供商是指自己生产、编辑或提供所有类型的信息或数据的平台。《互联网法》规定，内容提供商应为其在网络上提供的所有内容负责。

2) 第三方平台的责任

第三方平台（Hosting provider）是指提供或运营托管服务的提供商，类似 Twitter、Facebook 这类社交媒体，允许第三方在其平台上提供服务或内容。

《互联网法》规定，平台无需审核用户在平台上发布的内容，但要注意的是，平台有义务配合监管机构的审查及内容删除要求，并且应保存相关数据至少一年，最长不应超过两年。若违反相关规定，土耳其总统有权对企业处以 10 万 -100 万利拉的行政罚款。

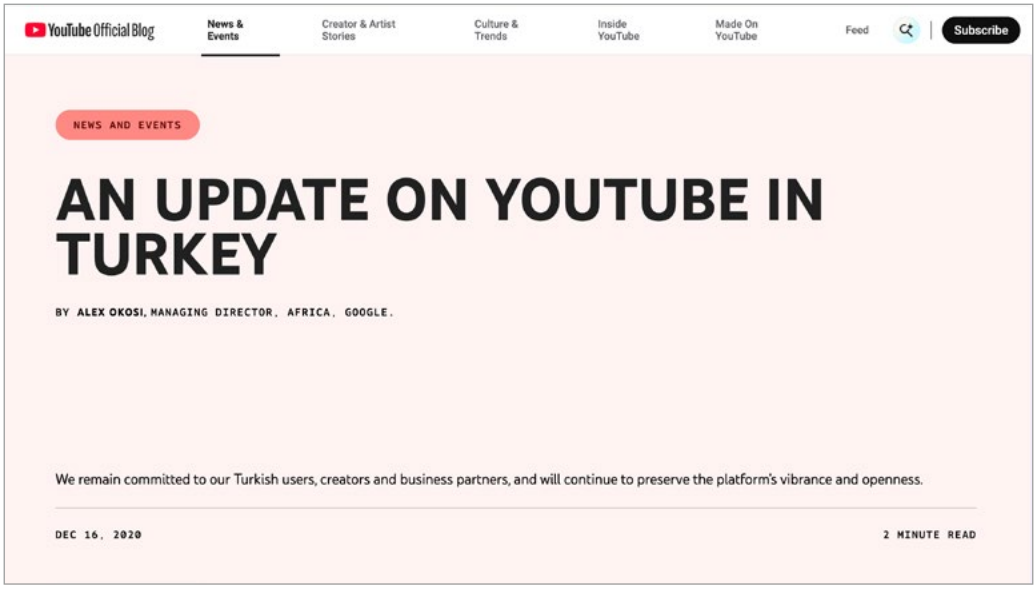
根据 2020 年最新通过的修正案，日活超过 100 万的国外社交媒体企业应在土耳其当地派驻至少一名工作人员，且必须为土耳其人，或在当地指派一个代理商，负责处理《互联网法》所规定的义务与责任执行工作，包括但不限于内容投诉。该人的联系方式必须告知监管机构，且必须在其网站上列出。如果企业未在收到政府通知 30 日内完成人员派驻，将会面临 1000 万利拉的罚款；若多次提醒均未果，企业会面临罚款、广告禁令以及削减 50%-90% 带宽等处罚。

在修正案通过后，土耳其当局在 2020 年对 Twitter、Facebook、YouTube、Instagram 和 TikTok 各罚款 4000 万利拉（约合 510 万美元），原因是它们没有按规定任命当地代表。之后，Facebook、TikTok 和 YouTube 已经在该国建立了所需的法人实体。

在过去的十几年里，YouTube 多次遭到了土耳其政府的排挤。自 2007 年以来，该网站因各种原因多次被禁和解禁。2020 年，土耳其修改了第 5651 号互联网法令，其中对社交媒体平台上的内容将被删除进行了规范。经过分析和验证，YouTube 同意将遵守修改后的法律。

> ***YouTube 表示：****“在过去的几个月里，我们一直在深入分析土耳其最近修订的第 5651 号互联网法令，该法令对社交媒体平台提出了新的要求，如任命当地代表、遵守删除 & 响应周转时间以及发布半年一次的透明度报告等。YouTube 尊重我们运营所在国家的法律和法规，同时坚持我们对言论自由、信息获取和透明度的承诺。我们已经找到了前进的方向，并将在不损害我们价值观的前提下，依法开始任命当地代表法律实体的程序。”*

YouTube 在官方博客上发布更新状态，表明其土耳其用户和内容创作者对该平台很重要，它将确保保持一个开放的平台。



图源：网络

《互联网法》规定第三方平台的其他责任还包括：

- *要求 48 小时内响应政府的内容删除请求，24 小时内响应用户个人的内容删除请求。如拒绝删除相应内容，应提供拒绝理由（500 万罚款）。*
- *社交媒体企业应每半年提供一份内容审查统计报告（1000 万罚款）。*
- *将用户数据储存在土耳其境内（尚不清楚具体执行情况）。*

3）接口提供商的责任

接口提供商无需对使用其接口的内容进行审查，但其应配合监管部门要求，封禁违规内容的接口；保存相关数据至少半年，最长不超过两年。

此外，2014 年 2 月的互联网法修正案还规定了第三方平台必须尽快删除内容的特殊情况，其规定，只要有任何个人或法人指控侵犯隐私，或网站内容被认为"歧视或侮辱特定社会成员"，相关机构机构便有权发出行政屏蔽令，互联网服务提供者必须在接到命令后 4 小时内屏蔽特定网址。尽管行政屏蔽令必须在 48 小时内交由法院复核，但屏蔽条件如此宽泛、模糊，为侵权适用与解释保留了自由裁量空间。

刑法第 299、301 条

土耳其刑法第 299 条规定，"侮辱土耳其总统将面临 1-4 年的监禁处罚"、"如果犯罪是公开实施的，则处罚将提高六分之一"。

第 301 条规定，"诋毁土耳其国家、共和国政府、大国民议会以及司法机构将面临 6 个月 -2 年的监禁"、"公开诋毁军方和警方也会面临同等刑罚"。

此前土耳其检察官遭到两名恐怖分子被劫持的照片在 Twitter 和 YouTube 上流传，土检察机关下令，全国上下封锁 Twitter 和 YouTube。在两家平台同意删除照片后，土耳其政府才对其解禁。

土耳其

AI 发展战略与监管政策

埃及正凭借区域领导力崛起为非洲 AI 中心。2025 年 1 月发布的《国家人工智能战略 2025-2030》是非洲最全面的 AI 政策框架之一，聚焦治理生态、科研人才、产业应用、国际合作四大支柱，计划建立国家 AI 委员会及监管沙盒。9 月推出 "人工智能大使" 计划，已培训超 1100 名司法人员，推广 AI 在案件审理、文本分析中的应用，目标 2030 年培养 3 万名 AI 工程师。

图源：埃及《国家人工智能战略 2025-2030》封面

国际合作是埃及 AI 战略的特色。2024 年塞西总统访华期间，双方签署 27 项合作协议，其中多项涉及数字经济领域。2025 年 6 月，中埃签署新一轮双边合作协议，特别强调加强 AI 政策协同与技术联动，标志着两国 AI 合作进入新阶段。这种开放姿态使其在有限资源下实现了技术追赶。监管框架强调包容性发展。2023 年发布的《埃及负责任 AI 宪章》提出五项核心原则：人类中心设计、透明度、公正、问责和安全。这种以伦理为导向的治理思路，帮助埃及在联合国教科文组织 AI 伦理评估中排名非洲第一。

其他互联网监管政策

自 2014 年塞西担任埃及总统以来，埃及政府从设立监管机构与立法两方面入手，逐步加大了对互联网的监管与控制。

在设立监管机构上，埃及政府于 2016 年设立了媒体管理最高委员会（SCMR）、国家新闻管理局和国家媒体管理局，通过这三个专门机构实现对埃及境内媒体的全面监管。

在立法上，埃及通过了多部通用或专门法律，使互联网行业执法变得有法可依，其中对互联网内容规定比较重要的是《反网络及信息技术犯罪法》、《新闻媒体管理法》以及《反恐怖主义法》。

《反网络及信息技术犯罪法》

《反网络及信息技术犯罪法》于 2018 年 8 月正式实施，是埃及首部规范互联网、社交媒体和网络服务提供者的法律，旨在打击极端分子和恐怖组织利用互联网传播极端思想。

该法律一共有 45 条，列举了各类型网络犯罪并细化了量刑的程度，对网络服务提供者的责任与义务提出了明确要求。

网络及通信服务供应商必须连续留存 180 天的详细系统记录，并有义务向相关政府监管机构提供特定用户及 IP 地址的数据信息，以应对网络犯罪和恐怖行动。

若网站被发现含有任何违反埃及现行法律、危害国家安全和经济的内容，调查机构可向法院提交封禁令，法院应在 72 小时内作出裁决。紧急情况下，调查机构可直接通知国家电信管理局，立刻要求服务商暂时封禁违规内容。服务商如果拒绝封禁违规内容，将会被处以一年以上有期徒刑，或者 50 万到 100 万埃及镑，或者两项并罚。

《新闻媒体管理法》

2018 年 8 月，埃及正式实施《新闻媒体管理法》，将新兴媒体也纳入监管范围，包括所有音视频节目、电影、广告、报纸，以及拥有超过 5000 名粉丝的个人网站、个人博客或个人电子账号。该法规定的媒体、社交平台等机构义务如下：

> ***禁止发布违规内容：****该法禁止新闻、媒体、网站出版或广播任何违反埃及宪法、职业道德、公共秩序和公共道德的内容；禁止教唆违法行为，或煽动歧视、暴力、种族歧视、仇恨以及极端主义；禁止发布虚假新闻、任何危害国家安全的内容。《新闻媒体管理法》对上述违规内容并没有给出明确的定义，监管机构媒体管理最高委员会也因此被赋予较大的自主裁量权。*
>
> ***许可证注册：****任何媒体机构或网站必须在获得 SCMR 许可证后，方可在埃及运营。*
>
> ***处罚：****第 17 条规定，散布虚假新闻、谣言，号召、煽动违法、暴力、仇恨、歧视、宗派主义、种族主义，或者威胁国家统一的，或者冒犯国家机关、损害国家利益的，或煽动公众、侮辱他人，或在未验证其真实性的情况下通过社交媒体传播信息，SCMR 可罚款 100 万埃及镑，或封禁页面。第 22 条规定，对不符合新闻媒体规则的军事、安全行动或恐怖事件的报道，SCMR 可以禁止相关内容的出版或广播，甚至采取临时封禁，暂停或取消许可证的处罚措施。*

相较《反网络及信息技术犯罪法》，《新闻媒体管理法》则进一步对媒体内容进行了限制，并给出了具体的日常监管和处罚措施，需要出海企业重点关注。

《反恐怖主义法》

因埃及地区经历多次政变、且极端势力、恐怖主义活动较多，埃及对涉及暴力犯罪、恐怖主义活动的监管在塞西政府组建后更为严格。2015 年，埃及通过《反恐怖主义法》，为埃及境内密集的反恐活动提供法律支持。《反恐怖主义法》明确提到，通过技术手段或互联网向恐怖分子、组织提供任何形式的帮助也被视为恐怖活动。

