

释放人工智能的想象力：生成式人工智能监管全球概览

作者：[Nancy A. Fischer](#)、[Elizabeth Vella Moeller](#)、[Jerry McNerney](#) 阁下、[Amaris Trozzo](#)、[Steven Farmer](#)、[盛佳](#)、[柯呈佶](#)、[廖钰燕](#)、[许春彬](#)、[Fred Ji](#)、[蔡文骏](#)

要点

- 美国：美国政府和国会正在开展制定人工智能监管法律的初步措施，并正在实施一些临时监管措施。例如，白宫最近宣布与七家著名的生成式人工智能公司达成自愿协议，提供最基础的安全、保障和公众信任方面的护栏。
- 欧盟和英国：预计欧盟将在 2023 年底前定稿《欧盟人工智能法案》，该法案将根据风险水平对人工智能使用进行分类。英国政府今年 3 月发布的一份白皮书则授权各个行业监管机构在其管辖范围内对人工智能进行监管，并阐明英国政府计划建立用于支持各个行业监管机构的中央职能机构。
- 中国：今年 7 月，中国发布了首部关于生成式人工智能技术的法规，该法规引入了内容监控、标记和数据来源等诸多方面的服务提供者应当承担的重大义务，同时强调服务提供者应签订明确权责的协议以保护用户个人信息。

2023 年 8 月 11 日

本文讨论了美国、欧洲、英国和中国关于生成式人工智能的最新立法进展。

美国

美国国会领导人正在加紧努力出台立法，指导各个政府机构对人工智能技术的监管。6 月，参议院多数党领袖查克·舒默（纽约州民主党人）公开宣布了 SAFE 创新框架，该框架为人工智能立法设定了优先事项，重点为：安全、可问责性、保护社会的基础和可解释性。该框架的目标是在不影响创新的情况下提供安全保障。尽管目前尚不确定法案能否在 2023 年通过，但我们预计美国国会将在本届会议的剩余时间里继续引入立法、举行听证会并组织人工智能论坛，以便在正式立法颁布之前为人工智能行业人士提供与代表他们的美国国会议员接触的机会。

此外，在 5 月 16 日的参议院司法委员会下辖的隐私、技术和法律小组委员会的听证会上，几位参议员表示支持成立一个新的联邦机构，专门监管人工智能，并可能针对人工智能技术实施许可证制度。也有人呼吁建立一个模仿国际原子能机构的国际人工智能监管机构。

5 月 23 日，白宫透露了在美国全国范围内推进人工智能技术研发和部署的三个新步骤。此外，美国科技政策办公室（OSTP）通过向公众征求意见，以制定一项全面的国家人工智能战略，重点是促进人工智能的公平和透明，同时最大限度地提高人工智能的效益。征求意见期现已结束，公众反馈的结果将公布，并使得 OSTP 进入推进国家人工智能战略的下一阶段。

美国各联邦机构也在与其管辖范围和立法权限范围内推进人工智能相关工作，其时常向受监管实体发布指导意见，解释相关联邦机构将如何将现有法律适用于任何人工智能违法行为。例如，联邦贸易委员会一直积极监管人工智能的欺骗和不公平行为，特别是关注《公平信贷报告法》、《平等信贷机会法》和《联邦贸易委员会法》等法规的执行。

欧盟

欧盟在被称作“人工智能法案”的人工智能法律制定工作中稳步推进，该法案已进入立法程序的最后阶段。欧盟的目标是在 2023 年底前通过该法案的最终稿，而该法案通过后可能须经过 24 个月的过渡期方能正式生效。

正在审议中的人工智能法案按照风险等级对人工智能应用进行了分类。法案禁止人工智能的某些用途（例如，在公共场所使用的实时生物识别监控系统和可能造成损害的潜意识技术），并对具有更高风险的应用实施更严格的监控和披露要求。

欧盟的目标是确保在欧洲开发和使用的的人工智能符合该地区的价值观和权利，包括确保人工智能获得人类监督、具有安全性、隐私性和透明度、遵从非歧视原则以及社会和环境福祉。目前法案规定的违约罚款可能高达违法者全球收入的 7% 或 4,000 万欧元。

有关欧洲《人工智能法案》的更多详细分析，请参阅[此处](#)。

此外，2022 年 9 月 28 日，欧盟委员会还提出制定一项新的法律，称为“人工智能责任指引”，旨在在人工智能领域适用非合同民事责任规则。拟议的法律（与《人工智能法案》密切相关）旨在为人工智能系统造成的损害制定统一规则，为受害者提供更广泛的保护，并通过提高对人工智能系统的保障来促进人工智能行业发展。它将解决与人工智能相关的具体举证困难问题，要求欧盟成员国授权该国法院下令披露有关特定高风险人工智能系统的相关证据。该拟议法律将影响人工智能系统的用户和开发者，明确开发者在人工智能系统出现故障时的责任，并便于人工智能系统相关犯罪的受害者进行索赔。关于该新法律的磋商正在进行中，尚不清楚何时通过。

英国

2023 年 3 月 29 日，英国政府发布了一份白皮书，阐释了其支持创新的人工智能监管措施。现有行业监管机构将有权在各自行业内监管人工智能，而无需制定新的法律或设立单独的人工智能监管机构。白皮书的重点是改进现有制度以对人工智能进行监管，并避免可能阻碍创新的严格立法。

