

谋求平衡：以务实姿态应对欧盟监管

2024年，生成式人工智能领域的监管与创新有望达到平衡，尽管有预测称欧盟法规或将阻碍市场发展。

健全有效的监管法规有利于释放市场潜力。就生成式人工智能而言，明确的法规有助于增强供应商、企业客户和最终用户投资该技术的信心。欧盟拟于2024年出台相关法规以规范自身市场，这将为其他地区的立法提供参考和示范，并为全球范围的监管奠定基础。

2024年，欧盟预计有两项法规将助力规范生成式人工智能的发展与应用，分别是自2018年起生效的《通用数据保护条例》(GDPR)¹和即将出台(有望于2024年初生效)的《人工智能法案》(AIA)。随着生成式人工智能引发了关于如何管理个人同意、更正、删除、减少偏见和版权使用等问题的讨论，组织和监管机构如何看待、执行和管理争议领域，可能将影响该行业的发展轨迹。

尽管存在潜在挑战，但行业与监管机构之间以公开透明对话的形式开展务实合作，有助于实现生成式人工智能监管与创新之间的平衡。该过程依旧延续2023年的讨论模式，欧盟及其他市场的监管机构纷纷参与。供应商遵循要求积极调整，监管机构亦力促行业创新。²如果2024年欧盟法规所提出的问题得以解决，同时强化核心技术优势，则该市场发展势必蒸蒸日上。

欧盟现行和拟出台法规将对生成式人工智能产生全球性影响

本报告将重点关注欧盟关于生成式人工智能的立法，因其很可能是首批达成的具有全球性影响的法规之一。³近年来，随着欧盟法规产生全球性影响，“布鲁塞尔效应”⁴已经显现，我们预计涵盖生成式人工智能的欧盟法规也将延续这一效应。⁵欧盟法规的域外效力将带来不同影响：

1. **欧盟法规直接适用于面向欧盟市场或用户的所有供应商。** 不合规组织或被处以巨额罚款。
2. **欧盟法规将为其他市场的立法提供参考和示范。** 例如，印度的《2023年数字个人数据保护法案》⁶以及巴西和美国加州的类似法规均以欧盟法规为参照。《人工智能法案》则为美国参议院两党的人工智能立法蓝图提供了参考，其中包括对人脸识别等高风险应用的许可以及基础模型训练数据的公开披露等内容。⁷

3. **跨国公司以及欧盟境内公司可在全球范围内采用统一的人工智能治理（包括针对生成式人工智能的治理）实践**，以在遵循欧盟法规的基础上，采用与领先实践相一致的、更标准化的治理方法。

到2024年第一季度，预计欧盟有关生成式人工智能的大部分法规将变得相对明确。

2024年，欧洲对生成式人工智能的监管方向将更加明确。届时，业界应能看到《人工智能法案》的议定文本，作为对GDPR的补充。⁸所有希望提供或部署生成式人工智能解决方案的公司都应在遵守GDPR的同时，密切关注《人工智能法案》的发展动向。

达成《人工智能法案》最终协议的过程分为三个阶段；在撰写本报告时，前两个阶段已经敲定，第三阶段也是最后阶段，有待欧盟理事会、欧洲议会和欧盟委员会的“三方会谈”结果。

- 欧盟理事会于2022年12月最终确定其立场，⁹彼时随着Chat GPT的问世，生成式人工智能才刚刚进入主流意识。
- 欧洲议会于2023年6月最终确定其立场，其中包括对生成式人工智能的具体规定。¹⁰本报告关于《人工智能法案》的援引主要反映的是当时状况。
- 《人工智能法案》的最终版本预计将于2024年初发布，其中可能包括欧洲议会立场的调整。《人工智能法案》生效后，将会留出两年的宽限适应期。