《反恐怖主义法》规定，任何人建立或使用通讯网站、网站或其他媒体宣扬恐怖主义思想或内容，旨在误导安全当局、影响任何恐怖犯罪的司法程序、在恐怖团体或其成员之间发布任务、或交换有关恐怖分子在国内外的行动或移动的情报，应处以五年以上苦役监禁。

任何人故意以任何方式发布、广播、展示或宣传关于国内恐怖主义行为或反恐行动的虚假新闻或谣言，违反国防部发表的官方声明，在无偏袒法定纪律处分之前提下，应处以不少于 20 万埃及镑和不多于 50 万埃及镑的罚款。

埃及律师穆罕默德 · 拉马丹曾因因侮辱总统和粗暴侵犯言论自由被判处 10 年徒刑，该判决依据埃及反恐法第 29 条规定，如果创建社交媒体账户宣传"恐怖主义"活动或损害国家利益，可判处最高 10 年监禁。

中东北非合规要点汇总：

1. **阿联酋：AI 领域需遵守数据本地化要求与 ISO 42001 认证，AI 生成内容需经中央系统预筛查。互联网内容严禁色情、亵渎宗教等违规信息，盈利性内容创作者需取得官方许可。**
2. **沙特阿拉伯：AI 系统需贯彻七项伦理原则，纯 AI 生成内容不享有版权保护。互联网内容需符合伊斯兰教法，禁止恐怖主义、色情、侮辱国家等内容，音视频活动需申请运营执照。**
3. **土耳其：AI 产品需通过本地技术法规认证，处理超百万用户数据需设立境内数据中心。互联网平台需 48 小时响应政府内容删除请求，日活超 100 万的外资平台需派驻本地代表。**
4. **埃及：AI 发展遵循负责任伦理宪章，聚焦治理与国际合作两大核心。互联网严禁传播恐怖主义、虚假新闻等内容，媒体机构及粉丝超 5000 的账号需取得运营许可证。**

不同地区相关合规法规附表

区域市场	特定国家	监管政策与法规	合规重点
东南亚	新加坡	《生成式AI治理模型框架》、《互联网操作规则》、《防止网络假信息和网络操纵法案》	价值观导向问题（仇恨言论、假新闻）、内容实质错误、AI幻觉
	越南	《数字技术产业法》、《网络安全法》、72号法令、刑法“反国家宣传罪”	反政府内容（绝对红线）、数据本地化、人类价值和社会道德
	泰国	《计算机犯罪法》、刑法“冒犯君主罪”、《1935赌博法》	冒犯君主（绝对红线）、反政府、赌博、透明度和责任
	马来西亚	《通信与多媒体法令》(CMA)、《煽动叛乱法》、《1953住家公开赌博法》	反感内容（色情、LGBT+）、煽动叛乱（王室、宗教、种族）、赌博、透明度与告知义务
	印度尼西亚	《反色情法》、《部颁条例5》(MR5)、《个人数据保护法》	色情内容（零容忍）、亵渎宗教、种族主义、假新闻、平台注册与数据接口提供
	菲律宾	《网络犯罪预防法》、《反儿童色情法》、《数据隐私法》	儿童色情（零容忍）、网络诽谤、赌博、恐怖主义内容
北美	美国	《儿童在线安全法案》(KOSA)、《儿童在线隐私保护法》(COPPA 2.0)、《删除法案》、各州AI法案	未成年人保护（绝对红线，含AI生成内容）、仇恨言论、枪支毒品、算法歧视
	加拿大	《在线危害法案》(C-63)、《人工智能与数据法案》(AIDA)、《在线流媒体法案》	7类危害内容（特别是儿童性剥削，需24小时删除）、仇恨言论（入刑）、本土内容保护（营收5%）
南美	巴西	《巴西人工智能法案》、《通用数据保护法》(LGPD)、《巴西互联网法》	儿童色情（零容忍）、种族歧视（入刑）、假新闻、数据隐私（LGPD）、必须配合司法命令（否则封禁）
欧洲	欧盟	《数字服务法》(DSA)、《通用数据保护条例》(GDPR)、《AI法案》	GDPR数据隐私、非法内容（特别是仇恨言论）、算法透明度、未成年人保护、风险评估报告
中东北非	阿联酋	《2031年阿联酋人工智能战略》、《反歧视法》、《媒体内容标准》、《互联网接入管理政策》	宗教与道德（色情、LGBT+、饮酒、赌博、亵渎神明）、政治敏感、数据本地化
	沙特阿拉伯	《政府基本法》（基于伊斯兰教法）、《反网络犯罪法》、《人工智能道德准则》、《AI相关作品版权原则》	伊斯兰教法至上（禁止饮酒、赌博、色情、LGBT+，违反教义即违法）。政治敏感（严禁侮辱国王、王储及国家体制，严禁煽动分裂）
	土耳其	《5651号互联网法》（及修正案）、《个人数据保护法》（KVKK）	本地代表制度、侮辱国家罪、数据本地化
	埃及	《反网络及信息技术犯罪法》、《新闻媒体管理法》、《反恐怖主义法》、《国家人工智能战略2025-2030》	反恐与假新闻、国家安全（禁止发布危害国家经济、破坏国家统一的内容，调查机构可直接申请封禁）

03

AI 时代泛娱乐出海：垂直场景的合规挑战与解决方案

Global Expansion of Digital Entertainment in the AI Era:
Compliance Challenges and Vertical-Specific Solutions

游戏行业：应用内内容与生态的双重风控挑战

在全球数字文化深度交融的浪潮中，游戏凭借独特的叙事魅力突破地域界限，成为联结亿万玩家情感的重要纽带。2019 至 2024 五年间，中国自研游戏的海外销售收入稳定突破 150 亿美元，向 200 亿美元的新台阶迈进，这不仅是产业规模实现量级跨越的直接印证，更标志着中国游戏产业已完成从产品"走出去"到构建"全球影响力"的质变升级。

尽管中国游戏出海成绩亮眼，但全球市场的复杂性也带来了多重风险挑战，主要集中在内容安全与应用内生态安全两大维度。

合规与风控挑战

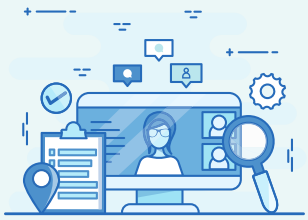
内容安全挑战：合规红线不可逾越

内容安全是游戏出海的首要门槛，受多语言环境、应用商店规则及地区法规差异影响，风险呈现多样化、复杂化特征。

多语言环境下的基础内容安全威胁

多语言环境是游戏出海内容安全的首要"试炼场"，不同语言的语义差异、场景属性与地区规范的叠加，让基础内容风控面临双重难题。

首先是场景的多维风险暴露，主要集中在公开场景与私密场景两大维度：



公开场景

涵盖个人资料（昵称、头像、简介）、世界频道等图文交互场景，易出现违禁、低俗、色情、涉政等基础违规内容。例如用户昵称可能暗藏谐音违禁词，头像或世界频道的图文可能包含宗教敏感元素，这类内容因传播范围广、曝光度高，极易触发平台审核与玩家反感。

私密场景

以用户私信为核心，风险针对性更强，常见广告导流、辱骂、色情诱导等行为，甚至形成"游戏内诱导加群→第三方平台诈骗"的完整黑产链路，直接导致玩家财产损失与品牌信任危机。

其次是识别与合规的双重挑战：

1. ***识别难度高***——*场景多样化且数量巨大，文本、图片等多模态内容分散在公开、私密场景中，信息量庞大且对审核的实时性要求极高；AI 技术的迭代让违规内容变种繁多，大模型生成的违禁内容可绕过传统关键词过滤，大幅增加审核成本；跨平台诈骗风险加剧，黑产通过游戏内诱导加群，转向第三方平台实施诈骗，玩家财产损失频发，严重动摇品牌信任根基。*

2. ***不同地区法规和文化差异***——*法律法规层面，各国对言论自由、非法内容的定义差异大，例如美国对言论的包容度与沙特阿拉伯的宗教言论管制形成鲜明对比，平台在全球运营中需同时应对互相冲突的法律义务，合规风险严峻；文化认知层面，不同文化对"有害"内容的感知存在根本差异，如 LGBTQ 宣传在部分地区是多元文化体现，在另一些地区则属于敏感内容，宗教敏感性、色情定义的差异更是巨大，这要求内容审核标准必须依据地区文化动态调整。*

海外应用商店内容审核规范

应用商店是游戏出海的"入场通道"，Google Play、App Store 等主流平台均有严格的内容审核标准，违规将面临下架、账户终止等严重后果。

应用商店	应用商店内容审核要求	违反后主要处罚
Google Play	• **儿童安全**：严禁CSAM、儿童诱骗、未成年人性化和性勒索等内容； • **不当内容**：严禁露骨性内容或色情材料（NSFW）、仇恨言论、暴力、暴力极端主义、敏感事件、危险产品（爆炸物、枪支、弹药、大麻/烟草和酒精 • **金融服务**：严禁欺骗性或有害金融产品和服务 • **非法活动**：严禁促进或宣传非法活动（非法毒品的销售和制造）	• **拒绝**：更新版本无法上架 • **移除**：应用及所有旧版本移除，无法下载 • **账户终止**：所有应用被移除，开发者无法发布新应用
App Store	应用程序不得包含**冒犯性、令人不安、意图引起厌恶或毛骨悚然的内容**，包括： **诽谤/歧视/恶意内容、暴力、武器和危险物品、露骨性内容或色情材料（NSFW）、煽动性宗教评论、虚假信息或功能**	• **应用移除**：从AppStore中移除 • **开除Apple开发者计划** • **账户终止**
Amazon	应用程序应为"家庭友好"，不得包含/宣传： **露骨性内容、仇恨或骚扰内容、医疗内容**：譬如生成某些药丸可治愈所有疾病、预测**胎儿性别、酒精、烟草**（包括电子烟）和**大麻、敏感事件、比特币和其他加密货币、非法和有害内容、暴力、恐怖主义和暴力极端主义**	**拒绝或下架**
Shopify	应用程序中不得包含**不当内容**，包括： **诽谤与粗俗、煽动性内容、非法活动、性内容、暴力、酒精/烟草与毒品**	• **停止应用发布** • **暂停应用发布或移除国家/地区** • **账户移除或撤销** • **终止协议**

Google Play：

- **审核核心要求：**严禁儿童性虐待材料（CSAM）、儿童诱骗、露骨性内容、仇恨言论、暴力极端主义、危险产品（枪支、爆炸物）、非法金融服务及毒品宣传等。

- **违规处罚：**分三级——拒绝版本更新、移除应用及所有旧版本、终止开发者账户（无法发布新应用）。

App Store：

- **审核核心要求：**禁止冒犯性、令人不安的内容（如诽谤歧视、暴力血腥、宗教煽动）、露骨性内容、虚假功能，且需符合"家庭友好"原则，严禁烟草、大麻、加密货币（如比特币）宣传等。

- **违规处罚：**应用从商店移除、开除 Apple 开发者计划、终止开发者账户。

- **其他平台共性要求：**均严禁酒精、烟草、暴力恐怖主义、非法内容，部分平台对医疗内容（如虚假治病宣传、胎儿性别预测）额外限制，违规可能面临"暂停应用发布""撤销账户"等处罚。

