Anjie Broad

全球生成式人工智能监管

研 究 报 告

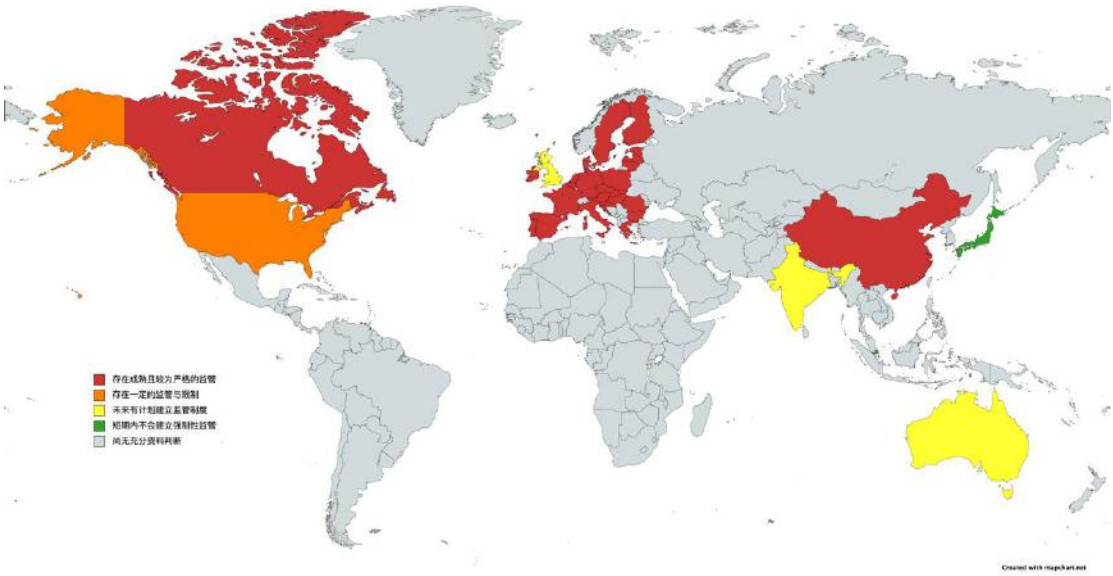
2023.08

目录

一、生成式人工智能监管地图

截止 2023 年 7 月 24 日，我们依据对全球生成式人工智能监管动向的梳理，绘制了监管地图：

1. 红色区域表明该地区目前已经或即将建立成熟且较为严格的监管制度，主要包括欧盟、加拿大和中国；
2. 橙色区域表明该地区目前对生成式人工智能产品存在一定的限制和监管要求，例如美国；
3. 黄色区域表明该地区目前还没有相关制度，但是监管机构未来有计划对生成式人工智能制定规则，例如英国、澳大利亚和印度；
4. 绿色区域表明该地区目前没有相关制度，且监管机构也倾向于在未来不采取强制性的监管措施，例如新加坡、日本、中国香港。



注：地图颜色仅表明不同地区在生成式人工智能领域法律监管的程度。

二、主要地区监管概述

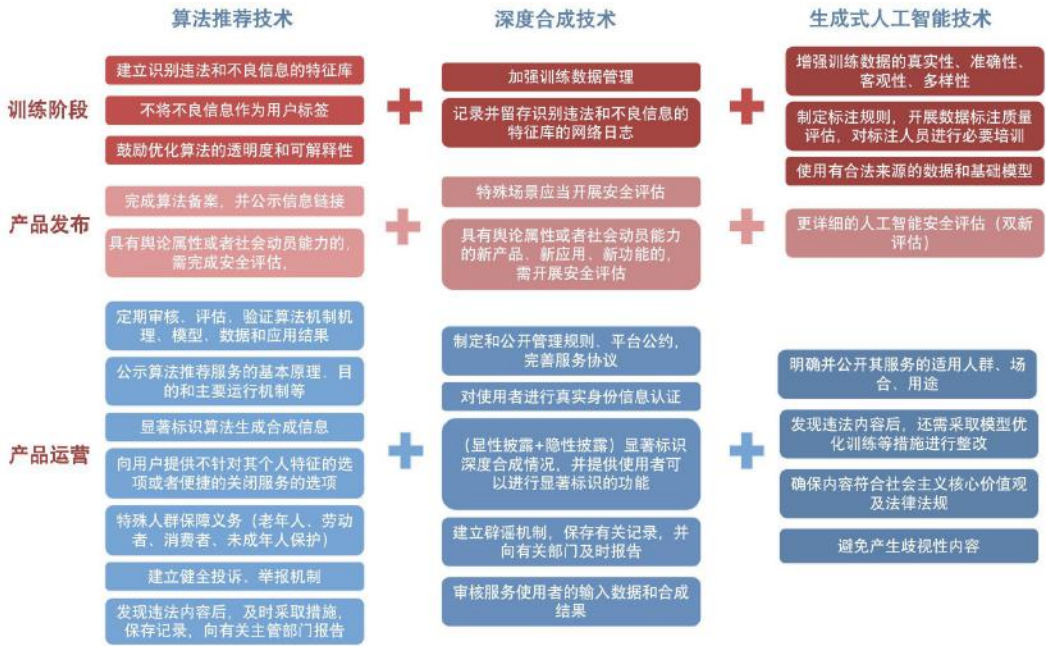
（一）中国

中国监管机构从 2021 年起加强了对人工智能的监管，并陆续发布了《互联网信息服务算法推荐管理规定》（“算法推荐规定”）《互联网信息服务深度合成管理规定》（“深度合成规定”）和《生成式人工智能服务管理暂行办法》（“生成式办法”）三部重要的部门规章。这三部规章都可能适用于向中国境内公众提供的生成式人工智能服务。而在未来，中国还可能制定层级更高的《人工智能法》，并落实科技伦理审查制度。

与此同时，这三部规范也确立了中国将采取“包容审慎的分类分级监管”原则。在未来，国家有关主管部门针对生成式人工智能技术特点及其在有关行业和领域的服务应用，将制定相应的分类分级监管规则或者指引。

目前负责算法备案、评估与监管的机构主要是网信办（中央网信办网络管理技术局）。与此同时，电信主管部门、公安部门也会依据职责负责监督管理工作。

针对生成式人工智能的主要合规义务：



（二）欧盟

欧盟现行人工智能立法仍主要集中在传统人工智能而非生成式人工智能上，但也逐渐触及生成式人工智能的问题，例如在最新的《人工智能法》草案中规定了通用及生成式人工智能相关基础模型提供者的义务。

欧盟的人工智能路径主张“以人为本（human-centric）”，在促进人工智能发展与创新的同时，构建监管体系以防范风险、保护公民基本权利和安全。

欧盟目前对于人工智能的监管和治理重点在于尊重人格尊严、个人自由和保护数据及隐私安全。最新的立法进展是 2023 年 6 月 14 日欧洲议会对《人工智能法》的修订。该法采用“基于风险的方法”（risk-based）将人工智能系统分成四类，制定了不同程度的合规要求。其中，生成式人工智能系统一般属于有限风险的人工智能系统，需遵守最低限度的透明度义务，但可能会因其适用的领域和生成的内容而落入高风险人工智能系统的范畴。[1] 与此同时，该法也规定了提供者、进口商、分销商、部署者、通用及生成式人工智能相关基础模型提供者等不同主体的合规义务。

根据侵权行为的性质、严重程度等因素，不遵守《人工智能法》的相关当事人可能被处以不同规模的行政罚款。

	个人	公司
将禁止的人工智能系统投放市场或投入使用	最高可被处 4000 万欧元罚款	罚款 4000 万欧元或上一财政年度全球年营业额的 7%（以较高者为准）
不遵守高风险人工智能系统关于数据治理、透明度和提供信息给用户的规定	最高可被处 2000 万欧元罚款	罚款 2000 万欧元或上一财政年度全球年营业额的 4%（以较高者为准）

[1] European Parliament News: https://www.europarl.europa.eu/news/en/headlines/society/20230601STO93804/eu-ai-act-first-regulation-on-artificial-intelligence

违反其他对于人工智能系统或模型的要求和义务	最高可被处 1000 万欧元罚款	罚款 1000 万欧元或上一财政年度全球年营业额的 2%（以较高者为准）
在答复请求时向公告机构和国家主管当局提供不正确、不完整或误导性的信息	最高可被处 500 万欧元罚款	罚款 500 万欧元或上一财政年度全球年营业额的 1%（以较高者为准）

历时两年有余，《人工智能法》草案的提出和修订体现了欧盟对人工智能的严格监管态度。尽管最新的草案修订稿对创新发展的考量增多，但总体而言，欧盟对人工智能的监管仍非常严苛。

在执法方面，近年来，已有多个包括意大利、法国、西班牙等多个欧盟成员国对 ChatGPT 展开调查。

在监管方面，欧盟目前没有专门针对人工智能监管的独立机构，主要由欧盟层面及成员国层面不同领域的主管部门联合监管和执法。最新的《人工智能法》草案要求成员国指定一个或多个主管机构，包括一个国家监管机构，负责监督《人工智能法》的应用和实施；并提议在欧盟层面建立由成员国和欧盟委员会的代表组成的新欧盟机构——“人工智能办公室”，以期为《人工智能法》的统一适用提供指导、协调联合跨境调查。此外，欧洲数据保护委员会(EDPB)、国家市场监督机构、海关、国家消费者保护当局会负责各自领域与人工智能有关的执法工作。

《人工智能法》中通用及生成式人工智能相关基础模型提供者的特殊义务：

基础模型的提供者	生成式人工智能系统中使用的基础模型的提供者 将基础模型专门用于生成式人工智能系统的提供者
• 识别、减少和缓解对健康、安全、基本权利、环境以及民主和法治的合理可预见风险 • 仅处理并纳入受适当基础模型数据治理措施约束的数据集 • 设计、开发基础模型以实现适当性能 • 利用公布后的协调标准设计、开发基础模型，减少资源浪费 • 基础模型的设计应具有测量、记录资源消耗的能力，技术上可行的情况下记录系统的部署、使用在其生命周期中可能产生的其他环境影响 • 起草广泛的技术文件和可理解的使用说明，使下游提供者能够遵守其义务 • 建立质量管理体系，以确保并记录其符合义务，并有可能进行实验以满足这一要求 • 在欧盟数据库中注册该基础模型	
/	• 训练、设计模型以防止生成非法内容 • 记录、公开提供受版权法保护的训练数据的详细使用总结 • 设计和开发旨在与自然人交互的人工智能系统，应使该系统、供应商或用户及时、清晰和可理解地向与系统交互的自然人披露交互的事实（除非从环境和使用背景中可以明显看出）、人工智能启用的功能、决策负责者、就损害寻求司法救济的现有权利和程序、寻求解释的权利 □ 不适用于：法律授权用于侦查、预防、调查和起诉刑事犯罪的人工智能系统，除非这些系统可供公众举报刑事犯罪

《人工智能法》中包括生成式人工智能在内的有限风险的人工智能系统的透明度义务：

与透明度有关的主体义务	
• 供应商 □ 确保旨在与自然人交互的人工智能系统的设计和开发方式能使自然人及时、清晰和可理解地知晓其正与AI系统交互，除非从环境和使用背景中可以明显看出这一点 □ 如相关，应告知人工智能启用的功能、决策负责者、就损害寻求司法救济的现有权利和程序、寻求解释的权利 □ 不适用于：法律授权用于侦查、预防、调查和起诉刑事犯罪的人工智能系统，除非这些系统可供公众举报刑事犯罪	• 人工智能系统未经用户同意生成或操纵看似真实或真实的文本、音频或视觉内容（深度伪造） □ 以适当、及时、清晰和可见的方式披露该内容是人为生成或操纵的，并且尽可能披露产生或者操纵该信息的自然人的姓名或者法人的名称 □ 不适用于：法律授权使用生成或操纵文本、音频或视觉内容的人工智能系统、行使《欧盟基本权利宪章》所保障的言论自由权和艺术和科学自由权所必需的人工智能系统 □ 如内容构成明显具有创造性、讽喻性、艺术性或虚构的电影、视频游戏视觉效果和类似作品或节目的一部分，则此义务仅限于以适当、清晰和可见的方式披露此类生成或操纵内容的存在
• 未被禁止的情绪识别系统或生物识别分类系统的用户 □ 以及时、清晰和以可理解的方式告知暴露在该系统下的自然人系统的运行情况 □ 在处理个人生物识别和其他数据前征得其同意 □ 不适用于：法律允许用于检测、预防和调查刑事犯罪的生物识别分类人工智能系统	