以下是欧盟关于生成式人工智能的立法所适用的一些特定术语。在《人工智能法案》中，关于生成式人工智能的关键组成部分和参与者类型定义如下：

- **基础模型 (FM)**：指经过大量数据训练的人工智能模型，如Open AI的GPT模型或谷歌的PaLM 2。¹¹该类模型可应用于各种不同任务，因此不同于狭义的人工智能模型。根据欧盟的定义，基础模型可用于GPAI或其他更具体的人工智能模型。¹²
- **通用人工智能 (GPAI)**：指具有广泛适用性的人工智能系统¹³，故其并无“预期使用目的”。¹⁴它可用于“多种环境”和“多种其他人工智能系统”。核心功能包括包括识别（如图像或语音）、生成（目前最常见的是文本或图像）、模式检测和翻译。
- **生成式人工智能**：指专门用于生成各种格式输出的人工智能系统。最著名的生成式人工智能应用包括Chat GPT、Snap AI、谷歌的Bard和微软的M365 copilot。

《人工智能法案》还对两类实体做出定义：

- **提供者**：指开发或接受委托开发人工智能系统，以期有偿或无偿向公众提供该系统的自然人或法人、公共当局、机构或其他团体。
- **部署者**：指任何自然人或法人、公共当局、机构或其他团体在其授权下使用人工智能系统。在某些情况下，部署者也可能是提供者，比如部署者将人工智能系统用于高风险应用。

接下来，我们将首先探讨众所周知的GDPR，然后分析即将出台的《人工智能法案》。

生成式人工智能与欧盟GDPR

预计生成式人工智能的发展与应用也须遵循GDPR关于个人数据处理的规定。2018年5月生效的GDPR¹⁵明确了“数据主体”的权利，数据主体是指其个人数据经处理后可用于识别其身份的个人。

欧盟法规的一项基本原则是，个人数据的使用应遵循相关法律法规，每项处理活动都要保证其合法性。¹⁶

这一要求似乎与基于基础模型的生成式人工智能的核心原则相冲突。每个模型都是在大量（且越多越好）的原始数据上训练而来。根据欧盟相关法律规定，这类数据的绝大部分（具体比例因模型而异）可能需要征得同意。最大的基础模型或基于数万亿字节（数千GB）或数千万亿字节（数PB）的数据训练而来。¹⁷GPT-3等早期模型即是基于570 GB的数据训练而来。¹⁸任何媒介（文本、图像、代码或其他）中的生成式人工智能应用，均可利用每个基础模型中的知识创建内容。

鉴于数据使用者众多，在必要情况下征得个人同意便成为了一项复杂工作。此外，由于每个基础模型支持实际上支持无限数量和范围的应用程序，因此为每个额外目的征求同意则更加不切实际。

不过，征得个人同意有时并非强制性要求。基于“合法利益”使用个人数据以训练生成式人工智能基础模型，也可称为具有充分的“合法依据”。¹⁹当处理数据存在充分理由并且是实现预期结果的唯一方法时，即存在合法利益²⁰。监管机构或希望看到相关组织已展开了适当评估，以在依据合法利益行事与保护个人权利和自由之间找到平衡。

此外，征得个人同意很可能需要付出“不相称的努力”。一种折衷之法可能是公开数据操作的存在及其目的。这是意大利监管机构Garante于2023年4月允许Open AI恢复服务所要求其采取的步骤之一。²¹Garante规定数据控制者（基础模型的指定负责人）有义务在广播和在线媒体上开展宣传活动。²²此举旨在告知用户其个人数据可能已被使用，并说明如何通过在线工具删除此类数据。

这有助于让监管机构对训练目的持积极看法，即训练旨在提高模型的推理能力，从而将其部署到生成式应用程序中（如Open AI的Chat GPT、stability.ai的Dream Studio或Adobe的Firefly）。

欧盟数据保护委员会（EDPB）或将在2024年进一步明确同意问题及其他有争议的领域²³。

GDPR关于更正、删除和被遗忘权的原则适用于生成式人工智能的基础模型

GDPR包括一系列与个人数据相关的权利。如果数据有误，个人可以要求更正。如果数据主体不再希望个人数据与某组织相关联或由该组织处理，则可要求其删除相关数据。自GDPR生效以来，这些权利已广为人知。处理此类请求可能会花费组织数千美元。

生成式人工智能的基础模型是在无数可能包含错误的网站上训练而来。训练过程是一个单一事件，期间错误也可能被纳入模型。通过重新训练模型，可以最准确地更新模型以反映更正或其他更改，但这需耗费大量的成本和时间。²⁴