根据全球数据安全平台 Pixalate 发布的报告，2025 年第二季度，Google Play 下架应用 16 万款，同比 2024 年 Q2 的 110 万款降低 85.59% 。这个数字背后，藏着出海的积极信号：Google Play 的上架政策在放松。

但政策放松不等于可以为所欲为。从下架应用的类型分析来看，以下几个领域依然是"重灾区"：

> ***广告欺诈类违规：****虽然整体下架量减少，但广告相关违规依然占比较高。特别是虚假点击、恶意跳转等行为，依然是 Google Play 重点打击的对象；*
>
> ***权限滥用：****过度申请用户权限，特别是与应用功能不符的敏感权限申请，依然容易被标记；*
>
> ***恶意代码：****任何形式的恶意代码都是零容忍；*
>
> ***内容违规：****涉及暴力、色情、政治敏感等内容的应用，依然是严格管控的范围。*

从被下架应用的类型来看，工具类和游戏类应用占比最高。这提醒平台：即使是看似"安全"的应用类型，也要严格遵守平台规则。

不同地区的数据合规处理挑战

根据联合国贸易发展组织（UNCTAD）统计 ，全球 77% 的国家（共 194 个国家）完成了数据数据安全和隐私立法或者已经提出法律草案。这对出海企业在合规方面提出了更高的要求和更大的挑战， 数据处理合规已成为游戏出海运营的"前置条件" 。

欧盟 GDPR：*2018 年 5 月生效，全球覆盖最广、最严格的数据隐私条例，不合规将面临"营业额 4% 或 2000 万欧元"的高额罚款。*

美国 CCPA（加州消费者隐私法案）：*聚焦加州用户数据权益，明确消费者对个人数据的查询、删除、opt-out（拒绝数据出售）权利。*

新加坡 PDPA（个人数据保护法）：*要求企业收集、使用个人数据前需获得用户同意，且需建立数据保护机制。*

中东（迪拜）《DIFC 数据保护法》：*针对迪拜国际金融中心内的企业，对个人数据的存储、传输、处理提出严格合规要求。*

应用内生态安全：黑产冲击侵蚀运营根基

游戏内生态的稳定直接关联玩家体验、付费意愿与商业收益，而黑产通过技术手段实施的批量薅资源、外挂作弊、广告导流等行为，正从多个维度侵蚀应用生态，给企业带来直接的经济损失与品牌损害。

虽本白皮书主题定位在内容合规与风控，但从企业的角度，黑产对账号的攻击风险同样会造成不必要的业务损失，游戏行业就是泛娱乐出海中最典型的黑产攻击对象，所以本章节会将风控范围从内容扩展到账号，皆在给出海的游戏企业展示更全面的风险类型，做更充分的风控准备。

黑产批量薅取游戏资源

黑产对游戏生态的冲击首先体现在对资源体系的破坏，批量薅取游戏资源是最常见且危害深远的行为。黑产借助群控设备、接码平台等技术工具批量注册虚假账号，这些账号在领取新人专属奖励后，会被包装成"初始号"低价倒卖；更有甚者通过模拟器搭配自动化脚本，快速刷取游戏内的抽卡、任务等资源，形成"自抽号"后流入市场。

业务影响

降低正常玩家付费意愿：黑产低价倒卖资源，导致玩家"首充意愿"下降，破坏付费生态。

造成资源浪费与运营成本上升

大量低质黑产账号拉低游戏活跃指标，增加沉默玩家促活成本。

损害品牌信誉

正常玩家无法公平获取奖励，易对游戏品牌产生负面认知。

云手机批量获取游戏资源

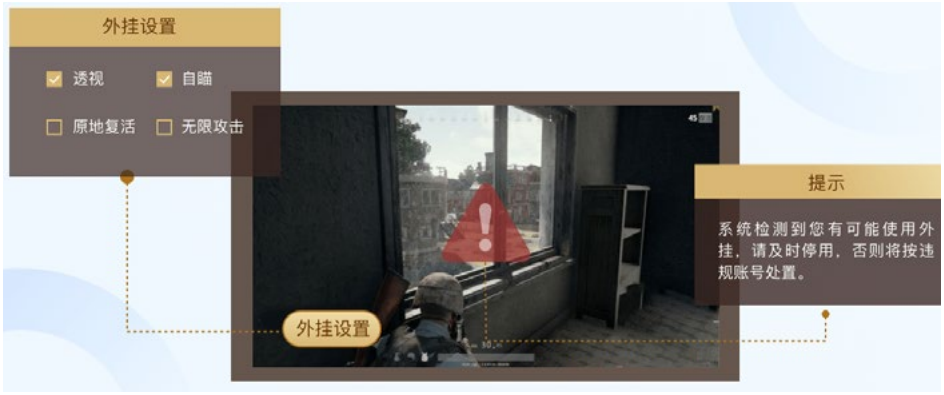
外挂工具破坏竞技公平

外挂作弊则从公平性层面摧毁游戏生态，尤其对竞技类游戏影响致命。随着技术的迭代，外挂已从早期的单一脚本作弊升级为"模拟器 + 外挂 +AI 自动化"的复合作弊模式，内存修改外挂可直接篡改游戏数据，图像识别类 AI 工具能模拟真实玩家操作轨迹实现自动瞄准、自动完成任务，这种高伪装度的作弊行为让传统检测手段难以识别。对游戏业务的影响主要体现在以下三点：

竞技环境恶化：*PVP 类游戏中外挂用户破坏公平性，导致正常玩家流失。*

经济系统失衡:*外挂自动化获取大量游戏资源，打破内循环经济平衡，影响玩家体验。*

付费转化率下降：*玩家可通过外挂获取付费内容，直接拉低整体付费意愿。*



游戏外挂示意

广告导流与竞品挖人：高价值用户流失

广告导流与竞品挖人则直接瞄准高价值用户，造成用户与收益的双重流失。

广告导流

游戏内充斥色情、博彩、竞品游戏的广告信息，诱导玩家添加第三方社群或跳转外部平台，高价值用户被精准引流，造成厂商用户流失与收益受损。

竞品挖人

竞品通过黑产工作室，以"注册送礼""高价收号"等诱人条件吸引用户，尤其针对付费能力强的高价值用户，直接冲击游戏营收根基

应对策略与能力要求

面对 AI 时代下游戏行业融合了内容、账号、业务生态的全链路风险，出海企业必须摒弃"点状防御"的旧思路，转而构建一个覆盖**用户全生命周期的"智能护航体系"**。该体系必须具备从被动拦截转向主动治理的能力，其核心能力要求可归纳为三大垂直支柱和一个全球化基座。

垂直能力支柱：全场景风险的精细化治理

基于游戏行业的风险场景（如聊天社区、注册登录、营销活动、公屏私信等），一个成熟的防御体系必须具备三大核心能力矩阵：

内容基础安全

这是应对"聊天 / 资料 / 社区"内容风险的基石。在 AIGC 的冲击下，内容审核不再是简单的关键词过滤，而必须进化为多模态的深度语义理解。

能力要求：

平台必须具备智能文本识别（理解隐晦辱骂、政治黑话、色情暗示）、智能图片识别（识别 AI 生成的违规图像、二维码变体）、智能音频识别（处理语音聊天中的违规内容）、智能视频识别（分析 UGC 视频）及智能网页识别（检测私信中的钓鱼链接）等全模态能力。

内容生态维护

内容层面的治理绝不是单纯的对风险内容的识别与拦截，内容标签对于音视频的智能推荐、分类和管理等及精细化运营，生态维护提供关键依据。

能力要求：

须具备对音视频内容进行标签化分类与分级的能力，对音频进行音色、性别、年龄、语种、场景音、伪造人声，等灵活化的识别能力；对视觉画面中出现的人脸、人体、画面主题、场所、logo、物品等内容及画质进行识别与标签标记。标签标记越准确，越精细，对智能运营与用户体验更友好。

账户安全

游戏生态的健康不仅依赖于内容合规，更深层次地取决于对"人"（即账号）的治理。AI 黑产的'智能化"和"仿真人"进化，使得传统基于规则的账号风控变得极易被绕过。因此，一套现代化的风控体系必须具备从源头到全链路的智能对抗能力。

能力要求：

源头注册防护（遏制黑产账号）： 风险治理必须前置在用户生命周期的最开端。平台必须具备强大的智能注册保护能力，在用户注册时，即时通过设备指纹技术精准识别模拟器、云手机、群控设备等"黑产设备"。结合 IP 信誉库、行为序列分析等 AI 能力，精准识别机器批量注册行为，在源头遏制"黑产账号"和"风险团伙"的涌入。

业务与营销风控（打击黑产工作室）： 针对"打金工作室"、"自动化脚本"、"羊毛党"等利用业务逻辑漏洞（如新手奖励、营销活动）进行获利的行为，风控体系必须具备业务层面的全链路分析能力。这要求系统能通过 AI 驱动的行为序列分析，实时监控资源获取、交易、消耗、登录模式等异常数据流，主动发现、预测并处置破坏游戏经济平衡的群体性欺诈行为。

高对抗安全防护（反外挂 / 作弊）： 针对"外挂"、"盗号"等高技术对抗风险，防御必须下沉。风控体系应具备端云一体的安全保护能力，通过端侧 SDK 实时监测内存修改、注入、变速、恶意调试等作弊行为，结合云端策略引擎，实现对盗号团伙和高阶外挂的精准打击，维护游戏的核心公平性。

社区生态治理（反引流 / 诈骗）： 针对"公屏 / 私信"中的"平台引流"、"色情引流"、"赌博引流"和"诈骗"等风险，风控系统必须将账号安全与内容安全能力打通。通过 AI 模型精准识别导流意图和变体广告，并对发布这些信息的黑产账号进行快速、自动化的处置（如禁言、封禁），实现对社区生态的闭环治理。

全球化基座：多语言支持与数据合规

上述三大能力支柱，必须建立在一个稳固的全球化基座之上，才能真正服务于出海业务。

多语言深度审核能力

出海业务的本质是本地化运营，内容风控的首要挑战便是语言。

覆盖广度： 风控体系必须支持全球主要语种，如中文、日语、韩语、印尼语、泰语、越南语、菲律宾语、马来语等东亚 / 东南亚语言，以及英语、西班牙语、葡萄牙语、俄语、法语、德语、意大利语等欧美主流语言，至少覆盖 10-15 个以上的主要国家语言。

理解深度： 更重要的是，系统不能依赖'先翻译再审核'的低效模式。它必须具备对原始文本（小语种）的直接语义理解和意图识别能力，能够精准捕捉本地化的俚语、黑话和文化禁忌。一个成熟的 AI 审核流程应当是：接收原始文本 -> 智能识别语种 ->（若为小语种）优先调用本地化分类模型，或辅以翻译引擎 -> 输出精准的本地化监管标签。