（三）美国

美国是目前人工智能技术的策源地和领导者。为了保持这一领导地位，美国政府将促进人工智能的创新和发展作为高度优先事项。

因此，美国主张监管需以促进人工智能负责任的创新（responsible innovation）为目标，应通过监管和非监管措施减少人工智能开发和部署的不必要障碍，同时保护美国的技术、经济和国家安全、公民自由、人权、法治、隐私和尊重知识产权等核心价值观。

目前美国对人工智能的关注重点在于公民权利、公民自由、数据与隐私安全及知识产权保护。

尽管美国的部分州已通过与人工智能有关的法案，但美国目前尚无专门规范通用人工智能或生成式人工智能的法律性文件。虽然两党议员尝试就不同人工智能事项提请立法，有趋势表明美国对人工智能的监管或将加强，但联邦立法监管目前仍处于非常早期的阶段。细分领域不具有法律强制性的非监管措施如主管部门的政策文件、指南、治理框架、试点项目和实验、行业自律、社会组织的共识等是目前人工智能企业在美国行事的主要参考和依循。

需要注意的是，美国政府有对于生成式人工智能相关的产品实施严格进出口管制的趋势。美国政府曾严格限制用于人工智能工作的顶级计算芯片、超级计算机和其他半导体产品对中国的出口。此外，2023 年 7 月 4 日，媒体报道美国政府还计划限制中国企业使用美国的云计算服务。[2]

美国目前没有专门针对人工智能监管的独立机构，各领域主管部门将继续在自己的职责范围内对生成式人工智能进行监管。

[2] https://www.reuters.com/technology/us-set-restrict-chinas-access-cloud-computing-wsj-2023-07-04/

为包含人工智能生成材料的作品申请版权保护的申请人有以下义务：

为包含人工智能生成的材料的作品申请版权保护的申请人的义务

- **提交新的作品登记申请**
  - ✓ 披露作品中包含的人工智能生成的内容
  - ✓ 简要说明人类作者对作品的贡献
  - ✓ 简要描述人工智能生成的内容，在申请表中将其明确排除
  - ✓ 不应仅仅因为在创作作品时使用了人工智能技术而将该技术或提供该技术的公司列为作者或合著者
- **已提交申请，申请中包含人工智能生成的内容**
  - ✓ 检查其向版权局提供的资料是否充分披露了人工智能生成的内容，否，则应采取措施更正信息

  办公室将在符合条件的情况下颁发新的补充注册证书，其中包含针对人工智能生成材料的免责声明
- **已提交但待决的申请**
  - ✓ 联系版权局的公共信息办公室（Public Information Office），报告其申请遗漏了作品包含人工智能生成材料的事实
- **已处理并已注册的申请**
  - ✓ 提交补充注册
  - ✓ 更新公开记录

近期与生成式 AI 相关的监管动态：2023 年 7 月 21 日，Amazon、Anthropic、Google、Inflection、Meta、Microsoft 和 OpenAI 向美国政府自愿作出一系列承诺。

七大AI巨头向美国政府作出的承诺

- **确保产品安全后才向公众推广**
  - ✓ 在其人工智能系统发布前对系统进行内部和外部安全测试（部分测试由独立专家进行）
  - ✓ 在整个行业以及与政府、民间社会和学术界分享有关人工智能风险管理的信息，包括安全方面的最佳实践、试图规避安全措施的信息以及技术合作
- **建立把安全置于首位的系统**
  - ✓ 投资网络安全和内部威胁防范措施，以保护自有的和未发布的模型权重
  - ✓ 只有在充分考虑了安全风险后，才能发布模型权重
  - ✓ 促进第三方发现和报告其人工智能系统中的漏洞
- **赢得公众的信任**
  - ✓ 开发强有力的技术机制（如水印系统），确保用户知道内容何时是人工智能生成的
  - ✓ 公开报告其人工智能系统的能力、局限性、适于及不适于使用的领域、系统的安全风险和社会风险（如对公平和偏见的影响）
  - ✓ 优先研究人工智能系统可能带来的社会风险（包括避免有害的偏见、歧视及保护隐私）
  - ✓ 开发和部署先进的人工智能系统

（四）加拿大

加拿大将自己定位为人工智能发展的领导者。为此，加拿大政府主张对人工智能采取灵活的监管方法，在促进人工智能系统设计、开发和使用的安全性的同时，维护加拿大民众的价值观，亦不致扼杀负责任的创新。

加拿大对人工智能的监管重点在于隐私保护。

目前加拿大没有明确针对人工智能的监管制度，加拿大的人工智能系统受到一般性的隐私、技术和人权法律的监管。虽然加拿大尚未处于制定全面的人工智能监管制度的阶段，但联邦和省两级都在采取行动弥补监管空白，以期让加拿大民众信任人工智能技术。2022 年 6 月提出的《人工智能和数据法》作为加拿大隐私领域立法改革的一部分，是加拿大针对人工智能进行监管的首次尝试。该法采用“基于风险的方法”（risk-based）对高影响人工智能系统的风险进行重点监管，具体标准有待后续法规进行界定。该法为人工智能系统负责人即设计、开发、使用、提供或管理人工智能系统运行的主体设定了合规义务，在相关活动中处理或者使用数据的主体也承担一定的透明度义务。违反该法的当事人可能面临行政罚款、行政诉讼甚至刑事制裁。

《人工智能和数据法》下人工智能系统负责人的合规义务：

任何人工智能系统负责人的一般义务	高影响人工智能系统负责人的特殊义务
• 依规评估该系统是否为高影响系统 • 依规就数据匿名化的方式、匿名数据的使用或管理制定措施 • 保存记录 ✓ 有关缓解措施（包括缓解任何伤害或偏差输出风险的有效性）的一般记录 ✓ 支持该系统（不）属于高影响系统的原因	
/	• 建立识别、评估和减轻因使用该系统而可能造成的伤害或偏差输出风险的缓解措施 • 建立措施以监督缓解措施的遵守情况 • 系统的使用导致或可能导致重大损害时：尽快通知 ✓ 被指定的加拿大枢密院的成员 ✓ 如无被指定的成员：创新、科学与工业部部长

高影响人工智能系统提供者、管理运行者的其他义务

- 在公开网站上发布系统的简单说明，内容包括
  - ✓ 系统的预期使用方式（提供者）/ 如何使用该系统（管理运行者）
  - ✓ 拟生成的内容类型、拟作出的决策、建议或预测
  - ✓ 现行的缓解措施
  - ✓ 法规规定的其他信息

违反合规义务的人工智能系统负责人将面临下表所示的处罚。

	个人	公司
经公诉定罪	法院酌情罚款	不超过 1 千万加元与其被判刑前的财政年度内全球总收入的 3%的罚款中更高额的罚款
经简易程序定罪	不超过 5 万加元罚款	不超过 500 万加元与其被判刑前的财政年度内全球总收入的 2%的罚款中更高额的罚款

在实践方面，加拿大的两个地区法院对于在法律意见书等材料中使用人工智能（一般是生成式人工智能技术）发布了实务指导。

在执法方面，2023 年 4 月，加拿大隐私专员办公室启动对 OpenAI 的调查，5 月下旬，多省数据保护机构与加拿大隐私专员办公室合作，对 OpenAI 展开联合调查。

在监管方面，加拿大目前没有专门针对人工智能进行监管的独立机构，主要由不同领域的主管部门监管和执法。正在审议的《人工智能和数据法》规定创新、科学与工业部部长负责执行该法；提议建立由新的人工智能和数据专员领导的办公室，作为支持该法监管和执行的专门知识中心。此外，加拿大检察署、加拿大隐私专员办公室等部门负责各自领域与人工智能有关的执法工作。

（五）英国

英国希望成为世界上建立和发展人工智能业务的最佳场所，以便将人工智能

的巨大潜力转化为英国的经济增长和社会效益。因此，英国主张建立支持创新（pro-innovation）、合比例（proportionate）、可信、适应性强、明确和协作的监管框架，以期为投资者、企业和公众提供信心。

英国不希望繁重的合规义务打击企业的创新意愿，因此没有专门针对人工智能进行立法，而主要通过发布指南、最佳实践等非监管措施为企业提供指导，鼓励促进人工智能领域负责任的创新，且对生成式人工智能这类新兴的人工智能的发展持较开放的态度。例如，财政部建议让创新者和企业家在监管沙盒（regulatory sandbox）中试验新产品或服务而无需担心罚款或责任风险，存在监管不确定性的生成式人工智能也可以使用监管沙盒。2023 年 6 月，英国内阁办公室发布了《公务员使用生成式人工智能的指引》，介绍了公务员使用生成式人工智能的一般原则。

英国对生成式人工智能风险的关注重点在于数据和隐私保护。教育部在《教育领域中的生成式人工智能》中明确指出，为保护个人隐私，个人信息和敏感数据不应输入生成式人工智能工具。英国信息专员办公室也提示开发或使用生成式人工智能的人负有数据保护的法定义务，他们需要考虑处理个人数据的合法依据为何、处理或控制数据的身份、通过数据保护影响评估、公开有关处理信息、降低安全风险、限制不必要的处理、尊重个人访问、纠正、删除等权利请求、是否完全使用生成式人工智能自动决策等问题。[3] 对于人工智能生成的作品，英国知识产权署在 2022 年 6 月公布咨询结果，认为暂时没有证据表明人工智能生成的作品是有害的，其仍然受到英国版权法的保护。[4]

在执法方面，2023 年 5 月，英国竞争与市场局对开发和使用人工智能基础

[3] Generative AI: eight questions that developers and users need to ask | ICO
[4] Artificial Intelligence and Intellectual Property: copyright and patents: Government response to consultation - GOV.UK (www.gov.uk)

模型涉及市场竞争和消费者保护的方面进行初步审查。[5] 2023 年 6 月，教育部启动公众咨询，寻求民众关于生成式人工智能在英国教育中的应用及其机遇与风险的看法。[6]英国信息专员办公室呼吁企业在使用生成式人工智能前“应对隐私风险”，同时宣布“将对企业是否遵守数据保护法律进行更严格的检查”。[7]

英国没有专门监管人工智能的独立机构，各领域监管机构在其执法范围内对相关行为进行监管。如英国信息专员办公室作为英国的数据保护机构将对人工智能相关主体处理数据的行为进行监管，英国竞争与市场局将对不公平竞争、损害消费者的行为进行监管等。

（六）澳大利亚

澳大利亚的人工智能方法希望确保人工智能可信赖、安全和负责任。

澳大利亚目前主要发布了人工智能的道德原则，尚未决定对人工智能进行硬性监管。

澳大利亚尚未制定专门针对人工智能的立法，所有人工智能的治理都依赖于现行法律和指南。

近期与生成式人工智能有关的动态是 2023 年 7 月，数字转换机构和工业、科学与资源部合作发布了《关于政府使用生成式人工智能平台的临时指南》，指出公共部门使用生成式人工智能需遵守四项原则：负责任地使用；透明度和可解释性；隐私保护和安全；决策以人为本、落实问责制。[8] 此外，澳大利亚知识产权局正在对生成式人工智能对知识产权的影响进行研究。[9]

---

[5] CMA：https://www.gov.uk/government/news/cma-launches-initial-review-of-artificial-intelligence-models
[6] https://www.gov.uk/government/consultations/generative-artificial-intelligence-in-education-call-for-evidence
[7] Don’t be blind to AI risks in rush to see opportunity – ICO reviewing key businesses’ use of generative AI | ICO
[8] https://architecture.digital.gov.au/guidance-generative-ai /
[9] Generative AI And The IP System – What Does It All Mean? | IP Australia