潜在解决方案是利用负反馈环路对模型进行微调²⁵。如果确定原始数据点是错误的，则可以改变应用于错误数据点的权重，以尽量减少该数据点再次出现的可能性。反馈环路并非万全之策，却是适当之法。不过，目前还不确定这种方法在可能需要删除大量数据的集体诉讼挑战中将如何发挥作用。

数据最小化原则和统计准确性

数据最小化原则是指, 个人信息的收集应严格限制在与完成特定任务相关且必要的范围内, 一旦任务完成, 应立即删除数据。²⁶该原则似乎与基础模型的作用原理相悖, 因为基础模型效能的高低与其能查询数据的多少挂钩。

不过, 如果对数据进行去个性化处理, 比如采用假名化 (用占位符值替换个人标识符, 可降低但不能消除数据泄露风险) 和匿名化 (删除标识符, 则数据不再能够识别特定个人) 等方法, 数据最小化原则或仍适用于生成式人工智能监管。^{27 28}使用这类方法有助于保持训练数据量, 但要完全匿名化却有一定难度。各组织应制定适当框架进行评估, 并向监管机构解释和保证其如何确定哪些处理是必要之举。

基础模型的大小与统计准确性有关, 这是《人工智能法案》拟议法规的一个要素。²⁹在人工智能背景下, 准确性是指生成输出的质量。对于基础模型而言, 高质量训练数据越多, 输出结果就应越准确。³⁰

接下来, 我们将探讨《人工智能法案》对生成式人工智能可能产生的影响。

根据欧洲议会达成的协议, 《人工智能法案》确立了基础模型的义务

如前所述, 欧盟议会于2023年6月最终确定其立场, 其中包括对生成式人工智能的具体规定。《人工智能法案》的最终版本预计将于2024年初发布, 其中可能包括欧洲议会立场的调整。

协议包括以下内容:

- 基础模型应在欧盟数据库中进行注册。
- 应对模型进行广泛测试, 保证其在整个预期生命周期内具有适当水平的可预测性、可解释性、可追溯性、安全性和网络安全性。
- 通过适当的设计、测试和分析, 识别并降低模型开发过程中的风险。
- 应就模型训练数据集建立完善的数据治理标准。应评估数据源的数据质量和偏见。
- 在模型的整个生命周期内, 应监控并尽量减少能源使用。
- 向下游提供者提供详细易懂的技术文件, 助其遵循合规要求。技术文件应在产品投入市场后十年内可用。
- 建立质量管理体系, 以确保并记录合规情况。

此外, 生成式人工智能系统所用基础模型的提供者, 以及将基础模型专门用于生成式人工智能系统的提供者应该:

- 遵循额外的透明度要求, 包括将输出注明为人工智能生成内容
- 确保生成内容符合欧盟法律
- 记录并公布受版权保护的训练数据摘要

偏见或可减少

《人工智能法案》旨在最大限度地减少人工智能系统中的偏见。这包括减少人类偏见。基础模型训练可能基于包含偏见的内容，例如包含性别、种族、性取向偏见的文本。

训练数据还可能包括语言偏见，大多数内容都是用英语撰写的，而从特定性别、种族、社会阶层、教育程度和收入群体的作者处获得的内容居多，也会产生额外偏见。³¹因此，用于训练基础模型的历史偏见可能会生成重复甚至加剧这些偏见的内容。

监管机构将要求使用多种方法减少偏见，包括加权或纳入可抵消偏见的合成数据。³²数据控制者（人工智能开发者或部署者）或被要求记录“可追溯性”，以解释所采取的步骤。³³

版权：需要进一步明确允许行为

2024年，版权内容的使用可能需要进一步明确。³⁴

欧盟现行法律可能允许将版权数据用于模型训练，特别是“不涉及复制行为的文本与数据挖掘应用实例，或属于临时复制行为的强制性应用的例外情形”。³⁵《人工智能法案》草案要求披露用于模型训练的版权内容。

欧盟近期通过《数字化单一市场版权指令》(Digital Single Markets Directive)³⁶引入了将文本与数据挖掘用于科研和商业合法用途的许可；不过，对于商业用途，用户有权“选择退出”该许可。包括几家媒体公司在内的内容所有者已经行使了这一权利，选择将其数据从人工智能训练中剔除。³⁷截至2023年4月，Stable Diffusion v3模型的训练已集中删除了超过十亿条数据。³⁸