跨地区数据安全与合规保护

数据安全是出海合规的"生命线"。所有风控行为都必须在各国数据主权的框架下进行。

数据场景梳理： 必须具备清晰梳理海量业务中数据利用场景的能力，在了解数据保护现状的前提下，安全地进行数据脱敏和合规化处理。

风险等级划分： 出海企业在处理跨境数据时，必须进行严格的合法性、正当性和必要性评估。风控解决方案必须具备合规咨询与风险划分能力，能根据不同国家（如欧盟 GDPR、美国 CCPA、巴西 LGPD）的监管强度和法律法规，提供定制化的数据治理建议。

全球化服务部署： 为满足各国（尤其是欧盟、俄罗斯、中东等）日益严格的'数据本地化'要求，风控服务商必须具备全球化的服务集群部署能力。在北美（如弗吉尼亚）、欧洲（如法兰克福）、东南亚（如新加坡）、中东、南美等核心区域设立本地化数据中心，是保障数据不出境、满足内容数据检测需求的行业标准，也是为全球玩家提供毫秒级风控响应的必要前提。

标杆案例一

某末日策略类游戏

客户介绍

这是一款策略向对抗类游戏，围绕特殊生存场景展开。玩家要在对抗中推进探索，依靠策略规划达成核心生存目标，该游戏表现突出，全球范围内的营收规模已接近十亿量级。

在泛娱乐出海的版图中，以"末日生存"为主题的 SLG（策略游戏）是一个极其吸金且竞争白热化的赛道。这类游戏（以某款近十亿量级营收的头部产品为代表）的成功，依赖于几个核心且差异化的特点：

1. **全球同服（Global Server）：** 其核心魅力在于构建了一个"微缩地球村"。来自美国、韩国、中东、欧洲、东南亚的玩家汇聚在同一个世界地图上，进行实时的资源掠夺和领土战争。

2. **高强度联盟社交（High-Stakes Alliances）：** 生存不是个人战，而是联盟战。玩家必须结成跨国、跨语言的庞大联盟，进行 24/7 的外交、谈判、背叛和协同作战。

3. **高 LTV（高生命周期价值）：** SLG 玩家的忠诚度和付费意愿极高，核心付费用户是所有游戏类型中价值最高的群体之一。

这种"全球同服、高社交、高价值"的独特模式，使其不仅是一个游戏，更是一个高强度的"全球社交实验场"和高价值的"数字经济体"。这也使其风控挑战变得尤为严峻和复杂。

内容风控挑战：世界频道的"外交风暴"与"文化地雷"

游戏的世界频道（World Chat）和联盟频道（Group Chat）并非普通闲聊，而是承载着联盟

外交、宣战、策略部署的核心功能。比如，当一个土耳其联盟的领袖需要和一个韩国联盟谈判时，任何一个词汇的误解都可能导致一场灾难性的服务器战争，造成高价值玩家的集体流失。

1. **多语言的“火药桶”：** 平台必须实时处理全球数十种语言（包括阿拉伯语、德语、法语、俄语、泰语等）的审核。更致命的是，风险点往往不在于标准脏话，而在于本地化的政治隐喻、宗教禁忌和种族俚语。

2. **不堪重负的人力：** 面对 7x24 小时涌入的全球信息流，企业的人工审核部门（即使是多语言团队）也迅速被淹没，高压之下，错漏频发。

账号风控挑战：生态系统的“寄生军团”与“经济侵蚀”

极高的玩家价值，使游戏成为 AI 黑产（Bot-driven threat actors）眼中的“金矿”。它们的核心目的不是破坏，而是“寄生”和“获利”。

3. **AI 黑产的“仿真人”账号：** 黑产利用自动化脚本和模拟器，24 小时批量注册账号、刷取游戏资源（打金工作室），再通过世界频道和私信，以低价向真实玩家兜售资源或进行诈骗引流。

4. **生态的“劣币驱逐良币”：** 这些“广告导流”信息严重污染了核心的社交场景（外交、战术频道），极大破坏了付费玩家的游戏体验和资产安全。黑产的“寄生”行为，正从内部侵蚀着这个十亿量级营收的数字经济体。

合规风控方案与战略成效：用“智能”重塑“秩序”

面对这场“外交”与“经济”的双线战争，让该企业意识到，传统的“人海战术”审核和“关键词”过滤已彻底失效。它们选择了一套“以智御智”的全链路智能风控解决方案。

1）重构多语言内容审核

应对策略

平台不再依赖"先翻译再审核"的低效模式，而是部署了覆盖全球主流语种的 AI 内容审核引擎。该引擎不仅能识别多语言的显性违规，更能精准理解隐藏在俚语、谐音、变体词背后的真实意图（如广告、辱骂、诈骗）。

价值体现：

AI 承担了全球聊天场景中 99% 以上的初审工作。人工审核成本骤降 75%，同时，由于AI对高风险政治、宗教内容的精准拦截，确保了平台在全球各区域（特别是中东、欧洲）的线上内容零监管事故。更重要的是，人类审核员被解放出来，从疲于奔命的灭火队"转变为处理复杂文化冲突的"高级专家"。

2）从源头净化黑产

应对策略：

针对"寄生军团"，平台没有在聊天频道"打地鼠"，而是采用了源头治理策略。在用户注册环节即引入了 AI 账号风险识别工具。

价值体现：

该工具通过分析设备指纹、IP信誉和行为序列，能在用户注册的瞬间，就精准识别出自动化脚本"和"模拟器"等黑产设备，并进行前置风险打标和自动化处置。

效果：

这场"源头打击"成效显著。游戏内的广告信息浓度下降了 83%，玩家相关的客诉量随之下降了 22%。这不仅大幅优化了用户体验，更重要的是，它切断了黑产的"寄生"链路，稳固了游戏内的经济平衡，有效保护了高价值玩家的利益，让游戏的核心生态得以重回健康和繁荣。

标杆案例二

某中世纪策略游戏

这是一款以中世纪奇幻世界为背景的策略 SLG 手游，玩家通过建设城堡、训练军队、争夺圣剑展开冒险。其在除中国外的市场收入已达数十亿元。

相比现代战争题材，该游戏资源获取和经济内循环机制更为复杂，强调 PVE（玩家对环境）与 PVP（玩家对玩家）的平衡。高价值付费用户通常是资源积累型玩家。在此特点下，黑产工作室可利用外挂脚本和模拟器，24 小时自动"打金"（刷资源），并以远低于官方的价格倾销，严重破坏了游戏经济平衡。同时，利用内存修改外挂在 PVP 中获取不公平优势。具体风控挑战表现为：

1. **资源号泛滥 ：** 游戏工作室通过外挂脚本挂机等非常规手段获取大量资源并低价售卖，导致玩家放弃官方渠道购买资源，平台收益受损。
2. **设备识别缺失 ：** 缺乏统一设备标识，无法跨平台关联多款游戏，难以高效治理多端作弊行为。

应对策略与成效：引入智能设备指纹技术，构建"黑产族谱"

应对策略：

该企业采用了一套企业级的智能设备指纹解决方案。该方案通过分析数百个设备参数，为每一台登录游戏的设备（无论是真机还是模拟器）生成一个全球唯一的、不可篡改的设备 ID。

价值：

风控系统能精准识别"一台设备登录多个账号"(群控) 或"一个账号在多个异常设备登录"（盗号）等行为，并自动将这些异常设备和账号关联，构建出庞大的黑产团伙画像。

结果：

系统日均自动封禁的黑产账号超过 15 万个，其中精准封禁的"打金工作室"账号超过 75%。游戏资源价格回归正常，付费玩家的公平性得到保障，游戏的生命周期得以极大延长。

音视频 / 社交行业：强内容属性下的监管与伦理挑战

音视频和社交行业作为用户生成内容（UGC）和实时互动（Live-Streaming）的核心载体，是泛娱乐出海领域中面临全球监管压力最严苛、风险对抗环境最复杂的赛道。其内容风险的性质已从传统的单一违规，升级为"多模态"融合、AIGC 驱动的系统性挑战，对平台的合规能力提出了前所未有的要求。

强监管与 AIGC 对抗：风控新挑战

音视频 / 社交平台的风险挑战可概括为**三大维度**

全球强监管环境下的平台责任升级与高昂代价

随着欧盟《数字服务法案》（DSA）、美国《儿童在线安全法案》（KOSA）等法规的收紧，平台不再是内容的"避风港"，而是被要求承担积极的审核和治理责任，将合规能力直接与企业的市场准入和生存权挂钩。

未成年人保护

不可触碰的绝对红线 平台需严格遵守 COPPA、GDPR 等未成年人保护法规。风险行为包括：诱导发送未成年人隐私信息（如通过私聊索要裸照、家庭住址），或利用未成年人账号牟利（如兜售未成年人视频、未成年人违规直播）。

> ***处罚示例：*** *触犯相关未成年人保护法，可能面临高额罚款，例如美国 FTC 对未成年人隐私违规的最高罚款可达 4.3 万美元 / 例。*

宗教与价值观合规

高敏感区域的刑事风险 尤其在东南亚和中东（MENA）市场，宗教和道德禁忌是绝对红线。风险集中于仇恨言论与诽谤（攻击特定信仰、神职人员或信徒），以及不当亵渎行为（对圣物、经典如《古兰经》或神圣场所进行恶意诋毁）。

> ***处罚示例：*** *在沙特阿拉伯，发布或传播煽动宗教仇恨的内容，可面临最高 5 年监禁和 300 万罚款；在印度尼西亚，亵渎宗教可面临最高 5 年监禁，且平台需对内容承担连带责任。*

未成年

场景特征：

- **诱导发送未成年人隐私信息：** 通过私聊向 18 岁以下用户索要裸照、家庭住址，或制作"未成年人付费"服务。
- **利用未成年人账号牟利：** 兜售未成年人视频、未成年人直播。

处罚示例：

触犯 COPPA、GDPR 等未成年人保护法，可能面临高额罚款（如美国 FTC 对未成年人隐私违规最高罚款 4.3 万美元 / 例）。

宗教

场景特征：

- **仇恨言论与诽谤：** 攻击伊斯兰教信仰、神职人员或信徒，发布带有歧视性、侮辱性或煽动暴力的内容
- **不当亵渎行为：** 对伊斯兰教的圣物、经典（如《古兰经》）或神圣场所进行恶意诋毁、涂改或侮辱

处罚示例：

沙特阿拉伯发布或传播煽动宗教仇恨的内容，可面临最高 5 年监禁和 300 万；印尼亵渎宗教可面临最高 5 年监禁，平台需对内容负责

色情

场景特征：

- **传播淫秽物品：** 发布、制作、复制或贩卖淫秽视频、图片等，或在评论区、私信中传播色情链接或信息
- **性暗示与挑逗：** 账户头像、昵称或简介含有露骨的性暗示，发布擦边球内容，或在直播中进行低俗挑逗，游走在色情边缘