澳大利亚没有专门针对人工智能进行监管的机构，各领域主管部门将在其主管范围内对人工智能进行监管。澳大利亚信息专员办公室负责执行隐私法律，有权对人工智能涉嫌侵害隐私的行为进行调查。澳大利亚工业、科学与资源部经常参与人工智能政策文件的制定。

（七）印度

印度没有人工智能监管的正式法律法规。与此同时，人工智能产品息息相关的数据保护领域，印度目前也迟迟没有通过多次修改的《个人数据保护草案》。目前，仅有《信息技术法案》（*Information Technology Act, 2000*）与《信息技术（合理的安全实践和程序及敏感个人数据或信息）规则》（*Information Technology (Reasonable Security Practices and Procedures and Sensitive Personal Data or Information) Rules, 2011*）要求数据处理者维持合理的安全措施和程序，制定隐私政策，在收集或传输敏感的个人数据或信息时获得同意，并将此类收集的数据告知接收者的数据主体。

但是，近年来，印度监管者一直在积极研究人工智能监管的议题。2023 年 3 月起，印度政府的国家人工智能门户网站 INDIAai 组织学术界的知名人士举办了三场圆桌讨论会，讨论研究生成式人工智能的影响、伦理和监管问题以及它给印度带来的机遇。[10]印度电子和信息技术部部长钱德拉塞卡(Rajeev Chandrasekhar)在接受媒体采访时表示，政府正计划监管人工智能(AI)，以保护公民免受伤害。[11]而 7 月 20 日，印度电信监管局(Trai)也建议成立印度人工智能和数据管理局

[10] https://indiaai.gov.in/research-reports/impact-opportunity-and-challenges-of-generative-ai

[11] https://www.livemint.com/economy/india-to-regulate-ai-for-user-protection-rajeev-chandrasekhar-11686335676103.html

(AIDAI)作为独立的监管机构。[12]

（八）日本

日本希望实现以人为中心、网络空间和物理空间高度融合从而可以促进经济发展、解决社会问题的“社会 5.0”（society 5.0）。为此，日本对人工智能的治理提出了“以人为本的人工智能社会原则”，该原则包括以人为本、教育、隐私保护、确保安全、公平竞争、公平、问责制和透明度、创新七个原则。

日本没有选择政府作为唯一制定、监督和执行人工智能规则主体的治理模式，而是选择对人工智能实行敏捷治理（agile governance）。这是一种多个利益相关主体共同治理的模式。不同的利益相关者包括政府、企业、个人和社区将对其所处的社会状况进行持续分析，明确意图实现的目标并设计各种系统实现这些目标，且对结果进行持续评估，以改进这些系统。

日本目前没有专门规制人工智能的法律，主要采用软法如政策性文件、指导性文件等对人工智能进行规范，鼓励人工智能用户和企业自愿采取行动以实现“以人为本的人工智能社会原则”。

日本目前对人工智能的监管关注点主要在于隐私和数据保护，主要聚焦医疗、农业、交通、教育等领域，其中医疗领域的规范最为成熟，配套监管措施也最多。

近期关于生成式人工智能的动态：2023 年 4 月，日本教育、文化、体育、科学和技术大臣 Keiko Nagaoka 在地方会议上证实，日本的法律不会保护人工智能数据集中使用的受版权保护的材料，“在日本，无论是非营利目的、营利目的、复制以外的行为，还是从非法网站等获取的内容，无论采用何种方法，都可以被

12 https://www.livemint.com/technology/tech-news/trai-issues-recommendations-on-ai-says-regulatory-framework-for-development-of-responsible-ai-urgently-needed-11689859911432.html

用于信息分析。”[13] 可见日本促进创新和进步、增强其在人工智能领域竞争力的决心。2023 年 6 月，日本个人信息保护委员会对 OpenAI 发出警告，由于担心 ChatGPT 收集敏感数据，该委员会可能会采取行动。日本个人信息保护委员会表示，OpenAI 只能在获得许可的情况下收集数据并应尽量减少收集行为，同时要求 OpenAI 考虑隐私保护和生成式人工智能的潜在好处之间的平衡。[14]

日本没有专门对人工智能进行监管的机构，各领域主管机构将在其各自职权范围内对人工智能进行监管。日本个人信息保护委员会负责保护包括个人信息在内的个人权益，人工智能处理数据的行为将落入该委员会的监管范围内。日本经济产业省作为负责提高民间经济活力、维护对外经济关系、保护经济产业发展的部门经常参与人工智能治理的政策、指南等文件的制定。

（九）新加坡

新加坡希望建立可信赖和负责任的人工智能生态系统。

新加坡目前没有考虑对人工智能进行监管[15]，尚无专门规制人工智能的法律。已有的文件主要是指南、实践指引等。新加坡发布了《模范人工智能治理框架》（第 2 版）及其配套文件《人工智能治理框架——机构实施及自我评估指引》《用例汇编》，帮助企业负责任地部署人工智能。

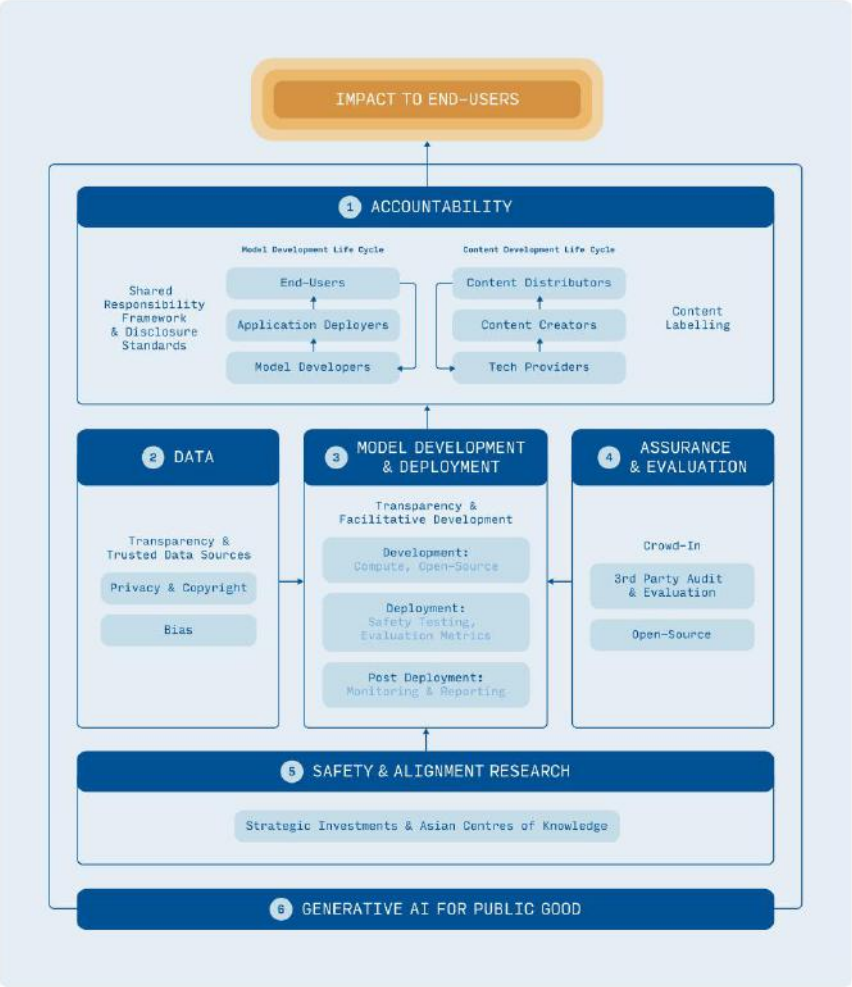
近期与生成式人工智能有关的动态：新加坡信息通信媒体发展局发布讨论文件《生成式人工智能：对信任和治理的影响》，为政府和企业的高级领导人提出了建立可信赖和负责任地使用生成式人工智能的生态系统的想法。

---

[13] 决算行政監視委員会分科会質疑を振り返る。-さいたかし（キイタカシ）/選挙ドットコム (go2senkyo.com)

[14] https://www.reuters.com/technology/japan-privacy-watchdog-warns-chatgpt-maker-openai-data-collection-2023-06-02/

[15] Singapore is not looking to regulate A.I. just yet, says the city-state (cnbc.com)



新加坡没有专门对人工智能进行监管的机构，现有主管部门将在其职能范围内对人工智能企业进行执法。如新加坡个人数据保护委员会将从数据保护的角度对人工智能的数据使用、处理等行为进行监管。

值得一提的是，新加坡资讯通信媒体发展局和新加坡个人数据保护委员会推出了全球首个人工智能治理测试框架和工具包——“A.I. Verify”。A.I. Verify 通过技术测试和流程检查可以帮助人工智能企业对其人工智能系统进行评估，验证系统是否符合国际公认的人工智能原则、系统技术的公平性、可解释性和稳健性以及根据企业合规需求定制报告等。

（十）中国香港

中国香港目前对人工智能的监管持非常包容和开放的态度。创新科技及工业局局长孙东在回应记者问题中说，留意到国家最近针对有关技术和应用提出新措施，特区政府一直保持高度关注，将继续持非常开放态度对待有关技术发展，适时作出应对。

近几年来，中国香港主要强调人工智能的开发和使用应合乎道德。从 2018 年到 2022 年，香港发布了包括《中国香港的道德问责框架》《开发及使用人工智能道德标准指引》《人工智能道德框架》和《应用人工智能的高层次原则》在内的多份指导性文件。这类不具有法律上约束力的政策性文件是香港目前规范人工智能的主要文件。

中国香港对人工智能的风险关注点主要集中在数据和隐私保护。

香港金融管理局在《应用人工智能的高层次原则》中表示，过分严格的规定可能会阻碍人工智能相关技术的进一步发展，因此只在指南中规定了设计和采用人工智能时应考虑的高层次的原则。可以预见，香港在短期仍将对人工智能保持非常宽容的态度。

中国香港地区没有设立监管人工智能的独立机构。实践中，个人资料私隐专员公署有权对人工智能开发和使用过程中的数据处理行为进行监管。创新、科技与工业局作为负责科技发展的政府部门，也经常参与到人工智能监管与发展政策、战略的制定中。

三、中国

（一）监管对象

这三部规章规制的对象“算法推荐技术”、“深度合成技术”与“生成式人工智能”，从技术角度解释和法律文义解释的视角看，存在包含关系。因此概念范围最狭窄的生成式人工智能服务可能需要同时满足三部规章的义务要求。[16]

其中，《生成式人工智能服务管理暂行办法》规制的是向中国境内公众提供的生成式人工智能服务。但有学者指出，“公众”一词目前尚未有明确的定义和解释，因此不能排除 to B 的服务将完全豁免于该办法的适用。同时，来源于中国境外向境内提供的生成式人工智能服务也受到该办法的管辖。

[16] 张凌寒，《深度合成治理的逻辑更新与体系迭代——ChatGPT 等生成型人工智能治理的中国路径》；葛梦莹，南钰彤，《专家解读|生成式人工智能、深度合成技术与算法推荐的关联与合规要点》

三部规章在适用范围的定义上均采取了“技术+提供服务”的定义模式

规章	技术	对技术的解释	提供的服务
《互联网信息服务算法推荐管理规定》	算法推荐技术	生成合成类、个性化推送类、排序精选类、检索过滤类、调度决策类等算法技术	互联网信息服务
《互联网信息服务深度合成管理规定》	深度合成技术	利用深度学习、虚拟现实等生成合成类算法制作文本、图像、音频、视频、虚拟场景等网络信息的技术	互联网信息服务
《生成式人工智能服务管理暂行办法》	生成式人工智能技术	具有文本、图片、音频、视频等内容生成能力的模型及相关技术	向中国境内公众提供生成文本、图片、音频、视频等内容的服务