最大的基础模型可归类为系统基础模型

《人工智能法案》旨在对人工智能的每项应用进行风险评估。这与基础模型的通用性质相悖。

然而，按照欧盟《数字服务法案》对在线平台和搜索引擎类型进行分类时所采用的方法来讲，系统基础模型（SFM，其影响代表系统风险的模型）与其他模型可能存在区别。³⁹是否可被归为SFM，可能取决于训练模型所需的计算资源数量、所用的输入类型和成本，以及其可能产生的市场影响。SFM或需承担更多的尽职调查义务。⁴⁰

另一个可能的结果是，《人工智能法案》或将建立起适用于所有基础模型的若干基本要求，例如就透明度和技术文件而言，如果基础模型用于高风险人工智能系统，则须遵循额外的要求。

小结

欧盟法规将发挥重要作用。它可能会产生域外效力和地区影响。乍看之下，欧盟法规关于数字服务的若干现行原则似乎极大阻碍了生成式人工智能的发展与应用。事实上，一些评论家可能认为欧盟的指导原则不适用于生成式人工智能监管。

生成式人工智能在未来几年的发展趋势及影响如何，目前仍是未知数。要确定其影响的规模和性质，可能还需要几年时间。2024年及以后，供应商和监管机构预计将希望通过加强合作，为消费者、企业、供应商和整个社会带来有利结果。各国政府清楚意识到依托监管沙盒等方式促进生成式人工智能创新发展的重要性。⁴¹

2024年，随着生成式人工智能应用的发展以及由此带来的法律挑战变得更加清晰，监管响应的方向或将变得更加明确。生成式人工智能仍属新兴领域之一，这或使现阶段的立法难以明确。可能仍有一些核心问题亟待解决，比如当生成式人工智能的提供者和部署者为两个独立实体时，应当如何划分双方责任。

作者

Paul Lee

United Kingdom

Lucia Lucchini

United Kingdom

Michelle Seng Ah Lee

United Kingdom

Valeria Gallo

United Kingdom

政策助力生成式AI市场创新发展

为了应对生成式AI产业的发展与变化,我国也迅速出台相关监管规定,如《互联网信息服务深度合成管理规定》、《生成式AI服务管理暂行办法》等,加强精细化治理与监管,逐步建立健全监督体系。在2023年7月,中国网信办联合有关部门发布了《生成式AI服务管理暂行办法》(办法),自8月15日起施行。《办法》将“促进生成式AI健康发展和规范应用”作为治理的第一视角。

放眼海外,各国都在基于既有的人工治理框架、制度与规范,设计针对生成式AI的治理方案。欧盟基于2018年起的适用的《通用数据保护条例》(GDPR)和即将出台的《欧盟人工智能法案》(AIA)助力生成式AI产业有序发展。与《欧盟人工智能法案》(AIA)以风险预防和责任追究为重心的立法不同,《办法》将生成式AI的创新与发展置于重要地位,鼓励探索生成式AI的应用与发展。

图:《欧盟人工智能法案》与《生成式AI服务管理暂行办法》区别

维度	欧盟	中国
监管思路	风险分级	包容审慎和分级分类
适用范围	包括属人、属地和实质多个管辖标准(域外适用)	中国境内
数据管理和隐私保护	与《通用数据保护条例》(GDPR)保持一致,强调个人信息合法性说明	基于网络安全法、数据安全法、个人信息保护法对生成式AI进行针对性设计,强调个人信息合法性及知识产权
责任主体	AI价值链上的多个参与主体	GAI服务提供者

来源:外部资料、德勤研究

最新出台的《办法》内容主要体现在以下几个方面:

- 第一 监管范围主要针对利用AIGC向中国境内公众提供生成文本、图片、音频等内容(生成式AI服务)的服务商。
- 第二 立足于生成式AI基础技术创新(如算法、芯片及配套软件平台)、多方行业参与者(如企业、教育和科研机构)协作、以及国际交流与合作(如参与生成式AI国际规则制定)等维度,促进技术发展。
- 第三 生成式AI服务治理主要聚焦使用具有合法来源的数据和基础模型,尤其是涉及知识产权,个人信息的数据。持续关注《网络安全法》、《数据安全法》、《个人信息保护法》等法律、行政法规和有关主管部门的相关监管要求。
- 第四 生成式AI提供者、服务者需遵守服务规范,其中提供者应依法承担网络信息内容生产者责任,涉及个人信息的,还需承担个人信息处理者责任。
- 第五 明确国家网信办、发改委、教育部、科技部、工业和信息化部、公安部、国家广电总局等部门参与生成式AI服务监管。

当前《办法》尚未对分类分级监管形成具体实施方案，但随着中国式人工智能相关技术和应用逐渐成熟，将逐步形成分行业、分场景、分应用的差异化监管模式。《办法》的出台，充分协调了发展与安全之间的关系，在释放出促进生成式AI技术发展明确信号的同时，还明晰了生成式AI技术发展路径和方向，引导企业在划定的范围加速开展技术创新和应用。展望2024年，我们预计生成式AI行业呈现以下四个发展趋势：

1. **政策更倾向于发展一侧，行业企业迎来更为包容的创新环境。**《办法》中重点突出了AIGC服务应用层的信息内容安全，对模型及其相关技术作为基础设施的功能研发活动保留了一定试错空间，行业企业可联合各科研机构，继续投入开发生成式大模型，探索生成式AI在多领域的应用。
2. **行业企业重心逐步向基础技术自主创新、多领域应用场景等转移。**行业企业应着重对生成式AI算法、框架、芯片及配套软件平台等基础技术实现自主创新模式。同时围绕自身优势构建行业大模型能力，满足差异化大模型服务需求，可重点从ToB端切入行业大模型领域，助力政企行业客户提质增效。
3. **生成式AI产业生态体系正在进一步完善。**在生态领域，《办法》特别提到支持构建产学研相结合的大模型生态，使各方参与主体积极参与到大模型领域国家标准、行业标准制定，协同参与制定相关国际标准。
4. **公共训练数据适度开放，引导行业企业开展更多相关数据训练活动。**《办法》中涉及扩大训练数据范围，以及推动生成式AI公共训练数据资源平台建设等内容，打消了提供者对使用训练数据的顾虑，将进一步赋能算力基础设施等领域，提供大规模智算和大模型开发平台服务，进而提升算力资源利用效能。

生成式AI技术尚在发展中，目前仍处于起步阶段。由于其技术模型在研发、使用、优化等过程中涉及数据提供方、开发者、服务提供者、使用者等不同主体，因此生成式AI的健康发展需要多方参与建设。当前我国已逐步对生成式AI应用划定发展“红线”，企业在未来应聚焦《办法》中的重要事项，并采取措施开展有效风险管理。

- **建立全面的生成式AI管理体系。**在大模型研发、生产和应用过程中，对生成式内容主动进行合理标记并在模型训练、提供服务等环节进行合规管理。
- **强化模型训练和数据安全。**企业在进行模型训练和应用时，确保训练数据的合规性和安全性，保障数据来源合法性，尤其需要强化对外部数据提供方的审查和管控。对数据流转情况，按规定履行评估申报、合同备案义务；同时提升训练数据多样性，避免从单一范围训练数据。
- **建立算法安全监测、漏洞报告制度。**通过一系列工具链对AIGC算法建立持续的安全监控和漏洞报告，提高对异常数据或行为的识别能力，及时响应用户反馈，并配合监管机构的合规监督检查，防止人工智能算法产生歧视、侵权、不正当竞争等行为。