处罚示例：

阿联酋制作、传播或储存色情内容最多可面临最高 100 万迪拉姆的罚款，平台需对内容负责

内容风险的多模态融合与隐蔽性进化

内容的风险不再是孤立的文本或图片，而是通过多维度的信息流进行隐蔽传播。

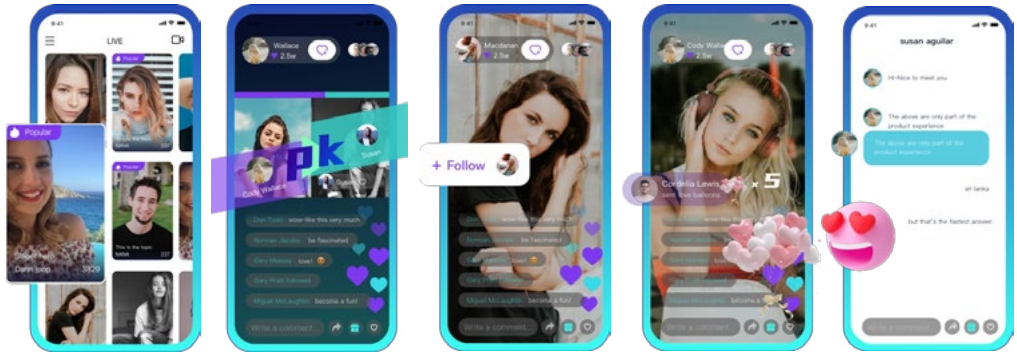
色情与软色情的隐蔽性：

除了显性的传播淫秽物品（贩卖视频、图片或传播色情链接）外，高对抗性的黑产账号利用性暗示与挑逗（如擦边球图片、露骨的昵称 / 简介）绕过关键词过滤。

直播 /UGC 场景的复杂性：

风险被嵌入到画面、音频、文本等多种模态中，使得传统基于单一维度（如只看画面）的审核极易失效。例如：

> ***内容质量风险：*** *衍生出大量空播、视频匹配无人脸、直播背景噪音等违规场景，利用低质量内容或违规场景本身进行试探。*
>
> ***作弊引流风险：*** *出现封面图质量问题（如画面不清晰、人脸被遮挡、涉嫌软广）或封面图差异问题（封面图与实际真人差异较大，或用户真人与相册内容差异较大），利用虚假形象进行引流。*



社交娱乐平台场景丰富

业务与流量风险的高频对抗：广告导流与生态寄生

社交和直播的高并发、强时效性特征，使其成为黑产进行金融诈骗、流量作弊和业务侵蚀的重灾区。

金融诈骗、广告导流与钓鱼欺诈：

黑产团伙利用直播间、私信、评论区等场景，以高频次、强实时性的方式进行"杀猪盘"、裸聊诱导、以及赌博、钓鱼等引流活动。其成功率高、溯源难度大，严重侵害了用户财产安全。

流量作弊、广告导流与业务侵蚀：

黑产利用批量注册账号、养号、群控等手段进行大规模的刷量、刷粉，并通过广告导流的方式引流到站外（如微信、Telegram 等），同时，他们会针对平台的营销活动权益进行黑产薅取和倒卖获利。

> ***风险案例：*** *黑产利用脚本批量注册账号后，薅取平台提供的"新客权益"（譬如钻石等虚拟货币），并低价倾销，不仅稀释了平台的正常流量，更以业务寄生的方式污染了社区环境，形成难以根除的对抗。*

杀猪盘诈骗示例

应对策略与能力要求：

面对强监管和高对抗环境，音视频 / 社交行业必须从传统的"人机接力"模式，向 AI 驱动的"智能体（Agent）决策"模式转变，构建一个覆盖内容、账号、业务、合规的全链路治理体系。

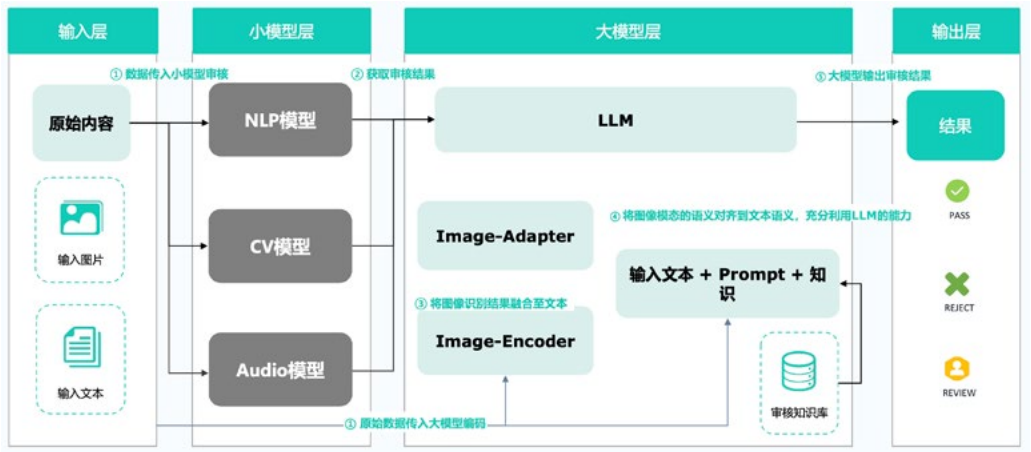
多模态融合治理与 AI 攻防：

平台需具备强大的多模态融合理解能力，能够将文本、音频、视频、图片等信息整合到一个统一的语义空间进行识别，以应对直播和 UGC 内容的复杂隐蔽性。

针对 AIGC 原生风险，必须建立 AI 攻防体系，AIGC 应用在上线前、上线中、上线后持续运营的全生命周期中，风险随业务推进呈现不同特征，既需满足政策合规硬性要求，也需应对用户交互、平台运营及舆情管理中的潜在隐患。AIGC 应用要想平稳、合规、健康的上线运营并行稳致远，就需要构建**「上线前安全评估 - 上线后风险防控 - 长期运营保障」**的全生命周期业务风控体系，大模型内生安全、账号风险识别、内容风险识别、舆情布控响应等多个风控要点为 AIGC 应用构筑全生命周期的安全保障。

智能体（Agent）决策系统：

取代传统的"规则 + 人力"审核流，引入具备复杂判断和自主学习能力的智能体决策系统。该系统能够实时整合来自内容、账号、地理位置、历史行为等多个维度的风险信号，基于各国和地区的动态合规策略进行全局风险判断，实现毫秒级的自动化处置，大幅提升风险处置的准确性和效率。



账号与业务风险的体系化防御：

针对黑产团伙，风控体系应实现从**渠道流量 - 注册登录 - 业务行为**的全面覆盖，精准识别渠道假量、虚假注册登录、营销欺诈、广告诈骗导流等风险问题。通过构建**账号画像、设备标签、行为标签**，利用**时域关联网络、频度策略，黑产聚集特征**等手段，有效识别黑产账号、虚拟手机号、虚假 IP、虚假设备，从源头保障平台的商业生态。

全球化合规集群与信创支持：

为满足 GDPR、LGPD 等数据不出境的合规要求和全球用户体验，必须构建全球分布式合规集群（如在欧洲、北美、中东等地），确保数据安全隔离和风控服务的低延迟。同时，平台需具备在特定市场（如中国）支持信创等国产化基础设施的能力，保障业务的自主可控和持续运营。

案例一

HOLLA Group

HOLLA Group 成立于 2016 年，是一家以 “打破物理屏障，加速全球线上实时沟通普及” 为使命的全球社交娱乐公司，核心业务聚焦海外市场，已构建起覆盖多元社交场景的产品矩阵 —— 旗下 HOLLA、Monkey、Hay 等多款产品，累计注册用户超 1 亿，每日完成数千万次随机配对聊天，成为连接全球用户的重要社交桥梁

在竞争激烈的海外社交赛道中，HOLLA Group 展现出强劲的市场爆发力：核心产品 Monkey 上线仅数月便跻身美国 Social Top 30 榜单，同时在全球超 67 个国家和地区的 Social 类应用中进入 Top 50；凭借产品创新力与用户价值，公司还先后入选 Google Sand Hill、Google Rising Star、Google Advantage 等谷歌精英 Program，成为海外实时社交领域公认的 “成长型黑马”，其业务规模与全球影响力正持续快速扩张。

海外社交合规：未成年人、文化差异、用户体验

随着 HOLLA Group 产品在全球市场的渗透，其实时社交模式下的内容安全与合规压力也同步凸显，具体面临三大核心挑战：

海外未成年管控 “零容忍”，合规红线高

欧美等核心目标市场对社交平台的未成年内容管控实施严格的“零容忍” 政策 — 无论是用户注册环节的未成年人身份识别，还是实时聊天中可能出现的诱导未成年人、向未成年人传播不良信息等行为，一旦违规将面临高额罚款、应用下架甚至法律追责，而 HOLLA 的随机配对模式进一步增加了未成年相关风险的防控难度，亟需精准的风险识别能力。

跨文化、跨区域审核标准差异大，技术积累要求高

HOLLA 用户覆盖全球超 100 个国家，不同种族、文化背景及地区的法律法规对内容审核的要求存在显著差异：例如部分地区对宗教相关内容敏感，部分国家对政治话题管控严格，还有的区域对“调侃式表达”的界定与主流市场不同。传统基于关键词的审核方式无法适配多元文化场景，需具备结合语义理解、文化语境的深度审核能力，这对内容审核的跨区域经验与技术沉淀提出极高要求。

数据合规与用户体验的平衡难题

海外市场普遍要求社交平台遵循“用户数据最少收集原则”，同时需符合 GDPR（欧盟通用数据保护条例）、CCPA（加州消费者隐私法案）等不同地区的数据传输与存储法规。而 HOLLA 的实时聊天业务对数据处理的“低延迟”有强需求 —— 若为满足合规过度限制数据流转，会影响配对聊天的实时性；若忽视合规风险，则可能触发隐私监管处罚，如何在“数据合规”与“用户体验”间找到平衡点，成为业务持续运营的关键。

针对 HOLLA 的海外业务场景与核心痛点，数美科技以“合规为基、体验为要”，提供覆盖“风险识别 - 策略适配 - 数据保障”的全链路内容风控方案，打造贴合实时社交场景的安全底座：

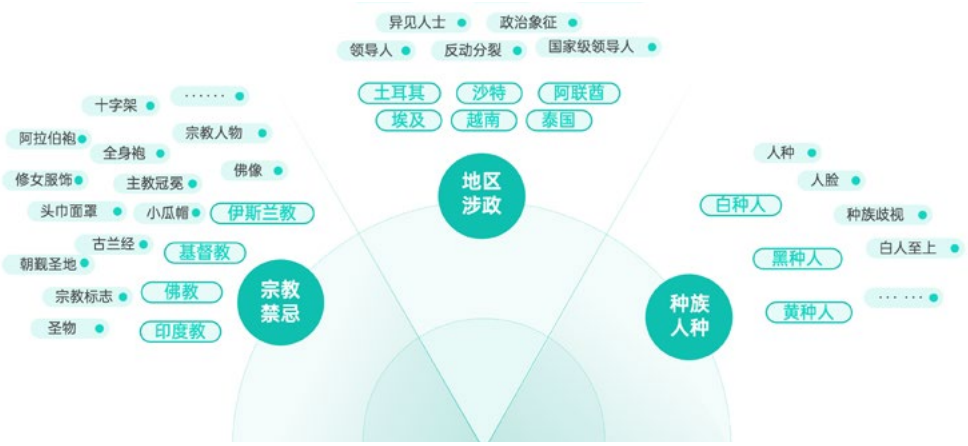
专项未成年保护体系，筑牢合规防线

数美构建了针对社交场景的未成年内容专项检测模型：通过多维度识别（如用户头像 / 昵称的未成年人特征、聊天文本中的年龄暗示、语音内容的稚嫩声线判断），精准定位未成年人用户及潜在的未成年人相关风险行为；同时，针对 “随机配对” 场景，设置 “未成年人保护触发机制”——若检测到一方可能为未成年人，将会限制高风险聊天内容（如暴力、色情相关表述）的传输，并向平台推送风险预警，确保符合海外市场对未成年保护的严苛要求。