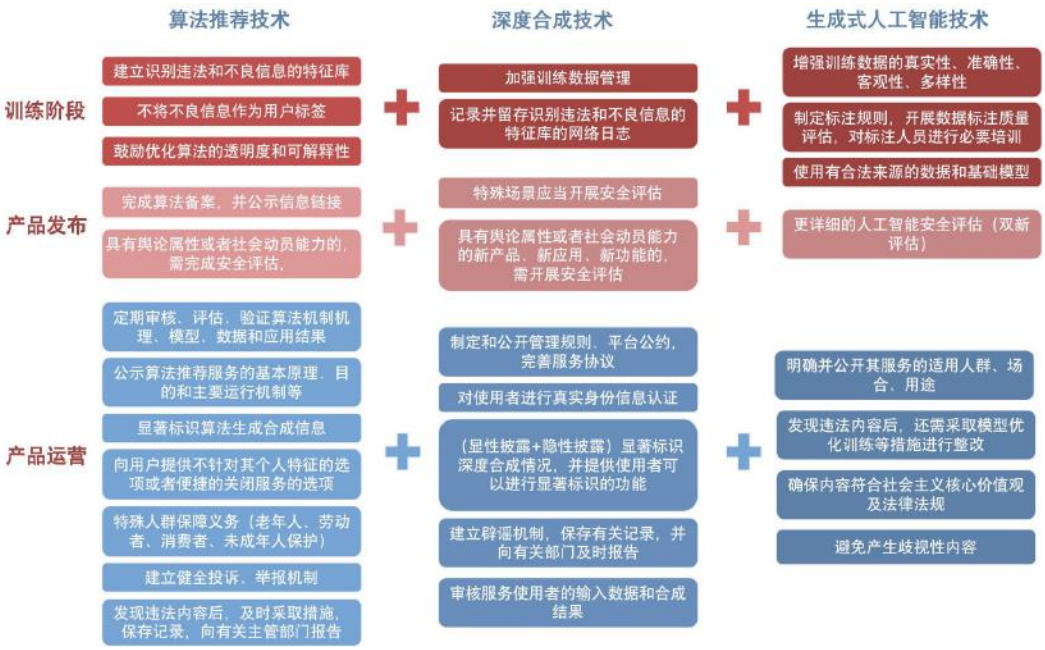
（二）监管部门

目前负责算法备案、评估与监管的机构主要是网信办（中央网信办网络管理技术局）。与此同时，电信主管部门、公安部门也会依据职责负责监督管理工作。

（三）合规义务要求

1.合规义务概览

如下图所示，面向中国境内的生成式人工智能服务需要同时满足《算法推荐规定》《深度合成规定》与《生成式办法》三部规章的要求。



2.安全评估义务

《生成式办法》要求提供具有舆论属性或者社会动员能力的生成式人工智能服务的，应当按照国家有关规定开展安全评估，并按照《互联网信息服务算法推荐管理规定》履行算法备案和变更、注销备案手续。

这是生成式人工智能服务上市前的必要流程，据悉目前主流的 GPT 产品如百度的文心一格、阿里的通义千问均处于安全评估阶段。

生成式人工智能的安全评估与普通的算法推荐和深度合成评估不同，评估负责单位是中央网信办网络管理技术局，评估的颗粒度更细，所需提供的材料也更多。目前具体的评估内容和流程均未公开。

（四）未来可能面临的监管挑战

1.人工智能法

2023 年 5 月 31 日，国务院办公厅发布了《国务院 2023 年度立法工作计划》，计划在 2023 年提请全国人大常委会审议人工智能法草案。目前尚无该法律具体

的信息，但学者猜测该法律可能会与欧盟的人工智能法案相似，具有更高的层级和更广的适用范围。

2.科技伦理审查

2021 年 12 月 24 日，第二次修订的《科学进步法》在第 103 条提出了要建立科技伦理审查机制。而在 2023 年 4 月 4 日，科技部官网发布公开征求对《科技伦理审查办法（试行）》，完善了科技伦理审查的流程和要求。

其中涉及数据和算法的科技活动，数据处理方案需要符合国家有关数据安全的规定，数据安全风险监测及应急处理方案得当；算法和系统研发需要符合公平、公正、透明、可靠、可控等原则。

根据办法草案，未来人工智能企业可能需要承担的合规义务有：

1）建立本单位的科技伦理（审查）委员会。

2）建立健全科技活动全流程科技伦理监管机制和审查质量控制、监督评价机制，加强对科技伦理高风险科技活动的动态跟踪、风险评估和伦理事件应急处置。

3）在开展科技活动前进行科技伦理风险评估或审查。科技活动负责人应向本单位科技伦理（审查）委员会申请伦理审查。

此外，根据科技部相关负责人在 2022 年《关于加强科技伦理治理的意见》新闻发布会的答复，国家科技伦理委员会正在研究制定科技伦理高风险科技活动清单，医学、生命科学和人工智能方面是重点领域。未来，开展科技伦理高风险科技活动将需要按规定进行登记。

四、欧盟

（一）监管框架

欧盟现行人工智能立法仍主要集中在传统人工智能而非生成式人工智能上，但也逐渐触及生成式人工智能的问题。在采用软法方法但不足以解决人工智能技术因其特定特征而带来的风险后，欧盟开始着手完善立法。以下是欧盟关于人工智能的主要立法、指南。

2019 年 4 月，欧盟委员会人工智能高级专家组发布了《可信赖的人工智能的道德指南》，为在欧盟设计、开发、部署、实施或使用人工智能产品和服务的利益相关者提供指引。指南道出了可信赖的人工智能的四项道德原则：尊重人类尊严、防止伤害、公平和可解释性，并提出了实现这四项原则的七项要求。

Figure 2: Interrelationship of the seven requirements: all are of equal importance, support each other, and should be implemented and evaluated throughout the AI system's lifecycle

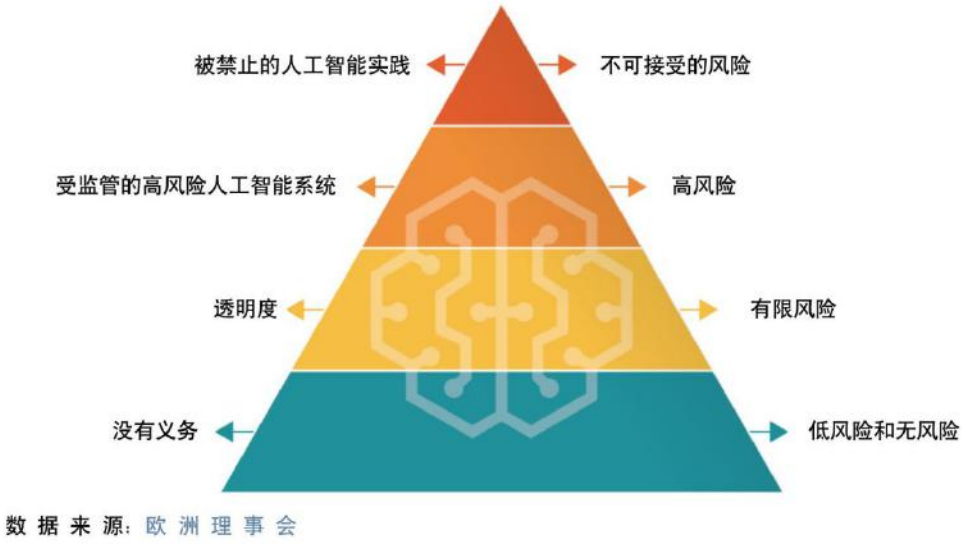
（《可信赖的人工智能的道德指南》提出的七项关键要求及其交互关系）

2020 年 2 月，欧盟委员会发布了《人工智能白皮书：欧洲实现卓越和信任的途径》，指出欧洲未来人工智能监管框架的关键是建立“卓越生态系统”和“可

信赖的生态系统”，保护基本权利、消费者权利等。

2022 年 9 月，欧盟委员会发布了《将非契约性民事责任规则适用于人工智能的指令的提案》，倡议推定人工智能造成的损害与受害者遭受的损失之间存在因果关系，赋予国家法院命令披露涉嫌造成损害的高风险人工智能系统证据的权力，以期确保受人工智能系统伤害的人与受其他技术伤害的人享有相同的保护水平。

2021 年 4 月，欧盟委员会提出了《人工智能法》草案，为欧盟人工智能产品和服务的开发、供应和使用制定了统一的法律框架。该提案建议在欧盟法律中引入技术中立的人工智能的定义，并根据“基于风险的方法”为不同类型的人工智能系统制定要求和义务。



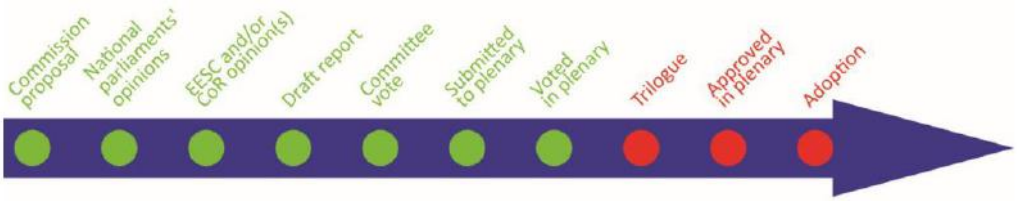
（《人工智能法》的分级监管示意图）

2022 年 12 月，欧盟理事会通过了其共同立场，建议缩小人工智能系统的定义；将禁止使用人工智能进行社会评分的禁令的适用范围扩展到私人行为者，并增加在公共场所允许使用“实时”远程生物识别系统这一例外情形；通过实施法律，对通用人工智能系统施加要求；增加新规以考虑人工智能系统可用于许多不同目

的（通用人工智能）的情况；简化《人工智能法》的合规框架，特别是加强人工智能委员会的作用。

2023 年 6 月 14 日，欧盟议会投票通过了最新的《人工智能法》草案。欧盟委员会、欧洲议会和欧盟理事会将进入“三方会谈”阶段。最新的草案再次作出了多项实质性修改，包括：

- 修改人工智能系统的定义，使其与亚太经合组织商定的定义保持一致
- 扩大被禁止的人工智能系统清单
- 限缩高风险人工智能系统的范围为“存在可能损害人类健康、安全、基本权利或环境的‘重大风险’”的人工智能系统，在欧盟部署该类系统的主体有义务进行基本权利影响评估
- 对通用人工智能[17] 和 ChatGPT 等生成式人工智能[18] 模型的提供者施加义务
- 赋予各国当局要求访问人工智能系统的训练和培训模型（包括基础模型）的权力。此外，成员国有权寻求加强公民权利以对人工智能系统提出投诉，并有权获得对基于高风险人工智能系统的决定的解释
- 为支持创新，研究活动和免费、开源人工智能组件的开发在很大程度上不受《人工智能法》的约束



（《人工智能法》最新立法进度图）

[17] 第 3 条（1）（1d）：可被用于和适应各种非专门和特定设计的应用的 AI 系统。
[18] 专门用于生成具有不同自治程度的复杂文本、图像、音频或视频等内容的人工智能

（二）一般合规义务

1.分层监管：系统要求

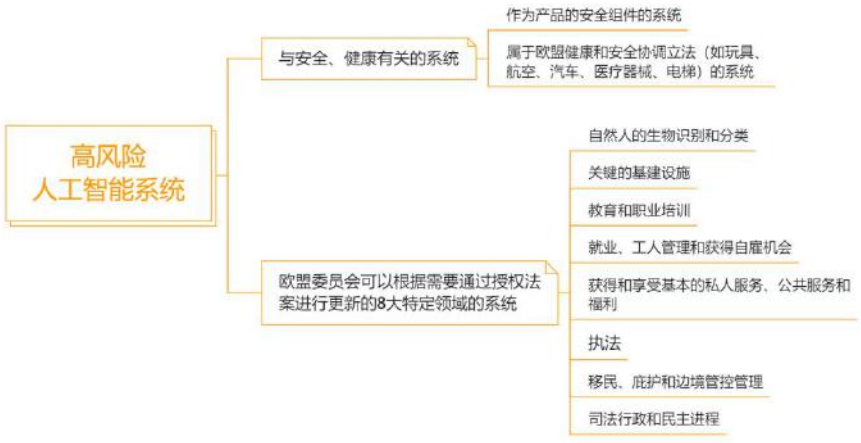
（1）不可接受的风险人工智能系统：全面禁止

《人工智能法》禁止以下类型的人工智能系统进入欧洲市场。

- 采用有害的操纵“潜意识技术（subliminal techniques）”的人工智能系统
- 利用特定弱势群体（身体或精神残疾）的人工智能系统
- 根据敏感或受保护的特征或基于对该特征的推断对自然人进行分类的人工智能系统
- 公共当局或代表其使用的用于社会评分目的的人工智能系统
- 在公共场所用于执法目的的“实时”远程生物识别系统，但在少数情况下除外
- 预测刑事或行政违法行为的风险评估的人工智能系统
- 无目的地从互联网或闭路电视录像中抓取面部图像、创建或扩展面部识别数据库的人工智能系统
- 在执法、边境管理、工作场所和教育机构等领域推断自然人的情绪的人工智能系统
- 通过“后”远程生物识别系统分析公共可访问空间的记录片段的人工智能系统