作者

钟昀泰
中国

尾注

1. 欧洲议会和欧洲理事会于2016年4月27日发布的关于在个人数据处理和此类数据的自由流动方面保护自然人并废除95/46/EC指令的第(EU) 2016/679号条例 (《通用数据保护条例》) : <https://eur-lex.europa.eu/eli/reg/2016/679/oj>
2. First regulatory sandbox on Artificial Intelligence presented, European Commission, <https://digital-strategy.ec.europa.eu/en/news/first-regulatory-sandbox-artificial-intelligence-presented#:~:text=A%20pilot%20of%20the%20first,renowned%20experts%20in%20the%20field> ; Spanish Ministry of Finance, <https://www.hacienda.gob.es/Documentacion/Publico/GabineteMinistro/Notas%20Prensa/2023/CONSEJO-DE-MINISTROS/22-08-23-NP-CM-Estatutos-Agencia-Inteligencia-Artificial.pdf>
3. China finalises its Generative AI Regulation, Data Protection Report, <https://www.dataprotectionreport.com/2023/07/china-finalises-its-generative-ai-regulation/>
4. The Brussels Effect and Artificial Intelligence, Centre for the governance of AI, <https://arxiv.org/pdf/2208.12645.pdf>
5. European framework on ethical aspects of artificial intelligence, robotics and related technologies, European Parliamentary Research Service, [https://www.europarl.europa.eu/RegData/etudes/STUD/2020/654179/EPRS_STU\(2020\)654179_EN.pdf](https://www.europarl.europa.eu/RegData/etudes/STUD/2020/654179/EPRS_STU(2020)654179_EN.pdf)
6. The Gazette of India, <https://egazette.gov.in/WriteReadData/2023/248045.pdf> ; Future of Privacy Forum, <https://fpf.org/blog/the-digital-personal-data-protection-act-of-india-explained/>
7. Senators Want ChatGPT-Level AI to Require a Government License, Wired, <https://www.wired.com/story/senators-want-chatgpt-ai-to-require-government-license/#>
8. European Parliament, <https://www.europarl.europa.eu/news/en/headlines/society/20230601STO93804/eu-ai-act-first-regulation-on-artificial-intelligence>
9. The road to the adoption of the regulation, European Union, <https://spanish-presidency.consilium.europa.eu/en/news/eu-pioneers-regulation-artificial-intelligence/#:~:text=EU%20countries%20approved%20the%20negotiating,to%20facilitate%20investment%20and%20innovation.>
10. European Parliament, <https://www.europarl.europa.eu/news/en/headlines/society/20230601STO93804/eu-ai-act-first-regulation-on-artificial-intelligence>
11. NVIDIA, <https://blogs.nvidia.com/blog/2023/03/13/what-are-foundation-models/>; Arxiv, <https://arxiv.org/abs/1810.04805> ; Google AI, <https://ai.google/discover/palm2/>
12. Ada Lovelace Institute, <https://www.adalovelaceinstitute.org/resource/foundation-models-explain-er/#:~:text=Notable%20examples%20are%20OpenAI%27s%20GPT,-%20but%20important%20-%20to%20regulate.>
13. Council of the European Union, [https://data.consilium.europa.eu/doc/document/ST-14278-2021-INIT/en/pdf](https://data.consilium.europa.eu/doc/document/ST-14278-2021-INIT/en/pdf;);