白皮书中提出的监管框架基于两个关键特征对人工智能进行了定义，即适应性和自主性。白皮书认为，通过参照这两项特征去区分人工智能，并设计监管框架应对这些特征所带来的挑战，英国立法者可以避免意外出现的新技术对监管框架的冲击。

白皮书还列出了监管机构在应对人工智能相关风险时应遵守的五项“以价值观为中心的跨行业”的原则。这些原则包括：(i) 安全、保障和稳健性，(ii) 适当的透明度和可解释性，(iii) 公平性，(iv) 问责制和治理，以及 (v) 可挑战性和补救。

白皮书列出的原则以经济合作与发展组织（OECD）基于价值观制定的人工智能原则的为基础。前者反映了英国政府对后者的承诺，后者促进了人工智能合乎道德的使用。白皮书采取以原则为基础的监管手段，目的是使框架更为灵活和符合比例性原则。虽然一开始不具有法律约束力，但英国政府预计，这些原则可能在未来变得具有执行性，这取决于人工智能技术的发展前景及其社会影响。

除了这些原则外，英国政府还计划设置中央职能机构，以支持监管机构发挥人工智能监督作用，并确保监管框架运作符合比例性原则且支持创新。白皮书没有提及由哪一个或哪些具体实体承担这些核心职能。

在发布白皮书的同时，英国政府还投资 200 万英镑资助了一个新的沙盒，让人工智能创新者能够在新的人工智能产品上市前对其进行测试。沙盒将使企业能够测试人工智能法规如何适用于其产品。

白皮书发布后，英国政府将继续与企业 and 监管机构合作，开始建设其规划中指出的核心职能机构。英国政府将公布一份人工智能监管路线图，并对有关白皮书提出的咨询做出回应。从长远来看，在白皮书发布后 12 个月或更长时间，英国政府计划完成所有核心职能机构的成立，支持监管机构应用跨部门原则，发布人工智能风险登记草案，开发监管沙盒，并发布评估框架效果的监测和评估报告。

中国

2023 年 7 月 13 日，中国国家网信办发布了最终版的《生成式人工智能服务管理暂行办法》（《生成式 AI 办法》）。该《生成式 AI 办法》适用于在中华人民共和国境内向公众提供的生成式人工智能服务，并主要监管利用人工智能技术进行的内容生成（生成式 AI 服务）。《生成式 AI 办法》明确将从事生成式人工智能技术研发和应用且未向公众提供服务的行业组织、企业、教育和研究机构、公共文化机构等排除在其适用范围之外。

《生成式 AI 办法》规定生成式人工智能服务提供者须承担多项重要义务，包括监控和控制其服务生成的内容。供应商应立即删除任何非法内容，对从事非法活动的用户采取行动，并向有关部门报告。此外，提供者必须用适当的标签标记生成的内容，并使用合法来源进行数据训练，同时尊重知识产权并获得个人信息处理的同意。《生成式 AI 办法》重申了中国现有的网络安全和个人隐私规则，规定保护用户的个人信息，禁止非法收集和共享可识别数据。

中国也可能采取以行业为导向的监管模式，由不同的政府部门在各自的特定领域监管生成式人工智能服务，预计将出台针对行业的人工智能法规和分类指南。

《生成式 AI 办法》是对《互联网信息服务算法推荐管理规定》（以下简称《算法规定》，2022 年 3 月 1 日起生效）和《互联网信息服务深度合成管理规定》（以下简称《深度合成规定》，2023 年 1 月 10 日起生效）的最新补充。

《算法规定》适用于在中国大陆使用算法推荐技术（包括但不限于生成合成、个性化推送、排序精选、检索过滤、调度决策等技术）提供互联网信息服务的任何实体。《算法规定》等规定要求，具有舆论属性或社会动员能力的算法推荐服务提供商（包括生成式人工智能服务提供商）应按照申请规定进行安全评估，并在开始提供服务之日起 10 个工作日内完成在线备案手续。

《深度合成规定》借助“深度合成技术”概念实现对中国大陆互联网信息服务提供行为的规范。“深度合成技术”的定义为“使用生成排序算法（如深度学习和虚拟现实）创建文本、图像、音频、视频、虚拟场景或其他信息的技术”。《深度合成规定》为深度合成服务提供商和技术支持者在数据安全和个人信息保护、透明度、内容管理和标签、技术安全等方面规定了一套全方位的责任。生成式人工智能服务提供商需要根据《深度合成规定》在内容（生成式人工智能生成的图片、视频或其他内容）上添加标记。

总结而言，《生成式 AI 办法》旨在结合中国现有的网络安全和数据隐私规则，建立确保生成式人工智能服务的使用负责任且透明的框架。这一框架为服务提供者保留了一定灵活性，同时也赋予其重大责任。中国政府重视出台鼓励人工智能创新和大规模工业应用的产业政策，而非限制人工智能技术的发展，这也在《生成式 AI 办法》中得以体现。

关于我们对中国《生成式人工智能服务管理暂行办法》的详细讨论，请参阅[《中国发布首个生成人工智能管理办法》](#)。

结论

全球的人工智能治理的战略多样化。欧盟专注于特定行业的监管，美国选择以联邦指导为主、地方调整为辅的去中心化方法，中国则优先考虑消费者透明度和全球人工智能标准的主导地位。企业需要制定有关人工智能伦理和产品合规性的全球立场，以遵守新的监管规定。