跨区域多模态审核能力，适配文化差异

为解决跨文化审核难题，数美打造了"多语种语义理解 + 区域化策略库" 双引擎：

> ***语言层面：****支持 40+ 海外语种，不仅能识别敏感词，还能捕捉隐喻、谐音等隐晦风险表述，进行深度语义分析（如不同地区对“不当内容” 的隐晦说法）；*
>
> ***文化层面：****助力 HOLLA 建立覆盖 67 个核心目标国家的 “区域化审核策略库”，整合当地法律法规、文化禁忌（如宗教敏感内容、区域政治红线），实现 “一地一策” 的精准适配；*
>
> ***模态层面：****针对 HOLLA 的 “文本 + 语音 + 图片” 多模态内容，采用 “跨模态协同审核” 技术 —— 例如结合图片中的场景特征与文本聊天语境，判断是否存在跨模态组合的风险（如看似正常的图片搭配隐晦的不良文本），避免单一模态审核的遗漏。*



标签示例

同时，为呈现不同市场的文化差异及其潜在风险，数美基于全球政策与文化差异，构建包含宗教符号（十字架、佛像）、种族特征（白种人、黑种人）、地域标识（阿拉伯袍、土耳其国旗）等海外特色标签体系，支持企业按地区法规自定义审核策略。通过标签细化训练模型，可将人机协同效率提升 50% 以上，有效应对出海内容合规挑战。

合规优先的数据处理方案，平衡体验与安全

数美严格遵循海外数据监管要求，提供合规化数据处理架构：

> ***数据收集：****采用"最小必要" 原则，仅采集审核所需的核心数据，如聊天文本片段、风险内容截图，避免冗余数据存储；*
>
> ***低延迟优化：****针对实时聊天场景，优化风控模型的计算效率，将单条内容的审核响应时间控制在毫秒级，确保风险检测不影响配对聊天的实时性，兼顾"合规"与"体验"。*

动态策略迭代机制，适配业务增长

数美以风控伙伴角色深度嵌入 HOLLA 的业务全周期，针对不同阶段需求提供适配性风控支持：

> ***业务增长扩张期：****在 HOLLA 核心欧美市场进入平稳增长阶段时，数美为其配置基础审核策略与未成年专项防控策略，快速满足当地"零容忍"合规要求；当 HOLLA 向东南亚、拉美等新兴市场扩张时，数美同步更新区域化审核策略库，确保风控能力与市场拓展节奏精准匹配，助力业务顺利落地新区域。*
>
> ***日常运营舆情应对期：****依托境内外全域监控渠道，数美可捕捉全球社交平台的热点风险（如突发不良话题传播），并在 24 小时内完成审核策略迭代，帮助 HOLLA 快速响应潜在风险，保障*日常运营稳定。

方案效果与价值：安全、效率、增长三重赋能

定制化风控方案落地后，HOLLA Group 的海外业务带来 "合规安全、效率提升、生态优化" 的多重价值，实现 "安全与增长" 的双向共赢：

1 合规风险显著降低，保障全球平稳运营

通过使用视觉模型和年龄划分标签，将人工年龄标注的成本降低了约 30%，降低了未成年人出现的风险，实现了专项未成年保护，让 HOLLA Group 在欧美等核心市场实现业务安全运营。同时不断迭代更新的跨区域合规审核方案，帮助 HOLLA 顺利进入多个新兴海外市场，业务扩张的 "合规门槛" 显著降低。

2 审核效率与准确度双升，构建健康的社交生态

方案落地后，HOLLA Group 的内容审核关键指标实现跨越式提升：

审核准确度：多模态协同审核与跨区域策略库的应用，使风险内容的误判率大幅降低，避免因文化差异导致的"误拦截"；

审核效率：毫秒级的实时审核响应，满足了实时聊天的低延迟需求，审核更高效精准

违规控制：平台整体违规内容量级降至合作前的 20%，其中未成年相关风险、跨文化敏感内容的违规量大幅下降，构建了更健康的社交生态。

3 用户体验与信任度提升，赋能业务增长

减少误拦截：跨文化审核的精准适配，避免了因文化差异导致的正常聊天内容被误拦截，用户投诉率下降，配对聊天的流畅性显著提升；

增强用户信任: 数美的合规化方案与未成年保护措施，帮助 HOLLA 在海外用户中建立"安全、可靠" 的品牌形象，核心产品 Monkey 的用户留存率提升，助力其在多个国家的 Social 榜单排名稳步上升。

HOLLA Group

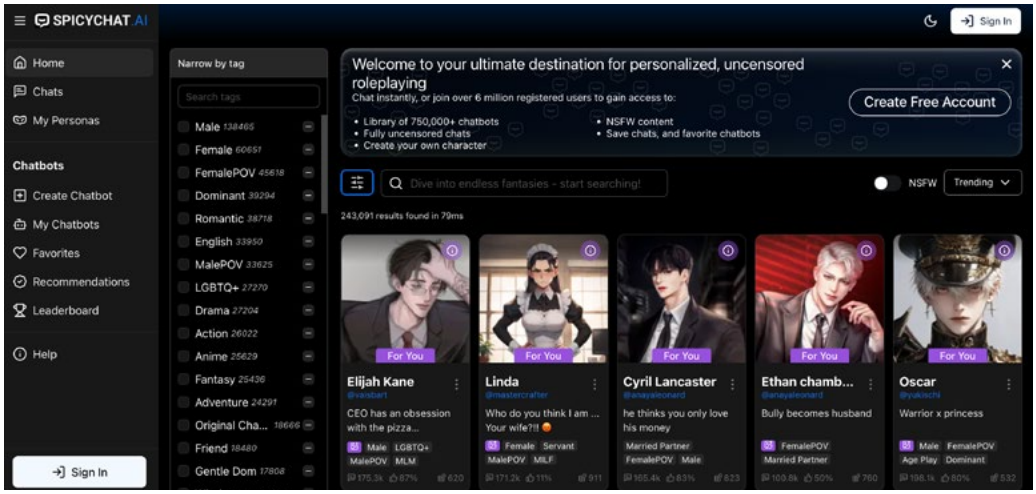
"数美为 HOLLA Group 产品能在海外平稳运营保驾护航，提供了必不可少的内容安全底座与合规保障 —— 其定制化的风控方案不仅帮我们解决了海外市场的合规难题，更让我们在'安全'与'体验'间找到了最佳平衡，为业务的全球增长注入了关键信心。"

— HOLLA Group CTO 张玉智

案例二

全球 Top10 AI 社交工具 SpicyChat.Ai

SpicyChat.AI 是一款以 AI 角色扮演聊天为核心的全球化 AI 社交平台，曾是登顶全球 Top10 的 AI 社交工具，以“Free Uncensored Roleplay Chat”（自由的、不设限的角色扮演聊天）为核心理念，鼓励用户自由创作多样化的内容，同时用户可以在这个平台上创建和定制自己的 AI 角色，进行沉浸式的角色扮演聊天体验。



SpicyChat.Ai 界面

具体来看，SpicyChat.AI 的“不设限”，体现在：

内容创作不设限：

用户可以自由创作多元化的内容，无论是浪漫故事、冒险经历还是其他创意想法，都可以在这个平台上得到展现。

角色扮演不设限：

用户可以创建自己想要的 AI 角色，这些角色可以是任何想象得到的形象，从历史人物到虚构的奇幻生物，甚至是用户自己设计的独特角色。

多语言创作不设限：

SpicyChat.AI 支持多种语言，使得全球不同地区的用户都能够在这个平台上进行创作和交流。

互动体验形式不设限：

用户可以与自己创建的 AI 角色进行互动，这种互动可以是对话、游戏或者其他形式的交流，为用户提供更丰富的体验。

现象级平台的平衡难点：自由与合规的角力

SpicyChat.AI 的成功源于平台的开放和创新，然而当自由创作的不设限遇到内容合规的红线，这场超过 750000 个虚拟角色的狂欢背后，隐形的内容安全挑战开始浮出水面。作为一家以"不设限"为核心理念的 AI 平台，SpicyChat.AI 的开放生态也面临着棘手的内容合规挑战：

1. 欧美市场的合规高压线：

未成年（underage）、性暴力描述（rape）、兽行（bestiality）和乱伦（incest）等敏感内容，不仅违反《儿童在线保护法案》（COPPA）等法规，更可能引发平台被下架的危机。

2. 用户体验与安全的博弈：

关键词屏蔽已无法应对复杂场景，例如，"我 17 岁"可能是年龄陈述，也可能是未成年内容的暗示；"虚构小说中的强奸情节"需要与真实违规内容区分，依赖关键词拦截会影响用户的自由创作，同时内容审核带来的延迟响应会让平台用户活跃度下降。平台在拦截高敏感内容的同时，必须保障用户流畅的创作与互动体验。

3. 多语种审核的复杂性：

英语用户虽占主流，但东南亚、拉美用户快速增长带来语言混杂问题，内容审核难度加大。

构建全球化内容风控方案：守护平台安全边界

真正的自由，始于对风险的精准判断。作为面向全球的内容审核专家，数美科技为 SpicyChat.AI 定制了"双端协同"的 AI 内容风控方案，精准响应业务需求：

Web 端 SpicyChat.AI：高拦截率与体验保障的高效融合

SpicyChat.ai 建立 le "严格拦截 + 体验优化"双合规机制。一方面，对 4 类高敏感内容进行严格拦截，确保 85% 的拦截率，从源头阻断违规内容传播。另一方面，通过智能算法优化审核逻辑，避免过度拦截对正常用户创作的干扰。例如，在涉及"强奸"内容时，不仅识别关键词"rape"，更能通过 NLP 模型对上下文语义的深度解析，判断 " 小说中的虚构强奸情节 " 与 " 真实违规描述 "，既守住安全底线，又保留用户创作空间。

移动端 PixelChat：适配应用商店规则的引流保障

受制于 Apple Store 与 Google Store 的严格限制，PixelChat 面临更严苛的内容审核环境——采用"高标准策略阈值"，在避免漏杀的同时，确保应用符合商店规范。移动端被定位为 SpicyChat·AI 的引流入口，并通过轻量化审核机制，在保障内容安全的前提下，维持用户流畅的访问体验，实现安全审核与业务增长的协同推进。

https://play.google.com/intl/zh-CN/about/developer-content-policy/

https://developer.apple.com/cn/app-store/review/guidelines/#objectionable-content

这套双端方案，支撑了 Spicychat.ai 在全球化的核心业务的合规：无论是欧美市场对高敏感内容的零容忍，还是应用商店的特殊规则，都能通过这套风控方案实现精准适配，真正做到"一地一策，一客一策"。

数据背后的价值：安全与体验的双重跃升

在稳定运营的月度分析中，SpicyChat.AI 平台上用户生成的内容会经过 17.25 亿次文本过滤服，2.83 千万条风险内容的拦截，日均拦截量达 10 万条以上。这些数字的背后，是全球化合规方案的三大能力的诉求：