（2）高风险人工智能系统：重点监管

《人工智能法》将高风险人工智能系统分为以下两大类。



该法要求高风险人工智能系统完成事前合格评估以及包括风险管理、测试、技术稳健性、数据培训和治理、透明度、人为监督和网络安全在内的其他要求。

事前合格评估	其他
• 适用欧盟现有产品安全立法的人工智能产品和服务（例如医疗设备） ✓ 通过第三方符合性评估	• 风险管理 ✓ 建立、实施、记录和维持相关风险管理系统，并定期更新 • 测试 • 技术稳健性 • 数据培训和数据治理 ✓ 系统应具有运行时自动记录事件的能力，该能力应符合最新的技术和公认的标准或通用规范 • 透明度 • 人为监督 • 网络安全
• 不适用欧盟立法 ✓ 自行开展符合性评估	
• 用于生物识别 ✓ 由“公告机构（notified body）”进行符合性评估	

高风险人工智能系统的透明度要求
• 设计和开发应确保其运行足够透明，使提供商和用户能够合理地了解系统的功能 □ 用户应能够适当地理解和使用系统，允许用户根据第68(c)条向受影响的人解释系统做出的决定
• 以适当方式（持久的媒介亦可）提供附可理解的使用说明，帮助操作和维护系统、支持用户的知情决策 □ 其中包括简明、正确、清晰、尽可能完整的、对用户而言合理相关、可访问且可理解的信息 ➢ 提供商及其授权代表（如涉及）的身份和联系方式 ➢ 系统的特点、能力和性能限制 ➢ 在事前合格评估时已由提供者预先确定（如有）的系统的变化、性能 ➢ 人为监督措施，包括为方便用户解释系统输出而实施的技术措施 ➢ 任何为确保系统在预期使用寿命内正常运行的必要的维护和保养措施，包括软件更新

（3）有限风险人工智能系统[19]：有限的透明度义务

有限风险的人工智能系统只需遵守最低限度的透明度义务，使用户能够在知情的情况下作出决定。在与应用程序交互之后，用户可以决定是否继续使用它。用户应该知道其正在与人工智能——包括生成或操作图像、音频或视频内容的人工智能系统（生成式人工智能）互动。[20]

（4）低风险或无风险人工智能系统：无额外法律义务

[19] 如与人类交互的系统（即聊天机器人）、情感识别系统、生物识别分类系统，以及生成或操纵图像、音频或视频内容的人工智能系统（即深度伪造）

[20] EU AI Act: first regulation on artificial intelligence | News | European Parliament (europa.eu)

虽然拟议的《人工智能法》未对低风险或无风险人工智能系统施加限制，但其设想制定行为准则，鼓励非高风险人工智能系统的提供者自愿适用高风险人工智能系统的强制性要求。

2.主体义务

（1）高风险人工智能系统相关主体义务

1) 提供者、部署者及其他各方

提供者的义务	部署者的义务
• 有理由认为投放市场或投入使用的系统不合规 ✓ 采取必要纠正行动使系统合规 ✓ 酌情撤销、禁用或召回该系统 ✓ 立即通知：分销商、进口商、提供该系统或将其投入使用的成员国主管当局（如涉及，还应通知部署者、公告机构，与部署者合作调查原因）	/
• 意识到系统存在对健康或安全或保护个人基本权利的风险 ✓ 立即通知提供或将其投入使用的成员国监管机构、公告机构，特别是不符合的性质和所采取的任何相关纠正措施 ✓ 立即通知分销商、进口商（如涉及，还应通知部署者、授权代表）	
• 建立质量管理体系并形成文件	
• 向授权代表提供必要的权力和资源	
• 与国家主管部门合作 ✓ 允许相关机构访问其可控制的系统自动生成的日志	

其中，基础模型的提供者、生成式人工智能系统中的基础模型的提供者或将基础模型专门用于生成式人工智能系统的提供者有义务在基础模型投放市场或投入使用后 10 年内，保留其技术文件并交国家主管机关使用。

在以下 3 种情形发生时，其他当事方将被视为提供者，承担使高风险人工智能系统合规的义务及上述《人工智能法》对提供者、部署者及其他各方施加的义务。同时，原供应商应向新供应商提供人工智能系统的技术文档和所有其他相关且合理可预期的信息能力、技术访问或基于公认的最先进技术的其他协助。

"视为提供者"
• 以其名义生产并与受欧盟立法约束的特定产品一起投入使用的人工智能系统的生产者
• 下列情形之一的经销商、进口商、用户或者其他第三方 ✓ 将其姓名/名称或商标置于已投放市场或投入使用的高风险AI系统 ✓ 对已投放市场或投入使用的高风险AI系统进行实质性修改，但该系统仍为高风险AI系统 ✓ 对未被划分为高风险的系统，包括已投放市场或投入使用的通用人工智能系统进行实质性修改，使该系统成为高风险AI系统
• 基础模型直接集成到高风险AI系统的基础模型提供者

2) 进口商

进口商的义务
• 有理由认为系统不合规、系伪造或附有伪造文件时 ✓ 不将系统投入市场
• 系统存在对健康或安全或保护个人基本权利的风险时 ✓ 将该风险通知提供者、市场监督机构
• 在系统及其包装或随附文件上注明名称、注册商号或注册商标、可联系的地址
• 当系统由其负责时，确保储存或运输条件不会影响系统的合规性
• 配合国家主管部门的任何行动，降低系统风险
• 与国家主管部门合作 ✓ 根据国家监管机构的合理要求，以易于理解的语言向其提供所有必要的信息和文件 ✓ 允许相关机构访问其可控制的系统自动生成的日志

3) 分销商

分销商的义务
• 有理由认为系统不合规 ✓ 不将其投入市场
• 根据其掌握的信息有理由认为其提供的系统不合规 ✓ 采取必要纠正措施，撤回或召回系统，或 ✓ 确保供应商、进口商或任何相关运营商酌情采取纠正措施
• 系统存在对健康或安全或保护个人基本权利的风险 ✓ 通知供应商或进口商（如涉及，通知相关的国家主管部门并提供详细信息）
• 当系统由其负责时，确保储存或运输条件不会影响系统的合规性
• 配合国家主管部门的任何行动，降低系统风险
• 与国家主管部门合作 ✓ 根据国家主管部门的合理要求，向该部门提供其拥有或可获得的所有信息和文件

4) 部署者

部署者的义务	
• 确保根据使用说明使用系统	• 保留其控制范围内的由该系统自动生成的日志
• 对系统实施人工监督，确保负责监督的自然人能力及资格适当、经过培训且拥有必要资源	• 配合国家主管部门与系统有关的任何行动，降低系统风险
• 如可控制数据输入 ✓ 确保输入与系统的预期目的相关且具有足够的代表性	• 部分高风险AI系统的部署者（协助）作出与自然人有关的决策 ✓ 告知自然人高风险AI系统的使用及其获得该法第68条c款所述解释的权利
• 根据使用说明监控系统的运行，并在相关情况下依法告知提供者	• 进行数据保护影响评估并发布其摘要
• 有理由认为按使用说明使用可能导致系统出现对健康或安全或保护个人基本权利的风险 ✓ 暂停系统的使用 ✓ 先通知提供者，再通知进口商或分销商和有关的国家监管机构	

（2）与透明度相关的主体义务

与透明度有关的主体义务	
• 供应商 □ 确保旨在与自然人交互的人工智能系统的设计和开发方式能使自然人及时、清晰和可理解地知晓其正与AI系统交互，除非从环境和使用背景中可以明显看出这一点 □ 如相关，应告知人工智能启用的功能、决策负责者、就损害寻求司法救济的现有权利和程序、寻求解释的权利 □ 不适用于：法律授权用于侦查、预防、调查和起诉刑事犯罪的人工智能系统，除非这些系统可供公众举报刑事犯罪	• 人工智能系统未经用户同意生成或操纵看似真实或真实的文本、音频或视觉内容（深度伪造） □ 以适当、及时、清晰和可见的方式披露该内容是人为生成或操纵的，并且尽可能披露产生或者操纵该信息的自然人的姓名或者法人的名称 □ 不适用于：法律授权使用生成或操纵文本、音频或视觉内容的人工智能系统、行使《欧盟基本权利宪章》所保障的言论自由权和艺术和科学自由权所必需的人工智能系统 □ 如内容构成明显具有创造性、讽喻性、艺术性或虚构的电影、视频游戏视觉效果和类似作品或节目的一部分，则此义务仅限于以适当、清晰和可见的方式披露此类生成或操纵内容的存在
• 未被禁止的情绪识别系统或生物识别分类系统的用户 □ 以及时、清晰和以可理解的方式告知暴露在该系统下的自然人系统的运行情况 □ 在处理个人生物识别和其他数据前征得其同意 □ 不适用于：法律允许用于检测、预防和调查刑事犯罪的生物识别分类人工智能系统	

（三）特殊合规义务

除以上一般合规义务外，特定主体还存在与知识产权、进出口相关的特殊合规义务。

供应商、部署者、分销商、进口商、通用及生成式人工智能系统相关基础模型的提供者在将高风险人工智能系统投放市场或投入使用前存在下图所示的义务。

1.提供者、部署者

提供者、部署者的义务	
• 确保系统合规	• 有符合要求的质量管理体系
• 确保遵守合格评估要求	• 草拟、保存系统的技术文件
• 在系统或随附文件上注明名称、注册商号或注册商标、地址和联系信息	• 保留其控制下的系统自动生成的日志
• 确保进行监督的自然人明确意识到自动化或确认偏差的风险	• 进行数据保护影响评估并发布其摘要
• 提供输入数据或其他相关信息的说明	• 起草欧盟符合性声明
• 对系统加贴CE标志	• 在欧盟数据库中注册系统
• 确保系统符合可访问性要求	• 根据国家监管机构的合理要求，提供所有必要的信息和文件证明系统合规
• 在欧盟以外建立的供应商：应书面指定在欧盟建立的授权代表	

2.进口商

进口商的义务	
• 验证系统具有CE标志、附有所需文件和使用说明	• 确保系统提供者、进口商（如有）已遵守该法规定的义务

3.分销商

分销商的义务
• 确保供应商已执行了适当的合格评估程序
• 确保供应商已起草了技术文件
• 确保系统带有CE标志、附有所需文件和使用说明
• 确保在欧盟以外建立的供应商已书面指定了授权代表

4.通用及生成式人工智能基础模型的提供者

基础模型的提供者	生成式人工智能系统中使用的基础模型的提供者 将基础模型专门用于生成式人工智能系统的提供者
• 识别、减少和缓解对健康、安全、基本权利、环境以及民主和法治的合理可预见风险 • 仅处理并纳入受适当基础模型数据治理措施约束的数据集 • 设计、开发基础模型以实现适当性能 • 利用公布后的协调标准设计、开发基础模型，减少资源浪费 • 基础模型的设计应具有测量、记录资源消耗的能力，技术上可行的情况下记录系统的部署、使用在其生命周期中可能产生的其他环境影响 • 起草广泛的技术文件和可理解的使用说明，使下游提供者能够遵守其义务 • 建立质量管理体系，以确保并记录其符合义务，并有可能进行实验以满足这一要求 • 在欧盟数据库中注册该基础模型	
/	• 训练、设计模型以防止生成非法内容 • 记录、公开提供受版权法保护的训练数据的详细使用总结 • 设计和开发旨在与自然人交互的人工智能系统，应使该系统、供应商或用户及时、清晰和可理解地向与系统交互的自然人披露交互的事实（除非从环境和使用背景中可以明显看出）、人工智能启用的功能、决策负责者、就损害寻求司法救济的现有权利和程序、寻求解释的权利 □ 不适用于：法律授权用于侦查、预防、调查和起诉刑事犯罪的人工智能系统，除非这些系统可供公众举报刑事犯罪

