为 AI 问题选择正确的政策解决方案

作者 ： Hodan Omaar 和 Daniel Castro | 2024 年 5 月 20 日

INTRODUCTION

政策制定者发现自己在要求他们迅速采取行动应对人工智能 （ AI ） 风险的呼声中。人们的担忧涉及一系列社会和经济问题，从人工智能取代工人和助长错误信息到威胁隐私、基本权利甚至人类文明。有些担忧是合理的，但有些则不是。有些人需要立即采取监管措施，但许多人没有。一些需要专门针对 AI 的法规，但大多数没有。识别哪些问题值得回应，以及它们需要采取哪些类型的政策行动，才能制定有针对性、有影响力和有效的政策，以应对人工智能带来的真正挑战，同时避免扼杀创新的不必要的监管负担。

本报告涵盖了对人工智能的 28 个普遍关注，对于每一个关注，描述了关注的性质，这种关注是否以及如何是人工智能特有的，以及什么样的政策回应 （ 如果有的话 ） 是合适的。可以肯定的是，还有一些可能被纳入的担忧，还有一些将在未来提出的担忧，但从对人工智能文献的回顾和人工智能监管行动的不断增长的语料库来看，这些都是政策制定者必须应对的主要担忧。本报告将 28 个关注的问题分为 8 个部分 ： 隐私，劳动力，社会，消费者，市场，灾难性情景，知识产权以及安全与保障。每个问题都可以提出自己的报告，但这里的目标是提炼每个问题的实质，并提供务实，清晰的回应。

对于每个问题， 我们将适当的政策响应分类如下：

追求监管是 ...

AI - specific： 对 **AI** 的一些担忧最好通过制定或更新专门针对 **AI** 系统的法规来解决。这些法规可能会禁止某些类型的 **AI** 系统， 创建或扩大对 **AI** 系统的监管， 或对 **AI** 系统的开发人员和操作员施加义务， 例如要求审核， 信息披露或影响评估。

一般： 对人工智能的一些担忧最好通过制定或更新法规来解决， 这些法规不是专门针对人工智能， 而是创建适用于各个行业和部门的广泛法律框架。这些法规的示例包括数据隐私法， 政治广告法和复仇色情法。

追求非监管政策是...

AI - specific： 对 **AI** 的一些担忧最好通过实施针对 **AI** 的非监管政策来解决。这些政策的例子包括资助 **AI** 研发或支持 **AI** 特定行业标准的开发和使用。

一般： 对人工智能的一些担忧最好通过实施非监管政策来解决， 这些政策不是针对人工智能， 而是关注人工智能系统运行的更广泛的技术和社会背景。这些政策的例子包括就业错位政策， 以减轻劳动力市场更加动荡的风险， 或改善联邦政策数据质量。

不需要政策

一些担忧最好通过现行政策或允许社会和市场随着时间的推移而适应来解决。政策制定者目前不需要实施新的监管或非监管政策。

CONTENTS

AI 担忧的政策需求概述

- 需要 AI 特定法规的担忧：

1.3. 人工智能可以实现政府监督。

3.6. AI 可能会做出有害的决定。

8.1. 人工智能可能导致欺诈和身份盗窃。

8.3. AI 可能会产生安全风险。

- 需要遵守一般规定的担忧:

1.1. AI 可能会在数据泄露中暴露 PII 。

1.5. AI 可以推断敏感信息。

1.6. AI 可能会帮助不良行为者骚扰和公开羞辱个人。

3.2. AI 可能会在选举中助长深度造假。

6.1. AI 可能使建造生物武器变得更容易。

7.3. 大赦国际可能侵犯宣传权。

- 需要 AI 特定非监管政策的担忧：

1.4. 人工智能可以实现工作场所监控。

3.3. AI 可能操纵选民。

3.5. 大赦国际可能使歧视长期存在。

6.2. AI 可能会创造出新颖的生物反应器。

6.3. AI 可能会变成像上帝一样的“超级智能”。

6.4. 人工智能可能导致能源使用失控。

7.1. 人工智能可能非法培训受版权保护的内容。

8.2. AI 可能导致网络攻击。

- 需要采取一般非监管政策的担忧：

1.2. AI 可以揭示训练数据中包含的 PII 。

2.2. AI 可能会使蓝领工人流离失所。

2.3. AI 可能会使白领脱臼。

3.1. 大赦国际可能存在政治偏见。

7.2. AI 可能会创建侵权内容。

- 不需要新政策的担忧：

2.1. AI 可能导致大规模失业。

3.4. AI 可能会助长不健康的个人附件。

4.1. AI 可能会加剧监控资本主义。

5.1. 人工智能可能使拥有关键投入的公司能够控制市场。

5.2. AI 可能会强化技术垄断。

1. 隐私

#	风险	政策需要政策解决方案	
1.1	AI 可能会在数据泄露中暴露个人身份信息。	一般规定	政策制定者应要求公司发布安全政策， 以提高消费者的透明度。国会应通过联邦数据泄露通知立法。
1.2	AI 可能会揭示敏感将军信息政策培训数据中。	非监管性包括在	政策制定者应该资助隐私和安全增强技术的研究， 应该支持行业主导的负责任的网络抓取标准。
1.3	人工智能可以实现政府监控。	特定于 AI 的法规	国会应指示司法部 （ DOJ ） 制定指导方针， 以供州和地方执法部门在调查中使用， 概述具体的用例和能力， 包括何时需要使用逮捕令，以及何时通知公众使用人工智能执法的透明度准则。
1.4	AI 可以实现工作场所监控。	特定于 AI 的非监管政策	政策制定者应帮助设定工作场所使用的 AI 技术的质量和性能标准
1.5	AI 可以推断敏感信息。	一般规定	政策制定者应该制定和制定全面的国家隐私立法， 以技术中立的方式解决数据驱动推理的风险。
1.6	AI 可能会帮助不良行为者骚扰和公开羞辱个人。	一般规定	国会应禁止非自愿分发所有色情图像， 包括在色情图像中复制个人肖像的深度假货， 并制定一项联邦法规，禁止复仇色情， 包括带有计算机生成图像的色情内容。

问题 1.1：AI可能在数据泄露中暴露个人信息

问题： 当有人未经授权访问数据时，就会发生数据泄露。例如，攻击者可能会规避安全措施以获取敏感数据，或者内部人员可能会不适当地访问机密信息。用户可以与人工智能系统（如聊天机器人）共享个人身份信息 (PII)，提供法律、金融或健康服务。如果发生数据泄露，这些对话的记录可能会被曝光和不当访问，从而泄露敏感信息。数据泄露的一个例子是 2023 年 3 月 OpeAI 的 ChatGPT 聊天机器人发生的大量报道事件。由于系统使用的开源库中的错误，一些用户能够从其他用户的聊天记录中看到标题。[1]

虽然人工智能系统确实可能遭受数据泄露，就像任何 IT 系统一样，但它们并没有造成或加剧潜在的隐私和安全风险。在过去的二十年中，数据泄露一直是不幸但经常发生的事情。2022 年，美国有近 1800 起数据泄露事件影响了数亿美国人。[2]

解决方案： 政策制定者应该解决更大的数据泄露问题，而不是只关注涉及 AI 系统的数据泄露。国会可以做的一件事是要求公司发布安全策略，以提高消费者的透明度。大多数公司发布隐私政策，为监管机构创建透明和负责任的机制，以确保公司遵守其既定政策。但是，对于信息安全实践，这种做法并不存在，这导致了模糊的标准，流行语的监管以及市场中的信息不对称。通过发布安全策略，公司将有动力描述他们已经采取的安全措施的类型，而不仅仅是声称采取了“合理的安全措施”。“这是政策制定者可以采取的改善私营部门安全做法的具体步骤。[3]

此外，国会应该通过数据泄露通知立法，优先于相互冲突的州法律。[4]所有 50 个州以及哥伦比亚特区、关岛、波多黎各和维尔京群岛都有数据泄露法律；然而，每个司法管辖区都有自己的一套规则，规定如何快速报告数据泄露或安全事件应报告给谁。这种不同要求的拼凑被子为消费者提供了明显的不平衡保护，并为公司创造了不必要的复杂局面，这些公司必须花费更多的时间来浏览这个模糊的法律领域，而不是实际保护消费者数据。[5]