1. 精准拦截，守护合规底线

针对未成年、强奸等重点内容，依托 AI 算法与专业审核团队的协同作业，实现从关键词识别到语义理解的全维度审核。例如，在处理涉及未成年内容时，不仅拦截明确提及"underage"的文本，更对暗示未成年人参与的模糊描述进行深度挖掘，确保违规内容无所遁形。

2. 智能审核，提升用户体验

通过上下文模型处理用户反馈内容，有效提升审核准确性。当用户对某条内容提出反馈时，系统不再孤立分析被反馈内容，而是结合前后文语境、用户历史行为等多维度信息综合判断，减少误判率。同时，基于整体描述的全局审核，避免因单一关键词触发的过度拦截，让用户在安全环境中享受创作自由。

3. 双端协同，赋能业务增长

Web 端与移动端的差异化方案，既满足平台核心安全需求，又为业务发展提供支持。PixelChat 作为引流入口，通过数美科技优化后的审核机制，在合规前提下保持用户获取的流畅性，实现安全审核与用户增长的良性循环。

Spicychat.ai 的崛起，印证了生成式 AI 社交的无限可能。而全球化的 AI 浪潮裹挟着文化冲突、法规差异与技术挑战汹涌而来，企业需要的不仅是合规，而是能洞悉人性、理解文化、预判风险的 AI 智能合规方案，平衡用户体验与风控安全。这个再次证明：真正的全球化内容风控，不是用同一把尺子丈量世界，而是让每种表达都能在安全中找到自由的边界。

04

泛娱乐出海企业合规行动指南与战略建议

Compliance Action Guide and Strategic Recommendations for Entertainment Enterprises Going Global

在 AI 技术浪潮和全球监管趋严的双重背景下，泛娱乐出海企业的合规策略正经历一场深刻的范式革命。传统的"事后补救"模式已无法应对 AI 驱动的高隐蔽性、强对抗性风险。合规不再是法务部门的孤立工作，而是必须内嵌于产品设计、技术架构和公司战略的"CEO 工程"。

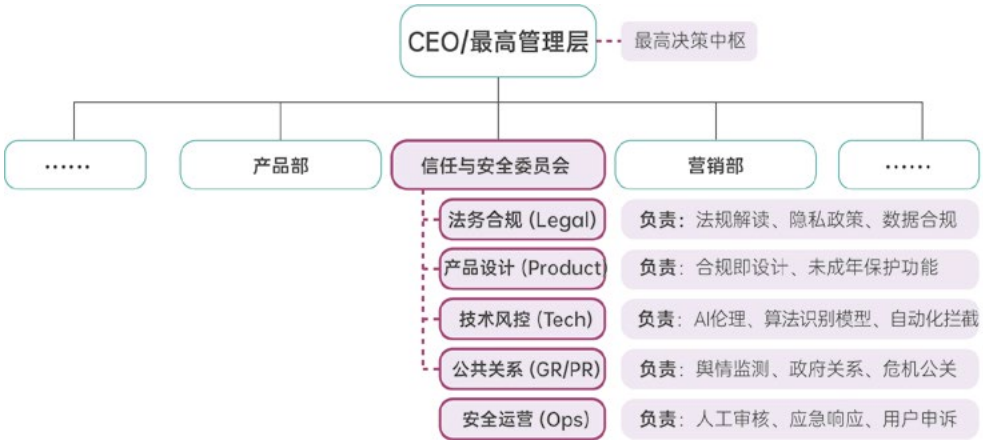
战略规划：从合规底线到信任资产

合规的最高境界不是规避罚款，而是将"安全可信"（Trust & Safety - T&S）塑造成品牌的核心竞争力。这要求企业在顶层设计上实现三个关键转变：

组织升级：建立跨职能的"信任与安全 (T&S)"中枢

传统分散在法务、运营、技术等部门的合规职能必须被整合，具体行动指南如下：

> ***设立 T&S 决策机构：*** *组建一个由法务、产品、技术（风控）、运营、公共关系（GR/PR）共同参与的跨职能"信任与安全委员会"，直接向 CEO 或最高管理层汇报。*
>
> ***职能定位：*** *该中枢不只是"灭火队"，更是"导航仪"。它必须在产品立项（T-0）阶段就介入，评估潜在的市场法规、数据隐私、AI 伦理和内容生态风险，确保产品"合规即设计"(Compliance by Design)。*



战略制定：量化“全球风险偏好”框架

CEO 必须回答一个核心问题：在不同市场，我们愿意承担多大风险？“一刀切”的策略是低效且危险的。所以建议的行动如下：

定义“零容忍区”： 明确全球一致的底线，如儿童安全（CSAM）、恐怖主义、极端暴力。这些领域必须投入最顶级的 AI 审核资源，实现“零漏放”。

划分“动态敏感区”： 针对不同市场的文化和政治差异，制定动态的风险偏好。

> ***举例：*** *在欧洲，对“仇恨言论”和“数据隐私 (GDPR)”采取“低风险偏好”；在中东，对“宗教亵渎”和“LGBTQ”内容采取“极低风险偏好”；在北美，对“软色情”的偏好度可能相对宽松，但对“枪支毒品”则需严格卡控。*

定义“业务风险区”： 针对“黑产薅羊毛”、“打金工作室”等业务风险，应根据 ROI（投入产出比）设定容忍度，目标是在不影响真实用户体验的前提下，将损失控制在可接受范围。

产品立项：嵌入 AIGC 时代的安全能力

在 AIGC 时代，合规前置化的内涵已发生改变。不止于增强安全意识和测试，更是合规融入产品设计环节的大势所趋：

强制执行 “双重影响评估”： 在任何新功能（尤其是 AIGC 功能）上线前，必须强制执行：

> ***数据隐私影响评估：*** *评估功能是否符合 GDPR、CCPA 的“数据最小化”和“用户知情权”原则。*
>
> ***内容安全风险评估：*** *评估 AIGC 功能是否可能被滥用于生成 Deepfake、仇恨言论、侵权内容或进行 Prompt 注入攻击（如第三章所述）。*

践行 “隐私即设计”： 从技术架构层面，默认开启数据加密、去标识化、匿名化等隐私保护技术，将用户数据泄露风险降至最低。

区域性内容治理最佳实践

战略的落地依赖于一个智能、敏捷、可扩展的治理执行体系。AI 时代，传统的人海战术已然失效，企业必须构建"AIOps（AI 驱动的智能运营）"能力。

分级治理策略：从手动配置到动态策略引擎

全球化运营需要一个能够实时执行"全球风险偏好框架"（4.1.2）的技术大脑，而非"一刀切"的审核规则：

启用"一国一策"的动态策略引擎： 风控后台必须支持基于用户地理位置（Region）、语言、年龄等标签，实时调用匹配的审核模型和策略。引入强大的、具备低代码 / 无代码配置能力的内容合规工具，是实现这种精细化运营的前提，它能让运营团队（而非技术团队）能够快速响应不同市场的法规变化。

构建全球分布式合规集群： 为满足各国（尤其是欧盟、俄罗斯、中东）日益严格的"数据本地化"要求和全球用户的低延迟体验，风控系统必须具备全球化的服务部署能力。在欧洲（如法兰克福）、北美、东南亚（如新加坡）等地部署本地化 AI 风控集群，是保障数据不出境、毫秒级响应的行业标准。

本地化运营：从人海战术到人机协同

AIGC 带来的内容井喷，彻底终结了"在海外建立大型人工审核中心"的传统模式，应更注重人机协同的全新模式：

应用"智能体"审核 99% 的内容： 行业最佳实践正在转向"智能体（Agent）"技术。Agent 不再是简单的'AI 初审"，而是能模拟人类专家，自主融合内容、账号、行为数据，并基于"动态策略引擎"做出决策的"数字审核员"，自动化处理海量、高并发的显性风险。

赋能 "本地专家"处理 1% 的疑难： 抛弃低效的"人工审核团队"，转而建立小而精的"本地文化专家组"（可自建或与本地合作伙伴共建）。他们不再负责海量审核，而是专注于处理 AI Agent 无法决断的、涉及本地俚语、亚文化、政治隐喻、宗教禁忌的 1% 疑难杂症。

构建 "人机回馈"的进化机制 ： 建立高效的回馈机制（Feedback Loop），本地专家的每一次判例，都将作为"高质量数据"反哺给 AI Agent，使其自主学习和进化，实现越用越准。这才是 AI 时代最高效、最具扩展性的本地化运营模式。

危机应对机制：从被动公关到主动透明

在欧盟 DSA 等法规下，"危机应对'已从被动的公关处理，升级为主动的、可审计的透明度管理:

1. **建立清晰的用户申诉渠道：** 这是 DSA 的强制要求，也是赢得用户信任的关键。平台必须提供清晰、易于访问的渠道，允许用户对内容删除、封禁等决策提出申诉。

2. **突发话题的实时感知与风控：**建立风险态势感知系统，能够实时捕捉热点事件、突发话题等潜在风险，并通过合规策略动态迭代快速应对，确保应急处置响应及时，快速应对日常及突发的合规风险。

3. **主动发布《透明度报告》：** 效仿头部平台，定期发布《透明度报告》，主动披露收到的政府 / 用户举报数量、内容处置类型、申诉处理情况、AI 审核准确率等，将"被动合规"转变为"主动透明"，以此构建强大的品牌护城河。

出海合规与本地化服务

数美科技：

提供 AI 风控技术与全球化风控服务能力

以全栈式风控与大模型能力，建立账号到内容的风险识别体系

助力企业满足全球不同区域的合规要求（如欧盟《数字服务法》、中东宗教文化规范），通过全栈式 AI 风控能力实现产品全生命周期风险识别，避免因违规导致的产品下架、巨额罚款。

借助大模型与精细化标签的差异化能力，在多语种、多文化场景下保持审核准召率，解决出海企业“本地化文化与法律法规理解不足、审核效率低”的痛点，同时维护品牌形象、保障用户体验。

内容风险识别：

智能文本 / 图片 / 音频 / 视频识别：有效识别各类场景中出现的涉黄、违禁、色情、辱骂、广告导流、等风险内容，支持 47 种语种（汉语、英语、阿拉伯语等），具备意图识别、语义正负向理解，AI 幻觉识别，多模态识别、生成式长文本识别的能力。

全栈式内容风控：实现全路径实时布控，覆盖内容生产、传播、互动全业务场景，有效识别产品全生命周期的内容风险；通过全流程运营体系（从风险感知到效果监控的迭代闭环），使内容审核准确率高达 99%，助力出海企业在东南亚、中东、欧美等多区域实现全场景内容合规，避免因区域文化差异导致的内容违规风险。

大模型审核能力：十年行业经验积累的千亿级样本沉淀，独立训练的审核大模型具备语义理解、情绪捕捉、意图研判及逻辑推理能力，在多语种复杂语境下（如阿拉伯语宗教敏感内容、东南亚语种民俗禁忌）仍能精准识别违规。模型迭代速度提升至小时级，相较传统方案识别准确率显著领先，解决出海企业 "多语种内容理解不足、审核不准确" 的痛点。