5.部署者

部署者的义务	
• 在工作场所将高风险AI系统投入服务或使用前 ✓ 咨询工人代表，以便根据指令2002/14/EC达成协议，并通知受影响的员工将受到的约束	• 首次使用八大特定领域的高风险AI系统：基本权利影响评估 ✓ 如无法确定缓解评估发现的风险的详细计划；避免将系统投入使用；及时通知提供商、国家监管机构 ✓ 除中小企业外的部署者：通知国家监管机构、相关利益相关者；尽可能让可能受系统影响的个人或群体的代表参与评估 ✓ 市场监管部门可基于特殊原因豁免此义务 ✓ 公共当局或联盟机构、机构、办公室或机构或代表其行事的部署者、被指定为守门人的企业的部署者：公布影响评估结果摘要
• 如部署者已被要求根据法规(EU)2016/679第35条或指令(EU)2016/680第27条进行数据保护影响评估 ✓ 基本权利影响评估应与数据保护影响评估一并进行（数据保护影响评估应作为附录发布）	

不遵守《人工智能法》的制裁：根据侵权行为的性质、严重程度、持续时间、

后果、人工智能系统的目的、受影响人员的数量及受影响程度等因素，相关当事人可能被处以不同规模的行政罚款。

	个人	公司
将禁止的人工智能系统投放市场或投入使用	最高可被处 4000 万欧元罚款	罚款 4000 万欧元或上一财政年度全球年营业额的 7%（以较高者为准）
不遵守高风险人工智能系统关于数据治理、透明度和提供信息给用户规定	最高可被处 2000 万欧元罚款	罚款 2000 万欧元或上一财政年度全球年营业额的 4%（以较高者为准）
违反其他对于人工智能系统或模型的要求和义务	最高可被处 1000 万欧元罚款	罚款 1000 万欧元或上一财政年度全球年营业额的 2%（以较高者为准）
在答复请求时向公告机构和国家主管当局提供不正确、不完整或误导性的信息	最高可被处 500 万欧元罚款	罚款 500 万欧元或上一财政年度全球年营业额的 1%（以较高者为准）

《人工智能法》还规定，成员国需制定罚则并采取一切必要措施，确保罚则得到适当、有效的执行。

（四）监管动态

近年来，已有多个包括意大利、法国、西班牙等多个欧盟成员国对 ChatGPT 展开调查。

时间	国家	监管机构	性质	措施	事由
2023.3.31	意大利	Garante	主动调查	限时中止令	合法性、透明度、未成年人数据处理、数据准确性
2023.4.11	法国	CNIL	投诉驱动	被动调查	合法性基础、数据

					权利行使
2023.4.14	西班牙	AEPD	主动调查	主动调查	透明度
2023.4.24	德国	BADI	主动调查	编制问卷	数据保护影响评估
2023.4.20	挪威	Datatilsynet	未开启调查	官方表态	/
2023.4.21	爱尔兰	DPC	未开启调查	官方表态	/

（欧盟成员国对 ChatGPT 的调查情况表[21]）

2023 年 3 月，欧洲消费者组织呼吁欧盟和各国安全、数据和消费者保护部门立即对 ChatGPT 和类似的聊天机器人展开调查。欧洲消费者组织副总干事 Ursula Pachl 说“像 ChatGPT 这样的生成式人工智能为消费者提供了各种可能性，但是人们对这些系统如何欺骗、操纵和伤害人们表示严重关切。它们也可以被用来传播虚假信息，延续放大歧视的现有偏见，或被用于欺诈。”[22]

2023 年 4 月 13 日，数据保护委员会决定成立一个专责小组（dedicated task force），以促进合作，就数据保护当局可能采取的执法行动交流信息。[23] 最近，该小组正在调查 ChatGPT。

（五）相关机构

欧盟目前没有专门针对人工智能监管的独立机构，主要由欧盟层面及成员国层面不同领域的主管部门联合监管和执法。

最新的《人工智能法》草案要求成员国指定一个或多个主管机构，包括一个国家监管机构，负责监督《人工智能法》的应用和实施；并提议在欧盟层面建立由成员国和欧盟委员会的代表组成的新欧盟机构——“人工智能办公室”，以期为

[21] 此表引自微信公众号“科技利维坦”《ChatGPT 在欧洲“被禁”的法律逻辑》一文，ChatGPT 在欧洲“被禁”的法律逻辑 (qq.com)
[22] Investigation by EU authorities needed into ChatGPT technology (beuc.eu)
[23] EDPB resolves dispute on transfers by Meta and creates task force on Chat GPT | European Data Protection Board (europa.eu)

《人工智能法》的统一适用提供指导、协调联合跨境调查。

在国家一级和跨境一级执行数据保护法的欧洲数据保护委员会（European Data Protection Board）将从数据保护的角度对人工智能的数据处理行为进行监管，使《通用数据保护条例》得到遵守。

国家市场监督机构负责进行市场监督、有责任在发现危险产品时采取适当措施。根据《人工智能法》，该机构将负责评估提供者、部署者、授权代表、进口商和分销商对高风险人工智能系统的义务和要求的遵守情况。不合规的人工智能系统或尽管合规但对人员健康或安全、基本权利或其他公共利益保护构成风险的人工智能系统可能被采取禁用、限用、撤回或召回等纠正措施。由《市场监督条例(EU)2019/1020》建立的欧盟产品合规网络（EU Product Compliance Network, EUPCN）是负责在非食品领域执行欧盟产品立法的国家当局之间的主要合作论坛，确保各产品部门市场监督当局间的合作和信息交流，确定欧盟一级市场监督的优先事项，并发起协调或联合举措。

根据查明的风险，海关可以在边境拒收产品、实施销售禁令或发布警告信息。如果产品已经在消费者手中，可以要求召回。所有这些措施都在“安全门”（safe gate）上报告，从而迅速向系统内的所有国家当局传播信息，方便其追踪危险产品、在本国采取措施，确保欧洲单一市场的安全。

国家消费者保护当局负责执行消费者保护法。欧盟委员会与欧洲数据保护委员会(EDPB)秘书处定期举办联合研讨会，讨论消费者和数据保护执法人员共同关心的案件，并进一步加强他们的交流与合作。2020 年 7 月，消费者和数据保护主管部门志愿者小组成立，以在国家和欧盟层面促进交流，建立最佳做法和执法经验。

五、美国

（一）监管框架

美国目前无论在联邦还是州一级均尚未专门针对通用人工智能或生成式人工智能进行正式立法。部分州通过了与人工智能有关的立法，主要涉及视频面试、与就业有关的自动决策、保护消费者权利、防止保险公司人工智能的算法歧视等问题。科罗拉多州、伊利诺伊州、佛蒙特州、阿拉巴马州等州成立了专门小组或委员会对人工智能进行研究。

以下是美国与人工智能高度相关的主要政策文件、治理框架、指南等。

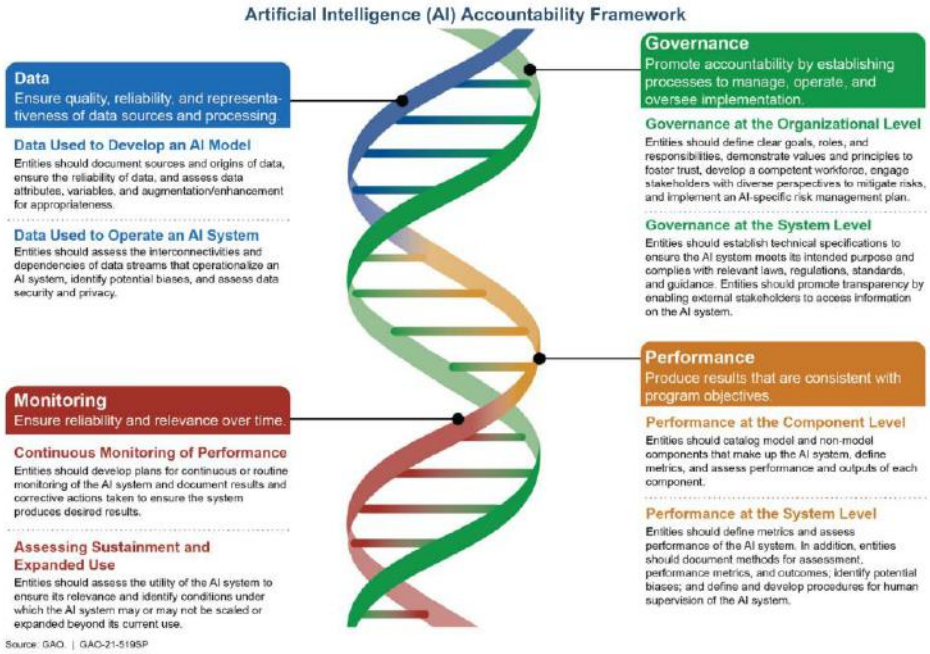
2019 年 2 月，美国总统行政办公室发布《保持美国在人工智能领域的领导地位》，主张在联邦政府、工业界和学术界推动人工智能的技术突破；减少人工智能技术安全测试和部署的障碍以便人工智能产业创新；培养公众对人工智能技术的信任和信心，以维持美国在人工智能领域的领导地位。

2020 年 11 月，美国总统行政办公室发布《行政部门和机构负责人备忘录——人工智能应用监管指南》。该指南用于规范在联邦政府以外部署的人工智能系统，旨在为所有联邦政府机构提供指导。指南要求，联邦机构必须避免不必要地阻碍人工智能创新和增长的监管或非监管行动。在法律允许的情况下，联邦机构在决定是否以及如何监管人工智能可能应用的领域时，应评估潜在监管对人工智能创新和发展的影响。

2020 年 12 月，美国总统行政办公室发布政策文件《促进可信赖的人工智能在联邦政府中的使用》，主张联邦机构应以培养公众信任和信心的方式设计、开发、获取和使用人工智能，同时保护隐私、公民权利、公民自由和美国价值观。

2021 年 6 月，美国政府问责局在《人工智能：联邦机构和其他实体的问责

框架》中围绕治理、数据、绩效和监督等原则确定了问责实践，以帮助联邦机构和其他机构负责任地使用人工智能。



（《人工智能：联邦机构和其他实体的问责框架》示意图）

2022 年 10 月，美国白宫科技政策办公室发布白皮书《人工智能权利法案的蓝图——让自动化系统为美国人民服务》，列出了人工智能的五项基本原则，对应美国人民都应享有的五项核心保护。

（五项核心保护）

2023 年 1 月，美国商务部下属机构美国国家标准与技术研究院发布《人工智能风险管理框架》及配套使用手册。框架对人工智能的风险进行了界定，通过

四大模块，帮助设计、开发、部署或使用人工智能系统的人更好地管理与人工智能有关的个人、组织和社会的风险。该框架的目标是提高人工智能系统的可信度——包括（人身、财产）安全、（系统）安全和弹性、（系统运行机制）可理解和（系统输出结果）可解释、隐私（保护）增强、有效和可信赖、隐私增强、透明及负责七大维度。[24]

（《人工智能风险管理框架》四大功能示意图）

2023 年 3 月，美国版权局发布《包含人工智能生成材料的作品的版权登记指南》，介绍了版权局如何将《版权法》的“人类作者要求”（Human Authorship Requirement）适用于此类作品的注册申请，并为申请人提供指导。版权局明确《宪法》和《版权法》的“作者”仅指人类，只有存在足够的“人类作者”要素的作品才能申请注册并获得版权保护，人工智能生成的内容应明确在申请外排除。版

[24] 引自微信公众号“清华大学智能法治研究院”《人工智能治理 | 美国 NIST《人工智能风险管理框架评述》一文，
https://mp.weixin.qq.com/s?src=11&timestamp=1691565786&ver=4701&signature=1nEYbt9wsPpmhIeQG*gU8PCuHv*O3vCZidb0V56ghyye-oJfFrWzyFXq1-u76ud4faE2bEdx0uFGcXIKxBmhEiaqjrTBTE9FjjbaEai5QaaLdikB6itbUFK6R4uoXvC8&new=1