14. OECD, <https://oecd.ai/en/wonk/eu-definition-gpais>
15. Council of the European Union, <https://artificialintelligenceact.eu/wp-content/uploads/2022/05/AIA-FRA-Art-34-13-May.pdf>
16. Council of the European Union, <https://www.consilium.europa.eu/en/policies/data-protection/data-protection-regulation/>
17. Norwegian Consumer Council, <https://storage02.forbrukerradet.no/media/2023/06/generative-ai-rapport-2023.pdf>
18. IBM and NASA team up to spur new discoveries about our planet, IBM, <https://research.ibm.com/blog/ibm-nasa-foundation-models>
19. NVIDIA, <https://blogs.nvidia.com/blog/2023/03/13/what-are-foundation-models/>
20. Can legitimate interest be an appropriate lawful basis for processing Artificial Intelligence training datasets?, Vrije Universiteit Brussel, <https://researchportal.vub.be/en/publications/can-legitimate-interest-be-an-appropriate-lawful-basis-for-proces>
21. Information Commissioner's Office UK, <https://ico.org.uk/for-organisations/uk-gdpr-guidance-and-resources/lawful-basis/a-guide-to-lawful-basis/lawful-basis-for-processing/legitimate-interests/#:~:text=The%20legitimate%20interests%20can%20be,The%20processing%20must%20be%20necessary.>
22. The Italian Data Protection Authority, <https://www.garanteprivacy.it/web/guest/home/docweb/-/docweb-display/docweb/9870832>
23. EDPB resolves dispute on transfers by Meta and creates task force on Chat GPT, European Data Protection Board, https://edpb.europa.eu/news/news/2023/edpb-resolves-dispute-transfers-meta-and-creates-task-force-chat-gpt_en
24. OpenAI's CEO Says the Age of Giant AI Models Is Already Over, Wired, <https://www.wired.com/story/openai-ceo-sam-altman-the-age-of-giant-ai-models-is-already-over/>
25. The AI Feedback Loop: Maintaining Model Production Quality In The Age Of AI-Generated Content, Unite.ai, <https://www.unite.ai/the-ai-feedback-loop-maintaining-model-production-quality-in-the-age-of-ai-generated-content/#:~:text=An%20AI%20feedback%20loop%20is,%2C%20development%2C%20and%20model%20improvement.>
26. European Data Protection Supervisor, [https://edps.europa.eu/data-protection/data-protection/glossary/d_en#:~:text=The%20principle%20of%20"data%20minimisation,necessary%20to%20fulfil%20that%20purpose.](https://edps.europa.eu/data-protection/data-protection/glossary/d_en#:~:text=The%20principle%20of%20)
27. Irish Data Protection Commissioner, <https://www.dataprotection.ie/en/dpc-guidance/anonymisation-pseudonymisation#:~:text=Where%20data%20has%20been%20anonymised,is%20still%20considered%20personal%20data.>
28. Information Commissioner's Office UK, <https://ico.org.uk/media/about-the-ico/consultations/4019579/chapter-3-anonymisation-guidance.pdf> ;

29. European Parliament, https://www.europarl.europa.eu/doceo/document/TA-9-2023-0236_EN.pdf
30. On the Opportunities and Risks of Foundation Models, Arxiv, <https://arxiv.org/pdf/2108.07258.pdf>
31. Tackling bias in large ML models: the role of synthetic data, Syntheticus, <https://syntheticus.ai/blog/tackling-bias-in-large-ml-models-the-role-of-synthetic-data>
32. Synthetic data 101: What is it, how it works, and what it's used for, Syntheticus, <https://syntheticus.ai/guide-everything-you-need-to-know-about-synthetic-data>
33. How to Address Data Bias in Machine Learning, Medium, <https://towardsdatascience.com/how-to-address-data-bias-in-machine-learning-c6a45db53b8d>
34. Legal Issues in Generative AI under Japanese Law, Lexology, <https://www.lexology.com/library/detail.aspx?g=68d490a1-3021-4040-afdd-90ae8fa69337>
35. European Parliament, <https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32019L0790>
36. Official Journal of the European Union, <https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:32019L0790&qid=1697472087050>
37. Kluwer Copyright Blog, <https://copyrightblog.kluweriplaw.com/2023/05/09/generative-ai-copyright-and-the-ai-act/>
38. Spawning lays out plans for letting creators opt out of generative AI training, TechCrunch, <https://techcrunch.com/2023/05/03/spawning-lays-out-its-plans-for-letting-creators-opt-out-of-generative-ai-training/?guccounter=1>
39. European Union's Digital Services Act, <https://www.eu-digital-services-act.com/#:~:text=will%20have%20to%3A-,1.,based%20violence%20or%20mental%20health.>
40. Adapting the European Union AI Act to deal with generative artificial intelligence, Bruegel, <https://www.bruegel.org/analysis/adapting-european-union-ai-act-deal-generative-artificial-intelligence#:~:text=The%20evolving%20AI%20Act,based%20on%20its%20intended%20use.>
41. First regulatory sandbox on Artificial Intelligence presented, European Commission, <https://digital-strategy.ec.europa.eu/en/news/first-regulatory-sandbox-artificial-intelligence-presented#:~:text=A%20pilot%20of%20the%20first,renowned%20experts%20in%20the%20field;Spanish>

致谢

The authors would like to thank **Nick Seeber, Lukas Kruger, Suchitra Nair, Ben Stanton, Robert MacDougall, Joanne Conway, and Isabel Parker** for their contributions to this article.

Cover image by: **Manya Kuzemchenko**