精细化标签体系：凭借独创的 1800 + 四级内容标签体系（含行业首创 "意图标签"），可深度拆解文本、音频、视频中的语义意图与潜在风险（如区分 "玩笑式表述" 与 "恶意煽动"），覆盖泛娱乐出海常见的文化禁忌、宗教敏感等场景，让内容合规从 "表层识别" 升级为 "深层意图管控"。

账号风险识别：

基于先进的人工智能技术和全球 SaaS 风控服务网路，深度融合全栈式模型体系，结合黑产情报、设备风险识别、虚假账号识别等，在业务员场景全流程布控，有效防营销欺诈，诈骗导流识别等风险行为，为企业的出海业务增长报价护航。

注册登录保护：防范批量注册、盗号、撞库等恶意账号行为，保障账号体系安全。

营销活动风险识别：识别营销活动中的作弊行为，如羊毛党薅羊毛、虚假参与等，确保营销资源有效投放。

渠道流量风险识别：检测渠道流量中的异常行为，如刷量、虚假用户导入等，助力企业精准评估渠道质量。

诈骗导流识别：识别账号或行为中的诈骗引流、恶意导流行为，保护用户财产安全与企业商业利益。

保障出海企业用户账号体系的安全合规，符合数据安全、用户隐私保护等法规要求（如GDPR），防范账号层面的恶意攻击与数据泄露风险。

确保营销活动、渠道投放的真实性与有效性，避免因虚假行为导致的资源浪费，同时符合广告投放、数据统计的合规性要求。

防范诈骗导流等恶意行为，维护企业声誉，保护用户权益，规避因用户受损引发的合规纠纷。

扬帆出海：

提供市场策略、本地化资源和行业咨询

扬帆出海四大核心业务矩阵

传媒活动：依托优质出海社群资源，打造高效、高质感的出海主题活动，链接百万出海从业者，助力拓展行业人脉与合作机会。

品牌 PR：构建覆盖多渠道的出海媒体传播矩阵，同时产出行业白皮书、年鉴等专业内容，稳居出海赛道头部资讯平台，是企业提升品牌声量、拓展行业影响力的核心传播渠道。

出海商务通：整合出海客情资源与 B2B 精准人脉网络，打通商务对接链路，助力企业高效推进客户转化，实现销售业绩进阶。

创投平台：全年精选 100+ 优质出海企业项目，组织超 100 场线上线下交流活动，搭建资本与创业者的对接桥梁，共享全球市场增长新机遇。

扬帆出海 PAGC 展会

作为广州官方认证的 “优秀展览项目”，PAGC 全球产品与增长展会是出海领域公认的 “趋势风向标”—— 连续 5 年领跑赛道，2024 年超 80% 参会者现场达成资源对接，更推动众多企业敲定年度核心合作。

PAGC 展会每年会在广州广交会展馆举办，往年，盛会都会吸引超 2 万人到场参会，2025 年，2 万 + 出海全链路精英、3000+C 级决策层，150 + 头部展商，现场带来 AIGC、短剧、游戏等赛道的前沿工具，14 场高价值活动由 100 + 大咖拆解增长方法论。

从行业白皮书发布到 “金帆奖” 标杆表彰，PAGC 不仅是资源对接场，更在重塑出海生态的合作规则，是企业锚定赛道机会、卡位全球市场的年度必赴盛会。

结语
CONCLUSION

文化输出，更要文化尊重

2026 即将到来，全球泛娱乐行业正在进入深度融合与高速增长的新阶段。对已经出海或正在计划出海的中国企业而言，国际市场依然蕴藏着巨大的增量空间——技术实力、内容创新与全球用户需求的契合，为中国企业提供了前所未有的历史机遇。然而，未来"大有可为"，前提是我们在加速迈向世界的过程中，不忽略责任、不忽略规则，更不忽略文化尊重。

内容出海不仅是商业行为，更是文化交流。中国企业要保持文化自信，但文化自信不意味着"自说自话"。真正的全球化，是理解差异、尊重多元，在地化表达、在地化价值。只有建立在尊重基础上的文化输出，才能形成长期信任，实现可持续增长。

与此同时，全球监管不断升级，内容健康度、未成年人保护、数据安全、广告合规、社区治理等，都成为企业能否走得稳、走得远的关键。合规不再是成本，而是能力；不再是阻力，而是竞争力。一个清朗、可信赖的内容生态，对企业本身、行业环境乃至国家品牌都是至关重要的长期资产。

出海清朗，AI 向善

AI 时代的到来，更对企业提出"技术向善"的新要求。AI 既能提升效率、加速全球化，也可能带来虚假内容、广告欺诈、未成年人风险等问题。负责任的 AI、透明的算法、对弱势群体的保护，是所有出海企业必须同步建设的能力。只有让技术在边界内发展，让创新在秩序中繁荣，企业才能真正把 AI 变成全球化的助推器。

未来的竞争，不只是速度之争，而是长期主义之争、治理能力之争、价值观之争。愿所有奔赴全球的中国企业，在文化自信中保持谦逊，在技术创新中坚守善意，在商业拓展中做到清朗可信，与全球用户共同建设一个更加健康、开放与可信赖的数字世界。

附录：政策法规索引

东南亚

印度尼西亚个人数据跨境传输规则 - 朔州市中级人民法院
https://sxszzy.shanxify.gov.cn/article/detail/2024/06/id/7995459.shtml

马来西亚成立国家人工智能政策法规办公室
https://baijiahao.baidu.com/s?id=1818201828898628225&wfr=spider&for=pc

"东南亚方式"：人工智能治理的第三条道路
https://mp.weixin.qq.com/s?_biz=MzUyNDc2Mjg2MQ==&mid=2247485304&idx=2&sn=4afc7346af1d6801ed5c3ca56cc705b3&scene=21#wechat_redirect

垦丁律师事务所《生成式人工智能海外合规白皮书 - 东南亚篇》
https://www.ishumei.com/new/consoleResearch/knowledge/18

泰国公布首部人工智能法案 重点关注风险管控
https://th.mofcom.gov.cn/zcfg/qtfg/art/2025/art_b54d92ecfb3640399454a0d7083cd034.html

泰国成立打击假新闻中心 加强互联网安全
http://www.cac.gov.cn/2019-11/22/c_1575956633135146.html

Thailand unveils 'anti-fake news' center to police the internet
https://www.reuters.com/article/us-thailand-fakenews-idUSKBN1XB48O

"Royal Thai Police", Thailand: Nation, Religion, King
https://www.nationreligionking.com/

Model AI Governance Framework for Generative AI
https://aiverifyfoundation.sg/wp-content/uploads/2024/05/Model-AI-Governance-Framework-for-Generative-AI-May-2024-1-1.pdf

Decree 13/2023/ND-CP on protection of personal data
https://vietanlaw.com/decree-13-2023-nd-cp-on-protection-of-personal-data/

美国：

Sensor Tower: State of Mobile 2025

白宫官网：White House Unveils America's AI Action Plan – The White House

FTC 官网：FTC Adopts Amended Children's Online Protection Act Rule | Wiley Rein LLP - JDSupra

复旦中美友好互信合作计划：美国观察丨"势在必得的竞争"：《美国 AI 行动计划》全局解析

人民法院报：《人工智能法律治理的国际实践》

《科学学研究》网络首发论文：《美国人工智能监管政策的复杂性及临近性研究——基于州域层面政策文本的主题模型》

加拿大：

Sensor Tower: State of Mobile 2025（加拿大 AI 应用市场规模与用户数据）

加拿大政府官网：Canada's Artificial Intelligence and Data Strategy（联邦 AI 战略文件）

加拿大隐私专员办公室（OPC）：OPC Announces Amendments to PIPEDA for AI Data Processing（2024 年 PIPEDA 修订内容）

刑法更新：https://www.lesaonline.org/criminal-law-updates-bill-c-63-online-harms-act/

加拿大司法部官网：Criminal Code of Canada - Sections 163.1, 319（儿童色情与仇恨言论相关刑法条款）

安大略省网站：
https://www.ontario.ca/page/ontarios-trustworthy-artificial-intelligence-ai-framework

加拿大遗产部：Online Safety for Children in Canada - 2025 Guidelines（未成年人在线安全指南）

加拿大公共安全部：Counter-Terrorism and Internet Platforms: 2024 Cooperation Report（反恐与平台合作报告）

国际治理创新中心（CIGI）：Canada's AI Governance: Balancing Innovation and Risk（加拿大 AI 治理分析报告）

观韬解读 https://www.guantao.com/page3997

巴西：

Sensor Tower: State of Mobile 2025

巴西参议院官网：https://www25.senado.leg.br/web/atividade/materias/-/materia/157233

巴西联邦议会官网：https://www.camara.leg.br/proposicoesWeb/fichadetramitacao?idProposicao=13853

巴西国家标准协会官网：https://www.abnt.org.br/normas/todas-normas/?q=NBR+17023-2024 ）

央视新闻：https://content-static.cctvnews.cctv.com/snow-book/index.html?item_id=5782427840033345566

36 氪：https://36kr.com/p/1721135448065

清华大学人工智能国际治理研究院：巴西人工智能发展与中巴人工智能领域合作现状研究

欧洲：

European AI Spending to Reach $144 Billion by 2028, Reflecting 30% CAGR, According to IDC
https://my.idc.com/getdoc.jsp?containerId=prEUR253256125

欧盟加强社交媒体监管
https://www.cac.gov.cn/2020-07/13/c_1596175852288032.htm

Implementation of the Audiovisual Media Services Directive
https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52023IP0134

REGULATION (EU) 2022/2065 OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL
https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32022R2065

DIRECTIVE OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL
https://eur-lex.europa.eu/resource.html?uri=cellar:be0b5038-3fa8-11eb-b27b-01aa75ed71a1.0001.02/DOC_2&format=PDF

REGULATION (EU) 2016/679 OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL
https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32016R0679

欧洲儿童网络安全隐私与数据保护机制
https://www.thepaper.cn/newsDetail_forward_25379397?commTag=true

欧盟迎来重要的网络版权保护新措施
https://ipr.mofcom.gov.cn/article/gjxw/lfdt/oz/bqoz/201904/1935368.html

欧盟加强社交媒体监管
https://www.cac.gov.cn/2020-07/13/c_1596175852288032.htm

中东北非：

UAE Launches AI-Powered "Digital Fraud Hunter" to Shield Against Online Scams
https://www.dayofdubai.com/articles/uae-launches-ai-powered-digital-fraud-hunter-to-shield-against-online-scams

向"西"走，去中东！海湾六国出海新趋势
https://mp.weixin.qq.com/s/fVddqgOljr5s8fJJP2exXw

Updated UAE Media Law Effective 29 May 2025
https://alsuwaidi.ae/updated-uae-media-law-effective-29-may-2025/

UAE's New Media Law Overview: What Media Companies and Influencers Need to Know
https://www.middleeastbriefing.com/news/uaes-new-media-law-overview-media-companies-and-influencers/

The United Arab Emirates' Government Platform
https://u.ae/en

用AI构建清朗可信赖的数字世界

联系我们：

400-610-3866

邮箱：pr@ishumei.com

官网：www.ishumei.com

数美科技公众号

扫码了解更多企业动态

扬帆出海公众号

了解更多出海资讯