权局将逐案审查人类在多大程度上创造性地控制了作品的表达并实际构成了“作者”要素。

2023 年 5 月，美国国会研究服务部发布《生成式人工智能与数据隐私：初探》的报告，对生成式人工智能的定义、数据使用、数据来源等情况进行介绍，建议国会完善隐私立法、规范数据收集、加大对替代技术方法的研究与开发。美国国会研究服务部建议立法设置如下机制：

- 通知和披露机制：要求开发者在收集或使用个人数据之前获得数据主体的同意并告知主体数据将用于何种目的
- 选择退出的机制：对于尚未公开的数据，生成式人工智能的开发者应向用户提供退出数据收集的选项
- 删除和最小收集机制：用户有权从当前数据集中删除自己的数据，或规定个人数据的最长保留期限

（二）一般合规义务

虽然美国尚未针对通用人工智能及生成式人工智能系统及相关主体设定合规义务，但相关人工智能企业应考虑主管部门在不同领域出台的指南、框架等文件，减少后续的调查和执法风险。

联邦贸易委员会消费者保护局局长在“人工智能和算法的商业指导”中提示企业在消费品中应用人工智能时应注意以下事项。[25]

[25] Using Artificial Intelligence and Algorithms | Federal Trade Commission (ftc.gov)

企业在消费品中使用人工智能和算法的注意事项	
• 透明 ✓ 收集敏感数据时保持透明 ✓ 利用消费者信息在信贷、就业、保险、住房或类似福利和交易中作出自动决策：需向消费者提供“不利行动”通知（“adverse action” notice），告知消费者其有权查看与其有关的信息并纠正不准确的信息	• 向消费者解释决定 ✓ 根据算法决策作出拒绝给予消费者有价值事物时，解释该决定 ✓ 使用算法向消费者分配风险评分，应披露影响评分的关键因素、重要性排序 ✓ 可能改变基于自动化工具的交易条款时，告知消费者
• 确保决定公平 ✓ 不歧视受保护阶层 ✓ 关注消费者保护的结果 ✓ 为消费者提供访问和纠正用于决策的信息的机会	• 确保数据和模型的稳健性及经验可靠性 ✓ 确保数据的准确性、最新性 ✓ 确保人工智能模型经过验证、再验证，以确保其按预期工作且不会非法歧视
• 对合法、合道德、公平和不歧视负责 ✓ 考虑数据集的代表性、数据模型的偏差、预测准确性、使用大数据的道德和公平风险 ✓ 防止算法被未经授权的人使用 ✓ 考虑问责制	

2023 年 7 月 21 日，Amazon, Anthropic, Google, Inflection, Meta, Microsoft 和 OpenAI 向美国政府自愿作出的一系列承诺亦可供企业参考。[26]

七大AI巨头向美国政府作出的承诺
• 确保产品安全后才向公众推广 ✓ 在其人工智能系统发布前对系统进行内部和外部安全测试（部分测试由独立专家进行） ✓ 在整个行业以及与政府、民间社会和学术界分享有关人工智能风险管理的信息，包括安全方面的最佳实践、试图规避安全措施的信息以及技术合作
• 建立把安全置于首位的系统 ✓ 投资网络安全和内部威胁防范措施，以保护自有的和未发布的模型权重 ✓ 只有在充分考虑了安全风险后，才能发布模型权重 ✓ 促进第三方发现和报告其人工智能系统中的漏洞
• 赢得公众的信任 ✓ 开发强有力的技术机制（如水印系统），确保用户知道内容何时是人工智能生成的 ✓ 公开报告其人工智能系统的能力、局限性、适于及不适于使用的领域、系统的安全风险和社会风险（如对公平和偏见的影响） ✓ 优先研究人工智能系统可能带来的社会风险（包括避免有害的偏见、歧视及保护隐私） ✓ 开发和部署先进的人工智能系统

[26] FACT SHEET: Biden-Harris Administration Secures Voluntary Commitments from Leading Artificial Intelligence Companies to Manage the Risks Posed by AI | The White House

（三）特殊合规义务

除以上一般合规义务外，特定主体还存在与知识产权、进出口相关的特殊合规义务。

根据美国版权局《包含人工智能生成材料的作品的版权登记指南》，为包含人工智能生成材料的作品申请版权保护的申请人在不同的申请阶段存在不同义务：

为包含人工智能生成的材料的作品申请版权保护的申请人的义务

- **提交新的作品登记申请**
  - ✓ 披露作品中包含的人工智能生成的内容
  - ✓ 简要说明人类作者对作品的贡献
  - ✓ 简要描述人工智能生成的内容，在申请表中将其明确排除
  - ✓ 不应仅仅因为在创作作品时使用了人工智能技术而将该技术或提供该技术的公司列为作者或合著者
- **已提交申请，申请中包含人工智能生成的内容**
  - ✓ 检查其向版权局提供的资料是否充分披露了人工智能生成的内容，否，则应采取措施更正信息

  办公室将在符合条件的情况下颁发新的补充注册证书，其中包含针对人工智能生成材料的免责声明
- **已提交但待决的申请**
  - ✓ 联系版权局的公共信息办公室（Public Information Office），报告其申请遗漏了作品包含人工智能生成材料的事实
- **已处理并已注册的申请**
  - ✓ 提交补充注册
  - ✓ 更新公开记录

在进出口方面，据路透社 2022 年 9 月 1 日的报道，芯片设计公司英伟达表示，美国官员要求其停止向中国出口两种用于人工智能工作的顶级计算芯片。[27]

2022 年 10 月 7 日，美国国家工业和安全局发布了《中国超级计算机和半导体产品出口管制公告》[28]，宣布对华人工智能和半导体技术出口管制政策。根据美国《出口管制改革法（2018 年）》的规定，在 2022 年 10 月 21 日至 2023 年 4 月 7 日间出口、再出口或从国内转让超级计算机给中国企业前，出口商、再出口

---

27 https://www.reuters.com/technology/nvidia-says-us-has-imposed-new-license-requirement-future-exports-china-2022-08-31/

28 Public Information on Export Controls Imposed on Advanced Computing And Semiconductor Manufacturing Items To The People'S Republic Of China (Prc) (Doc.Gov)

商或转让方须取得国家工业和安全局的临时通用许可证。此外，出口商、再出口商或转让方须保留接收方的名称、超级计算机在中国的目标地址及接收方总部的所在地。

据《华尔街日报》2023 年 7 月 4 日的报道，美国政府计划限制中国企业使用美国的云计算服务。[29]如果新规被采纳，亚马逊、微软等美国云服务提供商向中国公司提供使用先进人工智能芯片的云计算服务前将需先征得美国政府的许可。商务部预计在未来几周内实施这一限制。

（四）监管动态

2023 年 2 月 16 日，美国司法部和商务部工业与安全局成立了颠覆性技术打击队（Disruptive Technology Strike Force），用以防止、起诉逃避美国对人工智能、三维打印、半导体、先进生物科学等关键技术出口管制的行为。[30]

2023 年 5 月 13 日，美国总统科学技术顾问委员会成立了生成式人工智能工作组，帮助评估关键机遇和风险，并就如何尽可能公平、负责和安全地确保相关技术的开发和部署提供意见。[31]

2023 年 6 月 22 日，商务部部长宣布美国国家标准与技术研究所将成立生成式人工智能公共工作组，帮助研究所制定关键指南，进而帮助企业应对与生成式人工智能技术相关的特殊风险。[32]

2023 年 7 月 13 日，美国联邦贸易委员会启动对 OpenAI 是否违反消费者保护法的调查，要求 OpenAI 提供有关其处理个人数据、向用户提供不准确信息的

[29]https://www.wsj.com/articles/u-s-looks-to-restrict-chinas-access-to-cloud-computing-to-protect-advanced-technology-f771613
[30] Office of Public Affairs | Justice and Commerce Departments Announce Creation of Disruptive Technology Strike Force | United States Department of Justice
[31] PCAST Working Group on Generative AI Invites Public Input | PCAST | The White House
[32] Biden-Harris Administration Announces New NIST Public Working Group on AI | NIST

可能性以及“对消费者造成损害（包括声誉损害）的风险”的记录。[33]

（五）相关机构

美国目前没有专门针对人工智能进行监管的独立机构，各领域主管部门将继续执行现行美国法律及根据各领域特殊情况制定的政策、指南。

美国商务部工业和安全局执行《出口管理条例》适用范围内人工智能相关商品、软件和技术的出口和再出口的法律、法规和政策。

美国联邦贸易委员会通过执行《联邦贸易委员会法案》和许多其他法律法规，有权起诉人工智能侵犯数据隐私、数据安全、不公平竞争等行为。

美国消费者产品安全委员会依据《消费者产品安全法》等法律有权监管消费品。

美国消费者金融保护局负责制定、执行联邦消费者金融法律，保护金融市场中的消费者免受不公平、欺诈或滥用行为的侵害及歧视。该局曾发布通告，确认无论使用何种技术（包括人工智能），联邦消费者金融法律和不利行动要求（adverse action）都将适用。

[33] https://www.washingtonpost.com/technology/2023/07/13/ftc-openai-chatgpt-sam-altman-lina-khan/

六、加拿大

（一）监管框架

加拿大拥有在不同领域规范人工智能的法律框架，包括《个人信息保护和电子文件法》《消费者隐私保护法》《加拿大消费品安全法》《食品和药品法》《银行法》《刑法》《机动车辆安全法》《加拿大人权法》和各省人权法等。与人工智能高度相关的主要法律、政策文件如下。

2017 年，加拿大发布了世界上第一个国家人工智能战略《泛加拿大人工智能战略》，用于推进人工智能领域的研究和创新，助力加拿大成为人工智能领域的先行者和全球领导者之一。

2020 年 11 月 12 日，加拿大隐私专员办公室发布《人工智能监管框架：<个人信息保护与电子文件法>改革建议》。办公室希望在获得人工智能的好处的同时维护个人的基本隐私权。因此，办公室建议：开发人工智能系统的人员应确保在设计系统时保护隐私；个人有权获得帮助其理解人工智能系统对其作出的决定的解释，该解释应基于准确的信息，且不具有歧视性或偏见；个人应有权对自动产生的决策提出质疑；隐私和数据保护监管机构有权要求提供上述证据、发布具有约束力的命令及对无视隐私的企业实施经济处罚。

2022 年 6 月 16 日，加拿大创新、科学与工业部部长提出了 C-27 法案即《数字宪章实施法》，其中包括《人工智能和数据法》。《人工智能和数据法》是加拿大正式规范非联邦政府机构的人工智能系统的首次尝试，加拿大也因此成为欧盟以外的第二个正式将此类立法提交审议的国家。该法的关键定义和概念与欧盟《人工智能法》、亚太经合组织人工智能原则、美国《人工智能风险管理框架》等存在共通之处。该法的主要内容如下：

- 采用类似欧盟《人工智能法》“基于风险的方法”，重点规范在国际和省际贸易和商业活动中使用的高影响（high-impact）人工智能系统的相关风险，规定了相关主体的义务，以确保在可能引入风险的每个点上都实行问责制
- 创新、科学与工业部部长被授权管理和执行该法，其有权下令制作与人工智能系统有关的记录，并公布违规情况的有关信息
- 设立由新的人工智能和数据专员领导的办公室，作为支持该法监管和执行的专门知识中心。在该法生效并调整生态后，专员的职能将逐步从单纯的教育和援助转变为包括遵守和执行
- 通过制定新的刑法条款，禁止鲁莽和恶意使用人工智能、严重损害加拿大人及其利益的行为

2023 年 6 月 23 日，曼尼托巴省法院（Court of King’s Bench of Manitoba）发布《曼尼托巴省法院的实践指导：在法庭意见书中使用人工智能》，规定当人工智能被用于向法院提交的材料时，材料必须表明人工智能是如何被使用的。

2023 年 6 月 26 日，育空地区最高法院（Supreme Court of Yukon）发布《人工智能工具的实践指导》，要求在该法院的律师或当事人在进行法律研究或提交法律意见书的过程中就任何事项、以任何形式依靠了人工智能（如 ChatGPT 或任何其他人工智能平台）时，必须告知法院所使用的工具及其目的。

（二）一般合规义务

1.高影响人工智能系统的范围

《人工智能和数据法》草案并未对高影响人工智能系统进行界定，而是将具

体标准留给后续将出台的法规予以规定。

尽管目前尚无相关法规，加拿大政府认为下列因素是其审查人工智能系统是否具有高影响的关键因素。[34]

审查是否属于高影响人工智能系统的关键因素
• 基于预定目的和潜在意外后果，证明存在损害健康和安全或对人权造成不利影响的风险
• 潜在危害的严重程度
• 使用规模
• 已经发生的损害或不利影响的性质
• 由于实践或法律原因不能合理地选择退出该制度的程度
• 经济或社会环境不平衡，或受影响人员的年龄
• 风险受另一法律监管的充分程度

以下几类人工智能系统的潜在影响获得了加拿大政府的高度关注，或将落入高影响人工智能系统的范围。

- 影响服务获取或就业的筛查系统
- 用于识别和判断的生物识别系统
- 会大规模影响人类行为的系统
- 对健康和安全至关重要的系统

加拿大政府还指出，高影响人工智能系统的义务受到下列原则指导：

[34] The Artificial Intelligence and Data Act (AIDA) – Companion document (canada.ca)

高影响人工智能系统义务的指导原则	
• 人类监督和监控 ✓ 高影响人工智能系统的设计和开发必须使管理该系统运作的人员能够开展有意义的监督	• 透明度 ✓ 向公众提供关于如何使用高影响人工智能系统的适当信息（信息应该足以让公众了解系统的能力、限制和潜在影响）
• 公平与公正 ✓ 认识到建立高影响人工智能系统可能产生歧视性结果 ✓ 采取适当行动，减轻对个人和群体的歧视性结果	• 安全 ✓ 主动评估高影响人工智能系统，查明使用该系统可能造成的危害，包括可以合理预见的滥用 ✓ 采取措施减轻危害的风险
• 责任 ✓ 建立必要的治理机制，确保在使用高影响人工智能系统时遵守所有法律义务 ✓ 主动记录已实施的策略、过程和措施	• 有效性及稳健性 ✓ 高影响人工智能系统能够始终如一地实现预期目标 ✓ 高影响人工智能系统在各种情况下都稳定且有弹性

2.人工智能系统相关主体义务

（1）负责人的义务

《人工智能和数据法》规定，在国际或省际贸易和商业活动中设计、开发、提供、管理人工智能系统运行的人对该人工智能系统负责，并根据其系统是否为高影响人工智能系统，负有以下义务：

任何人工智能系统负责人的一般义务	高影响人工智能系统负责人的特殊义务
• 依规评估该系统是否为高影响系统 • 依规就数据匿名化的方式、匿名数据的使用或管理制定措施 • 保存记录 ✓ 有关缓解措施（包括缓解任何伤害或偏差输出风险的有效性）的一般记录 ✓ 支持该系统（不）属于高影响系统的原因	
/	• 建立识别、评估和减轻因使用该系统而可能造成的伤害或偏差输出风险的缓解措施 • 建立措施以监督缓解措施的遵守情况 • 系统的使用导致或可能导致重大损害时：尽快通知 ✓ 被指定的加拿大枢密院的成员 ✓ 如无被指定的成员：创新、科学与工业部部长

其中，高影响人工智能系统的提供者、管理运行者除上述义务外，还负有在公开网站上发布系统的简单说明的义务：

高影响人工智能系统提供者、管理运行者的其他义务
• 在公开网站上发布系统的简单说明，内容包括 ✓ 系统的预期使用方式（提供者）/ 如何使用该系统（管理运行者） ✓ 拟生成的内容类型、拟作出的决策、建议或预测 ✓ 现行的缓解措施 ✓ 法规规定的其他信息

对于如何评估和减轻高影响人工智能系统的风险，加拿大政府提供的如下示例可供参考：

受管制的活动	评估和减轻风险的措施实例
• 系统设计 ✓ 包括确定人工智能系统的目标以及实现这些目标的数据需求、方法或模型	• 对在特定场景下使用人工智能系统的潜在风险进行初步评估，并判断人工智能的使用是否合适 • 评估和处理数据集选择引发的潜在偏差 • 评估所需的可解释性水平，并相应作出设计决定
• 系统开发 ✓ 包括处理数据集、使用数据集训练系统、修改系统参数、开发和修改系统中使用的方法或模型，或测试系统	• 记录使用的数据集和模型 • 评估和验证，包括必要的再培训 • 建立人力监督和监测机制 • 记录适当的使用和限制
• 使系统可供使用 ✓ 部署功能齐全的系统，无论是通过开发者、商业交易、应用程序编程接口(API) 还是公开工作系统	• 保存关于如何满足设计和开发需求的文档 • 向用户提供有关数据集使用、限制和适当用途的适当文档 • 对系统的可用性进行风险评估
• 管理系统的运行 ✓ 在使用过程中对系统进行监督，包括开始或停止运行，在系统运行时监测和控制对其输出的访问，改变与系统运行有关的参数	• 适当地记录和监控系统的输出 • 确保充分的监控和人力监督 • 根据运行参数进行必要干预

（2）处理或使用数据的主体：透明度义务

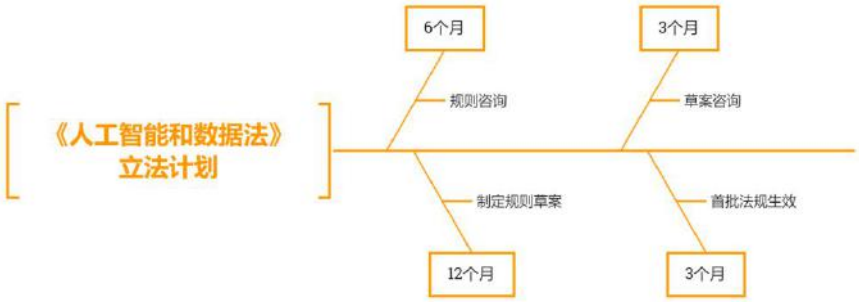
在国际或省际贸易和商业活动中，为设计、开发或使用任何人工智能系统而处理或提供与人类活动有关的任何数据的人承担着一定的透明度义务：

国际或省际贸易和商业活动中为设计、开发或使用人工智能系统而处理或提供与人类活动有关的任何数据的人
• 依规就数据匿名化的方式、匿名数据的使用或管理制定措施
• 保存记录 ✓ 有关缓解措施（包括缓解任何伤害或偏差输出风险的有效性）的一般记录 ✓ 支持该系统（不）属于高影响系统的原因

在设计、开发、提供、管理人工智能系统的过程中，处理或将匿名数据供他人使用的人也存在类似义务：

在国际或省际贸易和商业活动中，处理或将匿名数据供他人使用的人
• 依规就数据匿名化的方式、匿名数据的使用或管理制定措施

根据加拿大政府的立法计划，《人工智能和数据法》最早将于 2025 年生效。在生效后的最初几年，该法的重点将是教育、建立指南和帮助企业自愿遵守合规义务。政府打算在采取执法行动之前，给企业足够的时间适应新的监管框架。



（《人工智能和数据法》立法计划）

《人工智能和数据法》针对违反该法的行为规定了两种类型的处罚：行政罚款和对常规违规行为提起行政诉讼，以及针对刑事违规的单独机制。其中行政罚款的金额将根据即将出台的法规确定。

人工智能系统的负责人违反其合规义务属于常规违规行为，根据其是否是个人、经何种程序定罪，负责人将面临不同的处罚：

	个人	公司
经公诉定罪	法院酌情罚款	不超过 1 千万加元与其被判刑前的财政年度内全球总收入的 3%的罚款中更高额的罚款
经简易程序定罪	不超过 5 万加元罚款	不超过 500 万加元与其被判刑前的财政年度内全球总收入的 2%的罚款中更高额的罚款

除以上外，《人工智能和数据法》规定了 3 种与人工智能系统有关的刑事违

规行为：

刑事违规行为
• 知道或相信个人信息系直接或间接从下列罪行获得或得出，为了设计、开发、使用或提供人工智能系统而拥有或使用该个人信息 ✓ 在加拿大犯下议会立法或省立法机关规定的罪行，或 ✓ 在任何地方的作为或不作为，一旦发生在加拿大就会构成犯罪的罪行
• 在无合法理由的情况下，明知或放任可能对个人造成严重的身体或心理伤害或对个人财产造成重大损害的人工智能系统的使用，仍然供人使用该人工智能系统且该使用导致了上述伤害或损害
• 存在欺诈公众和对个人造成重大经济损失的意图，供人使用该人工智能系统且该使用导致了上述损失

构成上述刑事犯罪的人将受到下述制裁，并可能受到执法部门的调查及加拿大检察署(Public Prosecution Service of Canada)的起诉。

	个人	公司
经公诉定罪	法院酌情罚款 并/或 处 5 年以下（不包括 5 年）监禁	不超过 2500 万加元与其被判刑前的财政年度内全球总收入的 5%的罚款中更高额的罚款
经简易程序定罪	不超过 10 万加元的罚款 并/或 处 2 年以下（不包括 2 年）监禁	不超过 2000 万加元与其被判刑前的财政年度内全球总收入的 4%的罚款中更高额的罚款

（三）监管动态

2023 年 4 月 4 日，加拿大隐私专员办公室因一项声称 ChatGPT 在未经其同意的情况下收集、使用和披露个人信息的投诉，针对 OpenAI 展开调查。[35]

2023 年 5 月 25 日，加拿大隐私专员办公室宣布与魁北克省、不列颠哥伦比亚省和阿尔伯塔省的数据保护机构合作，对 OpenAI 在同意、公开和透明、访问、

[35] https://www.priv.gc.ca/en/opc-news/news-and-announcements/2023/an_230404/

准确性和问责制等方面是否符合加拿大隐私法展开联合调查。[36]

（四）相关机构

加拿大目前没有专门针对人工智能监管的独立机构，主要由不同领域的主管部门监管和执法。

正在审议的《人工智能和数据法》规定创新、科学与工业部部长在系统可能导致伤害或偏差输出时或相关主体可能违规时，可以命令相关主体出示记录以证明其遵守了该法规定、命令开展独立审计。在伤害风险迫在眉睫的情况下，部长可以责令停止使用某人工智能系统、公开披露关于违反该法或为防止伤害的信息。

加拿大检察署在相关人员涉嫌犯罪时，负责考虑是否对其提起诉讼。

加拿大隐私专员办公室一向负责监督加拿大隐私立法的遵守情况，对涉及侵犯隐私保护的行为有执法权。

[36] https://iapp.org/news/a/opc-announces-joint-federal-provincial-investigation-of-openai-at-iapp-cps-2023/

About the authors

作者简介

丁震宇

安杰世泽律师事务所丨合伙人

邮箱：dingzhenyu@anjielaw.com

丁震宇律师主要从事兼并与收购、私募股权投资方面的法律服务，对于区块链、大数据、人工智能等前沿科技领域的法律法规与业务实操有着深刻的理解与研究。丁震宇律师为复旦大学法学院法律硕士开设“数字法治前沿课程”，还曾在浦江法治论坛发表《中国主权数字货币(DCEP)的机遇及其全球监管合作》演讲。

丁震宇律师于 2005 年毕业于新加坡国立大学，获公司法与金融法专业法学硕士学位，于 2002 年获得复旦大学法学学士学位。丁震宇律师曾荣获 2022 年 Legalband 数字经济领域的中国律师特别推荐榜 15 强，2019 年至 2022 年连续获得《国际金融法律评论（IFLR1000）》在并购领域的重点推荐。

周博华

安杰世泽律师事务所丨律师助理

周博华于 2023 年获得北京大学法律硕士学位，于 2021 年获得复旦大学法学学士学位。周博华曾获得第一届安杰世泽学术潜力之星，主要从事数据合规、人工智能方面的法律服务。

*注：感谢实习生王祉恒对本研究报告提供的协助。

Anjie Broad
安杰世泽